JOURNAL OF CYBERSECURITY

*Journal of Cybersecurity*, 2019, 1–19
doi: 10.1093/cybsec/tyz002
Research paper

Research paper

# Content analysis of cyber insurance policies: how do carriers price cyber risk?

**Sasha Romanosky, Lillian Ablon, Andreas Kuehn and Therese Jones**

RAND Corporation, 1200 South Hayes St, Arlington VA, 22202

*Corresponding author: E-mail: sromanos@rand.org

Received 1 October 2018; accepted 20 December 2018

## Abstract

Data breaches and security incidents have become commonplace, with thousands occurring each year and some costing hundreds of millions of dollars. Consequently, the market for insuring against these losses has grown rapidly in the past decade. While there exists much theoretical literature about cyber insurance, very little practical information is publicly available about the actual content of the polices and how carriers price cyber insurance premiums. This lack of transparency is especially troubling because insurance carriers are often cited as having the best information about cyber risk, and know how to assess – and differentiate – these risks across firms. In this qualitative research, we examined cyber insurance policies filed with state insurance commissioners and performed thematic (content) analysis to determine (i) what losses are covered by cyber insurance policies, and which are excluded?; (ii) what questions do carriers pose to applicants in order to assess risk?; and (iii) how are cyber insurance premiums determined – that is, what factors about the firm and its cybersecurity practices are used to compute the premiums? By analyzing these policies, we provide the first-ever systematic qualitative analysis of the underwriting process for cyber insurance and uncover how insurance companies understand and price cyber risks.

**Key words**: cyber insurance; cyber liability; pricing cyber risk; thematic analysis; purposive sampling

## Introduction

Data breaches and security incidents have become commonplace, with thousands occurring each year and some costing hundreds of millions of dollars [1]. Consequently, the market for insuring against these losses has grown rapidly in the past decade (discussed more below). Cyber insurance is a broad term for insurance policies that address first and third party losses as a result of a computer-based attack or malfunction of a firm's information technology systems. For example, one carrier's policy defines computer attacks as a, "hacking event or other instance of an unauthorized person gaining access to the computer system, [an] attack against the system by a virus or other malware, or [a] denial of service attack against the insured's system".[1]

Although there exists a large, and growing, body of academic literature on cyber insurance,[2] it is almost exclusively theoretical, examining network externalities, asymmetric information and the viability of cyber insurance markets. While this work is necessary for understanding the antecedents of market success and failure, it does not examine the actual legal contracts (the insurance policies) upon which the theories and models are based.

Further, while insurance companies are often seen as the singular organizations with specialized ability to quantify and price operational risks,[3] there is almost no public information about how carriers actually assess – and differentiate – cyber risk across firms and industries, and particularly, how they compute prices for cyber insurance premiums. This lack of transparency in policies and practices is cited as one of the leading obstacles hindering adoption of

1 POL-35. Note that we will obfuscate the actual policy numbers and companies throughout this manuscript.

2 See the many references to cyber insurance research at https://econinfosec.org/weis-archive/ (15 August 2018, date last accessed).

3 The Cyber Incident Data and Analysis Repository (CIDAR) was an effort championed by DHS to leverage the capabilities that were growing within insurance carriers. See https://www.dhs.gov/cybersecurity-insurance# (17 September 2018, date last accessed).

cyber insurance,[4] and presents significant challenges for senior executives seeking to manage risks across their organizations because they are unable to effectively understand and compare coverages across insurance carriers. Moreover, the lack of transparency prevents these decision makers from using this information to implement security controls that could both reduce their operational costs, and improve their security posture.

Therefore, this research seeks to fill what we perceive to be a critical gap in the design, understanding, and purchase of cyber insurance underwriting by providing fundamental analysis and transparency of actual cyber insurance policies.

Since insurance in the USA is regulated at the state level, insurance carriers are required to file policies with state insurance commissions describing each new insurance product. These filings include the full text of the policy (coverage, exclusions, triggers, etc.), a security application questionnaire, and a rate schedule describing the formula for deriving insurance premiums. It is these filings that provide a unique opportunity to examine how insurance companies understand and price risks, and specifically, which business, technology and process controls (if any) are considered in rate calculations.

In this qualitative research, we seek to answer central questions concerning the current state of the cyber insurance market. Specifically, by collecting insurance policies from state insurance commissioners across New York, Pennsylvania, and California, we examine the composition and variation across three primary components: (i) the coverage and exclusions of first and third party losses which define what is and is not covered, (ii) the security application questionnaires which are used to help assess an applicant's security posture, and (iii) the rate schedules which define the algorithms used to compute premiums.

Below we provide a brief introduction to the size of the US market for cyber insurance, followed by a description of relevant literature. We then explain our research methodology, data, and results from the content analysis.

## The market for cyber insurance

The US cyber insurance market has grown rapidly over the past decade. With less than $1 billion in premium in 2012, some experts estimate that the US cyber insurance market will grow to $7.5 billion by the end of the decade [4], with others projecting $20 billion by 2025 [5, p. 24]. A recent survey of industry leaders found that 88% of respondents saw cyber as a "potentially huge untapped market" which they anticipated would grow faster than the rest of the property/casualty (P/C) insurance industry [6].

While the US market penetration may be more accelerated than other countries, only around one third of US companies have purchased some sort of cyber insurance [7], with significant variation in cyber insurance across US industry sectors. For example, barely 5% of manufacturing firms have cyber insurance coverage, whereas the healthcare, technology, and retail sectors have reached an adoption of close to 50% [8].[5] Yet, Marsh [10] reports cyber insurance growth rates of 27% across all industries, ranging from 6% in health care to 63% in manufacturing, for US-based clients in 2015.

The supply side of insurance is also growing very rapidly. While only a few firms were offering insurance products a decade ago, the National Association of Insurance Commissioners (NAIC) reported there to be around 500 carriers now offering cyber insurance [11].[6] Reports suggest that the US cyber insurance market is dominated by a handful of carriers, including American International Group, Inc. (AIG), accounting for approximately 22% of the market, Chubb Limited (CB) at 12%, and XL Group Ltd. (XL) at 11% [12], with ACE Ltd, Zurich and Beazley also providing coverage.

Average premiums are priced between $10 000 and $25 000,[7] with some carriers writing limits between $10 million and $25 million, and as high as $50 million [13]. However, as with most other insurance products, towers of cyber policies can be purchased in the event of extreme losses, and Airmic [14] suggests that limits of $200 million and $300 million exist for some industries.

## Related literature

This article is informed by two main streams of literature. The first is research on cyber insurance, which is almost exclusively theoretical [15, 16,17, 18, 19].[8] Overall, this body of work examines the incentives for firms to purchase insurance (demand side), the incentives for insurers to provide contracts (supply side), and the conditions necessary in order for a market to exist. The inevitable tension for firms, as many identify, is whether to invest in ex ante security controls in order to reduce the probability of loss, or to transfer the risk to an insurer [20]. In particular, Böhme and Schwartz [16] provide an excellent summary of cyber insurance literature, and define a unified model of cyber insurance consisting of five components: the networked environment, demand side, supply side, information structure, and organizational environment. Here, the network topology plays a key role in affecting both interdependent security and correlated failures. Their demand-side model considers the risk aversion of the insured, heterogeneity across wealth, impact, and defense and utility functions of firms, while the supply-side discussion considers, *inter alia*, the competitive landscape of insurers, contract design (premiums, fines), and the carrier's own risk aversion. Discussion of information structure relates to adverse selection and moral hazard, and finally, organizational environment describes issues such as regulatory forces that may exist to mandate insurance, require disclosure in the event of a loss, and the effect of outsourced security services and hardware and software vendors on a firm's security posture. Despite this body of work, however, none of it examines the form or content of actual insurance policies, or the pricing mechanism used by carriers.

In addition, there is some qualitative research on cyber insurance policies. In addition to conducting very rigorous theoretical modeling of an insurance market, Marotta et al. [21] provide an overview of covered loss areas across 14 carriers. Majuca et al. [22] mainly describe of the evolution of insurance policies since the late 1990s,

4 How, according to a leading insurance broker, a "lack of education about the loss exposures and available coverages to be a leading obstacle to cyber insurance sales" [2], and during a webcast by Deloitte, in response to the question, "Which do you think will be the biggest obstacle to convincing more buyers to purchase cyber coverage", over one-third of respondents ($n = 2094$) cited "lack of understanding about the risk and coverages available" [3].

5 Marsh [9] also reports an adoption rate of 16% across all industries; 8% in manufacturing; 50% in health care, but lower adoption numbers of 12% in communications, media, and technology, and 18% in retail/wholesale – for clients purchasing standalone cyber insurance in 2014.

6 Though, conversations with insurance practitioners suggest that while there may be 500 individual carriers, they represent subsidiaries from only about 70 insurance carriers.

7 Personal correspondence with an executive of a large insurance carrier.

8 This section appeared, in part, in Department of Commerce, Comments to Docket No 130206115-3115-01 1/10 by Sasha Romanosky, 26 April 2013.

as well provide an overview of covered losses from seven carriers, while Baer and Parkinson [23] review policies from six carriers. And Woods et al. [24] examine 24 self assessment questionnaires provided from insurance carriers.

And so, our research is also informed by qualitative research methods which guide us when examining, in a systematic and rigorous way, a corpus of documents. Specifically, the field of thematic analysis is an inductive (as opposed to deductive) research methodology used for "systematically identifying, organizing, and offering insight into patterns of meaning (themes) across a data set" [25]. In particular, "inductive" thematic analysis is used, "in cases where there are no previous studies dealing with the phenomenon, and therefore the coded categories are derived directly from the text data" [26]. This approach is appropriate for our research on cyber insurance since, to our knowledge, there is very little previous work that has rigorously examined each of the components of these policies.

Thematic content analysis is a rigorous methodology which has been used for decades and across many disciplines [27]. For example, Schwarcz [28] performs content analysis on a sample of homeowner insurance policies in order to measure the variation in coverage across insurance carriers, and Davis et al. [29] examined US state heath laws regarding prescription monitoring programs in order to determine the qualities of the law's intended purpose (such as related to countering misuse or abuse, or assisting with criminal investigations, etc.). Yu et al. [30] discuss the relationships between text mining (performed by machine learning techniques) and typical human-driven content analysis, providing dozens of examines of text mining across bioinformatics, business systems, engineering and education. And Ingle and Wisman [31] perform content analysis on teacher contracts in Kentucky to examine changes over time.

