Do Data Breach Disclosure Laws Reduce Identity Theft?

Sasha Romanosky, Rahul Telang, Alessandro Acquisti
Heinz School of Public Policy and Management
Carnegie Mellon University
{sromanos, rtelang, acquisti} @andrew.cmu.edu

ABSTRACT

In the United States, identity theft resulted in corporate and consumer losses of $56 billion dollars in 2005, with up to 30 percent of known identity thefts caused by corporate data breaches. Many states have responded by adopting "data breach disclosure laws" that require firms to notify consumers if their personal information has been lost or stolen. While the laws are expected to reduce identity theft, their full effects have yet to be empirically measured. We used panel data from the US Federal Trade Commission to estimate the impact of data breach disclosure laws on identity theft over the years 2002 to 2009. We find that adoption of data breach disclosure laws reduce identity theft caused by data breaches by, on average, 5.7 percent.

Keywords
Data breach disclosure, security breach notification, economics of information security, identity theft

INTRODUCTION

Data breaches occur when personally identifiable information such as names, social security numbers, and credit card numbers are accidentally lost or maliciously stolen. These breaches can result in hundreds of thousands (sometimes millions) of compromised records, and lead to identity theft and related crimes (Givens, 2000):[1] in the United States, identity theft resulted in corporate and consumer losses of around $56 billion dollars[2] in 2005 (Javelin Research, 2006). In an effort to reduce these crimes, many states have responded by adopting data breach disclosure (or "security breach notification") laws, requiring firms to notify individuals when their personal information has been compromised. However, to date, no empirical analysis has investigated the effectiveness of such legislative initiatives in reducing identity theft. In this paper, we use panel data gathered from the Federal Trade Commission (FTC) and other sources over a seven year time period to empirically examine this effect.

---

[1] Criminals use stolen personal information in many ways. For example they can incur fraudulent charges on existing accounts, or apply for new utilities (phone, electrical, television, Internet) and financial accounts (such as credit cards, mortgages, and loans).

[2] This value was calculated as the estimated number of identity theft victims in 2005 multiplied by the average amount stolen per victim: 8.9M victims * $6,383 stolen/victim = $56.6B. (Actual amount lost per consumer was $422 on average.)

In response to the recent publicity surrounding data breaches, much time and effort have been devoted to finding solutions to prevent breaches and help consumers avoid, or mitigate, any resulting harm. At least four US congressional hearings have convened to discuss how data breach disclosure laws may reduce identity theft (US Congress, 2005a, 2005b, 2005c, 2005d). In a testimony to the U.S. Senate, the chairman of the FTC testified, "The Commission believes that if a security breach creates a significant risk of identity theft or other related harm, affected consumers should be notified. Prompt notification to consumers in these cases can help them mitigate the damage caused by identity theft" (FTC, 2005, p10). Moreover, the US Government Accountability Office (GAO) has stated that "notification to the individuals affected …has clear benefits, allowing people the opportunity to take steps to protect themselves against the dangers of identity theft" (GAO, 2006). The US Security and Exchange Commission has proposed new security and privacy guidelines, including "requirements for notices to individuals […] intended to give investors information that would help them protect themselves against identity theft" (SEC, 2008). In addition, other countries have argued in favor of breach disclosures. For example, the UK Science and Technology Committee has claimed that "data security breach notification law would be among the most important advances that the United Kingdom could make in promoting personal internet security" (Science and Technology Committee, 2007).

As of December 31, 2009, 45 US states had adopted such data breach disclosure laws, as shown in Figure 1.[3]

[Insert Figure 1: Adoption of breach notification laws from 2002-2009]

many of these laws specifically addressed identity theft prevention. For example, California's law was intended "to help consumers protect their financial security by requiring that state agencies and businesses […] to quickly disclose to consumers any breach of the security of the system, if the information disclosed could be used to commit identity theft" (SB1386). Further, representative Simitian (CA) writes that the purpose of the California data breach disclosure bill was to, "provide assurance that when consumers are at risk because of an unauthorized acquisition of personal information, the consumer will know that he is vulnerable, and will thus be equipped to protect himself physically and/or financially" and moreover, to "provide an incentive to those responsible for public and privacy databases to improve their security" (Simitian, 2009, 1015). The Hawaiian law is even more direct: "[t]he purpose of this Act is to alleviate the growing plague of identity theft by requiring businesses and government agencies that maintain records containing resident individuals' personal information to notify an individual whenever the individual's personal information has been compromised by unauthorized disclosure" (SB2290). Montana's breach law is "an act adopting and revising laws to implement individual privacy and to prevent identity theft" (SB732).

While details of the legislations vary across states, their central themes are consistent. Specifically, they require notification a) in a timely manner, b) if personally identifiable information has either been lost, or is likely to be acquired, by an unauthorized person, c) and is reasonably considered to compromise an individual's personal information. The laws generally define a breach as the "unauthorized acquisition of computerized data that compromises the security, confidentiality, or integrity of personal information maintained by the person or business" (Hutchins, 2007). Personal

[3] For the purpose of this paper, we are not considering federal sectoral legislation such as the Gramm-Leach-Bliley Act (GLBA) as their effects are not identifiable with our econometric model.

information generally refers to an individual's name in addition to another piece of identifiable information such as driver's license, passport, or credit card number.

One differentiator among the state laws is the trigger, or threshold, by which notification must be made. At least twenty five state laws require notification when the personal information is reasonably assumed to have been acquired by an unauthorized party whereas other state laws require notification only if it is reasonable to believe the information will cause harm to consumers. The consequences of not complying include retribution by the state attorney general or a civil right of action (the ability for affected consumers to bring a lawsuit). Many states do not specify a maximum civil penalty. However, the Arizona and Arkansas laws allow a civil penalty not exceeding $10,000, whereas the limit is $25,000 in Connecticut and Idaho, and $500,000 in Florida. An important characteristic of these laws is that the residency of the consumer, rather than the location of the breach, drives disclosure. Therefore, a firm that incurs a data breach must comply with the state laws of each of their affected consumers. For example, if a retail firm based in Oregon suffers a breach that includes personal information of residents from California, the firm must notify those Californian residents. Of course, not all breaches affect consumers in every state. Breaches in state government agencies (e.g. DMVs), community colleges, schools and hospitals usually only affect residents of a single state. Even breaches by national firms (e.g. chain stores) may only compromise individuals (often employees) of a single state.

The rationales for these laws are contained within two phrases: "*Sunlight as a disinfectant*,"[4] and "*Right to know*." First, notification can "transform [private] information about firm practices into publicly-known information as well as alter practices within the firm" (Schwartz & Janger, 2007). Hence, by highlighting a firm's poor security measures, legislators hope to create an incentive for all firms (even those that have not been breached) to improve the protection of their data, thereby "disinfecting" themselves of shoddy security practices (Ranger, 2007). This, in turn, is expected to reduce the probability of breaches and resulting harm (including identity theft). In other words, since it has been shown that consumers lose confidence in firms who suffer breaches (Ponemon, 2005), proponents believe that the laws will force firms to internalize more of the cost of a breach through notification letters, customer support call centers, and mitigating actions such as marketing campaigns and free credit monitoring.

Second, this form of light-handed paternalism often represents a preferred approach to legislative enforcement compared with a "command and control" regime (Magat & Viscusi, 1992). Consumers feel that they have the right to be informed when firms *use* or *abuse* their information. Having being notified of a breach of their personal information, consumers could then make informed decisions and take appropriate actions to prevent or mitigate the impact of identity theft. For example, to lessen their risks, consumers who have been notified of a breach may alert their bank, their credit card merchant, the FTC, or law enforcement; they may close unused financial accounts; they may place a credit freeze or fraud alert on their credit report.[5] Notifications can also enable law enforcement, researchers, and policy makers to better understand which firms and business sectors are better (or worse) at protecting consumer and employee data. However, it may only be through legislation that firms acquire sufficient incentive to actually improve their practices to reduce the likelihood of future breaches and repair consumer confidence.

---

[4] This phrase is originally attributed to Justice Louis Brandeis, 1933, http://www.brandeis.edu/investigate/sunlight/, accessed 11/08/07.

[5] A fraud alert informs potential creditors that a consumer may have been a victim of identity theft. The creditor must then take additional measures to verify the identity of the consumer. A credit freeze prevents a creditor from checking a consumer's credit report, or opening new accounts.

Arguments in favor of such disclosure laws are compelling. However, it is unclear whether a data breach disclosure regime does, in fact, increase social welfare. While it may improve a firm's security practices, and help some consumers mitigate the risk of identity theft, on balance, it may only serve to burden them. First, firms must comply with multiple, disparate, and perhaps conflicting state laws. Next, if the probability of suffering identity theft following a data breach is, in fact, very low, then costs incurred as a result of the laws would be unwarranted: firms would be forced to notify consumers without benefit, and consumers would be needlessly freezing and "thawing" their credit reports (FTC, 2005, p10; GAO, 2007). Cate (2009) posits that, "if we think breaches really cause harm, then notices are too little. We're just shifting the burden to somebody else. If breaches do not cause harm ... then notices are an unnecessary cost." Cate (2005) also argues the consumers may become desensitized if they receive too many notices. Moreover, Lenard and Rubin (2005, 2006) argue that these laws are unnecessary for a number of reasons; they may impede e-commerce and stifle technological development by discouraging firms to innovate using consumers' personal information (or stop collecting it altogether); that the externality is not so grave, because most of the cost of identity theft and fraud is already born by the firms (businesses, banks, credit card issuers, merchants);[6] that firms may use self-regulated notifications as a market differentiator, and if notifications are sufficiently valued by the consumer, the market will react accordingly.

