Scientific Research Publishing

Journal of Mathematical Finance, 2022, 12, 46-70
https://www.scirp.org/journal/jmf
ISSN Online: 2162-2442
ISSN Print: 2162-2434

Pricing Cyber Security Insurance

Zhaoxin Lin, Travis R. A. Sapp, Rahul Parsa, Jackie Rees Ulmer, Chengxin Cao

College of Business, Gerdin Business Bldg. Suite 2330, 2167 Union Drive, Iowa State University, Ames, IA, USA
Email: zxlin@iastate.edu, trasapp@iastate.edu, raparsa@iastate.edu, jrulmer@iastate.edu, ccao@iastate.edu

How to cite this paper: Lin, Z.X., Sapp, T.R.A., Parsa, R., Ulmer, J.R. and Cao, C.X. (2022) Pricing Cyber Security Insurance. *Journal of Mathematical Finance*, **12**, 46-70.
https://doi.org/10.4236/jmf.2022.121003

Received: November 15, 2021
Accepted: January 15, 2022
Published: January 18, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

Abstract

Cybersecurity breaches may be correlated due to geography, similar infrastructure, or use of a third-party contractor. We show how a logistic regression may be used to estimate the probability of an attack where breaches may be correlated among firms up and down the supply chain. We also show how a Poisson regression may be used to estimate the number of records breached. Losses arising from cybersecurity breaches have an unknown distribution. We propose the stock price reaction to a breach as an objective measure of the loss in wealth sustained by the firm due to a breach. This loss measure reflects the immediate and long-term effects of a breach, including reputational effects and other intangible impacts that are otherwise more difficult to quantify. We examine stock returns for 258 cybersecurity breach announcements over 2011-2016 in order to obtain the empirical loss distribution. We find a five-day abnormal return of −1.44%. Seventy-one percent of these 258 announcements result in a negative abnormal return, and a gamma distribution provides an excellent fit to these losses. In addition to introducing a predictive model for correlated losses, our study shows how insurers can use either the empirical stock return distribution of losses or the per record cost of a breach in the pricing of cyberinsurance.

Keywords

Risk Management, Loss Distribution, Cyber Breach, Cyberinsurance, Insurance Premium, Correlated Losses, Event Study

1. Introduction

Hacking incidents and information security breaches in digital networks have risen to the top of corporate and governmental radar screens due to the volume and intensity of such incidents. Cybercrime costs are estimated at $6T worldwide for data breaches alone (Cybersecurity Ventures, [1]). This does not include the cost of ransom ware, denial of service attacks, intellectual property theft, or other types of cybercrime. While cybersecurity was once relegated to the lower levels of the IT enterprise within organizations, with the dramatic rise in incidents having widespread fallout, including increased operational, legal, and compliance costs, and significant negative publicity, cybersecurity now occupies a place of prominence among the top decision makers in organizations.

Cyberinsurance has been gaining increasing acceptance among firms, and this market is growing as more insurers are adding cyberinsurance policies to their offerings. However, estimating the premium continues to be a significant challenge in cybersecurity insurance. This requires that insurers understand the probability of a breach occurring. Estimating this probability is complex and is based on a number of factors. Cybersecurity breaches tend to be correlated among firms up and down the supply chain. The probability of a breach also depends upon the amount of resources the firm deploys in its IT budget. Other factors include whether the firm is diversified in its use of software. We model these factors in order to predict the likelihood of a breach.

Estimating the premium for cyberinsurance also requires that firms, and their insurers, understand the approximate value of their assets, both digital and physical. While the value of a server is straightforward to calculate, the value of the data on that server is much more difficult to estimate. The Ponemon Institute periodically provides estimates on the value of personally identifiable information. For example, in 2008, the estimated cost of a breached record was $202 (Ponemon Institute [2]) and the loss per company in 2016 was reported at $9.5M (Ponemon Institute [3]). As large as these numbers are, they do not cover all data that has value to a firm—for example, marketing plans, strategic planning documents, and other intellectual property such as patent applications and formulas. Also missing is the destruction of mission-critical infrastructure, the disruption of operations, and any reputational damage sustained by the firm, which devalues its trade name, and harms relationships with its customers and suppliers. All of these less tangible effects cause losses by reducing the ability of the firm to generate cash flows, thus reducing its value. Furthermore, the impact from a cyberattack can continue to reverberate over a period of many months or even years as stolen data surfaces and new legal costs are incurred. A 2017 study published by Deloitte [4] finds that the direct costs associated with data breaches are typically relatively small compared to these intangible costs that are more difficult to measure[1]. We provide a unique approach to estimating the total losses due to a cybersecurity breach. Our estimate is based on the abnormal stock market reaction to the announcement of cybersecurity breaches. This measure captures the market's unbiased estimate of the total losses to the firm from a breach.

[1] The study is titled "Beneath the surface of a cyberattack: A deeper look at federal sector impacts." The authors identify seven "above-the-surface" cyber incident costs: technical investigation, citizen/customer breach notification, post-breach citizen/customer protection, regulatory compliance, public relations, attorney fees and litigation, cybersecurity improvements. Other "below-the surface" costs they note are insurance premium increases, increases in the cost to raise debt, and national security losses (for federal agencies).

Firms have little clear guidance on the optimal amount of spending for data security. Gordon and Loeb [5] advise that firms should continue to spend on information security only up to where the marginal benefit is just greater than the marginal cost of the incident, although catastrophic breaches should be accounted for. A modification of this model by Gordon *et al.* [6] considers the impact of a breach in one firm either upon other firms that are connected by network, or upon society at large through damage of critical infrastructure. They estimate that firms routinely underinvest in cybersecurity from a socially optimal perspective when considering the presence of these negative externalities of security breaches in a connected world. Our paper builds on this insight and models the correlated aspect of breaches up and down the supply chain.

Conventional thinking is that cyberinsurance premiums will simply be priced too high for consumers for two primary reasons. First, there is a lack of historical claims data from which to estimate losses. As insurance providers do not want to overly expose themselves to risk, this lack of information will cause premium prices to be too aggressive in order to hedge against this risk. Another issue of concern to insurance companies is that of correlated losses. As specific operating systems and business applications dominate the business computing market, an attack in a previously unknown vulnerability in such a system, a zero-day attack or black swan event, could affect most of an insurer's clients. While reinsurance is a partial response to this problem, an attack that paralyzes wide swaths of the global computing infrastructure would be catastrophic in multiple dimensions. Insuring against such a threat has, at least theoretically, been seen as overly expensive due to the possibility of enormous losses.

Insurance providers now have some limited historical data available on attacks and threats. Various organizations, such as the Ponemon Institute, have created estimates of the direct financial costs of data breaches. Cyberinsurance is gaining wider acceptance, but still does not have the market share that its potential demand would suggest. Do firms still feel that cyberinsurance premiums are overpriced? An equally relevant question is the opposite—are cyberinsurance premiums underpriced? Based on conversations between insurance company managers and the authors, cyberinsurance policies are priced without extensive analysis of potential loss, but rather are priced to be competitive with other providers in the market. It may be observed that, since the cyberinsurance providers are still doing business and there has not been any substantial movement in premium prices, the industry may have achieved some form of equilibrium. However, bubbles in security markets can, for a time, look deceptively stable as well. This becomes largely an empirical question then, as more information is now available to evaluate pricing schemes.

One way in which insurance premiums can be priced is to add up the historical costs of responding to, and recovering from, a cybersecurity breach. The costs of any replacement hardware, software, overtime, consulting and contractor fees, legal fees, marketing, and customer and/or employee outreach, plus any potential business interruption can be totaled and averaged. One potential drawback of relying on only this data is that cybersecurity breaches can have much different impacts depending upon the industry of the firm and the type, magnitude, and duration of the breach. Also, a larger concern is that it is not always possible to distill all costs of a data breach into a direct cost, which is straightforward to verify and reimburse. Many costs are intangible. For example, some costs are reputational, causing a substantial decline in consumer and supplier trust in the firm, and these indirect costs can have the greatest long-run impact on the profitability of the firm.