## Research methodology and data collection

In the USA, insurance laws are statutorily enforced by the McCarran–Ferguson Act (15 U.S.C. §§ 1011-1015) which empowers states to regulate the "business of insurance", and which is overseen by a nonprofit organization called the National Association of Insurance Commissioners (NAIC). In the 1990s, NAIC developed an online electronic records system called SERFF in order to facilitate the "submission, review and approval of product filings between regulators and insurance companies".[9] The filed documents include the policy forms (description of coverage, triggers, and exclusions), application forms (the self-assessment questionnaires presented to clients in order to assess their security posture), rate information (equations and tables governing the pricing of premiums), and other supporting documentation required or requested by the state insurance commissioners. As of 2016, 49 states and 3900 insurance companies and filers all participate in SERFF (though not all states allow electronic filing). The adoption of this electronic filing system by multiple states ensures uniformity and consistency of filed documents across all states, and are made available to the public, in part due to state open records laws.[10]

There is a distinction regarding insurance regulation related to admitted versus nonadmitted markets. Carriers that seek to operate in an "admitted" market (which is the source of our data collection) must file their policies and rate schedules with the state insurance commissions and comply with all state regulations in order to be licensed in a given state. Alternatively, carriers may avoid some of the restrictions imposed by state insurance commissioners by selling insurance in the "nonadmitted" market (also known as excess or surplus insurance lines). While some suggest that a sizeable portion of US cyber insurance is sold in the nonadmitted market,[11] the NIAC estimates that $1.8 billion in annual premiums is written in this admitted market [32, p. 8].[12]

## Sample selection

As mentioned, the goal of this research is to provide transparency around the three main components of cyber insurance policies: coverage and exclusions, security questionnaires, and rate schedules. We therefore leverage a form of qualitative research called directed content methodology, or thematic analysis, which enables us to identify and categorize themes and concepts, and derive meaning and insights across a collection of policies.[13]

In order to determine the appropriate number of policies to examine, we employ a common form of qualitative nonprobabilistic sampling known as purposive sampling [27]. Sample size in purposive sampling is determined by a concept called thematic saturation, which is the point at which "no additional data are being found whereby the (researcher) can develop properties of the category. As [the researcher] sees similar instances over and over again, [she] becomes empirically confident that a category is saturated" [34]. Guest et al. [35] further defines thematic saturation as the "point in data collection and analysis when new information produces little or no change to the codebook" while at the same time, the observations are "selected according to predetermined criteria relevant to a particular research objective" [35] – such as in our case of studying cyber insurance coverage from a larger pool of all state insurance documents. Specifically, Guest et al. [35] state that "the size of purposive samples be established inductively and sampling continue until 'theoretical saturation'... occurs".

We estimate the full population of cyber insurance policies to be around 2000–3000, a number larger than this research effort is able to examine.[14] Analysis of state-level insurance regulation, as well as conversations with industry experts and regulators, suggests that for the purpose of this study, there should be no systematic variation across the states in the content of insurance policies. This is not to say that there would be no differences, but just none that would materially bias any results or conclusions. Therefore, for the purpose of data collection, we can reasonably consider all US states to be similar, thus supporting a pooled analysis.

9 See http://www.serff.com/about.htm (20 January 2017, date last accessed).

10 Most often, the actual documents are filed by underwriting analysts employed by the carrier, or outsourced to specialized firms, or third parties agents, filing on behalf of the carrier. For example, we found a number of instances where insurance carriers employ the services of a third party organization which developed model policies and premium rates. In effect, these carriers were outsourcing the creation of lines of insurance, to provide coverage for specialized books of business for which they would likely have no prior experience underwriting, such as for cyber security.

11 One insurance executive believed that as much as 90% of the cyber insurance market is with nonadmitted carriers [32].

12 In which they state, "The remainder of the report will provide figures filed for each category and explain assumptions used to arrive at the $1.8 billion in direct written premium by admitted insurers"[33, p. 8].

13 This approach differs from a summative analysis where one would already have an exhaustive list of appropriate keywords and focus on the use of specific terms of interest and how common they are as compared to other terms.

14 This estimate is based on extrapolating based on the total number from a small sample of states.

For our data collection, we used the online SERFF system managed by NAIC to search for policies using the keywords: "cyber", "security", and "privacy". We limited the search to the broad category of property and casualty (P&C) insurance since "cyber insurance" is not covered under a single line of business, but instead is distributed across multiple lines of property and casualty insurance. We collected only "approved" documents, and omitted those which were filed but rejected. In total, we downloaded and examined 235 filing dockets from New York, Pennsylvania, and California. These states were chosen because they are three of the largest states by population, and where we therefore expect to see many policies with the most variation, thereby improving our thematic saturation.

The dockets covered years from 2007 to 2017, though not all 235 dockets contained all documents of interest for this research, which we discuss more below. The policies came from both large and small carriers, such as AXIS, Berkshire Hathaway, CUMIS, Chubb, Everest, Famers, Federal Insurance Company, Great American, The Hartford Steam Boiler Inspection and Insurance Company, Philadelphia, QBE, Travelers, XL, Zurich, etc.

In addition, some large insurance carriers make their coverage and exclusion policies available online, and so we also collected policies from the public websites of 15 major insurance carriers. Security questionnaires and rate schedules were not available and therefore not included in our analysis.

## Code development for thematic analysis

Each insurance docket consisted of a zipped file, often containing dozens of individual documents which may include (i) the policy coverage and exclusions form, (ii) the security questionnaire, the (iii) rate schedule, in addition to other supporting documents. Each docket was examined individually, though as mentioned, not all documents were included in each docket.

The coding process then began as follows: first the principal investigator created a master codebook for each state (NY, PA, CA) and recorded the following metadata for each docket: the policy identifier (i.e. a unique identifier assigned by the state), state, submission date, the filing insurance company, the product name, the insurance line, and the insurance group. Coverage/exclusion forms, questionnaires, rate schedules and other relevant documents were then embedded into the master codebooks.

Next, two authors of this Article coded the coverage/exclusion forms, while one author each coded the security application, and rate schedule sections. Each team developed their own code book as they examined and processed their respective documents. The codebooks for each section were guided by an inductive approach that enabled investigators to identify themes and patterns within their respective documents [34,36]. For example, the codebook for the coverage and exclusion section coded the covered losses separately from the exclusions. The codebook for the security questionnaires coded each unique question, which were then grouped into major and minor categories, while the codebook for the rate schedules differentiated between distinct categories of rate pricing (discussed more below).

The authors followed common coding practices to first deductively anticipate initial coding variables, and then as each subsequent policy was examined, updated the codebook in order to capture unexpected findings (Bowen, 2009). The themes were adjusted to create new or collapsing redundant themes, as needed. Thematic analysis performed on these sorts of structured documents presents a particular benefit over analysis of very loosely structured content, such as human subject interviews. In interview situations, the subject may provide a response, then backtrack, become distracted, or take an unexpected tangent, leaving the coder to interpret or otherwise search for latent meaning in a body of text, and making coding more prone to measurement error. In our case, however, coding was relatively more objective and straightforward because it was a direct result of whether a topic is present, or not, in the policy document.

We begin the qualitative content analysis by examining the coverages and exclusions (immediately below), followed by the security application questionnaires, and then the equations and methods used to derive the premiums. Note that policy identifiers have been anonymized using "POL-#", where the "#" symbol is replaced by a unique identifier.

## What losses do cyber insurance policies cover and exclude?

Cyber insurance, like most insurance products, generally distinguishes between two broad loss categories, "first party" and "third party". First party losses relate to those directly suffered by the insured (i.e. the "first" party to the insurance contract), while third party liability relates to claims brought by parties external to the contract (i.e. the "third" party) who suffer a loss allegedly due to the insured's conduct.

Of the 235 policy dockets collected from Pennsylvania, New York, and California, 54 had complete coverage and exclusion forms (2 of which were duplicate) filed between 2009 and 2016. In addition, we collected 15 coverage and exclusion forms posted by large insurance companies (from the nonadmitted market), for a total of 67 unique policies.

Our coding process for this section was as follows. For each policy, and for both coverage and exclusion sections, we coded each new criteria as they appeared, extending the codebook to capture the main components as necessary. The codes were generally categorized as covering first or third party losses, such as computer attack, network security liability, and personal data compromise. We repeated this process for all 67 policies. Note again that coding was a fairly objective process, facilitated by the fact that these policies are quite standardized in format, helping to reduce subjective interpretation common to unstructured data (such as from interviews). Once complete, we identified a total of 17 covered losses, and 58 exclusions. As a validity check, 6 randomly selected policies (9%) were checked for accuracy. We achieved a reliability rate of 97% for covered losses (3 discrepancies among 6 policies * 17 codes), and a reliability rate of 94% for exclusions (18 discrepancies among 6 policies * 58 codes).

As shown in Figure 1, we found that the covered losses appeared more consistent across all policies, whereas exclusions were more varied. For example, after reviewing only 6 policies, 88% of the covered losses had been coded, and by the 37th policy, we reached full saturation (upper panel). That is, it only took 37 policies before we identified all covered losses from the policies in our dataset. By comparison, after 16 policies, we reached 71% saturation for exclusions, and achieved full saturation by the 60th policy (lower panel).

As a simple form of robustness check, we compared policies between admitted and nonadmitted markets in order to determine whether there were any systematic differences in terms of new covered losses or exclusions. Carriers from the nonadmitted

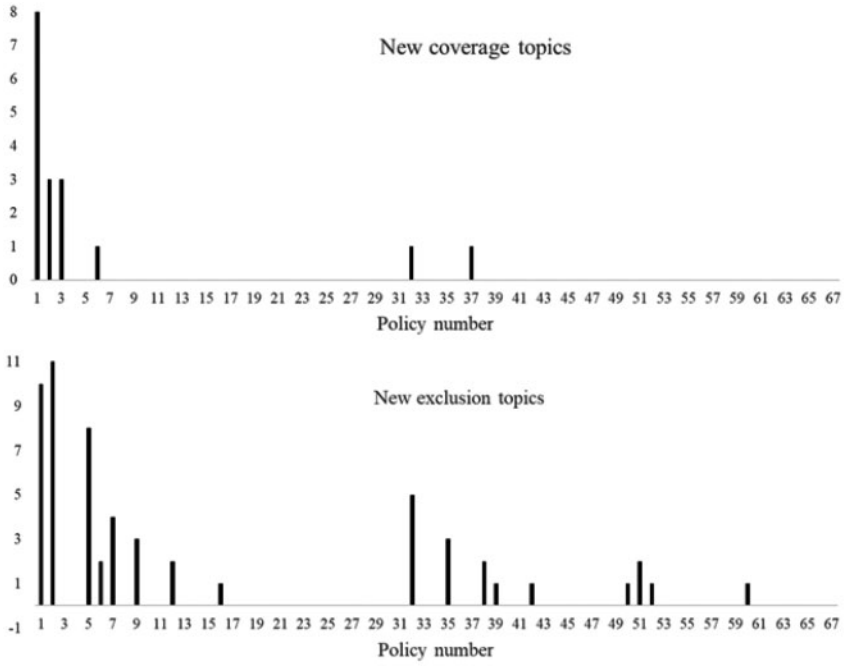


**Figure 1:** Identification of criteria over the course of reviewing policies

market were coded as policies 53–67 and as show in the upper panel of Figure 1, no new covered losses were coded, and as shown in the lower panel, only 1 new exclusion was coded (an exclusion for loss derived from an industrial control system, ICS/SCADA).

We find the consistency in coverage across policies to be surprising. From discussions with industry experts, the consensus is that there is so much variation across policies that examining a sample would provide no meaningful insights in industry-wide coverage. The results presented above, however, suggest that there is, in fact, a strong similarity for both coverage and exclusions across many policies and these states.[15]

Next, we describe the covered and excluded losses in more detail.

## Covered losses

Coverage for losses due to cyber incidents can be categorized in a number of different ways, and one familiar way is to differentiate between losses borne as a direct result of the incident (first party losses), and losses incurred as a result of litigation by alleged injured parties (third party losses). We discuss these more below, and then describe the most common losses overall.

## First party coverage

As mentioned, first party coverage includes losses incurred directly by the insured. For example, costs related to investigating the cause of a data breach or security incident, costs associated with restoring business services, the cost of notifying affected individuals, credit monitoring services, costs incurred from public relations and media services in order to communicate the event,[16] extortion and ransom payments,[17] and losses associated with business interruption.

In order to manage the various risks associated with these kinds of cyber incidents, carriers frequently assigned sublimits (and in some cases, distinct premiums), to groups of first party losses. For example, some policies differentiated among just a couple of categories, such as personal data compromise and computer attack.[18] Personal data compromise relates to the "loss, theft, accidental release or accidental publication of personally identifying information (PII) or personally sensitive information".[19] A computer attack relates to unauthorized access, malware attack, or denial of service (DoS) attack on any computer or electronic hardware owned or leased and operated by the policy holder.