In summary, these arguments present a stimulating debate as to whether data breach disclosure laws can reduce identity theft -- an impact that, to our knowledge, no one has attempted to empirically measure. The purpose of this manuscript is to investigate the effectiveness of data breach disclosure laws of reducing identity theft. Because of the compelling controversy surrounding the connection between adoption of these laws and identity theft, we hope to offer a relevant and timely contribution to the policy debate. Using panel data on identity theft gathered from the Federal Trade Commission and other sources from 2002 to 2009, we use state and year fixed effect regression analysis to empirically estimate the impact of data breach disclosure laws on the frequency of identity thefts due to breaches. We found that adoption of these disclosure laws reduce identity theft, on average, by 5.7 percent.

The next section in this paper provides background literature related to information economics and disclosure policies. The paper then presents the conceptual model behind our empirical approach, and the results of the data analysis. A discussion of the policy implications of our findings completes the manuscript.

RELATED WORK

Our paper draws from the literature on disclosure policies, the literature on information security economics, and the literature in environmental and crime policy.

Information Economics and Disclosure Policies

Many researchers have studied the effects of disclosure on market outcomes. For instance, Jin & Leslie (2003) investigated health information disclosure in the restaurant industry, and found that disclosing the hygiene quality of a restaurant increases health inspection scores and lowers the occurrence of food borne diseases. Moreover, disclosure becomes a credible signal to consumers, who respond by demanding cleaner restaurants. Mathios (2000) examined the effects of mandatory disclosure of food nutrition labels on salad dressing sales in a chain of New York grocery stores. He found that producers of salad dressings with the highest fat content suffer a greater decline in market share once

[6] As estimated by Javelin Research in 2003 (90.5 percent), 2005 (89.6 percent) and 2006 (93.7 percent)

forced to disclose nutrition information, relative to less fatty dressings. These studies provide some evidence of how information disclosure policies can affect firm behavior and improve market outcomes. A lengthy discussion of many disclosure policies related to healthcare, auto safety, public education and more can be found in Fung et al. (2007).

A number of studies have examined the financial impacts to firms that disclose a privacy or security breach. Campbell et al. (2003) find a significant and negative effect on the stock price of the breached company, but only for data breaches caused by "unauthorized access of confidential information." Cavusoglu et al. (2004) find that the disclosure of a security breach results in the loss of $2.1 of a firm's market valuation. Telang & Wattal (2007) find that software vendors' stock price suffers when vulnerability information in their products is announced. Acquisti et al. (2006) use an event study to investigate the impact on stock market prices for firms that incur a privacy breach, and find a negative and significant, but short-lived, reduction of 0.6 percent on the day when the breach is disclosed. Ko & Dorantes (2006) study the four financial quarters following a security breach, and find that, while breached firms' overall performance was lower (relative to firms that incurred no breach), their sales increased significantly (again, relative to firms that incurred no breach). Despite absence of more conclusive empirical findings on the effect of publicly disclosed data breaches, firms nevertheless appear to be making security and operational investments in the wake of disclosure laws (Samuelson Law, 2007).

Environmental Disclosure

There are strong precedents of disclosure legislation, and analysis of these policies in the United States. For example, the Food and Drug Administration (FDA) and the Environmental Protection Agency (EPA) have implemented regulations that require firms to notify consumers in case of an adverse impact of their products and services. A specific example of EPA efforts is the Toxic Release Inventory (TRI) program developed by the Environmental Protection Community Right to Know Act (EPCRA). Firms polluting above a certain threshold must report the quantity and type to the Environmental Protection Agency. In an analysis of firm outcomes, for example, Hamilton (1995) discovered that the first disclosure reduced firm stock price by 0.3 percent, or a loss of $4.1M in stock value on the day of the disclosure. Konar & Cohen (1997) found that after announcement of TRI, firms with the largest negative (abnormal) stock returns reduced their emissions the most.

Criminal Deterrence Policies

Estimating the effect of a law on criminal outcomes is a familiar research question in criminology. To be clear, though, the policies identified in this manuscript (data breach disclosure laws) are meant to influence safety and protection measures by potential victims of a crime (identity theft), rather than influence criminal behavior. However, for the purpose of this study, we gained valuable methodological insight from the approaches of criminology and policy evaluation. For example criminologists frequently seek to measure the deterrent effect of law, both in general (Blumstein et al., 1978; Levitt, 1995; Nagin, 1978; Robinson & Darley, 2003) and, specifically, regarding the effects of laws on crime (Lott & Mustard 1997; Black & Nagin 1998; Donohue & Ayres, 2003) and capital punishment (Mocan & Gittings, 2003; Wolfers & Donohue, 2006).

IDENTITY THEFT AND BREACH DISCLOSURES: A CONCEPTUAL MODEL

Impact of Data Breach Disclosure Laws

Consider two, separate effects of these disclosure laws: consumer precautions, and firm investment. The primary function of data breach disclosure laws is to force firms to notify consumers when their personal information has been lost or stolen. Ideally, as more consumers are notified, more will take precautions to reduce the risk and the costs of becoming a victim of identity theft. For example, they could notify their financial institutions to block transactions and cancel accounts, or apply credit freezes and fraud alerts.

On the other hand, firms have to incur significant tangible costs to notify consumers after a data breach. These may include replacement costs of credit cards (through bank negotiations) to proving free credit counseling, setting up 1-800 numbers, etc. Moreover, firms are also likely to suffer significant intangible costs in the form of negative reputation effects. Acquisti et al. (2007) show that repeated disclosure of data breaches and newspaper headlines could lead to a significant reputation impact and loss in share price. Ponemon (2005) suggests that consumers lose confidence in firms who suffer breaches. Hence, a secondary effect of the laws would be to induce firms to invest and improve their security controls - in order to avoid a data breach, and avert the direct and indirect costs associated with its notification. These investments may reduce the number of data breaches, thereby reducing the number of identity theft crimes due to breaches. And so both these effects (consumers taking precautions, and firm investing in better security) should reduce the incidence of identity theft.

In order to qualify the overall effect of data breach disclosure laws, however, it is important to note that identity theft originates from different sources. Disclosure laws should reduce identity thefts for situations where consumer data is controlled by *firms*, but should not directly reduce identity thefts due to – say - stolen mail or garbage. In a randomized phone survey conducted by Synovate (FTC, 2007b), 12 percent of identity thefts occurred as a result of interaction with firms, while another 56 percent of victims did not know the cause. In another survey conducted by Javelin Research (2006, p7), 35 percent of identity fraud was a result of information that was within the control of businesses.[7] And in 2007, researchers at the Center for Identity Management and Information Protection (CIMIP) at Utica College studied 517 identity theft cases from the US Secret Service (Gordon et al., 2007). For cases where the source could be determined (about half of the total 517), 26.5 percent originated from firms.

The Impact of Disclosure Laws on Breaches

Naturally, even prior to their impact on identity theft, a first-order effect of the laws should be to reduce the number of *breaches*. However, we note that the number of reported breaches is endogenously affected by the law as well: only after the laws are passed, firms are forced to disclose, and their breaches enter the statistics. This creates the impression that breaches increase following the enactment of the laws. In other words, analyzing the number of breaches directly is unlikely to provide useful results. As shown in Figure 2, it is apparent that the number of reported breaches has increased, as expected.

[Insert Figure 2: Data breaches from 2002-2009]

[7] The categories controlled by the firm are: Taken by a corrupt business employee: 15 percent, Some other way: 7 percent, Misuse of data from an in-store/onsite/mail/telephone transaction: 7 percent, Stolen from a company that handles your financial data: 6 percent.

Identity Theft Data

The most comprehensive public source for identity theft data are the consumer reports published by the FTC since 2002. The Identity Theft Act and Assumption Deterrence Act of 1998 led the FTC to establish the Identity Theft Data Clearinghouse in November 1999 to collect identity theft complaints from victims.[8] Consumer Sentinel is the web portal by which annual identity theft reports are made available to the public, and where law enforcement can further mine the data.

For our analysis, we used consumer reported identity thefts collected from the FTC for each state from the years 2002 to 2009. Since only annual data are published, we invoked the Freedom of Information Act to request monthly data. In our analysis, we aggregated the monthly data to 6-month periods (2 per year) for the years 2002 to 2009 (producing 800 observations).

One of the advantages of this data source is the consistency of data collection methodologies across states (without which our estimations could be erroneous).[9] On the other hand, the data is self-reported by victims. This is a familiar issue for criminologists, who often rely on various forms of self-reported crime data (e.g., Uniform Crime Reports and National Crime Victimization Surveys). The frequent under-reporting of crimes is often referred to as the "dark figure" (Biderman & Reiss, 1967) and represents a potential source of error. However, not only is the FTC (to our knowledge) the only source for cross-sectional (cross-state) time series identity theft data, but – more importantly – trends in FTC time-series identity theft data are consistent with other surveys by the Bureau of Justice Statistics (Baum, 2006, 2007), Synovate (FTC, 2003, 2007b), and Javelin Research (2006, 2007).

Summary statistics for total annual reported identity thefts based on the data we obtained through the FOIA request are shown in Table 1.

[Insert Table 1: Identity theft reports, 2002-2009]

In 2009, for example, California had the highest reported number of identity thefts of over 42,000 while North Dakota had the lowest, at 192. For comparison, relative rates of other reported crimes such as murder, robbery, burglary and motor vehicle thefts are shown in Table 2.

