Economics and Internet Security: a Survey of Recent Analytical, Empirical and Behavioral Research

Tyler Moore
and
Ross Anderson

TR-03-11

Computer Science Group
Harvard University
Cambridge, Massachusetts

Economics and Internet Security: a Survey of Recent Analytical, Empirical and Behavioral Research

Tyler Moore
Center for Research on Computation & Society
Harvard University
tmoore@seas.harvard.edu

Ross Anderson
Computer Laboratory
University of Cambridge
Ross.Anderson@cl.cam.ac.uk

1. Introduction

An economic perspective has yielded invaluable insights into the analysis and design of information security mechanisms. Systems often fail because the organizations that defend them do not bear the full costs of failure. This simple insight has profound consequences for a growing number of industries, and it extends to dependability as well as security. For instance, utilities reduce direct, measurable costs by routing control messaging over the Internet; this can raise the risk of service failure, whose costs are mainly borne by its customers. Another example comes from anti-virus software; since infected machines often cause trouble for other machines rather than their owners, expenditures on protection tend to be suboptimal. Online crime is growing rapidly; for example, the most recent British Crime Survey shows that more than twice as many citizens now fall victim to fraud each year as to traditional acquisitive crime such as house burglary and vehicle theft. There is no purely technical solution to growing vulnerability and increasing crime: law must allocate liability so that those parties in a position to fix problems have an incentive to do so. But at present it frequently does not; and this policy gap is widening as systems become global and acquire a myriad of competing stakeholders.

In this survey, we discuss the economic challenges facing information security in greater detail: misaligned incentives, information asymmetries and externalities. We then describe several key areas of active research: modeling attack and defense, breaches of personal information, the burgeoning underground markets for online criminal services, and the security of the payment system. We also describe the state of the art using three broad approaches: theoretical, empirical and behavioral analysis. Finally, because economic analysis has revealed significant barriers to the provision of information security, policy must play a role in any fundamental improvements. So we discuss proposed policy interventions. Researchers can make a significant impact by informing the policy debate in critical areas – which we try to identify.

2. Economic Barriers Facing Information Security

Information systems are prone to fail when the person responsible for protecting a system is not the one who suffers when it fails. For example, medical records systems are bought by hospital directors and insurance companies, whose interests in account management, cost control, and research conflict with

To appear in: T. Moore, R. Anderson, 2011. Internet Security. In: Peitz, M., Waldfogel, J. (Eds.), The Oxford Handbook of the Digital Economy, Oxford University Press.

patients' interests in privacy. Banks encourage customers to bank online to save in branch operating costs, even if attacks lead to frauds whose cost falls partly on customers. As pointed out by Anderson and Moore (2006), misaligned incentives between those responsible for security and those who benefit from protection are rife in IT systems. Consequently, any analysis of information security must begin with stakeholder incentives.

The second generic problem is asymmetric information. Most of the players in the infosec world have an incentive to exaggerate the harm: for instance, AT&T's Chief Security Officer Edward Amoroso testified to the US Congress in March 2009 that cyber-criminals' annual profit exceeds $1 trillion.[1] $1 trillion is about 7% of US GDP – bigger than the entire IT industry. The figure is implausible, yet similarly generous estimates are made by other firms in the security-industrial complex.

Victims, on the other hand, may under-report incidents. Banks do not want to undermine trust in their brand, and in most countries the banking industry will not even reveal aggregate figures for how much they are losing to online fraud[2]. Many other businesses do not want to report cyber-crime or cyber-espionage incidents out of fear over damage to their reputation and stock price. The lack of reliable data on the costs of information insecurity makes it difficult for firms to manage risk. One countermeasure has been security breach disclosure laws, now passed in over 30 U.S. States and under consideration by the European Parliament.

Asymmetric information also affects product quality. Akerlof's model of the "market for lemons" (Akerlof 1970) appears to apply to many security product markets (Anderson 2001). Vendors may assert their software is secure, but buyers cannot tell, and refuse to pay a premium for quality – so vendors invest in marketing rather than engineering. Bad security products drive out good ones from the marketplace. The primary countermeasure has come in the form of certification schemes such as the Common Criteria, but these have been repeatedly gamed.

The IT industry is characterized by many different types of externalities, where individuals' actions have side effects on others. We discuss two key types: network externalities and externalities of insecurity.

[1] http://commerce.senate.gov/public/?a=Files.Serve&File_id=e8d018c6-bf5f-4ea6-9ecc-a990c4b954c4

[2] UK banks do report aggregated fraud losses. In 2009, the total reported losses due to all forms of payment fraud were £440 million (approximately $641 million). Of that total, £59.7 million ($87 million) was attributed to online banking losses. Source: http://www.paymentsnews.com/2010/03/uk-card-and-banking-fraud-losses-down-28-in-2009-to-4403mm.html. David Nelson at FDIC has been trying to collect similar figures from US banks on a voluntary basis. He estimates that $120 million was collectively lost by US banks due to malware infections targeting online banking services. Source: http://www.computerworld.com/s/article/9167598/FDIC_Hackers_took_more_than_120M_in_three_months?source=rss_news. An estimate of the total annual proceeds from online crime may be in the neighborhood of the low billions of dollars.

The software industry tends toward dominant firms, thanks in large part to the benefits of interoperability and the resulting network externality: a larger network, or a larger community of software users, is more valuable to each of its members. Selecting a network or a software platform depends not only on its features and performance but also on the number of other people who already use it; such platforms typically create two-sided markets, where more users lead to more application software writers and other complementers. This helps explain the rise and dominance of Windows in operating systems, iTunes in online music sales and Facebook in online social networks. It also helps explain the typical pattern of security flaws. As a platform vendor is building market dominance, it must appeal to vendors of complementary products as well as to its direct customers. A secure platform would impose more burdens on complementary software developers, so security is neglected at the start, and only added once market dominance has been achieved. In addition, market races for platform dominance explain the philosophy of "Ship it Tuesday and get it right by version 3" which also leads to insecure software.

Network externalities also help explain why many attempts to provide more secure versions of Internet protocols, such as DNSSEC and S-BGP, have failed to achieve widespread adoption. Such protocols do not help much until many other users have also adopted them, so no-one wants to go first. The protocols that have succeeded, such as SSH and IPSec, have generally been those that provide adopting firms with internal benefits immediately.

Insecurity also creates negative externalities. A compromised computer that has been recruited to a network of compromised machines (a so-called botnet) will typically harm other computers more than the host. Botnets send spam, host phishing scams and other illegal content, launch denial-of-service attacks, and provide anonymous cover for attackers. In each case, the target is someone other than the host computer. In the event of cyber-attacks on critical infrastructure, the social cost of control systems failure (such as prolonged power outages) would greatly exceed the direct loss to a utility in terms of lost revenue. Because the private risks facing utilities are less than the social risks, we expect an underinvestment in protection. Finally, we must also consider the opportunity cost when people are afraid to use the Internet.

3. Key Problem Areas

We now discuss a series of key areas of active research in the economics of information security. The areas are:

- Modeling attack and defense
- Breaches of personal information
- Malware and botnets
- Payment system security

For each of these areas, we describe the key analytical, empirical and behavioral contributions, plus a discussion of opportunities for a research and policy agenda. Please note that, apart from the discussion on breaches of personal information, we do not cover the related topic of information privacy here.

3.1 Modeling Attack and Defense

Computer scientists used to design information security systems starting from assumptions about the adversary's capabilities. This approach worked for simple problems such as the design of cryptographic algorithms and protocols. However, now that machines are connected via the Internet to millions of other machines operated by competitors, we need to consider not just the capabilities of possible adversaries but also their motivation. Attackers and defenders are now seen as engaged in one or more strategic interactions, where the incentives to disrupt and protect systems matter most.

3.1.1 Analytical Research

In a seminal work, Varian (2004) analyzed the reliability of information systems using a public-goods game-theoretic framework. He noted that infosec can follow two canonical cases analyzed by Hirshleifer (1983): weakest-link or best shot. Hirshleifer told the story of Anarchia, an island whose flood defenses were constructed by individual families and whose defense depends on the weakest link, that is, the laziest family; he compared this with a city whose defenses against missile attack depend on the single best defensive shot. Varian added a third case, sum-of-efforts, where the city's defense depends on the total of all contributions by its citizens. He showed that sum-of-efforts is best, while weakest-link is worst of all.

These games naturally map to different infosec scenarios. For instance, a single software vulnerability introduced by a careless programmer error introduces a weakest link that can be exploited. Sometimes security depends on total effort – testing software for vulnerabilities depends on the sum of all testers' efforts, for example. There are also cases where security depends on the best effort – the actions taken by an individual champion such as a security architect.

Varian's analysis tells us that as more agents are added, systems become more reliable in the total-effort case but less reliable in the weakest-link case. In practical terms, this tells us that software developers should hire fewer, better programmers, many testers, and the best security architect they can find.

