Quantifying the Costs of Data Breaches

Siddharth Dongre, Sumita Mishra, Carol Romanowski, Manan Buddhadev

▶ **To cite this version:**

Siddharth Dongre, Sumita Mishra, Carol Romanowski, Manan Buddhadev. Quantifying the Costs of Data Breaches. 13th International Conference on Critical Infrastructure Protection (ICCIP), Mar 2019, Arlington, VA, United States. pp.3-16, 10.1007/978-3-030-34647-8_1 . hal-03364563

HAL Id: hal-03364563

https://inria.hal.science/hal-03364563v1

Submitted on 4 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Chapter 1

QUANTIFYING THE COSTS OF DATA BREACHES

Siddharth Dongre, Sumita Mishra, Carol Romanowski and Manan Buddhadev

Abstract Recent years have seen increases in the number of data breaches. This chapter attempts to quantify the impacts of data breaches in terms of the monetary costs incurred by providers and consumers. This is important because data breaches are a major factor when allocating funds for security controls. Case studies involving the Equifax incident in 2017 and the Target incident in 2013 are employed to demonstrate that the cost impacts of data breaches are significant for providers as well as consumers. The cost components in the overall cost function for providers and consumers are presented. Guided by open-source data, the cost components in the provider portion of the cost function are expressed as best-fit functions of time since the data breach. An important point in the cost quantification is that equal weights are assigned to the costs incurred by the provider and the consumers.

Keywords: Data breaches, cost analysis, providers, consumers

1. Introduction

The average cost of data breaches has increased by 6.4% during the past year, with an average increase of 4.8% in the cost of each stolen record [9]. These statistics point to a general increase in the cost impacts of data breaches. Clearly, it is imperative to understand the many aspects of data breaches in terms of their cost impacts.

A data breach is defined as an incident that leads to the loss or exposure of sensitive information. The focus of this chapter is on specific data breaches that have exposed personal information such as social security numbers, driver's license information, dates of birth, credit card numbers, telephone numbers and residential addresses, and/or other information that malicious entities could use to perpetrate activities such as identity theft and credit card fraud.

The root causes of data breaches vary from incident to incident. Most data breaches occur due to vulnerabilities in web applications hosted by providers or through cyber-espionage activities [9]. Since the majority of breaches have these two vectors, their costs appear to be more significant to providers than consumers.

Acquisiti et al. [1] have analyzed the impact of privacy breaches on the market value of providers. Their research demonstrates that a data breach has a statistically-significant negative impact on the market value of a company on the day that the breach is announced.

Romanosky [21] has analyzed the causes and costs of cyber incidents in an attempt to understand how companies should improve their security postures in order to reduce the risk of data breaches. He states that public concerns regarding data breaches are excessive compared with the financial impacts on companies.

Most research efforts, including the work of Acquisiti et al. [1] and Romanosky [21], analyze the cost impacts of data breaches on providers. Limited research has focused on the cost impacts on consumers. In contrast, the research described in this chapter considers the cost impacts from the perspectives of providers and consumers. Both providers and consumers have to pay to mitigate the negative effects of data breaches. For example, the Equifax data breach of 2017 cost the company approximately $439 million [18], but numerous Equifax consumers also paid a price by becoming victims of identity theft [16] that exposed them to financial losses.

This chapter presents a mathematical formulation that expresses the cost impacts of data breaches. The costs incurred by the provider and consumers have different components, all of which vary with time. Therefore, a cost function for a provider and consumers is developed, which incorporates multiple cost components and weights for the components that vary with time. In the case of providers, the component weights are derived from real data pertaining to the Equifax data breach of 2017 and the Target data breach of 2013. The two case studies were selected because they had significant, direct impacts on providers and consumers, and open-source data related to the breaches and their impacts was available.

