Quantitative Cyber-Security

Colorado State University
Yashwant K Malaiya
CS559
L23

CSU Cybersecurity Center
Computer Science Dept

Presentations/Final Report

Slides should be ready by Wed 11/18/20, but ..

- Post 24 hours in advance of the presentation in the designated canvas forum.
- Schedule will be announced later
- Peer reviews will be needed.

Final report is due on Wed 12/9/20.

Topics

- Probability of a breach
- Breach cost
- Impact of a breach on the stock price

Probability of a breach

- A key component of risk
  - Risk is meaningless without probability of the adverse event
- Very limited data and modeling effort at this time
  - Breaches do not happen regularly, thus the available data for a specific organization is not enough for modelling
  - Data collection is not systematic, many breaches may not be reported.
    - There will never be clean and complete data
  - This is an early phase of research

Breach Probability Model

A proposed model for the probability of a breach for the next

$$\text{P \{breach\}} = Fcountry * FBCM * Findustry * Fbreach_{cause} * Fencryption * Fprivacy * \alpha\, exp\,(-\beta x)$$

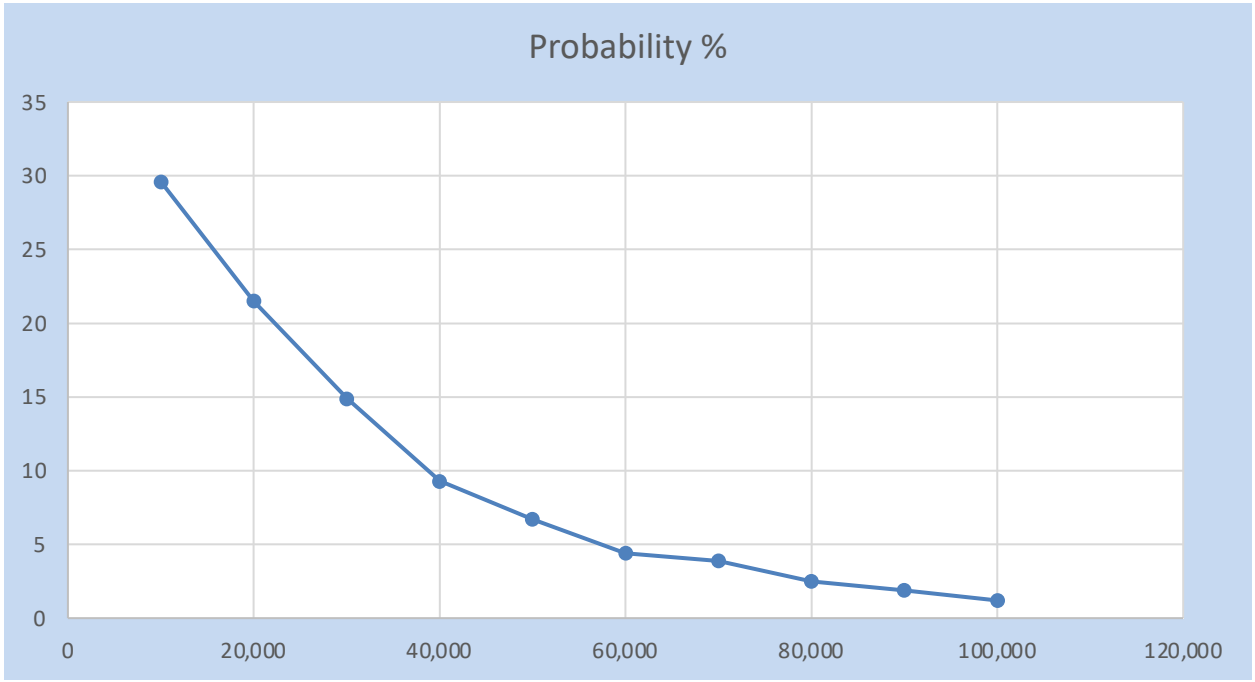
Where α = 0.4405, β = 4E-05, x the breach size Algarni, Malaiya 2015

- The values of the parameters may gradually change with time.
- Justification in the following slides.

Probability of a data breach by number of records lost

Over the next two years, involving minimum of 10,000 and maximum of 100,000 records.

Cost of a Data Breach Report 2019, IBM Security, study conducted by Ponemon Institute.



Exponential form

Colorado State University

Colorado State University

Quantitative Security

Colorado State University

Yashwant K Malaiya

CS 559

Costs of security breaches

CSU Cybersecurity Center
Computer Science Dept

Cost Metrics

Total Cost of a Breach =

Direct costs + Indirect costs – Recovered costs

Direct costs: funds spent directly

= Incident investigation cost

+ Customer Notification/crisis management cost

+ Regulatory and industry sanctions cost*

+ Class action lawsuit cost*

Indirect costs: lost business opportunity

= loss of goodwill, customer churn#

Recovered costs = Insurance recovery + tax break

* Post data breach response

Measured by the stock-market?

Cost Metrics

Total Cost of a Breach

= fixed costs + variable costs – recovered costs

$$\boldsymbol{Cost\ per\ Record} = \frac{Total\ cost\ of\ breach}{number\ of\ affected\ records}$$

- Fixed cost: regardless of the size of breach
- Variable costs depend on the number of records.
  - May not be linear because of economy of scale

Cost per record

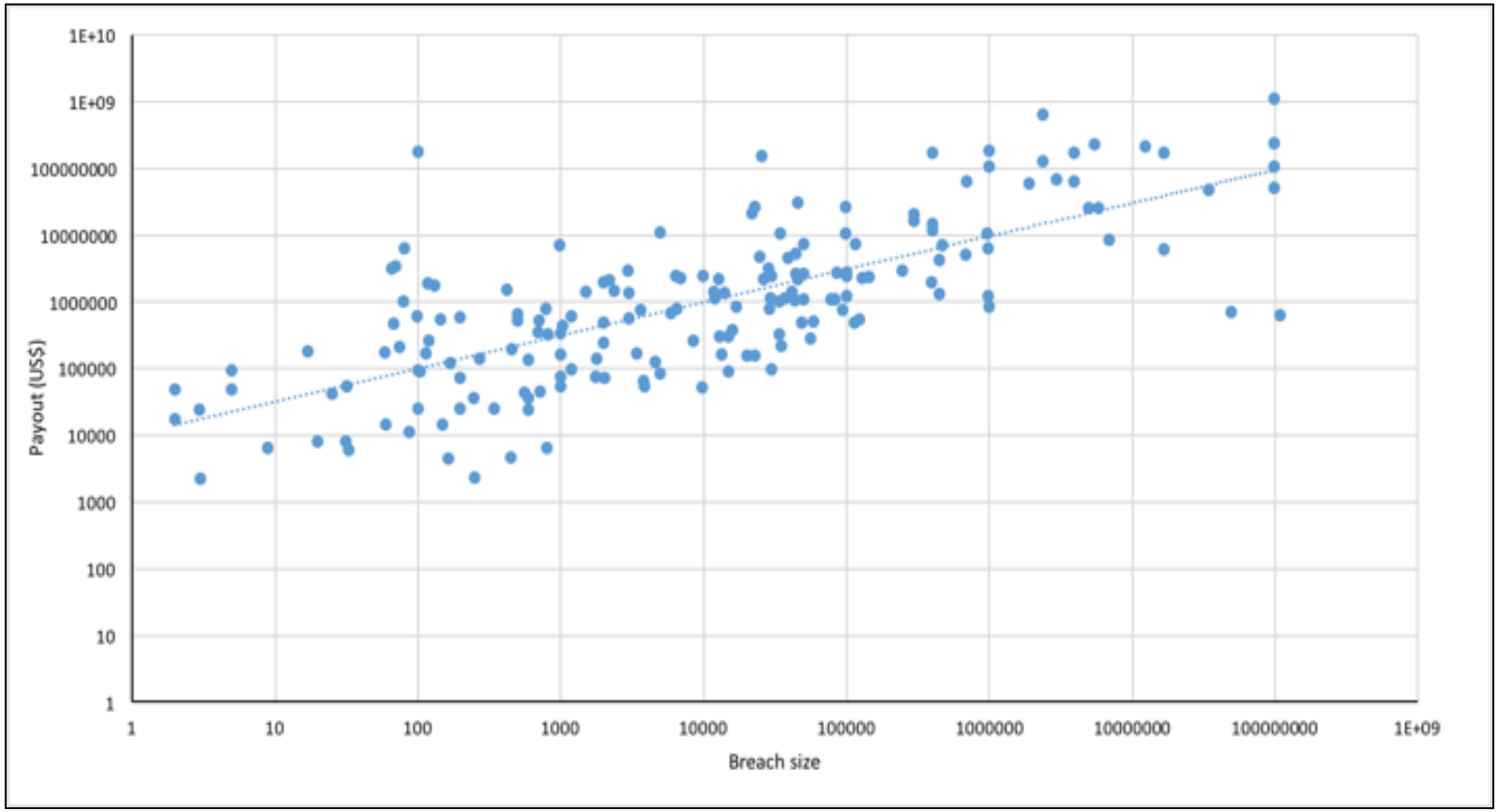
- Cost per record metric
- Partial costs
- Average costs?
- Available data
- Proposed model for Cost per record

Is there an average cost per record?

- Using averages make sense, at least for initial estimates
- The **law of large numbers**:
  - sample size grows, its mean gets closer to the average of the whole population.
- The **Flaw of Averages**:
  - $2 billion in property damage in North Dakota.
  - In 1997, the U.S. Weather Service forecast that North Dakota's rising Red River would crest at 49 feet.
  - Officials in Grand Forks made flood management plans based on this single figure.
  - The river crested above 50 feet, breaching the dikes, and unleashing a flood that forced 50,000 people from their homes.