[Insert Table 2: Reported offenses per 100,000 population, 2002-200]

Figure 3 shows reported identity theft rates increasing from 2002 until 2005, after which they decline slightly in 2006 and increase, then drop slightly in 2009. Prior to 2005, only California had adopted the law, while others followed in 2005 (n=8), 2006 (n=19), 2007 (n=8), 2008 (n=6) and 2009 (n=3).[10] Figure 3 shows the relative changes in reported identity theft rates for three groups: those that adopted in 2005, 2006 and the 5 states that, as of the end of 2009, had not adopted the law.[11] Trends for states that adopted between 2007 and 2009 show similar patterns and therefore we omit these for clarity.

---

[8] See http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=105_cong_public_laws&docid=publ318.105, accessed 11/02/09.

[9] For instance, underreporting would be problematic if the reporting patterns changed suddenly over time across states. If the reporting levels change uniformly across all states - which is likely the case with FTC data - these effects would be captured by our time dummies.

[10] States that adopted in 2005 were: Arkansas, Delaware, Florida, Georgia, North Dakota, Tennessee, Texas and Washington. States that adopted in 2006 were: Colorado, Connecticut, Idaho, Illinois, Indiana, Louisiana, Maine, Minnesota, Montana, Nebraska, Nevada, New Jersey, New York, North Carolina, Ohio, Oklahoma, Pennsylvania, Rhode Island, and Wisconsin. States that adopted in 2007 were: Arizona, Hawaii, Kansas, Michigan, New Hampshire, Utah, Vermont, Wyoming. States that adopted in 2008 were: Iowa, Maryland, Massachusetts, Oregon, Virginia, and West Virginia. States that adopted in 2009 were: Alaska, Missouri, and South Carolina.

[11] Alabama, Kentucky, Mississippi, New Mexico, and South Dakota.

[Insert Figure 3: Average identity theft rates from 2002-2009]

Reported identity theft rates for states that adopted the law in 2005 are slightly higher than others, while states that had not adopted any law (as of December 31, 2009) show the lowest overall average (we discuss the issue of potential endogeneity of the laws below). States that adopted in 2007 fall generally in between other groups. For comparison, we also include plots of identity theft rates and their changes centered around the year of adoption as shown in Figure 4.

[Insert Figure 4: Identity theft rates and Percent changes before/after law]

The left panel plots the identity theft rates and changes in identity theft rates (right panel) for three groups of states (those that adopted in 2005, 2006 and those that, as of 2009, had not adopted the data breach law). We include only these three groups for clarity and consistency with previous figures (only 2005 and 2006 provides 3 time periods before and adoption. Moreover, plots for states that adopted in 2007-2009 follow no observable pattern and therefore provide no additional insight). The x-axis represents the three time periods before adoption of the law (T-3, T-2 and T-1) and three time periods after adoption of the law (T=0, T+1 and T+2). For example, for states that adopted in 2006, T-1 represents data from 2005, while T=0 represents data from 2006. Data for states without the law have been centered around 2006.

First, to the extent we can draw inferences about these groups, t-tests of the difference of means reveals no statistical difference between the means of these groups. Next, the left panel suggests that identity theft rates are increasing before adoption of the laws for all groups but that even for those that adopted in 2006, rates continued to increase. Rates for states that did not adopt show a decline in period T+2, yet are still higher and show a gradual increase from T=0. Moreover, while identity theft rates for states without the law are lower during some periods, rates for all groups increased over time, with states that adopted in early (2005) showing the largest increase before and after.

These trends provide some initial insight into what may (or may not) be driving the changes in identity theft reporting. We scrutinized those changes using a fixed effect regression model, as described in the following sections.

Endogeneity

A practical concern with all empirical analyses that investigate the effect of a treatment (data breach disclosure laws) on an outcome (identity theft) is that of endogeneity. Specifically in this case, that adoption of the laws was not random, but instead driven by high levels of identity theft within each state. Below we first discuss key obstacles and drivers for state adoption of laws, then describe in detail the process by which California came to adopt the first breach disclosure law. Finally, we include a number of quantitative analyses that we believe suggest strong exogeneity.

Innovation and diffusion of state laws

It is a very old and common practice for state legislatures to observe and mimic another states' policies. Indeed, this practice of 'copycat' legislation refers to the 'diffusion' and 'innovation' of laws (modeled after the innovation-diffusion paradigm of technology). Walker (1969) claims that inertia and risk aversion are often obstacles to legislators writing new laws (the "innovation"), however, these issues quickly dissipate if the legislator can point to other states that have successfully adopted the law (the "diffusion"). As more and more states adopt the law, Walker claims that, "it may become recognized as

a legitimate state responsibility, something which all states ought to have. When this happens it becomes extremely difficult for state decision makers to resist even the weakest kinds of demands…once a program has gained the stamp of legitimacy, it has a momentum of its own" (Walker, 1969, 890). For example, Walker (1969) describes 88 separate state laws which were adopted across no less than 20 other states, prior to 1965, that cover industries from welfare, health, education, civil rights, labor, taxes, etc. Indeed, the practice extends as far back as the early 1800s in which new states joining the Union partially duplicated clauses of existing state constitutions (Tarr, 2000, 51).

The State Legislative Process

More specifically, the adoption of state laws is both a very random and structured process. A given legislature is presented with anywhere from hundreds to thousands of bills during in any one session, depending on its session length. For example, in an average 2 year session, around 5000 bills are introduced by state legislators. And once the bill is introduced (filed), it becomes part of the legislative machinery and therefore subject to the nuances, timelines, schedules, deadlines, together with all other bills. "Every state is unique in its method of legislative operations and in its lawmaking procedures. Individual states take pride in doing things their own way" (Neal, 2005).

First of all, new bills can only be introduced into the state legislature when it is in session and bills must be submitted before the filing deadline. Moreover, all bills must follow the same process: after being filed with the state legislature, bills are assigned to one or more committees (in each chamber) wherein they must endure multiple rounds of voting, amendments and public lobbying. Only if a bill passes a 3rd vote, will it be sent to the governor for signature. A further legislative reality, of course, is that some bills fail to acquire the necessary votes and die, requiring them to be re-introduced the following year.

To further complicate the process, once signed by the governor, states take a number of approaches when determining when new bills actually become effective (adopted). Generally, legislatures set the effective date to either the first calendar day of the legislative session (early January), or the start of the fiscal year (early July), whereas the California data breach bill simply became effective 6 months from the date of signing, in order to allow firms to prepare for its consequences.

Together, all of these structural processes provide strong variation (and therefore randomness) between states and their approach to filing, signing and enacting laws.

The California Data Breach Bill

As discussed, the legislative process is both highly structured and random: it is fraught with complexities, delays, political positioning, historical nuances and historical formalities. It creates both structure within a given state legislature but also variation across the 50 state legislatures.

Consider the case of the first data breach law in California co-introduced by representative Joe Simitian who, when comparing the ideal legislative process to reality, confessed that, "in truth, [the legislative process] is far more random, dramatic, and idiosyncratic than any flow chart could ever describe" (Simitian, 2009, 1009). For example, the California data breach bill became law because of, "a spelling error, an afterthought, an unrelated concern with digital signatures, a page three news story, rule of germaneness, the intellectual quirks of a lame-duck Senator, the personal experiences of 120 state legislators, and another bill altogether" (Simitian, 2009, 1009). As further justification for his motivations of writing and supporting the bill, Representative Simitian stated he wanted a bill that was well defined and very likely to succeed, "when you're a new state legislator, high prospect of passage is very important" (Simitian, 2009b).

Next, we address specific questions related to the adoption of these laws and how they vary across states.

Did all states pass the data breach law equally fast?

Let us consider three important dates of a bill: the date it is first filed by a state legislator (either house/assembly or senate representative), the data it is signed by the governor, and the date when it actually becomes effective. In Table 3, we show descriptive statistics regarding the time between filing and signature, signature and adoption, and the total time between filing and adoption. Note that we were unable to determine the filing date for 4 states (California, Colorado, Kansas, and New York), and all other dates were obtained from state legislatures.

[Insert Table 3: Delay between filing, signature and adoption]

First, notice the wide variation between filing and signature: less than a month for some states, while almost two years for other states. Moreover, some states took only two and a half months between filing and adoption, while another state took about 30 months. However, this only demonstrates the presence of variation in delays across states and does not address the possibility that states with higher rates of identity theft may systematically adopt laws more quickly, relative to states with low rates of identity theft. If we believe that adoption of data breach disclosure laws were, indeed, endogenous, then it should be true that those states with higher levels of identity theft would be much quicker at both signing and adopting them.

Do states with higher rates of identity theft adopt the laws more quickly?

Figure 5 illustrates the adoption durations for all states. The y-axis sorts states from highest to lowest identity theft rates (top to bottom). For example, Arizona, Nevada and Texas, had the highest rates of identity theft, while Iowa, Vermont, and North Dakota had the lowest.

[Insert Figure 5: Months to sign and adopt data breach laws

Again, to support the claim of endogeneity, we would expect to see that states with high rates of identity theft would be quicker at both signing and adopting the bill (identified as diamonds and circles, respectively). That is, we would expect to see data points generally contained within the oval region shown in the figure: that data points for states with high rates (at the top of the y-axis) would be positioned very close to the y-axis, while data points for states with low rates of identity theft (near the bottom) would be very far from the y-axis (to the right). Clearly, however, the data points are very scattered for states with both high and low rates of identity theft (supported by Table 1). Moreover, adoption of the bill does not appear to occur more quickly for states at the top, relative to states at the bottom.

