The Economics of Hacking

Hui, Kai-Lung and Zhou, Jiali

Hong Kong University of Science and Technology

19 September 2020

The Economics of Hacking

Kai-Lung Hui and Jiali Zhou

klhui@ust.hk

jiali.zhou@connect.ust.hk

Department of Information Systems, Business Statistics and Operations Management, School of Business and Management

Hong Kong University of Science and Technology

Acknowledgements: This is a draft of an article that has been accepted for publication by Oxford University Press in the forthcoming Oxford Research Encyclopedia of Business and Management (https://oxfordre.com/business) by Kai-Lung Hui and Jiali Zhou due for publication in 2020.

The Economics of Hacking

Abstract

Hacking is becoming more common and dangerous. The challenge of dealing with hacking often comes from the fact that much of our wisdom about conventional crime cannot be directly applied to understand hacking behavior.

Against this backdrop, this essay reviews hacking studies, with a focus on discussing the new features of cybercrime and how they affect the application of classical economic theory of crime in the cyberspace. Most findings of hacking studies can be interpreted with a parsimonious demand and supply framework. Hackers decide whether and how much to "supply" hacking by calculating the return on hacking over other opportunities. Defenders optimally tolerate some level of hacking risks because defense is costly. This tolerance can be interpreted as an indirect "demand" for hacking. Variations in law enforcement, hacking benefits, hacking costs, legal alternatives, private defense, and the dual use problem can variously affect the supply or demand for hacking, and in turn the equilibrium observation of hacking in the market. Overall, this essay suggests that the classical economic theory of crime remains a powerful framework to explain hacking behaviors. However, the application of this theory calls for considerations of different assumptions and driving forces, such as psychological motives and economies of scale in offenses, that are often less prevalent in conventional (offline) criminal behaviors, but that tend to underscore hacking in the cyberspace.

Keywords

hacker, hacking, cybercrime, supply, demand, law enforcement

As businesses and consumers increasingly digitize their transactions and activities, hacking of computing and data resources has become common and dangerous, creating even larger economic impacts than many conventional crimes (Kshetri 2006). Hacking encompasses many unique characteristics, making it more challenging to tackle than many conventional crimes in the physical world.

This essay advances and analyzes several key features, including law enforcement, hacking benefits and costs, legal alternatives, private defense, and the dual use nature, that set hacking apart from other conventional crimes, and that call for more focused research and development. The analysis extends classical economic theories of crime, viz., the market model of crime (Ehrlich 1996; Freeman 1999), to the cyberspace. It illustrates how such market regulation of hacking provides a powerful framework to understand and predict hacking prominence in view of the unique assumptions and driving forces, such as psychological motives and scale economies in offenses, that are often less prevalent in conventional (offline) criminal behaviors, but that tend to underscore hacking in the cyberspace.

Because this essay focuses exclusively on economic factors affecting hacking choices, it does not review studies that: (1) Do not involve hackers. For example, research about user misbehavior in organizations or security vulnerabilities due to under-investment of protection is not included in this review. (2) Do not focus on hacking behavior. For example, research about exchanges in illicit markets or underground communities, or that assumes exogenous hacking threats, is not included. (3) Do not focus on economic factors and analysis. For example, research emphasizing sociological or psychological theories, or hacking technology, is not included.

Furthermore, this essay covers general hacking activities but not selected hacking types or techniques. Readers interested in particular hacking types can refer to, for example, Cartwright, Castro, and Cartwright (2019) for ransomware, Rao and Reiley (2012) for spam, Ramzan (2010) for phishing, and Leontiadis, Moore, and Christin (2014) for search-engine poisoning.

The Market Model of Hacking

Consider a hacking "market" with three types of actors: Hackers, defenders, and the government. Hackers obtain benefits by attacking some digital assets, for example, by breaching a bank's database and using the stolen credit card data for illicit purchases. Defenders and the government want to protect the digital assets from hackers' attacks. The government imposes ex post punishment against hacking. Defenders take ex ante precautions (Katyal 2001). The term "cybercrime" is used to refer to malicious hacking throughout this essay.

A convenient starting point to analyze the interaction between hackers, defenders, and the government is the so-called "market model of crime" (Ehrlich 1996; Freeman 1999), which interprets criminal activities as a "product". The familiar supply and demand

framework in economics can then be used to study the equilibrium levels of cybercrimes in the market.

Cybercrime Supply

The classic model of criminal offense (Becker 1968) offers a practical way to model the supply of cybercrime. The model considers individual's choice between criminal activities and legal activities based on expected utility.

Specifically, let U_{c1} denote the hacker's utility of committing the cybercrime, which captures the expected illegitimate benefit from the offense net of any costs incurred in acquiring such benefit. Let U_{c2} denote the hacker's utility of committing cybercrime but getting apprehended and punished. In addition to the benefit and direct costs of committing the offense, U_{c2} is related to the severity of criminal sanction upon apprehension. In general, we expect $U_{c1} > U_{c2}$ because criminal sanction should decrease hackers' utility. Let U_{nc} denote the hacker's utility of abstaining from crime, which can be interpreted as the opportunity cost of cybercrimes.

