COST-EFFECTIVE FIRM INVESTMENTS IN CUSTOMER INFORMATION PRIVACY

Robert J. Kauffman  
W. P. Carey Chair in IS  
W. P. Carey School of Business  
Arizona State University  
rkauffman@asu.edu

Yong J. Lee  
Ph.D. Program in IS  
W. P. Carey School of Business  
Arizona State University  
leeyi@asu.edu

Ryan Sougstad  
Ph.D. Program in Info. and Dec. Sci.  
Carlson School of Management  
University of Minnesota  
r sougstad@csom.umn.edu

Abstract. Extensive personal information is gathered explicitly or implicitly when a customer interacts with a firm. Significant risks are associated with handling such personal information. Providing protection may reduce risk of misuse or loss of private information, but it imposes some costs on the firm and its customers. Risk is associated with improper handling of sensitive customer information. Profits from e-commerce that are earned when there is improper use of private customer information are subject to lawsuits, restitution and other undesirable outcomes. So a firm will want to ensure appropriate privacy protections are in place to safeguard customer information. We present a profit optimization model for customer privacy protection investments considering the potential value implications that arise. We employ a profit-at-risk approach based on value-at-risk methods from financial economics.

Keywords: Financial economics, information privacy, privacy protection, profit-at-risk, value-at-risk.

1. INTRODUCTION

The ubiquity of the Internet, the expansion of data storage space, and revolutionary increases in computing power have enabled firms to collect and analyze massive amounts of wide-ranging customer information at very low costs. Extensive personal information can be gathered either explicitly or implicitly when a customer interacts with a firm. Such information is gathered together to profile users and provide targeted services, such as tailored advertisements, discount offers and so on. However, significant risks are associated with handling such personal information of customers. Firms are responsible for proper treatment of such personal information. When it is misused, it may cause significant impact on the firm’s reputation, and may result in large financial losses to the firm. For example, a recent incident involving privacy breaches with a major U.S. retail business that lost credit card information for more than 94 million customers’ credit card information resulted in US$62 million in costs and losses by the card issuer as a result of the breach. To be successful in e-commerce, many firms now implement various methods to protect their customer’s private information in order not to be the victims of such breaches. Privacy concerns now have become a major obstacle to attracting customers to e-commerce [6]. Firms collecting personal information from consumers should consider the costs and risks of privacy protection and strive to offer both effective technology solutions and the security policies associated with them.

Firms face the risks of improper access to, errors with, and theft of their customers’ private information [17]. As a firm invests more for privacy protection, such risks likely will be reduced. However, due to the unforeseen nature of such risks, they may not be controlled completely, even if a firm implements all possible protections. Therefore, it is important for a firm to understand the factors affecting the value of the firm’s decision to invest in privacy protection. Public policy and regulatory studies have overlooked the need to understand the protection of customer privacy in economic terms. Although the right to privacy should be protected, economic theory predicts that it may not be optimal for a firm to always protect it fully.

Providing protection methods for customers may reduce the risk of misuse or the loss of private information, but at the same time, it imposes some costs on the firm as well as on customers. Some of protection methods involve direct costs to customers. For example, the use of hardware solutions may require users to purchase technology to access services securely. An example is PayPal, which sells a small keychain device that provides a login code that enables the validation of a user’s login.

Other solutions may not involve direct monetary costs to customers, but some indirect costs may occur. The required time and installation effort, the learning costs for use, and dealing with malfunctions of the software solutions are examples of such indirect costs. Customers’ perceptions about the value of privacy protection may vary. The customer may choose to use the protection services, depending on how much they value privacy, and how much the services cost.

Firms, on the other hand, incur implementation costs to provide protection services. Although the result of mishandling of private information can be costly, most firms will choose to implement only some of the available customer privacy protections. In most cases, implementing all of the available protections is prohibitively expensive. Firms must balance the cost and risk associated with privacy
breeches against the investment required to find a reasonable level of privacy protection.

We will answer the following questions. What are the cost-effective investment options for firms to protect privacy? What are the factors that drive firms to invest in privacy protection? How can we identify which investment level choices will yield optimal profits? We develop a profit optimization model for privacy protection service implementation that considers the risks associated with implementation. We use a profit-at-risk approach based on value-at-risk methods from financial economics. We show that the optimal investment choice is based on the expected reduction of risks due to implementing protection and the amount of costs the firm is able to pass on to its customers with a price premium. We also provide model-based evidence for why firms may not choose to implement full protection for their customers.

