Why Companies Change Privacy Policies: A Principal-Agent Perspective

Wallace Chipidza  
Baylor University  
wallace_chipidza@baylor.edu

Dorothy Leidner  
Baylor University  
University of Lund  
dorothy_leidner@baylor.edu

Debra Burleson  
Baylor University  
debra_burleson@baylor.edu

Abstract
Companies often update their privacy policies, but it is unclear whether these updates strengthen or weaken user privacy. Using a principal-agent theoretical lens, we analyze over 200 changes to privacy policies of five prominent Internet companies over a period of fifteen years. We find that over time, privacy policy updates are more likely to weaken, rather than strengthen, user privacy on the Internet. Moreover, analysis of pilot data suggests that companies are more likely to update their privacy policies if they experience negative revenue growth over successive financial reporting periods. Since our results show that changes in privacy policies benefit companies at the expense of user privacy, we conclude that the principal-agent problem exists in the information privacy arena.

1. Introduction
In a widely publicized story, Gamestation, a British video game company, revealed that it legally owned the souls of thousands of its customers [5]. Unbeknownst to customers, the company had hidden a “soul-selling” clause in its terms of use as an April’s Fool prank. Although the company included an option for the customer to opt-out of the agreement and receive £5 in return, close to 90% of customers agreed to the terms of use. By playing this elaborate prank on its customers, Gamestation made a subtle but very serious point about user behavior on the internet: very few users know what they are agreeing to when they click “I accept”.

Internet users have to sacrifice some privacy in order to access a service. The company's privacy policy governs this exchange and determines just how much privacy the user will lose from accessing the service. Although privacy policies were initially crafted to eradicate the information asymmetry between companies and their customers, companies rarely identify the third parties with whom they share user information. As a result, most Internet users do not know exactly how many and which entities have access to their personally identifying information, and what they intend to do with it. Not surprisingly, opposition to third-party tracking is emphatic: independent privacy watchdog TRUSle estimates that 85% of Internet users object to tracking for the purposes of ad targeting [36].

The overwhelming opposition to measures that restrict user privacy by companies stands at odds with actual user behavior on the Internet. Past research has shown that companies routinely sell user information to third parties, send cookies to user devices and profile users for the purposes of ad targeting ([32], [31]); yet users continue to access these services at increasing rates [37]. This suggests that users might not know what is contained in the privacy policies they consent to. Previous studies on company privacy policies have generally given a snapshot of the global state of user privacy at a particular point in time ([11], [20]). This study significantly departs from that norm and instead analyzes changes in privacy policies over time. By investigating longitudinal patterns, we reveal a more useful picture of how companies respond to the privacy concerns of customers, if at all.

In our analysis, we seek to answer the following research questions: 1) What are the factors that influence companies to update their privacy policies? and 2) What is the nature of most privacy policy updates? By answering these questions, we can reveal how the threats to user privacy have evolved with time, and how companies respond to user privacy concerns. In this regard, our study addresses the lack of privacy research at the organizational level of analysis as highlighted in past privacy research [3]. Moreover, our first question answers the call by privacy literature reviewers [30] to investigate “the why related to privacy” so that we can achieve a more complete understanding of the factors that lead to loss of user privacy on the Internet.

Our study reveals that companies are more likely to make privacy policy changes that weaken, rather than strengthen, user privacy on the Internet. We also find that companies update privacy policies in response...
to declining levels of revenue. Lastly we reveal that technology companies behave as agents to multiple groups of principals, a situation which complicates privacy policies and generally disadvantages the consumer in an environment characterized by an ineffective monitoring system. We use our findings to formulate a set of propositions that can be used as a foundation towards building a theory of why companies change privacy policies.

The rest of the paper is organized in this manner: first, we present the theoretical underpinning of our study; we proceed to describe the method of data collection and analysis before presenting our findings, formulating a set of propositions, discussing our contribution in detail and concluding.

