IS THERE A COST TO PRIVACY BREACHES? AN EVENT STUDY[1]

Security and Assurance

Alessandro Acquisti
Carnegie Mellon University
acquisti@andrew.cmu.edu

Allan Friedman
Harvard University
allan_friedman@ksgphd.harvard.edu

Rahul Telang
Carnegie Mellon University
rtelang@ andrew.cmu.edu

Abstract

While the literature on information security economics has begun to investigate the stock market impact of security breaches and vulnerability announcements, little more than anecdotal evidence exists on the effects of privacy breaches. In this paper we present the first comprehensive analysis of the impact of a company's privacy incidents on its market value. We compile a broad data set of instances of exposure of personal information due to failures of some security mechanism (hacking, stolen or lost equipment, poor data handling processes, and others) and we present the results of various empirical analyses, including event study analysis. We show that there exists a negative and statistically significant impact of data breaches on a company's market value on the announcement day for the breach. The cumulative effect increases in magnitudes over the day following the breach announcement, but then decreases and loses statistical significance. We also present regression analyses that aim at disentangling the effects of a number of factors on abnormal stock returns due to reported breaches. Finally, we comment on the differences between the impact of the security breaches already described in the literature, and the privacy breaches described here.

Keywords: Privacy, Information Security, Economics, Finance, Event Studies

[1] This research was supported by CMU Berkman Faculty Development Fund and by CMU CyLab. The authors would like to thank Steve Frank, Jutta Williams, and Franklin Ho for their research assistantship, and participants to workshops at ETRICS and WEIS for useful feedback.

Introduction

A common motivation for organizations to invest in information security is to safeguard their confidential data as well as their customers' personal information. Over the past few years, privacy incidents have been announced frequently enough to question whether organizations have the necessary incentives to safeguard consumer information. This paper moves towards an understanding of these incentives by measuring market reaction to privacy breaches. Several recent studies (Campbell et al. [2003], Hovava and D'Arcy [2003], Cavusoglu, Mishra, and Raghunathan. [2002], Kannan, Rees, and Sridhar [2004], and Telang and Wattal [2006]) have explored the impact that announcements of various *security* breaches have on a company's market value. Such announcements have often–but not always–a significant and negative impact. Much less explored is the link between a company's stock price and incidents that affect the *privacy* of its customers, employees, or partners. These incidents may be related to security vulnerabilities in a company's server, but extend well beyond the events already studied in that literature: they may include the intentional but illegal sale of consumer data; the loss of equipment containing sensitive consumers' or employees' information; or other instances in which companies were found in flagrant intrusions or violation of other parties' private data. Little is known about the more general consequences of such incidents on other measures of a company's performance - such as sales, revenues, or profitability.

Currently, there is only scattered evidence about the price companies pay for their privacy debacles. In a few notable but rare instances the offending company was subject to public outrage and hard to quantify reputation losses (consider, for instance, the Amazon "price discrimination" experiment - see Acquisti and Varian [2005]; or the RealNetworks case - see Acquisti [2004]). In the case of Choicepoint, after involuntary allowing criminals to access over 163,000 consumer credit reports, the company was forced to pay a $15 million in penalties (against $143 million 2005 earnings; Choicepoint [2006]). In other cases, companies have been punished harshly by the stock market: ChoicePoint's own stock tumbled in a few weeks after the incident from $46.01 to $37.64; while Internet advertising company DoubleClick lost 20% of its market value in March 2000 after the storm generated by privacy concerns associated with its acquisition of Abacus Direct, an offline consumer data company specialized in purchasing habits (Sakalosky [2002]).

Such extreme market consequences however are not common and often short-lived. A year after the incident, ChoicePoint stock was trading at pre-breach levels. Data collection, processing, and retention continue to grow as business activities. Although recent surveys report a negative attitude among consumers against firms that have exposed their information (Ponemon [2005]), there is little field evidence that the professed retaliation against offending companies has actually taken place.

The complex chain of consumers' data manipulation further complicates attempts to understand the consequences of privacy incidents. Some companies that have compromised consumers' data do not interact directly with those consumers; they are intermediaries or infomediaries (Hagel and Rayport [1997]), shielded from immediate consumer reaction. In addition, as news of privacy invasions and data breaches become more and more common, no related noticeable spike has yet been detected among identity theft and frauds (see Cate [2005]). Claims to the FTC about identity theft have held steady from 2004 to 2005. Moreover, the daily bulletin of data exposure may have started generating audience fatigue and reduced attention. Finally, announcements of privacy breaches in the media are often scattered and vague, often exhibiting a progression of articles and wires that slowly clarify the nature and scope of the breach.

This paper identifies a set of privacy incidents that stem from various forms of security, policy, or business failures, and discusses their possible consequences. We focus our empirical analysis on a specific type of violation - data breaches - and on available stock market data, although we have also gathered and analyzed data related to privacy breaches in organizations that are not traded (from government entities to universities). Specifically, we perform cross-tabulations, event study analyses, and regression analyses on the stock prices of public companies involved in data breaches. We show that there exists a negative and statistically significant impact of data breaches on a company's market value on the announcement day for the breach. However, we also show that such impact is lower than those observed in the literature for security vulnerabilities (from viruses, to hacking, to exploits discovered in applications). This dichotomy implies a different market perception, and possibly a discounting of the consequences of incidents that directly compromise consumers' data compared to those that affect consumers or corporations through viruses, denial-of-service attacks, and software vulnerabilities.

The rest of the paper is organized as follows. In Section 2, ,we discuss the relationship between privacy and security, as well as the possible consequences of a privacy incident for a firm's profitability and performance. In Section 3 we lay out the hypotheses that drive our research. In Section 4, we focus on a particular form of privacy incident, consisting of breaches of consumers' privacy - be it stolen, hacked, illegally sold, or unscrupulously handled data. In that section we presents the results of various empirical analyses, including an event study approach. In Section 5 , we analyze our initial results and present some basic regression models. In Section 6 we discuss current paths of analysis.

Theory of Privacy incidents

As the amount of personal data stored and processed by organizations increases, so too does the complexity of the information systems required for its safe-keeping. Such data trend is correlated with an increase in the number of privacy incidents. We broadly define a privacy incident as an event involving misuse of individuals' personal information. This misuse can consist of illegal sale, or usage, or lack of protection. It can be criminal, commercial, or ultimately innocuous. It can be intentional or unintentional. It can involve customers', partners', or employees' data.

Privacy incidents add a unique dimension to discussions of privacy. As noted in countless articles, attempts to precisely define the concept of "privacy" have only given mixed results – the concept and its definition often remain ambiguous, changing with the perspective of the observer. Introna and Pouloudi (1999) subsume many definitions under the idea of "freedom from judgment of others," but data disclosure does not always create such a potential. The unifying feature of privacy incidents is the violation of certain expectations about how data will be handled. Moor (1997) unifies several privacy theories into a concept of "control/restricted access" where an individual has an expectation of being able to control the flow of personal information, and restrict access where appropriate. Similarly, Smith, Milberg, and Burke (1996) show that unauthorized access and secondary use are two of the four primary factors in the concern for information privacy, a finding that remains robust across time (Stewart and Segars [2002]). Proper treatment of consumer information is part of an "implied social contract" with the customer (Milne and Gordon [1993]). A promise of fair information practice can override strong consumer preferences against sharing information (Culnan and Artmstrong [1999]) so a violation of that promise is a breach of the privacy conceptualization of "control/restricted access."