Our research question is two-fold. First, how should insurance companies price cybersecurity insurance premiums, in an environment where breaches among firms are often correlated, in order to manage their exposure in this market? Second, how can insurance companies offer contracts that cover the total loss to the firm, which includes not only the direct costs of a breach, but the harder-to-estimate intangible costs that harm the value of the firm, such as loss of reputation, disruption of operations, and so forth. In this paper we outline a general framework for pricing cybersecurity insurance in an environment where cyber breaches may be correlated across firms. We model the number of breaches using IT budgets, and we account for the number of breaches occurring in the firm's suppliers and distributors. We also account for the number of vendors of software used by the firm. In our model, we assume that the number of breaches follows a Poisson distribution. We show theoretically how to estimate losses and then we empirically estimate losses based on a Poisson regression. We also estimate the probability of a breach using logistic regression.

An important innovation of our approach is that we present a method to model the total costs due to a breach, which includes both direct and indirect costs to the firm. To our knowledge, ours is the first attempt to capture these intangible costs and encapsulate them into an estimate of the total loss to the firm. Our approach to estimating losses for the purpose of pricing cyberinsurance premiums is drawn from the stock market reaction to cybersecurity breaches. Assuming a reasonably efficient stock market, new information is quickly assessed by investors as to the overall long-run impact on the firm's ability to create value. The shareholders then trade on this information, causing price to rapidly change in response to the perceived impact on firm value. An abnormal negative stock price reaction after the announcement of a cybersecurity breach should encompass the market's best assessment of the impact of the breach to the value of the firm. Therefore, we propose the cumulative abnormal return around the breach announcement from a large sample of firms as an excellent source from which to estimate the distribution of total losses and to provide premium pricing information to cyberinsurance providers.

For pricing cyberinsurance products, we provide a methodology to measure the loss distribution and show how these products may be priced in a world where breaches are potentially correlated among firms. Our predictive model brings together a number of relevant factors that are found to be correlated to a cyber security breach at a firm. We measure the cumulative abnormal returns from publicly-traded firms having a cybersecurity breach announcement in order to gauge the magnitude of losses from a breach. We use this objective and quantifiable measure to model cyberinsurance premium pricing and to assist insurers in their efforts to provide a fairly-priced cyberinsurance product. Our model should also be of value to firms in terms of adding appropriately-priced cyberinsurance to their cybersecurity protection, detection, and recovery processes. Finally, we demonstrate the use of the model through an example firm.

The remainder of this paper is organized as follows. The next section provides the relevant background and literature review on cybersecurity breach costs and cyberinsurance. Section three describes the cybersecurity insurance pricing model. Section four describes our data sample, method, and results for the event study. In Section five we model and empirically estimate both the probability and number of cybersecurity breaches using regression techniques. The last section concludes and provides future directions for research.

2. Background and Related Literature

2.1. Cybersecurity Breach Costs

Cybersecurity breaches can wreak havoc on organizations and their stakeholders. Many researchers have explored the economic costs to firms as a result of cybersecurity breaches. One line of research examines the public market reaction by conducting an event study (Garg *et al.* [7]; Campbell *et al.* [8]; Cavusoglu *et al.* [9]; Acquisti *et al.* [10]; Goel *et al.* [11]; Hovav *et al.* [12]; Gordon *et al.* [13]; Morse *et al.* [14]; Spanos and Angelis [15]; Rosati *et al.* [16]). The underlying assumption in this methodology is that, if financial markets are rational, the disclosure of the data breach will quickly and fully be reflected in the value of the firm. Acquisti *et al.* [10] report a negative, significant decrease in stock prices at the announcement of data breaches, an effect which accumulates over several days and then dissipates in the following months. Cavusoglu *et al.* [9] link the drop in stock value at the announcement of a breach to several characteristics such as firm type, firm size, and the year the breach occurred to help explain the cross-sectional variation in abnormal returns produced by security breaches.

The literature also reports different types of data breaches result in different market responses. For example, Hovav *et al.* [12] find that the market does not penalize companies that experience a denial-of-service (DOS) attack. They find that only "internet-specific" companies' market prices are sensitive to the news of a data breach, but not others. Morse *et al.* [14] report that different types of security breaches can have a disparate impact on a firm's value, as the market seems to penalize breaches that could have been avoided with reasonable precautions by the affected company. Overall, through an event study approach, these studies show that information security breaches tend to adversely affect the market value of the affected firms.

2.2. The Market for Cyberinsurance

Information security, or more broadly, cybersecurity, is a multi-faceted challenge. A fundamental problem is that security is often built into technologies as an afterthought, instead of being developed as part of the initial design process. This applies to nearly all the technology used within and across organizations, including purchased systems, open-source systems, and custom systems.

There are many reasons for this lack of built-in security. There are significant time-to-market pressures on software and other technologies. This pressure results in less time for robust quality assurance processes, which, in theory, could result in fewer vulnerabilities or defects. Consumers have typically rewarded firms for increased functionality over security. Increased functionality leads to greater complexity, in turn, leading to a higher probability of errors or defects in the software. Additionally, vendors often will prioritize ease-of-use over security, particularly when it comes to new installations, leaving to the consumer to adjust security settings to a higher and more appropriate level of protection. Unfortunately, these adjustments are often overlooked in the installation process, leaving systems critically vulnerable to attacks from both outside and within the organization.

Firms have mostly moved to a risk management approach for information security management, as it simply is unsustainable, if not impossible, to secure against all information security incidents. Professional associations, government agencies, and academic researchers have suggested various frameworks for managing information security risk, but most consist of a basic risk assessment process of identifying assets, estimating the value of those assets, estimating the likelihood of loss to those assets, and then allocating resources to protect the assets in line with the annualized loss expectancies. This process is known as an Annualized Loss Expectancy (ALE) approach.

Zero-day exploits, issues with deriving appropriate probabilities and likelihoods, and situational complexity all contribute to the challenges with the approach. While the ALE process and the various frameworks all have flaws, they do provide valuable insight into managing the risk associated with cybersecurity threats. There are also many ways in which firms approach this problem, including looking deeper at-risk management outcome options. Once firms have an appreciation for the risk presented by their technology assets, they can opt to manage their risk through accepting risks, transferring risks, and mitigating risks. Firms have traditionally focused on mitigating or reducing risk through various information security technologies such as anti-virus software, firewalls, intrusion detection and prevention systems, access control systems, penetration testing, backup and recovery systems, disaster recovery planning, and other technologies and strategies. Firms also accept substantial risks related to information security activities, both explicitly and implicitly. Firms are more likely to disclose the acceptance of some level of information security risk through their financial reporting mechanisms, in the voluntary risk disclosure section of 10-K reports (annual reports) to the Security Exchange Commission (SEC) for publicly-traded firms in the US (20-F for firms whose stock is traded on international exchanges). Frequently, firms implicitly accept risks by hoping that they will not fall victim to an attack or by denying that they are at risk. Unfortunately, this implicit acceptance does not absolve the firm of responsibility for attacks or other information security crises.

An area of increasing attention is the transference of risk for information security concerns to other parties. This transference may be to outside contractors for management of various components of the technology infrastructure with the goal of improving information security, as well as purchasing insurance policies to offset potential costs of increasingly likely information security breaches.

The potential for the use of cybersecurity insurance primarily comes from the framework provided in Gordon, Loeb, and Sohail [17]. Their framework essentially consists of assessing risk, mitigating risk via controls, transferring the residual risk via insurance, and then maintaining a steady state of acceptable risk. Several issues quickly arose in the application of this framework, notably the availability, coverage, and pricing of cyberinsurance premiums.