2. Cost Function

The cost impacts of a data breach can be broadly expressed as a function of time $C(T)$. Specifically, this cost function is the sum of the costs incurred by the provider and by consumers, $C_p(T)$ and $C_c(T)$, respectively, which are also functions of time. The time T denotes the number of months elapsed since the breach was discovered. Unique weights $W_p \in [0, 1]$ and $W_c \in [0, 1]$ are assigned to the costs incurred by the provider and by consumers, respectively, based on the relative impacts of the two cost perspectives. Thus, the costs incurred due to a data breach at time T months after the breach is given by:

$$C(T) = W_p C_p(T) + W_c C_c(T) \tag{1}$$

Each term in Equation (1) is expressed as the sum of the individual cost components for the provider and consumers, $C_{pi}(T)$ and $C_{cj}(T)$, where $1 \leq i \leq N$ and $1 \leq j \leq M$, and N and M are the numbers of cost components incorporated for the provider and consumers, respectively.

Thus, the costs incurred by the provider and by consumers are given by:

$$C_p(T) = \sum_{i=1}^{N} C_{pi}(T) \tag{2}$$

$$C_c(T) = \sum_{j=1}^{M} C_{cj}(T) \tag{3}$$

Each term $C_{pi}(T)$ and $C_{cj}(T)$ can be further expressed as the sum of the costs incurred each month, which varies with time $t \in [0, T]$ expressed in months:

$$C_{pi}(T) = \sum_{t=0}^{T} C_{pi}(t) \tag{4}$$

$$C_{cj}(T) = \sum_{t=0}^{T} C_{cj}(t) \tag{5}$$

Equations (1) through (5) can be combined to yield the following overall cost function for the provider and consumers:

$$C(T) = W_p \sum_{i=1}^{N} \sum_{t=0}^{T} C_{pi}(t) + W_c \sum_{j=1}^{M} \sum_{t=0}^{T} C_{cj}(t) \tag{6}$$

where the weights are based on well-defined cost component values C_{pi} and C_{cj} for the provider and consumers, respectively. These cost component values vary on a case by case basis. In this work, the cost component values are assigned based on case studies involving the 2017 Equifax and 2013 Target data breaches.

3. 2017 Equifax Data Breach

Equifax is one of the leading credit reporting agencies along with TransUnion and Experian. It provides important services that determine the creditworthiness of consumers based on their credit histories. The information provided by Equifax is used by lenders to decide whether or not to issue credit lines to consumers and to determine the appropriate credit limits.

In July 2017, Equifax became the victim of one of the largest data breaches in history [7]. The breach was traced to a vulnerability in Equifax's web application systems, which were developed using the Apache Struts 2 framework [13].

In March 2017, a few months before the breach, Apache announced a vulnerability in its technology. However, many users, including Equifax, did not

apply the patch. The vulnerability enabled an unknown entity to remotely access Equifax's web application servers and run malicious programs, eventually extracting sensitive data belonging to more than 145 million consumers. Credit card numbers of more than 209,000 consumers were compromised. Private information such as social security numbers, driver's license numbers and dates of birth were also exposed.

Equifax reportedly handled the data breach in an irresponsible manner. It did not notify the affected consumers until two months after the breach was discovered. Equifax executives sold nearly $2 million in stock before the breach was disclosed; however, a special company committee cleared the executives upon finding that they did not know about the breach when they made the transactions [4].

Equifax stock lost billions of dollars within a few months of the announcement of the breach, demonstrating the major impacts that data breaches can have on providers. However, numerous innocent consumers became victims of identity theft and credit card fraud as a result of the breach. Indeed, the Equifax breach is a lesson about the significant impacts that data breaches can have on consumers.

3.1 Components Affecting Data Breach Costs

An analysis of corporate filings and news reports in the aftermath of the Equifax data breach identified several components that may affect the costs incurred by providers. Data from Equifax quarterly reports was used to derive the cost function for each component. The cost function formulas were obtained by applying machine learning algorithms to the available data.

Earnings Loss from Customer Dissatisfaction. Equifax reported that its earnings were affected by customer dissatisfaction – its net income fell 27% to $96.3 million in the third quarter of 2017 [2]. It is safe to assume that the loss in earnings due to customer dissatisfaction is the highest immediately after a breach and decreases gradually over time.