Are states with higher rates of identity theft adopting the laws in earlier years?

However, we may still be concerned that states with higher rates of identity theft are systematically adopting the laws in earlier years compared with states of lower rates. Figure 6 shows the date of adoption for each state, again sorted by state with highest identity theft rates (top) to lowest (bottom) on the y-axis. The rightmost points represent those five states that by 2009 had yet to adopt a breach disclosure law.

[Insert Figure 6: Date of adoption]

For endogeneity to be a concern, we would expect to see that the states on the upper part of the figure would be first adopters (leaders), while the states with lower levels would take longer to adopt (followers). Again, there is no systematic indication that states with higher identity theft are adopting the law sooner than states with lower identity theft.

In further analysis, we compare the identity theft rates across groups of states that adopted in different years (and those that had not adopted). Whether we average the rates over all years for each group, or compare rates for each individual year, we find no statistical difference between any two groups. We also performed a cox proportional hazard model to estimate the average probability of a state adopting the law and again find no statistical evidence that identity theft affects the probability of adoption of these laws.

In sum, we find no systematic correlation between the identity theft and the speed at which a law is passed, nor do we find statistical evidence of high-identity theft states passing laws in earlier years compared with low- identity theft states.

DATA ANALYSIS

Panel Data: Basic Model

We now specify our econometric model to analyze how adoption of laws affects identity theft. To identify the effect of law, we exploit the panel nature of our data and employ state and time fixed effects. Thus, our basic estimating model has the form:

$$idtheft_{st} = \beta_0 + \beta_1 hasLaw_{st} + \sum_i \rho_i \; related_{st} + \sum_j \delta_j economic_{st} + \gamma \; fraud_{st} + \theta_s + \lambda_t + \varepsilon_{st} \qquad (1)$$

idtheft is the reported identity thefts in each 6-month period in state *s* at time *t*. For robustness, we provide results for both identity theft rate (crimes per 100,000 population) and the log of identity theft.

$hasLaw_{st}$ is a dummy variable, coded as 1 (one) if the state has adopted the law and zero otherwise. This dummy captures the effect of law on the identity theft rate. The dates of the adoption of data breach notification laws (between January 1, 2002 and December 31, 2009) were obtained from state legislature websites. For the purpose of analysis, we are interested in the date the law became effective rather than the date the law was passed. As described, we code adoption during each 6 month time period for a number of reasons. First, this is the smallest time frame by which we expect firms would be able to improve their security practices. Next, by way of legislative procedures, state legislatures generally set the effective date for laws to be either the beginning of the calendar year (Jan 1st), or the beginning of the fiscal year (Jul 1st). Indeed, for our sample 30 of the 45 states adopted the laws either exactly on, or within 1 month after either of these dates. 7 more states had effective dates within two months of a new period, in which case we coded the law has having being adopted in that period. For the remaining 8 states (those that adopted 3 or more months into the period), the law was coded starting the following period. E.g. if adoption occurs more half way through a 6-month period, we set the adoption to occur in the next period.

$Related_{st}$ represents credit-related laws that may also affect (prevent) identity thefts. One such legislation is the credit freeze law. These laws enable consumers to apply access control to their credit reports, thereby preventing firms with whom they have no prior agreement to make credit inquiries. If an

[13] Note that it will not prevent victimization if the attacker uses an existing account.

attacker is trying to open a new account that requires a credit check, they will be stopped and this kind of identity theft will be prevented.[13] The Fair and Accurate Credit Transactions Act (FACTA)[14] is a federal legislation that was passed as a response to identity theft. It allows individuals to request a free annual credit report. This legislation was enacted over the period from December, 2004 to September, 2005 beginning with west coast states and ending with east coast states. A variable was coded as 1 (one) if the law existed in a given state/time and 0 (zero) otherwise.

$Economic_{st}$ is a vector of state-level economic and demographic controls, as are commonly used in crime analysis (Lott & Mustard, 1997; Donohue, 2004; Wolfers & Donohue, 2006), such as the log of population, per capita income, and the average unemployment rate over each 6 month period (16 periods total).

As discussed above, there are many causes of identity theft that are not due to data breaches, such as lost or stolen wallet, home computer intrusions, phishing, and so forth. We used "Fraud," as recorded by the FTC, as proxy for these other sources. Fraud data is collected, managed and reported in a virtually identical method as identity theft and includes such activities as shop-at-home/catalog sales, prizes/sweepstakes, internet auctions, and foreign money offers.

State population data were obtained from the US Census bureau. Unemployment rates were collected from US Department of Labor, Bureau of Labor Statistics. Personal income was gathered from the Bureau of Economic Analysis of the US department of commerce. Fraud data was collected from the FTC. With the exception of population, which is only available annually, all data is available either monthly (identity theft, fraud, unemployment rate, adoption of related laws) or quarterly (income). In the case of population, we linearly extrapolated the missing data point as the average of the two adjacent years. For example, the first 6-month period in 2008 was computed as the average of the second 6-month period in 2007 and the first 6-month period in 2008.

θ_s and λ_t are state and time fixed-effects and ε_{st} is the familiar error term. This state, time fixed effect model is widely used in the literature to examine the effect of a policy intervention (Bertrand et al., 2004). State fixed effects allow us to control for unobserved state specific factors and time dummies allow us to control for time trends. Thus the unbiased effect of *haslaw* can be identified from variation across state *and* time. Regressions are estimated with heteroskedastic robust standard errors clustered-corrected by state. Descriptive summary statistics for these variables are provided in Table 4.

[Insert Table 4: Descriptive statistics]

While our conceptual framework identifies mediating variables (for example, individuals who are notified and firm investments) the empirical model focuses on observable variables which ultimately affect the outcome of interest, identity theft. This practice is not unfamiliar. For instance, researchers who study the effects of concealed gun laws recognize mediating effects, but relate the dependent and independent variable through observable control variables (Lott & Mustard, 1997; Black & Nagin, 1998; Cleary & Shapiro, 1999), and those who study capital punishment are often interested in the deterrent effect on murder rates (Dezhbakhsh & Shepherd, 2004). In both cases, the analysis of the treatment effects acknowledges the mediating but unobservable factors, but uses crime as the dependent variable, and the effect of law and other economic and demographic controls as the independent variables. In

[14] See http://www.ftc.gov/opa/2004/11/facta.shtm, accessed 10/07/07

addition, as previously discussed, there is a crucial relationship between data breach disclosure laws and identity theft that legislators have drawn, and which provides the specific motivation for this analysis.

Note that, as discussed above, identity thefts occur for various reasons –one of which is the result of data breaches. An important consideration for both our conceptual model and empirical estimation, therefore, is whether we are measuring the change in identity theft caused by *all sources*, or change in identity theft caused by data breaches only. In reality, data aggregation may cause our standard errors to increase, but will not lead to a biased estimate. Consider a dependent variable (Y = total identity theft) consisting of two elements: identity thefts caused by data disclosure (y_1), and identity thefts caused by other reasons (y_2). If the effect of law (say, x_1) is to reduce only y_1 but not y_2, then the preferred regression is:

$$y_1 = \beta_0 + \beta_1 * x_1 + \varepsilon_1 \qquad \text{(A1)}$$

However, we do not observe y_1, but only $Y = y_1 + y_2$. Hence, the estimated model is:

$$Y = \gamma_0 + \gamma_1 * x_1 + \nu \qquad \text{(A2)}$$

The question is: how significantly biased is γ_1 from β_1? To estimate this, note that from A1 we have:

$$E[y_1 \mid x_1] = \beta_0 + \beta_1 * x_1$$

From A2 we have:

$$\begin{aligned} E[Y \mid x_1] &= E[(y_1 + y_2) \mid x_1] \\ &= E[y_1 \mid x_1] + E[y_2 \mid x_1] \end{aligned}$$

As long as y2 is independent of x1 (which is by construction):

$$= E[y_1 \mid x_1] + E[y_2]$$

This implies:

$$\begin{aligned} E[Y \mid x_1] &= \beta_0 + \beta_1 * x_1 + E[y_2] \\ E[Y \mid x_1] &= (\beta_0 + E[y_2]) + \beta 1 * x_1 \end{aligned} \qquad \text{(A3)}$$

Comparing A3 with A2, notice that $\gamma_0 = (\beta_0 + E[y_2])$ and $\gamma_1 = \beta_1$. Thus, γ_1 represents an unbiased estimate of the effect of law (though it will suffer from higher standard errors). If a covariate is correlated with y_2, then it would indeed be biased. In summary, even though our dependent variable reflects identity thefts due to reasons other than data breaches, we will still achieve unbiased estimates when these crimes are uncorrelated with the effect of law. This implies that the estimates we obtain reflect the effect of law on identity theft due only to breaches, and not identity thefts due to other causes, such as lost or stolen wallets.

Panel Data: Extended Model

The basic model in Eq. (1) estimates the average effect of law. We now extend that model to gain an understanding into how the laws may have differential effects.

Lagged law. It is conceivable that the effect of the laws increases as firms invest in security measures over time. To test this, we introduce three lagged dummies, *d1PerOld*, *d2PerOld*, and *d3PerOld*, representing 1 (6 months), 2 (one year) and 3 or more (1.5 years+) periods after the law is adopted, respectively.