His model nicely captures the externalities often present in information security investment decisions. Security level is a public good, while defensive costs are private[3]. Several papers have extended Varian's model. Grossklags, Christin and Chuang (2008) generalize the class of games to include an arbitrary number of players with a choice between self-protection and self-insurance. In the former case, players invest in measures to protect against attack, such as firewalls, antivirus and patching. In the latter, they invest in mechanisms to mitigate the effects of compromise, such as data backup. The former are still public goods and help protect other players, whereas in the latter the gains are entirely private and insecurity can have negative consequences for others, such as a web server that has been compromised to host fake phishing pages.

They also introduce a novel game called "weakest-target". Here the overall security level is still determined by the weakest player, as in the weakest-link game. But the only players harmed are those with investment at the minimum level; all the others escape. This models attackers who only

[3] Others, notably Kunreuther and Heal (2003), have modeled the externalities which arise when security investment is undertaken privately but outcomes are correlated.

compromise the “low-hanging fruit”. For example, spammers do not care which computers they compromise – they just need to capture enough to send their spam. So they often use exploits that only work on unpatched, older versions of Windows.

With these additional extensions to Varian’s model in place, the authors find that welfare in the Nash equilibria often is only slightly less than the social optimum. Frequently, the socially optimal outcome is to invest in self-insurance rather than protection, or in the weakest-target game, to select “sacrificial lambs” to appease attackers while protecting the majority.

These three papers model defense strategically, but not attack. Several papers have attempted to study the strategic interaction of attackers and defenders. Danezis and Anderson (2005) found that peer-to-peer systems are more resilient to censorship attempts when the principals serve and defend content that they are interested in, rather than being assigned content randomly. The trade-off they study is analogous to the tension between social diversity and solidarity in modern societies.

Fultz and Grossklags (2009) study the security games of Varian (2004) and Grossklags et al. (2008) while explicitly modeling the interaction between attacker and defender. The attacker can choose the number of defenders to target, as well as the force of attack. The trade-offs for the attacker are now between the cost of attack, likelihood of detection and value of successful attack. A range of equilibria are possible. When self-insurance is cheaper than protection, the attacker operates with full force, while the defender insures or takes the loss. In weakest-link games, the attacker attacks but the defender only partially protects. For weakest-target and best-shot games, no pure symmetric equilibrium exists.

Böhme and Moore (2009) take a different approach to modeling adaptive attack and defense. In some ways, this is a bit more realistic than the games discussed so far. Defenders respond to attacks by plugging known holes; yet, as soon as one flaw is fixed, another weak point is often identified and exploited. For example, attackers construct botnets to send spam, distribute malware and host illegal content such as phishing websites. Attackers concentrate on the most irresponsible ISPs, moving on to others only after the ISP cleans up its act or is shut down (Moore and Clayton 2007, Day et al. 2008). Likewise, technical countermeasures to payment card fraud have evolved over time, causing fraudsters to adopt new strategies as old weaknesses are fixed. For example, when UK banks migrated to PIN verification of transactions rather than signatures, in-person retail fraud declined while overseas ATM fraud and card-not-present fraud skyrocketed (APACS 2007).

Böhme and Moore’s model captures this dynamic interaction: attackers find the weakest link, defenders fix the problem, attackers find new holes which are then plugged, and so on. A single defender protects an asset of value against *n* possible threats. Each threat can be warded off by investing in its corresponding defense. While the defender knows the costs for each defense, he may be less certain of the cost of different attacks. To model this uncertainty, the threats are ordered *1 to n* by increasing *expected* cost of attack. Figure 1 (left) illustrates how expected and actual costs of defense may vary in general, while Figure 1 (right) applies the model to the security of payment card systems as an example. Face-to-face retail fraud (F2F) might reasonably be seen as the weakest link in the payment card environment; its reduction in the UK following the adoption of EMV (“Chip and PIN”) cards supports this view. Similarly, the banks correctly anticipated that losses due to credit cards lost or stolen inside the UK would drop once PINs were required for use. One area where the banks' expectations were not met is with ATM fraud on UK cards outside the UK. It turned out that fraudsters could easily clone stolen UK cards and use them in foreign ATMs; hence the true cost of attack was lower than expected. Likewise, the banks' losses due to card-not-present fraud (CNP) were much higher than forecast; unsurprisingly, many banks then decided to deploy readers that verify PINs.

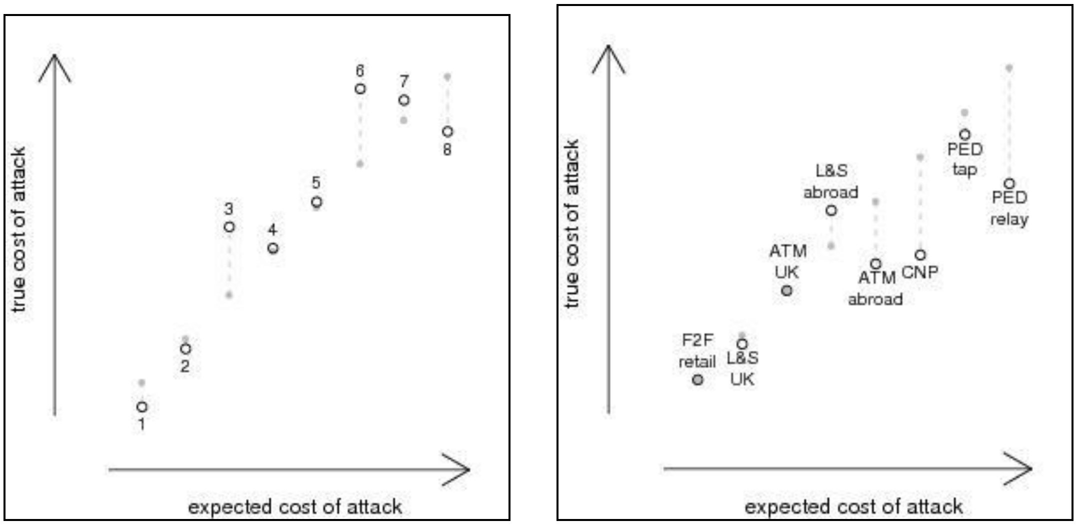

Figure 1: Expected vs. true cost of attack in the iterated-weakest link model.

The model is a repeated game; in each round, the defender decides which, if any, of the *n* threats to protect against. The attacker identifies and carries out the cheapest exploit, and only operates as long as it is profitable to do so.

In the *static case*, where the defender only gets one chance to protect a system, increasing uncertainty about which link is weakest causes the defender to protect more assets, but when uncertainty is too high, the defender does not know which asset to protect and so chooses to protect none. In the *dynamic case*, where repeated defensive investments are allowed, an uncertain defender may initially protect few assets and wait for the attacker to "identify" the weakest links to be fixed in later rounds. Hence, it can be quite rational to wait until attacks actually happen and then fix the stuff that fails.

Rather than having many defenders independently protecting a single asset, this model considers a single firm with many assets to protect. Furthermore, this model examines defender uncertainty rather than externalities. It turns out that defender uncertainty can be just as powerful an explanation for under-investment as externalities; both can lead to compelling market failures.

Thus far, we have discussed models where attack and defense are carried out by two mutually exclusive groups, the "good" guys and "bad" guys. In fact, the distinction between attacker and defender is blurred in the context of cyber warfare. As the United States collects responsibility for cybersecurity at a national level under the unified Cyber Command, a single organization has now assumed responsibility for defending domestic Internet infrastructure and cyber resources, and deterring or attacking enemies through offensive operations. Moore, Friedman and Procaccia (2010) crafted a game-theoretic model that explores the strategic interactions of actors capable of both attack *and* defense.

The vulnerability-stockpiles game explores the scenario of a military cyber-organization that discovers a previously undisclosed software vulnerability. Is it better to pass the information to the relevant software vendor and improve everyone's security, or would it be more prudent to keep the vulnerability hidden and develop a zero-day exploit to be saved for an offensive mission against an enemy? In the game, two players look for zero-day exploits. The players' relative technical prowess is captured by a variable parameter. Upon discovery, they can either choose to stockpile or disclose the vulnerability. If one player stockpiles a vulnerability undiscovered by the other, then that player gains and the other suffers. Society also suffers an additional cost because systems remain insecure. If either player chooses to defend, then both players are protected but neither gains as much as they would from a successful attack. If both players discover the vulnerability and choose to stockpile, then neither wins, but society still suffers.

A range of equilibria is possible, depending on the relative sophistication of the players and the extent to which they internalize the social cost of insecurity caused by undisclosed vulnerabilities. If the players ignore all social costs of undisclosed vulnerabilities, then they choose to pursue aggressive stockpiling regardless of their technical prowess. As more social costs are internalized, however, the weaker player chooses to defend rather than stockpile. We anticipate that many future analyses will continue to examine trade-offs between attack and defense.

3.1.2 Empirical Research

The academic literature on security economics now has a reasonable number of analytical models. Recently, researchers have been looking to confirm them via empirical evidence of attacker adaptation as well as cooperation between defenders.