Violations of the above implie social contracts and expected information practices can take many forms. In 1999, RealNetworks admitted that it monitored the listening habits of its users, despite failing to mention this in its privacy policy. This incident was a violation of the understanding users had about what the company termed legitimate use: since it was not in the policy, their monitoring was a form of misuse. When DoubleClick (an internet marketing firm that tracks users across websites) moved to purchase Abacus Direct (an offline purchasing data company), many objected; the point of contention was the linking of online and offline identities. Amazon's brief attempt to offer differential prices based on tracked consumer data was also met with strong opposition; here the problem was an application of data for price discrimination, which many people inherently find objectionable (Kahneman, Knetsch, and Thaler [1986]). Choicepoint's misuse was blatant, as the company authenticated untrustworthy parties to access credit reports. Failure to protect data can be less severe, such as the recent spate of lost backup tapes by Time-Warner and others, but still evinces a significant gap between the expectations of data protection and the execution by the controlling party.

Each of these privacy incidents has different cause and can have a different set of effects. The common theme is the misuse of personal information. A subset of privacy incidents can be attributed to a failure of information security. Here, a privacy incident is "security motivated" when one of many mechanisms designed to protect personal information fails. This failure could be technical (a direct attack by malicious actors), managerial (failure to patch a known vulnerability), organizational (incomplete protections), or human (data left on a stolen laptop). In the examples above, DoubleClick and Amazon took active steps to misuse information which they saw as legitimate, so these incidents would not be examples of security failure; Choicepoint and Time-Warner had a failure of protection, where the security of personal information was inadequate or incomplete. Note that not all the incidents are the result of an intentional attacks by malicious entities on data or systems - in fact, those (which are the focus of the current information security economics literature and existing event study approaches) tend to be the minority.

Consequences of incidents

Later in this paper we examine a particular consequence of a privacy incident: the effect of announcement of data breaches on stock value. However, each incident can have numerous consequences, with complex repercussions. To date, there exists little solid research on the effects of privacy incidents, although anecdotal evidence can be combined with basic surveys to highlight some ramifications. Of course, every incident will be different, and will affect the firm and the consumer accordingly.

Even if the consequences of a violation are significant for a firm, the cost to consumers may not be fully considered by the firm as part of its risk. In this case, a privacy incident is a negative externality that natural incentives do not correct. Moreover, a lack of complete information about how companies protect personal information provides insufficient signals for consumers to correct this imbalance themselves.

Consumers

The predominant harm for consumers following a breach is the risk of impersonation, fraud, or identity theft. A full exploration of this topic is outside the scope of this paper. One often-cited Federal Trade Commission (FTC) survey puts the number of victims in 2003 near 10 million and the total economic losses in the billions of dollars (Federal Trade Commission [2003]). However, as Cate (2005) points out, these figures include a wide range of activities and may not be accurately sampled or reported. The Identity Theft Resource Center (2005) follows a sample of self-reported victims and reports very high costs in terms of time and money. Apart from the issues of sample bias, this survey claims that less than 50% of the perpetrators were unknown to the victims, indicating that corporate privacy incidents may not play the largest role in this type of fraud. Nonetheless, even when no actual fraud directly follows from a breach, the victim may be at an increased risk of phishing (Reilly [2006]), a form of spam attack intended to gain the victim's personal data. In addition, the standard precautions that experts advise consumers to take to avoid or following a privacy incident (Privacy Rights Clearinghouse [2006]) impose nontrivial costs in time and effort - and do not guarantee protection or reparation.

Consumers suffer less tangible harms as well. Perceived privacy risk can be as important as real privacy risks, and demand commensurate protection (Raab and Bennett [1998]), so even the fear of privacy harms can be counted as a negative consequence of the loss of control and access restriction discussed above. Expectations matter, and the consumer suffers when they are violated. In his taxonomy of privacy, Solove (2006) notes the legal and ethical distinction between mere data disclosure and a breach of confidentiality; the latter marks a violation of trust that entitles the victim to some recompense.

Firms

While responsibility for a privacy incident can rest with any data-controlling entity (see below), the analysis in this paper focuses largely on for-profit firms (other institutions may have a similar subset of consequences). The most immediate punishment for a privacy incident is a direct penalty from a governing institution. In the US, the Federal Trade Commission has jurisdiction in this area. The FTC can fine a firm responsible for breaches, or recommend expensive process overhauls to prevent future incidents. Choicepoint was forced to pay a $15 million fine (Federal Trade Commission [2006]). Often, however, the FTC does not fine the responsible firm (such as in the case of Petco, following a 2003 breach - Federal Trade Commission [2004]). Settlements usually do not require the company to admit responsibility, saving it from further liability. Yet, liability is another consequence of privacy incidents that can become significant. After a 2003 incident, BJ's Inc. did not face fines, but several civil suits from affected banks, for a total of $13 million. BJ's Inc itself sued IBM. In this case and others, the privacy incident affected client or partner firms and could, in the worst case, terminate the relationship. Finally, notification of affected consumers and accompanying recovery assistance such as a hotline represents non trivial expenses.

Beyond the immediate costs, a privacy incident can create long term indirect consequences. Privacy has been identified as a principle component of trust (Camp [2003]) and vice versa (Ponemon [2003]). An incident can damage a customer or partner relationship built on trust. Rhee and Haunschild (2006) show that when company admits a mistake (e.g.. through a product recall), public trust can be lost, and that reputation loss can have measurable ramifications for the company's market share. Ponemon (2005) notes that consumers retain a negative impression of responsible firms and will alter their consumption patterns, although it must be verified whether survey data of this sort can capture actual behavior. Still, a firm might face higher insurance premia for liability after a breach, and future business partners might be less inclined to trust the firm.

Quantifying Privacy Consequences

Identifying consequences of a privacy incident is difficult enough, but quantifying these consequences is remarkably complex. Gellman (2002) argues that studies of privacy losses tend to understate the effects of privacy losses, while Lenard and Rubin (2005) and Cate (2005) argue that many estimates are overstated. There is no, in short, consensus about estimated costs, or even the best method (Svensson [2003]). None of the above-mentioned calculations includes the costs of preventing a privacy incident: the net effect of a breach remains an open question.

A common simplifying tool employed throughout the literature is to use stock price as a proxy for the consequences of various events for a company. A potential decline in stock price is itself an adverse consequence of a privacy incident. Notwithstanding the known limitations of analyses based on such valuations of stock market prices (from the debated efficiencies of stock market, to the actual consequences of market oscillation that may be only temporary), it is of interest to contrast how the market reacts to privacy breaches, to its reaction to security breaches. The earlier examples of DoubleClick and ChoicePoint offer anecdotal evidence that stock price can reflect news of privacy incidents: in the two studies that are closest to ours, Garg et al. (2003) use eight incidents to show that a breach of credit card information has a large negative effect on stock price, while a theft of customer information has a negligible impact; while Campbell et al. (2003) find that 11 instances of attacks on confidential data (of which 9 involved consumers' data) have a statistically significant impact on a company's evaluation.