Kesan, Majuca, and Yurcik [18] examine the cyberinsurance market and show the social welfare increases with the development and maturation of the cyberinsurance market. Bohme [19] argues that the usage of a few dominant infrastructure platforms, such as the Microsoft enterprise technology stack, leads to correlated losses from information security incidents, which are difficult to insure against. Therefore, cyberinsurance policies would be priced to reflect such losses leading to high premium prices. Ogut, Menon, and Raghunathan [20] also argue that the interdependent nature of firms' IT security infrastructure have implications for the development of a mature cyberinsurance market, ultimately affecting insurance premium pricing. Zhao and Xue [21] suggest a "captive insurance" approach to managing insurance costs and to address the information asymmetry and incentive compatibility concerns. Such "captive insurance" would essentially be provided by a consortium of firms that collectively self-insure, which is hypothesized to provide greater incentive alignment and information sharing than would occur with a third-party insurance provider. Zhao, *et al.* [22] explore two new arrangements: risk pooling agreements and managed security services, which attempt to overcome issues with correlated losses. They report that risk pooling agreements can serve as a complement to cyberinsurance.

From the insurance consumer perspective, Bandyopadhyay, Mookerjee, and Rao [23] show that under certain circumstances information asymmetry causes cyber-insurance products to be over-priced and therefore not consumed. However, Srinidhi, Yan, and Tayi [24] examine the potentially positive role of insurance in reducing the probability of financial distress among firms in light of potential information security breaches.

Cyberinsurance typically covers losses due to cyberattacks and other incidents such as malware, viruses, and user error. Such policies also typically include coverage for ensuing legal and compliance fees, costs of identity theft monitoring, and even ransom payments. Most major insurers are offering some type of cyberinsurance policy product, with many being policies resold from Hartford Steam Boiler, a division of Munich Re.

While many in the cybersecurity community equate recovery with incident response, disaster recovery, and business continuity, the availability of cyberinsurance is an additional tool for managing information and digital risk. Cyberinsurance can cover losses related to cyberattacks and incidents, and is becoming widely available. There are two possibilities with the current level of premiums. One is that insurers overprice premiums in order to cover the poorly understood risks from correlated attacks and catastrophic breaches. In fact, it may be that the premiums are priced too low, because insurance companies want to gain market share in this newer market, but have not yet had to pay out claims at a level that would cause great concern of overexposure to losses.

Pricing of cybersecurity premiums and coverage limits should, in theory, be the result of actuarial analysis of risks based on historical data and calculated projections. However, given the challenges inherent in estimating the true costs of cybersecurity breaches and the factors that lead to higher amounts of risk, estimating the premium remains much more art than science. Currently, cyberinsurance policy premiums are usually priced in accordance to similar policies in the marketplace. Insurers may also conduct a subjective analysis of IT readiness and competence as well as overall corporate culture in a semi-customized approach to pricing. Insurers are wary of overpricing premiums relative to competitors, given the relative immaturity of the market. Also, cybersecurity claims apparently have not been so onerous that insurers are limiting or refusing coverage to certain customers, such as occurs with insurers limiting or refusing to write policies to homeowners who own coastal properties in hurricane-prone areas.

3. Cyberinsurance Pricing Model

We present two models for estimating losses from an attack. The first estimates total losses based on the public stock market response. The second model provides aggregate loss estimates based on each breached record using the traditional actuarial model. The distinction between these estimates is in how the losses are estimated and what they cover. The first loss measure estimates the immediate and long-term effects of a breach, including reputational effects and other intangible impacts that are otherwise more difficult to quantify. The aggregate loss model estimates the total claim amount made to the insurance company.

3.1. Method 1: Total Loss Model

In this method, we directly estimate the total loss to the company since, for an insurance company, the interest is ultimately in the amount the insurer must pay if a breach occurs, and not on the number of records breached or individual losses. Here, we propose an innovative method to assess the total loss to the company using the event study CAR. This abnormal stock loss to the shareholders reflects the decline in value to the firm due to the full perceived impact of the cyber breach, reflecting both direct costs and intangible costs.

The announcement of a cybersecurity breach typically has a significant impact on the firm's stock price. This cumulative abnormal return (CAR) can provide a reference point for cybersecurity premium pricing and planning for limits on claims. Let X be a Bernoulli random variable with $X = 1$ denoting a breach and $X = 0$ denoting no breach. Let CAR denote the amount of loss if a breach occurs. Then the total loss to the insurance company is given by

$$\text{The expected total loss} = E(\text{Total Loss}) = p * E(\text{CAR}) \tag{1}$$

and the variance of the total loss is given by

$$Var(\text{Total Loss}) = p * Var(CAR) + p * (1-p) * E(CAR)^2 \tag{2}$$

where $p = P(X = 1)$.

In Sections 4 and 5 that follow, we describe in detail the models to estimate the CAR and its distribution. In Sections 6.1 and 6.2, we describe in detail the model to estimate $p = Pr(X = 1)$.

3.2. Method 2: Aggregate Loss Model

Alternatively, we will assume the collective risk model, a commonly used model in the insurance industry, gives the aggregate loss to the company. Let N denote the number of records breached. For each breached record, let the random variable U denote the loss to the insurance company for an individual claim, *i.e.*, the claim amount for each breached record. Unlike the market CAR, this random variable, U, measures only the claim amount and does not include intangible losses such as reputational effects and associated value loss from the breach. Since the losses are positive, we will model them using a positively skewed distribution. Note: We follow standard practice and model $N \sim$ Poison (λ). Then, the aggregate loss to the company is given by

$$E(\text{Aggregate Loss}) = E(N) * E(U) \tag{3}$$

and the variance of the total loss is given by

$$Var(\text{Aggregate Loss}) = E(N) * Var(U) + Var(N) * E(U)^2. \tag{4}$$

Since $N \sim$ Poisson (λ) which means $E(N) = Var(N) = \lambda$. Equation (4) can be rewritten as

$$Var(\text{Aggregate Loss}) = \lambda * \left(Var(U) + E(U)^2\right). \tag{5}$$

An ideal distribution for the individual losses, U, would be Gamma or Pareto since losses are positively skewed, generally with a heavy tail. In this paper, we will use the Gamma distribution to model U. While we do not have data to model U, in Section 5, we present a rationale to infer this distribution and estimate the parameters of the Gamma distribution (in the real world, insurance companies will have the data to model this). In Section 6.3, we present the model to estimate the number of breaches.

4. Estimating Total Losses

4.1. Event Study Data Sample and Method

For the event study, which we use to estimate total losses, we examine cyber breaches occurring during the years 2011 to 2016. Our initial list of breaches is obtained from Privacy Rights Clearinghouse (PRC, https://www.privacyrights.org/data-breaches), which provides information about breach type, affected population and files, public announcement date, and gives an incident description. PRC is a nonprofit organization which has collected a comprehensive list of publicized data breaches involving loss of personally identifiable information (e.g., social security numbers, bank account information, emails, driver's license numbers, and medical information). PRC uses various sources including attorney generals' offices, media announcements, government agencies, and other discontinued databases such as DataLossDB since 2005. The initial list contained a total of 4787 breaches, and most of these breached firms were privately owned. We removed all firms which were not publicly traded. For each remaining observation in the dataset, we searched the Internet for any possible public announcements of the breach earlier than the date listed by Privacy Rights Clearinghouse, but found most dates to be accurate. For each breached firm, we also searched for possible confounding events around the cyber breach announcement date, such as announcements of earnings or takeovers. Any sample observation with a confounding announcement within the event window was removed. In addition, after analyzing the breach descriptions, we removed the following non-relevant events: 1) encrypted data loss, 2) not cyber-related (for example, payment drop-box broken into), 3) not a breach (initial report later contradicted), 4) confounding event (lawsuit over trade secrets theft). After applying these filters, the final sample contains 258 incidents.

