In this Comment, I provide a brief overview of what I believe is a useful framework for understanding different forms of government interventions (policies), and how they can be applied to IT security in order to protect critical infrastructure. I also describe cyberinsurance, and discuss how this market (not government) mechanism may hold great promise for improving the security posture of the private sector. I begin by addressing what I believe is a relevant and as of yet unanswered question regarding the current state of IT security investment.

Are Firms (and Consumers) Investing Enough in IT Security?

I believe this is a relevant question to first ask because it addresses a fundamental issue and appropriately frames the problem. Consumer advocates, and security and privacy professionals would likely all agree that, “no, companies are not spending enough on IT security.” And in some sense this makes sense: just look at all the data breaches that have been reported. Clearly, we need only count the number of data breaches, privacy intrusions, and software vulnerabilities in order to show that companies are not spending enough to protect consumers’ data and produce safe applications. I see two problems with this argument, however.

First, it assumes that consumers cannot -- or should not -- take measures to protect their own data and computing devices. Like pedestrians looking to cross a busy roadway, do consumers, themselves, not also bear some responsibility for taking appropriate precautions when browsing the internet, and protecting their personal data? Certainly, there are many circumstances when individuals are harmed through no fault of their own (e.g. hack of one’s personal data without consent causing identity theft). However, there are also many situations where individuals are or could be empowered to take measures to protect their data (e.g. practicing proper browsing habits, appropriate disclosure of personal information, password hygiene, laptop and data record storage, etc). And so any discussion of how to induce companies to invest in security measures should also be accompanied with a similar discussion of how to induce consumers to take appropriate security and privacy precautions.

My second concern is that the argument (that companies do not invest enough) implicitly assumes that a world where companies did properly invest would experience zero data breaches or security incidents. Effectively, “appropriate” security, by that argument, implies “absolute” security. But as we have heard many times, perfect security is only achievable with zero utility (i.e. an unpowered computer is perfectly secure but entirely useless). Therefore, if we recognize that
perfect security is neither practical nor efficient, and we instead seek to have both companies and individuals bear some responsibility for their actions, and that each should invest in an efficient level of precaution (i.e. one that balances the incremental costs with incremental benefits), then this would describe a world in which data breaches and security incidents existed.

My point is simply that the existence of security incidents could, in fact, reflect a state of efficient security investment. Just because we see some volume of data breaches or security incidents does not necessarily imply that companies (or individuals) are not already spending an efficient amount on data security. Yes, increased spending may reduce security incidents, but that additional (marginal) cost of investment may be larger than the (marginal) benefit from that investment. And so if it is efficiency we seek (who would choose to support waste, after all?), then before we answer the question of “how should we get companies to invest more?” we must first ask, “should we get companies to invest more?”

In sum, my point is that the goal of a policy maker should be to optimize -- not minimize -- security incidents. That is, it should seek to balance the (incremental) cost of a security measure with its (incremental) benefit. And so, the existence of data breaches and security incidents does not necessarily imply that companies are not investing in an efficient level of security. In fact, an absence of security breaches would likely imply excessive spending. Moreover, people, just like companies, are self-interested. We make decisions to maximize our welfare, and we should not expect companies to do otherwise (or blame them when they do). Therefore, it is also reasonable to consider that individuals, just as with companies, should also bear some responsibility for protecting their digital assets.

Agreed, this problem is more complicated when firms impose harms on other firms, or on individuals without compensating them. In these cases, government intervention can be useful. Next, I discuss some of these interventions which, again, should be viewed as ways to optimize security incidents, not eliminate them.

**Government Policies that Reduce Risky Behavior**

There are many forms of government intervention (aka public policies, laws, regulations) that can be adopted in order to induce companies (and individuals) to invest in an appropriate level of IT security, and optimize (not eliminate) any harms suffered because of their actions.

Three such interventions are ex ante safety regulation (mandated standards), information disclosure, and ex post liability. Ex ante regulation is considered a heavy-handed prevention mechanism that enforces a minimum standard of care. These mandated standards are useful when the harms are thought to be either catastrophic or miniscule, and affect large numbers of individuals. Obviously, catastrophic harms (e.g. nuclear disasters) are worth preventing through regulation, but so are minor harms that are distributed across thousands or millions of individuals. The reason is that the harm to any one individual may be small, but on aggregate, it is worth the burden of increased prevention. Regulation is also useful when the source of the harm is unknown. For example, in the case of health risks (pollution) or data breaches, often the offending company may not be known, and so it becomes worthwhile to mandate specific

---

1 Certainly there is evidence of security incidents or data breaches occurring because of lack of software patching, sloppy controls or behaviors – and these instances should be addressed – but that represents a separate matter.