Consequences of security vs. privacy incidents

While there is little work on economic costs of privacy breaches, more work has been done on economics costs of security breaches. Prior event study analyses on information security have focused on the change in market value of firms whose systems are breached (Cavusoglu et al. [2002] and Kannan et al. [2004]). These studies show that announcements of a security breach negatively impact the CAR (Cumulative Abnormal Return) of firms whose information systems have been breached. Campbell et al. (2003), quoted above, conduct a similar event study and find that only the impact of confidentiality related security breaches is negative and significant, while the impact of non-confidentiality related security breaches is not significantly different from zero. Hovava and D'Arcy (2003) show similar results by finding that Denial of Service (DoS) type attacks are not associated with any significant loss in value for firms. Telang and Wattal (2006) show that software vendors suffer when vulnerabilities in their products are published in major newspapers.

The motivation and results of this paper differ from the above work in three ways. First, the focus is explicitly centered on personal information, rather than the security product or corporate networks. Second, security and privacy are only partially overlapping sets: in this paper we use focus on privacy breaches, extending outside conventional security breaches, to include organizational and human failures – and therefore also policy and business violations. Finally, we compile the largest dataset yet used to examine the economic effects of security risks, spanning many industrial sectors and including responsible third parties. This allows us to build a clearer model of the cost of privacy breaches.

Hypotheses

The previous literature on event studies and security breaches has shown that certain types of breaches negatively affect a company's stock market valuation. Since privacy breaches can be associated with liability, fines, reputation, and other costs, we expect a similar impact:

H1: *A company suffers a loss in market value whenever a privacy breach is announced*

It is also possible to forecast, following Telang and Wattal (2006)'s findings, that in the case of privacy breaches the source of the announcement has an influence on the magnitude of the effect:

H2*:* *The magnitude of the negative CAR (Cumulative Abnormal Return) will be larger whenever a privacy breach is reported in national media rather than local or industry outlets*

Because a company's privacy costs may also depend on the number of individuals whose information has been compromised, we expect a monotonic relation between the number of affected parties and the stock market impact:

H3*:* *The magnitude of the negative CAR will be increasing with the number of individuals affected by the privacy breach*

In addition to the above hypothesis, there are other aspects of the relation between privacy breaches and stock market valuations that we intend to study. In particular:

1. Whether a "privacy fatigue" is emerging, as the number of breaches grows and consumers' attention possibly diminishes (in ongoing regression, we test whether breaches that took place after the *ChoicePoint* debacle are met by similar or different severity in the stock market)
2. Whether privacy breaches determine similar or different reactions in the market place as security breaches.

Privacy breaches: Data and analysis

In order to study somewhat homogenous and comparable data, we focused our attention on data breaches for publicly traded companies, defined as instances in which the data of consumers, employers, or third parties associated with a company traded on a public market was exposed through:[2]

- Bad security practices
- Hacker attacks
- Insider attacks
- Computer or data thefts
- Lost data or equipment

[2] We have, however, also gathered and analyzed data about privacy breaches affecting other, non-publicly traded organizations - such as government entities and universities. In the Appendix we present additional tables about privacy breaches for such organizations. That additional data, while of interest, cannot of course be used in an event study approach.

- Other (e.g. illegal sale or handling of individual data)

The exposure in question may have been caused by insider attacks, hacking, stolen or lost equipment, or by voluntary but illegal sale of data. To compile as complete a list of privacy incidents as possible, we mined and search for announcements of privacy breach events in news databases like *Lexis-Nexis* and *ProQuest*, as well as online compilations (*Choicepoint*, Inc [2006], Wall [2006], and mailing lists and dedicated blogs such as Adam Shostack and Chris Walsh's `http://www.emergentchaos.com` and `http://attrition.org`).

We considered each event case by case, searching the several outlets discussed above. For each event, we had to determine the precise date announcement in order to look for a market response. Given the complex nature of the news cycle and how information propagates, it is not always immediately clear when a breach was announced. This is further complicated by the pattern of disclosure, where new developments sometimes modify (almost always increase) the scope of the initial announcement. When several companies were involved in a data breach, each company was treated as a separate event. Several events had to be discarded because the sequence of publication dates or news propagation did not allow us to be confident about a certain event's data. Since the model assumes that the event is the primary change in stock value during the event period, we abandoned a few event data when unrelated major announcements (such as earning projections or mergers) for the company of interest fell in a window of days close to the event day.

After this cleaning, we were left with 79 breach events. Each event in our data set was coded for a number of data type available in most announcements. The mean number of individuals affected by a breach was around 1 million, while the median breach disclosed 95,000 subjects' data. Tables 1 to 4 provide some aggregate statistics about our dataset.

Table 1: Distribution of privacy breaches in publicly traded companies (NYSE and NASDAQ) by year.

Year	Number of incidents
2000	4
2001	4
2002	1
2003	3
2004	3
2005	51
2006 (until March 1st)	13

Table 2: Distribution of privacy breaches in publicly traded companies by subject exposed.

Type of subject	Number
Customers	49
Employees	10
Third party data	18
Other/undetermined	2

Table 1 shows the distribution of reported occurrences across time. Note the sharp increase in 2005: a California state law mandating the report of data breaches affecting residents went into effect on January 1, 2005 (California SB1386/AB-700). The sudden increase of announced incidents suggests that the rate public disclosure rose, rather than the number of actual breaches. We also distinguish, in Table 2, among the subjects of the data disclosed -

whether it referred to a firm's customers, employees, or if the company had no direct relationship with the data subject. The majority of breaches involve customer data, which could magnify reputation consequences.

We further separate the announcements by cause of data breach. Of course, these classifications are necessarily artificial, since each event has its own detail, and we must use the information provided in the initial announcement. When a trusted party or an outside attacker is specifically mentioned, we label it as an insider or hacker attack, respectively. If the data was lost on stolen hardware, we note that as a separate attack, since the attacker could have valued the equipment rather than the data. When no evidence of theft is available, we simply note that the data is "lost". The category of "bad security practices" comprises a wide range of internal mistakes that allowed data to be accessed, whether faxing account data to a wrong number or failing to secure a sensitive web page adequately. Most of the breaches are due to mishandling of security practices or data, or inappropriate physical defense against thieves (see Table 3). By themselves hackers' attacks do not appear to be the main source of data problems in our sample.

Gleaning the exact amount of personal data leaked in each breach from breach announcements is also often hard, although most announcements offer some details. Social security numbers are the data more often compromised in these breaches, followed by credit card information (Table 4).

Table 3: Distribution of privacy breaches in publicly traded companies (NYSE and NASDAQ) by type.

Type of breach	**Number**
Bad security practices	24
Hacker	9
Insider attack	8
Computer or hardware theft	18
Lost data	12
Other/undetermined	8

Table 4: Distribution of privacy breaches in publicly traded companies (NYSE and NASDAQ) by type of information exposed.