The national effect. While the majority of breaches are, indeed, confined within a state, any diffusion across states may nevertheless reduce the power of our test. We use two measures to control for this. First, we weight identity theft by interstate commerce activity in 2002 as a proxy for how connected a state is with other states. Ideally, we would include this as an explanatory variable in our econometric model, however only cross sectional (not panel) data were available. Second, we interact the *hasLaw* dummy variable with the percentage of all American states that have adopted the law by that time (*Law*PercStatesWLaw*). The *haslaw* dummy can now be interpreted as the effect of law when no other states have adopted these laws. If the effect is national, then we should find that once a few states adopt the laws, then the marginal impact of law may reduce considerably.

Differential effect of law across the states. It is reasonable to consider that the effect of the laws may be different across the states. The Bureau of Justice, National Crime Victimization Survey on Identity Theft (Baum, 2007) reported greater levels of identity theft for households with higher incomes and in more urban locations. Hence, we create two indicator variables, high income and urbanization. We first set a dummy variable equal to 1 if the state's income is greater than the median income from 2009 ($37,124). States coded as high-income in this period remain high-income in all time periods. We then interact this high income dummy variable with the breach law (*Law*HighIncome*). Next, using data on percent urbanization for each state,[15] we set an indicator variable equal to 1 if the state's percent urbanization is greater than the mean of 68.8 percent (results are unchanged using the median urbanization). We then interact urbanization with the state's adoption of the law (*Law*Urban*).

Strictness of Law. In the basic model, we have assumed that all breach disclosure laws are homogenous. In the extended model, we relax this assumption, and consider that some laws may be stricter if they exhibit the following properties: are acquisition-based (forcing more disclosure from a lower threshold of breach), cover all entities (businesses, data brokers and government institutions), and allow for a private right of action (i.e. individual or class action law suits). Based on the examination of state laws, we classify 11 states as having stricter laws: California, Hawaii, Maryland, Massachusetts, Minnesota, Rhode Island, Tennessee, Vermont and Virginia, We then interact strictness with the state's adoption of the law (*hasLaw*Strict*) to compare states with strict and non-strict laws.

RESULTS

Effect of Law on Identity Theft

The results of the regression in Eq. (1) (the Basic Model) are shown in Table 5. The dependent variable in Columns 1 and 2 is identity theft rate (crimes per 100,000 population), however, since we are also interested in the effect of the law on the change in identity theft, the dependent variable in Columns 3 and 4 is the log of total identity thefts. The variable of interest is *hasLaw*, the effect of data breach

[15] See http://allcountries.org/uscensus/37_urban_and_rural_population_and_by.html, accessed 01/10/08.

disclosure laws. Results are shown first using just state and time fixed effects (Columns 1 and 3), then with the full set of explanatory variables (Columns 2 and 4).

[Insert Table 5: Effect of law on identity theft, Eq. (1)]

Given that we are, in fact, most interested in the effect of the law on the change in identity theft, the dependent variable in all subsequent specifications is the log(identity theft) and therefore the coefficients are interpreted as percent changes in identity theft. Further, we believe that the log-level may provide a more intuitive interpretation of the average effect of law and not dependent on a given year's identity theft rate or the method of averaging. Log-level specifications also provide a better fit for the data as demonstrated by the R^2 values in Table 5.

We also report the results of the extended model in Table 6. Column 1 of Table 6 describes the results of the basic model and we extend it in column 2 (lagged law), column 3 (weighting the identity theft by state's commerce; controlling for the national effect). To avoid clutter, we do not report the interaction of law with state specific effects, interaction of law with other states adopting the law, and strictness of law. These effects are statistically and economically insignificant. All specifications use cluster-corrected standard errors by state and include time dummies for 16 periods, though we do not report those estimates to improve readability.

[Insert Table 6: Effect of law on identity theft, Eq. (2)]

Overall, we expected a negative coefficient for all of the law-related variables, indicating that their presence would reduce the numbers of identity thefts. In the Basic Model (column 2), the coefficient of law is -1.368, and marginally significant at the 10 percent level. Using the overall average identity theft rate (per 6-month period) of 32.8, the estimate suggests that, on average, adoption of data breach disclosure laws reduces the identity theft rate by about 4.1 percent (1.368/32.8). The log-level specification, however, which also provides a better fit for the data, suggests that adoption of the law reduces identity thefts by 5.7 percent, on average, and is significant at the 1% level. Because of the stronger fit for the data, further specifications are provided using the log-level outcome variable.

Column 1 in the Extended Model shows the effect of the lagged adoption of law, and suggests that there is no significant change after 6 months, whereas 12 months after adoption identity theft decreases by about 3 percent and is significant at the 1% level. However, the law appears to have no effect 6 or 18 months after adoption.

The dependent variable in column 2 weights the identity theft rate by the percentage of interstate commerce as an attempt to compensate for consumer reports in one state that could have actually occurred in another state. The interpretation of the coefficient is unchanged from previous specifications and in fact exhibits the same magnitude and statistical significance as the simple log-level specification.

Specification 3 tests the marginal effect of more urban states. The coefficient of interest is the interaction between *urban* and *hasLaw*. The results suggests that, indeed, the data breach laws reduce identity theft in more urban states by just over 9 percent, relative to less urban states.

As mentioned above, we also examined the impact of law at a national level, in states with higher average per capita income and stricter laws, yet we found no significant results, suggesting that the laws in higher income states do not reduce identity theft relative to their complement. Moreover, stricter laws are not found to reduce identity thefts more than weaker ones.

Robustness

A further consideration of disclosure laws is that they may produce a conflicting (opposing) effect by increasing consumer awareness - what we call an *awareness bias.* Since identity theft rates are based on self reported information, the passage of law may increase consumer awareness, causing more people to report incidents. We attempt to control for this awareness bias by using the Google archive search feature to search for the phrase "identity theft" in two state newspapers for each state over each 6-month period from 2002 to 2009. We find that our results are robust when we control for this awareness bias.

We also performed the panel data robustness analysis as described by Dugan (2002) and we present the graphical results in Figure 7.

[Insert Figure 7: t-statistics for per capita and log identity theft]

The y axis represents the t-statistics from regressing identity theft on the full set of covariates and state and time fixed effects (Columns 2 and 4 in Eq(1)). The left panel refers to per capita identity theft while the right panel refers to log(identity theft). Each box plot represents the distribution of t-statistics as we omit each of the 50 states at a time (50 observations per boxplot). Then, we do this 6 times, first omitting data from 2005, then 2006, 2007, 2008, 2009. The "X" represents the inclusion of a year's data, while the "0" represents the exclusion of that year. For example the left box plot (in both panels) represents the distribution of t- statistics when we omit data from 2005 (0XXXX). The rightmost boxplot (in both panels) includes data from all years (XXXXX). The plots are presented on the same y-axis scale for easier comparison.

First, we notice that the t- statistics for per capita identity theft is generally smaller compared with log(identity theft), though the outliers are more pronounced. The outliers for most boxplots (in both panels) are those states which have never adopted the law (Alabama, New Mexico, Mississippi and South Dakota). Interestingly, the upper outliers (those which reduce the t-stat are generally Alabama, New Mexico and Mississippi) while the lower outlier is generally South Dakota.

Generally, this method is most useful for testing whether our average coefficient results are driven by particular states or years. For example if the boxplots for all years except 2009, were tight around -1.0, but then the boxplot for 2009 was much lower (say, -5.0), then this would be cause for concern. However, we see no such extremes in either panel, though, clearly the results for log(identity theft) in the right panel are stronger. Overall, we believe this provides further evidence that our results are robust to state and year outliers.

DISCUSSION

Our research analyses the impact of data breach disclosure laws on identity theft. We used a standard difference-in-difference approach commonly used in literature and have controlled for various limitations in the data. We find that adoption of these laws reduce identity theft due to breaches by a statistically significant amount of 5.7 percent, on average. To place this in context, recall that the average amount stolen from consumers in 2005 was $6,383 (Javelin, 2006). The mean number of identity theft reports over 2002-2009 was 238,791. Given a mean reduction of 5.7% (and 95% CI of 0.098 and 0.017), this provides a mean reduction in the cost of identity theft by $86.9 million (0.057 * $6,383 * 238,791; with a confidence interval between $149.5 million and $25.3 million).

We do not find any significant relationship regarding the strictness of these laws on identity theft, nor do we find any significant effect of the laws in regions of higher population. While we generally do not find evidence of the laws gaining strength with time, we do find some evidence that the laws were effective in a short term (6-12 month) period. This could be explained by a temporary heightened awareness by consumers of the notifications, causing them to briefly take more precautions. Perhaps, then, as more notices are sent, and without noticeable signals of the effect of their actions, consumers would became desensitized and ignored further notices.

The lack of otherwise economically stronger findings may be due to a number of factors. One obvious explanation is that the laws are simply not particularly effective at reducing the number of identity theft victims either because of lack of consumer or firm action.