The four data points in Figure 1 show Equifax's net income (earnings loss) figures for four consecutive quarters after the breach. Based on the variation of net income (earnings loss) C_{p1} in millions of dollars over time t in months, the following best-fit function was obtained to express the costs due to customer dissatisfaction as a function of time:

$$C_{p1}(t) = 165.39 - 33t + 3.42t^2 \tag{7}$$

where the parameters $a = 165.39$, $b = -33$ and $c = 3.42$ are specific to the provider, in this case, Equifax.

Market Capitalization Loss from Investor Nervousness. After the breach was publicly announced, Equifax stock value fell sharply because nervous shareholders sold their holdings. Equifax's market capitalization

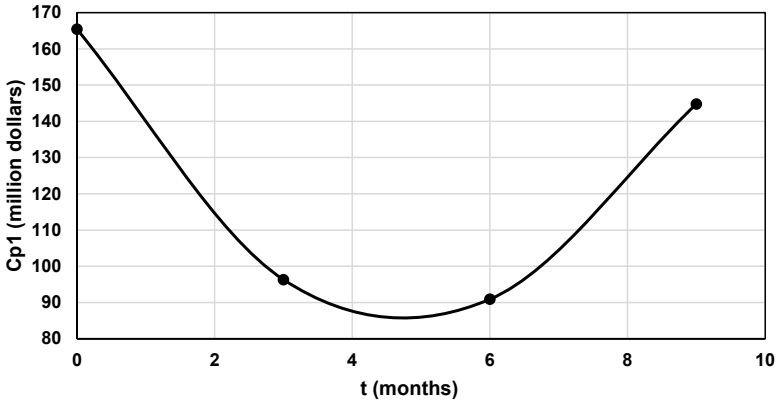

Figure 1. Variation in Equifax's costs (earnings loss) from customer dissatisfaction.

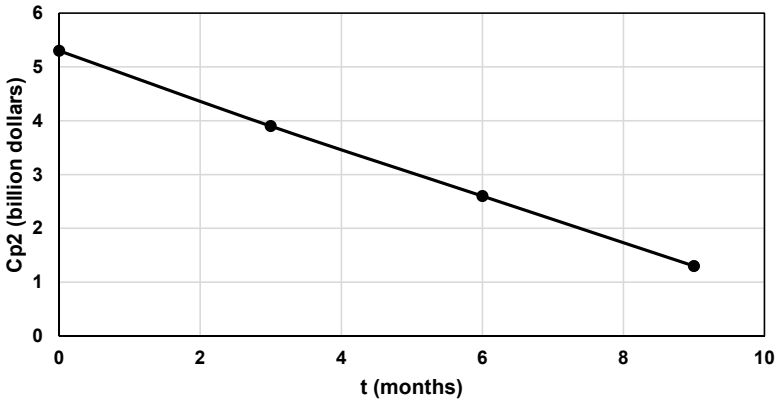

Figure 2. Variation in Equifax's market capitalization loss.

dropped by $5.3 billion [19]. However, after the initial slump, the stock value gradually increased over the next three quarters.

The four data points in Figure 2 show Equifax's market capitalization losses from four consecutive quarterly reports after the breach. The costs associated with this component decrease linearly with time. Based on the variation in the market capitalization loss C_{p2} in billions of dollars over time t in months, the following best-fit function was obtained to express the costs due to investor nervousness as a function of time:

$$C_{p2}(t) = 5.3 - 0.44t \tag{8}$$

where the parameters $c = 5.3$ and $m = -0.44$ are specific to Equifax.

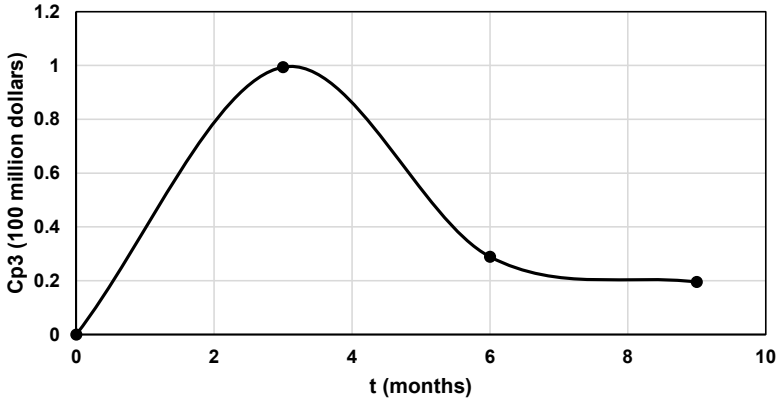

Figure 3. Variation in Equifax's legal and investigation fees.