Type of information	**Number**
SSN	41
Credit card	18
Complete credit record	9
Other personal information	10
Other/undetermined	1

Event studies

To measure the economic costs of privacy breaches to a firm, we adopted an event study methodology. Our methodology follows closely from prior event study analysis. The implicit assumption in this methodology is that the financial markets respond to news that affect a security's value, so change in stock price is a good proxy for the impact of a given event. Campbell et al. (2003) present an useful summary of the event study analysis highlighting the history as well as the commonly followed methodologies. Event study methodologies are well accepted for studying the implications of public announcements on stock prices. Hendricks and Singhal (1996) study the impact of quality award winning announcements on the market value of firms and observe positive abnormal returns generated by winning a quality award. They further note that awards given by independent organizations and announcements by small firms are more likely to have a significant impact on the firms' market value.

In an event study approach, the stock price of the firm is explained by the event after controlling for trends and volatility. In particular, if a firm suffers from privacy breach then it may incur financial losses (fines and penalty, loss of reputation) which should reflect in its stock price. Thus, stock prices on the days surrounding the event can capture the impact of that event and measure the economic cost of such privacy incidents.

Models

An event study assumes that returns on a stock are significantly impacted by an event of interest (in our case, the event of interest is the vulnerability disclosure announcement). The period of interest for which we observe the event is known as the event window. The smallest event window is one day (day of the announcement or day 0).[3] In practice, the event window is often expanded to include two days (day 0 and day 1)[4] to capture the effect of price announcements made after the close of the markets on a particular day. Sometimes researchers include a day before the announcements to incorporate any information leaks about the event.

In our study we focus on a one-day event window (day 0). Hendricks and Singhal (1996) cite two reasons to use a one-day event period. First, a shorter event period permits a better estimation of the effects of information of stock prices since it reduces the possibility of other confounding factors not related to the announcement. Second, it also increases the power of the statistical tests. Abnormal returns are defined as the difference between the actual return of the stock over the event window minus the expected return of the stock over the event window. To compensate for a delay in the news cycle and to understand the long term effects, we also examine a short period following the event day.

The expected return on the stock is calculated in several ways. We focus on the market model, which assumes a stable linear relation between the market return and the return on the stock. We also verify our results using other methods, such as the market-adjusted method and the mean-adjusted method. The coefficients of the linear model are calculated by choosing a portion of the data as the estimation window. The estimation window closely precedes the event window. In our case, we use standard estimation window of 92 actual trading days (more than 120 solar days), from day -100 to day -8. There are three main methods followed in the event study methodology (Campell et al. [1997], Hendricks and Singhal [1996]) to estimate the abnormal returns.

The Market Model

In the market model, the abnormal returns are estimated as follows:

$$ARit = Rit - \alpha i - \beta iRmt \quad (1)$$

Rit for a stock is the percent change in the stock price at time *t*, $(Pit - Pit-1)/Pit-1$, where *i* denotes the event (*i*=1,2, …N), *m* denotes the market, and *t* denotes the day of the event (e.g., *t* = 0 denotes the day of the vulnerability announcement.). *ARit* denotes the abnormal return of event *i* at time *t*, *Rit* denotes the actual return and *Rmt* denotes the market return at time period *t*. The abnormal return is defined as the difference between the actual return and the normal return. This is the part of the actual return that cannot be explained by market movements and captures the effect of the event. Depending on whether the stock of interest was traded on NASDAQ or NYSE, we used the appropriate market index. We use ordinary least squares regression to estimate the coefficients for the above regression.

[3] If an announcement is made on a day when the markets are closed, we consider the next day when the markets open as day 0.

[4] Day 1 is the day after the announcement.

The quantities of interest for general analysis are the day-specific average returns and aggregate returns over time. The mean abnormal return is the mean across all observations on day *t* of the event. The cumulative abnormal return for the event CAR is defined as the sum of the abnormal returns over the event window.

The Market Adjusted Model

In this model, the event window returns are compared to an expected return of the market only over the event period, so the abnormal returns are given as:

ARit = Rit − Rmt (2)

where the terms have the usual meaning as in the Market Model.

The Mean Adjusted Model

In this case, the returns are compared to the mean market return over the event period. Abnormal returns are now given as:

ARit = Rit − Ri (3)

where *Ri* is the mean return on the stock which made a vulnerability announcement during event *i*, over the duration of the estimation period (in our study, that means 92 days).

Test statistics

Brown and Warner (1985) have presented a comprehensive analysis of suitable test statistics for the abnormal mean return. Similarly to the security case, since data breaches may have been disclosed by more than one company on a given day, our statistics should allow for event day clustering. A *t* statistic proposed by Brown and Warner (1985) takes in to account event day clustering and cross-sectional dependence in the security specific excess returns (see also Telang and Wattal [2006]):

$$t = \frac{\overline{A}_t}{\sqrt{S^2_{\overline{A}}}} \qquad (4)$$

where $S^2_{\overline{A}} = \frac{1}{T-1}\left(\sum_{s=1}^{T}(\overline{A}_s - \overline{\overline{A}})\right)$ and *T* is the number of days in the estimation period and $\overline{\overline{A}} = \frac{1}{T}\left(\sum_{s=1}^{T}\overline{A}_s\right)$

The null hypothesis is that the abnormal returns are not significantly different from zero. Under the null hypothesis, the abnormal returns are independent and identically distributed and normal with a mean of zero and the variance given by the variance of abnormal returns over the estimation period.

Event study results

We present here the results of the event study on the 79 events met the independence conditions discussed above. We used an estimation window from *t* −100 to *t* −8. We used a forecast window from *t* −7 to *t* +10.

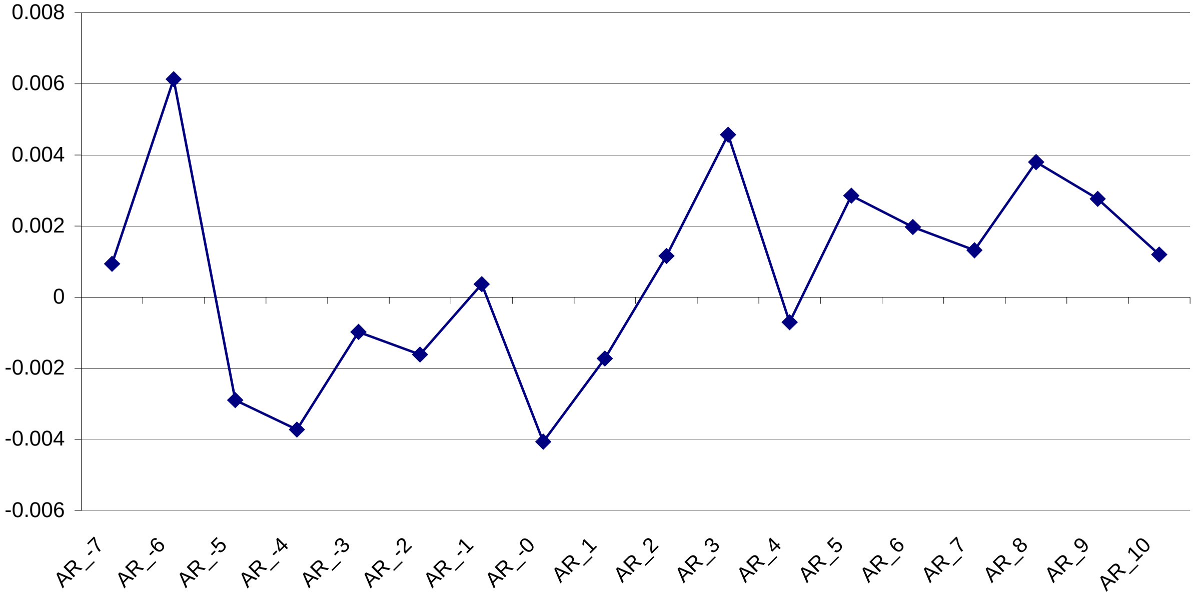

Figure 1: Mean abnormal returns from 7 days before the event through 10 days after

To strengthen our results, we tested all the three models discussed above (market model, market adjusted, and mean adjusted) and we calculated both mean and median values as well as percentages of negative abnormal returns on event day. We performed parametric and non parametric tests to evaluate our results.