Consumer inaction may likely be a result of behavioral decision biases such as *optimism bias* (consumers perceiving their chances of suffering identity theft to be very low), *rational ignorance* (consumers believing the cost of taking precautions outweighing any benefits they may receive), and *status quo bias* (consumers' own inertia inhibiting them from anticipating the consequences of identity theft and responding) (Loewenstein et al., in preparation). Magat & Viscusi (1992) argue that disclosure legislation will only be effective if the human element is considered. They claim that consumers are not always rational decision makers and that notices "must convey information in a form that can be easily processed, and in an accurate and meaningful way that will enable individuals to make informed decisions." For example, there is evidence that very few disclosure letters inform consumers of the data that was actually compromised or provide customer support contact information (Samuelson Law, 2007). In addition, fewer than 10 percent of the 163,000 consumers availed themselves of free credit monitoring services following the Choicepoint breach (Brodkin, 2007) and another study found that 44 percent of identity theft victims ignored breach notification letters (FTC, 2007b). A recent Ponemon survey discovered that 77 percent of respondents claimed to be concerned or very concerned about loss or theft of personal information, but only 47 percent of respondents took advantage of free or subsidized credit monitoring services (Ponemon, 2008).

On the other hand, managers of firms may also believe their probability of suffering a breach is small enough that they may still not fully appreciate (and therefore internalize) the associated penalties. Or, they may estimate the net direct and indirect costs of breaches to be quite small, compared to the investments necessary to significantly decrease the probability of those breaches. For example, Choicepoint incurred a total of $26 million in fines and fees (Vijayan, 2008) - and they survived, with their assets (consumer personal information) being valuable enough to become a recent acquisition target by Reed Elsevier (the parent company of LexisNexus; see Nakashima et al., 2008). In addition, TJ Max reported costs of $178M for a breach that was disclosed in early 2007 and involved over 47 million customer records. Despite this, they enjoyed a quarterly increase in profits by 47 percent one year later (Kaplan, 2008).

Furthermore, if the vast majority of identity theft does not originate from data breaches (either because the information is simply lost and will never be used maliciously, or because credit card companies reimburse consumers for their loss) then the maximum effectiveness of these laws is inherently limited.

It is also conceivable that limitations in the FTC data may restrict our inferences about the true effect of law. However, reported crime data is commonly used as a proxy for actual crimes in empirical studies. Moreover, effects such as awareness bias common to all states (say, from a nationally syndicated news program or nationally circulated online or printed magazine) would be captured in our regression

by time fixed effects. Similarly, unobserved state variables such as race or income which could potentially influence identity theft rates would be captured by state fixed effects.

A broader issue relevant to policy makers is whether there are other means by which this law could (and should) be evaluated. Environmental disclosure laws often measure a deterrent policy by their effectiveness at reducing not just the frequency of incidents, but also the severity of incidents and a firm's compliance with the regulation (Cohen, 2000). Therefore, it is possible that these disclosure laws could help reduce the severity of the crimes (as measured by consumer losses or type of identity theft), or compliance, as measured by the improvement in a firm's security practices. Indeed, studies have shown that a victim loses less money the sooner they become aware of fraudulent activity (FTC, 2007b; Javelin Research, 2006). Javelin claims that losses are 21 percent lower when consumers detect identity theft within the first week, and 65 percent lower when consumers detect the crime within a year. Moreover, they claim that average consumer costs declined in 2005 by 37 percent ($422).

CONCLUSION

As information security and privacy concerns rise, we will increasingly see legislation used as a tool for consumer protection, generating policy debates and significant lobbying. In this paper, we investigated the effects on identity theft rates of increasingly popular, though contentious, data breach disclosure laws. Despite many US states having adopted these laws since 2003, we have not seen any empirical work that examines their efficacy. Using panel data from 2002 to 2009 for 50 states, we conducted an empirical analysis to examine whether these laws have reduced the identity thefts. We found that the passage of law had reduced identity theft by about 5.7 percent.

Clearly, it appears that the effectiveness of data breach disclosure laws relies on actions taken by both firms and consumers. Firms can improve their controls; however, once notified, consumers themselves are expected to take responsibility to reduce their own risk of identity theft – something which only a minority appears to be doing. It may be that only with time we will see more firms internalize the costs of breaches (and ensuing identity theft), more consumers respond to the risks, and the victimization rates decline.

Proper research on the effectiveness of data breach disclosure laws is hampered by a relative scarcity data. Hoofnagle argues that the current collection of identity theft records is not sufficient, and that banks and other organizations should be required to release identity theft data to the public for proper research (Hoofnagle, 2007). We certainly agree with this view. To the extent that sampling and awareness biases can be reduced, it will allow researchers to more accurately measure the impact of disclosure laws. Moreover, we believe that the better collection of identity theft victimization, consumer and firm losses, and changes in firm behavior will be valuable information for researchers, policy makers and consumers.