2. Theoretical Foundation

2.1. The Principal-Agent Problem

The principal-agent problem arises out of a relationship where one party (the agent), performs a certain function on behalf of another party (the principal), but the latter cannot observe the actions of the former [15]. The two parties share the risk and the payoff that results from the performed action. As identified in economic theory, the ingredients for a principal-agent problem are threefold. First, the setting must involve an agent performing a function on behalf of a principal. Second, the actions taken by the agent are not fully transparent to the principal (moral hazard). Third, the agent has more information concerning these actions than the principal (information asymmetry). Because the agent possesses an information advantage over the principal, he may take actions that benefit him but hurt the principal.

The information privacy arena contains all the ingredients of the principal-agent problem. In the relationship between the user and the company, the company is the agent and the user is the principal because the user entrusts the company with custody of his information. In maintaining custody of personal information, the company has the ability to make decisions that impact the user’s privacy. Because the information privacy interests of a company and its users may be misaligned, and even diametrically opposed, the company will exploit its custody of personal information in ways that erode user privacy.

Moral hazard stems from a misalignment of interests between the user and company: whereas the user wants to limit information reuse and unauthorized access to personal data, the company aims to maximize revenue and minimize costs. The company has incentive to expand its revenue base by selling data to third parties and may reduce costs through sub-optimal investment in information security. Underscoring this latter point is the prevalence of data breaches in the USA and elsewhere ([6], [8]). Since the principal is unable to observe the protection mechanisms placed by the agent in order to safeguard user data, and the company has incentive to reuse information and limit information security costs, moral hazard exists in the Internet privacy context.

Privacy policies reduce, but do not completely eradicate, information asymmetry between consumers and online service providers. The cost of reading a privacy policy is non-trivial: reading a single privacy policy requires an average of 10 minutes. Individuals have relationships with many organizations. To read the privacy policies of each organization with whom an individual transacts would require significant time. In the United States, the national opportunity cost for reading privacy policies amounts to $781 billion per annum [20]. Because privacy policies are updated several times a year, this cost might actually be understated. Consequently, companies retain a significant information advantage over their customers concerning their actual practices.

Even if every user would read all the privacy policies, some information would still remain hidden. Policies typically include stipulations that information may be shared with third parties, affiliates and law enforcement; however, the identities of third party entities are seldom revealed. Hence, the company knows more about the companies with access to a user’s data than the user does. In addition, websites are often accessible in multiple countries. Since privacy policies normally clarify that companies have to accede to data access requests by law enforcement, the potential exists that repressive governments may demand personal data in order to quell dissent. In many cases, the company may not even inform the user about such data access by the governing authorities [13]. Certainly, this is a clear example of hidden action that the agent knows about but the principal does not. The user-company relationship satisfies the twin criteria of moral hazard and information asymmetry and therefore exemplifies the principal-agent problem.

In the following section we identify the most prominent threats to user privacy from the information privacy literature. These threats will serve as input criteria for our data coding.

2.2. Threats to User Privacy

Consumers have certain expectations when they supply personal data to companies. Early research into information privacy showed that consumers are concerned about how companies use their data [9]. In
particular, consumers find selling of data to third parties objectionable even when such practices are legal. Profiling through combining various pieces of information from disparate sources elicited negative reactions from consumers. Such uses of data are examples of hidden action that constitute moral hazard in the principal-agent relationship.

The information privacy interests of consumers and firms do not always align. Son and Kim [32] posited that the use of customer relation management software for targeted marketing could conflict with consumers’ desire for information privacy. They proceeded to identify categories of user responses to privacy threats: refusing to use the service, stopping to use the service, disseminating negative word-of-mouth reviews to friends and family, and complaining to the company and to third party companies. These responses indicate that there is a cost incurred by companies when users perceive threats to their privacy.

Individuals benefit from a personalized experience when using various websites; however, this gain comes at a cost to personal privacy because personalization requires users to share more personal information with the service provider. A study investigating the tradeoff between online personalization and privacy found that in making the decision to accept or reject personalization, consumers seek to maximize their utility by minimizing the cost of ceding their privacy to online service providers [2]. If the cost of threats to individual privacy exceeds the benefit from personalization, then the consumer receives negative utility. This suggests that individuals with higher concerns for privacy are less willing to be profiled for personalized service.