Figure 1 summarizes the abnormal return values over the $t-7$ to $t+10$ window (CAR_-0, represents the event day - the day when the data breach was announced). The negative values on day 0 are the largest and most significant in the window considered. After that, the values start increasing and the abnormal return become positive on day 3.

Table 5 pins down specific values for the abnormal return on event day (which is identical to the abnormal returns on that day). The different models give comparable results highlighting the mild (around half percent) negative impact of breaches announcements on the companies' stock prices.

Figure 2 presents the actual *cumulative* results, or the accumulated difference between the breached companies and the projected market returns, starting at *t* -5. From this figure, the sudden drop in returns from the event is more evident. Starting on day 2, however, positive abnormal returns increase the CAR. The values after that point become positive but no longer significant (see also Table 6).

Table 5: Cumulative abnormal returns on event day (t values in parentheses)

AR t=0	Market Model	Market Adjusted Model	Mean Model
Mean Abnormal Return	0.41 (1.97)	0.52 (2.45)	0.35 (1.62)
Median Abnormal Return	0.22	0.26	0.15
Percentage below zero	58%	58%	59%

Table 6: Cumulative abnormal returns over different periods. Double star represents significance at the 0.05 level.

Day	-1	0	0 to 1	0 to 2	0 to 5	0 to 10
CAR	0.03	-0.41**	-0.58**	-0.46	0.21	1.3

Discussion

Data breaches seem to have a moderate but statistically significant negative impact on a firm's stock value. The day prior to the event the mean abnormal return is positive. On the even day and the day after, the mean abnormal return is negative and reaches cumulatively a value close to -0.6%. These values are robust under different model specifications. The mean single day loss of 0.4% represents a large expected drop in market value, as the average firm loses 0.4% of its value from day *t* -1.

The differences between the mean and median abnormal return values deserves further attention. On the event day, both the statistics are negative and significant. However, the mean values tend to be systematically higher than the median ones. This suggests the presence of strong outliers that are driving some of the negative performance after the announcement of a breach. However, the percentage of negative abnormal return for a firm object of such announcements is well above 50% (from 58% to 59% depending on the model adopted).

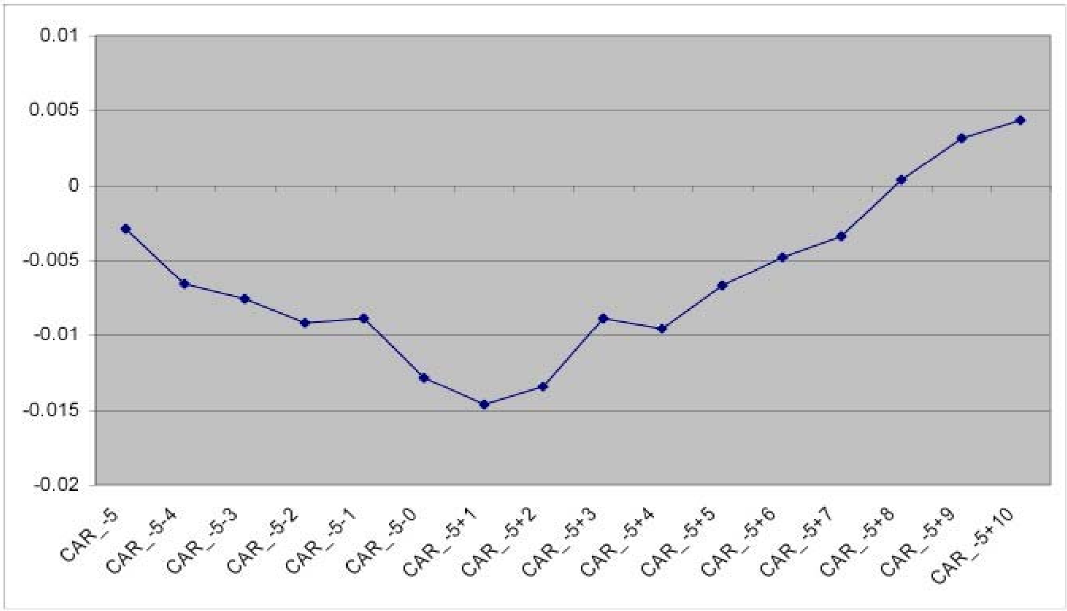

Figure 2: Cumulative Abnormal Returns from 5 days before through 10 days after the event

In theory, this should be a net loss of over $140 million. However, looking at each company's abnormal returns and applying that difference to each firms' market value produce an estimate of only $9,953,968. Clearly, some firms with large market capitalizations do not suffer as strong adverse effects from an announcement. This somewhat surprising finding suggests that the penalty for a privacy breach may be an anticipation of an absolute value of consequences, which would hurt smaller companies more. A preliminary exploration does not strongly support this hypothesis, since AR is barely correlated with capitalization value (Pearson product-moment correlation coefficient=0.08), even when compared to the correlation between AR and stock price (cc=0.18). Future work to explore this question is discussed below.

An aspect of this study that deserves more attention is how the impact of announcements changes over time. Anecdotally, companies like ChoicePoint and DoubleClick have suffered well after the first initial day following the announcement, and some of their worst losses have appeared even a few days later. This is possibly related to the fact that data breaches tend to be more confusing - their magnitude, implications, and nature are more complicated and often not immediately tractable - than 'pure' security breaches. For this reason it is possible both that the media keeps infusing new information into the market place, but also that the marketplace takes more time to evaluate and incorporate those news into market prices than in presence of security breaches.

To better understand why the market reacts to privacy breaches, we examine a range of factors. Looking at each attribute alone, in Table 7, it appears that retail firms suffer more and to a greater degree. This could be explained from the reputation effect. Note that switching costs are often lower from one retailer to another than other consumer firms (i.e. banks), so the privacy reputation of a consumer goods merchant could drive away more business than a financial institution, even if the stakes are higher. If the breach announcement had evidence of a malicious actor deliberately trying to access data, the negative effect of the breach is higher, while if a third-party company (such as a data processor) can be blamed for the breach, the market appears to react more favorably. Both of these results conform to what one might expect. The number of victims looks to be correlated with a negative market reaction, but only for very large data breaches.