Let p be the probability of apprehension and conviction, which is related to the strength of law enforcement. As a rational decision maker, the hacker will commit a cybercrime if and only if its benefit from committing the offense exceeds its cost. Mathematically, this condition can be expressed as

$$(1-p)U_{c1} + pU_{c2} > U_{nc}. \qquad (1)$$

The left-hand side of Equation (1) is the hacker's expected utility from committing the cybercrime. This expected utility is increasing in U_{c1} and U_{c2} and decreasing in p when $U_{c1} > U_{c2}$, which is generally the case. Equation (1) is more likely to be satisfied when the right-hand side, U_{nc}, is small. Realistically, hackers are more likely to commit an offense if the opportunity cost of cybercrime is low (for example, when they are jobless and so have more time to spend on hacking).

By Equation (1), a hacker is more likely to commit cybercrime when the return on cybercrime, $(1-p)U_{c1} + pU_{c2}$, increases. If we plot the return on cybercrime against the amount of cybercrime that hackers are willing to commit, we obtain an upward-slopping "supply" curve. Figure 1 shows an illustrative supply curve, SS. In Figure 1, the horizontal axis, Q, represents the amount ("quantity" in a conventional supply curve) of cybercrime. The vertical axis, P, represents utility gained ("price" in a conventional supply curve) by the hacker from committing the offense.

Cybercrime Demand

It is less straightforward to define the "demand" for cybercrime. In principle, no people or organization would desire crime, but everyone has a certain degree of tolerance for crimes because, realistically, defending against crimes is costly (Mukhopadhyay et al. 2013). This is especially the case for cybercrimes, the defense of which requires

substantial domain knowledge and effort. For example, it is impractical for a company to constantly monitor how its employees handle phishing emails. Most companies would conduct some training sessions and then let employees tackle phishing emails on their own, leaving a possibility ("tolerating") that some employees may inadvertently respond to and get cheated by a phishing email.

Similar to conventional crime economics theories (e.g., Ehrlich 1996; Freeman 1999), the tolerance of cybercrime is interpreted as cybercrime "demand". The term "demand" does not imply defenders desire cybercrimes; instead, it merely reflects the situation where defenders have to tolerate some cybercrimes because it is practically infeasible, or too costly, for them to eliminate all cybercrimes.

In conventional markets, price serves as a wealth transfer instrument that regulates the exchange of goods and services. In the case of cybercrimes, wealth is moved between the defender and the hacker. For example, when a hacker breaches into a company's server to steal customer information, the company suffers a loss in terms of reputation damage, customer redress and compensation, or fines payable to regulatory agencies. The hacker can benefit from fame, selling valuable customer information such as credit card data, or directly using the data for fraudulent behaviors or exploitations, such as email scams or business process compromises.

If hacking increases the defender's loss of wealth, the defenders' tolerance ("demand") for cybercrime decreases and they will have a greater incentive to reduce the loss by spending more on protection, such as deploying more resources to detect and prevent intrusion, conducting security audits to reduce system bugs, or cybersecurity awareness training. Therefore, if we plot the loss of wealth due to cybercrime against the amount of cybercrime that the defender is willing to tolerate, we obtain a downward-slopping "demand" curve. Figure 1 shows an illustrative demand curve, DD.