Moore and Clayton (2007) collected data on one attack-defense arms race, namely the removal of phishing websites that impersonate banks. Phishing is the criminal activity of enticing people to visit websites that impersonate genuine bank websites and dupe visitors into revealing passwords and other credentials.[4] One common countermeasure is to remove the imitation websites promptly.

The authors monitored several thousand websites, finding heterogeneity at several levels. While many banks were impersonated, some banks were targeted much more frequently than others. While the median time to remove phishing websites was 20 hours, they found that take-down speed was highly variable. In fact, the removal times for phishing websites matched a skewed lognormal distribution: most websites were removed within a few hours, but a substantial minority slipped through the cracks and survived for many weeks. The variation did not appear to be random, however. Instead, attackers systematically identify weak targets far faster than defenders can plug the holes.

[4] Although a wide range of companies have been subject to phishing attacks, most are financial institutions; for simplicity, we use the term "banks" for firms being attacked.

Moore and Clayton (2007) described how one leading group of online fraudsters, the rock-phish gang, operates. Periodically, the gang picks a new domain name registrar and registers hundreds of web domains, using false names and stolen credit cards. As registrars will take down a domain name quickly if it looks abusive – such as a misspelling of a bank's name – the rock-phish gang chooses innocuous domain names. It takes time for banks to convince a naïve registrar that an ordinary-looking domain is being used for criminal purposes.

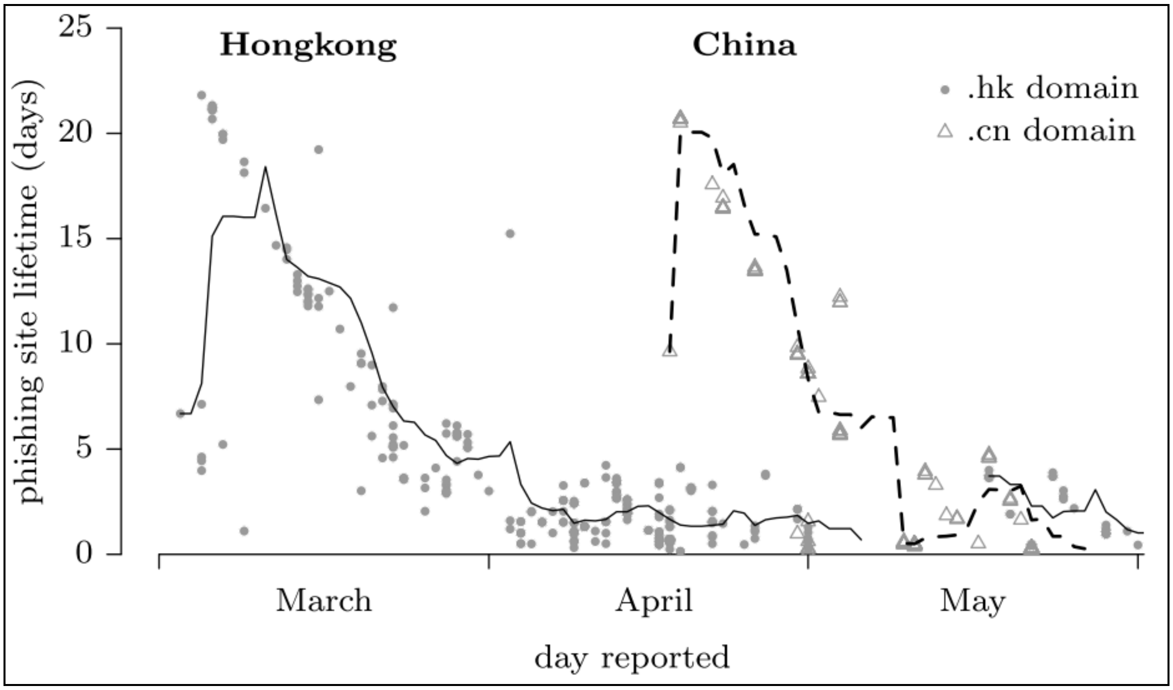

Figure 2: Take-down latency for phishing attacks targeting different registrars in Spring 2007.

Figure 2 presents scatter plots of phishing website lifetimes based on data reported by Moore and Clayton (2007). Both .hk (Hong Kong) domains (left) and .cn (China) domains (right) lasted much longer in their first month of use than in later months. The gang targeted Hong Kong domains first in March 2007, followed by Chinese domains in May 2007 after the Hong Kong authorities wised up. Other researchers also documented how websites hosting malware move from one registrar to the next (Day et al., 2008). These empirical findings inspired the analytical iterated weakest link model of Böhme and Moore (2009) described previously.

Unfortunately, it is not only the domain registrars that struggle to cooperate when defending against attacks. Although some banks deal with phishing website removal in-house, most hire specialist take-down companies. Such companies — typically, brand-protection firms or information security service providers — perform two key services. First, they are good at getting phishing websites removed quickly, having developed relationships with ISPs and registrars across the globe and deployed multilingual teams at 24/7 operation centers. Second, they collect a more timely and comprehensive listing of phishing URLs than banks normally can.

Most take-down companies view their URL feeds as a key competitive advantage over banks and competitors. However, recent work has shown that their feeds have large gaps in coverage. Moore and Clayton (2008) examined six months of aggregated URL feeds from many sources, including two major take-down companies. They found that up to 40 percent of phishing websites remained unknown to the company with the take-down contract despite being known to others. A further 29 percent were discovered by the responsible take-down company only after others had identified them. By measuring the missed websites' longer lifetimes, Moore and Clayton estimated that 120 banks served by these two companies collectively risk at least US$330 million per year by failing to get their take-down contractors to share information.

Commercial take-down contractors cross borders with impunity, but life is not so simple for law enforcement. A generation ago, the adoption of the motor car from the 1930s to the 1950s meant that a burglar could steal from a house in a town where he was not known to the police, and be home in time for breakfast; that led to radical changes in police organization and methods. Online crime mostly consists of industrial volumes of petty offenses, committed by perpetrators in one country against victims in another. International police cooperation was simply never designed to cope with this. The relevant international agreement, the Council of Europe Convention on Cybercrime, is fairly limited in scope. However, Wang and Kim (2009) do find evidence of cooperation among treaty signatories. They find that the average number of cyber attacks originating from countries after joining the treaty falls between 15-25%. While the reduction could conceivably be attributed to explicit cooperation between signatories, the more likely explanation is that signing the treaty signals that a country has committed resources to fighting cyber crime.

A few empirical studies have examined how attackers select their targets. Attackers compromise web servers in order to host fraudulent content, such as malware and phishing websites. Moore and Clayton (2011) found that many attackers use search engines to identify potentially vulnerable hosts to target. Attackers craft targeted search terms that pick out a particular program, or version of a program, which the attacker can subvert. For instance, the search term "phpizabi v0.848b c1 hfp1" returns websites running software suffering from an unrestricted file upload vulnerability. Using logs from websites compromised by phishing attackers, they found that at least 18% of website compromises are triggered by the use of carefully crafted search terms.

They also found that 19% of phishing websites are recompromised within six months, and the rate of recompromise approximately doubles if the website has previously been identified through web search. These findings provide additional support to the conception of attackers pursuing the "weakest target".

Ransbotham (2010) studied the diffusion of vulnerabilities in open and closed source software. There are long-running debates over which approach yields software with fewer vulnerabilities. Open source software benefits from inspection by a wider community of testers, while closed source restricts access to those who might uncover flaws, regardless of motive. Ransbotham steers clear of this argument and instead tries to answer the question whether attackers target open source software more than closed source for attacks due to its relative ease of access. By looking at two years' worth of intrusion detection alerts, he finds that open source software is targeted more frequently, and furthermore that attackers convert vulnerabilities into exploits faster for bugs affecting open source software.

3.1.3 Behavioral Research

Information security professionals know that the user is often the weakest link. The famous hacker Kevin Mitnick managed to break into telephone exchanges using "social engineering" – tricking people to give away passwords (Mitnick 2002). Similarly, phishing attacks exploit many people's lack of ability to tell genuine websites from forgeries. One area that is beginning to be explored is whether there are psychological and behavioral explanations for why some people fall for internet attacks and others do not. Sheng et al. (2010) analyzed the demographics of who falls for phishing attacks using an online survey. They found that women were more likely to fall for phishing scams than men, as do people aged 18-25 when compared to the general population.

Motivating end users to improve their security practices requires a mixture of psychology and economics. Adams and Sasse (1999) challenge the notion that users are being careless when they flout recommended rules for choosing safe passwords. Instead, they find that people reject policies incompatible with their work procedures, and they lack awareness of threats. Herley (2009) goes further: he argues that users rationally reject security advice, because the burden of complying with security procedures often outweighs any gain in protection. For instance, if everyone took the time to read the URLs of all websites they have visited to look for phishing attacks, the aggregate loss in productivity would outweigh the money actually stolen.