Table 1 provides some selected characteristics of the breached firms in our sample. The average market capitalization of a breached firm is $56 billion, while the median firm has an equity value of only $18 billion. The average breached firm has $10 billion in cash on hand, while the median firm has only $2 billion. The prior year's stock performance on average is 15.3%, for the sample of breached firms, though the bottom quartile had negative stock performance. Figure 1 shows the breakdown of the sample by industry. We note that finance and insurance related firms are the most frequent targets, comprising nearly a fourth of our sample firms. Companies involved in software and Internet businesses are tied for second. It is worth noting that hacking accounts for nearly half of all of the breaches, followed by unintended document loss. Insider breaches and stolen portable device breaches rank third and fourth, respectively.

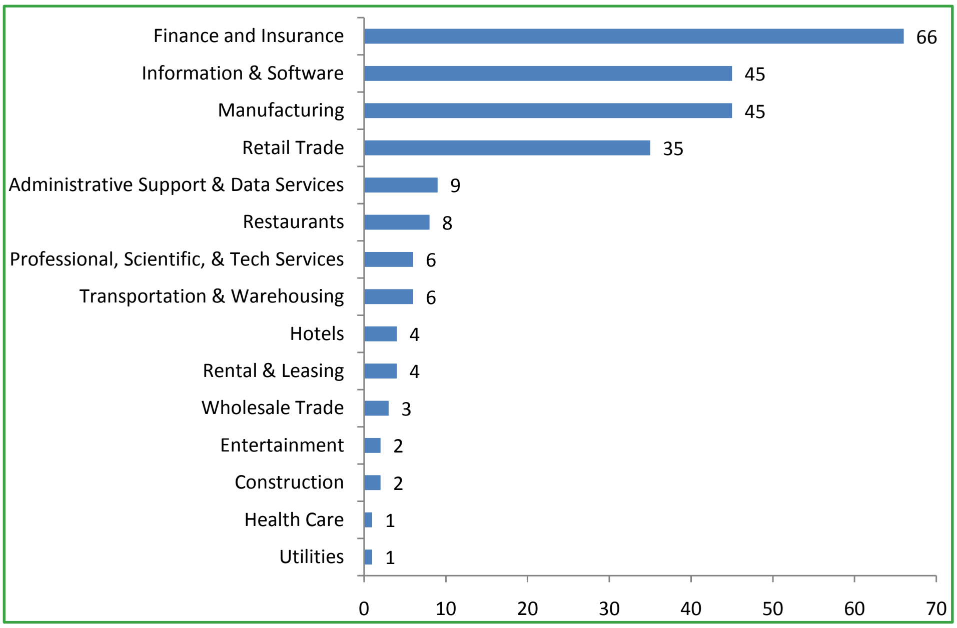

Figure 1. Industries of sample firms. The figure shows the industry for 237 of the 258 sample firms that suffered cyber breaches during 2011-2016 and for which a North American Industry Classification System (NAICS) code was available in COMPUSTAT. The classification summarized and reported here is based on the first two digits of the NAICS code.

Table 1. Sample description.

	Market Cap ($M)	Sales ($M)	Profit Margin (%)	ROE (%)	Cash ($M)	Prior Year Stock Return (%)
Mean	56,443	46,338	5.57	12.58	10,185	15.28
Standard Deviation	91,030	66,340	0.53	0.54	22,926	31.30
25th Percentile	3756	3197	3.14	6.52	334	−1.74
Median	18,736	17,902	7.06	12.14	1979	14.33
75th Percentile	74,120	72,312	16.69	21.49	8620	32.45

The table reports characteristics for the sample of 258 firms publicly announcing cybersecurity data breaches over the years 2011-2016. Some firms appear in the sample more than once.

To perform the event study and calculate the abnormal stock return at the breach announcement, we use the capital asset pricing model (CAPM) and the CRSP value-weighted market index. Daily stock closing prices and number of shares outstanding are from CRSP. We focus on the cumulative abnormal returns (CARs) over a five-day [−2, +2] event window surrounding the announcement, but also report results for a shorter three-day event window [−1, +1].

4.2. Event Study Results

The event study results are presented in Table 2. The average decline in stock price over a three-day window surrounding the announcement is 1.18%, with a

Table 2. Abnormal returns from cybersecurity breach announcements.

	3-day	5-day	41-day
Mean CAR (%)	−1.18	−1.44	−1.44
	(−7.30)	(−6.75)	(−2.53)
Median CAR (%)	−0.79	−1.01	−1.10
	(7.40)	(7.52)	(2.64)
Mean CAR ($ mil)	−446.78	−586.69	−1129.67
Median CAR ($ mil)	−58.00	−77.24	−58.16
# of positive CARs	76	75	109
% positive CARs	29%	29%	42%

The table reports cumulative abnormal returns (CARs) from a sample of 258 public announcements of cybersecurity data breaches over the period 2011-2016. Abnormal returns are reported for three windows surrounding the announcement day: [−1, +1], [−2, +2], and [−10, +30]. T-statistics are in parentheses.

t-statistic of 7.30. Stock price falls 1.44% over the five-day window surrounding the breach announcement, with a *t*-statistic of 6.75. In order to gauge the economic impact of the losses, we convert percent returns into dollar values by multiplying each 5-day CAR by the market capitalization of the firm on the day of the announcement. The average dollar loss due to a cyber breach announcement is $587 million. There are some very large firms in our sample, which skews the dollar distribution to the right, so the median dollar loss is smaller, but still substantial, at $77 million. Interestingly, we note that 29% of the CARs are positive around the announcement, indicating that some types of breaches are either considered insignificant by the market, have already been impounded into price, or do not show up in the measurement due to confounding effects of other unidentified information arriving simultaneously.

To gain a longer-term perspective on the impact of cyber breach announcements, we trace the cumulative abnormal returns starting from 10 days before the announcement to 30 days afterwards. The results are summarized in Figure 2. The plot shows a large drop at the announcement and then a relatively flat graph extending to the right. The average impact on stock price over [−10, +30] is a CAR of −1.44% (this happens to coincide with the 5-day CAR value). We conclude that there is no over-reaction or reversal and the loss in value appears to be permanent.

Due to the enormous intangible component and the associated lack of actuarial data, total losses arising from a cybersecurity breach have an unknown distribution. Therefore, we propose to shed light on the distributional form of these losses by using the stock market impact at the announcement for our large sample of firms. Out of 258 breach announcements, 71% suffer some amount of abnormal decline in value. We condition our analysis on these 183 firms that had a negative stock price change. Figure 3 displays the empirical frequency distribution of the losses, to which a gamma distribution is an excellent fit. The fitted

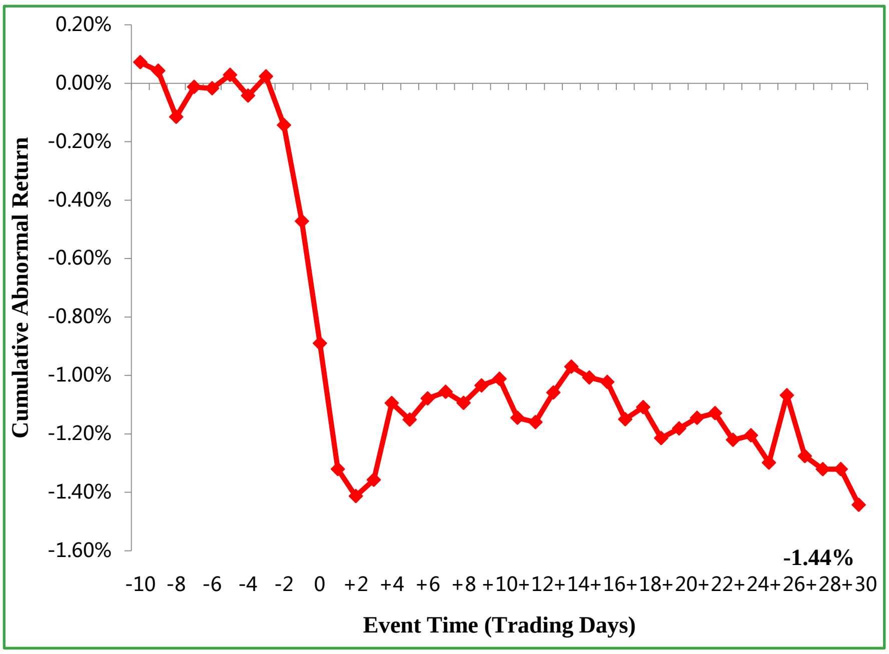

Figure 2. Cybersecurity Breach Announcement Cumulative Abnormal Returns. The figure shows the daily cumulative average abnormal return for all 258 sample firms in event time over the window [−10, +30]. Day 0 is when the cybersecurity breach was publicly announced.