Legal and Investigation Fees. The four data points in Figure 3 show Equifax's legal and investigation fees from four consecutive quarterly reports after the breach. Equifax spent $99.4 million in fees during the final quarter of 2017 and $28.9 million during the first quarter of 2018 [8, 17, 23]. The costs associated with this component start low, increase gradually and finally decrease again, which exhibits the characteristics of a Gaussian curve.

Based on the variation in the costs associated with legal and investigation fees C_{p3} in hundreds of millions of dollars over time t in months, the following best-fit Gaussian function was obtained to express the legal and investigation fees component as a function of time:

$$C_{p3}(t) = 1.38 \times e^{-(t-3.93)^2/2(1.17)^2} \tag{9}$$

where the parameters $a = 1.38$, $b = -3.93$ and $c = 1.17$ are specific to Equifax.

Customer Services. The four data points in Figure 4 show Equifax's customer services costs from four consecutive quarterly reports after the breach. Equifax paid approximately $64.4 million for customer support services during the final quarter of 2017 and the payments went down to $4.1 million during the first quarter of 2018 [23]. This cost component decreases gradually with time in a manner similar to earnings loss due to customer dissatisfaction. Based on the variation in the costs associated with customer services C_{p4} in millions of dollars over time t in months, the following best-fit function was obtained to express the customer services cost component as a function of time:

$$C_{p4}(t) = -0.14 + \frac{64.54}{2^{t/0.76}} \tag{10}$$

where the parameters $a = -0.14$, $b = 64.54$ and $c = 0.76$ are specific to Equifax.

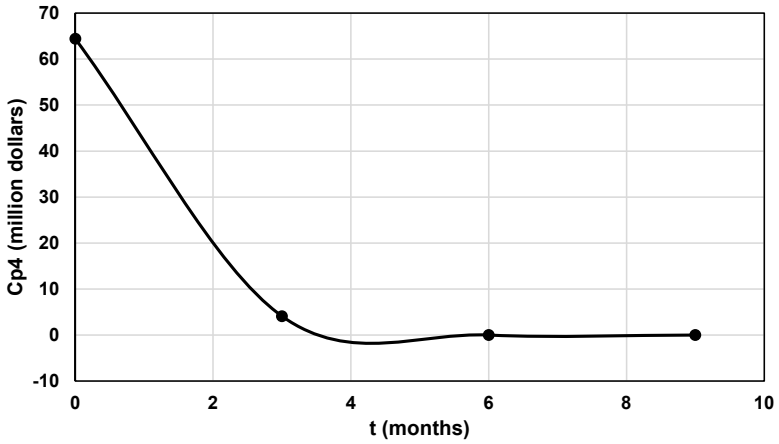

Figure 4. Variation in Equifax's customer services costs.

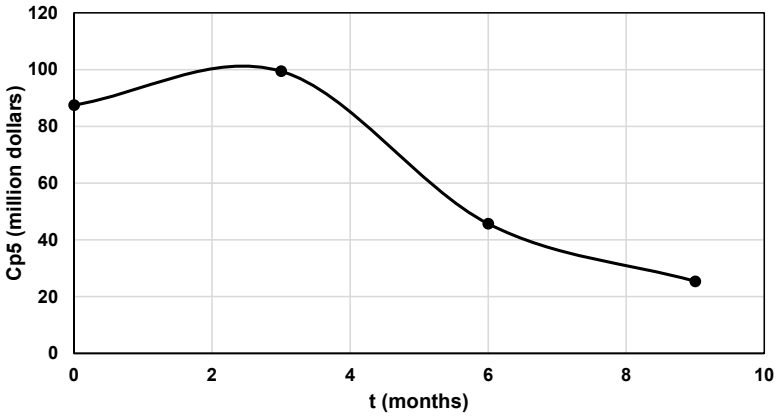

Figure 5. Variation in Equifax's information technology security upgrade costs.