Many end users do choose to defend themselves, though. Wash (2010) interviewed people to find out why, and found a number of persistent "folk" models that informed their actions. People distinguish between "viruses" and "hackers". In the first mental model, one has to download the virus intentionally and run it; in the second model, one can be caught out by visiting shady sites or opening shady emails. The user's dominant mental model directly influences which countermeasure is selected: buying antivirus software, or taking care about which websites are visited. Wash argues that regardless of whether the models are "correct", they should be used as a basis for policy and design.

End users are not the only defenders whose motivations affect outcomes. The lifetimes of illegal websites depend on who has an incentive to do something about them. According to Moore and Clayton (2008), phishing websites are removed faster than just about any other form of illegal online content. This is because banks are highly motivated to remove any website that impersonates them. By contrast, other illegal activities such as online pharmacies do not appear to be removed at all.

It turns out that most banks' strategies are suboptimal. They focus on removing only those websites that attack them directly, but ignore a key component of the phishing supply chain: mule recruitment. Phishermen recruit "money mules," dupes who launder stolen money to accounts under the criminals' control. The mules receive bank transfers from the victims' accounts and pass them on to the phishermen using a non-revocable payment method, typically Western Union. Because the fraudulent bank transfers are often reversed, the mule often ends up out of pocket rather than the bank – so banks lack an adequate incentive to crack down on mule recruitment. There is also a collective-action problem: it is hard to tell which bank will suffer from any given mule-recruitment campaign.

Moore and Clayton report that mule-recruitment websites stay online for two weeks on average, compared to only a few hours for phishing websites. Even though mule recruitment harms banks directly, and appears to be one of the bottlenecks in the phishermen's operations, not one of the banks or take-down companies pursues them actively. Typically, only vigilante groups such as “Artists Against 419” try to take these sites down, and even they treat them as low priority because they see the mules as complicit . If the banks were acting rationally, their trade associations would spend a lot more on taking down the muleteers.

3.1.4 Research and Policy Agenda

Empirical work on attack and defense has received far less attention in the literature than analytical models have, so there is ample scope for more. For example, the spread of security-breach disclosure laws (described next) are starting to make available the datasets needed to test the theories.

Many other good research opportunities may come from studying particular applications. This is why the remaining sections of the paper discuss problems such as malware and payment system security.

3.2 Breaches of Personal Information

A flurry of privacy breach-notification laws have been adopted in 44 states, led by the state of California in 2002[5]. Both public and private entities must notify affected individuals when personal data under their control have been acquired by an unauthorized party. The law was intended to ensure that individuals are given the opportunity to protect their interests following data theft, such as when 45 million credit card numbers were stolen from T.J. Maxx’s information technology systems[6]. Breach-disclosure laws are also designed to motivate companies to keep personal data secure. Researchers have extensively analyzed the effectiveness of breach-disclosure legislation, using both theoretical and empirical approaches.

3.2.1 Analytical Research

Information disclosure has two primary motivations. First is the view, articulated by Louis Brandeis, that “sunlight is the best disinfectant”: bringing unfortunate events to light can push firms to clean up their act. Second, disclosure can be motivated by a sense of the community’s “right to know”. The Emergency Planning and Community Right-to-Know Act of 1986 forced manufacturers to disclose to the EPA (and, consequently, the public) the amount and type of toxic chemicals released into the environment. The aggregated data, known as the Toxic Release Inventory (TRI), has been effective in reducing the amount of toxic chemicals discharged into the environment (Konar and Cohen 1997). The success of the TRI helped motivate the state of California to adopt the breach-notification legislation.

[5] California Civil Code 1798.82

[6] http://www.businesswire.com/portal/site/home/permalink/?ndmViewId=news_view&newsId=20071130005355

Romanosky and Acquisti (2009) study the circumstances under which information disclosure is a more effective policy remedy than assigning ex-post liability to the responsible firm or using ex-ante safety regulation. Mandatory notification of data breaches should reduce the incidence of data breaches for two reasons. First, it incentivizes firms to invest in countermeasures that reduce the likelihood of a breach. Second, it helps consumers affected by a data breach take precautions that mitigate the effect of disclosure (e.g., credit-report monitoring). Unlike ex-ante safety regulation, data breach disclosure does not compel companies to invest in any countermeasures. And unlike ex-post liability, firms do not have to compensate consumers whenever they expose personal information. Both factors might lead to a higher incidence of breaches under a disclosure-only regime.

In their analysis of mandatory disclosure, Romanosky and Acquisti balance the positive effects of transparency and consumer empowerment in reducing breaches against negative costs of consumer actions and firm compliance, customer service and reputation damage. They conclude that no policy intervention is strictly better than the others, and that each has drawbacks. The authors call for the collection of more accurate information on the frequency of violations and the harms imposed in order to determine the optimal policy response. Romanosky, Sharp and Acquisti (2010) propose an economic model that formally examines when mandatory information disclosure is the optimal outcome. They find that the higher costs imposed on firms by disclosure regimes can trigger lower overall social costs by inducing firms to increase their investment in care. They also find that when firms are only required to compensate consumers minimally, the added costs of complying with disclosure are necessary to reduce social costs.

Given the risks of data breaches, one may wonder why firms choose to retain personal data at all. In fact, the collection of personal data is often designed to reduce a different type of risk: payment fraud. Companies often argue that the collection of personal information (address information, purchase histories, biometric information, etc.) helps them determine whether an attempted purchase is legitimate or fraudulent. Roberds and Schreft (2009) construct a model where firms trade off the risk of payment fraud and data breaches. They find that when participants cannot cooperate on protecting personal information, they choose to overcollect personal information and underinvest in protecting it.

3.2.2 Empirical Research

Breach-disclosure legislation has created an extensive source of public data on security incidents. Several researchers have taken advantage of it to carry out promising analyses; they have often found evidence that the information disclosure requirement has both punished violators and reduced harm compared with when breaches did not have to be reported.

Campbell et al (2003) conducted an event study on information security breach reports appearing in popular newspapers prior to the adoption of breach-disclosure laws. They found that the stock prices of listed companies suffering breaches were more likely to fall when the breach involved a loss of confidential information. Acquisti, Friedman, and Telang (2006) subsequently conducted a study examining only privacy breaches after breach-notification laws became common. They found a statistically significant negative impact on stock prices following a reported breach. Meanwhile, Romanosky, Telang, Acquisti (2008) examined identity theft reports obtained from the FTC from 2002 to

2007. Using time differences in the adoption of state breach disclosure laws, they found an average 6.1% reduction in fraud rates following each state's adoption.

Why is the reduction so small? Firms report a much greater awareness to the risk of losing personal information and claim to have increased investment in precautions such as hard drive encryption (Mulligan and Bamberger 2007). A plausible explanation is that inadvertent data leakage by firms is but one cause of fraud, so mandatory data breach disclosure can only be a partial solution. Fraud can occur in many other ways, particularly via online criminal networks that steal credentials from consumer computers (Moore et al. 2009). Obtaining meaningful data on dependent variables is often difficult, with security incidents as elsewhere in econometrics.

Consequently, researchers must often combine disparate data sources to answer questions empirically. For instance, the breach disclosure laws adopted by many states offer an exclusion whenever the lost data has been encrypted. This aims to encourage companies to take reasonable precautions. But is it sound policy? Miller and Tucker (2010) pull data from several sources to answer this question for the U.S. health sector. They combine reports of privacy breaches at hospitals with a database indicating whether particular hospital systems use encryption to protect patient data. Their conclusion is striking: disk encryption does not reduce data loss. In fact, the use of disk encryption appears to induce careless behavior by employees leading to a higher incidence of data breaches. They also find that the use of extensive database technologies for managing electronic health records is correlated with higher rates of data loss. These results challenge the wisdom of breach-disclosure laws offering exclusions for encryption.

3.2.3 Behavioral Research

Do consumers take rational countermeasures when they receive a notice of a breach? Romanosky and Acquisti (2009) outline several behavioral issues for consumers in reacting to breach notifications. First, it is often impossible for consumers (or firms for that matter!) to correctly assess their risk following notification of a breach. So informing them may not always trigger the right response. What happens if they stop buying goods online? Such an overreaction is certainly possible. Second, breach notification could impose significant transaction costs on consumers, prompting them to call their bank or the companies that leaked the information. Third, consumers might not be rational in their responses: they cannot be expected to account for all possible outcomes of data disclosure and accurately characterizing the risks of each.

3.2.4 Research and Policy Agenda

Empirical research on privacy breaches is likely to continue, as they are currently the best source of public data on information security incidents. Measuring the effect of breach notifications on firm behavior is one promising area, and we also need to understand the effect on consumers better.

It is also natural to wonder if other types of information security incident might be dealt with via mandatory notification. In other information security incidents, it is often firms that lack information rather than consumers. If a bank does not disclose that several business customers have lost millions of dollars due to the compromise of online banking credentials, its other business customers remain ignorant to the need to take precautions. Might information asymmetries across firms be tackled using the same mechanisms as between consumers and firms?