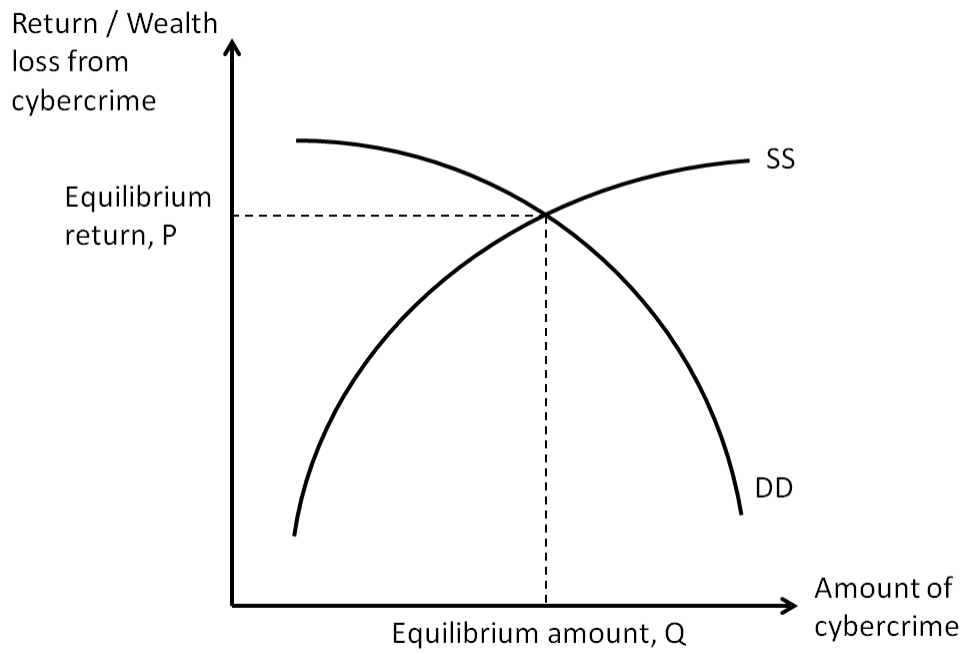

Figure 1. The market model of hacking

Empirically, the demand curve can be approximated by surveying companies about their willingness to spend resources in protecting themselves at different hypothetical loss levels due to hacking. Estimating the supply curve is more challenging as it is often difficult to identify hackers. Nevertheless, it can be approximated by polling the hacker community (e.g., active participants in online hacker forums) about their inclination to commit cybercrimes against different levels of returns. Statistical techniques such as randomized survey responses (e.g., Warner 1965; Kwan, So and Tam 2010) could be used to elicit truthful answers on sensitive topics such as hacking.

In any case, the market model of hacking provides a convenient and intuitive theoretical tool to analyze potential factors relevant to hacking. More importantly, it helps assess how hacking and cybercrime could be influenced in a predictable direction.

Market Equilibrium

The demand curve and supply curve in Figure 1 together determine the equilibrium amount of cybercrime and return on/wealth loss due to cybercrime (Q and P in Figure 1). Consider, for example, when the expected wealth gain from a cybercrime exceeds the equilibrium level. Hackers would supply more cybercrimes than what defenders would tolerate. Defenders would then step up the protection to reduce their wealth loss. This would continue until the wealth lossin cybercrime and level of cybercrime return to the equilibrium level. Similarly, if the wealth loss is lower than the equilibrium level, then defenders would tolerate more cybercrimes than what hackers would supply. Defenders would spend less in protection to save costs. This would make the system

less secure, which raises the potential return on cybercrime. The supply and demand of cybercrime would again return to the equilibrium level because of the "regulation" of the return/wealth loss arising from the cybercrime.

Note that this stylistic equilibrium analysis categorically groups hackers into one single entity and defenders into another single entity. Realistically, hackers vary in terms of demographics and hacking interests. Defenders may comprise firms, customers, and security service providers. The "hackers" and "defenders" above need not refer to a single actor; instead, they are abstract entities that facilitate systematic analysis of how the choices of each of these hackers and defenders affect the equilibrium observation of cybercrimes in the market.

Separating the consideration of cybercrime demand and supply allows us to illustrate how changes in external factors affect the equilibrium cybercrime level. For example, consider a costless cyber insurance that covers part of the loss in a data breach (e.g. Böhme and Schwartz 2010; Bandyopadhyay and Mookerjee 2019). Because defenders' loss from cybercrime is reduced due to the cyber insurance's compensation, defenders will tolerate more cybercrimes at any levels of wealth losses. This implies the cyber insurance will shift the demand curve to the right (i.e., it makes defenders more tolerating of hacking because the hacking impact is smaller). The new equilibrium level of cybercrime will become Q' in Figure 2. Importantly, this "demand shift" will also increase the return on cybercrime from P to P' in the equilibrium.

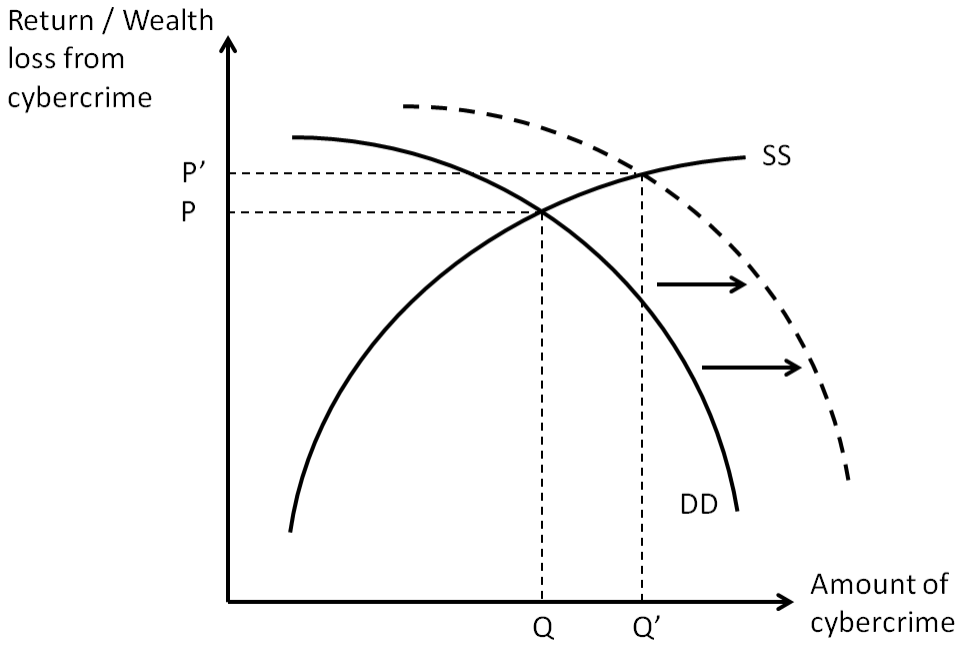

Figure 2. Demand curve shift due to a costless cyber insurance

Similarly, consider what happens when the returns from other (non-hacking) activities

become more attractive. Less cybercrimes will be provided at any return levels because some "would-be" hackers will now change to perform these non-hacking activities instead of hacking. The supply curve will be shifted to the left. The new equilibrium crime level will become Q'' as shown in Figure 3.