Information Technology Security Upgrades. The four data points in Figure 5 show Equifax's information technology security upgrade costs from four consecutive quarterly reports after the breach. Immediately after the breach was announced, Equifax incurred a one-time charge of $87.5 million, which was presumably spent on incident response and disaster recovery [2]. During the last quarter of 2017, a portion of the $99.4 million spent on fees was due to information technology security upgrades; the upgrade costs dropped to $45.7 million during the first quarter of 2018 [23]. This cost component starts high immediately after the data breach and decreases gradually.

Based on the variation in information technology security upgrade costs C_{p5} in millions of dollars over time t in months, the following best-fit Gaussian

function was obtained:

$$C_{p5}(t) = 99.65 \times e^{-(t-1.79)^2/2(3.79)^2} \tag{11}$$

where the parameters $a = 99.65$, $b = -1.79$ and $c = 3.79$ are specific to Equifax.

4. 2013 Target Data Breach

Target is one of the largest departmental store chains in the United States. It specializes in fast-moving consumer goods. In December 2013, Target became the victim of a massive data breach in which nearly 40 million credit and debit card numbers, and nearly 70 million personal information records were stolen [22].

Several security firms analyzed the data breach to determine the root causes. Their reports state that poor network segmentation, a mistake on Target's part and malicious actions by an adversary contributed to the massive data breach. The adversary reportedly installed BlackPOS malware on point-of-sale terminals to collect sensitive user information, especially credit and debit card numbers. The stolen information was discovered being sold on black market websites [11].

The data breach exposed numerous consumers to identity theft and credit card fraud. It is another example of how the impacts of a data breach on consumers are just as significant as those on the provider.

4.1 Components Affecting Data Breach Costs

Since the Target data breach was announced, several reports have been released that estimate the losses incurred by the company. This section discusses the components that affect the costs incurred by Target as a provider.

Earnings Loss from Customer Dissatisfaction. Target's profits reportedly fell by $440 million during the final quarter of 2013, i.e., immediately after the data breach [15]. In the final quarter of 2014, Target reported a net loss of $2.6 billion during the one year after the breach [20]. It can be assumed that this cost component (earnings loss) reached its maximum value in the first quarter after the data breach and decreased sharply over the course of a year.

The four data points in Figure 6 show Target's net income (earnings loss) figures for four consecutive quarters after the breach. Based on the variation of net income (earnings loss) C_{p1} in billions of dollars over time t in months, the following best-fit Gaussian curve was obtained to express the cost due to customer dissatisfaction as a function of time:

$$C_{p1}(t) = 0.91 \times e^{-(t-7.22)^2/2(3.81)^2} \tag{12}$$

where the parameters $a = 0.91$, $b = -7.22$ and $c = 3.81$ are specific to Target.

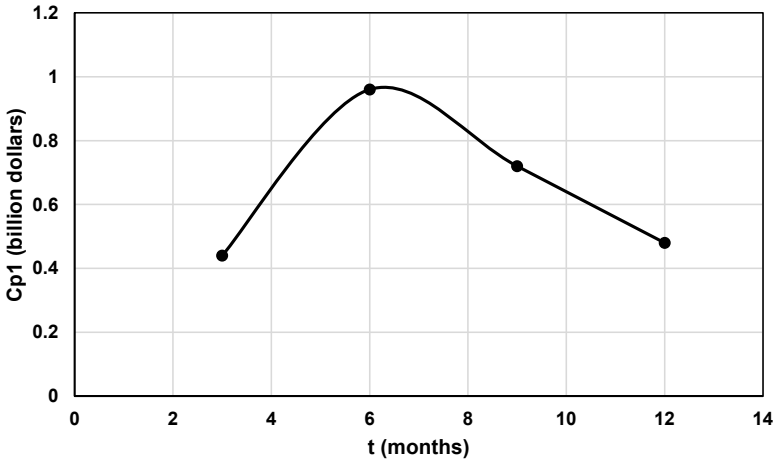

Figure 6. Variation in Target's costs (earnings loss) from customer dissatisfaction.