The Flaw of Averages, Sam Savage, Harvard Business Review, Nov. 2002

The breach cost vs. breach size



Verizon 2015 data, the claim amount vs. breach size
(ranges from single digits to 108 million records)

Colorado State University

The breach cost vs. breach size

- Our proposed model

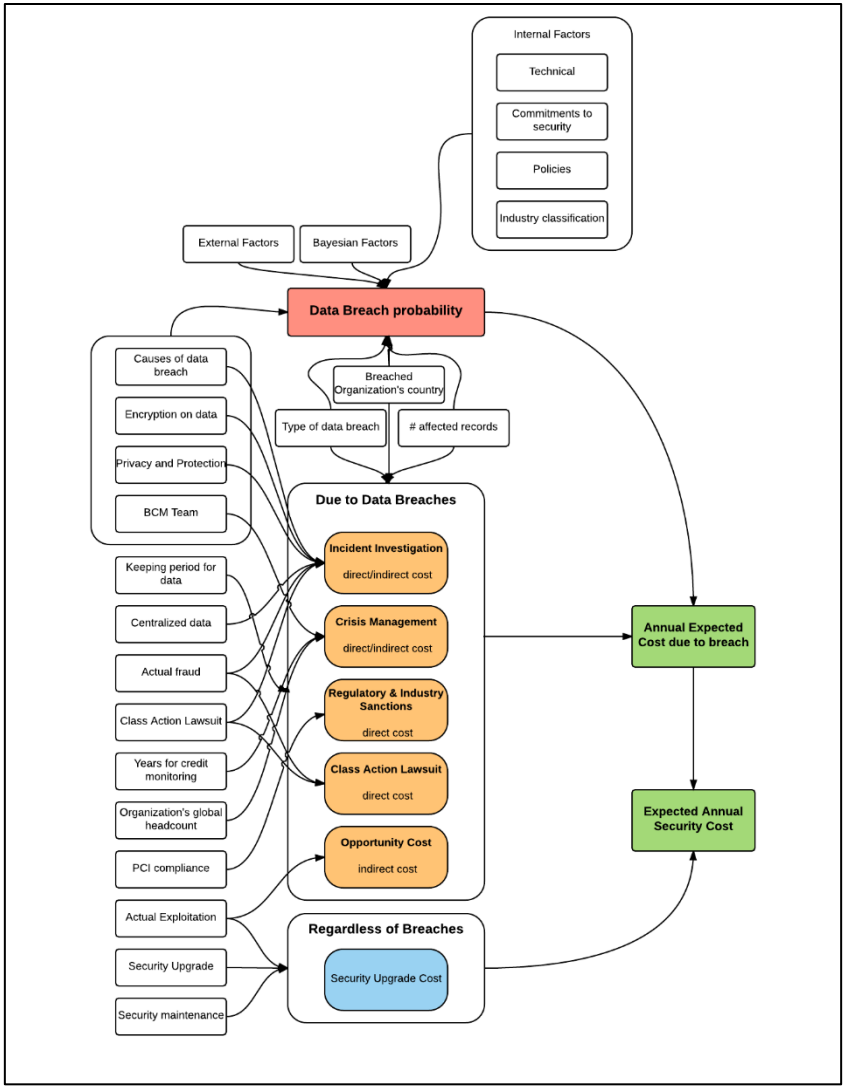
$$\boldsymbol{Total\ breach\ cost} = a\ *\ size\ \hat{}\ b$$

for breach sizes bigger than or equal to 1000 records

- Nonlinearity caused by ***economy of scale,*** *thus b should be < 1.*
- Thus

$$\boldsymbol{Cost\ per\ record} = a\ *\ (size)\ \hat{}\ (b-1)$$

Overall risk evaluation model



Colorado State University

Models for Partial costs

- Details in Abdullah Algarni's dissertation: Quantitative economics of security : software vulnerabilities and data breaches, CSU

- $\boldsymbol{Investigation\ cost\ per\ record}$
$$= [a * (size)^{b-1}\ for\ factors\ 4,5,6] * Fbreach\_cause * Fencryption * Fprivacy$$
- $\boldsymbol{Crisis\ Management\ Cost\ per\ Record}$
$$= [a * (size)\ ^\wedge\ (b-1)\ for\ factor\ 11] * FBCM$$
- $\boldsymbol{Sanctions\ cost\ per\ record}$
$$= a * (size)\ ^\wedge\ (b-1)\ \ for\ factor\ 14$$
- $\boldsymbol{Class\ Action\ Lawsuit\ Cost\ per\ record}$
$$= a * (size)\ ^\wedge\ (b-1)\ for\ factor\ 15\ and\ 16$$
- ***Opportunity cost: considered separately***

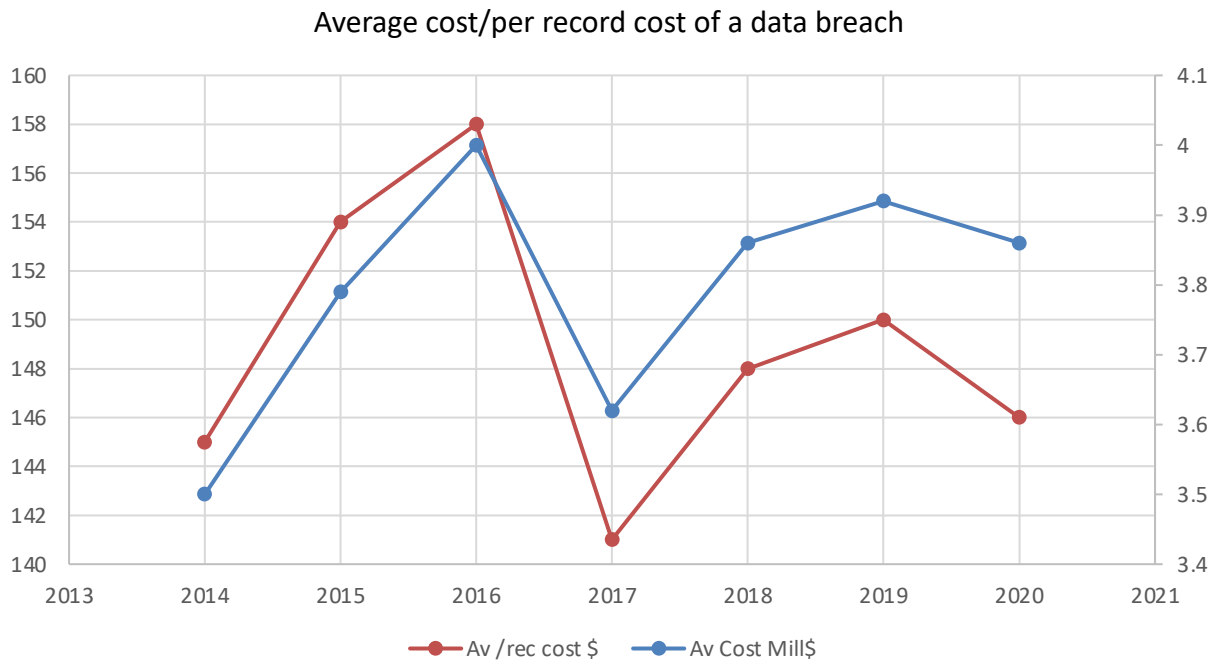
2020 Data

Ponemon Global Cost of Data Breach Study 2020

- 3,400-99,730 records
- Excludes mega-breaches, considered separately

Average total cost of a data breach

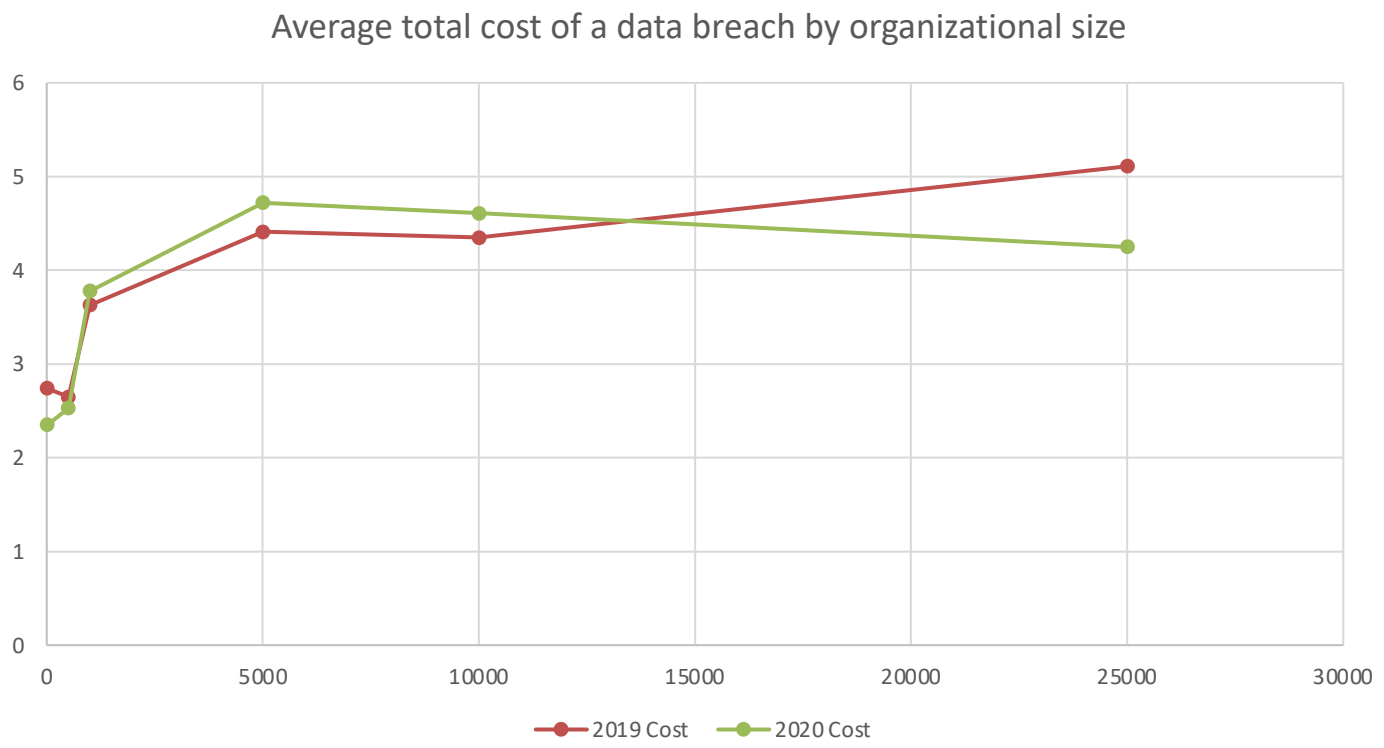
Per record cost: US$, Total cost measured in US$ millions
(Ponemon Global Cost of Data Breach Study 2020, 3,400-99,730 records. Excludes mega-breaches)