However, more sophisticated – or perhaps, risk averse – policies differentiated among more coverage areas, each with their own sublimits. For example, POL-30 distinguished among the following groups as shown in Table 1.

15 In a number of cases, carriers would declare that firms from certain industries were ineligible to receive coverage. These industries included firms from adult business and gambling or gaming industries. In other cases, carries specifically excluded organizations involved in the sale or distribution of products regulated by the Bureau of Alcohol, Tobacco and Firearms, those involving use of pornographic data or images, or those with greater than 25% revenues generated from online sales. In one very restrictive case, the carrier considered firms within the following industries to be ineligible: education, healthcare, finance, government, publishing, data storage, website design, firms with websites containing information related to children, healthcare, entertainment/gambling, or sale of contraband or counterfeit items.

16 See POL-126.

17 See POL-127

18 CyberOne policies commonly had these few number of differentiators. For example, POL-1covered both *personal data compromise* and *computer attack* for 1st party coverages, and *network security liability* 3rd party coverage, provided by CyberOne. Many other CyberOne policies (e.g. POL-17, POL-23, POL-47, POL-49) just included coverages for *computer attack* and *network security liability*, so it may be the case that separate coverage for *personal data compromise* is considered something additional.

19 And, if PII is involved, it must result in or have the possibility in resulting in the fraudulent use of such information.

**Table 1:** First party coverage sublimits

| Coverage area | Description |
|---|---|
| Data Compromise Response | "Provides coverage for specified expenses arising from a personal data compromise involving personally identifying information of affected individuals. Affected individuals may be customers, clients, members, directors or employees of the insured entity." |
| Identity Recovery | "Provides coverage for Identity Recovery caused by an identity theft of an identity recovery insured first discovered during the policy period." |
| Computer Attack | "Provides coverage for specified expenses arising from a computer attack on the computer system." |
| Cyber Extortion | "Provides coverage for the cost of an investigator retained in connection with the extortion threat and coverage for any amount paid by the insured in response to the threat." |

**Table 2:** Third party liability sublimits

| Liability | Description |
|---|---|
| Data Compromise | "[Provides] coverage for defense and settlement costs in the event that affected individuals or a government entity sue the insured because of a personal data compromise." |
| Network Security | "Provides coverage for defense and settlement costs in the event that a third party claimant sues the insured because of:<br>• The breach of third party business information<br>• The unintended propagation or forwarding of malware<br>• The unintended abetting of a denial of service attack<br>• The inability of an authorized third party user to access the insured's computer system." |
| Electronic Media | "Provides coverage for defense and settlement costs in the event that a third party claimant sues the insured alleging that the insured's electronic communications resulted in defamation, violation of a person's right of privacy, interference with a person's right of publicity or infringement of copyright or trademark." |

## Third party liability coverage

As mentioned, third party liability covers the cost of defending against public or private litigation, settlements, judgments, or other rulings, as well as fines, fees, and settlements stemming from these lawsuits. For example, POL-35's network security liability coverage covers costs due to, "a civil action, an alternate dispute, a resolution proceeding or a written demand for money" as a result of "a [t]he breach of third party business information, [t]he unintended propagation or forwarding of malware, [t]he unintended abetting of a denial of service attack".[20]

Similarly with first party losses, coverage is available, and limits are distributed, across multiple kinds of claims. For example POL-30 distinguished between liability (brought by either a private or public action) due to a data compromise, network security incident, and electronic media as shown in Table 2.

### Most common covered losses

Figure 2 shows the top 10 most common covered losses.

Beyond the generalities defined above, below we describe a number of notable categories from the analysis.

**Cost of claims expenses, penalties**

This includes legal claims expenses related to penalties, defense and settlement costs. For example, POL-20 expressed how expenses would be paid for violation of timely disclosure of breach notice laws, regulatory and defense penalties, payment card (PCI) Fines, claims against the reputation of anyone or any organization, the invasion of privacy, or any claims against website content to include copyright and plagiarism.

**Public relations services**

Coverage for public relations (PR) costs appeared in the vast majority of policies, though sometimes came with restrictions.[21] For example, some policies only covered costs associated with advertising or special promotions, or in situations when a data privacy wrongful act had occurred, while other policies limited the total dollar amount of coverage, or excluded any costs directed to employees, or when affected individuals had already been notified.

**Notification to affected individuals (e.g. credit monitoring)**

Some policies are specific in terms of the kinds of services that can be provided to affected individuals – supplying a list of programs from which the policyholder must choose. For example, POL-22 requires that credit monitoring, identity monitoring, and fraud resolution services coverage only apply if Experian is used (specifically, Experian's ProtectMyID Alert, Family Secure, and DataPatrol).

**Computer forensic investigation costs**

Expenses for computer forensic services (i.e. examining computer systems for indicators of malware or malicious activity) sometimes

20 POL-111 covers first party losses stemming from crisis management expenses, security breach remediation and notification, computer restoration expenses, funds transfer fraud, extortion, and business interruption, as well as third party losses from network and information security liability, communication and media liability, and regulatory defense expenses. (POL-97) include coverage for, "Loss of Digital Assets, Non-Physical Business Interruption and Extra Expense, Cyber Extortion Threat, Security Event Costs, Network Security and Privacy Liability Coverage, Employee Privacy Liability Coverage Electronic Media Liability Coverage, Cyber Terrorism Coverage".

21 As with the wide variation of types of coverage offered, specific elements like Public Relations services were listed under a variety of names: from the broad "computer attack coverage" to the specific "privacy breach expense coverages", "privacy notification costs", and "data compromise response expenses".

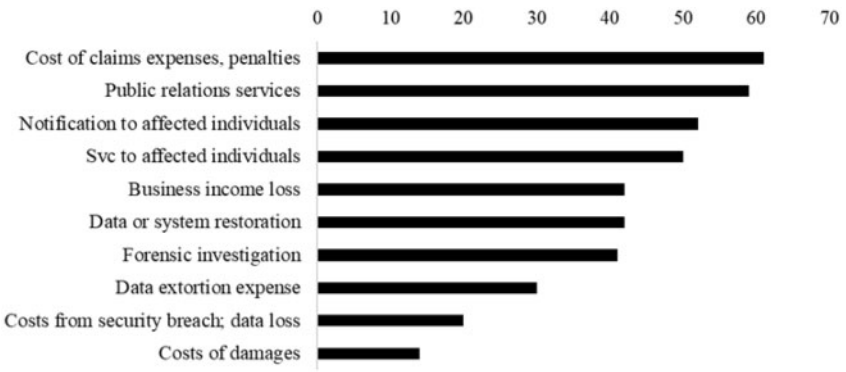


**Figure 2:** Most common covered losses

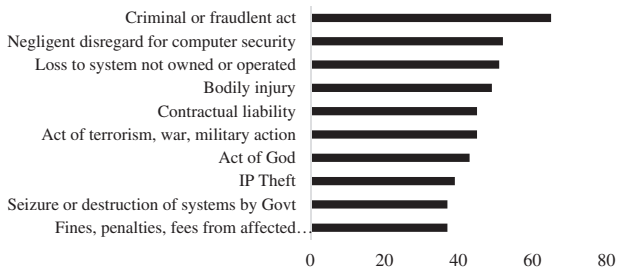


**Figure 3:** Most common exclusions

included the costs of computer expert services, and POL-22 noted that these expenses are specifically to be used in the case of disclosure of personally identifiable information (PII). For example, POL-22 states that, "If the incident involves an electronic security breach requiring computer expert forensic and investigation services … we will pay the costs of a computer security expert selected by you in consultation with our Breach Response Services Group from the program's list of approved security experts".

#### Items both covered and excluded

While about two-thirds of the policies covered expenses for data restoration, data re-creation, and system restoration, others explicitly excluded costs incurred to examine or correct a deficiency. For example, "cost[s] to research or correct any deficiency" (POL-49), or costs associated with the inspection, upgrading, maintenance, repair, or remediation of a computer system (POL-8; POL-19). Other expenses covered by many of the policies examined included business income loss, data extortion expenses, and forensic (computer) investigation.

### Exclusions

Figure 3 shows the 10 most common exclusions among the policies examined.

The exclusions most commonly observed were those not necessarily directly related to the cyber realm, but instead criminal, fraudulent, or dishonest acts, errors or omissions, intentional violation of a law, any ongoing criminal investigation or proceedings, and payment of fines, penalties, or fees. Several policies provide additional exclusions for infringement of patents, disclosures of trade secrets or confidential information, or violations of securities laws. We also found exceptions to the exclusions given certain circumstances (which themselves might have exclusions too). For example, in POL-22, any claims or losses arising from any deceptive or unfair trade practices are not covered – unless the claim results from the theft, loss, or unauthorized disclosure of PII, but only if no one involved in the deceptive or unfair trade practices participated or colluded in the theft, loss, or unauthorized disclosure.[22]

Other exclusions related to matters of physical harm (e.g. bodily injury, electric or mechanical failure, fire, smoke, wind or Act of God, release of pollutants), aspects of liability suits (e.g. nonmonetary relief, and expenses resulting from the propagation or forwarding of malware on hardware or software created, produced, or modified by the policy holder for sale, damages related to employment discrimination, contractual liability, theft of intellectual property), and losses to systems out of the policyholder's control (e.g. loss to the Internet, ISP, computer, or system not owned or operated by the policyholder). As mentioned previously, expenses for extortion or from an act of terrorism, war, or a military action were covered in rare cases, but mostly noted as exclusions.[23]

Other rare but notable exclusions included, collateral damage (i.e. malware, denial of service attack, or intrusion not directly aimed at the policyholder), failure to disclose a loss of PII if an executive of the firm was aware of such a loss, and salaries, benefits, expenses of employees.[24]

While we found no substantial differences in coverage between state policies and those of large carriers, there were some differences in exclusions, as shown in Table 3.

## Summary

Analysis of the covered and excluded losses highlights a number of important insights. First, as with all lines of insurance, there is a clear distinction between first and third party losses (i.e. costs borne by the firm directly, versus those incurred through litigation) which become relevant for establishing dollar values on limits and

22 As can be seen by this description, parsing out the nuances in the policies can be a challenge: exclusions include exceptions that have their own exceptions buried in them.

23 Note that the difference between losses not directly covered, and losses explicitly excluded is referred to as "silent" coverage, and is of great concern for carriers because it represents potentially claimable losses stemming from unanticipated events. For example, consider a patient death caused by someone hacking an insulin pump in a hospital. Even if a medical malpractice policy does not specifically cover this loss, it may be claimed – unless there is a specific exclusion. And so, omission of a covered loss without a specific exclusion reflect these kinds of "silent losses".

24 This exclusion was seen in four policies. Recall that one policy specifically included this as part of their coverage.

**Table 3:** Exclusions found more commonly/rarely in large carriers vs. state policies

| Rarely seen in large carriers | More common in large carriers |
|---|---|
| • Cost to research/correct deficiency<br>• Suit from propagation or forwarding of malware<br>• Nonmonetary relief | • Property damage<br>• Seizure/destruction of systems/data by government<br>• Natural elements<br>• Unlawful collection or sale of information<br>• Unsolicited dissemination of communication<br>• Unfair trade<br>• Intellectual property theft |

sublimits. Further, as seen from the most common covered losses in Figure 2, the top four relate to what are essentially cleanup costs. That is, indirect costs borne by the firm in order to comply with laws, manage the firm's reputation, and reduce further expenses following a breach. Whereas, the other costs (e.g. business income, data restoration, forensic investigation, etc.) are those directly associated with the cyber incident. One may speculate that this is because cleanup costs are more expensive (and/or more quantifiable) relative to direct costs, and therefore, exist because of increased demand by applicants. However, limited survey evidence suggests that direct and indirect costs are relatively equal.[25]

As consumers and firms adopt more technology and connected devices, there will likely be revisions to losses explicitly covered or excluded by cyber insurance policies. For example, one policy (POL-24) noted that expenses due to defects or deficiencies of the insured product were not covered. However, with the increase of the Internet of Things (IoT) devices, distributed denial of service (DDoS) attacks leveraging IoT devices, code reuse among products, and nonstandardized software security practices of developers, exclusions may well become more frequent. And while policies discussed traditional computers, networks, and systems, there was no explicit mention of emerging risks from mobile devices, drones, IoT devices, and the growing interdependencies of critical infrastructure.