APPENDIX

Figures

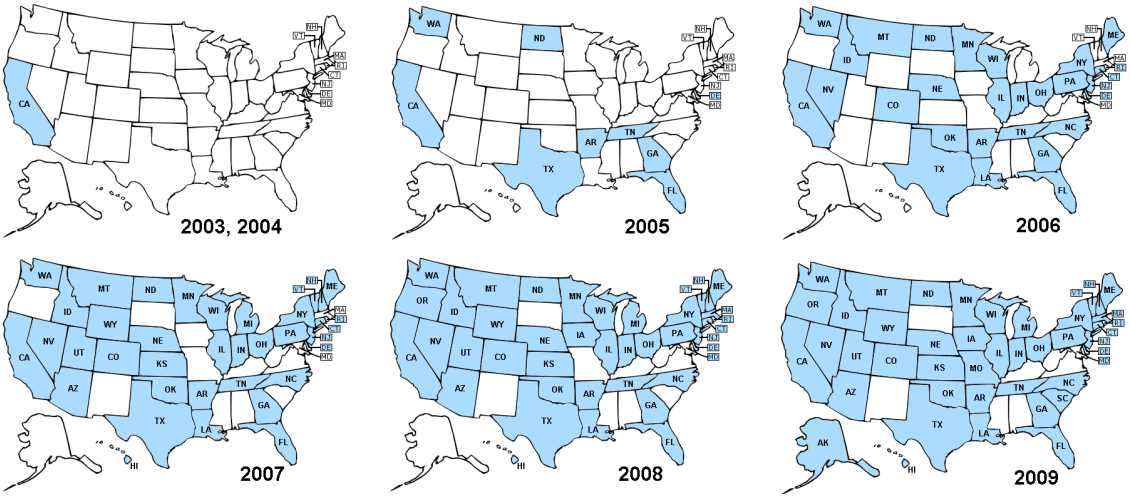

Figure 1: Adoption of breach notification laws from 2002-2009

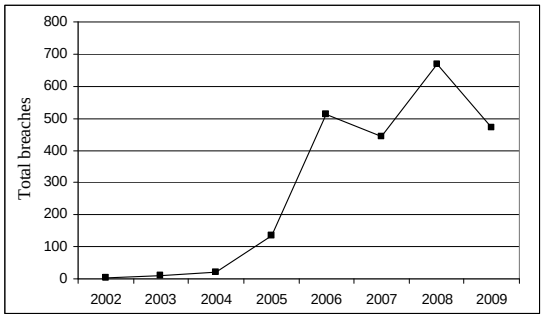

Figure 2: Data breaches from 2002-2009

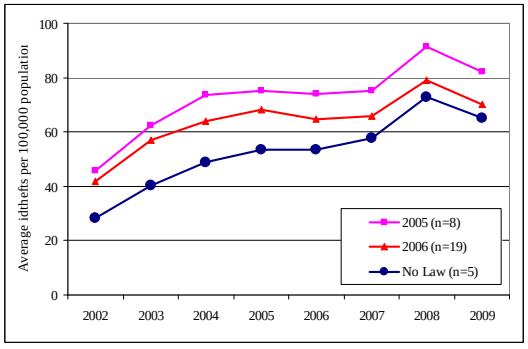

Figure 3: Average identity theft rates from 2002-2009

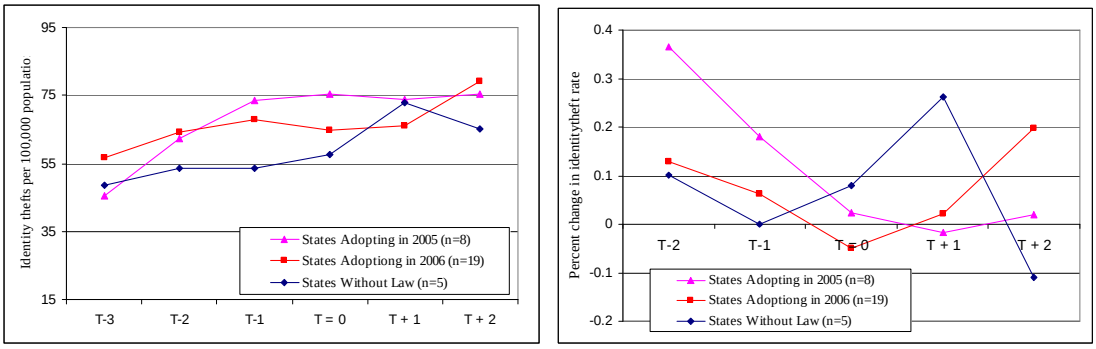

Figure 4: Identity theft rates and Percent changes before/after law

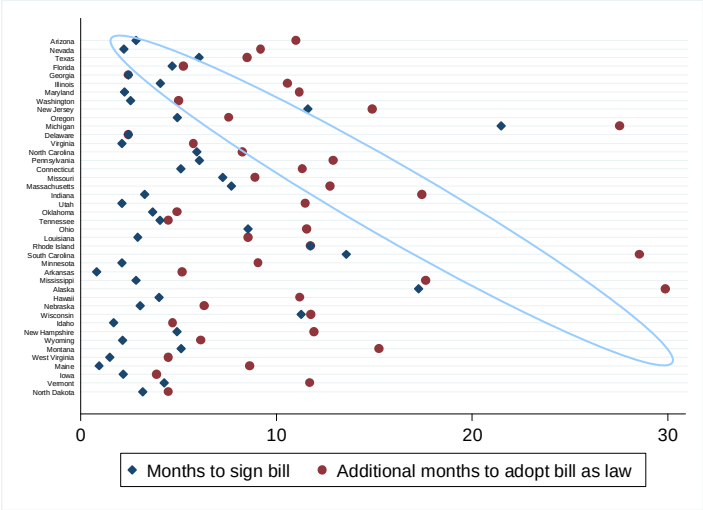

Figure 5: Months to sign and adopt data breach laws

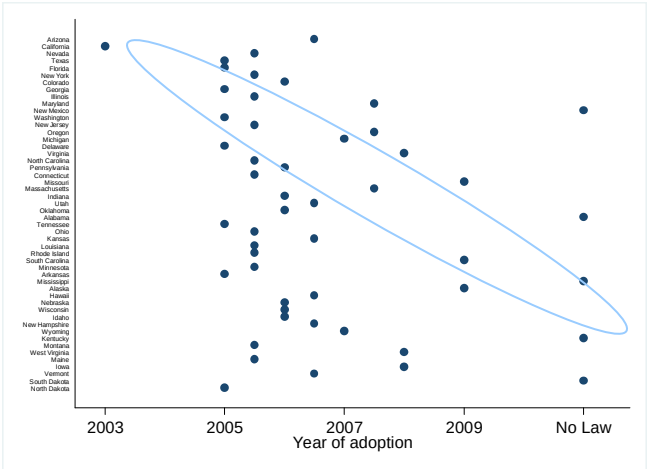

Figure 6: Date of adoption

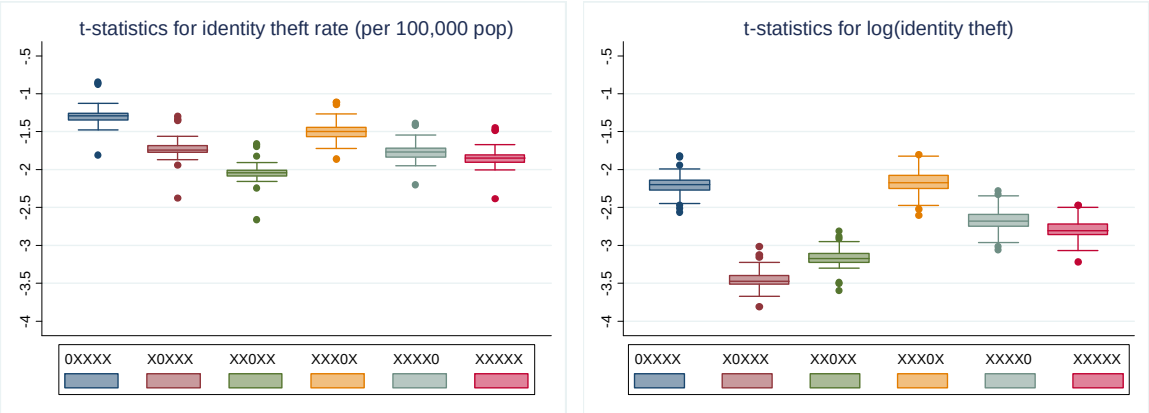

Figure 7: t-statistics for per capita and log identity theft

Tables

Table 1: Identity theft reports, 2002-2009

Year	Total	Average	Stdev	Min	Max	IDtheft rate	% change in rate
2002	154,327	3,087	5,059	81	30,782	53.7	
2003	207,116	4,142	6,576	127	39,500	71.5	33.0%
2004	239,037	4,781	7,520	179	43,900	81.7	14.3%
2005	247,747	4,955	7,676	158	45,180	83.9	2.7%
2006	238,627	4,773	7,228	178	41,415	80.1	-4.6%
2007	250,597	5,012	7,662	182	44,020	83.3	4.0%
2008	300,184	6,004	9,047	227	50,930	98.8	18.7%
2009	265,876	5,318	7,794	192	42,239	86.8	-12.2%

Table 2: Reported offenses per 100,000 population, 2002-2009

Year	Murder	Robbery	Burglary	MV Theft
2002	5.6	146.1	747.0	432.9
2003	5.7	142.5	741.0	433.7
2004	5.5	136.7	730.3	421.5
2005	5.6	140.8	726.9	416.8
2006	5.7	149.4	729.4	398.4
2007	5.6	147.6	722.5	363.3
2008	5.4	145.3	730.8	314.7

Table 3: Delay between filing, signature and adoption

	Mean	Median	Stdev	Min	Max	n
Filing - Signature	5.3	4.0	4.5	0.8	21.5	41
Signature-Adoption	5.1	4.2	3.8	0.0	15.0	45
Filing- Adoption	10.2	9.1	6.4	2.4	29.9	41

Table 4: Descriptive statistics

Variable (per 6-month period)	**Mean**	**Std. Dev**	**Min**	**Max**
Log(identity theft)	6.97	1.32	3.58	10.18
Identity theft (rate)	32.00	13.49	5.67	84.74
Identity theft (total)	2379.39	3709.80	36	26374
Has data breach law	0.38	0.48	0	1
Has FACTA	0.63	0.48	0	1
Has Credit Freeze Law	0.34	0.48	0	1
d1PerOld (6 months old)	0.05	0.22	0	1
d2PerOld (12 months old)	0.05	0.22	0	1
d3PerOld (18 months old)	0.05	0.22	0	1
Per capita income	35,547	6,701	23,019	66,690
Unemployment rate	5.42	1.73	2.37	14.37
Log(population)	15.11	1.01	13.11	17.43
Fraud rate (per 100,000)	82.28	45.35	16.80	323.28

Table 5: Effect of law on identity theft, Eq. (1)

	(1)	(2)	(3)	(4)
Dependent variable	idtheft rate	idtheft rate	log(idtheft)	log(idtheft)
Has Law	-0.928	-1.358*	-0.050*	-0.057***
	(0.818)	(0.733)	(0.026)	(0.020)
Has FACTA		1.869**		0.029
		(0.757)		(0.018)
Has CreditFreezeLaw		0.996		0.021
		(0.702)		(0.021)
Income per capita		0.000		-0.000
		(0.000)		(0.000)
Unemployment rate		0.145		0.005
		(0.425)		(0.010)
ln(population)		6.095		0.132

		(12.684)		(0.327)
Fraud per capita		-0.029*		-0.002***
		(0.016)		(0.000)
State and Time Fixed Effects	Yes	Yes	Yes	Yes
Constant	29.058***	-77.518	6.852***	4.554
	(0.364)	(194.260)	(0.014)	(5.031)
Observations	800	800	800	800
R-squared	0.752	0.759	0.848	0.859
Number of stateid	50	50	50	50

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6: Effect of law on identity theft, Eq. (2)

Dep var: log(idtheft)	(1)	(2)	(3)
VARIABLES	Lagged	Interstate	Urban
hasLaw		-0.057***	-0.009
		(0.020)	(0.026)
Law_Urban			-0.092***
			(0.027)
d1PerOld	-0.017		
	(0.014)		
d2PerOld	-0.031***		
	(0.011)		
d3PerOld	-0.024		
	(0.015)		
State and Time Fixed Effects	Yes	Yes	Yes
Full set of explanatory variables	Yes	Yes	Yes
Constant	4.962	8.604*	2.801
	(5.127)	(5.031)	(4.033)
Observations	800	800	800
R-squared	0.857	0.859	0.865

Number of stateid	50	50	50

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

REFERENCES

Acquisti, A., Friedman, A. & Telang, R. (2006, June 26-28). Is There a Cost to Privacy Breaches? An Event Study. Fifth Workshop on the Economics of Information Security.

Baum, K. (2006). Identity Theft, 2004. Bureau of Justice Statistics Special Report, NCJ 212213.

Baum, K. (2007). Identity Theft, 2005. Bureau of Justice Statistics Special Report NCJ 219411.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-in-Differences Estimates? The Quarterly Journal of Economics, MIT Press,. 119(1), 249-275.

Biderman, A. D. & Reiss, Jr., A. J. (1967). On Exploring the 'Dark Figure' of Crime. Annals of the American Academy of Political and Social Science, Vol. 374, Combating Crime, 1-15.

Black, D. & Nagin, D. (1998). Do Right-to-Carry Laws Deter Violent Crime? The Journal of Legal Studies, 27(1), 209-219.

Blumstein, A., Cohen, J. & Nagin, D. (1978). Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates, Report of the Panel of Deterrence and Incapacitation. Washington, D.C.: National Academy of Sciences.

Brodkin, J. (2007, April 10). Victims of ChoicePoint data breach didn't take advantage of free offers. Network World. Available at http://www.networkworld.com/news/2007/041007-choicepoint-victim-offers.html. Access date September 12, 2009.

Campbell, K., Gordon, L. A., Loeb, M. P. & Zhou, L. (2003). The economic cost of publicly announced information security breaches: empirical evidence from the stock market. Journal of Computer Security, 11, 431-448.

Cate, F. (2005, February 27). Another notice isn't answer. USA Today. Available at http://www.usatoday.com/news/opinion/2005-02-27-consumer-protection-oppose_x

.htm. Access date October 15, 2009.

Cate, F. (2009, March 6). Presentation at BCLT/BTLJ 2009 Symposium. University of California Berkeley, Berkeley Center for Law & Technology.