Colorado State University

Average total cost of a data breach by organizational size

- Note economy of scale



Colorado State University

Types of records compromised

(Ponemon Global Cost of Data Breach Study 2020, 3,400-99,730 records. Excludes mega-breaches)

Types of records compromised	Percent	Cost/rec	Cost/rec in malacious attack
Customer PII	80	150	175
Intellectual property	32	149	171
Anonymized customer data	24	147	163
Other corporate data	23	143	151
Employee PII	21	141	150

Partial Costs

		Cost in $million in category					
Category	**Percent**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Lost business	39.4	1.57	1.63	1.51	1.45	1.42	1.52
Ex-post response	28.8	1.07	1.1	0.93	1.02	1.07	0.99
Notification	6.2	0.17	0.18	0.19	0.16	0.21	0.24
Detection and escalation	25.6	0.98	1.09	0.99	1.23	1.22	1.11

Detection and escalation: Activities that enable a company to reasonably detect the breach of personal data either at risk (in storage) or in motion and to report the breach of protected information to appropriate personnel within a specified time period.

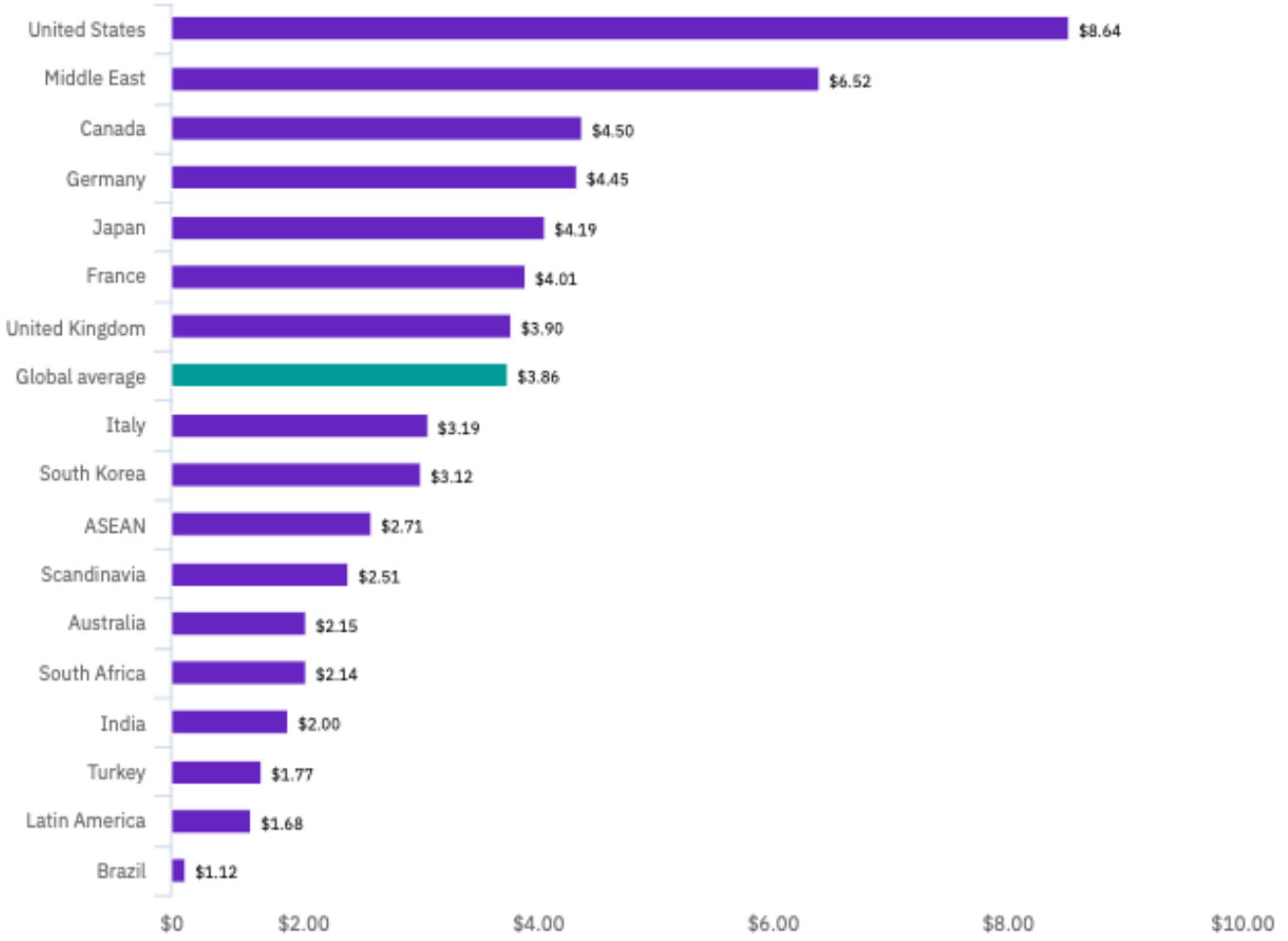
Notification: Activities that enable the company to notify individuals who had data compromised in the breach (data subjects) as regulatory activities and communications. Also included are costs that relate to communication with data protection regulators and other related parties.

Post data breach response: Processes set up to help individuals or customers affected by the breach to communicate with the company, as well as costs associated with redress activities and reparation with data subjects and regulators.

Lost business: Activities associated with cost of lost business including customer churn, business disruption, and system downtime. Also included in this category are the costs of acquiring new customers and costs related to revenue loss.

Ave total cost of a data breach by country / region

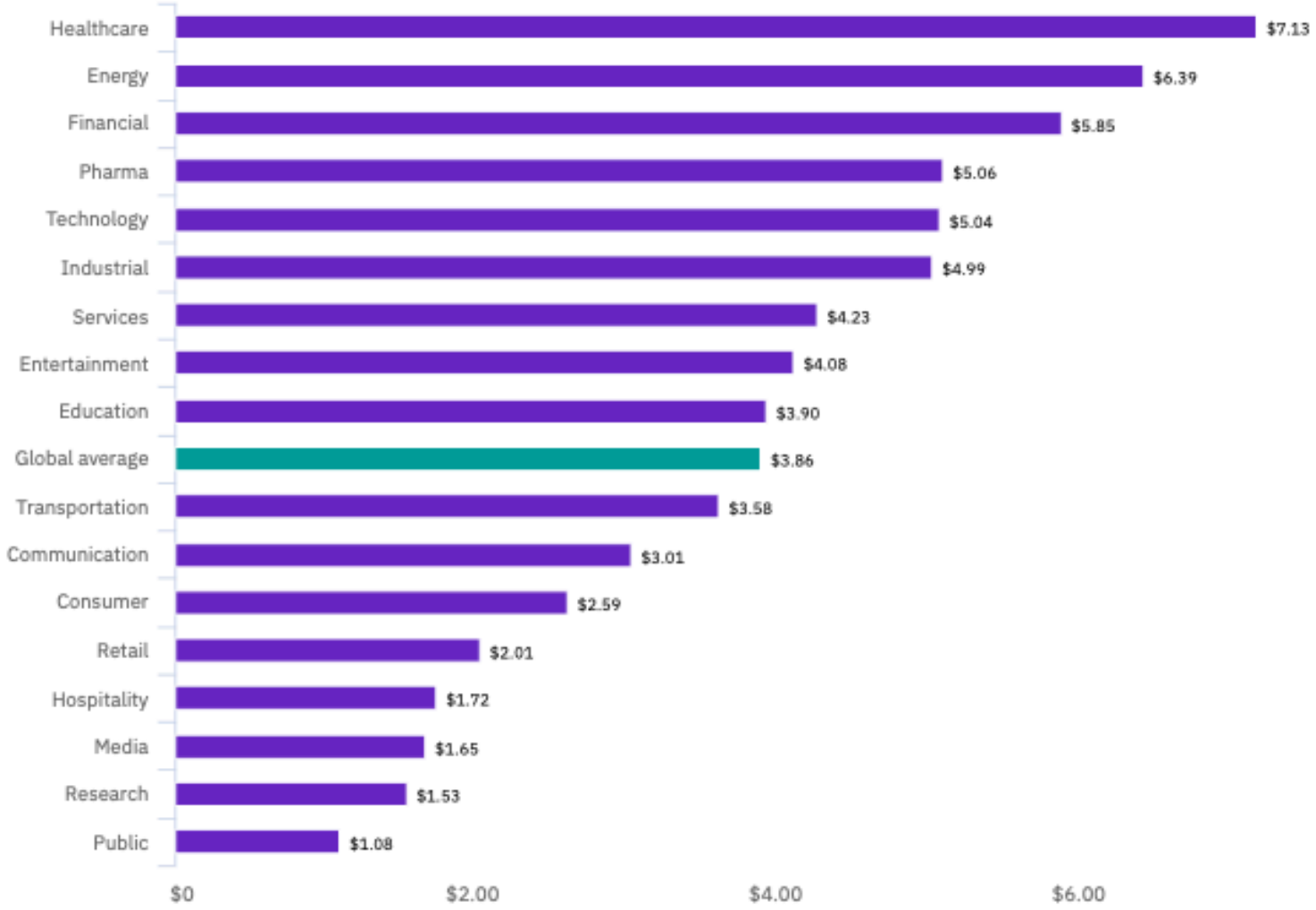
Measured in US$ millions (Ponemon Global Cost of Data Breach Study 2020)