Legal Fees and Lawsuit Settlements. Target reportedly made settlements totaling more than $153.9 million through May 2017, almost four years after the breach. The major costs incurred by Target during this period were [14]:

- $10 million to settle a class action lawsuit by consumers in March 2015.
- $19 million to MasterCard in April 2015.
- $67 million to Visa in August 2015.
- $39.4 million to banks and credit unions in December 2015.
- $18.5 million to settle actions by 47 state governments in May 2017.

Figure 7 shows the variation in Target's lawsuit settlement costs over a two-year period starting eighteen months after the breach. It is a classic example of how the costs incurred by a provider due to legal actions arising from a data breach are considerable over a long period of time. However, due to the unpredictable nature of legal settlements, it is difficult to express the associated costs as a function of time. The only statement that can be made is that the legal costs are significant over a long period of time.

Other Expenses. Target's 2016 annual financial report estimated that its total costs due to the data breach were $292 million. The annual breakdowns were $17 million in 2013, $145 million in 2014 and $39 million in 2015; information about the 2016 costs was not provided [14]. These figures cover the expenses incurred for incident response and forensics, disaster recovery and information security upgrades.

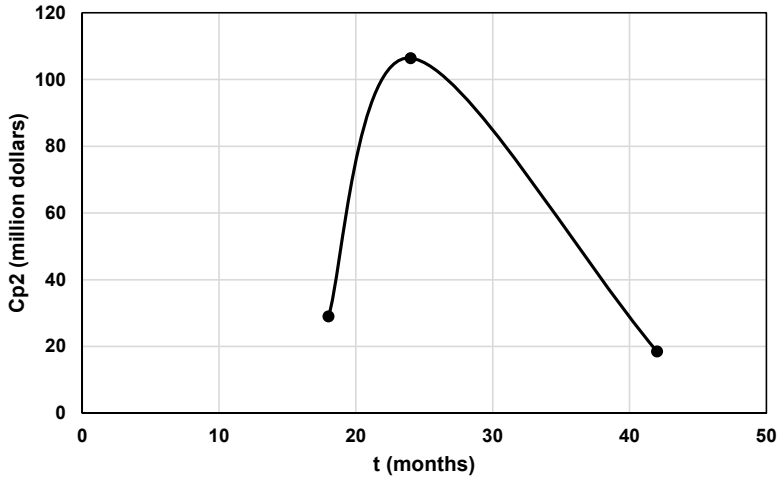

Figure 7. Variation in Target's lawsuit settlement costs over four years.

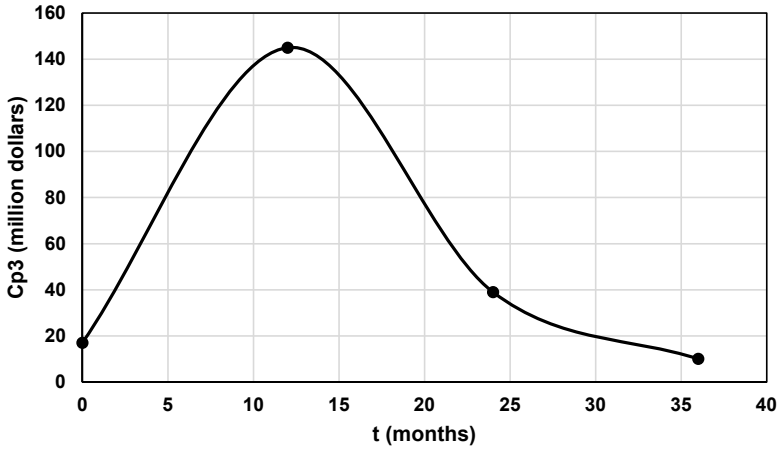

Figure 8. Variation in Target's other costs over three years.