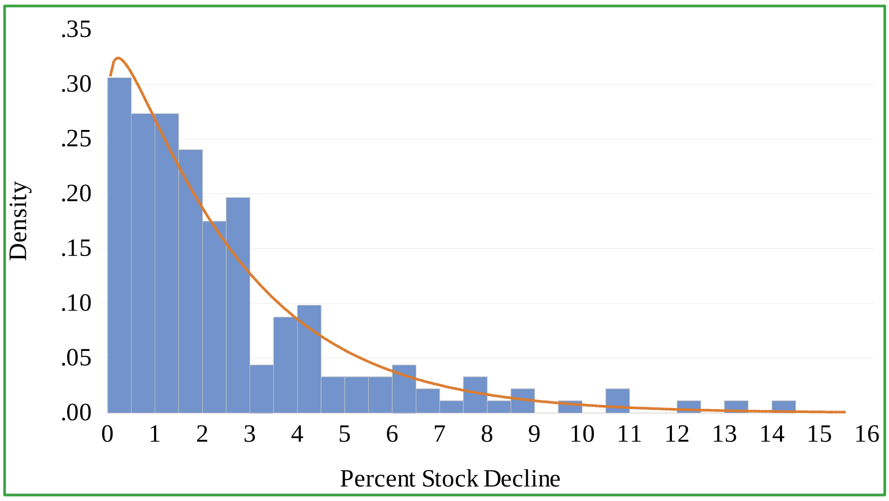

Figure 3. Empirical total loss distribution with gamma fit. The figure shows the empirical frequency distribution of the 183 negative 5-day CARs, displayed here as positive losses, with a gamma distribution fit to the data and overlaid onto the empirical frequency plot. The fitted gamma has parameters $\alpha = 1.09$ (z-stat = 10.78) and $\theta = 2.37$ (z-stat = 8.57) and a log likelihood value of −356.19.

gamma has parameters $\alpha = 1.09$ (z-stat = 10.78) and $\theta = 2.37$ (z-stat = 8.57). We test the hypothesis of a gamma distribution using three common distributional tests: Cramer-von Mises, Watson, and Anderson-Darling. All three tests fail to reject the hypothesis that the losses follow a gamma distribution, each showing p-values in excess of 0.25. Although our paper relies upon the empirical distribution of CARs to estimate total losses, there is an important implication of the gamma fit, which we highlight in the modeling section below. Specifically, our research validates the use of a gamma distribution to model individual firm total losses. In addition, this also confirms the distribution for individual losses, U (used in Section 3). This is true because if Individual Losses = U_i ~ Gamma (α_i, θ) then CAR = Total Losses ~ Gamma $(\sum \alpha_i, \theta)$. Based on this fact, we not only validate the distribution of U but will also estimate its parameters.

5. Modeling Data Breaches

5.1. Data and Method

In this section, we turn to our empirical analysis using actual data of firm breaches to predict the likelihood of a cybersecurity breach. This is a key piece of information as an input to the pricing model. We describe how we model data breaches in the firm, including the likelihood of a breach, and the number of records that are breached. We use multiple data sources, which include data on actual data breaches as well as firm IT structure and IT budget, in order to construct our models.

We study the impact of data breaches that have occurred in the firm's value chain in the past three years on the likelihood of data breaches occurring in the firm. The value chain includes customers (downstream industries) and suppliers (upstream industries). First, we define a focal firm's downstream and upstream industries using the Bureau of Economic Analysis (BEA) Input-Output (IO) Use table. Using the IO Use table from BEA, we construct a continuous Vertical Relatedness Index (VRI) for every industry pair. For each industry pair i and j, the VRI is defined as the dollar value of input from industry i in order to produce one dollar of industry j's output. Suppose the focal firm belongs to the IO industry i. Any industry j that generates a VRI higher than 5% is considered a downstream (*i.e.* customer) industry to the focal firm in industry i (Fan and Lang [25]). Similarly, suppose the focal firm belongs to the IO industry j. Any industry i that generates a VRI higher than 5% is considered an upstream (*i.e.* supplier) industry to the focal firm in industry j.

Second, we calculate the number of data breaches in each industry-year. We obtain data breaches that have occurred from 2010 to 2018 from the Privacy Rights Clearinghouse (PRC) website (https://www.privacyrights.org). In total, we are able to match 937 data breaches to Compustat firms. These data breaches are further aggregated to industry (using each Compustat firm's primary industry) and year level. We use the number of breaches that have occurred in the past three years in the downstream industries and upstream industries as our independent variables to estimate the impact of prior data breaches that have occurred in the value chain on the data breach likelihood in the focal firm using logistic regression.

The number of compromised records is a useful indicator of the severity of a breach. This is also typically the basis on which insurance companies will pay out in the event of a data breach. Accordingly, we also aggregate the number of compromised records in the downstream industries as well as in the upstream industries in the past three years to estimate the number of compromised records in the focal firm year. We estimate the number of breached records according to a Poisson regression model.

The amount of IT investment at a firm is correlated with the firm's ability to repel an attempted cyber breach. Therefore, we investigate the impact of IT investment on the likelihood of data breaches occurring at the focal firm in our regression models. As an independent variable, we include per employee IT budget, and we also explore different categories of the firm's IT budget: communication, hardware, service, software, storage IT budget. Other explanatory variables include the number of software vendors used at the focal firm, which indicates how well-diversified the firm is in its software exposure, and the cross-sectional standard deviation of the IT budget across different sites in the firm, which indicates imbalance in allocating resources.

All IT investment related information comes from the Harte-Hanks CI (HH) database. The HH database is the most comprehensive data source regarding enterprise IT investments. This data source is used by many published information systems research papers (Dewan and Ren [26], Jia *et al.* [27]). The HH database from 2010 to 2018 provides information technology related information on 743,735 unique sites (different physical locations), covering 37,792 5-digit zip codes. Specifically, the HH database provides the total IT budget (in USD), size (employment and revenue), physical location (zip code), and SIC code for each reported site. HH data is collected yearly by interviewing each site's IT managers.

5.2. Predicting Firm Breaches Using Logit

We model firm cyber security breaches using a logistic regression, where the dependent variable, Firm Breach, takes the value of 1 in the case of a breach and 0 otherwise. Our model is given by

$$\begin{aligned}\text{Firm Breach} = \alpha &+ \beta_1(\text{Up/Downstream Breaches}) + \beta_2(\text{IT Budget/Emp}) \\ &+ \beta_3(\text{Software Vendors}) + \beta_4(\text{Std Dev}) + \Gamma'(\text{Controls}) \\ &+ \text{Time F.E.}\end{aligned} \tag{6}$$

where Up/Downstream Breaches is the number of breaches in the last three years either in the firm's supplier industries or in the firm's consumer industries, IT Budget/Emp is the firm's IT budget on a per employee basis, Software Vendors refers to the number of vendors of software used by the firm, Std Dev refers to the standard deviation of the IT budget across sites in the focal firm. We use a standard set of controls based on firm financial variables. We control for firm size (market capitalization), growth prospects (Tobin's Q), profitability (ROE), sales growth, liquidity, D/E ratio, M/B ratio, P/E ratio, and cash flow. The financials are lagged by one period. We also include year fixed effects in our models, controlling for time trends.