One specific area that might benefit from mandatory disclosure of incidents is online banking and payment card fraud (discussed at greater length in Section 3.4). Currently, most countries do not collect and publish the data needed to answer very basic questions: is online identity theft increasing or decreasing? Are online banking and e-commerce less safe than transactions in the real world? The main argument in favor of public disclosure is that the financial industry does not internalize all the costs of insecurity. In some countries, such as the UK and South Africa, banks often refuse to reimburse fraudulent transactions, blaming the consumer for "carelessness" instead. Meanwhile, US consumers have perhaps the world's strongest protection in this respect, in which Regulations E and Z limit liability for unauthorized transactions. But businesses do not experience the same level of protection, and merchants are expected to share responsibility for covering the costs of fraud. Because banks do not internalize all these costs, the public deserves a fair and transparent accounting of who pays what share. Disclosing aggregated loss figures, as well as a breakdown of the number and average loss of incidents for both consumers and businesses, may be a reasonable policy intervention. Such information can help answer questions such as how many people experience online fraud, whether any groups pay a disproportionate share, and whether this changes over time.

3.3 Malware and Botnets

Malware is frequently used to steal passwords and compromise online banking, cloud and corporate services. It is often used to place infected computers into a "botnet": a network of thousands or even millions of computers that is used to carry out a wide range of wicked services under the control of an attacker. The services include sending spam, committing online-advertising fraud, launching denial-of-service attacks, hosting phishing attacks, and anonymizing attack traffic.

Many botnets are used to simply send spam. For example, the Reactor Mailer botnet ran from 2007-2009, at its peak sending more than 180 billion spam messages per day, 60% of the world's total (Stern 2009). At least 220,000 infected computers participated each day. The Zeus botnet, by contrast, includes key logger software to steal online credentials; it is estimated to be as large as 3.6 million computers[7]. Botnets can also be used to carry out denial-of-service attacks, where, the operator directs the bots to make connections to the same target website and overload it. For example, botnets were employed to carry out the denial-of-service attacks in Estonia[8] and Georgia[9].

3.3.1 Analytical Research

Malware is often designed to harm others rather than its host. Camp and Wolfram (2004) observed that information security is characterized by externalities: infected computers are brought under the control of attackers who use them to carry out scams. The Internet's global addressability and the concentration in operating systems and application software let criminals launch relatively unsophisticated, undirected attacks against everyone. They pick off only the weakest targets, such as PCs running out-of-date

[7] http://www.computerworld.com/s/article/9177574/Big_botnets_and_how_to_stop_them

[8] http://www.wired.com/politics/security/magazine/15-09/ff_estonia?currentPage=all

[9] http://www.zdnet.com/blog/security/coordinated-russia-vs-georgia-cyber-attack-in-progress/1670

software; but there are enough of these for criminals to build botnets comprised of hundreds of thousands of computers.

Other attackers tailor their attacks to a particular high-value target. For example, criminals send emails with malicious attachments to people working in the payroll departments of small businesses, using the correct name of the person and firm[10]. Once the attachment is downloaded, it installs a keystroke logger that steals online banking credentials. The same techniques were used to install malware in computers in the Dalai Lama's private office and siphon off sensitive data to China. There is a big gap between high-volume, scalable attacks and targeted ones, which Herley (2010) explores using a simple economic model. Scalable attacks work even when only a few actually fall for the scam. Non-scalable attacks work when the distribution of target value is highly skewed, and attackers can select juicy targets. Herley argues that as wealth, fame and political influence all follow a power-law distribution, only those individuals on the tail of the distribution should worry about being targeted.

Reducing the proportion of computers susceptible to malware and cleaning up computers that have been infected requires coordination. Responsibility should rest with the least-cost avoider. Assigning indirect intermediary liability to a third party may be necessary. According to Lichtman and Posner (2004), a number of conditions can make indirect liability attractive. First, the bad actors could be beyond the reach of the law, either because they cannot be identified or because they could not pay up even if caught. Second, high transaction costs could make it infeasible to design contracts that allocate responsibility directly. If either of these conditions is met, two additional factors should be considered. First, indirect liability is attractive when a third party is in a good position to detect or prevent bad acts. Second, indirect liability is useful when the third party can internalize negative externalities by reducing the incidences of bad acts.

Lichtman and Posner argue that these conditions hold for ISPs in the context of information security. (While the authors consider information security in general, their logic can be applied to the particular case of malware cleanup.) First, the miscreants actually carrying out the infections are typically beyond the reach of the law. Second, ISPs are in a good position to detect and clean up computers infected with malware. They can often detect that computers are infected by inspecting network traffic. They regulate a computer’s access to the broader Internet. Finally, they alone can associate network traffic with customer contact details. Consequently, ISPs are in an unrivaled position as intermediaries. Lichtman and Posner argue for ISPs to take on strict liability for the actions of their customers’ computers. In other words, they suggest simply making the ISPs liable for malware-infected customers, and let them choose how they carry that burden.

3.3.2 Empirical Research

Computer security researchers have recently been documenting the actions of online criminals. One approach is to study reports of transactions between criminals. Thomas and Martin (2006) studied online black markets which trade stolen credentials. Criminals take on specialized roles, from those who discover vulnerabilities to those who cash out and launder compromised bank accounts. Thomas and

[10] http://www.bankinfosecurity.com/articles.php?art_id=1732

Martin describe a horizontally integrated market of specialists that focus on particular comparative advantages. Phishermen create and maintain fake bank websites. In turn, they hire spammers who send out emails with links to the fake banks. The spammers contract with botnet herders, to send fake emails. Once credentials are harvested, phishermen recruit money mules, dupes who launder stolen money from victim accounts. Franklin et al. (2007) monitored the public chat channels used by online criminals to contact each other and gathered extensive data on credit card fraud, spamming, phishing, and the sale of compromised hosts. They found, for instance, over 100,000 unique credit card numbers advertised on the channels over 7 months.

Other researchers have empirically measured the prevalence of particular classes of attacks. The "drive-by-download" is a common method of malware installation, where miscreants compromise popular websites so that when an unsuspecting user visits the website, it secretly downloads and installs malware on to her computer. In one study, researchers at Google found 3 million drive-by-download URLs, and furthermore that 1.3% of Google's incoming search queries return at least one drive-by-download link in their results (Provos et al. 2008). According to Rajab et al. (2010), around 15% of malware distributed this way installs fake antivirus software that prompts users with endless warnings to pay up or risk attack. Moore and Clayton (2007) studied phishing websites that impersonate banks. They found that a substantial proportion of phishing websites are hosted on botnets. By studying the logs of websites that were compromised and loaded with phishing pages, they estimated that approximately 800,000 people worldwide fall victim to phishing annually. Using a different methodology (an optional toolbar add-on for Internet Explorer that counted duplicate logins on different websites), Florencio and Herley (2007) estimated approximately twice the number.

Finally, some researchers have begun to take a more active role in measuring malicious activity online. For example, Kanich et al. (2008) infiltrated a large botnet and altered the spam emails sent out so that they linked to a benign duplicate website under the researchers' control. They answered a long-standing question: how many people respond to spam? They found that only 28 sales resulted from 350 million spam emails advertising pharmaceuticals – a conversion rate of 0.00001 percent. Meanwhile, Motoyama et al. (2010) studied the market for solving CAPTCHAs – the mechanisms that separate humans from computers by requiring translation of jumbled characters. They collected data on the business model of human CAPTCHA solvers by both purchasing CAPTCHA-solving services and participating in CAPTCHA-solving. The market-clearing price for solving 1,000 CAPTCHAs has steadily fallen from $10 in 2006 to $.50 today. According to the authors, at these prices, paying people in poor countries is cheaper than relying on image-processing software.

Such innovative surveillance of the underground economy will continue to be necessary. However, it raises a number of legal and ethical questions. For instance, in the spam study, the authors were careful to ensure that the only spam they monitored would have been sent out anyway. By participating in CAPTCHA solving, the researchers might have fallen foul of the US Computer Fraud and Abuse Act; the researchers argue that they did not because they did not make use of the solved CAPTCHAs. In any event, researchers engaged in active data collection must tread carefully. For a broader discussion of the ethical challenges facing computer security researchers, see Dittrich et al. (2010).

Most of the research just mentioned has been descriptive in nature. Only recently are we beginning to see econometric analyses of collected empirical data. Van Eeten et al. (2010) present empirical evidence that ISPs are indeed important intermediaries. They use the amount of observed email spam emanating from ISPs as a proxy for botnet activity, which is reasonable given that nearly all email spam is sent by botnets. They relied on two main data sets: a record of 138 million IP addresses observed to be sending spam from 2005-2008, and a mapping from IP addresses to the 200 largest ISPs in 40 countries. They found high variability in ISP performance: the 10 worst offenders accounted for 30% of all spam sources. Yet the explanation for such poor performance is not simply size: controlling for the size of the customer base, large ISPs do a better job of cleaning up botnets than small ISPs. This could be because large ISPs can invest in automated detection and remediation that have high fixed costs but low marginal costs.