Figure 8 shows the variation in Target's other costs (for incident response and forensics, disaster recovery and information security upgrades) over a three-year period following the data breach. Based on the variation in the other costs C_{p3} in millions of dollars over time t in months, the following best-fit Gaussian function was obtained to express the cost component as a function of time:

$$C_{p3}(t) = 148.57 \times e^{-(t-13.44)^2/2(6.47)^2} \tag{13}$$

where the parameters $a = 148.57$, $b = -13.44$ and $c = 6.47$ are specific to Target.

5. Cost Impacts on Consumers

Data breaches expose sensitive consumer information such as social security numbers, driver's license information, dates of birth, credit card numbers, telephone numbers and residential addresses. Consumer information of this nature can be exploited to perpetrate identity theft and other fraudulent activities that can have devastating financial impacts on consumers.

It is posited that consumers as a whole incur costs that are comparable to those incurred by the provider as a result of data breaches. Therefore, the weights assigned to the costs incurred by the provider and consumers in Equation (6) are equal, i.e., $W_p = W_c = 0.5$. Thus, the overall cost function is given by:

$$C(T) = 0.5 \sum_{i=1}^{N} \sum_{t=0}^{T} C_{pi}(t) + 0.5 \sum_{j=1}^{M} \sum_{t=0}^{T} C_{cj}(t) \tag{14}$$

The following sections discuss four components of the cost function for consumers.

5.1 Identity Theft and Credit Card Fraud Costs

Many consumers whose personal data has been exposed by a breach become unwitting victims of identity theft and credit card fraud. In 2016, 15.4 million consumers were victims of identity theft or fraud and they collectively lost more than $16.2 billion. These figures went up in 2017 with 16.7 million victims losing $16.8 billion in total. On average, every consumer who becomes a victim of identity theft or fraud loses more than $1,000 a year [10]. The costs include notary fees and fax, copying, postage, mileage and calling charges incurred to address identity theft or fraud. The costs also include loss of income as a result of taking time off from work to handle the problems.

5.2 Protection and Monitoring Costs

The exposure of personal information puts consumers at risk of becoming targets of identity theft and credit card fraud. Consumers are urged to enroll in credit monitoring and identity protection services, which cost $120 to $300 annually [3].

5.3 Legal Fees

Victims of data breaches have the right to file lawsuits against providers that may be responsible for the breaches. Attorney expenses vary, but are they still relatively high [5]. Consumers who live in small towns and rural areas may be charged $100 to $200 per hour by experienced attorneys. In metropolitan

areas, attorney fees are $200 to $400 per hour. Attorney fees for complicated data breach cases that require technical expertise are even higher.

5.4 Other Costs

Consumers who are victims of data breaches are highly susceptible to identity theft and credit card fraud. Most victims are unaware that fraudulent activities are being perpetrated until it is too late; there are cases where even minors have become victims of identity theft or fraud [6].

Identity theft victims should consider freezing their credit, which prohibits credit reporting companies from disclosing their credit histories. This also prevents malicious entities from opening fake credit card accounts in their names. Credit freeze requests can cost consumers $2 to $10 per credit reporting agency [24]; several states now ensure that credit freeze requests are free [12].

Consumers who become victims of identity theft face the following severe consequences:

- Difficulty securing credit cards and loans.
- Difficulty securing home mortgages and home rentals.
- High credit card interest rates.
- Difficulty securing jobs.
- Psychological impacts such as distress and anxiety.

6. Conclusions

This research is the first attempt to quantify the costs of data breaches for providers and consumers. This is important because data breaches are a major factor when allocating funds for security controls. The cost components in the overall cost function for the provider and consumers have been identified. Guided by open-source data, the cost components in the provider portion of the cost function have been expressed as best-fit functions of time elapsed since the data breach. An important point in the cost quantification is that equal weights are assigned to the costs incurred by the provider and the consumers.

Future research will attempt to formulate cost components in the consumer cost function as functions of time. This effort will be theoretical as opposed to empirical because of the lack of data pertaining to consumer costs over time.