2 Think of fire codes standards that define thickness of walls and materials used in clothing or homes.
controls to prevent avoidable harms. However, its effectiveness is hampered when the regulated inputs are only loosely correlated with the harmful outputs. For example, mandating two factor authentication and encryption on Health IT systems may be ineffective if health professionals share login information and remained logged in to applications, causing medical identity theft. That being said, monitoring compliance to a regulation may well be easier for an enforcement agency than estimating the amount of harm caused by a harmful event. For example, it may be much easier to verify that a company has implemented basic security controls and passed a security audit, than to measure the total amount of harm caused by a breach of these controls which then lead to a subsequent attack on another company.

Disclosure, on the other hand, can be thought of as a corrective mechanism that empowers individuals to avoid potential harms. These are thought to be light-handed forms of regulation that impose less of a burden on companies because of the requirement to disclose (nothing more, or less). However, cognitive biases may instead burden individuals receiving the notice, preventing them from acting, or causing them to over-react. For example, breach disclosure notices may be ineffective because consumers simply aren’t sure what specific actions they should take to prevent any tangible loss. Alternatively, disclosure of information that is too frequent or too horrifying (over-disclosure) may have the opposite effect. For example, disclosure of vicious neighborhood or university crime data may cause people to unnecessarily fear for their lives or cease socially beneficial activities.

Finally, ex post liability is meant to allow injured parties to recover any losses through civil litigation. The cost of actual litigation and the threat of future litigation are also expected to force companies to internalize any losses, inducing them to increase prevention measures, such as more security controls. Liability, however, becomes ineffective if the harms are either not cognizable, verifiable, or when the injuring party is unknown. Identity theft and privacy harms, for example, often suffer from these limitations. While sometimes granting standing for plaintiffs in cases of credit monitoring, courts frequently dismiss claims for increased potential risk of identity theft, or for the fear of future harms. See Romanosky et al (in review) for a full analysis of data breach litigation. However, a liability regime has the advantage that it allows firms to manage their IT security (generally, any prevention measures) on their own, in ways that are most efficient for them. When they do properly bear the cost of their actions, they will naturally seek ways to reduce their total cost of any harmful behavior.

See Romanosky et al. (2009) for a full discussion of how these three interventions have been applied to breaches of personal consumer information.

Taxes, subsidies, and nudging, are additional methods of inducing efficient behavior. Taxes and subsidies are often thought to produce equivalent outcomes whether a policy maker taxes bad behavior or subsidizes good behavior (and in this way, each are considered efficient policies). However taxes (or any form of sanction) will be less efficient as the cost of applying those sanctions increases. For example, a subsidy can simply be paid to those who comply, while noncompliance must be detected and enforced, which can be costly. On the other hand, a sanction may be preferred to a subsidy if the threat of sanctioning is credible, because the desired behavior is achieved at no cost (Wittman, 1984). Some have even suggested a “reversible reward” in which a subsidy is offered for compliance, but then in the event of non-compliance, that same reward is used to penalize (perhaps through litigation or other sanction) the company, thus doubling the incentive mechanism (Ben-Shahar and Bradford, forthcoming 2013).

3 See Romanosky et al (2010) for a discussion of consumer over-reaction to data breach disclosure laws.
4 See Ben-Shahar & Schneider (2011), and Ho (2012) for more discussion on the limitations of disclosure.
Further, simply taxing an injurer for the total amount of harm caused is inefficient when the injured party is themselves able to avoid the harm at lower cost (this is the familiar Coasian “low cost avoider” response to Pigouvian taxes; Romanosky et al. (2010)). For example, it is more efficient for ISPs to block malware on their networks when the cost to them of protection is lower than it is for consumers. However, this is a simplistic view because it assumes that a single action by either party would entirely prevent the harm. The reality is that each party can invest in a continuous range of activities to prevent many forms of harm and that the policy decision is not to hold one party or the other entirely accountable, but rather to evaluate each decision consecutively in order to understand who can act most inexpensively.