Results from the logistic regressions are presented in **Table 3**. We see in Model I that an upstream breach in the firm's supplier industries within the prior three years has a significant effect on the probability that the firm gets breached. Also, as expected, with an increase in the firm's overall IT budget per employee, the probability of a breach goes down. When there are multiple software vendors to the firm, the probability of a breach increases. So, diversifying software vendors in an effort to minimize downtime may not be a helpful strategy. The probability of a breach increases with an increase in the volatility of the IT budget across sites at the firm. This suggests that some sites may form a weak link in the chain. From the financial control variables we see that firms that are large, high-growth, and with high cash flow, tend to have a higher probability of being breached by a cyber attack.

In Model II we replace the firm's overall IT budget as an explanatory variable with multiple categories of specific IT spending. We find that higher spending on software, and storage is linked to a greater likelihood of a breach. Spending on services leads to a lower probability of a breach. Higher spending on communication and hardware makes no difference.

In Model III we find that a downstream breach in the firm's consumer industries within the prior three years has a significant effect on the probability that the firm gets breached. The firm's overall IT spending again loads negatively on the probability of a breach, and is highly significant. The number of software vendors as well as the standard deviation of the IT budget across sites at the firm are both positive linked to the probability of a breach. When looking at the expanded categories of IT spending in Model IV, the only significant coefficient is on spending for storage. This is consistent across Models II and IV and seems to suggest that firms with greater data storage needs are more prone to cyber breaches.

Table 3. Logistic regressions explaining firm breaches.

	I	II	III	IV
Upstream breach	0.0343***	0.0352***		
	(0.01)	(0.01)		
Downstream breach			0.0190***	0.0184***
			(0.00)	(0.00)
Overall IT budget	−2.47e-06***		−2.44e-06***	
	(0.00)		(0.00)	
Comm budget		−0.0000701		−0.0000869
		(0.00)		(0.00)
Hardware budget		−0.000098		−0.0000825
		(0.00)		(0.00)

Continued

Services budget		−2.51e-05*		−0.0000194
		(0.00)		(0.00)
Software budget		3.49e-05*		0.00003
		(0.00)		(0.00)
Storage budget		0.00173***		0.00154***
		(0.00)		(0.00)
Software vendors	0.0699***	0.0698***	0.0717***	0.0716***
	(0.01)	(0.01)	(0.01)	(0.01)
Std dev of site IT budget	0.0536**	0.0644***	0.0489**	0.0588***
	(0.02)	(0.02)	(0.02)	(0.02)
ROE	−0.00257	−0.00295	−0.00199	−0.00221
	(0.02)	(0.02)	(0.02)	(0.02)
Sales growth	0.617**	0.700***	0.669***	0.749***
	(0.25)	(0.24)	(0.24)	(0.24)
Liquidity	−0.153**	−0.154**	−0.134*	−0.136*
	(0.08)	(0.07)	(0.07)	(0.07)
Debt/Equity ratio	0.00254	0.00272	0.00208	0.00219
	(0.00)	(0.00)	(0.00)	(0.00)
Market to book ratio	0.000288	0.000207	0.000474	0.000415
	(0.00)	(0.00)	(0.00)	(0.00)
PE ratio	0.000429*	0.000435*	0.000410*	0.000415*
	(0.00)	(0.00)	(0.00)	(0.00)
Market Cap	8.54e-06***	8.98e-06***	8.25e-06***	8.63e-06***
	(0.00)	(0.00)	(0.00)	(0.00)
Tobin's Q	0.0139	0.0147	0.0208	0.0219
	(0.04)	(0.04)	(0.04)	(0.04)
Cashflow	0.169***	0.159***	0.172***	0.161***
	(0.04)	(0.04)	(0.04)	(0.04)
Constant	−4.508***	−4.563***	−4.438***	−4.476***
	(0.13)	(0.13)	(0.12)	(0.13)
R-square	0.100	0.105	0.095	0.100
Observations	12,543	12,543	12,543	12,543

The table displays results of logistic regression where the dependent variable takes the value 1 if the firm had a cyber security breach in a given year, and a 0 otherwise. Up/Downstream Breach is the number of breaches in the last three years either in the firm's supplier industries or in the firm's consumer industries, Overall IT Budget is the firm's IT budget on a per employee basis, Software Vendors is the number of vendors of software used by the firm, Std Dev is the standard deviation of the IT budget across sites in the focal firm. Financial controls are lagged one period.

Overall, we learn that no company is an island; what happens in the surrounding industries matters. When your suppliers or consumers get breached, it is much more likely that your firm will get breached. This is an important finding that establishes the correlated nature of cyber breaches empirically. Spending on IT also matters; the more you spend per capita, the less likely you are to get breached—unless it is spending for data storage. Diversification across software vendors at your organization only serves to increase the chance of a breach; this gives more opportunities for weaknesses to be exploited. Finally, larger, high-growth firms with abundant cashflow tend to be the victims of successful breaches. They are likely more frequently the target of attacks. We will use Model I to estimate the probability of a breach which in turn will be used to estimate the total losses, using Method 1 of Section 3.1.

5.3. Predicting Number of Records Breached Using Poisson Regression

We next want to predict the number of records breached in a cyber security attack of a firm. We model the number of records breached using a Poisson regression, where the dependent variable, Records, is the number of records reported by the company as being breached in the attack. Our model is given by

$$\begin{aligned}\log(\text{Records}) = \alpha &+ \beta_1(\text{Up/Downstream Breached Records}) \\ &+ \beta_2(\text{IT Budget/Emp}) + \beta_3(\text{Software Vendors}) \\ &+ \beta_4(\text{Std Dev}) + \Gamma'(\text{Controls}) + \text{Time F.E.}\end{aligned} \tag{7}$$

where the independent variables are the same as described for the model in (6) above.

Results from the Poisson regressions are presented in Table 4. In Model I we see that more breached records in the supplier industries within the last three years leads to fewer records being breached at the firm. This may indicate that, though the firm has fallen victim to a breach, it has taken appropriate steps to

Table 4. Poisson regressions explaining number of records breached.

	I	II	III	IV
Upstream breached records	−5.87e-07***	−5.31e-07***		
	(0.00)	(0.00)		
Downstream breached records			−8.02e-08***	−5.19e-08***
			0.00	0.00
Overall IT budget	4.87e-07***		5.28e-07***	
	(0.00)		(0.00)	
Comm budget		1.63e-05***		1.73e-05***
		(0.00)		(0.00)
Hardware budget		0.000378***		0.000401***
		(0.00)		(0.00)