In light of the threat that personalization, targeted marketing and profiling entail for individual privacy, various advocates for individual control over privacy choices have encouraged organizations to incorporate ethical concerns into the design of privacy programs [11]. The reuse of submitted information and unauthorized access to customer data pose significant threats to privacy [11]. Because the average privacy policy takes 10 minutes to read, even if users are averse to an invasion of their individual privacy, they might not take the time to read the full policy. In fact, the longer and more tedious the policy, the more likely that users will not read the policy and, hence, accept a compromise to their privacy. Table 1 summarizes the threats to user privacy.

### 3. Methodology

To answer our research questions, we analyzed privacy policies from Amazon, Dropbox, Facebook, Google and Microsoft. We chose these organizations because they are market leaders in their respective Internet industries: Google in search, Amazon in e-commerce, Dropbox in cloud storage, Facebook in social networking and Microsoft in user retail software.

<table>
<thead>
<tr>
<th>Threat</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reuse of information</td>
<td>Culnan [9], Culnan &amp; Armstrong [10]</td>
</tr>
<tr>
<td>Unauthorized access</td>
<td>Culnan &amp; Williams [11]</td>
</tr>
<tr>
<td>Personalization</td>
<td>Awad and Krishnan [2]</td>
</tr>
<tr>
<td>Use of cookies</td>
<td>Soltani, et al [31]</td>
</tr>
<tr>
<td>Profiling/Ad Targeting</td>
<td>Son and Kim [32]</td>
</tr>
</tbody>
</table>

Of these companies, only Google [14] maintains an archive of previous policies on its website to enable customers to view changes to the policies. For the rest of the companies in the list, we retrieved archives of privacy policies from the WayBack Machine at Archive.org [18] and the TOSBack repository hosted at Github [34]. Our study focuses on the period 1999-2015.

We loaded every privacy policy version into a Github repository. Each time someone makes a change to a file in the repository, the Git software generates a unique hash for the commit using the SHA-1 algorithm. This system-generated hash output allowed us to reconstruct a version of the file associated with that change. Using Git commands on a UNIX terminal, we were able to recreate all the versions of the privacy policies that constituted our sample.

Once we had all the versions of the privacy policies in place, we sorted them in order of commit date. This enabled us to compare successive versions with the diff utility program in UNIX and hence capture all the changes between versions. For example, if the commit history shows two successive versions of the Dropbox privacy policy file with hashes 2lx56v51 and 3c4d3aa3, the command to compare the versions is:

```
~$ git diff --c 2lx56v51:.dropbox/Policy.txt 3c4d3aa3:.dropbox/Policy.txt
```

In total we analyzed 218 privacy policy changes. The breakdown of the sample was as follows: 39 changes from Facebook, 50 from Microsoft, 93 changes from Google, 25 from Dropbox and 11 from Amazon. We coded the data based on the threats shown in Table 1 and how companies responded to them. Two researchers independently coded the data; the inter-rater reliability was 95%.

We collected financial data covering the period 2008 to 2015 from Statista [33] and company websites in order to explore possible factors that influence
companies to change their privacy policies. Specifically, we recorded the quarterly gross revenue attained by each company over the eight-year period. We compared the revenue data from seventeen financial quarters with policy updates to an equal number of quarters, chosen at random, where privacy policies had not been updated. Then we analyzed the relationship between the change in revenue and the decision to update the privacy policy by performing a binomial logistic regression between two variables: 1) the preceding upward change in the growth calculated as following, where $R_k$ is the revenue accumulated and $G_k$ is the upward rate of change in growth in period k:

$$G_n = (R_n - R_{n-1}) / R_{n-1}$$

and 2) a dummy binary variable that indicated whether there had been a privacy policy update or not in that quarter; 1 meant yes and 0 meant no.