2. THEORY

2.1. Value-at-Risk

We address the issue of cost effective investment choices value-at-risk (VaR) methods. VaR measures the estimation of the maximum loss on portfolio with a given level of confidence. VaR methods have been widely used in financial services to evaluate asset portfolios in the presence of risk factors. Kauffman and Sougstad [8] applied value-at-risk methods in the contexts of information technology (IT) services contract design, and introduced the concept of profit-at-risk, the maximum expected profit that can be expected in the presence of risk. We apply a similar construct for estimating maximum expected profit that a firm earns from customer privacy protection, with a risk of the misuse of private information.

Value-at-risk is typically calculated for a single time period to gauge the likelihood of loss at the 95% confidence level [2]. The confidence level of 95% indicates that, on average, there is a 95% chance of the loss of the expected loss on an asset being lower than the value-at-risk value that is calculated. The method provides comprehensive risk measurement to analyze financial risks – it can deal with risks from any source. Risk in a value-at-risk model is represented by the volatility of the underlying costs or revenues, and the resulting expected value outcome. It is expressed using the standard deviation. This permits the maximum possible value of the portfolio in the presence of risk to be estimated.

This method provides a reasonably accurate estimate of risk at a reasonable cost [7]. It is widely used in financial risk analysis as it provides a quantitative measure of risk that can be directly compared with expected profit. The risk measured by the VaR method is defined as the volatility of unexpected outcomes associated with the value of assets or liabilities of interest [2]. For privacy protection, risk is associated with the improper handling of customer privacy information by the firm.

2.2. Privacy Protection

We distinguish between privacy protection and security. Privacy issues are applicable whenever data is exchanged or processed which is explicitly related to individuals. Security issues deal with loss, theft, and unauthorized use of the data in information systems [14]. Privacy deals with data that directly relate to people, who may lose control of their identities, their financial situation, or other aspects of their lives. Security issues also pertain to sensitive data that do not constitute sensitive personal information. They emphasize protection methods and technology. So there is an overlap between the two. Protecting a firm's operational data is a security issue; implementing authentication methods to protect customer's bank account may be a privacy issue as well as security issue.

For minimizing costs and achieving high profitability, protecting privacy is just another element of overhead from a firm's perspective – until a loss of private customer information occurs. Some secondary use by the firm of its customers’ private information could yield increased revenues, by attracting more customers using profiling. The more information a firm has about its customer, the better will be the service it can provide, and so its customers will be better off too [12]. Many customers are concerned when their personal information is analyzed in this way by firms [17]. It is difficult to protect their information because the process of collecting, storing and analyzing personal information occurs across the spectrum of a firm's operations. Also, it is hard to predict the ways that a customer’s private data may be misused. This is because the scope of customer privacy and private information is rather loosely defined, and there will be significant differences in how different individuals value it [15, 17].

As concerns about the improper use of personal data have increased, governments around the world have suggested adherence to fair information principles [1]. It offers guidelines for firms’ self-regulation practices, as well as for governmental legislation. As one of the most widely used guidelines for self-regulation, the principles suggest that a firm or a web site should provide notice to its customers or visitors whenever their personal information is collected or used. They also suggest that a firm or a web site should provide an opt-out choice to customers, whether their private information is used for analysis or passed on to a third party. The fair information
principles also provide guidelines for a firm to control access to the private information of its customers [13]. Ashrafi and Kuilboer [1] examined the way privacy is treated in the context of fair information principles. In a survey of the top 500 interactive companies in the U.S., the authors found that retailers and travel agencies have been complying fairly well with the principles. They also found that larger companies and retail stores tend to be at the top of the list among firms that have adopted these guidelines. However, they also found that few firms spend enough resources on privacy protection to comply.

3. MODEL DEVELOPMENT

3.1. Background

We consider a firm that provides several data privacy protection methods to its customers. They include software, hardware implementations and other non-technical privacy policies. Examples of software protection methods include intrusion detection add-ons, encryption software, secure socket layer connections, public key infrastructure, antivirus software and so on. Random number generators, smart cards and fingerprint readers are examples of hardware implementations for privacy protection. Besides such software and hardware solutions, firms also offer other means of privacy protection to their customers. Examples of non-technical means include providing opt-out choices while collecting personal information, conducting transactions through known third-party payment systems such as PayPal or Google Checkout, using one-time auto-generated credit card numbers that is not linked to personal information.