Table 7: Cumulative abnormal returns (Day 0) by industry, type of breach, subject, responsibility, cause, and numbers affected

Breach Classification		# Events	Abnormal Return	% Negative
Industry	Retail	14	-0.01570	71.43%
	Other	24	-0.00206	66.67%
	Finance	26	0.00048	53.85%
	Data processor	14	-0.00509	28.57%
Data Misuse	Attack evidence	33	-0.00870	57.58%
	No attack evidence	41	-0.00104	65.85%
Data Subject	Third party	17	-0.00597	64.71%
	Employee	10	0.00186	60.00%
	Customer	51	-0.00459	54.90%
Responsibility	Third party responsible	17	0.00158	23.53%
Breach Cause	Laptop / tape	29	-0.00286	65.52%
	Hacker / insider	17	-0.01098	47.06%
	Bad security practices	24	-0.00112	62.50%
# Affected	Less than 100,000	31	0.00202	54.84%
	100,000-500,000	22	-0.00458	54.55%
	More than 500,000	9	-0.02656	77.78%

Further exploration of Figure 3 highlights what appears to be a small size effect. The data itself is quite noisy, however, and the linear relationship is a poor fit. Still, this finding offers some limited support for our third hypothesis, which argues that the CAR should be a function of privacy costs, which in turn should be a function of the number of individuals affected. One could also anticipate a time effect. Either investors become inured to the problem of data breaches, or concern is compounded as companies fail to learn from each others mistakes. Privacy theory does not offer much guidance between these two alternatives, and the data is equally ambiguous, as the passage of time appears to have very little effect on abnormal stock returns (Figure 4). The absence of a time effect allows us to confidently use the full data sample in the following analysis.

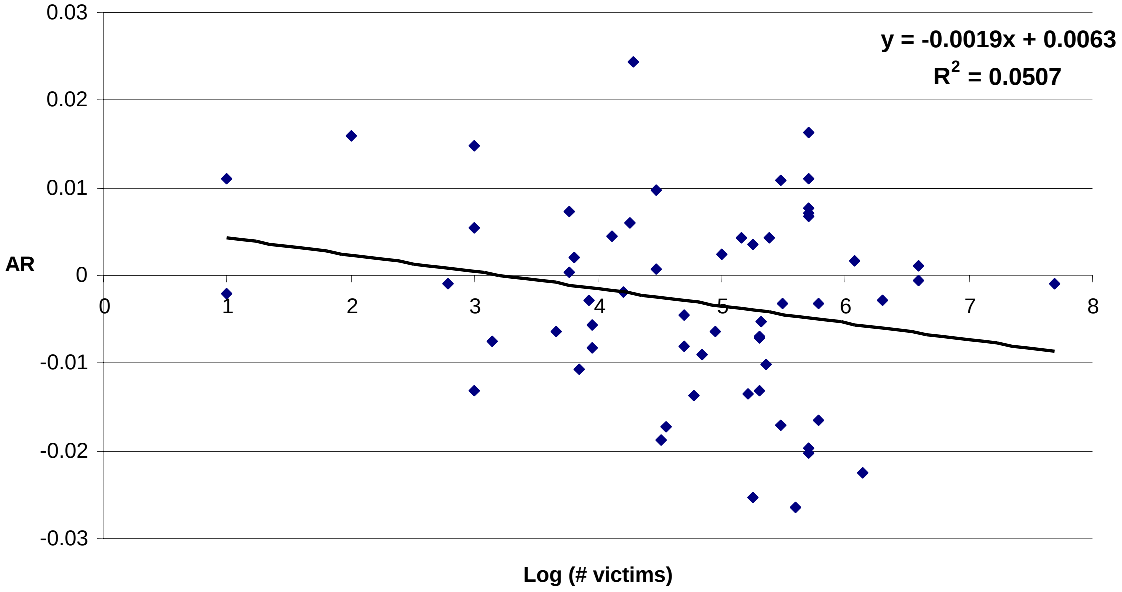

Figure 3: Larger breaches appear to have a negative effect on Day 0 Abnormal Returns

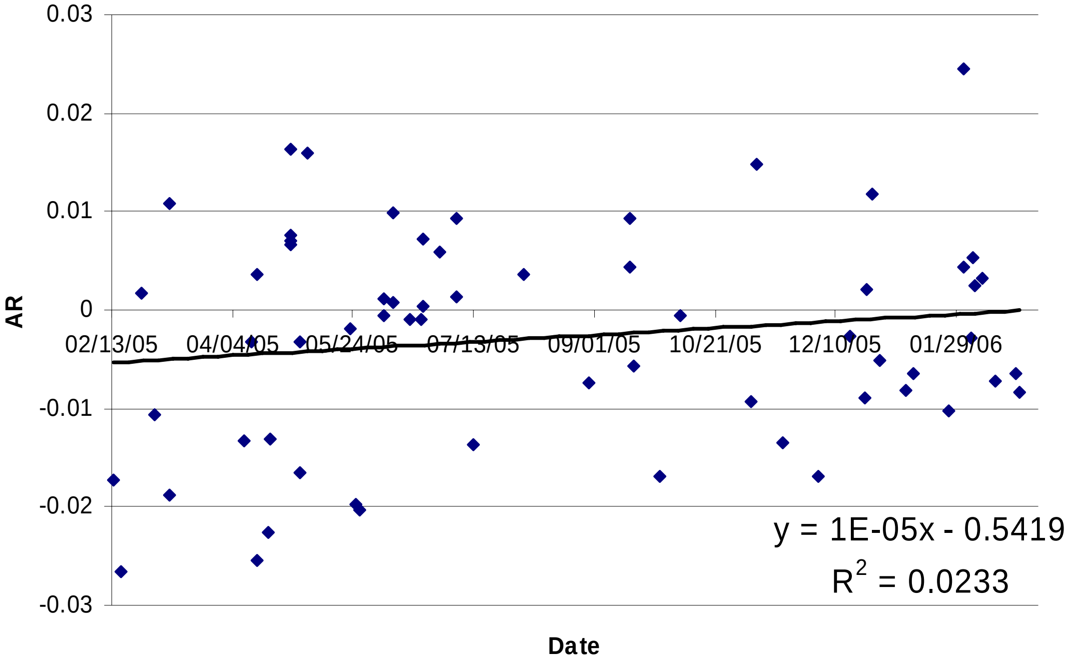

Figure 4: There does not appear to be a time effect

Regression analysis

The data presented above suggests a high level of variation across observed privacy incidents, and what could explain it. Of course, simple tabulated data could fail to account for a variety of competing factors. We use a simple regression model to examine whether any feature of the breaches can predict the market reaction. Both the nature of the firm and the details of the breach could affect how investors viewed the event. Table 8 summarizes a regression model predicting that abnormal returns on day 0 are a function of the number of victims of a breach, whether the breach was announced in a major paper or news wire, the industry of the firm involved, whether the breach was announced clearly as part of a malicious attempt to gain access to data, the subject of the data, whether a third party (outside the company) was responsible for the security failure and measure of the company's size, in market capitalization. This same model is also used to test factors predicting the change in total value of each stock.