Perhaps carriers recognize the increased likelihood of being a victim of collateral damage, and have therefore decided to exclude coverages from claims resulting in this (over half of the policies we examined excluded any claims related to war, military action, or terrorist action; and almost half of the policies excluded claims related to extortion or ransom [although approximately a third did include coverage for extortion or ransom]). We might expect that more policies in the future will include similar exclusions, as the likelihood increases (along with the cost to recover). Indeed, the matter of how malicious cyber incidents may or may not trigger these "act of war" exclusions" is currently a hotly debated issue.[26]

## Security questionnaires

The next component of cyber insurance policies to be examined is the security questionnaires. These questionnaires are provided by the carriers, and are ostensibly designed to solicit a comprehensive understanding of (or at least reasonable approximation to) the overall security posture of the applicant. Moreover, the questions should help to "differentiate" risks across a portfolio of applicants.

Of the 235 insurance dockets we downloaded and analyzed, 31 had questionnaires.[27] In eight cases, multiple questionnaires were included in a policy and in cases where the questionnaires were distinct (because they were written for different types of applicants,[28] or used different questionnaires for application and renewal), they were coded separately, generating a total of 45 questionnaires. We then found 11 cases of duplicate questionnaires, which we omitted from the analysis. This resulted in 34 unique, coded questionnaires.

Each questionnaire was analyzed in depth, and compared against existing questions and categories in the codebook. While most questions were straightforward to code (e.g. "does the applicant adhere to a particular technical standard?"), some required additional scrutiny in order to differentiate between related questions. Therefore, as is standard practice, coding was done using an iterative process involving adding new questions, or merging/splitting existing questions based on the growing understanding of distinct topics and categories (e.g. capturing new subcategories, such as Management policies, Privacy policies, and Technology policies). For validity, the investigator revisited the codebook to compare and adjust the coding, where necessary. A sample of 10 policies (22%) were then checked for accuracy, with 5 discrepancies found.[29]

In total, we identified 118 different topics, some of which were very detailed (e.g. "does the applicant deploy intrusion detection systems (IDS) or intrusion preventions systems (IPS)?") while others were quite broad (e.g. asking about general "business information"). However, many questions expressed similar themes, such as those pertaining to business information, data type, and questions regarding the compliance with PCI/DSS standards or the deployment of antivirus systems. Therefore, the 118 unique topics were organized into 14 subcategories, from which 4 main themes were created: Organizational, Technical, Policies and Procedures, and Legal and Compliance. Figure 4 illustrates the number of questions for each subcategory. For example, the Data Collection and Handling subcategory contained 11 unique questions, while the IT Security Budget/Spending subcategory had only 2. Overall, the Organization category had 35 questions, the Technical category had 21, and the Policies and Procedures and Legal and Compliance had 51 and 11 questions, respectively.

As shown in the left panel of Figure 5, after reviewing just 3 questionnaires, 78% of all 118 questions had been coded, and by

25 See Ponemon 2014 Cost of Data Breach Study: United States, Table 3, p18 showing a 51%/49% split between direct (investigations and forensics and lost customer business), and all other indirect costs.

26 For example, see a general discussion at https://policyholderinformer.com/2017/02/21/the-art-of-cyber-war/, and a specific incident regarding claims filed as a result of the 2017 NotPetya ransomware attack https://www.techlawx.com/blog/notpetya-insurance-coverage-dispute.

27 Note that these policies may not include the same subset of policies examined in the previous section.

28 For example, POL-31 wrote separate and distinct questionnaires for Technology Professionals, Accounting and Financial Professionals, and Small Firm Accounting and Financial Professionals.

29 1–5/(10*118).

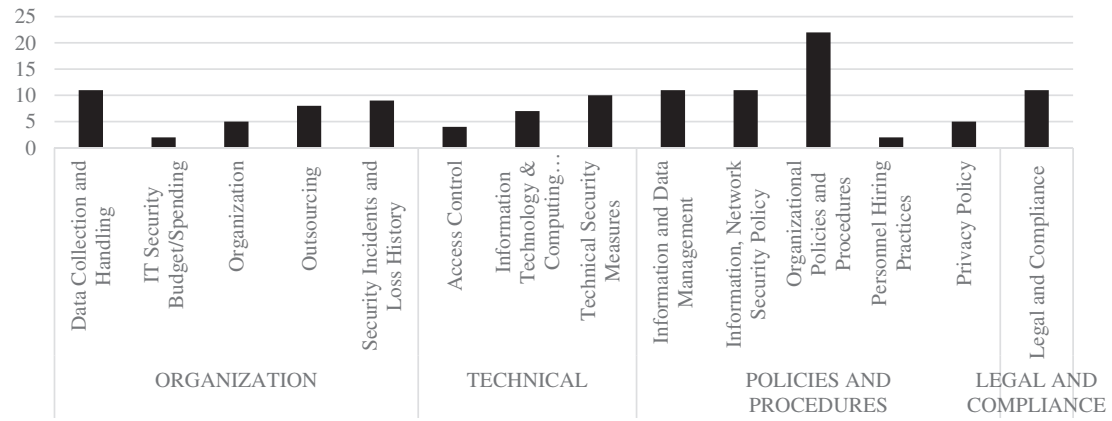


**Figure 4:** Number of unique questions per subcategory

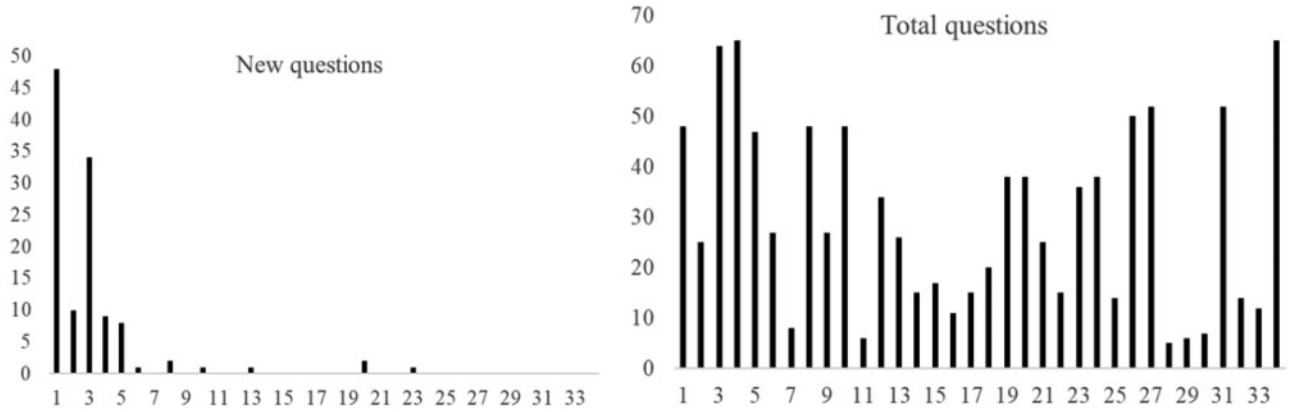


**Figure 5:** Question criteria

the 23rd questionnaire, we achieved 100% saturation. In regard to the "total" number of questions per document, there was considerable variation as shown in the right panel of Figure 5. In some cases, the questionnaires were quite long, with almost 70 questions, whereas others only included a few (the median number of questions was 26).

## Organizational

### Organization

The applications typically begin by collecting basic information about the company, such as the type of business and the industry sector in which the company operates, as well as financial information about revenues and assets. In a few cases, the questionnaires asked the company to submit an audited annual statement. For example, POL-7 asked for a "copy of most recent financial statements (10-K, annual report, etc.)".

To assess the operation of a business, POL-9 and POL-5 gathered information about the applicants' clients, including questions about the largest and most significant clients, the size of their contracts, and the duration of the project and relationship with the clients. POL-5 asks the applicant to provide "details on the Applicant's top three (3) revenue-producing clients or projects during the last fiscal year", and POL-9 asks to "list the Applicant's five largest clients", including value and length of contract.

Information is also collected about the company's past and current insurance coverage, including selected deductibles, and exclusions, if applicable.

### Data collection and handling

There was a concerted effort to understand the kinds of sensitive or confidential information that the applicant collects, stores, processes, or for which it is otherwise responsible. Of particular interest is PII, confidential client information, or corporate intellectual property, such as SSN, credit/debit card numbers, driver license, email addresses, IP addresses, financial and banking information, medical records, protected health information (PHI) as well as intellectual property, and trade secrets. For example, POL-18 asked, "what Third Party electronic information the Applicant collects or stores: 'Medical/Health Information', 'Credit Card Information', and 'Personally Identifiable Customer Information, other than Credit Card or Medical/Health Information'".

In comparison with the "technology and infrastructure" category these questions focus on the kind of data an applicant is managing. This suggests that carriers focus on data and the potential loss at risk. This possibly explains why relatively little information is collected about the technology and infrastructure landscape, or at least suggests that this category is less relevant when assessing an applicant's risk of filing a claim.

### Outsourcing

Questionnaires also addressed how the applicant manages its relationships with outscoring providers and the services the applicant relies on to conduct business. Given that it is common to outsource services and use third party service providers, these questions were relatively common. Questionnaires asked the insured to list the outsourced services and provide the names of providers, and some even provided a comprehensive list for the applicant to select. For example, POL-22 asks whether, "the Applicant outsource[s] any part of the Applicant's network, computer system or information security functions".

Questionnaires further assessed whether a security, privacy, and/or risk assessment was performed on the third party provider. The history of the third party providers is assessed, with regard to whether they were subject to privacy or security breaches in the past. Further, contracts between the insured and the third party were examined, such as whether they were structured in a way to hold third parties liable for losses resulting from data and security breaches, or whether they included an indemnity clause to transfer risk to a third party. For instance, POL-18 asks "Does the Applicant's contract with the service provider(s) state that the provider: (a) Has primary responsibility for the security of the Applicant's information?; (b) Has a contractual responsibility for any losses or expenses associated with any failure to safeguard the Applicant's data?" In some instances, the questionnaire asked whether the insured requires the outsourcing provider to have sufficient cyber insurance to minimize any liability a customer can claim that results from an incident at the outsourcing provider (e.g. data or security breaches at the site of the outsourcing provider).

### Incident loss history

In almost all questionnaires, the insurer collected information about the applicant's experience with regard to past security incidents. While the formulation and framing of the questions varied across the questionnaires, in essence, the following issues were addressed: (i) past data and security breaches and their impact; (ii) privacy breaches and loss of confidential information that triggered the notification of customers and/or employees; (iii) circumstances that could lead to an insurance claim; (iv) lawsuits and claims that are the result of an IP infringement; (v) extortions through the means of cyber, investigations by a regulatory or administrative agency. While other insurance companies often included multiple lengthy questions with regard to the security incident and loss history, POL-26 only asked, "Has the Applicant had any computer or network security incidents during the past two (2) years?"[30]

### IT security budget and spending

IT security budget and spending provides insights into how much an insured invests in its information and IT security. However, IT security budgeting and spending was addressed in one questionnaire, only. POL-18 asked "What is the Applicant's aggregated budget for system security" and "How is the system security budget allocated among: (a) prevention of security incidents; (b) detection of security incidents; (c) response in security incidents, all in percentage".