Nudging has become a very popular form of public policy (Thaler and Sunstein, 2009). It is a form of choice architecture that specifically exploits (rather than ignores) human cognitive biases in order to achieve outcomes that are thought to be in the best interests of the individual. For example, if students are more likely to fill up on foods that are presented first in a cafeteria lineup, then simply presenting healthier foods before fattening ones should create healthier plates, without eliminating personal choice. Indeed, there is no reason nudging cannot be applied to the private sector for the purpose of appropriate IT security investment. After all, companies (and government agencies) are run by humans.

The previous section outlined alternative policy interventions that can be adopted to induce better corporate behavior. As part of this discussion of risk prevention and mitigation, we next consider insurance. Specifically the market for, and use of, cyber insurance as a means of risk management.

**Cyberinsurance**

Since 2003, data breach disclosure (security breach notification) laws have helped notify individuals of the compromise of their personal information as well as shine a light on the data collection practices of companies (Romanosky et al, 2010). However, they have had a very powerful, though unintended consequence of increasing the demand for cyberinsurance. Because of this increased exposure, companies are facing increased loss due to breaches – losses for which they prefer to insure against.

Indeed, after a decade of underwriting tens of thousands of cyberinsurance policies, carriers are able to better assess a company's risk of loss, and more accurately price policies. Frequency and severity of loss data, together with improved analytics allows carriers to price policies that more accurately reflect an insured's expected loss.

Below, I discuss existing scholarship on cyberinsurance, and highlight some of the available data on this market.

**Existing cyberinsurance literature**

With few exceptions, the academic cyberinsurance literature consists of strictly theoretical papers that examine the viability of cyberinsurance markets. Overall, the literature examines the incentives for firms to purchase insurance (demand side), the incentives for insurers to provide contracts (supply side), and the conditions necessary in order for a market to exist. The inevitable tension for firms, as many identify, is whether to invest in security controls in order to reduce the probability of loss, or transfer the risk (cost) to an insurer.
The defining characteristics of cyberinsurance are: interdependent security, correlated failure, and information asymmetry. Some of these properties are common to all insurance markets, while others -- and their combined effects -- are unique to the risks of networked computing systems and cyberinsurance. First, *interdependent security* reflects the degree to which the security of one computer network is affected by the compromise of another system (the breached system is said to impose a negative externality on the victim). For example, the security of the DCA airport in Washington, D.C. may be compromised if luggage from SFO is not properly screened (Kunruther and Heal, 2009). Second, *correlated failures* is the systematic failure of multiple disparate systems due to a single event. For example, a computer virus spreading across networks, or a single malicious actor attacking many companies. (Notice the loss is further amplified by interdependent security.) Finally, *information asymmetry* in the context of insurance reflects the familiar moral hazard and adverse selection problems (i.e. companies behaving more risky when fully protected from loss; and insurance carriers not being able to differentiate between high and low risk clients).

It should be emphasized that while there are ways of reducing information asymmetries (we discuss these below), insurance carriers are mainly concerned with *correlated failures* because it defines the degree to which a breach by one firm (a potentially insured client) affects another, and therefore any indemnities paid. On the other hand, firms are mainly concerned with *interconnected nodes* because this determines how a failure by, say, a business partner, may affect them. However, as Bohem and Schwartz (2010) point out, the commonality is interconnected computing systems. To be clear, these simultaneous losses can occur in a number of ways. For example, interconnected IT systems enable computer viruses to propagate across disparate client networks; a single entity can direct a coordinated cyber-attack at multiple companies; or a security breach at an IT cloud service provider creates outages for all of its customers. Further, with the surge in big data business opportunities, companies are being seduced into collecting and analyzing massive amounts of consumer-level data -- data that becomes susceptible to compromise, the subject of litigation, and possibly an insurance claim.

Bohem & Schwartz (2010) provide an excellent summary of cyber insurance literature and define a unified model of cyberinsurance that consists of 5 components: the networked environment, demand side, supply side, information structure, and organizational environment. First, the network topology plays a key role in affecting both interdependent security and correlated failures; i.e. consider the difference in impact between extremes of independent computers versus a fully connected computing network. Their demand-side models consider the risk aversion of the insured, heterogeneity across wealth, impact, and defense and utility functions of firms. The supply-side discussion considers, among other properties, the competitive landscape of insurers, contract design (premiums, fines), and the carrier’s own risk aversion. Discussion of information structure relates to adverse selection and moral hazard. Finally, organizational environment describes issues such as regulatory forces that may exist to mandate insurance, require disclosure in the event of a loss, and the effect of outsourced security services and hardware and software vendors on a firm’s security posture.