On a national level, they find evidence that countries participating in cooperative defensive agreements experience fewer infections. Countries that have signed up to the London Action Plan, a 2004 agreement among 27 countries worldwide to foster cooperation among law enforcement in fighting spam, suffer 12 percent fewer infections on average. Similarly, signing up to the Council of Europe Convention on Cybercrime is negatively correlated with spam activity. This corroborates earlier research by Wang and Kim (2009), who found evidence that CoE Cybercrime Convention signatories have fewer reports of malicious activity observed by the SANS Internet Storm Center.

3.3.3 Behavioral Research

Malware authors certainly consider human behavior when crafting their exploits, from phishing attacks that take advantage of our inability to reliably distinguish legitimate websites from forgeries, to fake antivirus software that scares people into sending money to criminals. While many scams are applied globally, some appear only in particular cultures. Christin et al. (2010) describes "one click frauds" that target Japanese users. While visiting a pornographic website, a user may be presented with a link that, once clicked, will explain that the user has entered a binding contract that must be paid. Analyzing a sample of 2,000 public reports of fraudulent attempts over more than three years, the authors find that the top 8 criminals account for over half of all scams. Each scam typically asks for around $500 from each victim, not coincidentally the typical amount of "pocket money" Japanese salarymen are allocated from the monthly household budget. By constructing a simple economic model, they estimate that the average attacker earns upwards of $100,000 per year using the scams. Notably, the authors find that the scam is completely local to Japan, and that the infrastructure used to host the scams is not used for any other online frauds.

3.3.4 Research and Policy Agenda

Worldwide, ISPs in many countries are experimenting with different ways of enlisting in the fight against malware. In the US, Comcast now automatically notifies customers when they are infected via a browser pop-up that links to removal instructions[11]. The scheme relies on customers to clean themselves up, which sometimes works on types of malware detectable by automated tools such as Microsoft's MSRT. Inevitably, though, malware is often not removed by users after they have been notified. For these cases, Comcast has partnered with Symantec to offer a remediation service by a

[11] http://security.comcast.net/constantguard/

skilled technician for $100. Australian ISPs recently announced a notification-based effort for all its ISPs[12].

Another ISP-based option is to place infected computers into "quarantine". Once in quarantine, users are required to download and install anti-virus software and malware removal tools. They are then only permitted to rejoin the wider Internet once the security software is installed and the computer passes a network-based scan for malware. Quarantine is considerably more expensive than notification-only-based interventions, because more customer-support calls are made. Some ISPs use quarantine systems only for a minority of affected customers. Recently Dutch ISPs announced a signed agreement to notify and quarantine affected customers[13]. This collective action is designed in part to allay the fear that customers might switch providers rather than disinfect their machines.

Yet the most common response for an ISP to notification that customers are infected with malware is to do nothing. Why? The incentive to intervene is weak (van Eeten and Bauer 2008). ISPs are not affected that much by infected customer machines, apart from perhaps being chided by their peers if their network is sending too much spam; but they face real costs by intervening. The big-ticket item is the cost of dealing with the phone calls that come in after sending out notices or placing customers into quarantine. So many ISPs prefer to keep quiet.

Government intervention is being considered in some countries. If the cost of customer support is truly the greatest impediment, then the German government's decision to coordinate and subsidize a nationwide call center seems like a reasonable response (Karge 2010). Under this plan, ISPs will identify infected customers and pass along the information to the call center. Clayton (2010) describes a proposal under consideration by Luxembourg to subsidize the cost of voluntary cleanup whenever a customer has been notified of infection. In contrast to these "carrot"-based incentives, "sticks" might also be tried. Anderson et al. (2008) recommended that the European Commission introduce fixed penalties for ISPs that do not expeditiously comply with notifications of compromised machines present on their networks. Fixed penalties offer two distinct advantages over assigning strict liability: they reduce the burden on the victim to demonstrate losses and they reduce uncertainty on the responsible party for the amount of damages owed for misconduct.

Another way in which policy makers may coordinate defense is in the aggregation of infection notifications. When surveyed, Dutch ISPs revealed that they only notified and/or quarantined around 10% of infected customers (van Eeten et al. 2010b). The ISPs claimed that they had notified all customers that they knew were infected. As it happens, their lists of infections were very incomplete. Data incompleteness is a widespread problem in information security (Moore and Clayton 2008b), as firms often jealously guard any collected incident information as trade secrets. To combat this trend, the Australian Internet Security Initiative now aggregates data on compromised machines into a feed and passes it along to Australian ISPs.

There are many great opportunities for enterprising researchers to study the effects of different policies

[12] http://iia.net.au/images/resources/pdf/iiacybersecuritycode_implementation_dec2010.pdf

[13] http://www.darkreading.com/blog/archives/2009/09/dutch_isps_sign.html

and to identify what works best for different circumstances.

3.4 Payment System Security

Online crime is mostly monetized by abusing the financial system. So any study of Internet security must also consider the payment system. Payment network operators (e.g., Visa and MasterCard) involve many parties, each with different responsibilities, protections and obligations. To complete a transaction, four entities are typically involved: the cardholder, the cardholder's bank (also known as the *issuer*), the merchant, and the merchant's bank (also known at the *acquirer*). When a cardholder makes a purchase, the issuer authorizes the transaction and settles with the acquirer, who pays the merchant. If a transaction turns out to be fraudulent, responsibility depends on the circumstances. We discuss the different outcomes in detail below; where responsibility falls has a huge effect on the incentive to protect against fraud.

3.4.1 Analytical Research

Payment networks are examples of a two-sided market, where two types of users are served by a common platform: cardholders and merchants. Cardholders choose the payment network that serves the most merchants, while merchants only want to accept cards that serve the most customers.

A considerable academic literature has emerged on two-sided markets in general, and the study of payment networks in particular. Much of the effort has focused on the socially optimal interchange fees that should be paid by the acquirer to the issuer (Rochet and Tirole 2006). Additionally, competition among platforms does not necessarily improve social welfare (Chakravorti 2010). For instance, competition for cardholders can lead issuers to offer substantial rewards and other incentives; but issuers may then raise interchange fees to acquiring banks, which in turn raises merchant fees. So merchants can end up subsidizing lavish reward programs for a minority of cardholders – and insecure card-issuer practices. For these reasons, the recent US financial regulation bill empowers the Fed to cap interchange fees.

What are the security implications of the two-sided market structure of the payment system? First, payment networks have cultivated very successful credit card platforms with millions of participating merchants and cardholders. The value of this existing user base is enormous, and presents a significant barrier to new entrants. Having already invested heavily in a less secure payment technology and achieved market dominance, existing payment networks may be reluctant to invest further in security. In other words, payment security is difficult for a new entrant to improve. It is also hard to sell new technologies to incumbents: the adoption of EMV ("chip and PIN") smartcard technology in Europe, Canada and elsewhere took over twenty years of lobbying by the smartcard industry, and extensive government subsidy of development costs.

Second, there are many difficulties in assigning responsibility for reimbursing and protecting against fraud (MacCarthy 2010). In the US, consumers are protected from liability for unauthorized charges on their accounts (credit cards are covered by the Truth in Lending Act of 1968, implemented by the Federal Reserve as Regulation Z, while debit card holders are covered by the Electronic Funds Transfer Act, implemented through Regulation E). Instead, the obligation to repay is allocated between banks and merchants. For frauds occurring in face-to-face transactions, issuers normally foot the bill, rather than merchants. For card-not-present transactions (online and phone), however, the merchant often has to pay. Banks and merchants have continued to fight over the details. The Payment Card System Data Security Standard (PCI DSS) is designed to improve the security of payment systems at merchants; merchants who fail to get PCI accreditation are assigned liability for fraud. However, when fraud still happens, merchants or their payment processors can lose accreditation retrospectively. Merchants complain of the high costs of compliance and argue that PCI DSS is nothing more than a thinly veiled, bank-led liability shift from issuers to merchants.

In other countries, consumers are less well protected. Several years ago banks in Europe adopted the EMV protocol, in which each bank card becomes a smartcard that verifies PINs entered by the cardholders. One consequence of this upgrade has been a liability shift. Disputed transactions are blamed on consumers whenever the bank believes the PIN was used, and otherwise on the merchant. EMV was expected to reduce fraud, but fraud has in fact risen. One reason may be that the banks, which operate the payment system and are thus best placed to prevent fraud, now have less incentive to work hard at this.