Interestingly, the equilibrium return on cybercrime will also increase from P to P''. This is because, in the presence of more attractive substitute activities, only those hackers who expect higher returns from cybercrimes will continue to hack. Hackers who expect lower returns from cybercrimes will quit and spend their time on substitute activities that give them higher returns than hacking. As a result, cybercrimes with low expected returns will not be performed, leading to an increase in the observed cybercrime return level in the equilibrium.

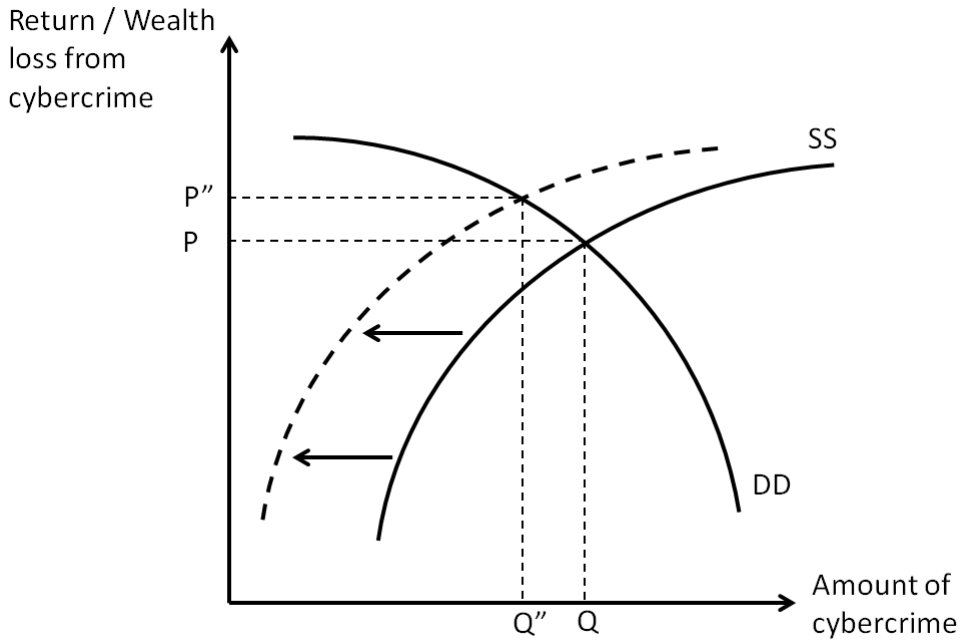

Figure 3. Supply curve shift due to more attractive legal jobs

Factors Shaping Cybercrime Provision

The supply and demand framework provides a convenient tool to analyze the factors shaping cybercrimes. By variously influencing supply and demand, these factors affect the equilibrium cybercrime levels through regulating hacking returns. Note that most of the factors affecting cybercrimes are contained in classical crime theories too (e.g., Becker 1968; Stigler 1970). However, different emphases and analysis are needed to articulate their significance in the cybercrime setting. Hence, the remainder of this essay focuses on the unique characteristics and influences of these factors.

Law Enforcement: Deterrence's Difficulty

We start from the celebrated economics of crime literature, which analyzes how law enforcement affects crime choices. There are two broad strands in this literature. The first focuses on normative analyses of the optimal enforcement of law, in particular, four policy choices – sanctioning rule, form of sanction, magnitude of sanction, and probability of detection (Polinsky and Shavell 2007). The second strand, the positivist analyses, focuses on assessing the responsiveness of crimes to sentencing (Chalfin and McCrary 2017; Freeman 1999). In general, the lesson from this strand of research is that increasing either the severity or certainty of punishment deters criminals.

The cyberspace is a novel regime that challenges the traditional justice system in terms of identification, investigation, and prosecution. One major challenge is that many cybercrime victims do not report their security breaches or incidents to legal authorities (Kshetri 2006; Police Executive Research Forum 2018). One reason for such under-reporting is that they could be unaware of the victimization because modern computer viruses, trojans, or malicious spyware are designed to secretly take control of the victims' systems for other mischievous actions, such as to form part of a botnet or, lately, to mine Bitcoin. The victims often do not have enough expertise to detect security compromises. Organizational victims, on the other hand, often intentionally under-report because they fear the victimization may affect customers' and stakeholders' impression of their ability and, in turn, their reputation. The head of the U.S. Federal Bureau of Investigation's Cyber Division estimated that only about 10 percent of all security incidents were reported (Police Executive Research Forum 2014).