Continued

Services budget		−4.68e-05***		−4.75e-05***
		(0.00)		(0.00)
Software budget		−6.99e-05***		−7.51e-05***
		(0.00)		(0.00)
Storage budget		0.000376***		0.000327***
		(0.00)		(0.00)
Software vendors	0.0761***	0.0774***	0.0810***	0.0840***
	(0.00)	(0.00)	(0.00)	(0.00)
Std dev of site IT budget	0.0734***	0.0867***	0.0722***	0.0858***
	(0.00)	(0.00)	(0.00)	(0.00)
ROE	0.0468***	0.0475***	0.0476***	0.0485***
	(0.00)	(0.00)	(0.00)	(0.00)
Sales growth	1.535***	1.548***	1.541***	1.557***
	(0.00)	(0.00)	(0.00)	(0.00)
Liquidity	−0.0286***	−0.0270***	−0.0416***	−0.0398***
	(0.00)	(0.00)	(0.00)	(0.00)
Debt/Equity ratio	0.0192***	0.0192***	0.0194***	0.0194***
	(0.00)	(0.00)	(0.00)	(0.00)
Market to book ratio	−0.00647***	−0.00647***	−0.00654***	−0.00654***
	(0.00)	(0.00)	(0.00)	(0.00)
PE ratio	3.68e-05***	−0.000140***	5.15e-05***	−0.000118***
	(0.00)	(0.00)	(0.00)	(0.00)
Market Cap	4.12e-06***	5.44e-06***	3.95e-06***	5.10e-06***
	(0.00)	(0.00)	(0.00)	(0.00)
Tobin's Q	0.135***	0.136***	0.137***	0.138***
	(0.00)	(0.00)	(0.00)	(0.00)
Cashflow	−0.440***	−0.429***	−0.490***	−0.480***
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	10.73***	10.62***	10.67***	10.55***
	(0.00)	(0.00)	(0.00)	(0.00)
R-square	0.191	0.203	0.192	0.205
Observations	12,418	12,418	12,418	12,418

The table reports results of regressions where the dependent variable is the number of data records breached in a firm year. Up/Downstream Breached Records are the number of breached records in the last three years either in the firm's supplier industries or in the firm's consumer industries, Overall IT Budget is the firm's IT budget on a per employee basis, Software Vendors is the number of vendors of software used by the firm, Std Dev is the standard deviation of the IT budget across sites in the focal firm. Financial controls are lagged one period.

safeguard as much data as possible because of attacks in related industries. Increased spending on IT per employee is correlated with more records breached. The direction of this variable is unexpected and seems counterintuitive. It may be that larger firms spend more per employee on IT and also have more data records at risk. This is consistent with the market capitalization coefficient that shows that large firms have more records breached. A larger number of software vendors to the firms leads to more records breached. Higher volatility in the IT budget across firm sites is also correlated with more records breached. We see from the financial controls that large, high-growth, more profitable firms that have less cash flow tend to suffer more breached records.

In Model II we use specific categories of the IT budget. We find that more spending on communication, hardware, and storage lead to an increase in number of breached records. Spending on services and software are correlated with fewer breached records. In Model III more breached records in the downstream industries within the last three years leads to fewer records being breached at the firm. Similar to our finding for upstream breaches, recent downstream breaches may be a tipoff to the firm to better protect its data records based on a perceived threat. Overall IT spending correlates with more breached records. We suspect this is due to large firms both spending more per capita and having more data records at risk. A larger number of software vendors to the firms and higher volatility in the IT budget across firm sites are also correlated with more records breached. Model IV uses specific categories of IT spending and these results mirror those of Model II.

Overall, our model for predicting the number of records breached is quite satisfactory. In Model I, for example, all of the coefficients are significant at the 1% level and the R-squared is 0.191. The coefficients generally make sense. In the case of per capita IT spending leading to more breached records, we believe this is an artifact of firm size. Also, IT spending is more nuanced as Model II illustrates with the breakdown among categories. All coefficients in Model II are significant at the 1% level and the R-squared is 0.203. We will use this model in Method 2 from Section 3.2 to illustrate our pricing model.

5.4. A Numerical Illustration

We will present an example to predict the losses using the two methods presented in Section 3. Consider the firm Amdocs Limited (ticker: DOX). It has the following attributes: upstream industries had four data breaches in the past three years with 41,803 compromised records, the per employee IT budget was $8,477, the firm uses four different software vendors, the normalized standard deviation of site per employee IT budget is 0.47, the ROE is 0.0642, sales growth is 0.0215, liquidity is 0.21, D/E ratio is 0.00034, M/B ratio is 1.767, PE ratio is 14.159, market cap is $5.359B, Tobin's Q is 1.526, cash flow is 0.699. We will use the logistic regression results of Model I in Table III for illustration. These data result in a predicted probability of a breach of 0.0196. We further adopt the Poisson regression results of Model II in Table IV and obtain a predicted number of breaches of 58,964.

We first apply the Total Loss Model of Equations (1) and (2) for illustration. This was labeled Method 1. In Section 4, we showed that CAR ~ Gamma (α = 1.09, θ = 2.37). That implies,

$$E(\text{CAR}) = 1.09 * 2.37 = 2.5833$$

$$Var(\text{CAR}) = 6.122$$

$$E(U^2) = 12.8$$

Now applying Method 1 given in Equations (1) and (2) in Section 3.1, we get

$$E(\text{Total Loss}) = p(\text{Breach}) * E(\text{CAR}) = 0.0196 * 2.5833 = 0.051$$

$$\begin{aligned} Var(\text{Total Loss}) &= p * Var(\text{CAR}) + p * (1-p) * E(\text{CAR})^2 \\ &= 0.0196 * 6.122 + 0.0196 * (1 - 0.0196) * 2.58^2 \\ &= 0.248235 \end{aligned}$$

We have to adjust the above values to the market capitalization of the firm to put them into dollar values. The current market capitalization of this company is \$10.34 billion. Using this amount, the expected loss and variance are:

$$E(\text{Total Loss}) = 0.051 * 10.34 = \$523541911$$

$$Var(\text{Total Loss}) = 0.248235 * 10.34^2 = \$26540213805$$

$$\text{Std Dev}(\text{Total Loss}) = \$162912$$

We next use Equations (3) and (4) of the Aggregate Loss Model described in Section 3.2 for illustration (this was labeled Method 2). From the results shown in Section 4.2 and **Figure 3**, we observed that a gamma distribution provides a good fit to the total losses. This provides empirical validity to the use of a gamma distribution assumption for the individual losses.[2] We will assume U ~ Gamma Distribution (α = 8, θ = 25). From the IBM Ponemon Institute study, the average cost of a breached record is about \$200, with substantial variation. The parameters we have chosen are calibrated to reflect these characteristics. Applying these to the Aggregate Loss Model from Section 3.2, we get

$$E(U) = 8 * 25 = \$200$$

$$Var(U) = 25 * 200 = 5000$$

$$\text{Std Dev}(U) = 70.71$$

Using the set of independent variables given above for the example firm, Table IV Model II gives us the estimate of $\lambda = E(N) = 58964$. Accordingly, we have

$$E(\text{Aggregate Loss}) = E(N) * E(U) = 58964 * 200 = \$11792800$$

and the variance of the total loss is given by

[2] Although not a rigorous proof, this inference is based on the property that the sum of gamma random variables is also gamma distributed. Hence, if we observe that the total losses are gamma distributed, we may reasonably infer that the individual loss distributions for each firm are gamma.

$$
\begin{aligned}
Var(\text{Aggregate Loss}) &= \lambda * \left(Var(U) + E(U)^2\right) \\
&= 58964 * \left(5000 + 200^2\right) \\
&= \$2653380000 \\
\text{Std Dev}(\text{Total Loss}) &= \$51511
\end{aligned}
$$

Note that the values for the Aggregate Loss Model (Method 2) are much smaller than those for the Total Loss Model (Method 1). This is to be expected since the CAR in the Total Loss Model includes intangible costs as well as direct costs. The intangible costs of a breach comprise a much larger component of the total costs of a breach than the direct costs. The Aggregate Loss Model is simply measuring the per record loss.

6. Conclusions and Future Directions

The Internet has not only revolutionized commerce and information exchange, but has ushered in new types of crime and potential liabilities with which companies must grapple. The potential losses stemming from a cyber data breach can be substantial, and not all firms are in a position to self-insure against such a large loss. This is where cyberinsurance can play an important role in the firm's risk mitigation strategy. The cyberinsurance market is still relatively new, however, insurance companies are still trying to gain a better understanding of the nature and size of potential losses from cyber breaches in order to price these products.