The firm will not choose to use all of these methods due to customer preferences and incompatible environments, and concerns about profit maximization. Previous research suggests that although some customers are sensitive about privacy, others are less concerned [4]. Different computing environments also prevent customers from adopting all of the possible methods of protection. Some methods require high-powered computing with advanced technology, some are dependent on the operating systems that customers have implemented, and some require knowledge of computing.

3.2. Model Specification

Model parameters are defined in Table 1.

Modeling Assumptions and Definitions. We assume that the firm offers a chosen mix of privacy protection services from a known pool of services. Customers prefer more protection to protect their privacy. Likewise, the firm selects a protection level to maximize its profit subject to a pre-set level of risk to maintain some overall level of profitability.

<table>
<thead>
<tr>
<th>Table 1. Modeling Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NOTATION</strong></td>
</tr>
<tr>
<td>$S$</td>
</tr>
<tr>
<td>$H$</td>
</tr>
<tr>
<td>$R(S)$</td>
</tr>
<tr>
<td>$C(S)$</td>
</tr>
<tr>
<td>$PaR$</td>
</tr>
<tr>
<td>$h(s)$</td>
</tr>
<tr>
<td>$\alpha$</td>
</tr>
<tr>
<td>$T$</td>
</tr>
<tr>
<td>$K$</td>
</tr>
<tr>
<td>$g(x)$</td>
</tr>
<tr>
<td>$f(n)$</td>
</tr>
<tr>
<td>$R(n,x)$</td>
</tr>
<tr>
<td>$Q(S, \alpha)$</td>
</tr>
</tbody>
</table>

Note: We use the notation $PaR$ when we refer to modeling values in our development and analysis of this problem, and reserve the term *profit-at-risk* for the higher-level concept and methodology.

In our model, we will assume that there are a number of protection methods available, and, the protection level $S$ that a firm provides can be thought of as being continuous from a low percentage to a high percentage of protection, or $0 \leq S \leq 1$. The firm chooses a protection level $S = 1$ when it provides data privacy protection as much as is possible based on current methods of computer security auditing. A protection level $S$ is 0 implies no spending to protect customer data privacy. The customer should not expect much protection, but nevertheless there will be only a likelihood of a loss, not the guarantee of it. In reality, government regulation may require some minimal level of protection level. If that is the case, the government restriction would represent $S = k$, where $k$ is a lower bound for the security level.

A useful approach was demonstrated by Kauffman and Sougstad [8]. They employ a *profit-at-risk constraint* in their analysis of IT service contracts. Their methodology considers the lowest acceptable level of profits with a given confidence level over a given time period. We view the managerial problem in terms of a portfolio of financial assets, which makes it amenable to analysis with value-at-risk methods. We consider the set of protection services as a portfolio that a firm can build. Service level $S$, this way, represents the proportion of protection services chosen among the possible protection services.

We assume total costs consist of two distinct categories: implementation costs, and estimated losses at a given confidence associated with privacy events. We use the words *estimated losses* for financial losses directly associated with the firm’s legal obligations and remediation measures on privacy handlings,
as well as any estimated indirect costs such as loss in reputation and trust. Implementation costs are costs related to develop and deploy the services to provide protection services to customers, and are not subject to uncertainty. Instead, implementation costs tend to be known by the firm, so it is more a matter of whether they choose to spend them. We further assume that implementation costs are linear in $S$, so implementation costs increases as $S$ increases.

We further assume that the protection spending is subject to variance in its capacity to provide customer privacy protection. Consider an e-commerce transaction. From accepting an order from a customer to the final delivery, a firm faces a number of potentially harmful outcomes with respect to the use of its customers’ data that need to be handled in advance of their occurrence. Although there are no certain outcomes, it may be possible for the firm’s managers to estimate the extent of the financial losses that may occur with different levels of probability.

Jorion [7] suggested two parameters that are useful to model the losses due to operational risk. Frequency is the number of occurrence at a given time horizon, and severity is the measure of the size of the loss once it occurs. He also suggested deriving a probability distribution function of losses by aggregating the frequency and severity distribution. It may be helpful to think of this in terms of the variance in protection efficacy here. The risk of financial loss associated with improper use of private information or loss of customer data is associated with investments made by the firm to protect customer data privacy. So we capture this idea in the distribution of potential losses (which are defined in terms of their mean and variance) that the firm will face when it is not able to control its customers’ data privacy.