Even when legal authorities receive security incident reports, formidable challenges exist in collecting the evidence for criminal proceedings. Obtaining the electronic trails of a cybercrime often requires adequately-trained and equipped investigators, who are not readily available in the legal system. In addition, to enhance the quality of evidence, investigators must trace the original sources of communications to connect the hackers and cybercriminals with real persons in the physical world. This is challenging as the anonymity and complexity of Internet communications allow cybercriminals to easily disguise and hide themselves from law enforcement agencies.

Challenges also exist in prosecution and criminal sanctions. Law enforcement tends to prioritize crimes according to offense severity. In most cases, the loss needs to exceed certain thresholds or be significant enough before investigation would start (e.g., Police Executive Research Forum 2014; Cárdenas et al. 2009). But cybercrimes often encompass multiple small or low-impact victims (Moore, Clayton, and Anderson 2009). Furthermore, the global nature of cybercrimes means a sophisticated offender can operate from locations outside the jurisdictions proscribing her/his activities. Territorial boundaries and differences in legal systems bring severe operational challenges in apprehending and prosecuting international cybercriminals.

These challenges limit the deterrence effect of law enforcement against cybercrimes, which is consistent with prior empirical studies. For example, the U.S. congress passed

the *Controlling the Assault of non-Solicited Pornography and Marketing Act of 2003* (CAN-SPAM Act) in 2003. This law criminalizes spamming and imposes penalties on violations. Despite the claim made by the U.S. FTC that the volume of spam had decreased and begun to level off (Majoras et al. 2005), research on the effectiveness of the CAN-SPAM Act is at most inconclusive (e.g., Kigerl 2016; Lee 2005; Grimes 2007). Similarly, Png, Wang, and Wang (2008) find limited evidence to support domestic law enforcement can deter attacks within the country. Hackers simply displace the crimes to other locations. In particular, the U.S. cybercrime law may have caused hackers to initiate attacks from other countries instead of the U.S.

Referring to the supply and demand framework, effective law enforcement increases hackers' risks of committing cybercrimes. By equation (1), it would decrease U_{c2}, making the crime less attractive. *Ceteris paribus*, it should shift the supply curve leftward, causing the equilibrium crime level to decrease. If, however, the law cannot be effectively enforced due to challenges in identification, investigation, and prosecution, then U_{c2} would increase, which would shift the supply curve to the right. Therefore, the effectiveness of deterrence through law enforcement would shape the equilibrium level of cybercrime and the return on cybercrime by distorting the supply, viz. hackers' incentives to hack.

As one example, in the context of Convention on Cybercrime, Hui, Kim, and Wang (2017) find that international cooperation reduces distributed denial of service (DDoS) attacks, a popular cyber-attack in which hackers try to overload systems by flooding the systems with superfluous requests, by at least 11.8 percent. A similar effect does not exist if the participating countries make a reservation on international cooperation, meaning when they are less committed in assisting each other in global enforcement, which would inevitably make cybercrimes more attractive to hackers.

Note that law enforcement can also affect defenders' behaviors, leading to an impact on equilibrium crime level via demand-side considerations. For example, Romanosky, Telang, and Acquisti (2011) find that data breach disclosure law helps reduce identity thefts by an average of 6.1 percent. Referring to Figure 1, mandating disclosure of security breaches would shift the demand curve leftward because defenders will tolerate less cybercrimes in view of the potential reputation damage due to security breach disclosure. The shift of the demand curve leftward would lead to both lower cybercrime level and a smaller return on cybercrime.

Benefits of Hacking: Financial vs. Non-Financial gains

Uncovering hackers' motivation is instrumental to understanding their behaviors. In particular, why would they choose to hack in the first place? The classical theory of crime assumes we can summarize the benefits or losses associated with criminal activities by some monetary terms (see, e.g., Becker 1968). It can be used to effectively model criminal behaviors when the criminals are motivated financially or by other

material benefits, as in the case of most conventional crimes such as burglary or robbery.

However, ample self-reported evidence suggests that hackers can be psychologically motivated. For example, Schell and Dodge (2002) surveyed more than 200 hackers, among whom 36 percent said they hack to "advance network, software, and computer capabilities," 34 percent claimed "to solve puzzles or challenges," and 8 percent said to "expose weakness in organizations or in their products." By contrast, only 4 percent of the surveyed hackers indicated their primary motivation to be "make money." Other studies suggest hackers may hack for curiosity, fun (e.g., Turgeman-Goldschmidt 2005; Xu, Hu, and Zhang 2013), retaliation, espionage, learning (e.g., Barber 2001), ideology (e.g., Coleman 2012; Jordan and Taylor 2004), demonstrating skills (e.g., Coleman and Golub 2008), and fame (e.g., Thomas 2002; Taylor 2012).