3.4.2 Empirical Research

Banks in Spain, Britain, the Netherlands, France and Australia regularly disclose aggregate information on payment card fraud. In 2009, for example, UK banks lost ₤440 million (approximately $641 million) due to all forms of payment fraud, while ₤59.7 million ($87 million) was attributed to online banking in particular[14]. US banks do not regularly publish fraud losses. While the information disclosed is not directly comparable, Sullivan (2009) examined the published data and, in an attempt to come up with comparable national statistics, pieced together several third-party estimates to arrive at an estimate for US fraud as well (Sullivan 2010). He found that Spain and Australia experienced the lowest rates of fraud in 2006, only $.022 and $.024 per $100 of transactions respectively, whereas the UK and US fared far worse, losing $.086 and $.092 respectively. What might explain the differences? First, Spanish and Australian payment networks work hard to identify fraud by studying transaction histories (Spain was the first country to move to universal online authorization of credit card transactions). Second, in the US and UK, more consumers purchase goods online, where transactions are riskier.

Another potential explanation is the security of payment cards themselves. Here, though, the results are counterintuitive. Europe has adopted the EMV protocol, which is more secure technically than the magnetic strip system used in the US. Yet overall fraud levels did not fall in the UK after the introduction of EMV in 2005, but actually rose from 2005–8. As described in Section 3.1.1 and illustrated in Figure 1 (right), criminals have shifted their tactics to weaker points in the payment system, such as foreign ATMs and card-not-present transactions. Additionally, EMV was developed by the payment networks in a closed fashion and not subjected to an open review process. Outside researchers have subsequently identified a number of significant flaws in the UK implementation. This includes finding that terminals which had been certified as secure turned out not to be, triggered by incentive failures in the certification regime (Drimer et al. 2008). Additionally, researchers discovered an attack which enables cards to be verified without entering the correct PIN (Murdoch et al. 2010). Since 2008 there has been a

[14] http://www.paymentsnews.com/2010/03/uk-card-and-banking-fraud-losses-down-28-in-2009-to-4403mm.html

slight fall in fraud reported by banks but this appears to be the consequence of liability being shifted to merchants and cardholders.

3.4.3 Behavioral Research

Given the complex interactions between the stakeholders in payment networks, it is important to consider the knock-on effects of liability rules. By offering strong protections to consumers, Regulations E and Z helped stimulate adoption of credit and debit cards. Some observers worry about moral hazard: cardholders acting carelessly because they know they will be reimbursed for fraud (Douglass 2009). Unfortunately, no empirical research has been conducted to test the claim. It is certainly tempting for payment networks to try to trim their obligations now that cardholder adoption is so high. But would this risk systemic failure? There is empirical evidence of moral hazard by banks when consumer protection is weaker. Anderson (1993) reported that UK banks suffered higher costs due to ATM fraud and security measures than US banks at a time when UK consumer protection was notably weaker. After UK regulators allowed banks to deny consumer claims of ATM fraud, the banks became careless; consumers experienced more fraud and eventually banks had to pay more too. We may be going round that circuit once more and on a larger scale now that EMV has given European banks the confidence to be more aggressive in denying liability.

3.4.4 Research and Policy Agenda

Sullivan (2009) calls for the periodic public disclosure of payment card fraud rates in the US, as is already being done in several countries; Anderson et al (2008) call for disclosure to be mandated in all Member States of the EU. As mentioned in Section 3.2.4, there are several benefits to disclosure. First, it lets policy makers scope the problem of online payment fraud, and see whether it is getting better or worse. Second, disclosure can uncover which parties are bearing the cost of fraud. Sullivan (2010) estimates that in the US, card issuers bear the brunt (59% to the merchants' 41%). In France, by contrast, losses were more evenly shared between issuers and merchants. Given compliance rules that put a greater responsibility on merchants to reimburse fraud, a transparent accounting of each party's share is necessary. Further research is needed to empirically test claims by merchants that efforts such as PCI DSS favor issuers.

Finally, the amount consumers and businesses are paying should also be tracked, especially in countries where they may be liable. Researchers could attempt to measure the effects of moral hazard on cardholders, if any exists. There are also opportunities to study the demographics of victims of fraud. Some evidence from the UK (Lea et al. 2009) suggests that fraud victims who end up out of pocket tend to come from at-risk groups, while other surveys show the wealthy suffer more fraud overall (Moon et al.,2010).. A rigorous examination of the data on victims could be valuable to policymakers.

Merchants and card issuers find it difficult to coordinate their defensive investments. Competing proposals for security improvements are being touted, from adopting EMV in the US to mandating merchant use of end-to-end encryption. The US government has its own ideas for improving the security of online transactions[15], but it is not clear how they might succeed with or without existing payment networks. Given the dominance of today's card networks, the outlook for outside innovation to improve

[15] http://pindebit.blogspot.com/2010/04/national-strategy-for-secure-online.html

security of the payment system is not good. The one realistic hope on the horizon is that the introduction in 2011–12 of mobile wallets – payment instruments based on hardware "secure element" (SE) chips in mobile phones – might not just refresh payment technology but also realign incentives. For example, banks and phone companies should logically cooperate to revoke lost or stolen phones and reissue credentials on new ones; banks actually want compromised credentials to be revoked, and phone companies actually want stolen phones to be replaced. However, in the one country where mobile payments have already been deployed (Japan), coordination is not good; so perhaps some industry or even government coordination might be helpful here (Anderson 2011).

4.Conclusion

In this survey, we have discussed how economic analysis has turned out to be an extremely powerful tool for engineers and policymakers concerned with information security, and more generally with the dependability of complex socio-technical systems. Systems tend to fail when the people who defend them are not the people who suffer when they fail. We have discussed in some detail how misaligned incentives, information asymmetries and externalities are rife in the specific case of online scams, malware and frauds against payment systems. We also described several areas of active research in these areas: models of attack and defense; analytical and empirical results for privacy breaches, malware, and payment fraud; and policy options. However this is only the beginning. There are many more systems on which we are all coming to depend. There will be much scope for the lessons of security economics to be applied to the smart grids of the future, to online social networks, and to all sorts of other emerging technological artefacts that can only be made dependable if we can engineer them so that their millions of stakeholders have the right incentives.

References

Acquisti, A., Friedman, A., Telang, R., 2006. Is There a Cost to Privacy Breaches? An Event Study. Proceedings of the International Conference on Information Systems (ICIS).

Adams, A., Sasse, M. A., 1999. Users are Not the Enemy. Communications of the ACM 42(12), pp. 40-46.

Akerlof, G. A., 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. Quarterly Journal of Economics 84(3), pp. 488–500.

Anderson, R., 1993. Why cryptosystems fail. Proceedings of the 1st ACM Conference on Computer and Communications Security, pp. 215—227. Anderson R., 2011. Can We Fix the Security Economics of Federated Authentication? Security Protocols Workshop, Cambridge, March 2011.

Anderson, R., 2001. Why Information Security is Hard—An Economic Perspective. Proceedings of the 17th Annual Computer Security Applications Conference, pp. 358–65.

Anderson, R., Böhme, R., Clayton, R., Moore, T., 2008. Security Economics and the Internal Market. European Network and Information Security Agency. Available at http://www.enisa.europa.eu/act/sr/reports/econ-sec/economics-sec/at_download/fullReport

Anderson, R., Moore, T., 2006. The Economics of Information Security. Science 314(5799), pp. 610–13.

Böhme, R., Moore, T., 2009. The Iterated Weakest Link - A Model of Adaptive Security Investment. Proceedings of the 8th Workshop on the Economics of Information Security (WEIS). Available at http://weis09.infosecon.net/files/152/paper152.pdf

Camp, L. J., Wolfram, C. D., 2004. Pricing Security: A Market in Vulnerabilities. In Economics of Information Security, Vol. 12, Advances in Information Security, ed. L. Jean Camp and Stephen Lewis, pp. 17–34. Boston: Kluwer Academic Publishers.

Chakravorti, S., 2009. Externalities in Payment Card Networks: Theory and Evidence. Proceedings of the 3rd Federal Reserve Bank of Kansas City Payments Conference, pp. 99-124.

Christin, N., Yanagihara, S., Kamataki, K., 2010. Dissecting One Click Frauds. Proceedings of the 17th ACM Conference on Computer and Commmunications Security (CCS), pp. 15—26.

Clayton, R., 2010. Might Governments Clean-up Malware? Proceedings of the 9th Workshop on the Economics of Information Security (WEIS). Available at http://weis2010.econinfosec.org/papers/session4/weis2010_clayton.pdf

Dittrich, D., Leder, F., Tillmann, W., 2010. A Case Study in Ethical Decision Making Regarding Remote Mitigation of Botnets. In: Sion, R., Curtmola, R., Dietrich, S., Kiayias, A., Miret, J., Sako, K., Sebe, F. (Eds.), Lecture Notes in Computer Science (LNCS) 6054, Springer, pp. 216-230.

Drimer, S., Murdoch, S., Anderson, R., 2008. Thinking Inside the Box: System-Level Failures of Tamper Proofing. Proceedings of the IEEE Symposium on Security and Privacy, pp. 281-295.

Florencio, D., Herley, C., 2007. Evaluating a Trial Deployment of Password Re-use for Phishing Prevention. Proceedings of the Anti-Phishing Working Group Ecrime Researchers Summit (eCrime), pp. 26-36.