Although no reliable estimates of the average value of personal information per Internet user exist, a 2011 study by the OECD proposed various methods for estimating its monetary value [23]. The OECD recommends using the average market capitalization per user as a proxy for the monetary value of personal data. We obtained this data from company websites and Statista [33]. Although this method has a limitation in that a relatively small proportion of technology companies are publicly traded, it still provides a useful measure of how much companies gain by accumulating users. A second method involves calculating the cost of a data breach per user as a proxy for the monetary value of data. We sourced security breach cost data from the Ponemon Institute ([17], [26], [27]). We used these two methods to evaluate trends in the monetary value of personal data.

4. Findings

Even though we present the relative frequencies of the different categories of policy changes in Table 2 below, there is a limitation in quantifying the changes: a number of small inconsequential changes can easily be surpassed by a single policy update in impact. There were common trends across versions of privacy policies belonging to the different companies. Over time, it’s clear that technical capabilities improved and this resulted in longer and more complicated privacy policy documents. The privacy policy of the companies grew an average 190 words per year. Consequently, because the cost of reading a single privacy policy rises with time, the information asymmetry between companies and their users continues to widen. Changes to privacy policies reflected advances in technology; in addition to cookies, companies now use pixel tags (also known as web beacons), to track user activity on their websites. Furthermore, the importance of legal compliance rose with time; for example, Google’s original privacy policy does not make an explicit reference to compliance with law enforcement, but all current privacy policies within our sample reflect the fact that companies currently comply with requests from law enforcement agencies. As time progressed, the implications of the principal-agent problem became clearer and consumers became more concerned with their information privacy resulting in the creation of privacy watchdogs like TRUSTe that mediate privacy-related disputes between users and companies. Moreover, companies realized that they had to be in compliance with regulatory regimes in different jurisdictions. As a result, all privacy policies of the companies within our sample now specify that they comply with the EU-US Safe Harbor framework.

<table>
<thead>
<tr>
<th>Nature of Change</th>
<th>Relative Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restrictive</td>
<td>61</td>
</tr>
<tr>
<td>Empowering</td>
<td>29</td>
</tr>
<tr>
<td>Law Enforcement/Regulatory</td>
<td>7</td>
</tr>
<tr>
<td>Watchdog Monitoring</td>
<td>3</td>
</tr>
</tbody>
</table>

We describe the four categories of privacy policy changes in detail in the following subsections.

4.1. Restrictive Changes

Restrictive changes refer to any changes that increased the chances for hidden action by companies or their affiliates and third parties. The more personal data collected by the company, the greater the potential for hidden action. Examples of personal data that can be used for hidden action include email address, physical address, phone number and date of birth. From the consumer’s point of view, selling of data to third parties or sharing data with affiliates threatens users’ privacy. As more entities gain access to a consumer’s information, that user’s privacy is eroded. The erosion may be realized in many ways, including unwanted phone calls, unsolicited emails and ad targeting when browsing the Internet. Furthermore, third parties may also have different privacy policies that govern their use of personal data. Thus, there is no upper limit to the number of privacy policies that a
user has to read in his interactions with a single website in order to fully understand what can happen to his personally identifying data.

We present the subcategories of restrictive changes that we found in the following subsections.

### 4.1.1. Affiliates and third parties

By far, the majority (61%) of changes to company privacy policies weakened user privacy. Of these, expansion of access to affiliates and non-affiliated third parties dominated the pool. Acquisitions by technology companies are common [28] and usually involve sharing of existing user data with the joining entities. For example Facebook, which has been in existence for only 11 years, has acquired 53 companies [4]. Since each acquired company potentially has access to all the data under the control of the parent company, a considerable erosion of user privacy occurs as a result of a merger or acquisition.

Furthermore, since new entities require new portals of access to user data, acquisitions and partnerships potentially open up new lines of attack for unauthorized access. A case in point is the widely publicized Target data breach of 2013, which occurred through a third party [38]. As companies share data with more third parties in pursuance of more targeted advertising or outsourcing specialized services such as storage and testing, user data ends up in multiple, previously unanticipated places.

Third-party applications also threaten aspects of the user experience other than privacy. According to Pathak, et al. [24] up to 66% of the energy used in free apps is spent executing code from third-party advertising sites. When computing resources are scarce, for example when battery levels are depleted or network bandwidth is narrow, third-party applications can significantly degrade the user experience, hence reducing the utility obtained by the user. As a result, third-party applications exacerbate the principal-agent problem by disproportionately rewarding the company at the user's expense.