The prior literature does not provide systematic guidance on how to analyze or predict criminal behaviors when criminals seek psychological benefits, which can be the case for hackers. For example, because of the economies of learning and specialization, hackers should focus on selected victims in every crime to leverage their experiences and expertise, just as a mugger familiar with robbing banks is more likely to rob a bank in his next crime. Empirical evidence has shown that hackers tend to seek variety, i.e., they prefer to attack victims in different regions using different hacking methods across attacks (Ooi et al. 2012). Such behavior is difficult to rationalize if hackers are purely profit-driven, but can be better explained if they are assumed to have a psychological desire or preference for variety seeking.

Among the psychological motivations, fame and peer recognition are often among the most prominent factors (e.g., Thomas 2002; Taylor 2012). Leeson and Coyne (2005) provide an example to formalize the analysis of fame-driven hacker behaviors. They suggest that the attacks generated by fame-driven hackers depend on the interactions between a supply force and a demand force. The supply of attacks from fame-driven hackers is determined by the cost of hacking and the number of hackers who desire fame in the underground hacker community. The demand for attacks is determined by the hacking community's reaction, which specifies how the community responds to different hacking quantities with fame. For example, one may expect more hacking is recognized with more notoriety and less hacking with less notoriety.

Notwithstanding the existence of non-financial hacking incentives, the market model of hacking presented in Figure 1 can help analyze the impact of non-financial incentives. Such non-financial incentives should shift the supply curve to the right because, being motivated non-financially, hackers would commit more cybercrimes at every level of return. Due to the right shift of the supply curve, the equilibrium level of expected return/wealth loss would decrease whereas the equilibrium level of cybercrime would increase. This seems consistent with the anecdotal observation that many cyber-attacks feature a high volume of attacks, but the nature of the offense is often minor, not causing overly-significant disruption or losses to the defenders.

Costs of Hacking: Economies vs. Diseconomies of Scale

Cybercrime is highly scalable. A physical bank robber may be able to hit one or two banks in a week. Cyber-attacks, such as computer viruses, worms, phishing, or ransomware, can target 1,000 if not 10,000 or even more devices at once.

Png and Wang (2009) use two analytical models to formalize the differences between one-to-many attacks, known as mass attacks or opportunistic attacks, and other (more conventional) one-to-one attacks, known as targeted attacks. In mass attacks, such as phishing scam or ransomware attacks, the effort exerted by a hacker applies to a big group of users. Attacking more targets does not add much cost to the hacker due to economies of scale, as such mass attacks usually exhibit low or even negligible marginal costs. By contrast, in targeted attacks, such as DDoS attacks or system intrusions, the effort exerted by a hacker applies to a particular victim. The marginal cost of attack against each additional victim would increase because the attack is more customized, meaning the total cost of attack increases with the number of targets. Attacking more targeted victims also increases the probability of apprehension (Fultz and Grossklags 2009) because each of these targeted, customized attacks may leave footprints for law enforcement agencies to track the hackers. Therefore, targeteds attacks often exhibit diseconomies of scale.

The difference in cost structure may affect how hackers choose targets. Ransbotham and Mitra (2009) review the postings from Usenet groups and find that three broad dimensions of target attractiveness, viz. being tangible and iconic and having reprisal value, are key hacker considerations in targeted attacks. By contrast, in mass attacks, hackers do not need to pre-select victims but are motivated to find a large number of targets. Hence, the likelihood of a particular firm becoming a target increases with the so-called "passive Internet presence," i.e., the number and functionality of Internet connections, and the volume and richness of Internet activities.

In general, the economies of scale associated with low or negligible marginal costs tend to shift the supply curve rightward. This should be the case for cyber-attacks, especially mass attacks, as well. Referring to equation (1), the low marginal cost of attacks tends to increase U_{c1} and U_{c2}, which makes attacking more preferred. Given the expected gain from each victim, a hacker committing a mass attack would obtain a higher total net benefit. Therefore, s/he will be motivated to commit more cybercrimes. Referring to Figure 1, the right-shift of the supply curve implies the mass attack market will have more cybercrimes with a lower equilibrium wealth loss for each attack. This does seem to be the case for mass attacks such as phishing scams and ransomware attacks.

Another way to model hackers' costs is to allow for a fixed cost of exploiting a new vulnerability (e.g., cognition and opportunity costs). In this setting, Allodi, Massacci, and Williams (2017) study how hackers make temporal choices of vulnerabilities. Mass attackers may exploit only one vulnerability per software version and be slow in

introducing new vulnerabilities into their arsenals. Empirically, this theory can explain the heavy-tail distributions of some exploited vulnerabilities (Mukhopadhyay et al. 2019). In particular, 5% of exploited vulnerabilities is responsible for about 95% of the attacks in some software (Allodi 2015).