Fultz, N., Grossklags, J., 2009. Blue versus Red: Towards a Model of Distributed Security Attacks. In: Dingledine, R., Golle, P. (Eds.), Lecture Notes in Computer Science (LNCS) 5628, Springer, pp. 167-183.

Franklin, J., Perrig, A., Paxon, V., Savage, S., 2007. An Inquiry into the Nature and Causes of the Wealth of Internet Miscreants. Proceedings of the ACM Conference on Computer and Communications Security (CCS), pp. 375–388.

Grossklags, J., Christin N., Chuang, J., 2008. Secure or insure? A Game-Theoretic Analysis of Information Security Games. Proceedings of the 17th International World Wide Web Conference (WWW), pp. 209-218.

Herley, C., 2009. So Long, and No Thanks for the Externalities: the Rational Rejection of Security Advice by Users. Proceedings of the New Security Paradigms Workshop (NSPW), pp. 133-144.

Herley, C., 2010. The Plight of the Targeted Attacker in a World of Scale. Proceedings of the 9th Workshop on the Economics of Information Security. Available at http://weis2010.econinfosec.org/papers/session5/weis2010_herley.pdf

Hirshleifer, J., 1983. From Weakest-Link to Best-Shot: the Voluntary Provision of Public Goods. Public Choice 41, pp. 371—386.

Kanich, C., Kreibich, C., Levchenko, K., Enright, B., Voelker, G. M., Paxson, V., Savage, S., 2008. Spamalytics: An Empirical Analysis of Spam Marketing Conversion. Proceedings of the ACM Conference on Computer and Communications Security (CCS), pp. 3–14.

Karge, S., 2010. The German Anti-Botnet Initiative. OECD Workshop on the role of Internet Intermediaries in Advancing Public Policy Objectives. Available at http://www.oecd.org/dataoecd/42/50/45509383.pdf

Konar, S., Cohen, M., 1997. Information as Regulation: The Effect of Community Right to Know Laws on Toxic Emissions. Journal of Environmental Economics and Management 32(1), pp. 109-124.

Kunreuther, H., Heal, G., 2003. Interdependent Security. Journal of Risk and Uncertainty 26(2—3), pp. 231—249.

Lichtman, D.G., Posner, E.A., 2004. Holding Internet Service Providers Accountable. U Chicago Law & Economics, Olin Working Paper No. 217. Available at SSRN: http://ssrn.com/abstract=573502

MacCarthy, M., 2010. Information Security Policy in the U.S. Retail Payments Industry. Proceedings of the 9th Workshop on the Economics of Information Security, Cambridge, MA. Available at http://weis2010.econinfosec.org/papers/panel/weis2010_maccarthy.pdf

Miller, A., Tucker, C., 2010. Encryption and Data Loss. Proceedings of the 9th Workshop on the Economics of Information Security, Cambridge, MA.

Moon D., Flatley J., Hoare J., Green B., Murphy R. Acquisitive Crime and Plastic Card Fraud – Findings from the British crime Survey 2008–9. Home Office Statistical Bulletin 2010, At http://www.homeoffice.gov.uk/rds/pdfs10/hosb0810.pdf

Moore, T., Friedman, A., Procaccia, A., 2010. Would a “Cyber Warrior” Protect Us? Exploring Trade-offs Between Attack and Defense of Information Systems. Proceedings of the 13th New Security Paradigms Workshop (NSPW), pp. 85—94.

Moore, T., Clayton, R., 2007. Examining the Impact of Website Take-down on Phishing. Proceedings of the Anti-Phishing Working Group eCrime Researchers Summit, pp. 1–13.

Moore, T., Clayton, R., 2008a. The Impact of Incentives on Notice and Take-down. In: M. Eric Johnson (Ed.), Managing Information Risk and the Economics of Security, pp. 199–223, Springer.

Moore, T., Clayton, R., 2008b. The Consequence of Non-cooperation in the Fight against Phishing. Proceedings of the Anti-Phishing Working Group eCrime Researchers Summit, pp. 1–14.

Moore, T., Clayton, R., 2011. The Impact of Public Information on Phishing Attack and Defense. Communications & Strategies 81 (to appear).

Motoyama, M., Levchenko, K., Kanich, C., McCoy, D., Voelker, G.M., Savage, S., 2010. Re: CAPTCHAs -- Understanding CAPTCHA-Solving from an Economic Context. Proceedings of the USENIX Security Symposium.

Mitnick, K., 2002. The Art of Deception: Controlling the Human Element of Security. Wiley, New York.

Mulligan, D., Bamberger, K., 2007. Security Breach Notification Laws: Views from Chief Security Officers. Samuelson Law, Technology & Public Policy Clinic, University of California-Berkeley School of Law. Available at http://www.law.berkeley.edu/files/cso_study.pdf

Murdoch, S. J., Drimer, S., Anderson, R. J., Bond, M., 2010. Chip and PIN is Broken. Proceedings of the 31st IEEE Symposium on Security and Privacy, pp. 433-446.

Provos, N., Mavrommatis, P., Rajab, M., Monrose, F., 2008. All Your iFrames Point to Us. Proceedings of the USENIX Security Symposium, pp. 1—15.

Rajab, M.A., Ballard, L., Mavrommatis, P., Provos, N., Zhao, X., 2010. The Nocebo Effect on the Web: An Analysis of Fake Anti-Virus Distribution. Proceedings of the 3rd USENIX Workshop on Large Scale Exploits and Emergent Threats (LEET).

Ransbotham, S., 2010. An Empirical Analysis of Exploitation Attempts based on Vulnerabilities in Open Source Software. Proceedings of the 9th Workshop on the Economics of Information Security (WEIS). Available at http://weis2010.econinfosec.org/papers/session6/weis2010_ransbotham.pdf

Roberds, W., Schreft, S., 2009. Data breaches and Identity theft. Journal of Monetary Economics 56(7), pp. 918—929.

Rochet, J., Tirole, J. 2006a. Externalities and regulation in card payment systems. Review of Network Economics 5(1), pp. 1-14.

Romanosky, S., Acquisti, A., 2009. Privacy Costs and Personal Data Protection: Economic and Legal Perspectives of Ex Ante Regulation, Ex Post Liability and Information Disclosure. Berkeley Technology Law Journal 24(3).

Romanosky, S., Sharp, R., Acquisti, A. 2010. Data Breaches and Identity Theft: When is Mandatory Disclosure Optimal? Proceedings of the 9th Workshop on the Economics of Information Security (WEIS). Available at http://weis2010.econinfosec.org/papers/session1/weis2010_romanosky.pdf

Romanosky, S., Telang, R., Acquisti, A., 2008. Do Data Breach Disclosure Laws Reduce Identity Theft? Proceedings of the 7th Workshop on the Economics of Information Security (WEIS). Available at SSRN: http://ssrn.com/paper=1268926

Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L. F., Downs, J., 2010. Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions. Proceedings of the 28th International Conference on Human Factors in Computing Systems, pp. 373-382.

Stern, H., 2009. The Rise and Fall of Reactor Mailer. Proceedings of the MIT Spam Conference. Available at http://projects.csail.mit.edu/spamconf/SC2009/Henry_Stern/

Sullivan, R., 2009. The Benefits of Collecting and Reporting Payment Fraud Statistics in the United States. Payment Systems Briefing, Federal Reserve Bank of Kansas City.

Sullivan, R., 2010. The Changing Nature of U.S. Card Payment Fraud: Industry and Public Policy Options, Economic Review, Federal Reserve Bank of Kansas City, issue Q II, pp. 101-133.

van Eeten, M., and Bauer, J. M., 2008. The Economics of Malware: Security Decisions, Incentives and Externalities. OECD Science, Technology and Industry Working Paper No. 2008/1

van Eeten, M., Bauer, J. M., Asghari, H., Tabatabaie, S., Rand, D., 2010. The Role of Internet Service Providers in Botnet Mitigation: An Empirical Analysis Based on Spam Data. Proceedings of the 9th Workshop on the Economics of Information Security (WEIS). Available at http://ideas.repec.org/p/oec/stiaaa/2010-5-en.html

van Eeten, M., Asghari, H., Bauer, J. M., Tabatabaie, S., 2010. Internet Service Providers and Botnet Mitigation: A Fact-Finding Study on the Dutch Market. Technical report, Delft University of Technology. Available at http://www.rijksoverheid.nl/bestanden/documenten-en-publicaties/rapporten/2011/01/13/internet-service-providers-and-botnet-mitigation/tud-isps-and-botnet-mitigation-in-nl-final-public-version-07jan2011.pdf

Varian, H., 2004. System Reliability and Free-Riding. In: Camp, L. J., Lewis, S. (Eds.) Economics of Information Security, vol. 12 of Advances in Information Security, pp. 1–15. Kluwer Academic Publishers.

Wash, R., 2010. Folk models of home computer security. Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS), pp. 1-16.