Let $f(n)$ be the probability distribution function for the loss frequency $n$, and $g(x)$ be the probability distribution function for the loss severity $x$. Then, the probability distribution function of loss can be written as $h(S) = \int g(x|n)f(n)dn$ [7]. We assume that increases in $S$ will reduce the frequency of loss, and so $f(n)$ decreases as $S$ increases. $h(S)$ follows an exponential distribution with parameters for frequency and severity. Overall then, $h(S)$ will represents expected losses due to privacy breaches.

The Base Case. We assume that a firm cannot price its products or services differently based on the total costs of providing a given level of protection to its customer. The more data privacy protection the firm invests in, the more the costs of implementation will be. The result is that revenue $R(S)$ is constant. In other words, the revenues associated with high-variance expected losses from less complete data privacy protection will be the same as those for the revenues and losses associated with the more complete data privacy protection.¹ When a firm provides more protection, transactions are less risky. So costs are increasing as protection becomes more complete.

Recent privacy breaches show the payoffs to customer data privacy protection are significant. Once a breach occurs, a firm is exposed to direct costs for investigations and informing customers. A firm also may face indirect costs on settlements and legal actions. Loss of trust and reputation are hard to restore fully. Thus, even though the probability may be low, privacy breaches are perceived to be very risky. So the perceived risk of no protection is greater than risk of providing some level of privacy protection.

Our model is defined as:

$$\max \pi = R(S) - C(S) - h(S)$$

s.t. $PaR = R(S) - C(S) - h(S) - Q(S, \alpha) \forall \alpha \geq k$

$$0 \leq S \leq 1$$

where $h(S) = \int g(x|n)f(n)dn$ and $Q(S, \alpha)$ represents the losses associated with the confidence level, $\alpha$. Again, $h(S)$ is the probability distribution function of losses considering their frequency and severity.

We also assume that the firm will only consider positive expected profit, where $R(S) > C(S)$. Moreover, we assume that the firm prefers more profit over less profit. In fact, the firm will prefer the highest minimum profit level at a given confidence level. Thus, they prefer more to less profit-at-risk.

3.3. Model Validation

A value-at-risk model can be tested and validated by backtesting, or systematically comparing historical $PaR$ measures with the subsequent returns [7]. Given the confidence level $\alpha$, one can test a model by counting how many times the actual loss exceeds the model’s $PaR$, after counting the number of exceptions out of total sample size. When the total sample size is large enough, a standard $t$-test can be applied to derive type 1 error or type 2 error, rejecting correct model or accepting incorrect model. To validate the model, one should consider both error types. Even if type 1 error is low, a high type 2 error means that model accuracy could be improved by increasing

¹ We admit that this is somewhat unrealistic assumption as more protection provided to customers ought to increase customers’ trust and the firm’s reputation. Also added trust and reputation will have a positive effect on customer demand and the firm’s pricing decisions [3, 5, 10, 11]. To obtain our initial results though, we only require that the revenue function be non-decreasing in the level of investment in security $S$ and be concave though. We will relax the fixed revenue assumption later.
confidence level $\alpha$. A manager can then adjust the model according to previous data. In addition to providing information on the model’s accuracy, backtesting may also inform the firm’s data collection initiatives. Managers will better understand the domain to which this approach is best suited, and what metrics provide the best decision support.

4. OPTIMAL PRIVACY PROTECTION

We assume that the probability distribution of losses at any level of $S$ is decreasing and concave. This matches the standard loss distribution in microeconomics. However, as $S$ increases, the probability of loss decreases, so the probability distribution shrinks toward 0. Maximum loss at a given confidence level is represented as a deviation from the expected loss. Figure 1 shows the probability distribution of losses for no protection ($S = 0$), and Figure 2 shows the losses at full protection ($S = 1$). Figure 3 is the midway point, when the firm gives a customer only one-half of the available protection ($S = 0.5$).