Finally, some studies have evaluated hackers' responses to interventions that raise the cost of hacking. Theoretically, a higher cost of hacking shifts the supply curve leftward, and so it should reduce cybercrimes. Empirical findings support this prediction. For example, Brunt, Pandey, and McCoy (2017) find attacks launched through DDoS-for-hire services reduced after a payment intervention requiring buyers to use Bitcoin for the transactions. The cost of attacks could have risen due to the difficulty in purchasing Bitcoins and the threat associated with Bitcoin storage and use. In the illegal market, McCoy et al. (2012) report how interrupting sellers' payment networks, which raises sellers' monetization costs, can be effective in suppressing such illegal businesses by causing, for example, reduction of illegal products sales, pursuit of riskier payment mechanisms, and drops in consumer conversion (in terms of people initiating but not eventually completing a transaction).

Legal Alternatives: White Collar vs. Marginal Jobs

The classical theory as outlined in the market model of hacking assumes offenders choose whether to commit a criminal act by weighing its benefits against the benefits from alternative activities. An increase in the pay from alternative activities increases the offenders' opportunity cost, and so should reduce the probability of them to commit the criminal act. A large empirical literature has examined how unemployment, wage, job training, and other career developments affect crimes in conventional markets (see Chalfin and McCrary 2017 for a detailed review).

According to this theory, how much potential offenders respond to the attractiveness of alternative opportunities depends on their earnings potential in the legal markets (Raphael and Winter-Ebmer 2001). Conventional crimes and cybercrimes are committed by distinct social groups whose earnings potential in the legal markets are likely to be different. Most conventional crimes are committed by marginal members of the society, who are often disproportionately drawn from groups with low incomes and poor employment prospects. By contrast, people with comparative advantages at cybercrimes tend to be educated and capable, but they may live in societies with poor job prospects or ineffective law enforcement. The literature refers to crimes committed by a person of respect and high social status as "white collar crimes" (Posner 1979).

Because of the demographic differences between conventional criminals and hackers, marginal jobs that constitute attractive opportunities for conventional criminals may not attract hackers. Referring to equation (1), increasing the attractiveness of legal alternatives would increase U_{nc}. This should shift the supply curve of cybercrime leftward. At each level of the return on cybercrime, fewer hackers would be willing to

commit cybercrimes. But the extent of the shift should depend on the type of legal alternatives. For example, white collar jobs might have a greater effect on this shift than marginal jobs because they fit hackers' demographic profiles better. Similarly, jobs opportunities in information technology-related sectors might have a greater effect on this shift than jobs opportunities in other ("blue collar") sectors.

Few studies have examined the connection between legal labor market conditions and cybercrimes. Empirical works such as Png, Wang, and Wang (2008) and Garg, Camp, and Kanich (2013) document a significant impact of labor market conditions on cybercrimes. However, these studies suffer two limitations. First, their empirical data allow them to identify only the origins of attacks but not the hackers. It is well known that cyber-attacks often come from computers under control by hackers from other geographical locations. Second, most of these research findings point to correlations, not causal relationships.

Private Defense: Positive vs. Negative Spillover

So far the discussion has been focusing on supply-side factors that influence hacking. This section turns to the role of private defense, which is related to defenders' tolerance, or "demand", for cybercrimes. The term "private defense" refers to defenders' actions to protect themselves and avoid crime victimization (cf. government or law enforcement actions to protect the general public).

Research on private defense in conventional crime settings is scant. The available studies are primarily interested in two questions. One is whether private protection deters criminals (Ayres and Levitt 1998; Cook and MacDonald 2011). The other is how private defense of one party affects the security of other parties, a phenomenon commonly called "spillover effect" in economics (Kunreuther and Heal 2003). The evidence suggests certain private defense measures, such as installing stolen vehicle tracking systems (Ayres and Levitt 1998) or establishing business improvement districts that strengthens local safeguards (Cook and MacDonald 2011), can help deter crimes. Also, there exists positive spillover effects in that the use of these defense measures benefit all defenders, including those not using them, because they increase the uncertainty of committing the crimes. However, we do not know if these findings would hold for other private defense measures (Chalfin and McCrary 2017).

The spillover effect also exists in the cyberspace. In particular, cybercrime studies have argued, theoretically, that the direction of spillover critically depends on the technology that relates individual efforts to the outcomes (Varian 2004). The direction of spillover will in turn affect hackers' strategies (Fultz and Grossklags 2009). In a "sum of effort" scenario, such as an attack that tries to slow down the distributed transfer of a file on a peer-to-peer network, there is a positive spillover effect. Adding more users will make the system more reliable and safer. Florêncio and Herley (2013) argue this is the key reason why many individuals in a large network are not attacked even when they have

poor security. By contrast, in a "weakest link" scenario, such as a hacker trying to breach a closed network by locating a vulnerable computer, there is a negative spillover. Adding more users will make the system less secure from attacks.

In general, the socially-optimal level of private defense should be set at a level where the marginal benefit of protection equals the marginal cost of protection. The spillover effect, however, is often ignored by individuals. Hence, an individual's equilibrium choice of private defense may not align with the true cost or benefit of that private defense for the society. In other words, the omission of spillover effects may lead to either too much or too little defense, implying market inefficiency.