Colorado State University

Average total cost of a data breach by industry

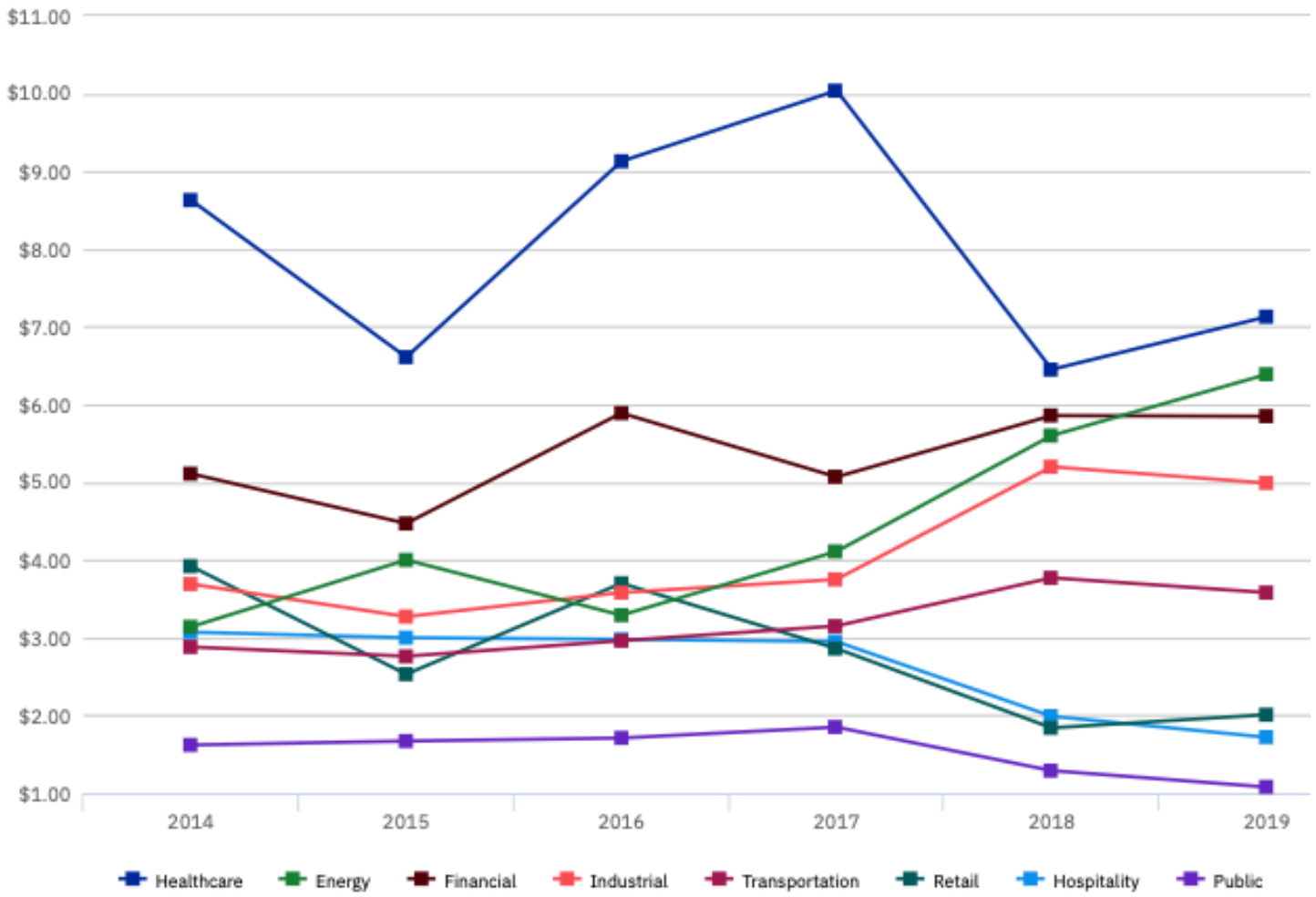
Measured in US$ millions (Ponemon Global Cost of Data Breach Study 2020)



Colorado State University

Trend in average total cost of a data breach in eight industries

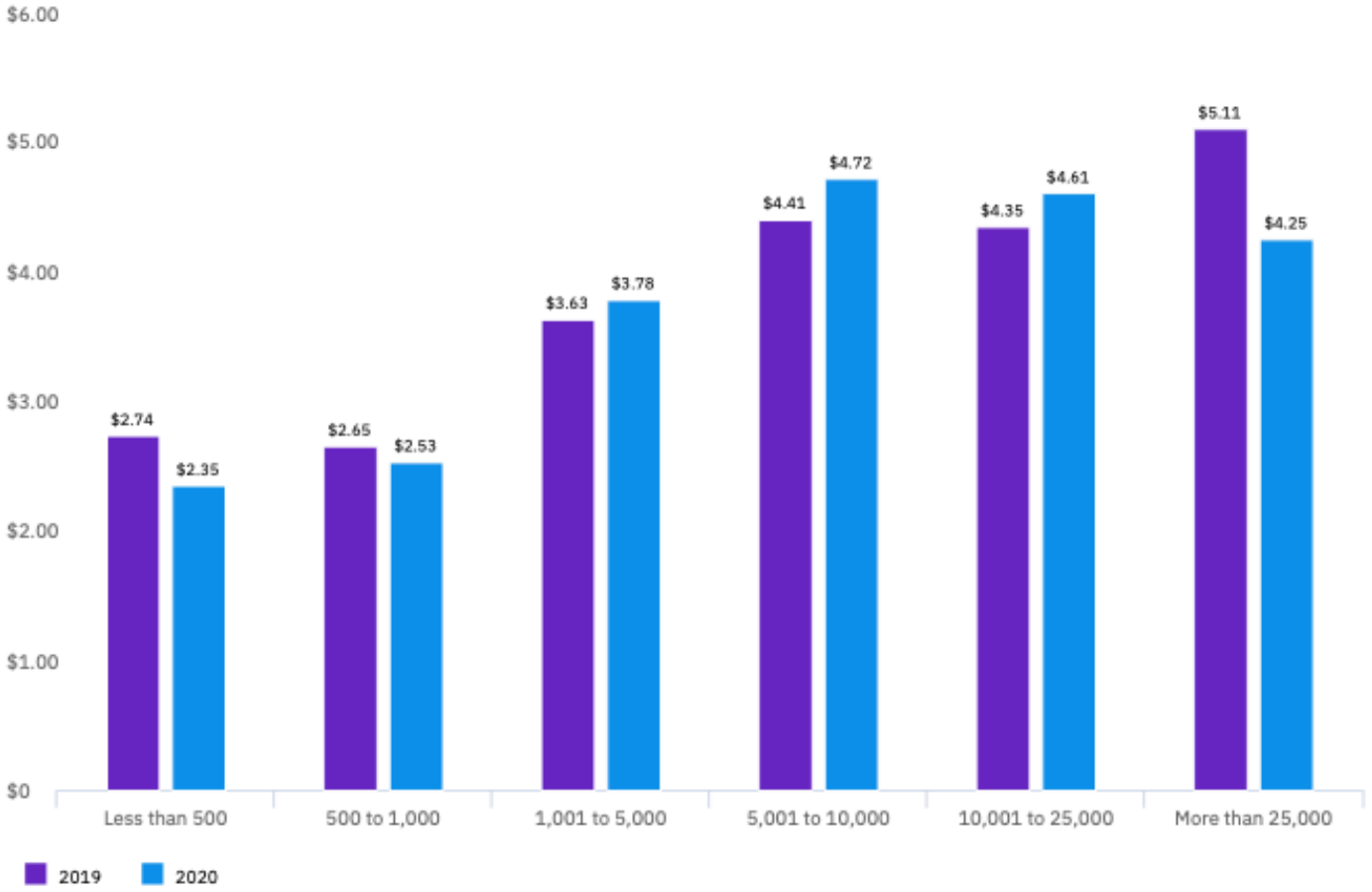
Measured in US$ millions (Ponemon Global Cost of Data Breach Study 2020)



Colorado State University

Average total cost of a data breach by organizational size

Measured in US$ millions (Ponemon Global Cost of Data Breach Study 2020)