References

[1] A. Acquisti, A. Friedman and R. Telang, Is there a cost to privacy breaches? An event study, *Proceedings of the Twenty-Seventh International Conference on Information Systems*, article no. 94, 2006.

[2] Agence France-Presse, Massive data breach has cost Equifax nearly $90 million, November 11, 2017.

[3] Consumer Reports, Don't get taken guarding your ID. Do-it-yourself safeguards are just as effective as paid services, September 8, 2014.

[4] Federal Trade Commission, The Equifax Data Breach, Washington, DC (`www.ftc.gov/equifax-data-breach`), 2018.

[5] D. Goguen, How, and How Much, Do Lawyers Charge? Lawyers.com (`www.lawyers.com/legal-info/research/how-and-how-much-do-lawyers-charge.html`), 2019.

[6] K. Grant, Identity theft isn't just an adult problem. Kids are victims, too, *CNBC*, April 24, 2018.

[7] S. Gressin, The Equifax Data Breach: What to Do, Federal Trade Commission, Washington, DC (`www.consumer.ftc.gov/blog/2017/09/equifax-data-breach-what-do`), September 8, 2017.

[8] M. Heller, Equifax hack could cost well over $600M, *CFO Magazine*, March 5, 2018.

[9] IBM Security and Ponemon Institute, 2018 Cost of a Data Breach Study: Global Overview, Cambridge, Massachusetts and North Traverse City, Michigan (`www.ibm.com/security/data-breach`), 2018.

[10] Javelin, Identity fraud hits all time high with 16.7 million U.S. victims in 2017, according to new Javelin Strategy and Research study, Press Release, San Francisco, California (`www.javelinstrategy.com/press-release/identity-fraud-hits-all-time-high-167-million-us-victims-2017-according-new-javelin`), February 6, 2018.

[11] B. Krebs, Who's selling credit cards from Target? *Krebs on Security* (`www.krebsonsecurity.com/2013/12/whos-selling-credit-cards-from-target`), December 24, 2013.

[12] K. Lobosco, Congress just made credit freezes free, *CNN*, May 22, 2018.

[13] J. Luszcz, Apache Struts 2: How technical and development gaps caused the Equifax Breach, *Network Security*, vol. 2018(1), pp. 5–8, 2018.

[14] V. Lynch, Cost of 2013 Target data breach nears $300 million, *Hashed Out* (`www.thesslstore.com/blog/2013-target-data-breach-settled`), May 26, 2017.

[15] Marketwatch, Target's profits down $440M after data breach, *New York Post*, February 26, 2014.

[16] K. McCoy, Equifax data breach: What's changed since last year's huge hack of personal information? *USA Today*, September 7, 2018.

[17] J. McCrank and J. Finkle, Equifax breach could be most costly in corporate history, *Reuters*, March 2, 2018.

[18] PYMNTS, Equifax breach to cost total of $439M (`www.pymnts.com/news/security-and-risk/2018/equifax-cost-275m`), March 5, 2018.

[19] V. Reklaitis, Equifax's stock has fallen 31% since breach disclosure, erasing $5 billion in market cap, *MarketWatch*, September 14, 2017.

[20] J. Roman, Target breach costs: $162 million. Response expenses continue to grow following 2013 incident, *BankInfoSecurity* (www.bankinfosecurity.com/target-breach-costs-162-million-a-7951), February 25 2015.

[21] S. Romanosky, Examining the costs and causes of cyber incidents, *Journal of Cybersecurity*, vol. 2(2), pp. 121–135, 2016.

[22] X. Shu, K. Tian, A. Ciambrone and D. Yao, Breaking the Target: An Analysis of the Target Data Breach and Lessons Learned, arXiv:1701.04940 (arxiv.org/abs/1701.04940), 2017.

[23] Titanadmin, The cost of the Equifax data breach? $242 million and rising, SpamTitan, Tampa, Florida (www.spamtitan.com/blog/cost-of-the-equifax-data-breach-242-million-rising), April 27, 2018.

[24] F. Williams, How credit freezes work and what they cost, CreditCards.com, Austin, Texas (www.creditcards.com/credit-card-news/credit-report-freeze-1282.php), September 13, 2017.