Our study contributes to this goal in two ways. First, we have proposed a model that accounts for the interconnected nature of companies, such as firms within the same industry, or firms using the same third-party data vendor. This model captures the correlated nature of cyber intrusions and data breaches, which tend to occur in clusters of companies within a short time period and can compound losses to an insurer when multiple clients are struck. Second, firms suffer both direct and indirect losses due to a breach, and researchers and analysts have argued that it is the indirect losses that are both more substantial, yet harder to quantify. If a firm wished to insure against their total losses, we provide concrete guidance by proposing the distribution of negative stock price returns from breach announcements as a proxy for this loss distribution. The stock price response is an objective, verifiable, immediate measure of the market's perceived loss in firm value, now and in the future, in present value terms, from the announced data breach. We find that for the 71% of firms in our sample suffering a stock price decline at the announcement, a gamma distribution well approximates this distribution of losses. An insurer could even choose to write an individual policy that pays out based upon the firm's measured abnormal stock loss at the public announcement. In short, these damage amounts may be used to inform cyberinsurance premium pricing and damage caps. As the cyberinsurance industry segment continues to grow and mature, this approach to assessing damages should be extremely valuable to insurers in providing useful cyberinsurance products at a fair price to their customers. Firms can also use this loss information to improve their cybersecurity event response processes and decide whether cyberinsurance fits within their risk mitigation strategy.

There are some potential limitations to this research. While the efficient markets hypothesis undergirds this study, to the extent that the market fails to correctly assess the damage to the firm and its overall value resulting from a data breach, this measure of total loss will be less accurate. A second consideration is that cyberinsurance policies are marketed to firms of various sizes, from sole-proprietorships to Fortune 50 companies. Our findings are based upon larger, publicly-traded firms. The results are only assumed to be scalable. Also, the larger firms can likely handle losses from a cybersecurity event better than smaller firms, simply due to resources available, such as access to reserves and capital markets. These topics all remain interesting avenues for future research.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Morgan, S. (2017) Cybersecurity Ventures. Cybercrime Report, Sponsored by the Herjavec Group.

[2] Ponemon Institute (2017) Responsible Information Management.
https://www.ponemon.org
http://www.ponemon.org/rim-council-faqs-1

[3] Ponemon Institutue (2016) 2016 Cost of Cyber Crime Study & the Risk of Business Innovation.
http://www.ponemon.org/local/upload/file/2016%20HPE%20CCC%20GLOBAL%20REPORT%20FINAL%203.pdf

[4] Deloitte (2017) Beneath the Surface of a Cyberattack: A Deeper Look at Federal Sector Impacts.
https://www2.deloitte.com/us/en/pages/risk/articles/hidden-business-impact-of-cyberattack.html

[5] Gordon, L.A. and Loeb, M.P. (2002) The Economics of Information Security Investment. *ACM Transactions on Information and System Security*, **5**, 438-457.
https://doi.org/10.1145/581271.581274

[6] Gordon, L.A., Loeb, M.P., Lucyshyn, W. and Zhou, L. (2015) Externalities and the Magnitude of Cyber Security Underinvestment by Private Sector Firms: A Modification of the Gordon-Loeb Model. *Journal of Information Security*, **6**, 24-30.
https://doi.org/10.4236/jis.2015.61003

[7] Garg, A., Curtis, J. and Halper, H. (2003) Quantifying the Financial Impact of IT Security Breaches. *Information Management & Computer Security*, **11**, 74-83.
https://doi.org/10.1108/09685220310468646

[8] Campbell, K., Gordon, L.A., Loeb, M.P. 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.
https://doi.org/10.3233/JCS-2003-11308

[9] 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**, 70-104. https://doi.org/10.1080/10864415.2004.11044320

[10] Acquisti, A., Friedman, A. and Telang, R. (2006) Is There a Cost to Privacy Breaches? An Event Study. *ICIS* 2006 *Proceedings*, Milwaukee, 10-13 December 2006, 94.

[11] Goel, S. and Shawky, H.A. (2009) Estimating the Market Impact of Security Breach Announcements on Firm Values. *Information & Management*, **46**, 404-410. https://doi.org/10.1016/j.im.2009.06.005

[12] Hovav, A. and 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. https://doi.org/10.1046/J.1098-1616.2003.026.x

[13] Gordon, L.A., Loeb, M.P. 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. https://doi.org/10.3233/JCS-2009-0398

[14] Morse, E.A., Raval, V. and Wingender Jr., J.R. (2011) Market Price Effects of Data Security Breaches. *Information Security Journal*: *A Global Perspective*, **20**, 263-273. https://doi.org/10.1080/19393555.2011.611860

[15] Spanos, G. and Angelis, L. (2016) The Impact of Information Security Events to the Stock Market: A Systematic Literature Review. *Computers & Security*, **58**, 216-229. https://doi.org/10.1016/j.cose.2015.12.006

[16] Rosati, P., Cummins, M., Deeney, P., Gogolin, F., van der Werff, L. and Lynn, T. (2017) The Effect of Data Breach Announcements beyond the Stock Price: Empirical Evidence on Market Activity. *International Review of Financial Analysis*, **49**, 146-154. https://doi.org/10.1016/j.irfa.2017.01.001

[17] Gordon, L.A., Loeb, M.P. and Sohail, T. (2003) A Framework for Using Insurance for Cyber-Risk Management. *Communications of the ACM*, **46**, 81-85. https://doi.org/10.1145/636772.636774

[18] Kesan, J., Majuca, R.P. and Yurcik, W.J. (2005) Cyberinsurance as a Market-Based Solution to the Problem of Cybersecurity—A Case Study. *Workshop on the Economics of Information Security* (*WEIS*), Cambridge, 1-3 June 2005.

[19] Bohme, R. (2005). Cyber-Insurance Revisited. *Workshop on the Economics of Information Security* (*WEIS*), Cambridge, 1-3 June 2005.

[20] Ogut, H., Menon, M. and Raghunathan, S. (2005) Cyber Insurance and IT Security Investment: Impact of Interdependent Risk. *Workshop on the Economics of Information Security* (*WEIS*), Cambridge, 1-3 June 2005.

[21] Zhao, X. and Xue, L. (2009) A Framework of Using Captive Insurance to Streamline IT Control and Compliance Management. *Journal of Information Privacy & Security*, **5**, 27-43. https://doi.org/10.1080/15536548.2009.10855868

[22] Zhao, X., Xue, L. and Whinston, A. (2013) Managing Interdependent Information Security Risks: Cyberinsurance, Managed Security Services, and Risk Pooling Arrangements. *Journal of Management Information Systems*, **30**, 123-152. https://doi.org/10.2753/MIS0742-1222300104

[23] Bandyopadhyay, Mookerjee, V.S. and Rao, R.C. (2009). Proposed Contracts Tend to Be Overpriced Because Insurers Are Unable to Anticipate Customers' Secondary Losses. *Communications of the ACM*, **52**, 68-73. https://doi.org/10.1145/1592761.1592780

[24] Srinidhi, B., Yan, J. and Tayi, G.K. (2008) Firm-Level Resource Allocation to Infor-

mation Security in the Presence of Financial Distress. Washington State University, School of Economic Sciences Working Paper Series WP 2008-17.

[25] Fan, J.P. and Lang, L.H. (2000) The Measurement of Relatedness: An Application to Corporate Diversification. *The Journal of Business*, **73**, 629-660. https://doi.org/10.1086/209657

[26] Dewan, S. and Ren, F. (2011) Information Technology and Firm Boundaries: Impact on Firm Risk and Return Performance. *Information Systems Research*, **22**, 369-388. https://doi.org/10.1287/isre.1090.0261

[27] Jia, N., Rai, A. and Xu, S.X. (2020) Reducing Capital Market Anomaly: The Role of Information Technology Using an Information Uncertainty Lens. *Management Science*, **66**, 979-1001. https://doi.org/10.1287/mnsc.2018.3235