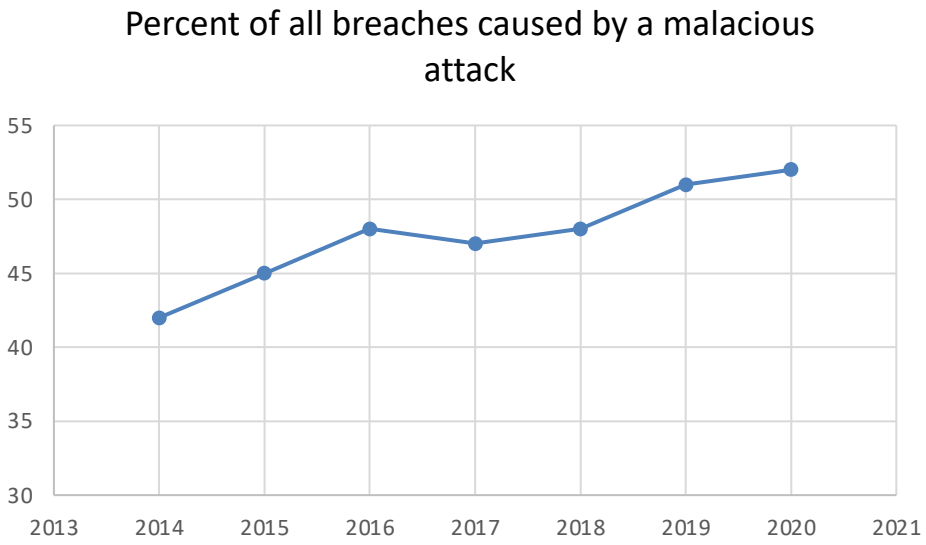


Colorado State University

Data breach root cause breakdown in three categories

Measured in US$ millions (Ponemon Global Cost of Data Breach Study 2020)

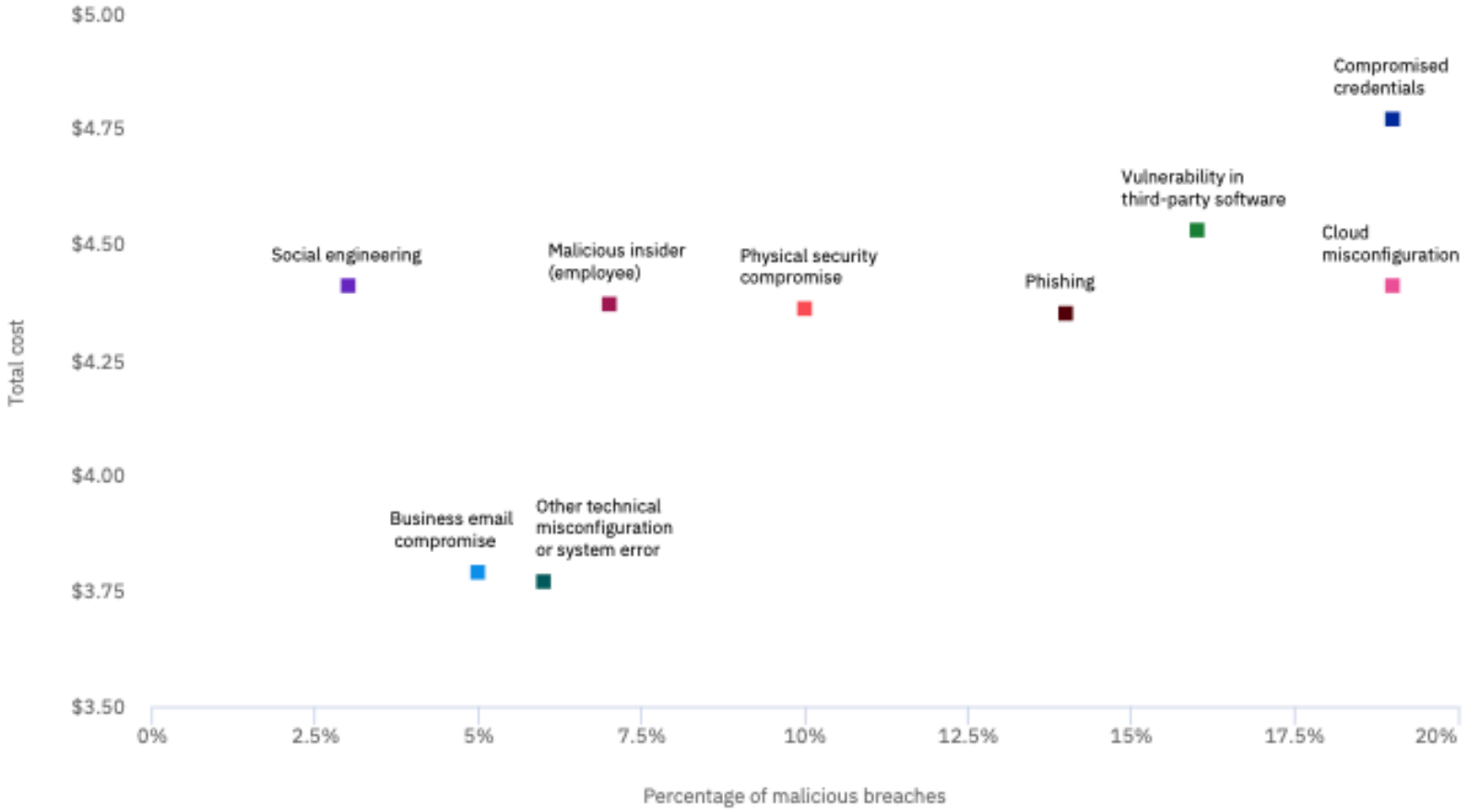
Root cause	Frequency	Av total cost $ mill
Malicious attack	52%	4.27
System glitch	25%	3.38
Human error	23%	3.33



Colorado State University

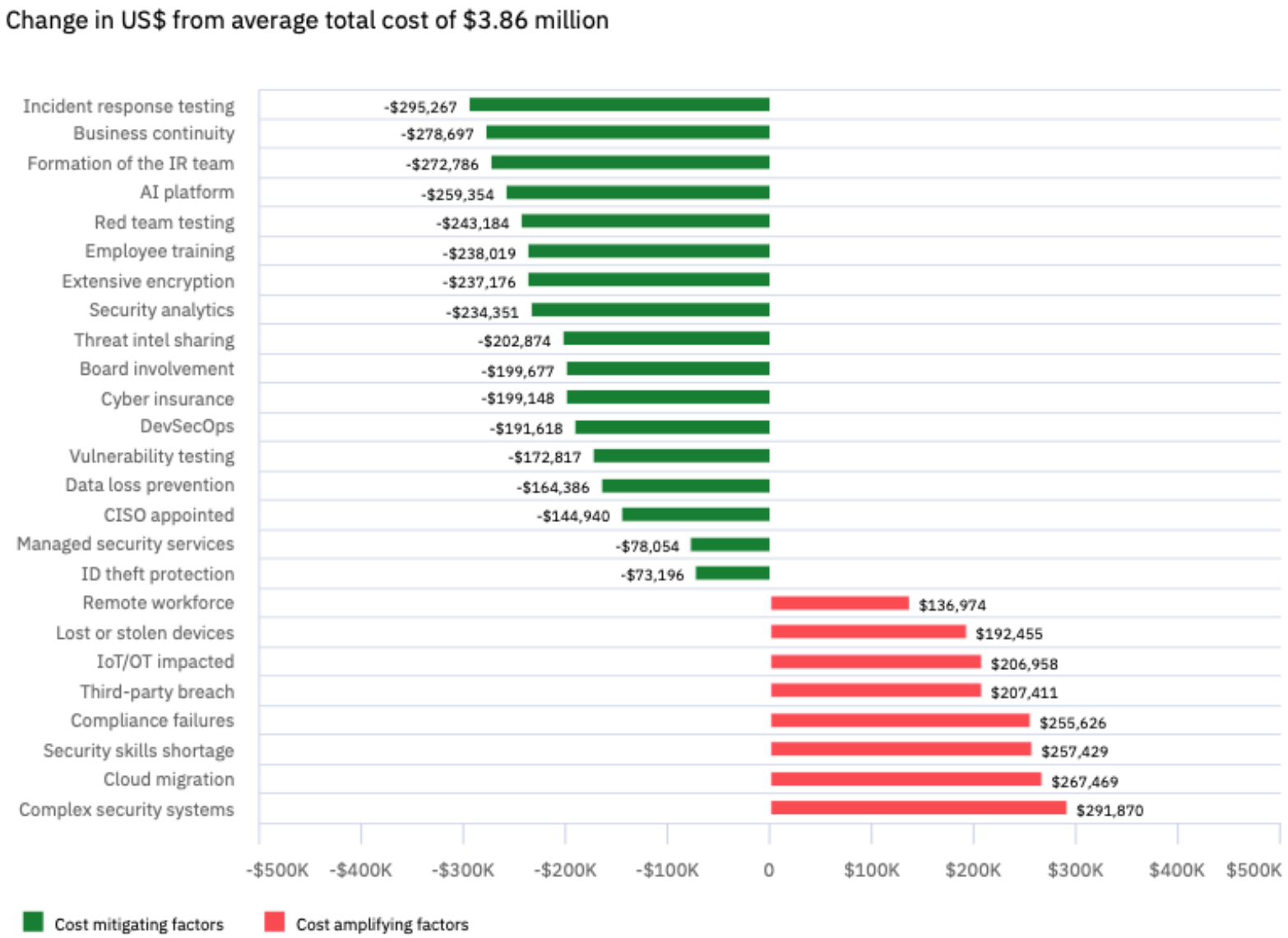
Av cost and freq of malicious data breaches by root cause vector

Measured in US$ millions (Ponemon Global Cost of Data Breach Study 2020)