### 4.1.2. Geographical expansion

A non-trivial proportion of policy changes reflect how technology has lowered geographical barriers to entry; because software development teams can collaborate across continents using virtualization software [7], companies typically outsource a lot of computing services such as storage and processing to other countries. Although this increased globalization creates new opportunities for technological progress and expansion, the phenomenon also presents challenges for user privacy. For example, in July 2004, Google added the following change to their privacy policy: “We may store and process personal information collected on our site in the United States or any other country in which Google Inc. or its agents maintain facilities. By using our services, you consent to the transfer of your information among these facilities, including those located outside your country.” As a result, a user’s personally identifying information can be exported to other countries, a situation which significantly threatens user privacy.

### 4.1.3. Narrowing user involvement

In general, our study shows that user privacy is eroded as time progresses. In fact, in no case did user privacy strengthen over time, rather changes to privacy policies had the overall effect of weakening user privacy. Several changes to companies’ privacy policies demonstrate this trend. As early as 2007, Amazon informed users that they could access its website anonymously through the use of third party tools. However, it soon removed that information from its policy and expanded access to customer data to third party advertising sites. And in 2009, Facebook decided to democratize its policy change process. Under the new framework, users would be allowed to comment and vote on proposed changes to the privacy policy under the **Opportunity to Comment and Vote** section: “If we receive more than 7000 comments concerning a particular change, we will put the change up for a vote. The vote will be binding on us if more than 30% of all active registered users as of the date of the notice vote.” However, in November of 2012 the company discontinued the democratic framework with the following amendment to its privacy policy under the amended headline **Opportunity to Comment** section: “After the comment period, if we adopt any changes, we will provide notice (for example, on the Facebook Site Governance Page or in this policy) of the effective date” hence effectively excluding user input from the policy changing process. And in yet another example of how companies degrade user control over their privacy is the amendment of Bing’s privacy policy in 2014. Whereas Microsoft previously promised to respond to user requests with the following declaration: “We will respond to requests to access or delete your personal information within 30 days”, the company subtly changed the language to the more vague “We will make reasonable efforts to respond to your web request to access or delete your personal information within 30 days” in the next version of the policy. As a result, a user who requires urgent deletion of his information from Bing has no guarantee that his request will be granted, or even responded to, in a specified time period. Users have little recourse when they desire their information to be deleted but the company does not respond.

### 4.1.4. Personalization

One way that personal data is collected is through user preference features. Such features are used to provide personalization of websites according to the user preferences, but at the same time,
they require users to share information with the service provider. This information may then be shared with third parties, resulting in hidden actions and threatening user privacy. The following passage in Amazon’s pre-2007 privacy policy shows this tradeoff between anonymity and personalization in action: “A number of companies offer utilities designed to help you visit Web sites anonymously. Although we will not be able to provide you with a personalized experience at Amazon.com if we cannot recognize you, we want you to be aware that these tools exist.”

4.2. Empowering Changes

Empowering changes refer to changes in privacy policies that strengthened user privacy. In the online privacy context, the company can provide the ability to opt-out of sharing data with other parties. Encryption, which blocks access to data from unauthorized users, is another example of an empowering change. For example, Dropbox, the market leader in cloud storage, encrypts user files on its servers by default. With encryption, even if there is unauthorized access through a security breach for example, the data cannot be decrypted and hence user privacy is preserved.

Our study revealed that companies also reveal information that helps customers take control of their privacy. For example, as reflected in its privacy policy update in 2005, Google requires the user to “opt-in consent for the sharing of any sensitive personal information.” All the companies surveyed now use the HTTPS protocol as the default so as to protect their customers from Internet channel eavesdroppers, thereby shielding customers’ sensitive data from possible theft. However, while this may protect consumers from illegal theft of their personal data, it does little to prevent unwanted, but legal, uses of their personal data by affiliates of the organization. Indeed, organizations appear motivated to ensure consumers against theft of data as indicated by the fact that 10% of all the changes that we classified as empowering contained references to information security. However, this motivation might stem as much from the desire to prevent bad publicity as the desire to protect consumers from unwanted uses of their data.