## Technical

### Information technology and computing infrastructure

Understanding the technology and infrastructure landscape of an insured would seem to be a relevant factor to consider in the risk assessment. Yet, only a few insurers cover this aspect in their questionnaire. When they did, only a few questions were posed, such as the number of computing devices, the number of IP addresses, or websites. For instance, POL-26 asked, "What is the Applicant's total number of IP addresses?" while POL-18 asks "List all website URL's and static IP addresses utilized by the applicant and its subsidiaries". In a few cases, policies asked whether the business' critical software was developed in-house. In another case, POL-52 inquired whether the insured segregated its IT systems that store and process PII from other parts of the network, "Are systems, applications and supporting infrastructure that collect, process, or store personal information segregated from the rest of the network?"

Information about the technology and infrastructure landscape would clearly help a carrier understand, if only at a basic level, the overall attack surface of a potential insured and, with more information, help assess their overall information security risk posture. However, it seems that only very rudimentary information is collected.

### Technical security measures

Questions regarding technical measures to protect against data theft and intrusions were found in most questionnaires. These included questions concerning the kinds of tools used to secure the applicant's networks and computers, including antivirus software to perform scans on email, downloads, and devices to detect malicious files or processes; IDS/IPS to detect possible intrusions and abnormalities in networks; and firewalls. POL-7 for instance, asks "Do you utilize firewall and intrusion prevention measures for your network and computer systems?" Encryption for data at rest and in motion was a technical measure that was often mentioned in the questionnaires. In its questionnaire, POL-7 asks, "Do you use commercial grade technology to encrypt all sensitive business and consumer information transmitted within your organization or to other public networks?" and "Do you use commercial grade technology to encrypt all sensitive business and consumer information at rest within your systems?" Some questions also focused on mobile devices, while VPN and two-factor authentication were less frequently listed as technical measures.

From our analysis, questions regarding such technical measures were present in almost all applications. However, there was considerable variation in the types of questions that addressed technical measures.

### Access control

Access control addresses the means and policies to secure user access, including the assignment of designated rights for users to resources. It attempts to restrict the access to sensitive data on a need to know basis. POL-54 asks, for instance, "Does the Applicant physically protect access to dedicated computer rooms and/or servers?" Beyond matters of access and users rights/privileges, questionnaires addressed whether processes were in place to revoke user rights and privileges once users terminated or left the organization.

30 Where "incident" was defined as "any unauthorized access or exceeding authorized access to any computer, system, data base or data; intrusion or attack; the denial of use of any computer or system; intentional disruption, corruption or destruction of electronic data, programs or applications; or any other incidents similar to the foregoing? – Note: if the answer to Question III is 'Yes', please attach a complete description of the incident(s), including whether the Applicant reported the incident(s) to law enforcement and/or the Applicant's insurance carrier".

Furthermore, this includes the monitoring of unauthorized access to or large download of sensitive data, as well as remote shutdown and data wipe out capabilities for computers. Again, POL-54 asks "Does the Applicant utilize remote shutdown of employee laptops?"

## Policies and procedures

### Information and data management

This category includes questions with regard to the applicant's data management practices – the number of records held, whether the applicant sells or shares sensitive information (i.e. PII) with third parties, and whether it processes information for third parties, including the processing or storing of credit or debit card transactions. For example, one insurer in questionnaire POL-22 asks whether, "the Applicant process or store personally identifiable information or other confidential information (including but not limited to payment information) for third parties".

The most common question in this category was whether a data retention and destruction policy existed. For example, POL-54 asks "Does the Applicant maintain procedures regarding the destruction of data residing on systems or devices prior to their disposal, recycling, resale or refurbishing?" Interestingly, the questions do not exclusively address digital data, but rather, data management is conceived more broadly to also include written records that warrant protection (e.g. handling of sensitive information such as client or human resource information, etc.).

The need for a corporate policy for record and information management and a classification system that determines what data must be protected was only expressed in a few questionnaires. In only one instance, did an application inquire whether the responsibility for records and information management was assigned to a senior executive.

### Employee, privacy, and network security subcategories

Questions concerning an applicant's privacy policy, and information and network security policy were common but varied in detail. In some instances, the questionnaires assessed details of how a policy was implemented and tested, and whether a policy was reviewed by the legal counsel and approved by the board of directors. POL-9, for example, asks "Does the Applicant have Security and Privacy Policies that are updated continually and implemented and, are there policies and procedures in place to ensure the Applicant is in compliant with requirements that govern the Applicant's industry?" If the applicant answers yes, the questionnaire continues to ask "If 'Yes' have the policies been reviewed by a qualified attorney?"

While privacy, and information and network security policies were the most common policies mentioned in the surveyed questionnaires, usage policies for the internet, social networking, and/or email were mentioned. Less common were policies for software development (i.e. the use of secure coding standards) and password policies (e.g. the use of strong encryption).

However, aside from these, the questions did not cover the substance of a particular policy (i.e. what should be in those policies, and how should they regulate particular issues) but rather only tested their existence. In numerous cases, the questionnaires asked whether the responsibility of privacy and information and network security and their respective policies are assigned or "owned" by a Chief Privacy Officer (CPO) role and a Chief Information Security Officer (CISO) role, respectively. In most questionnaires, the CPO and/or CISO roles were explicitly stated, in rather few cases was it referred to as responsibilities assigned to an individual. For instance, in POL-9 asks "Does the Applicant have a designated person that is responsible for the management, implementation and compliance of the Applicant's security and privacy policies and procedures".

### Organizational security policies and procedures

In addition to technical measures that are implemented to protect the information system in the daily business operation, organizational measures and procedures describe a set of measures to maintain and strengthen information security. Questions in this category related to penetration testing, vulnerability scanning, assessment, and management. Further, questions related to security and privacy assessment conducted by internal first parties or external third parties were asked, as were measures with regard to physical security (e.g. physical access control to computing facilities). For instance, POL-18 asks "Does the Applicant run vulnerability scans or penetration tests against all parts of the Applicant's network? If 'yes' how often are the tests run?" The applicant can then indicate the frequency by checking the box for "Daily, Weekly, Monthly, or Greater than Monthly". Several questionnaires assessed whether a business continuity plan (BCP), disaster recovery plan, as well as an incident response plan (IRP) were in place. Extended questions were concerned about the assignment of, and approval by, senior executives for the BCP and IRP. Further questions addressed data backup procedures as well as training with regard to information security procedures.

## Legal and compliance

Over the years, a variety of laws and regulations on the federal and state level, as well as industry standards have emerged that aim to protect consumers from the consequences of cyber incidents and data breaches. These laws, regulations, and standards are widely acknowledged in the questionnaires. Almost every questionnaire includes language about HIPPA, PCI/DSS, and GLBA, but also other US federal and state laws. In some but not all cases, the questionnaires ask to provide metrics about how well the respective standards are implemented and adhered to. PCI/DSS as an industry standard for payment processing was prominent in many questionnaires. Further, questions concerning PCI/DSS commonly exhibit a significant amount of detail. For example, one insurer asks: "How many credit or debit card transactions does the Applicant process annually?" and then continues to collect information about whether the applicant: "(a) Mask[s] all but the last four digits of a card number when displaying or printing cardholder data; (b) Ensure[s] that card-validation codes are not stored in any of the Applicant's databases, log files or anywhere else within the Applicant's network; (c) Encrypt[s] all account information on the Applicant's databases; (d) Encrypt[s] or use tokenization for all account information at the point of sale; or (e) Employ[s] point-to-point encryption, starting with card swipe hardware."[31]

## Summary

So far, this analysis begins to provide transparency into the information that carriers are concerned about when assessing cyber risk. For example, we observe an emphasis on the amount of data (i.e. number of records) and the type of data (i.e. sensitive and confidential

31 POL-22.

data) managed by the firm. The focus on sensitive data, particularly those to debit and credit card transactions and the detailed questions concerning PCI/DSS standard compliance is not surprising given that in the past decade data protection industry standards and data breach laws have developed and have been widely institutionalized in the USA.

On the other hand, there is little attention given to the technical and business infrastructure, and their interdependencies with environment in which the applicant is operating. These rather technical areas could provide further insights into the risk situation and security posture of an applicant. With regard to organizational processes and practices, it was surprising that risk management and IT security management as corporate functions and processes did not receive more attention.

It is noteworthy, however, that standards and frameworks for information technology management, such as the ITIL and COBIT are not mentioned, and in only one instance was an ISO standard mentioned. Also, the recently developed NIST Cybersecurity framework[32] is not mentioned, though from conversations with carriers, they are beginning to integrate it into these questionnaires.

Only in one instance, did a questionnaire asked about the size of the IT/information security budget and how it is spent with regard to prevention, detection, and response to security incidents. This finding was surprising given the amount of money spent on IT and information security could serve as a useful indicator for security maturity.

In addtion to the analysis described above, we did not observe any substantial changes in policy length, style, or composition over time. Conceivably, carriers may develop institutional knowledge that would lead them to improve and refine the questions ove time, or, perhaps the questions would be found to be too generic, requiring more details solicited from applicants.

## How do carriers price cyber insurance?

We mentioned earlier that insurance is regulated at the state level, and state laws require that insurance rates are not "excessive, inadequate, or unfairly discriminatory".[33] "Excessive" implies that the premiums are not priced unreasonably high, "adequacy" implies that the premiums are high enough in order to support the business for the carrier, and "discriminatory" implies that any price differences appropriately reflect variation in actual risk across firms.[34] But what are firms charging, and how do carriers determine these prices?

In this section, we examine the forms and equations used by insurance carriers to price cyber risks (formally known as "rate schedules").[35] We first examine justifications that carriers provide to state auditors when determining pricing policies, and then analyze the pricing schemes used to compute premiums. We conclude this section by showing the actual equations used to derive those premiums.

Coding in this section was accomplished in two steps. First, the principal investigator (PI) searched through each policy docket for files containing rate schedules or and any written justification of the premium calculation process. Second, for each policy that included justification of the premium calculation process, the text was copied and pasted into the master codebooks (previously described). In addition, a new codebook was generated in which the PI coded the type of policy, and the factors used to price the final premium, such as industry, claims history, etc., and where available, the number of security questions posed. As a validity check, all policies were reviewed a second time to ensure they were coded properly, and to identify any duplicates (of which 3 were found).

We first discuss the rate schedule justification, followed by the premium equations.

### How much do carriers know about cyber risk?

Of the 235 dockets examined, 56 included explanations for the state insurance auditor concerning the carrier's approach for deriving premiums. It is in these documents that we observe the process by which insurance pricing is conducted, and what information carriers may have in order to price cyber risk. From our analysis, we detected five main themes that carriers used for determining prices: (i) relied on external sources, (ii) estimated or guessed, (iii) looked to competitors, (iv) leveraged the experience of their own underwriters, and (v) adapted prices from other insurance lines.

Overall, many carriers began by stating how "cyber" is a relatively new insurance line, and that they have no historic or credible data upon which to make reliable inferences about loss expectations (e.g. "Limitations of available data have constrained the traditional actuarial methods used to support rates", POL-11).