Colorado State University

Impact of 25 key factors on the average total cost of a data breach 2020



Colorado State University

Security automation trends and effectiveness

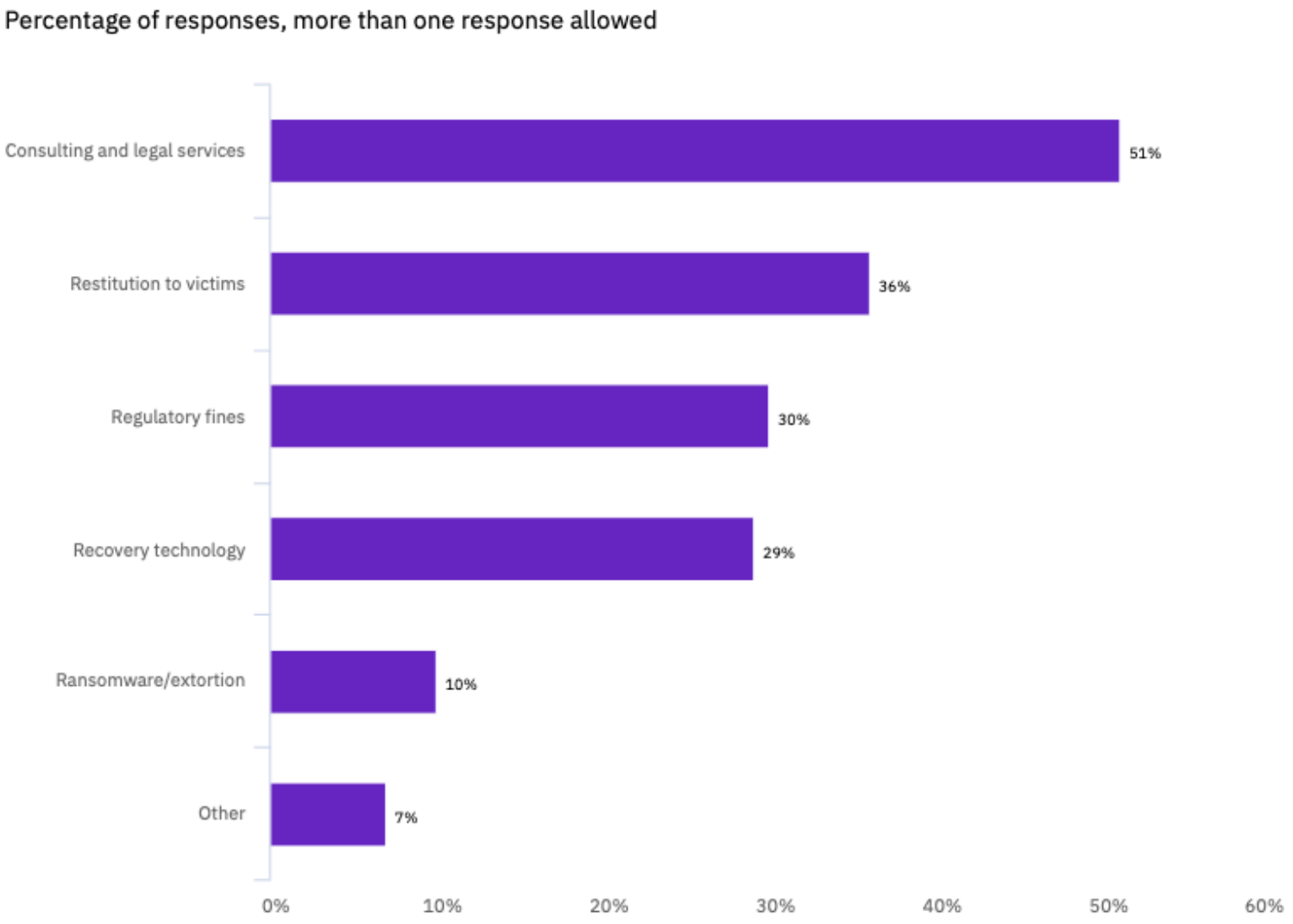
	Trends %			Cost of a breach (mill$)		
Automation level	**2018**	**2019**	**2020**	**2018**	**2019**	**2020**
Fully deployed	15	16	21	2.88	2.65	2.45
Partially deployed	34	36	38	3.39	3.86	4.11
Not deployed	51	48	41	4.43	5.16	6.03

- Security automation refers to enabling security technologies artificial intelligence, machine learning, analytics and automated orchestration.
- Note the significant reduction on the cost of breaches.

Average time to identify and contain a data breach

- Average days to identify in 2020: 207
- Average days to contain a breach in 2020: 73
  - Total 280 days
- Factors:
  - mild yearly trend
  - Country/region: Germany: 160 to Brazil:
  - Industry: Healthcare: 329 to Financial: 233
  - Root cause: malicious attack: 315, others 239-244
  - Security automation: Full: 234 to None: 308
- Impact 2020:
  - Breach lifecycle < 200 days: $3.21 m$
  - Breach lifecycle > 200 days: 4.33 m$

Types of costs recovered using cybersecurity insurance claims



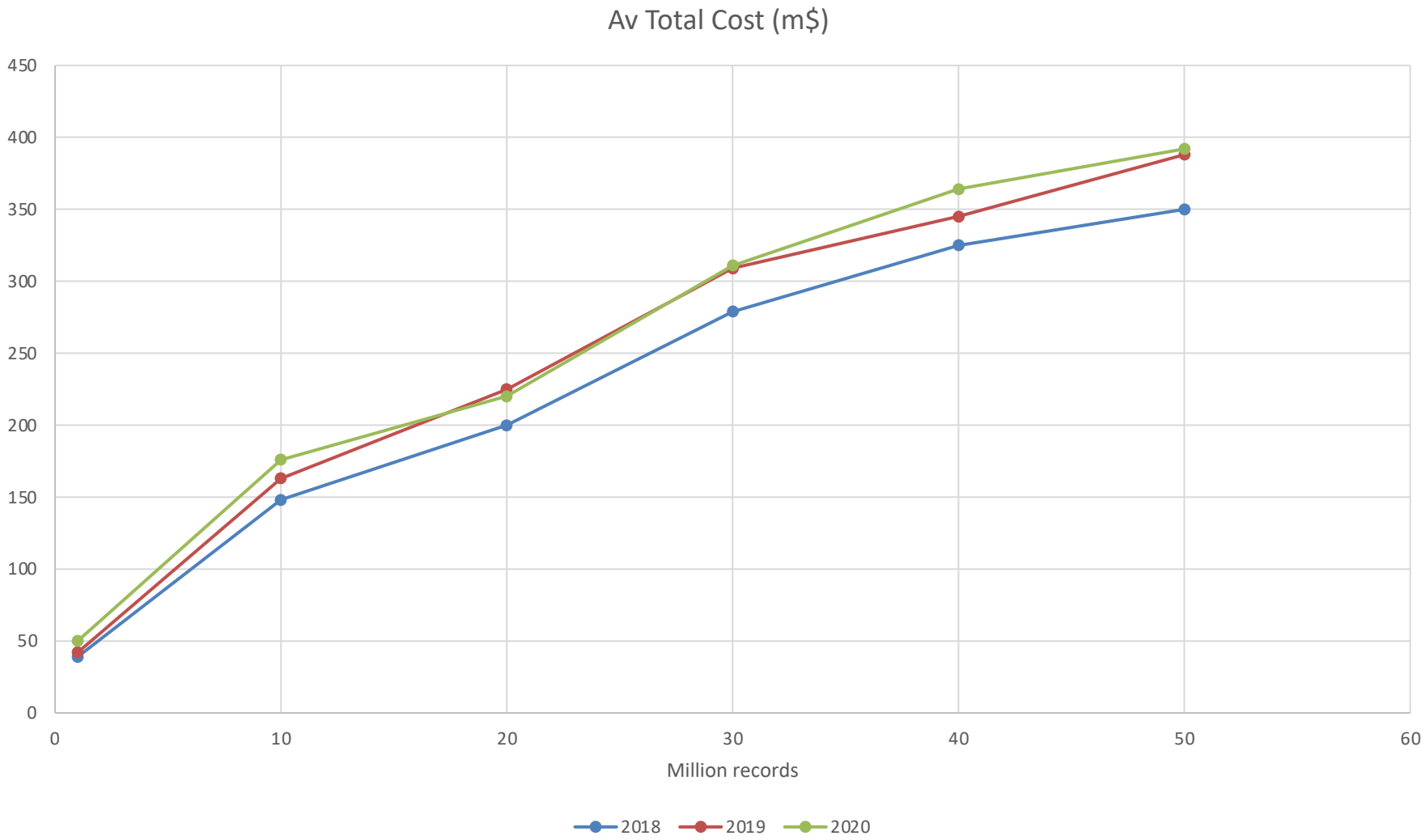
Colorado State University

Average cost of a malicious data breach by threat actor type

Measured in US$ millions (Ponemon Global Cost of Data Breach 2020)