Furthermore, advances in privacy-enhancing tools enable companies to equip their users with the ability to encrypt their data, enable anonymous browsing and turn off location tracing when using certain sites. For example, Google and Bing search sites grant users the ability to browse the Internet anonymously. In general, companies included privacy policy changes to reflect these advances, but these empowering changes were seldom the default setting on user accounts and users might not even be aware that their data was not encrypted or that location tracing was turned on by default. Ultimately, a proven method of resolving the principal-agent problem is through information, teamwork and co-operation between the two parties [22]. Empowering changes entail providing such information to consumers to reduce asymmetry and cooperating with them to control the collection and distribution of their personal information.

4.3. Law Enforcement/Regulatory Changes

As state and federal laws are promulgated or amended, companies respond to these legislative developments by updating their privacy policies. When conducting legal investigations, government agencies may request information belonging to certain users; such requests typically remain invisible to the affected user and hence the information gap between the user and the company is widened.

4.3.1. Legal compliance. Compliance with law enforcement emerged as a prominent theme in changes made to privacy policies by companies. However, it is only recently that companies have begun to make explicit references to legal obligations in their policies. For example, Google was in operation for five years before adding the following change to its privacy policy under the Information Sharing heading: “We may share … information (if) we conclude that we are required by law or have a good faith belief that access, preservation or disclosure of such information is reasonably necessary to protect the rights, property or safety of Google, its users or the public.” As pointed out in an earlier section, the preceding statement does not specify boundaries as to which jurisdictions Google will cooperate with law enforcement.

4.3.2. Privacy protection and legal enforcement. Companies have to maneuver a delicate balancing act between complying with law enforcement and placating user privacy concerns. For example, Dropbox changed its policy to explicitly inform customers that it would comply with any requests for files and user data from law enforcement, and that it would even decrypt user files in the process. At the same time, the company explained that any files already encrypted before storage on Dropbox would be safe from law enforcement’s prying eyes, hence informing users of a workaround regarding this concern. This change underscores the tricky waters that companies tread in acting as agents to multiple groups of principals.

4.4. Watchdog Monitoring Changes

Three percent of the policy updates in our study contained references to watchdogs that mediate
disputes between users and companies concerning potential infringement of user privacy rights. In the US, the three major watchdogs TRUSTe, Better Business Bureau and JAMS monitor how technology companies protect customers’ privacy. Companies that meet these standards receive a seal of approval. Some of the changes in this category further confuse matters, however; for example, Dropbox’s privacy policy update in February 2014 contained the following change: “JAMS is the independent organization responsible for reviewing and resolving complaints about our Safe Harbor compliance.” Previously, TRUSTe had been the watchdog responsible for resolving disputes between Dropbox and its users. There was no explanation as to why Dropbox changed its watchdog. In this regard, the privacy policy change does little to narrow the information gap between the user and the company: a user can reason that Dropbox changed its watchdog to one that would be more amenable to its preferences in a dispute with a user.

Since the changes explored in this section fall into different categories, it is important to understand how companies decide which changes to include in their policy updates. In the following section we propose a set of factors that may influence companies to change their privacy policies.

5. Why do Companies Change their Privacy Policies?

Our data provides insights into three possible reasons that organizations make such changes.

5.1. Declines in Revenue Growth

The sharing or selling of data to third parties constituted a plurality of restrictive changes to privacy policies. This suggests that revenue growth is a strong motivator for privacy policy updates.

As stated in the method, we analyzed the relationship between updating the privacy policy and the percentage upward change in growth. Because we modeled the decision to update the privacy policy as a binary dummy variable, we employed the binomial logistic regression method in SPSS to measure the likelihood that a company will update its privacy policy given a change in revenue growth rate. We found the relationship significant (β = -0.14, p < 0.05, n = 34). As shown by the Nagelkerke R² statistic [35], the logistic regression model accounts for 32% of the variance in a company’s decision to update its privacy policy. Specifically, for each percentage point decrease in revenue between successive quarters, the odds that a company will update its privacy policy increase by 13%. These results are consistent with the principal-agent perspective. As companies experience declining or stagnating revenues, they are incentivized to exploit the personal information within their custody in order to advance their financial interests, even at the expense of customer privacy. As a result, we propose that:

P1: A company is likely to update its privacy policy document if it experiences a decline in revenue growth over successive financial reporting periods.