In a number of cases, though, carriers employed the services of other companies to help develop premiums, or additionally it collected industry, academic, or government reports themselves that contained basic loss data. For example, POL-50 stated:

> Frequency was derived from data gathered from the 2011 Computer Security Institute Computer Crime and Security Survey and from the HSB/Ponemon survey. Severities were calculated for three of the sub-coverages (data restoration, data re-creation and systems restoration) using data drawn from the HSB/Ponemon survey and from the 2003 Graziado Business Review which were then combined with dollar amounts that represented the costs of repairing various kinds of covered damages. These costs were obtained from a variety of IT repair resources, including surveys and published rates.[36]

32 See https://www.nist.gov/cyberframework (9 August 2018, date last accessed).

33 This phrase is universal across state agencies and represents the spirit of state insurance regulation.

34 See https://www.iii.org/es/article/regulation-modernization (12 September 2018, date last accessed).

35 "Rate schedules" or "rate development" are industry terms of art for the forms used to price premiums. Also, note that these schedules were not available for all policies acquired.

36 The same carrier also wrote, "HSB interviewed several lawyers that focus their practices in the cyber area and asked them to quantify, for each kind of dispute, how much it costs to take it to summary judgment, what percentage of disputes go beyond summary judgment, how much it costs to take the dispute to trial, etc. This expert elicitation process produced the severity estimates". While another carrier wrote, "According to a recent study commissioned by the Federal Trade Commission, 90% of all ID theft out of pocket expenses are $1,200 or less. We believe that the availability of case management restoration services will reduce this severity to approximately $230. The same FTC-commissioned report suggests a frequency of 3.66%. Thus, our loss content is expected to be $8.42. Loss-related expenses (toll-free helpline and case management service) are expected to be $3.00, resulting in a total IDR loss cost of $11.42. We added the loss costs together and applied our expense and profit load of 65.6% to arrive at our gross premium of $1,913.91" (POL-30).

In other cases, carriers used other public information, which was augmented with additional sources or their own, limited experience. For example one carrier wrote, "We reviewed the rates for a less robust cyber product developed by Hartford Steam Boiler ('HSB') for the same types of accounts we are targeting[,] and then at a composite rate of the carriers writing more expansive cyber coverage for larger and more technologically sophisticated accounts. These two rates then became the two outside points of reference for establishing our rates" (POL-61).

Or, in some cases, the carrier would appear to guess [e.g. "The base retentions were set at what we believe to be an appropriate level for the relative size of each insured" (POL-6)], while many carriers employed what (limited) experience they had (e.g. "Rates for this coverage have been developed based upon the experience and judgment of our underwriters, claims personnel, and actuaries" (POL-25).

Further, in a number of occasions, we observed that carriers based their rates on the pricing of their competitors. For example, POL-36 states "the rates for the above-mentioned coverages have been developed by analyzing the rates of the main competitors as well as by utilizing our own judgment", and POL-31 states, "the program base rates and rating variables were based on a competitive review of the marketplace and underwriting judgment". While this may seem like an odd practice, discussion with insurance professionals suggest that this is, indeed, a common and appropriate occurrence.

In only a few cases were carriers confident enough in their own experience to develop pricing models, for example, one carrier wrote, "Underwriters collectively have over 40 years' experience in e-commerce, cyber, privacy and network security liability insurance. The collective knowledge of underwriters, including a deep understanding of competitive rates and feedback from the wholesale and retail brokerage industry, was used to establish rates for the program" (POL-2).

In a number of instances, we observed how carriers would turn to other insurance lines to price premiums because of their lack of data. One carrier admitted, "We are not using claim counts as the basis for credibility because we have not experienced any claims over the past three years" (POL-73). And in such cases carriers would base cyber risks on other insurance lines. For example, "Loss trend was determined by examining 10 years of countrywide Fiduciary frequency and severity trends. Because CyberRisk is a developing coverage we chose to use Fiduciary liability data because it has a similar limit profile and expected development pattern" (POL-43). Other carriers also leveraged loss history from other insurance lines, "the Limit of Liability factors are taken from our Miscellaneous Professional Liability product" (POL-25), and "Base rates for each module of this new product were developed based on currently filed Errors and Omissions and Internet Liability rates" (POL-104).

Regardless of the formal (and sometimes very informal) methods used in the underwriting process, it appears that state regulations require that carriers be vigilant about ensuring fair and accurate pricing. This is done, in part, by ensuring the underwriters are empowered to adjust premiums appropriately, when necessary [e.g. "The rating modifiers . . . allow the underwriter to debit or credit the policy premium based on the unique attributes of an insured. These modifiers reflect objective criteria associated with the cyber risks and controls of an insured" (POL-6)]. And further, this required concrete advisors by insurance auditors, where one auditor wrote, "Please be advised that the company is required to maintain statistical data, including incurred losses and loss adjustment expenses, on reported and unreported and outstanding and paid categories, on this program separate and apart from its other coverages. In addition, the experience should be reviewed annually, and appropriate rate revisions filed, (POL-49)" to which a number of carriers replied, "[w]e will monitor our book's performance as we develop our own experience to ensure that our product remains competitive and profitable" (POL-63).

Next we examine the actual rate schedules and analyze the methods used to price cyber insurance premiums.

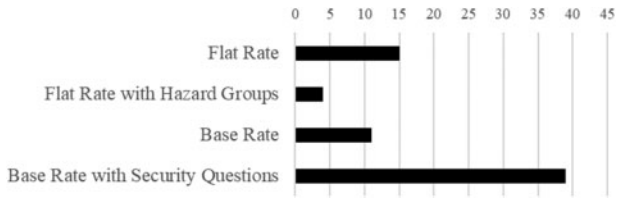


**Figure 6:** Rate schedule categories ($n = 69$)

## How do carriers assess cyber risk?

Of the 235 total policies examined, 72 contained a rate schedule, 3 of which were duplicates.[37] The 69 remaining forms were then segmented into 4 categories according to how they priced the premium. First, we distinguished between "flat rate pricing", and "base rate pricing". The flat rate pricing approach, as the name suggests, provides a single rate to all insured, regardless of their size, or any specific security controls by the insured, while the "base rate" pricing approach uses a series of lookup tables and modifiers to compute the premium, such as modifiers relating to the applicant's standard insurance criteria (e.g. limits, retention, claims history, etc.), and the applicant's industry. In addition, for each of the flat rate and base rate pricing structures, we also identified policies that incorporated either basic hazard metric (coded as "flat rate with hazard groups"), or information about the firm's security technologies, practices, and procedures (coded as "base rate with security questions").

The relative distribution of categories from our dataset is shown in Figure 6. Overall, there were 15 flat rate policies with 4 more that also used hazard groups (for a total of 19 flat rate policies). Of the base rate policies, there were 11 standard base rate policies, and an additional 39 that incorporated questions related to the firm's security posture (for a total of 50 base rate policies).[38]

We next examine each of these categories in detail.

37 We also detected a number of policies which used a similar format as one another, but varied in the numbers used, and so these policies were counted individually.

38 Note that we found two duplicate security-enhanced policies, and two flat-rate policies which were similar, but not identical. In addition, while a number of forms used similar formatting or templates, they varied in their pricing.

**Table 4:** Simple rate development

| Coverage | Frequency | Severity | Expected loss (lost cost) | Profit load | Premium |
|---|---|---|---|---|---|
| Computer attack | 0.20% | $49 800 | $99.60 | 35% | $153 |
| Network security liability | 0.17% | $86 100 | $147.23 | 35% | $227 |

**Table 5:** Base premiums by revenue

| Revenue (in millions) | Annual gross base premiums |
|---|---|
| $0–$10 | $1913.91 |
| $10–$20 | $2602.92 |
| $20–$50 | $3502.46 |
| $50–$100 | $5224.98 |

## Flat rate

The simplest approach to computing premiums was a fixed price for first and third party coverage to all insureds. While this approach offers a quick method for establishing premiums, it affords no differentiation by firm or industry. We found these polices generally offered to smaller companies. For example the CyberOne policy, developed by Insurance Services Organization (ISO), is used by many smaller insurance companies and offers first- and third-party premiums as shown in Table 4.

The frequency column reflects the probability of a given loss event, computed annually, while the severity reflects (presumably) the mean annual loss. The expected loss column is the product of the first two columns, and the profit load is the standard method by which insurance carriers cover costs and expenses. Note that while this approach was used by many carriers, there was some variation across carriers with regard to the frequency, severity, profit loading, and therefore premiums as they incorporated other information.

For context, in other research that examined the cost of cyber incidents, the median cost was found to be $170 000, with a probability of loss around 0.6% across the top 10 most risky industries, producing an expected loss of $1020 – considerably higher than the values shown above [37].

The final premium is then a function of the expected cost, and the profit load of 35%. From the policies examined in this research, profit loading ranged from 25% to 35%. Factoring in the profit loading then produces the final premiums of $153 and $227 for computer attack coverage and network security liability coverage, respectively.[39] Note that these premiums typically apply to policies with limits of $100 000 and deductible of $10 000.

Overall, this approach is simple and straightforward. However, it relies entirely on estimates of frequency and severity of cyber events and litigation costs.

## Flat rate with hazard groups

Even though only four policies we encountered fit into this category, we include it here for completeness. As with the standard flat rate policies, these policies define a fixed price premium, but with a single modifier based on a generic hazard group assigned by business type (again, typically for small businesses). For example, POL-46 described the following hazard groups:

39 That is $99.60 / (1 – 0.35) = $153.

**Table 6:** Retention by asset size

| Asset size (in millions) | Base rate | Base retention |
|---|---|---|
| to $100 | $5000 | $25 000 |
| $100 to $250 | $7000 | $25 000 |
| $250 to $500 | $8500 | $50 000 |
| $500 to $1000 | $11 000 | $100 000 |
| $1000 to $2500 | $14 000 | $150 000 |
| $2500 to $5000 | $16 500 | $250 000 |
| $5000 to $10 000 | $20 000 | $250 000 |
| $10 000 to $25 000 | $26 000 | $500 000 |
| $25 000 to $50 000 | $35 000 | $500 000 |
| $50 000 to $75 000 | $41 000 | $1 000 000 |
| $75 000 to $100 000 | $45 000 | $1 000 000 |

**Table 7:** Variation in premiums for $100m in sales or assets

| Policy | Premium |
|---|---|
| POL-55 | $3300 |
| POL-41 | $3500 |
| POL-56 | $3965 |
| POL-37 | $4000 |
| POL-6 | $5000 |
| POL-88 | $6, 000 |
| POL-32 | $7500 |
| POL-33 | $42 000[a] |

[a]For a $0 retention.

**Table 8:** Limits factor

| Limits | Factor |
|---|---|
| $500 000 | 0.809 |
| $1 000 000 | 1.000 |
| $2 000 000 | 1.132 |
| $3 000 000 | 1.245 |
| $4 000 000 | 1.371 |
| $5 000 000 | 1.405 |

- "Low Hazard – Insured has a website for informational purposes only or small amounts of sales from manufacturers whose main distribution channel is through retailers.
- Medium Hazard – Insured conducts business, at least partially, over their website and/or retain credit card numbers as well as other potentially sensitive information.
- High Hazard – Insured conducts a potentially large portion of their business through their website or retain sensitive information such as social security numbers or have some combination of both."

Examples of low hazard businesses, per this policy, were accounting offices, automobile shops and barber shops, while

**Table 9:** Claims history

| Category | Min | Max |
|---|---|---|
| Very favorable | 0.75 | 0.85 |
| Favorable | 0.9 | 0.99 |
| Average | 1.0 | 1.0 |
| Slightly unfavorable | 1.01 | 1.15 |
| Materially unfavorable | 1.16 | 1.25 |
| Very unfavorable | 1.26 | 1.4 |
| Extremely unfavorable | 1.41 | 1.7 |

businesses considered high hazard were electronics stores, and home improvement stores.