Table 8: Regression analysis

	Abnormal returns		Change in total market value	
	Linear coeff	Std Err	Linear Coeff	Std Err
Intercept	-7.690	4.112	866.11	925.89
Breach with > 100,000 subjects	* -1.203	0.670	-143.75	150.77
Reported in major paper or wire service	0.052	0.760	269.63	171.14
Firm Sector: retail	-0.458	0.992	-201.03	223.39
Firm Sector: finance	0.191	0.884	78.22	198.95
Firm Sector: data processor	0.139	1.130	-160.63	254.36
Breach as part of an active attack to get data	-0.388	0.680	-39.83	153.20
Data subject: Customer	-0.123	1.073	-58.91	241.53
Data subject: Employee	-0.154	1.260	14.87	283.77
Responsibility for breach attributed to 3rd party	0.335	0.819	71.18	184.36
Company size: log(total stock value)	** 0.803	0.398	-96.06	89.62

*Abnormal returns measured in percentage points, change in market capitalization measured in $ millions. Firm sector reference: "other"; data subject reference: "third party". * p > .1, ** p > .05*

When all firm and breach attributes are included, very few of the variables have any significant predictive power. There is scant support for Hypothesis 2, since the channel of initial report has no effect on the size of abnormal return or change in market value. Our third hypothesis, however, looks promising. Controlling for other factors, a breach of more than 100,000 subjects will reduce the return on stock price by 1.2% (p = .077). On the other hand, the larger a company is, the less impact a data breach will have, with each order of magnitude in total market capitalization increasing the expected return on the event to increase by .8% (p=047). None of the other variables have significant predictive power, nor do any of the parameters in the market value model. Neither industry, data subject, nor breach responsibility is indicative of how the market will react. Interestingly, there was also no market response about whether the announcement indicated that the breach was part of an active attack to get the data. Several methods of measuring a time effect were also added to the model without significant effect. As we continue to add to the dataset and identify outliers or inappropriate data, we hope to further refine this these regression models to better understand why the market responds differently to similar privacy incidents.

Comparison with other event studies

Any model is dependent on the underlying assumptions, and the conclusions of this paper are dependent on a theory of firm valuation through market reaction. These methods have been used to study a variety of phenomena related information technology and consumer trust. We present a summary these studies' findings in comparison to this paper in Table 9. The impact of privacy breaches of stock market valuations is somewhat less than that observed for security vulnerabilities (from viruses, to hacking, to exploits discovered in applications). This difference may signal a different market perception, and possibly a discounting of the consequences of incidents that directly compromise consumers' data compared to those that affect consumers or corporations through viruses, denial-of-service attacks, and software vulnerabilities, and calls for further investigation. Still, our findings are within an order of magnitude of recent studies on the negative effects of vulnerability disclosure (Telang and Wattal [2004]), and product recall announcements, as well as the positive effects of a quality award (Hendricks and Singhal [1996]). This latter comparison might validate the underlying theory that a privacy breach might indicate a shift in reputation and trust.

Table 9: Summary of related event studies

Classification of Event Study	Authors	Time Period	# Events	Window	CAR
Impact of a Data Breaches on Firms	Acquisti, A, Friedman, A, and Telang R	2000-2005	79	0 → 1	-.58 %
Impact of Vulnerability Disclosures	Telang R and S Wattal (2004)	1999-2004	146	0 → 1	-0.65%
Impact of Security Breaches on Firms	Campbell K, Gordon LA, Loeb MP and L Zhou (2003) *(Personal data accessed)*	1995-2000	11	-1 → 1	-5.4%
	Campbell K, Gordon LA, Loeb MP and L Zhou (2003) *(all security breaches)*	1995-2000	43	-1 → 1	-1.9%*
	Cavusoglu H, Mishra B and S Raghunathan (2004)	1998-2000	66	0→ 1	-2.1%
	Hovav A and J D'Arcy (2003)	1998-2002	23	-1 → 1	Not Significant
Impact of Auto Recall Announcements	Jarrell G and S Peltzman (1985)	1967-1981	116	-1 → 1	-0.81%
	Davidson WL III and DL Worrell (1992)	1968-1987	133	-1 → 1	-.68%
Impact of IT Investment Announcements	Chatterjee D, Richardson VJ and RW Zmud (2001)	1987-1998	96	-1 → 1	1.16%
	Im KS, Dow KE and V Grover (2001)	1981-1996	238	-1 → 0	Not Significant
	Dos Santos BL, Peffers K and DC Mauer (1993)	1981-1988	97	-1 → 0	Not Significant
Impact of Winning a Quality Award	Hendricks KB and Singhal VR (1997)	1985-1991	0	0	0.59%

*Not significant at the 10% level

Conclusions

We have presented the first comprehensive analysis of the impact of a company's privacy incidents on its market value.

We have gathered and combined privacy incidents and breach data from different sources for the 1999-2006 (until March) period. Our event study shows that there exists an impact for privacy violations. This impact is statistically significant and negative, although it is short-lived. The difference in our mean and median results suggests that a number of outlying firms are driving significant portion of the negative results. One possible explanation is that larger firms not only are more visible, but their trust reputation, built over time, can be more significantly affected by negative reports about their privacy practices.

In order to further study this conjecture, our current research is extending the regression analysis to examine the determinants of firm-specific stock returns and to understand and contrast the impact of "pure" security breaches compared to privacy ones. Specifically, we are testing the relation between abnormal returns, breach announcements, and a number of factors including channel of announcement, type of company, attack, and data breached, and severity of the breach. Obviously, we are limited by a small sample, but we are monitoring current breaches as they develop to extend our dataset. The longitudinal nature of this data enables us to detect time trends, to determine whether regular announcements generate privacy fatigue or if a standard develops to punish those who fail to safeguard personal information after several years of disclosure.

We also plan to further study implications of privacy violations that go beyond stock market impacts and that include non-breach related invasions - such as intrusive policies or faulty services in the realm of personal information. This will allow us to further analyze the additional data presented in the Appendix.

References

A. Acquisti. Privacy and security of personal information: Economic incentives and technological solutions. In Jean Camp and Stephen Lewis, editors, The Economics of Information Security, 2004.

A. Acquisti and Hal R. Varian. Conditioning prices on purchase history. Marketing Science, 24(3):1–15, 2005.

S. Brown and J. Warner. Measuring security price performance. Journal of Financial Economics, 14:3–31, 1980

S. Brown and J. Warner. Using daily stock returns: The case of event studies. Journal of Financial Economics, 1985

California SB1386/AB-700. California Civil Code, Sections 1798.29 and 1798.82-1798.84. Passed in 2002, in effect Jan 1 2005.

L. J. Camp. Design for trust. In Rino Falcone, editor, Trust, Reputation and Security: Theories and Practice. Springer-Verlang, 2003.

K. Campbell, L. Gordon, M. Loeb, and L. Zhou. The economic cost of publicly announced information security breaches: Empirical evidence from the stock market. Journal of Computer Security, 11(3):431–448, 2003.

J. Campell, W. Andrew, and A. MacKinlay. The Econometrics of Financial Markets. Princeton University Press, 1997.

F. H. Cate. Information security breaches and the threat to consumers. The Center for Information Policy Leadership at Hunton & Williams LLP, 2005.

H. Cavusoglu, B. Mishra, and S. Raghunathan. The effect of internet security breach announcements on market value of breached firms and internet security developers. In International Journal of Electronic Commerce, volume 9, 2002.

D.Chatterjee, V. J. Richardson, and R.W. Zmud. Examining the Shareholder Wealth Effects of Announcements of Newly Created CIO Positions, *MIS Quarterly,* 25(1), 43-70, 2001.