Fig. 1. Probability Distribution of Losses at $S = 0$

![Fig. 1. Probability Distribution of Losses at $S = 0$](image1)

With the profit-at-risk approach, we consider the maximum loss that occurs at a given confidence level. The dotted vertical lines in Figure 1 represent the maximum costs at 95% confidence on the probability distribution, marked by the point D. (We similarly label this in Figures 2 and 3 with E and F.) In conventional approaches to profitability analysis, only the expected values would be considered. Thus, if we only consider revenue in an expected sense, without considering the distribution of losses, then the case of no protection ($S = 0$) will always be preferred as the immediate costs will be the least. Of course, this breaks down as soon as data privacy-driven losses are considered, and so our recommendation is to consider revenues and costs in a more sophisticated manner.

Fig. 3. Probability Distribution of Loss at $S = 0.5$

![Fig. 3. Probability Distribution of Loss at $S = 0.5$](image2)

With the assumption that the reduction in expected losses decreases in $S$, the marginal benefit from investing in more data privacy protection decreases. So beyond some point, adding more protection will not improve perceived security. For our analysis, we will further assume that implementation costs $C(S)$ increase in $S$ at a greater rate than expected losses $h(S)$ decrease in $S$, that is $dC/dS > dh/dS$. Thus, total costs, which consist of both implementation costs and expected losses, increase in $S$. The expected total costs for the case of no privacy protection ($S = 0$) will be the lowest, followed by some protection ($0 < S < 1$) and full protection ($S = 1$). Since we assumed fixed revenue at any level of $S$, total costs are the only factor that affects profit.

The distribution of total costs will be narrower as $S$ increases. Figure 1 shows a probability density function for $S$ with risk depicted as the distance between the expected loss (Point A) and the maximum loss at the 95% confidence level (Point D). We see that this distance is wider than in the case of maximum protection ($S = 1$) in Figure 2 (Points C and F). So even though profit is decreasing in $S$, PaR is affected by the data privacy protection investments made to reduce risk. We further note that the profit associated with the mean value of total costs is always greater than profit-at-risk when the maximum loss is considered at a given confidence level for any value of $S$, with $0 \leq S \leq 1$. Figure 4 shows implementation costs, expected losses and total costs for the data privacy protection investment interval $0 \leq S \leq 1$.

Expected losses $h(S)$ are decreasing in $S$ and convex. Note that the total costs are subject to the interaction between the expected losses and implementation costs. With a continuous and convex total cost function, the profit function will be continuous and concave. From this, we can derive the profit-maximizing service level $S$. The first-order condition is $dC(S)/dS - dh(S)/dS = 0$; profit will be at a maxi-
mum when the marginal cost of implementation is equal to the marginal reduction of expected loss.

**Fig. 4. Implementation Costs and Losses, 0 \leq S \leq 1**

Consider the profit-at-risk function again. It is concave since maximum total costs at a confidence level in Figure 4 are convex and continuous. Optimizing $\text{PaR}$, we have $dC(S)/dS - dh(S)/dS - dQ(S,\alpha)/dS = 0$. Since the risk level falls as $S \to 1$, and the first-order condition for the severity of the loss $Q$, will be $dQ(S,\alpha)/dS < 0$. Thus, the point of maximum service-level $\text{PaR}^*$ will be greater than the profit-maximizing value $S^*$, that is, $\text{PaR}(S^*) < \pi(S^*)$. See Figure 5.

**Fig. 5. Profit and $\text{PaR}$ for 0 \leq S \leq 1**

We next assert two propositions that enable us to characterize the efficiency of protection associated with a given data privacy protection choice. We distinguish between the inefficient protection interval and the efficient protection level.

- **Proposition 1 (Inefficient Protection Interval).** When $(d\pi/dS) / (d\text{PaR}/dS) > 0$, there is always a better data privacy investment choice available that will maximize profit and/or profit-at-risk. This is the inefficient protection interval.

- **Proposition 2 (Efficient Protection Interval).** When $(d\pi/dS) / (d\text{PaR}/dS) < 0$, so that profit is decreasing while $\text{PaR}$ is increasing, investments in data privacy protection at level $S$ are subject to tradeoffs between profit and risk. We define this interval as the efficient protection interval.

A firm will not choose interval $0 \leq S \leq S'$ in Figure 5 as it does not satisfy profit-maximizing strategy of the firm; both profit and profit-at-risk are increasing in the interval. So investments in data privacy protection in this interval are inefficient: there is always a better investment level $S$ that yields more profit with less risk. This is very similar to the idea of an efficient frontier in portfolio management for financial assets. Similarly, $S^* \leq S \leq 1$ is inefficient; both profit and profit-at-risk are decreasing in this interval. So the data privacy protection investment level $S$ clearly matters in terms of efficiency. Within the interval of $S^* \leq S \leq S^*$, a firm may choose from more profit with more risk $S^*$, less profit with less risk $S^*$, or somewhere in the middle, with $S < S^*$.