Cavusoglu, H., Mishra, B.& Raghunathan, S. (2004). The Effect of Internet Security Breach Announcements on Market Value: Capital Market Reactions for Breached Firms and Internet Security Developers. International Journal of Electronic Commerce, 9(1).

Cleary, J. & Shapiro, E. (1999). The Effects of "Shall-Issue" Concealed-Carry Licensing Laws: A Literature Review. Information Brief, Minnesota House of Representatives, Research Department.

Cohen, M. A. (2000). Empirical Research on the Deterrent Effect of Environmental Monitoring and Enforcement. Environmental Law Reporter, 30, 10245-52.

Dezhbakhsh, H. & Shepherd, J. (2004). The Deterrent Effect of Capital Punishment: Evidence from a 'Judicial Experiment.' American Law and Economics Association Working Paper No. 18.

Donohue, J. & Ayres, I. (2003). Shooting Down the 'More Guns, Less Crime' Hypothesis. Stanford Law Review 51.4.

Donohue, J. (2004). Guns, Crime, and the Impact of State Right-to-Carry Laws. Fordham Law Review 73.

Dugan, L. (2002). Identifying Unit-Dependency and Time- Specificity in Longitudinal Analysis: A Graphical Methodology. Journal of Quantitative Criminology, 18(3):213-237.

Federal Trade Commission. (2003). FTC Identity Theft Survey Report: 2003. Federal Trade Commission and Synnovate.

Federal Trade Commission. (2005, June 16). Data Breaches And Identity Theft. Prepared Statement of the Federal Trade Commission Before The Committee On Commerce, Science, And Transportation, U.S. Senate.

Federal Trade Commission. (2007a). Consumer Fraud and Identity Theft Complaint Data: January-December 2006. Federal Trade Commission.

Federal Trade Commission. (2007b). FTC Identity Theft Survey Report: 2006. Federal Trade Commission and Synnovate.

Fung, A., Graham, M., &Weil, D. (2007). Full Disclosure: The Perils of and Promise of Transparency. Cambridge University Press.

Gordon, G. R., Rebovich, D.J., Choo, K & Gordon, J. B. (2007). Identity Fraud Trends and Patterns: Building a data-based foundation for proactive enforcement. Center for Identity Management and Information Protection (CIMIP), Utica College.

Government Accountability Office. (2006). Testimony Committee on Government Reform, House of Representatives; Preventing and Responding to Improper Disclosures of Personal Information. Statement of David M. Walker, Comptroller General of the United States, GAO-06-833T.

Government Accountability Office. (2007). Data Breaches Are Frequent, but Evidence of Resulting Identity Theft Is Limited; However, the Full Extent Is Unknown. GAO publication GAO-07-737.

Givens, B. (2000). Identity Theft: How It Happens, Its Impact on Victims, and Legislative Solutions. Written Testimony for U.S. Senate Judiciary Subcommittee on Technology, Terrorism, and Government Information.

Hamilton, J. T. (1995). Pollution as News: Media and Stock Market Reactions to the Toxics Release Inventory Data. Journal of Environmental Economics and Management, 28(1), 98-113.

Hoofnagle, C. J. (2007). Identity Theft: Making the Known Unknowns Known. Harvard Journal of Law and Technology, Vol. 21.

Hutchins, J. P. (ed). (2007). Data breach disclosure laws - State by State. American Bar Association.

Javelin Research. (2006). Identity Fraud Survey Report: 2006. Javelin Strategy & Research.

Javelin Research. (2007). Identity Fraud Survey Report: 2007. Javelin Strategy & Research.

Jin, G. Z. & Leslie, P. (2003). The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards. Quarterly Journal of Economics, 118(2), 409-51.

Kaplan, D. (2008, February 20). TJX reports soaring profits one year after breach disclosure. SC Magazine. Available at http://www.scmagazineus.com/TJX-reports-soaring-profits-one-year-after-breach-disclosure/article/107072/. Access date October 20, 2009.

Ko, M, & Dorantes, C. (2006). The impact of information security breaches on financial performance of the breached firms: an empirical investigation. Journal of Information Technology Management, 17(2).

Konar, S. & Cohen, M. A. (1997). Information As Regulation: The Effect of Community Right to Know Laws on Toxic Emissions. Journal of Environmental Economics and Management, Elsevier, 32(1), 109-24.

Lenard, T. M. & Rubin, P. H. (2005). Slow Down on Data Security Legislation. Progress Snapshot 1.9. The Progress & Freedom Foundation.

Lenard, T. M. & Rubin, P. H. (2006). Much Ado about Notification. Regulation, 29(1), 44-50,

Levitt, S. D. (1995). Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error? NBER Working Paper No. W5268.

Loewenstein, G., John, L. & Volpp, K. (in preparation). Using decision errors to help people help themselves. In E. Shafir (Ed.), The Behavioral Foundations of Policy. Princeton, NY: Princeton University Press.

Lott, Jr., J. R. & Mustard, D. B. (1997). Crime, Deterrence and the Right-to-Carry Concealed Handguns. Journal of Legal Studies, 26(1), 1-68.

Magat, W. A. & Viscusi, W. K. (1992). Informational approaches to regulation. MIT Press.

Mathios, A. (2000). The Impact of Mandatory Disclosure Laws on Product Choices: An Analysis of the Salad Dressing Market. Journal of Law and Economics, 43(2), 651-77.

Mocan, H. N. & Gittings, K. (2003). Getting Off Death Row: Commuted Sentences and the Deterrent Effect of Capital Punishment. Journal of Law and Economics, 46(2), 453-78.

Nagin, D. (1978). General Deterrence: A Review of the Empirical Evidence. In Alfred Blumstein, Jacqueline Cohen, and Daniel Nagin (eds.), Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime. Washington, D.C. National Academy of Science.

Nagin, D. (1998). Criminal Deterrence research at the outset of the twenty-first century. Crime and Justice: A Review of Research, 23, 1-42.

Nakashima, E. & O'Harrow, R. Jr. (2008, February 22). LexisNexis Parent Set to Buy ChoicePoint. The Washington Post. Available at http://www.washingtonpost.com/wp-dyn/content/article/2008/02/21/AR2008022100809.html. Access date November 1, 2009.

Neal, T. (2005). Learning the Game: How the Legislative Process Works. National Conference of State Legislatures.

Ponemon Institute. (2005). National Survey on Data Security Breach Notification. The Ponemon institute.

Ponemon Institute. (2008). Consumer's Report Card on Data Breach Notification. ID Experts and The Ponemon Institute.

Ranger, S. (2007, September 3). Data breach laws make companies serious about security. Silicon.com. Available at http://management.silicon.com/itdirector/0,39024673,39168303,00.htm?r=1. Access date September 7,2009.

Robinson, P. H. & Darley, J. M., (2003). The Role of Deterrence in the Formulation of Criminal Law Rules: At Its Worst When Doing Its Best. 91 GEO. Law Journal, 949-1002.

Samuelson Law, Technology, & Public Policy Clinic. (2007). Security Breach Notification Laws: Views from Chief Security Officers. University of California-Berkeley School of Law.

Security and Exchange Commission. (2008). Part 248—Regulation S–P: Privacy of Consumer Financial Information and Safeguarding Personal Information; Proposed Rule. Available at http://www.sec.gov/rules/proposed/2008/34-57427fr.pdf.

Science and Technology Committee. (2007). Personal Internet Security. House of Lords, Science and Technology Committee, 5th Report of Session 2006–07, HL Paper 165–I.

Schwartz, P. & Janger, E. (2007). Notification of Data Security Breaches. 105 Michigan Law Review 913.

Simitian, J, (2009). *How a Bill Becomes a Law, Really*. Berkeley Technology Law Journal, 24(3), 1009-1017.

Simitian, J, (March 6, 2009) 13th Annual Symposium: Security Breach Notification, Berkeley Center for Law & Technology, Berkeley Technology Law Journal.

Tarr, G. Alan. (2000). Understanding State Constitutions. Princeton University Press.

Telang, R. & Wattal, S. (2007). An Empirical Analysis of the Impact of Software Vulnerability Announcements on Firm Stock Price. IEEE Transactions on Software Engineering paper, 33 (8), 544-57.

U.S. Congress. (2005a). Identity Theft: Recent Developments Involving the Security of Sensitive Consumer Information: Hearing Before the Senate Comm. on Banking, Housing, and Urban Affairs.

U.S. Congress. (2005b). Assessing Data Security: Preventing Breaches and Protecting Sensitive Information: Hearing Before the House Comm. on Financial Services.

U.S. Congress. (2005c). Securing Electronic Personal Data: Striking a Balance Between Privacy and Commercial and Governmental Use: Hearing Before the Senate Comm. on the Judiciary.

U.S. Congress. (2005d). Securing Consumers' Data: Options Following Security Breaches: Hearing Before the Subcommittee on Commerce, Trade, and Consumer Protection of the House Comm. on Energy and Commerce.

Vijayan, J. (2008, January 29). ChoicePoint to pay $10M to settle breach-related lawsuit. Network World. Available at http://www.networkworld.com/news/2008/012908-choicepoint-to-pay-10m-to.html. Access date November 1, 2009.

Walker. (1969). The diffusion of innovations among the American states. American Political Science Review 63:880-899.

Wolfers, J. & Donohue, J. J. (2006). Uses and Abuses of Empirical Evidence in the Death Penalty Debate. CEPR Discussion Paper No. 5493.