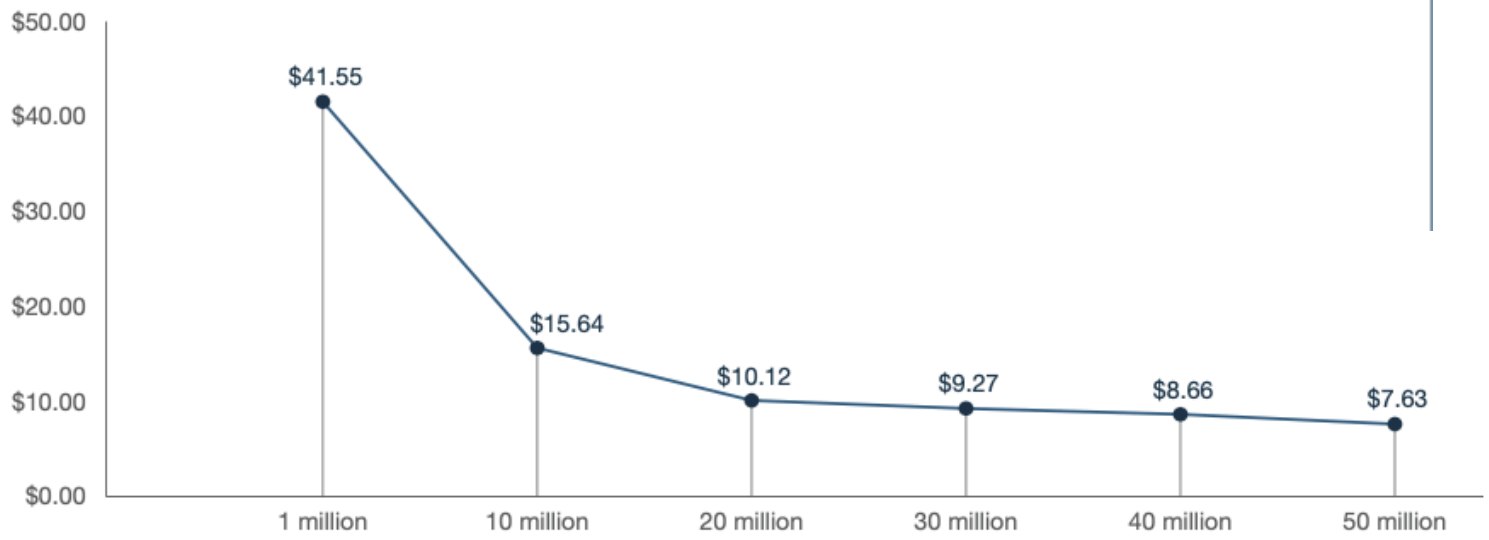
Actor type	Fraction %	Cost
Nation state	13	**$4.43**
Unknown	21	$4.29
Hacktivist	13	$4.28
Financially motivated	**53**	$4.23
Average malicious attack		$4.27

Mega breaches (>1 million records)



Colorado State University

Per capita cost of a mega breach



- At 50 million records, we estimate a per capita cost of $7.63. Per capita cost flattens out beyond 50 million records.
- From 2018 Cost of a Data Breach Study: Global Overview, IBM/Ponemon

National Costs of security Incidents

- This needs to be studies further in detail.

National Component	Illustrative cost
National Economy	**IP** $360B Gingrich 2016
National Security	Stuxnet-type attack $1Trillion
National Democracy	Clinton campaign: 1.2B, DNC breach

Quantitative Cyber-Security

Colorado State University
Yashwant K Malaiya

CS559

Breaches and Stock Price

CSU Cybersecurity Center
Computer Science Dept

Cost Metrics

Total Cost of a Breach =

Direct costs + Indirect costs – Recovered costs

Direct costs: funds spent directly

= Incident investigation cost

+ Customer Notification/crisis management cost

+ Regulatory and industry sanctions cost*

+ Class action lawsuit cost*

Indirect costs: lost business opportunity

= loss of goodwill, customer churn#

Recovered costs = Insurance recovery + tax break

Lost business opportunity

- Should reflect in the stock price
  - Stock price = f(earnings, growth)
- How do the data breach incidents influence market value?
- Subject of numerous studies

Examples of research results

Authors	Sample size	Period	Result
Garg et al. (2003)	22	1999–2002	Found that on average the loss to a company was $17–28 million as compared to some other reported estimates of between $50,000 to $2 million per incident.
Kannan et al. (2007)	72	1997–2003	No significant impact on the firms was detected on the analysis of both short- and long-term reactions.
Gatzlaff and McCullough (2010)	77	2004–2006	The overall effect of a data breach on shareholder is negative and statistically significant. The firms with higher market-to-book ratios experience greater negative abnormal returns.
Yayla and Hu (2011)	58	1994–2006	Pure e-commerce firms experienced higher negative market reactions than traditional bricks-and-mortar firms. Also found that DoS attacks had higher negative impact than other types of security breaches.

Spanos, Angelis 2015 Metastudy

- A number of papers conducting studies were found by the systematic search of bibliographic sources.
- Published in 2003-2015
  - 28 studies that have been performed to analyze the *Impact of Security Breaches to Breached Firms*, 25 (89.3%) presented a negative impact while in 20 studies (71.4%) this negative impact was found significant.
  - Two studies that analyzed the *Impact of IT Security Investments to Firms that Invest*, the results were positive and statistically significant.

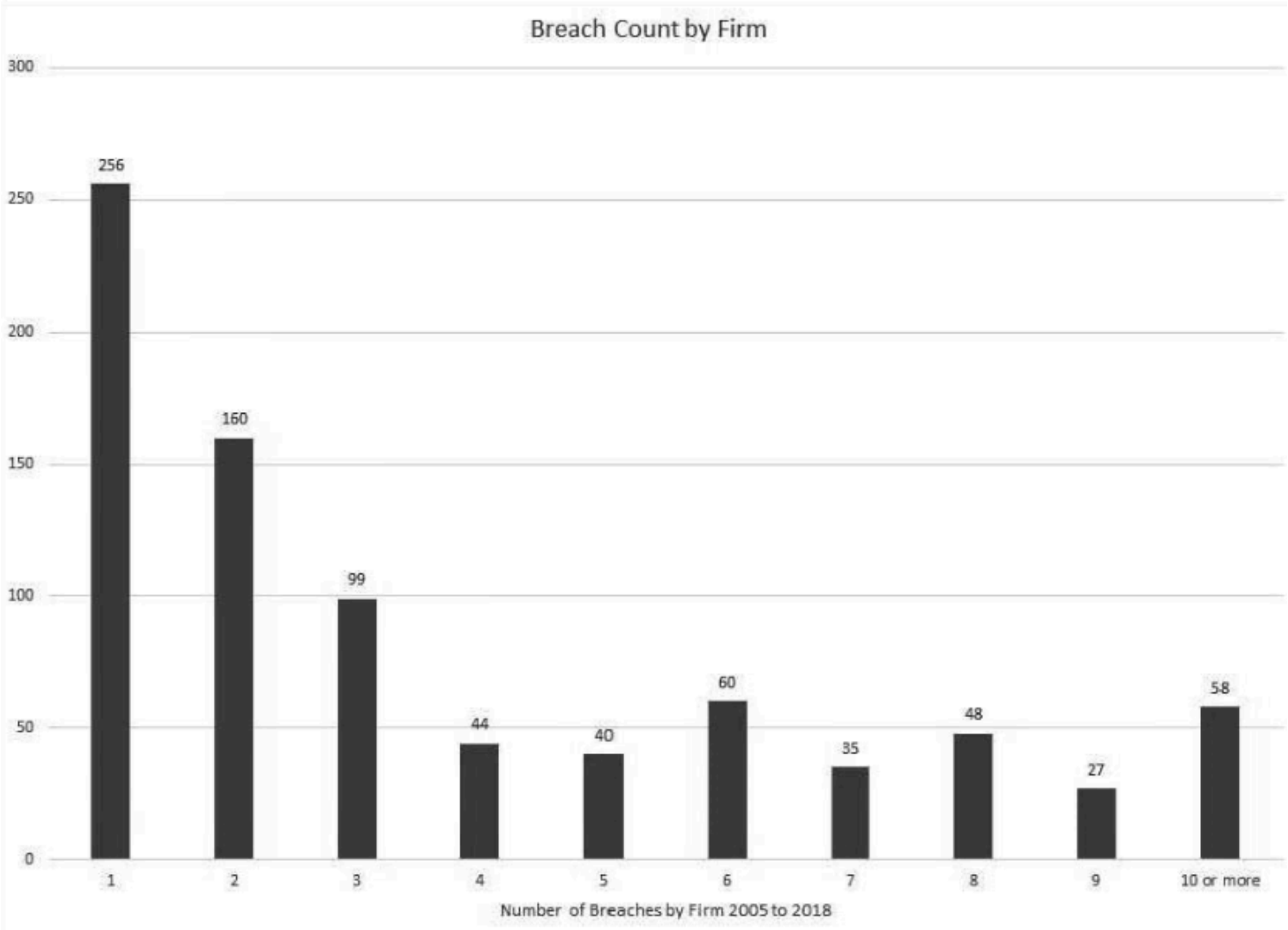
The impact of information security events to the stock market: A systematic literature review,
2015

The (Lack of) Economic Impact of Data Privacy Breaches

- Richardson, Smith, Watson, 2019
  - H1: On average, data breaches have a negative impact on short-term and long-term stock market returns.
  - H2: After a data breach, future company performance is negatively affected.
  - H3: On average, there will be an increase in audit fees and other fees around a data breach.
- 1,165 + 458 data breaches between 2004-2018
- Results: H1, H2, H3: all not supported