5.2. Increasing Monetary Value of Data

Another reason for changing policies might be the increasing monetary value of personal data. Increasing value of data incentivizes companies to exploit the personal data within their control. The principal-agent problem resident within the information privacy arena means that companies have an information advantage over their users and can take actions that are hidden to their users. These conditions allow companies to expand the collection and exploitation of personal data without fearing customer backlash. As a result, as the monetary value of personal data increases, companies are more likely to alter their privacy policies for their advantage.

As companies develop more data collecting and inferring capabilities, they acquire the ability to exploit the data. For example, over the past few years, geographical location algorithms have become more accurate, thereby affording companies new opportunities to generate revenue through more precise location-based advertising. Consequently, a non-trivial portion of privacy policy changes in the last few years was focused on how the user can control his location-based technology settings. For example, Google previously mentioned that it had the capability to retrieve “approximate user locations”, but edited that portion to “exact location” in a later version of the file. The value of personal data, as measured using the proxies described in section 3, has been steadily increasing at a rate of between 2% and 3% per year since 2008, and by up to 12% since 2013 [17]. Because companies want to maximize revenue, they update privacy policies to reflect changes in the manner in which they use personal data. We therefore propose:

P2: Companies are likely to update their privacy policies if the monetary value of personal data increases.

5.3. Informing Users

A lot of privacy policy changes reflected the intended role of the policy document, i.e. informing
users of the environment that houses their personal data. Phenomena that exhibit the principal-agent problem often involve a trade-off calculation: the agent has to find an optimal payoff ratio that works to his advantage whilst not completely alienating the principal. In the privacy context, the payoff is in the form of user privacy; hence the agent determines just how much user privacy to take away from the principal. If the company annexes so much privacy that the user notices, then users may switch to other services. Furthermore, as privacy concerns gain prominence, updates to company privacy policies are receiving increasing levels of scrutiny in the media. Therefore, at a minimum, the company’s privacy policy must perform its intended role of narrowing the information gap by informing users of the internal policies that govern the use and storage of personal data.

Moreover, policy updates also serve to assure users that companies adhere to the regulatory frameworks that govern the use of personal data, such as the US-EU Safe Harbor Program. Lastly, some policy changes can be seen as implicit reminders to users to not use company services for criminal activity, as evidenced by the 7% of changes that explained that companies comply with law enforcement requests for personal data, leading us to propose that:

P3: Companies are likely to update their privacy policies if there is a change in the internal, legal and/or regulatory frameworks that govern information privacy.

6. Discussion

6.1. Contributions and Implications

The principal-agent perspective postulates that where there is information asymmetry and misaligned interests, the agent will aim to maximize his self-interest. If the principal is unable to monitor the actions of the agent, then the agent has incentive to advance his interest, often to the detriment of the principal. In the privacy arena, this translates to companies adopting measures that financially benefit their bottom lines since their actions are essentially hidden to the users. Therefore, given a binary choice to adopt or not to adopt a policy that benefits the company but weakens user privacy, the company would adopt the policy.

We have shown that the relationship between users and companies exhibits the principal-agent problem in the arena of privacy policies. The high time cost of reading a single privacy policy does little to reduce the information asymmetry between users and companies. Since privacy policies keep getting longer over time, this information asymmetry is likely to widen. Furthermore, over time, companies are more likely to institute policy changes that restrict user privacy as opposed to strengthening it. This can be explained by the increase in the monetary value of personal data over time. As companies develop more powerful data collection capabilities, companies are incentivized to change policies to those that enable them to more fully exploit personal data. However, such changes will necessarily be detrimental to user privacy since data is transformed to revenue by selling it to third party entities for advertising purposes.