The existence of an inefficient protection interval suggests that a firm will not benefit from providing data privacy protection when it an insufficient amount of money. There is a minimum threshold of privacy protection that reduces risk and supports a desirable margin to achieve a profit.

5. ANALYSIS

We next analyze data privacy protection with: low anticipated risks and an efficiency protection investment level; high implementation costs, when the firm passes on implementation costs; and when the privacy protection investment that is made is complete. The idea of completeness of data privacy investment suggests that the firm has covered all reasonable risks. The investments reflect the firm understands up to the limit of managers’ bounded rationality on what they know about the threat sources.

5.1. Low Anticipated Risks, Efficient Protection

We earlier noted that it is important for a firm to protect its customers’ privacy properly, but at the same time, it is difficult to protect it completely. The process of collecting, storing and analyzing personal information occurs across the spectrum of a firm’s operations. Also, the possible means of infringement are unpredictable. Also, the scope of privacy or private information is loosely defined and subject to variability across consumers with heterogeneous preferences [15, 17]. Due to the ambiguity of privacy protection, many firms underestimate the risks of privacy protection. Now let us revisit the data privacy protection optimization problem. See Figure 6.
The difference between the maximum profit and maximum profit-at-risk is derived from the presence of the severity of loss distribution \( Q(S, a) \). Underestimating the risk that is present is associated with a smaller variance in \( Q \), thus reducing the efficient protection interval. Figure 6 shows an example with a smaller variance in \( Q \). Now compare this with Figure 5. The efficient interval is narrower with a smaller variance in \( Q \). Though \( S^* \) does not change, \( S^\hat{a} \) is lower in Figure 6. This suggests that underestimating the anticipated risk level will result in a lower protection level, with less investment in data privacy protection.

### 5.2. High Implementation Costs

Alternatively, we can consider what happens if the costs associated with the data privacy protection implementation are high, keeping the risk factors constant. Figure 7 shows the case where the costs of implementation \( C(S) \) are high, while it provides the same risk reduction as our base case. Due to the higher costs, the inefficient protection interval is wider and the value of PaR at all levels is decreased. Clearly, implementation cost plays an important role for the firm to make an optimal investment decision.

### 5.3. Passing Costs on to Customers

In terms of the goal of minimizing costs, protecting privacy is just another element of overhead from a firm's perspective. The firm that does not protect its customers’ information may lose standing with them, when there data are lost (e.g., the T. J. Maxx case, or other recent cases involving credit card data and social security numbers). Many justifications for investing in solving privacy problems have been suggested. Most of the studies relate protecting privacy with building trust and reputation though \([15, 17]\). This is an incomplete view, based on how we have conceptualized this problem.

To maintain profitability while providing proper privacy protection, a firm may pass on the costs associated with implementation to its customers by increasing the prices of its products or services. Let us assume that there is no change in demand based on the range of the price change. Interestingly, it may even be possible for a firm to offer a higher level of protection, charge more, and engender more trust and a stronger reputation in the market. This will allow a firm to benefit from the price premium it charges.

Consider two scenarios. The first is where the firm passes on some of the implementation costs to its customers. Figure 8 shows the changes in PaR and profit when the firm passes on 50% of the data privacy implementation costs to its customers. A second scenario is to pass 100% of its implementation costs to its customers, as shown in Figure 9.
plementation costs are fully passed on to customers at the point of maximum profit and maximum profit-at-risk: this is \( S = 1 \), or full investment. This is due to the fact that the costs are fully passed on to customers, and, thus, there are no losses or additional costs to the firm for implementing it. So even at the 50% of full investment level, as shown in Figure 8, the efficient protection interval is right-shifted toward \( S \rightarrow 1 \).

Another related observation is that the efficient protection interval starts out lower than in the base case, and is wider too. This is obvious as the firm suffers less from the costs associated with providing protection services. However, it is well to keep in mind that the amount of implementation costs that can be passed on to customers without diminishing their demand will be limited after some price premium level. So if the implementation costs are large, or the price premium from added trust and enhanced reputation is small, not all of the implementation costs will be able to be passed on to customers.