As mentioned, risk management is often framed as a trade-off between investing in controls that reduce the average loss of a security event, and insuring against a loss. Indeed, Ehrich and Becker (1984) show that as insurance becomes more affordable, there is less incentive to invest in self-protection (IT security) measures. At an extreme, if the price of insurance were very inexpensive, companies would be very unlikely to protect themselves against any kind of loss. Conversely, as insurance becomes more expensive, companies become more willing to self-protect (the price of insurance becomes much higher relative to any security measures). Ehrich and Becker (1984) also suggest that the demand for insurance is increasing in the size of the loss, and decreasing in
probability of loss. That is, companies are more willing to insure against larger, less frequent loss events. Whether or not this is true should be a testable question.

The purpose of this section has been to briefly highlight some of the important contributions of academic literature on cyberinsurance and identify the key issues and tradeoffs that firms and insurance carriers face. Next, we examine the limited set of publicly available data on cyberinsurance and cyberinsurance policies.

**What do we know about the cyberinsurance market?**

Industry data is still quite difficult to obtain, making for limited insights into the cyberinsurance market. However, the following data were available from industry surveys and reports.

Average premiums appear to be priced between $10k - $25k, with some carriers writing premiums between $10m - $25m, and as high as $50m (Betterly Report 2012). Other reports suggest typical premiums of $100k for a limit of $10m (Airmic, 2012).

While deductibles may be as low as $5k, deductibles between $500k and $1m are common for companies with $1b or more in assets. Some can even be as high as $25m.⁶

One leading carrier claims that the average limits of cyber policies increased by 20% to $16.8 million, relative to 2011, with some purchases of $180 and $200 million in 2011 and 2012, respectively (Marsh, 2013; Airmic, 2012). This increase in limits was driven by communications, media, and technology companies and induced through federal regulation, such as HIPAA, and data breach disclosure laws. As with most other insurance products, towers of cyber policies can be purchased in the event of extreme (> $20m) losses, and Airmic (2012) suggests that limits of $200m - $300m exist for some industries.

While it appears that no one firm knows the total amount of annual premiums currently written, estimates range from $500m (Airmic, 2012) to $1b (Betterley 2012), and growth appears to be driven by smaller and medium sized companies. One leading US insurance company states that purchase of cyber insurance policies increased 33% in 2012, relative to 2011, and was driven mainly by the Services (“professional, business, legal, accounting, and personal”) and Education industries, and smaller and mid-sized firms (Marsh, 2013). Others suggest a more moderate growth between 10% and 25%, with some outliers extending to 100% or more (Betterly Report 2012).⁷ Further, despite the suggestion that breach response costs are becoming higher than expected, some estimates suggest that in order to attract even more business (or perhaps as a response to better informed actuarial data) carriers maybe reducing their loading factor by 5% (Betterly 2012).

**Cyberinsurance policies and security self-assessment forms**

Generally speaking, cyber policies are consistent across carriers in regard to the types of losses that are covered, the triggers of the loss, the types of data that are covered, and exclusions. In addition, some policies are standalone products, while others are built into existing corporate policies like Errors and Omission (E&O). In order to address adverse selection, most carriers also require clients to provide a security self-assessment form. This consistency among policies is

---

⁵ Personal correspondence between the author and an executive of a large insurance carrier.

⁶ Id.

⁷ Certainly these percentage increases would be more meaningful given absolute base rates, which we unfortunately were not able to obtain.
helpful for both clients and brokers when comparing policies across carriers. Consequently, it also suggests that carriers compete mainly on price (premiums, retention and limits), rather than quality (coverage, exclusions, etc).


Cyberinsurance policies cover three general categories of loss: 1st party losses, regulatory fines and fees, and 3rd party liability. 1st party coverage includes losses stemming from outages or business interruption costs incurred due to a data breach, privacy violation, or security incident. For example, breach notification costs, credit monitoring, PR, forensic investigations, call center support, business interruption, and in some cases even extortion. Regulatory fines, and fees cover sanctions brought by state or federal agencies (e.g. by the FTC or SEC). 3rd party liability coverage includes settlements, judgments, and defense costs due to civil litigation. Naturally, cyber policies also include many exclusions, such as: discrimination, criminal or deliberate acts, patent infringement or violations of trade secrets, or acts of war, invasion or insurrection.8

In regard to the security self-assessment forms, carriers appear to be very interested in all forms of administrative, technical, and physical controls related to, for example, application system and network access; storage, transmission, and destruction of confidential data; employee awareness training; password hygiene; 3rd party contracts; security management and audits; application patching; disaster recovery/business continuity; incident response; etc.9

What Can Cyberinsurance (and Insurance Carriers) do to Improve IT Security?