Our logistic regression model shows a significant relationship between a decline in revenue growth and a company's decision to update its privacy policy; this strongly suggests that in their agency, companies seek to maximize their utility at the expense of user privacy. When revenue growth stagnates, privacy policy changes may act as signals that the company is seeking to pursue more aggressive revenue accumulating tactics. This should be true for publicly traded firms and companies that are contemplating their IPOs; such signals could encourage investor interest and hence benefit the company. On the other hand, companies may also publish new versions of privacy policies in order to assure users that their data is well protected, especially when the frequency of data breaches increases. Since more users mean more potential revenue [21] companies adding empowering changes to their policies are also motivated by the desire to maximize revenue.

Our third finding is that the user privacy situation is complicated by the fact that companies also act as agents for third-parties, affiliates (through sourcing data for them) and law enforcement in an environment where the user has little to no bargaining power. As a result, in order to satisfy the demands from their other principals, some user utility is sacrificed.

Although various resolution tactics for the principal-agent problem in Internet privacy have been proposed and implemented, particularly the watchdog mechanism, the company-user relationship is still heavily skewed towards companies. One clear source of the imbalance is the information advantage held by the companies – by increasing the cost of reading the privacy policy, the company effectively renders the information hidden to the user. We recommend that companies follow the examples of Google and Twitter and publish the actual changes between versions of their privacy policies. Such an action would reduce the cost of information search from the user's point of view.

Furthermore, the watchdog system is not without fault; websites certified by TRUSTe, the most
prominent trust authority, are twice as likely to be untrustworthy compared to uncertified sites [12]. As exemplified by Dropbox’s change of its preferred watchdog from TRUSTe to JAMS in 2014 (section 4.4), it is difficult to establish the effectiveness of the watchdog monitoring system since a company can unilaterally switch to a watchdog that serves its purposes better. This suggests that the existing monitoring system is not effective in resolving the information privacy principal-agent problem.

Another way to tip the scales more equitably between the company and the user is through adoption of the opt-in, rather than the opt-out mechanism. Humans exhibit behavioral inertia and are likely to accept the default privacy settings [19]. Users that actually desire to share information with the company, perhaps in exchange for a more personalized experience, can do so consciously if the default setting is to opt-in. For example, the default setting for the Bing browser is for third party advertising sites to have access to user data, unless they explicitly opt-out of sharing. On the other hand, Google grants users the ability to opt-in if sensitive user information is to be shared. The company can also signal seriousness in protecting user privacy through providing incentives such as discounts and store credits for users who opt-in to programs that collect sensitive user information. Such an approach to privacy expands control to the user and can help resolve the principal-agent problem.

Lastly, we have formulated a set of propositions as groundwork towards building a theory of why companies change privacy policies. The propositions reflect that privacy policies, at a minimum, still perform their intended role i.e. describing the environment that shelters user data. However, changes to privacy policies reveal motivations such as maximizing revenue generation that may be detrimental to user privacy. We believe that if tested on a large scale, our propositions will shed light on company motivations for altering privacy policies.

6.2. Limitations and Future Work

This study has some limitations: first, we analyzed changes from a handful of technology companies selected because of their prominence. Even by that criterion, the list of companies analyzed was not exhaustive. Future work could expand the company input set and employ quantitative tools to validate the scalability of our findings. Second, we only analyzed privacy policies of technology firms; however, some of the most visited websites are situated in other industries, for example the New York Times and various bank websites. Including other industries in our analysis will paint a more holistic picture of privacy on the Internet. Lastly, this study would benefit from a more comprehensive search for causality.

7. Conclusion

Although many and varied, few studies on Internet privacy have focused on longitudinal trends. We have shown that over time, updates to privacy policies are more likely to be restrictive as opposed to being beneficial to user privacy. We have also demonstrated a significant link between a decline in a company’s revenue over successive quarters and the decision to change privacy policies. Therefore, company concerns over revenue growth are closely associated with actions that restrict customer privacy. As such, companies are likely to sacrifice customer privacy in pursuit of increased revenue and hence they do not always act in the best interests of customers.

8. References