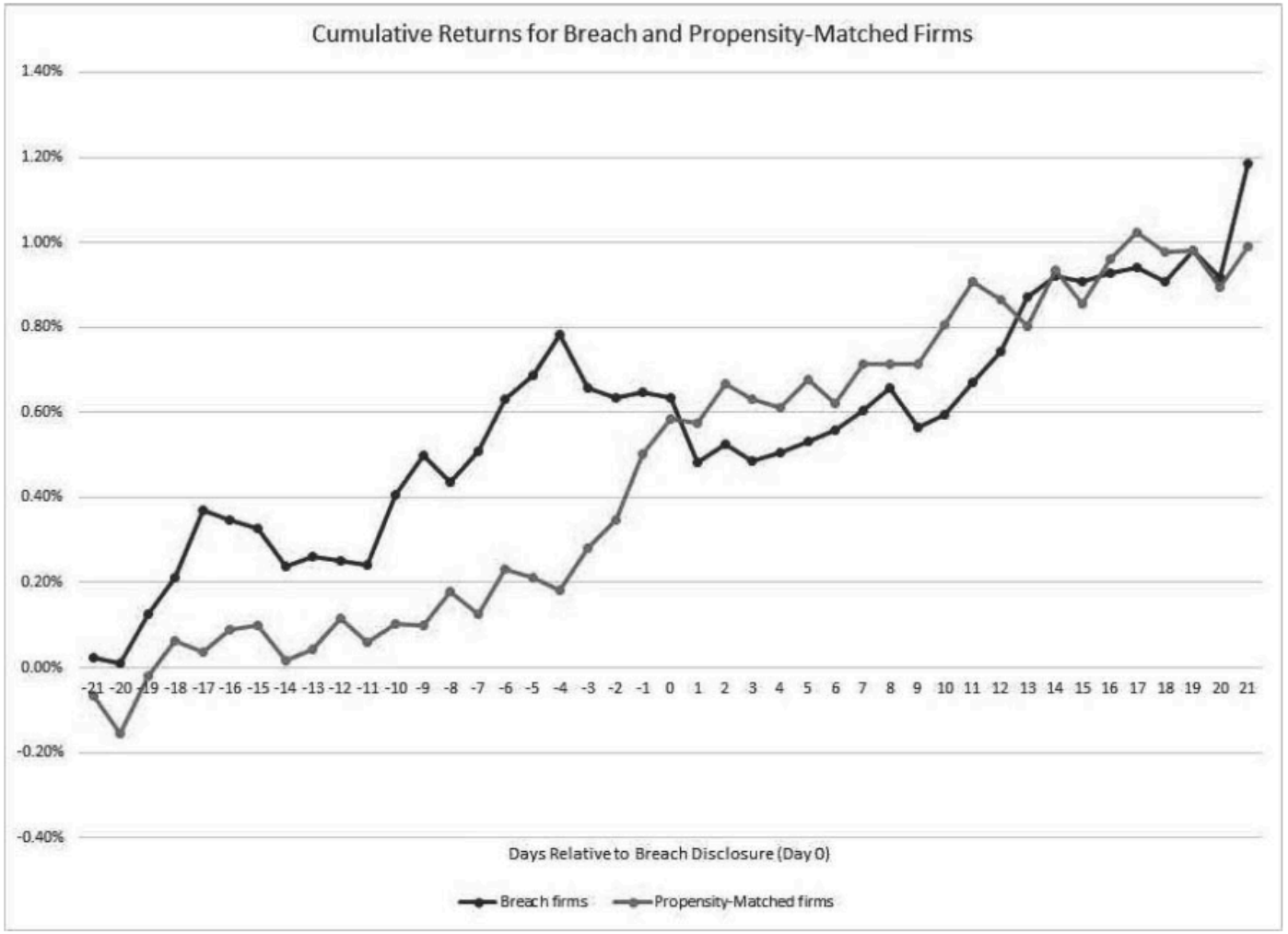
Much Ado about Nothing: The (Lack of) Economic Impact of Data Privacy Breaches, 2019

Breaches by Firm (2005-18)



Colorado State University

Cumulative Returns around Breach Disclosures



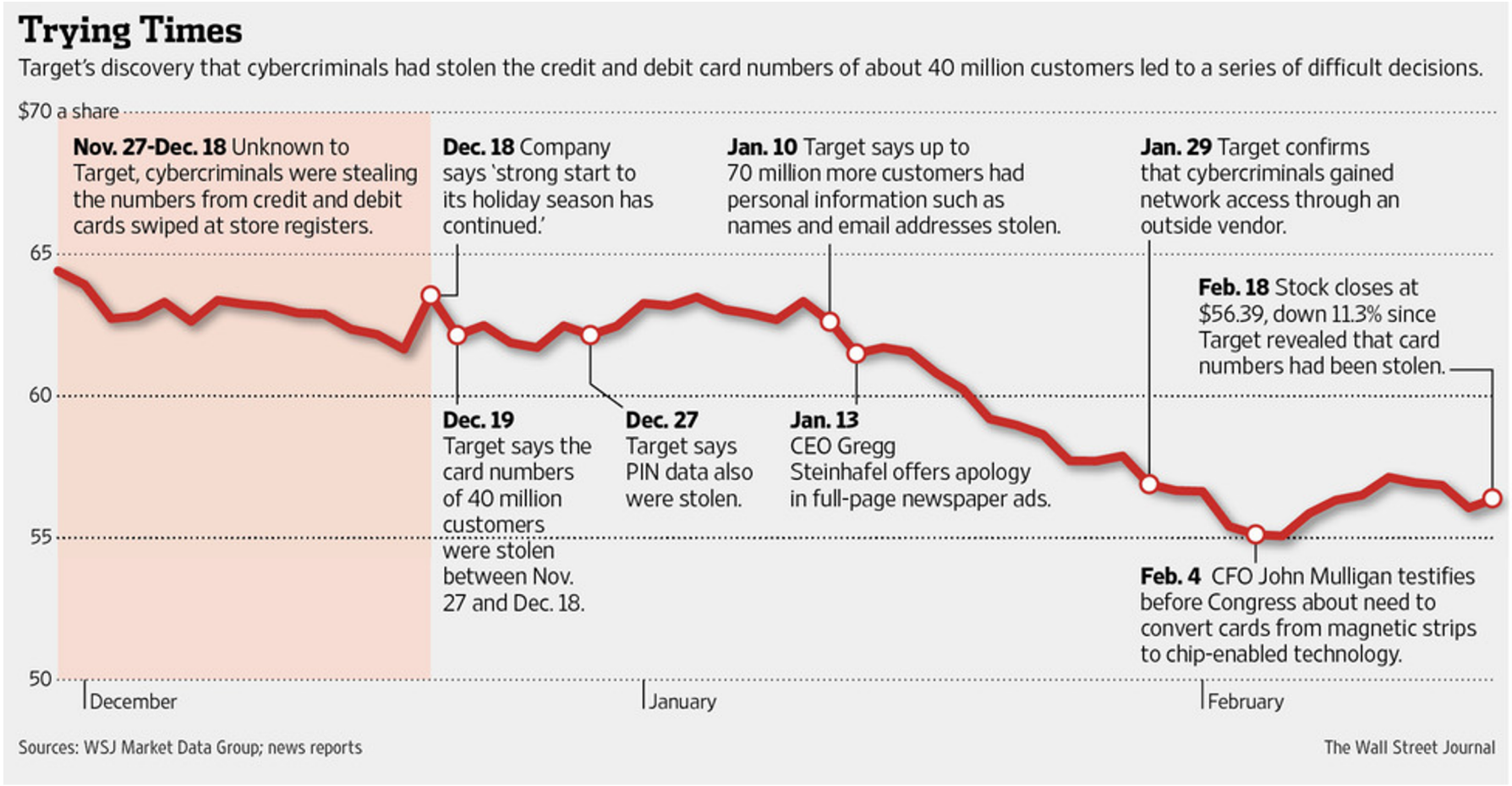
Cumulative daily returns for Breach and Propensity-Matched firms over the period starting 21 days before and ending 21 days after breach disclosure.

Colorado State University

Chang, Gao, Lee 2020 **Hypotheses**

- **Hypothesize 1 (H1).** *The announcement of a data breach has a negative effect on the short-term market value of the breached company.*
- **Hypothesize 2 (H2).** *The announcement of data breach has a negative effect on the long-term market value of the breached company.*
- **Hypothesize 3.1 (H3.1).** *The size of the data breach is positively associated with a higher negative return on the short-term market value of the breached company.*
- **Hypothesize 3.2 (H3.2).** *The size of the data breach is positively associated with a higher negative return on the long-term market value of the breached company.*

Target breach



Colorado State University

Chang, Gao, Lee 2020

- North American companies listed on the stock market from 1 January 2003 to 31 December 2014. 147 original samples.
- Buy-and-hold abnormal returns (BHARs) model
  - comparing companies' strategic buy-in stocks over a period of time to the benchmark portfolio (market index)
- Hypothesis 1: significant negative abnormal returns within the event window of [0, 0] and [0, 1].
  - When the event company declared that it suffered a confidential information breach, the event led to an average stock price fell of −0.23%.
  - The abnormal return for trading day after the event was −0.41% .

Chang, Gao, Lee 2020

- Hypothesis 2: significant negative impact on the long-term market value of the firm.
  - The average abnormal returns for the
    - 12 months to buy-and-hold after an event is –8.88%.
    - The BHAR is −32.69% and −32.16% for 24 months and 36 months after the event. Supported.
- Hypothesis 3.1: the size of the data breach is positively associated with a higher negative return on the short-term market value of the breached company.
  - The explanatory power of the model on the short-term market value is more than 7%. Supported.

Chang, Gao, Lee 2020

- Hypothesis 3.2: the size of the data breach is positively associated with a higher negative return on the long-term market value of the breached company.
  - negative relationship between the buy-and-hold abnormal return performance and event size in event period [0, +11], event period [0, +23], and event period [0, +35] exist.

Observations

- Numerous studies using actual data
- There is some impact of security breaches, generally small relative to the annual earnings
- Issues:
  - Market swings caused by many factors, along with noise
  - Abnormal return: assumes comparison with normal return, using an index (such as Nasdaq Composite, over 2500 stocks, dominated by Apple, Amazon, Microsoft)
  - Data-based analyses, mechanics of movement are not considered