As we have stated, it is commonly held that protection from harm induces reckless behavior (moral hazard). And this is seen as a key obstacle to a robust cyberinsurance market (indeed, any insurance market). However, despite this apparent truism, it may be possible for cyberinsurance carriers to reduce this effect, and even improve firms’ security posture.

Moral hazard is predicated on the assumption that compensation for a covered loss is immediate, complete, and costless to obtain. Ehrich and Becker (1984) suggest, however, that when relief is burdensome– either in regard to time, money or embarrassment -- the effects of moral hazard are diminished. Two common examples of devices that exploit this are deductibles and co-payments. Together, they reduce moral hazard by forcing the insured to bear some of the cost of a claim. Consider also unemployment insurance. The act of claiming unemployment benefits is sufficiently unappealing that it greatly increases the burden of seeking relief, thereby reducing one’s incentive to voluntarily exit from the labor market. Though, certainly it should not be the practice of an insurance company to make legitimate claim filing unnecessarily laborious.

In addition, car insurance companies offer discounts for usage-based performance (pay-as-you-go). Drivers are rewarded for taking fewer, shorter trips, in less risky areas, at less risky times of the day. Conceivably, a similar form of usage-based performance could be applied to firms seeking cyberinsurance. For example, discounts to firms that employ a limited (not infinite) data retention policy. Many other mechanisms can (and do) exist to ensure that policy holders maintain appropriate protection measures ex ante, as well as ex post, thus reducing careless activity (see Ben-Shahar and Logue, 2012 for a complete list).

While all of these mechanisms may help induce appropriate (security) behavior, it should be clear that they are invoked only once a policy has been purchased. That is, these behaviors and rewards enjoyed only by those with a policy. For example, discounts for safe driving only encourages better behavior for insured drivers, not for pedestrians, or drivers without insurance.

So the question remains, can insurance help increase safety for everyone? Indeed, yes. It has happened before. Ben-Shahar and Logue (2012) provide clear and detailed examples of how insurance companies have sponsored research for improvements in fire, car, construction and home safety standards – standards that have later become industry best practices and, in some cases, ex ante regulation. These best-practices can then also be used in ex post liability procedures in order to establish fault (or no-fault) in negligence claims.

In this regard, the incentives by all market actors are aligned. Carriers have an obvious incentive to promote stronger self-protection measures because safer customers file fewer claims, making them more profitable. Since firms are profit seeking, and suffer from ambiguity aversion, they prefer to know which regulations and controls should be adopted (if only to avoid litigation or sanctioning). Consumers are also made better off when their data are not compromised and they don’t suffer identity theft. And finally, efficient use of security controls maximizes social welfare, which is appealing to policy makers.

**Recommendations**

Many people state that insurance carriers lack sufficient actuarial data in order to create hazard classes of insured. I dispute these claims. Carriers have been writing cyberinsurance policies for over a decade now, and with thousands of data breaches and hundreds (if not thousands) of security and privacy lawsuits (Romanosky, in review), carriers have surely acquired sufficient data.

This affords insurance companies with a unique advantage over any other entity – even government agencies -- when it comes to assessing the benefits of different IT security controls. Recall, the critical question is: which security controls are most effective at reducing risk? Is it better to have a firewall or an IDS? Is two-factor authentication really better than single-factor? If so, by how much, and how much of a discount should it enjoy? To date, no single firm or government agency has been able to answer this basic, yet fundamental, question. But I believe it can be answered.

Insurance companies are perfectly positioned because they possess the necessary data. Using their security assessment forms, policy and claims data, they can correlate the security controls of an insured with loss outcomes. Therefore, with sufficient data, the carrier could rank order security

---

10 See also Braunberg (2013).
controls by effectiveness. They can, in effect, determine which IT controls measures are most effective at reducing loss. These answers would be invaluable at driving IT security research, the market for cyberinsurance, and ultimately the security posture of US critical infrastructure.

The Department of Commerce could facilitate this effort in two separate ways. First, reaching out to insurance companies and encouraging them to participate in academic research that would enable researchers to identify firm characteristics and security controls that are most strongly associated with risk reduction. Second, work with carriers to use these results to help create and drive a set of best-practices that could ultimately become industry standard.
References