5.4. When the Firm Chooses Full Protection

The efficacy of two different approaches, self-regulation policy and government regulation in protecting privacy, has been a subject of debate. For example, in the United States, state or federal legislatures have passed many regulations to protect consumer privacy. However, even well-intentioned legislation can have negative consequences [16]. But, as Ashrafi and Kuilboer [1] argued, self-regulation is not always feasible, and firms will act to self-regulate only within the availability of their limited resources. Studies in this category suggest that it is not enough to adhere to self-regulation since firms would choose to adopt the level of privacy protection that suits their convenience. Although the regulations offer detailed guidelines for privacy protection, they typically are either too broad in their contexts of application, and do not consider the characteristics of specific industries, or not specific enough for the regulations to be enforceable. However, due to growing concerns about consumer privacy and an increase in privacy regulations in the United States economy, firms are increasingly expected to maintain higher standards of online privacy [9].

Let us revisit our data privacy protection interval analysis results once again, with this broader discussion in mind. For example, consider the protection interval between \( S' \) and 1. In this interval, both profit and \( PaR \) decrease as \( S \) increases. So as profit-at-risk decreases, the firm will provide more protection to its customers in this region. The costs required to implement the data privacy protection offset the benefits of protection that more investment would provide. Our model shows that a firm will only choose \( S = 1 \) when a firm can fully pass on the implementation costs for data privacy to its customers.

From the security perspective of customers alone, more data privacy is preferred, and full protection would be beneficial. The reality is different though. From an economic standpoint, efficient protection may mean less-than-best full protection. Firms must consider the costs of implementation and operation, as well as the costs associated with the risk. We have seen that costs will outweigh the benefits at some a specific point \( S' \) in the analysis space. This finding agrees with prior empirical research [1], which shows that most firms do not implement to achieve full privacy protection.

6. SEVEN STEPS FOR MODEL APPLICATION

We present a process description to illustrate how managers make an investment decision in Textbox 1.

Textbox 1. The Model-Based Decision Process

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Collect parameter values. Manager should collect parameter values and derive reasonable functional forms for ( R(S), C(S) ) and ( h(S) ). Historical and benchmarking data will be helpful.</td>
</tr>
<tr>
<td>2.</td>
<td>Define confidence level ( a ). Defined this based on managerial strategy. A 95% confidence level may be adequate initially, as it can be adjusted after backtesting methods are applied.</td>
</tr>
<tr>
<td>3.</td>
<td>Derive the profit-maximizing service level ( S ). After setting up the model, a profit-maximizing service level ( S ) can be derived based on the first-order condition ( dC(S)/dS = -dh(S)/dS = 0 ).</td>
</tr>
<tr>
<td>4.</td>
<td>Choose optimal service level ( S' ). After deriving the maximum profit and ( PaR ), the optimal service level ( S' ) within the efficient protection interval can be chosen.</td>
</tr>
<tr>
<td>5.</td>
<td>Validate model. By backtesting, a manager can validate the model using previous data. If it is unacceptable, then change the confidence level ( a ) or the profit-maximizing service level ( S' ), and go to Step 2 or Step 4 and repeat the analysis.</td>
</tr>
<tr>
<td>6.</td>
<td>Make the investment decision. Do this based on the model’s output. Other managerial conditions and market conditions also should be considered. If external conditions prompt a change, the model can be re-validated by repeating Step 5.</td>
</tr>
<tr>
<td>7.</td>
<td>Adjust model. Periodically test/adjust for market changes.</td>
</tr>
</tbody>
</table>

Since our approach offers flexibility to adjust to external market conditions and is testable, managers can justify their decisions, and consider other factors not included in the base model.

7. CONCLUSION

We close with contributions and limitations.

7.1. Contributions

Our main contribution is to suggest a profit-at-risk approach that is derived from value-at-risk methods in financial economics to evaluate the optimality of firm-level decisions to protect customer data privacy. It is well known that there is a trade-off between security cost and protection, but to the best of our knowledge, this is a the first attempt at applying risk-return trade-offs in a privacy protection investments.
Even though the probability of a privacy breach may be very low for most of the firms, the associated risk position may be quite high because of potential extreme losses. Thus, some firms invest in privacy protection services to reduce their risk exposure. It is unclear how much a firm should invest though. Spending too much will cause a financial burden; while spending too little exposes it to unacceptable losses. We addressed this issue, and provided in-depth analysis of the dynamics of privacy protection investments via risk management methods.