## Base rate

As mentioned, 50 of the policies in our dataset used a base rate pricing model.[40] That is, the base premium is assessed as a function of the insured's annual revenues or assets (or, with some niche products, number of employees or students). This base premium is then multiplied by variables relating to standard insurance and industry-related factors. We describe the base rate approach first, followed by the standard insurance properties, and then the industry-related factors.

The factor that assigns the greatest influence on the premium is the base asset value or revenues of the applicant's firm. For example, Table 5 provides one example of how a policy initially defines the premiums, as a function of firm revenue.

And Table 6 shows premiums with associated retention (deductible) by asset size for POL-6.

Table 7 provides a sense of the variation in premiums found in our analysis across carriers, for a firm with $100 million in sales (or assets), a $1 million limit and $10 000 deductible. Notice the range from just $3300 to over $7500 (with one policy charging a drastically higher premium, but with $0 retention/deductible). These prices, of course, would not reflect the final price, but it does present one perspective in pricing.

### Standard insurance factors

Standard insurance factors include variables such as changes to the limits or deductible (retention) of a policy. For example, the greater the limits, or the smaller the deductible, the larger will be the premium, as shown in Table 8.

In addition, the premium will be modified based on factors such as coinsurance, time retention, prior acts, extended reporting period, and business interruption. Co-insurance adjusts for whether the insured carries coverage with other carriers. Time retention and extended reporting period adjust for the length of time an insured signs the contract, and is decreasing in the duration of the insurance contract.

Historical claims refers to the number of times the insured has suffered an incident and filed a claim in past years. Premiums typically increase about 10% for each event.[41] However, one carrier (POL-33) provides a more descriptive offering for claims history, as shown in Table 9.

Here, the min and max values provide a range of pricing modifiers for the insurance underwriters based on a list of considerations defined in the policy, such as: the number and size of claims made annually, history of litigation against the insured, and any "corrective measures implemented to limit the same wrongful acts from occurring again" (POL-33). That is, based on firm characteristics, the underwriter may choose to multiply (either reduce, or increase) the premium by as much as 0.85 up to 1.7.

**Table 10:** Business interruption

| Industry | Waiting period (hour) | Business interruption charge (%) |
|---|---|---|
| Auto dealership | 10 | 10.0 |
| Automotive services | 10 | 10.0 |
| Domestic services (e.g. plumbers, electricians, gardeners) | 8 | 5.0 |
| E-commerce | 24 | 50.0 |
| Education – colleges/universities/higher education | 8 | 5.0 |
| Professional services (excluding legal services) | 12 | 25.0 |
| Realtor – commercial/residential | 10 | 10.0 |
| Restaurant | 10 | 10.0 |
| Retail | 24 | 50.0 |
| Sports clubs/gyms | 8 | 5.0 |
| Telecommunications | 24 | 50.0 |

A few policies provided coverage for business interruption in the event of a data breach or security incident. For example, POL-2 defined the additional cost of business interruption as shown in Table 10.[42]

### Industry classification

Next, carriers attempt to control for risks to the insured based on the industry in which it operates. However, from the policies examined in this research, there was no consistency regarding approach, or any consensus on what the insurance industry would consider the "most" risky.

POL-18 assigns the energy, entertainment and hospitality sectors a weighting of 1.0 (meaning no adjustment – essentially neutral risk), while firms in the accounting, advertising, construction, manufacturing industries receive a weighting of 0.85 (less risky), and firms in the bio-tech, data aggregation, gaming, and public sectors receive a weighting of 1.2 (more risky). How these relative weightings are determined is unclear and never described in the policies we examined.

Another carrier distinguished among four hazard classes with premium modifiers as shown in Table 11.[43]

Notice how Class 2 (firms which process financial information) is assigned the datum, implicitly stating that SSN and high volume transaction information are considered more risky. Further, notice that four significant digits are used for these factor weightings. It is unclear how these figures were derived, or whether such precision at all accurately reflects true risk.

Another carrier (POL-32) takes a more aggregate approach by differentiating nonprofit, for-profit, and only a few other industries, as shown in Table 12.

40 Not to be confused with a statistical base rate or the base rate fallacy.
41 POL-32.
42 For brevity, we only show a sample of the full table.
43 POL-30.

**Table 11:** Four hazard classes

| Class | Description | Factor |
|---|---|---|
| 1 | Businesses whose primary personal information is relative to employees | 0.804 |
| 2 | Businesses that keep financial or account number information on individual customers but do not keep customers' Social Security numbers | 1.000 |
| 3 | Businesses with customers' Social Security numbers | 1.497 |
| 4 | Entities that collect and store a high volume of particularly sensitive personal information, are at high risk of loss or theft of that information and are subject to structural restraints on their security spending | 1.905 |

**Table 12:** Industry risk

| Industry classification factor | Weighting |
|---|---|
| Nonprofit, nonmedical | 1.0 |
| For profit, manufacturer | 1.5 |
| For profit, wholesale | 1.5 |
| For profit, nontechnical service provider | 1.5 |
| Computer consultants | 2.0 |
| System integration | 2.0 |
| Software manufacturer | 2.0 |
| Retail | 3.0 |
| Healthcare | 3.0 |
| Accountants | 3.0 |
| Financial | 4.0 |
| Large risk (over $250M revenue) | 5.0 |
| All other | 3.0 |

**Table 13:** Basic security modifiers

| Category | Modification | | |
|---|---|---|---|
| | Below Avg | Avg | Above Avg |
| Privacy controls | 1.20 | 1.00 | 0.80 |
| Network security controls | 1.20 | 1.00 | 0.80 |
| Content liability controls | 1.20 | 1.00 | 0.80 |
| Laptop and mobile device security policy | 1.10 | 1.00 | 0.90 |
| Incident response plan | 1.10 | 1.00 | 0.90 |

Next we examine the factors seen in the most sophisticated and detailed policies – those that account for information security controls by the applicant.

## Base rate with security questions

The most sophisticated approach used by 39 (57% of) policies examined in our dataset accounted for characteristics of the applicant's information security controls when determining the final premium pricing. Adjustments based on the applicant's actual security posture vary widely across polices, ranging from basic risk categories to more detailed metrics. One very simple approach (POL-44) considers broad categories of data protection, and adjusts based on qualitative ratings above or below what one may consider to be "average" maturity of controls, as shown in Table 13.

While simple (and possibly appropriate), this particular policy provides no guidance on how an underwriter is supposed to assess an applicant based on these properties. For example, there is no rubric provided as to differentiate "Below Average" from "Above Average" or even what would be included in a firm's collection of privacy controls.

**Table 14:** Security factor weighting[a]

| Rating | Weighting |
|---|---|
| Excellent | 0.75–0.85 |
| Good | 0.85–1.00 |
| Fair | 1.00–1.25 |
| Poor | 1.25–1.50 |

[a]*Source*: POL-64. See also POL-41.

A slightly more detailed and thoughtful approach was found in POL-64 which differentiates a firm's overall security posture along six dimensions (factors): data classification, security infrastructure, governance, risk and compliance, payment card control, media controls, and computer system interruption loss. Each factor provided four qualitative options (poor, fair, good, excellent) with a weighting as shown in Table 14.

The benefit of this approach relative to other simpler or more complex approaches is that it affords a reasonable tradeoff between specificity and practicality. For example, other polices adjust the premium based on specific answers to self-assessment questionnaires (whether the firm uses two-factor authentication, industry standard firewalls, proper best practices), it is highly unlikely that any insurance underwriter would know the marginal reduction in risk that any of these provide. The information simply does not exist to determine a meaningful answer. Therefore, this approach affords the underwriter the ability to investigate a firm's controls and make reasonable assessments. This policy also intelligently provides useful scoring rubrics for each category. For example, the data classification category describes the following:

> The Data Classification Factors are determined by assigning a hazard group factor which is based on the type(s) of data handled, processed, stored or for which the Insured is otherwise responsible for safeguarding. Examples of Data Types are credit card numbers, financial account information and/or personal health information. The appropriate factor should be applied multiplicatively. What type of data is processed, stored or maintained by or on behalf of the insured? Can the data be used to create a false identity, i.e., SSN, DOB, or not, i.e., e-mail address, passwords? Is the data subject to regulation (federal or state), i.e., protected health information (PHI) under HIPAA or driver's license numbers (PII) under state notification laws, etc. Does the data include corporate confidential information of a third party, such as trade secrets and intellectual property?

Other policies took a different approach and adjusted the premium based on the firm's responses to questions from the security questionnaire. For example, POL-6 included the following adjustments such as shown below.

| "(3) Is the disaster recovery plan tested at least annually? Answer YES to | Factor |
|---|---|
| Three of the above questions | 0.80 to 0.90 |
| Two of the above questions | 0.91 to 0.99 |
| One of the above questions | 1.00 to 1.05 |
| None of the above questions | 1.06 to 1.15 |
| (4) Did the total number of targeted computer attacks increase, decrease or remain unchanged in the past 2 years? | Factor |
| Decrease | 0.85 to 0.95 |
| Unchanged | 1.00 |
| Increase | 1.10 to 1.20 |
| (5) Are penetration tests conducted on the insured's network at least annually? | Factor |
| Yes | 0.85 to 0.95 |
| No | 1.10 to 1.20" |

## How are premiums finally computed?

From the policies examined (excluding for the flat rate pricing policies), once the base asset/revenue value is determined, the final premium is computed as the linear product of each of the factors contained in the rate schedule. POL-20 describes the process as, "Pricing is calculated by applying modification factors to a base premium. The modification factors are determined by various criteria including the Limit of Liability and Deductible purchased, the coverage enhancements or restrictions negotiated with the insured, and the risk's financial characteristics. All modification factors are multiplicative, unless otherwise indicated". As previously described, some policies may only consider a few factors, while others may include many. For example, the premium in POL-20 is computed as:

Premium = [Base Premium] x
[Loss Rating] x
[Professional Experience] x
[Longevity of Operations] x
[Use of Written Contracts] x
[Risk Characteristics] x
[Prior Acts Factor] x
[Coverage Adjustment] x
[Deductible]

While the formula in POL-6 is composed of 6 groups of factors and 13 separate security-related questions, producing a final expression of:

Premium = (Section 1 Base Rate) x
(Section 2 Industry Factor) x
(Section 3.1 Increased Limits Factor) x
(Section 3.2 Retention Factor) x
(Section 3.3 Coinsurance Factor) x
(Section 6 Third-Party Modifier Factors)

POL-100 further extends expands the security properties, producing the following expression:

Final Premium = (Third Party Liability Base Rate) +
(First Party Costs Base Rate, if elected) x
(Limit Factor) x
(Retention Factor) x
(Data Classification Factor) x
(Security Infrastructure Factor) x
(Governance, Risk and Compliance Factor) x
(Payment Card Controls Factor) x
(Media Controls Factor) x
(Computer System Interruption Loss Factor, if applicable) x
(Retroactive Coverage Factor) x
(Claims/Loss History Factor) x
(Endorsements Factor, if applicable)

## Summary

In this section, we examined a sample of cyber insurance rate schedules and achieved three main insights. First, we provide exposure of how insurance carriers justify the prices they charge, and what, exactly, they know (and do not know) about how to price cyber risk (i.e. guessing, using competitor pricing, leveraging other lines of business). Second, we identified the pricing strategies used by carriers (flat rate, flat rate with hazard groups, base rate, and base rate with security questions), and third, we provide transparency regarding the factors used by carriers in pricing risk (base rate, industry, and a series of data security and privacy modifiers), and presented the actual algorithms used to price premiums.