Choicepoint, Inc. 2005 disclosures of U.S. data incidents. URL http://www.privacyatchoicepoint.com/common/pdfs/Data_ Disclosures_2005.pdf, 2006.

Choicepoint, Inc. United States Securities and Exchange Commission Form 10-K, Commission file number 1-13069, 2006.

M. J. Culnan and P. K. Armstrong . Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation *Organization Science*, 1999

W. L. Davidson III and D. L. Worrell. The Effect of Product Recall Announcements on Shareholder Wealth *Strategic Management Journal,* 13(6), 467-473, 1992.

B. L. Dos Santos, K. Peffers, and D. Mauer. The Impact of Information Technology on the Market Value of the Firm, *Information Systems Research,* 4 (March), 1-23, 1993.

Federal Trade Commission. Identity theft survey report. URL http: //www.ftc.gov/os/2003/09/synovatereport.pdf, 2003.

Federal Trade Commission. Consent order with Petco animal supplies, *FTC File 032-3221*. URL http://www.ftc.gov/os/caselist/0323221/ 041108agree0323221.pdf, 2004

Federal Trade Commission. Stipulated final judgement and order in US v Choicepoint, *FTC File 052-3069*. URL http://www.ftc.gov/os/caselist/ choicepoint/choicepoint.htm, 2006.

A. Garg, J. Curtis, and H. Halper. Quantifying the financial impact of IT security breaches. *Information Management and Computer Security*, 11(2):74–83, 2003.

R. Gellman. Privacy, consumers, and costs - How the lack of privacy costs consumers and why business studies of privacy costs are biased and incomplete. http://www.epic.org/reports/dmfprivacy.html, 2002.

J. Hagel and J. Rayport. The coming battle for customer information. Harvard Business Review, 75(1), 1997.

K. Hendricks and V. Singhal. Quality awards and the market value of the firm: An empirical investigation. Management Science, 42(3):415–436, 1996.

K. Hendricks and V. Singhal. Delays in new product introductions and the market value of the firm: The consequences of being late to the market. Management Science, 43(4):422–436, 1997.

A. Hovava and J. D'Arcy. The impact of denial-of-service attack announcements of the market value of firms. Risk Management and Insurance Review, 6(2):97–121, 2003.

Identity Theft Resource Center. Identity theft: The aftermath. URL http://www.idtheftcenter.org/aftermath2004.pdf, 2004.

K. S. Im, K. E. Dow, and V. Grover. Research Report: A Reexamination of IT Investment and the Market Value of the Firm – An Event Study Methodology, *Information Systems Research,* 12(1), 103-117, 2001.

L. Introna and A. Pouloudi. Privacy in the information age: Stakeholders, interests and values. *Journal of Business Ethics* , 22(1):27-38, 1999.

G. Jarrell and S Peltzman. The Impact of Product Recalls on the Wealth of Sellers, *The Journal of Political Economy,* 93(3), 512-536, 1985.

D. Kahneman, J. Knetsch, and R. Thaler. Fairness as a constraint on profit-seeking entitlements in the market. American Economic Review, 76, 1986.

K. Kannan, J. Rees, and S. Sridhar. Reexamining the impact of information security breach announcements on firm performance. Working paper, Carnegie Mellon University, 2004.

T. M. Lenard and P. H. Rubin. An economic analysis of notification requirements for data security breaches. Progress on Point, The Progress and Freedom Foundation, 2005.

G. R. Milne and M. E. Gordon. Direct mail privacy-efficiency trade-offs within an implied social contract framework. *J. Public Policy Marketing* 12 2 206–15, 1993.

J. H. Moor. Toward a Theory of Privacy in the Information Age, *Computers and Society*, September 1997, pp. 27-32, 1997.

L. Ponemon. The 2003 Privacy trust survey. Ponemon Institute, LLC and The CIO Institute of Carnegie Mellon University, 2003.

L. Ponemon. What does a data breach cost companies? Ponemon Institute, LLC, 2005.

Privacy Rights Clearinghouse. Facts sheet 17b: How to deal with a security breach. URL http://www.privacyrights.org/fs/ fs17b¬SecurityBreach.htm, 2006.

C. D. Raab and C. J.. Bennett. *The Distribution of Privacy Risks: Who Needs Protection*? The Information Society 14(4):253-262, 1998.

T. Reilly. Press release: AG warns consumers: Do not give your private financial information to anyone calling or emailing pretending to be investigating the Boston Globe security breach. URL http: //www.ago.state.ma.us/sp.cfm?pageid=986&id=1602, 2006.

Rhee, M. and P R. Haunschild. The Liability of Good Reputation: A Study of Product Recalls in the U.S. Automobile Industry . Organization Science 17:1 p. 101, 2006

M. Sakalosky. Doubleclick's double edge. ClickZExperts, September 3, 2002. URL http://www.clickz.com/experts/crm/analyze_data/ article.php/1455141.

H. J. Smith, S.J. Milberg, and S. Burke. Information privacy: Measuring individuals' concerns about organizational practices J. MIS Quarterly. Minneapolis, Vol. 20, Iss. 2; p. 167, 1996.

D. J. Solove. "A Taxonomy of Privacy" . University of Pennsylvania Law Review, Vol. 154, No. 3, p. 477, January 2006

K. A. Stewart and A. H. Segars, An empirical examination of the concern for information privacy. instrument, Information Systems Research 13(1), 2002.

A. Svensson. Analysing information systems security. Department of Informatics, School of Economics and Management, Lund University, 2003.

R. Telang and S. Wattal. Impact of software vulnerability announcements on the market value of software vendors - An empirical investigation. Working Paper, Carnegie Mellon University, 2006.

B. Wall. Bill wall's list of computer hacker incidents. URL http: //www.geocities.com/SiliconValley/Lab/7378/hacker.htm, 2006.

Appendix

The analysis presented in the body of this paper focuses exclusively on the privacy breaches by publicly traded firms. However, the underlying motivation is to examine the effects on organizations of all privacy breaches and, ultimately, all privacy incidents. We argue in this paper that privacy incidents comprise a broad set of situations enabling the misuse of personal information. Since many public, non-profit and privately held organizations and institutions also have the responsibility to safeguard personal information, a complete exploration of the problem should include these stakeholders. Moreover, data misuse spans beyond accidental breaches to include intentional misuse through active sharing or unauthorized collection. The collection of breach announcements used in this paper is a subset of the larger collection of privacy incidents we have compiled for further analysis. The tables below summarize this broader dataset.

Table 1A: Distribution of all privacy incidents by year (up to March 2006)

Year	Number of Incidents
1999	1
2000	6
2001	4
2002	3
2003	3
2004	8
2005	151
2006	30

Table 2A: Distribution of all privacy incidents by type of incident

Source of privacy incident	Number of Incidents
Computer or tape theft	32
Data Mishandling	39
Illegal Data Selling	6
Hacks and exploits	80
Insider attacks	9
Intrusive Business Practices	13
Password Compromises	3
Software flaws with potential for data exposure	24

Table 3A: Distribution of all privacy incidents by economic sector

Sector	Number of Incidents
Commerce	28
Finance	31
Information	10
IT	51
Medical	11
Government	17
Educational	47
Other	9