Our model can be implemented in a real-world decision system. As it is tested, managers will see the benefits for gathering privacy-related security data, potentially sharing this data among firms. This will inform metric development and show managers the best approaches for data collection.

We have shown that optimal investment decisions for privacy protection depend on the tradeoff between the cost to implement the service and the reduced risk due to the service rendered. Practitioners may benefit from our model as it is directly applicable. By applying reasonable values to each attribute, it is possible to come up with an optimal investment choice for the firm. The profit-at-risk method allows setting up a confidence interval that the manager thinks appropriate. Conversely, managers also can test the effects of small changes in each attribute by performing sensitivity analysis on the model’s elements. This approach would further justify the investment.

We also have shown that there is a minimum protection level that the firm will need to achieve to make its investment in customer privacy protection effective. We called this the inefficient protection interval, and within it a firm incurs higher costs of implementation and operations relative to the associated risk reduction that the protection services provide.

Ambiguity in understanding what constitutes effective privacy protection, many firms tend to underestimate the risks. When a manager believes that the risk is likely to be low, the firm will invest less. Another factor that affects the investment level is the cost of implementing and operating the protection. Our model confirms that express lesser concern about risk, then the optimal value of PaR will occur at a lower level of protection. This implies that proper evaluation of expected risk is important to set the confidence level associated with privacy breaches.

Our model also confirms why firms do not spend fully for privacy protection. Since the risks that companies and their customers face with changing technology, developments in hacking, and evolving effectiveness of firm-level information security, it is likely that a firm and its customers may not be completely secure, even if the firm invests fully in data privacy protection. Thus, investing to achieve the maximum level of protection often is not an optimal strategy. Firms are better off when they consider maximizing profit-at-risk instead, and then investing accordingly.

One way to manage profitability in the presence of privacy protection investments is for the firm to pass all or some of its implementation and operation costs to its customers. Although this is rarely the first thing that most managers think about—just as it is not the first thing we have considered in our base model for data privacy protection investments—it is a necessary consideration. Consumers are likely to be willing to pay for this protection to some degree. Our model shows that firms should provide more data privacy security when they share protection investments with their customers. This is not always true, of course. Other factors, such as high price competition, price elasticity of market demand, and so on, should be considered during the decision-making process.

7.2. Limitations

We conclude by pointing out some caveats and limitations. In our model, the quality of all services is weighted equally. Adding any services will incur the same cost and revenue effect. Only the continuous quantity of services matters in the model. Some service elements may not cost the same amount though. Examples are the provision of payment services using secure third-party payment methods, opt-out choices on promotional mailing lists, and opt-out option to not save customers’ private information that is used during transactions. Different kinds of protection require different implementation costs. For the analysis we have conducted, this issue does not affect our results. Our model deals with the average costs incurred. So any implementation costs that deviate significantly from average costs will affect our policy recommendations and results.

We also assume every customer has homogeneous perceptions about the value of privacy protection. We only consider their average concerns about how well privacy protection works. There is a tradeoff between privacy concerns and secondary use of private information though. Customers will benefit from using a more sophisticated approach to profiling, even though they may lose some privacy in exchange for the benefits. There is a wide spectrum of attitudes among customers. Some willingly trade off privacy for better services due to profiling. Others value privacy more, no matter what improvement in service can be created. Culnan [4], more generally, found that there is a trade-off between the potential benefits of direct marketing, and consumer concerns about
privacy. She concluded that a customer who is more interested in the utility of direct marketing will value her privacy less than one who has a lower level of utility for direct marketing. Overall, Wang et al. [18] have suggested that many consumers the kind of information that is captured in profiling to be private information. They feel uncomfortable when merchants improperly collect, store and analyze such information for marketing purposes, and share such information with others, without their permission.

Finally, our model relies on estimates of the frequency and magnitude of future losses. We expect firms to build predictive models based on historical data to estimate future losses when customer information is compromised. Third-party firms which provide security benchmarks, auditors, and standards bodies may be good sources of such data.

REFERENCES

Acknowledgments. We appreciated input from the minitrack co-chairs, and the anonymous reviewers. Rob Kauffman thanks the W. P. Carey Chair for partial support.