From our analysis, the first and most important firm characteristic used to compute insurance premiums was the firm's asset value (or revenue) base rate, rather than specific technology or governance controls. This appears to be the single most common proxy for firm size, and therefore risk.

While some carriers have sophisticated algorithms for premium estimates, policies that cater to small business are very simple. In addition, premiums that capture third party losses (i.e. liability coverage) are generally more costly than those associated with first party losses, suggesting that carriers expect legal actions to be more expensive relative to direct losses suffered by the insured.

While a few carriers incorporate specific information collected from the policy's security self-assessment forms, many policies used more generic security risk categories (e.g. high, med, low). And while many policies incorporate industry factors into the underwriting process, no explanation or justification for how the actual risk weighting is provided. Further, the industries listed rarely match standard coding schemes like SIC or NAICS.

Beyond the specific equations, however, it is unclear which level of sophistication of premium calculation is optimal for the firm, and is best able to assess an applicant's risk. Indeed, this remains an outstanding issue among carriers.

## Limitations

There are a number of important limitations to this research. First, our analysis and conclusions reflect results based a sample of all insurance policies covered in the USA. Naturally, this suggests that further analysis incorporating more policies across more states may reveal additional or even different results. However, that being said, based on our previous examination of state regulations pertaining to the insurance industry, we have no *a priori* reason to believe that there will be any material differences across the USA, or that our findings would change in any material way.

The second potential limitation concerns the matter of admitted versus nonadmitted markets. If it were true that most cyber insurance coverage were written in the nonadmitted markets (i.e. markets that we do not observe), this would reduce the generalizability of our results beyond just the nonadmitted market. In effect, we would only be observing a sample of the overall population of polices. However, based on our preliminary analysis of the coverages and applications, we see no material differences in policies between these markets. That being said, we are unlikely to observe the rate schedules and algorithms for many policies within the nonadmitted market.

## Conclusion

This research has presented the first rigorous thematic analysis of cyber insurance policies filed by insurance companies with state insurance regulators. We collected over 235 policies from New York, Pennsylvania, and California, as well as policies posted publicly on carriers' websites, and separately examined three main components: the coverage, the application questionnaires, and the rate schedules.

Overall, we find that there is a very strong similarity (more similar that expected) across the covered losses, with more variation in exclusions. This suggests that carriers, by and large, are somewhat consistent in identifying cyber perils, and have a certain amount of confidence in their ability to price these risks.

The questionnaires, as part of the required regulatory filings by insurance firms in the admitted market, give interesting insights about what information is (and is not) collected. For example, they request what seems to be an appropriate balance between technical, organizational, and policy/procedure questions, as shown in Figure 4. While there is no formal theory by which to gauge the relative balance of questions, a visual inspection of the relative number of questions does not reveal any outlier categories. On the other hand, they do not provide insights about whether and if what additional information insurers may acquire from third-party providers to assess risk beyond the level of a single insured entity (e.g. industry and market risk regarding cybersecurity). For instance, risk information about the security posture of third-party service (e.g. cloud) providers and intermediaries that an insured relies on, may be difficult to obtain from a single insured entity. Yet, an insurer may have interest in the risk posture of a service provider that accumulates risk across multiple insured entities. A cloud computing provider may be such an example due to the dependencies of multiple insured entities covered by a single insurer. Such risk information may be elicited from other sources than a security questionnaire.

Finally, regarding the rate schedules, we found a surprising variation in the sophistication (or lack thereof) of the equations and metrics used to price premiums. Many policies examined used a very simple, flat rate pricing (based a single calculation of expected loss), while others incorporated more parameters such as the firm's asset value (or firm revenue), or standard insurance metrics (e.g. limits, retention, coinsurance), and industry type. More sophisticated policies included information regarding information security controls and practices as collected from the security questionnaires.

In defense of the insurance carriers, however, while the equations may indeed be rudimentary, and based on very subjective inputs, there is no authoritative source for cyber risk assessment. Indeed, proper assessment, and quantification of an organization's security posture is something that the information security industry has been struggling with for decades, and which, to this day, remains elusive.

By examining these components of insurance contracts, we hope to provide additional transparency and insights into this growing market of cyber insurance.

## Acknowledgments

We would like to thank Adam Hamm, Igor Mikolic-Torreira, Elizabeth Petrun Sayers, Lori Uscher-Pines, participants of the 2017 Workshop on the Economics of Information Security (WEIS), the 2017 Research Conference on Communications, Information and Internet Policy (TPRC), this journal's area editor, and the anonymous reviewers. We would also like to thank RAND's Institute for Civil Justice, and the William and Flora Hewlett Foundation for their generous support.

## References

1. Takahashi D. IBM Security Study: Mega Data Breaches Cost $40 Million to $350 Million. https://venturebeat.com/2018/07/10/ibm-security-study-mega-data-breaches-cost-40-million-to-350-million/ (17 September 2018, date last accessed).
2. Rough Notes. Specialty lines markets: The need is now standard. Can the product become so?, 2017. http://roughnotes.com/can-cyber-insurance-standardized/ (10 September 2018, date last accessed).
3. WSJ, Cyber Insurance: How to Address Obstacles to Growth https://deloitte.wsj.com/cfo/2018/03/21/cyber-insurance-how-to-address-obstacles-to-growth/ (10 September 2018, date last accessed).
4. Price Waterhouse Coopers. Insurance 2020 & Beyond: Reaping the Dividends of Cyber Resilience, 2015. https://www.pwc.com/gx/en/insurance/publications/assets/reaping-dividends-cyber-resilience.pdf (10 September 2018, date last accessed).
5. Allianz. *A Guide to Cyber Risk Managing the Impact of Increasing Interconnectivity, Global Corporate & Specialty,* 2015. https://www.allianz.com/v_1441749600000/media/press/document/other/Allianz_Global_Corporate_Specialty_Cyber_Guide_final.pdf (10 September 2018, date last accessed).
6. Insurance Information Institute. Industry Leaders Expect Commercial Lines to Grow at Greater Pace than Personal Lines; Cyber to Lead the Way, I.I.I. Survey Finds. III. New York. http://www.iii.org/press-release/industry-leaders-expect-commercial-lines-to-grow-at-greater-pace-than-personal-lines-cyber-to-lead-the-way-iii-survey-finds-012317 (24 January 2017, date last accessed)
7. Aon Benfield. *Insurance Risk Study - Growth, Profitability, and Opportunity,* 2014. http://thoughtleadership.aonbenfield.com/documents/20140912_ab_analytics_insurance_risk_study.pdf (11 September 2018, date last accessed).
8. Willis. That's the Board and They're Asking about Cyber Risk. Willis Insights, 2014.
9. Marsh. Benchmarking Trends: As Cyber Concerns Broaden, Insurance Purchases Rise. 2015.
10. Marsh. Benchmarking Trends: Operational Risks Drive Cyber Insurance Purchases. 2016.
11. NAIC. Early NAIC Analysis Sheds Light on Cybersecurity Insurance Data. Washington, D.C., 2016. http://www.naic.org/Releases/2016_docs/cybersecurity_insurance_data_analysis.htm (7 September 2018, date last accessed).
12. Fitch Ratings. Fitch: U.S. Cyber Insurance Premiums Total $1B Per New Supplemental Filing, 2016. https://www.fitchratings.com/site/pr/1010744 (20 January 2017, date last accessed).

13. Betterley R. *The Betterley Report: Cyber/Privacy Insurance Markey Survey* 2012. The Betterley Report.
14. Airmic. Airmic Review of Recent Developments in the Cyber Insurance Market & commentary on the increased availability of cyber insurance products. London: Airmic.
15. Johnson B, Böhme R, Grossklags J. Security games with market insurance. In Baras JS, Katz J and Altmann E. (eds), *Decision and Game Theory for Security*. Lecture Notes in Computer Science 7037. Berlin Heidelberg: Springer, 117–30, 2011.
16. Böhme R, Schwartz G. Modeling cyber-insurance: towards a unifying framework. In: *Workshop on the Economics of Information Security (WEIS)*. Cambridge, MA: Harvard University, 2010.
17. Böhme R. Towards insurable network architectures. *Informat Technol* 2010;**52**:290–93.
18. Böhme R, Kataria G. Models and measures for correlation in cyber-insurance. In: *Workshop on the Economics of Information Security (WEIS)*. UK: University of Cambridge, 2006a.
19. Böhme R, Kataria G. On the limits of cyber-insurance. In: Fischer-Hübner S, Furnell S and Lambrinoudakis C (eds), *Trust, Privacy and Security in Digital Business (TrustBus/DEXA 2006)*. Lecture Notes in Computer Science 4083. Berlin Heidelberg: Springer, 2006b, 31–40.
20. Ehrlich I, Becker G. Market insurance, self-insurance, and self-protection. *J Polit Econ* 1972;**80**:623–48.
21. Marotta A, Martinelli F, Nanni S, *et al*. A Survey on Cyber-Insurance, 2015. http://www.iit.cnr.it/en/node/36039 (31 December 2018, date last accessed).
22. Majuca RP, Yurcik W, Kesan JP. The Evolution of Cyberinsurance, 2006. https://arxiv.org/abs/cs/0601020 (31 December 2018, date last accessed).
23. Baer WS, Parkinson A. Cyber insurance in IT security management. *IEEE Security & Privacy* 2007;**5**:50–56.
24. Woods D, Agrafiotis I, Nurse JRC, *et al*. Mapping the coverage of security controls in cyber insurance proposal forms. *J Internet Serv Appl* 2017;**8**:8.
25. Braun V, Clarke V. Thematic analysis. In: H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (eds), *APA Handbook of Research Methods in Psychology*, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological. Washington, DC, US: American Psychological Association, 2012, 57–71.
26. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. *Nursing Health Sci* 2013;**15**:398–405.
27. Guest, G, MQueen KM, Namey, E. *Applied Thematic Analysis*. Sage Publications, 2012.
28. Schwarcz D. Reevaluating standardized insurance policies. *Univ Chicago L Rev* 2011;**78**:1263.
29. Davis CS, Johnston JE, Pierce MW. Overdose epidemic, prescription monitoring programs, and public health: a review of state laws. *Am J Public Health* 2015;105:e9–e11.
30. Yu CH, Jannasch-Pennell A, DiGangi S. Compatibility between text mining and qualitative research in the perspectives of grounded theory, content analysis, and reliability. *Qualit Report* 2011;**16**:730–44. Available at http://nsuworks.nova.edu/tqr/vol16/iss3/6. Last accessed September 10, 2018.
31. Ingle WK, Wisman RA. Extending the work of Cowen and Fowles: a historical analysis of Kentucky Teacher Contracts. *Educ Policy* 2018;**32**:313–33.
32. Greenwald J. *Specialty Market Keeps Grip on Cyber Risk*. *Business Insurance*, 2016. http://www.businessinsurance.com/article/00010101/NEWS06/912310137/Specialty-market-keeps-grip-on-cyber-risk (13 February 2017, date last accessed).
33. NAIC. Report on the Cybersecurity Insurance Coverage Supplement, 2017. http://www.naic.org/documents/committees_ex_cybersecurity_tf_report_cyber_supplement.pdf (10 September 2018, date last accessed).
34. Glaser, B, Strauss G, L. *Anselm. The Discovery of Grounded Theory: Strategies for Qualitative Research*, 1967. Chicago: Aldine Publishing Company.
35. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. *Field Methods* 2006;**18**:59–82. doi: 10.1177/1525822X05279903.
36. Bowen GA. Document analysis as a qualitative research method. *Qualitative Research Journal* 2009;**9**:27–40.
37. Romanosky S. Examining the costs and causes of cyber incidents. *J Cybersecurity* 2016;**2**:121–35. https://doi.org/10.1093/cybsec/tyw001.