Although spillover effects could lead to too much (when more users bring more benefits, as in the case of the "sum of effort" configuration) or too little (when more users bring negative spillover, as in the case of the "weakest link" configuration) protection, two other types of market failure unambiguously lead to too little protection. One is misaligned incentives. For example, companies usually do not have incentives to protect customer data in full as they do not bear the full costs of a data breach (e.g., Lee et al. 2013; Hui et al. 2019). Another market failure arises from information asymmetry, also known as the "lemon problem" (Akerlof 1970). For instance, in the software market, customers have no reason to pay for more secure software as most of them cannot evaluate the software's security. Hence, software vendors do not have incentives to invest in more security. These two types of market failures predict insufficient protection, which provides greater incentives for hackers to commit cybercrimes. For a thorough discussion of these market failures, see Anderson and Moore (2006).

Empirically, the first-order question is whether private defense deters cybercrimes. Nagle, Ransbotham, and Westerman (2017) analyze the data from 480 Fortune 500 enterprises and find that the number of open ports in a firm is associated with higher incidents of botnet activities, potential exploitations, and unsolicited communications, providing direct evidence that private defense (exhibited by the number of open ports in an enterprise) does affect the risks of cyber-attacks.

However, care should be taken in concluding that private defense always leads to less cybercrimes. As a counter example, Miller and Tucker (2011) find that using encryption to protect customer data increases (instead of decreases) the risk of data breach in the healthcare sector. This surprising result is driven by an increase in employee dishonesty- and carelessness-driven security breaches. The encryption seems to be a substitute of the employees' self-precaution. Similarly, in evaluating the security effects of meaningful-use attestation, a U.S. certification mechanism that fosters patient data protection, Kwon and Johnson (2018) find that although attested hospitals observe fewer data breaches in the short term, the number of data breaches does not decrease in the long-term. They suggest this attestation only helps hospitals establish policies and procedures to combat known attacks, but may impede quick responses and exploratory activities that address evolving security threats to the extent that it can be bad for the

hospitals' security. They also find this substitution effect exists between regulatory pressure and proactive security investments. Overall, these findings suggest one needs to consider the interactions between different defense measures before concluding whether certain private defense will lead to less cybercrimes (Hui et al. 2012).

Other empirical works have compared the deterrence effects of different security practices. Ransbotham (2010) compares the security of open source software and closed source software. He finds open source software vulnerabilities are at greater risk of exploitation, diffuse more rapidly, and attract greater volumes of exploitation attempts. Vasek, Wadleigh, and Moore (2015) find that using content management systems (CMS) with higher market shares increases the risk of web-server compromise, probably because such CMS become a more attractive target for miscreants. Surprisingly, they find that this conclusion also applies to different versions of the same software – popular software versions tend to be targeted more often than less popular versions even though the popular versions are newer and up-to-date. These findings offer practical insights on how defenders can make better choices among different defense measures.

The Dual Use Problem

Dual use refers to the setting where an activity has both positive and negative uses (Katyal 2001). Forbidding the activity itself would forfeit its good purpose. A well-known example of dual use technology in the conventional crime literature is the availability of guns (for a detailed review, see Lott 2013). On the positive side, allowing citizens to carry guns should have a deterrence effect that raises the expected risks and costs of criminal activities. On the negative side, legalizing guns would increase the likelihood that any particular dispute turns into a gun battle.

Because dual use activities affect both the supply and demand of crimes and could generate positive and negative consequences, their net impact is largely an empirical question. Many policy choices that intervene in the creation, use, storage, access, and communication and dissemination of cybercrime-related information face this "dual use" problem.

For example, one well-known controversial topic is vulnerability disclosure. On the one hand, disclosing vulnerability enables users to take precautions that can prevent or reduce security risks. On the other hand, hackers may make use of such information to attack users before users install any protection measures. Overall, Arora, Nandkumar, and Telang (2006) find that hackers exploit patched vulnerabilities (disclosure and patched) more than secret vulnerabilities (without disclosure). They most often exploit published vulnerabilities without patches.

This finding suggests hackers get information from vulnerability disclosure and patch releases, and so vulnerability information needs to be disseminated carefully. To give an advantage to security professionals, two other vulnerability disclosure mechanisms have been proposed. One is the so-called "limited disclosure". In this mechanism, some

vulnerability infomediaries, such as the Computer Emergency Response Team (CERT), iDefense, and Tipping Point, control the timing of vulnerability disclosure. Typically, these infomediaries will accept vulnerability reports from vulnerability identifiers (could be ordinary software users or security researchers). However, after receiving vulnerability reports, they immediately inform only the affected software vendors and wait for appropriate safeguards to be put in place before making the vulnerability information public. Mitra and Ransbotham (2015) find that, compared with limited disclosure, full disclosure (making the vulnerability public immediately) expedites the onset of attacks against a vulnerability, increases the penetration of attacks among target systems, increases the risk of first attacks, and shifts the volume of attacks corresponding to a vulnerability to an earlier stage in its life cycle. Therefore, limited disclosure provides practical advantages to benign security professionals.

The other mechanism is to create a "vulnerability market", which provides security professionals with monetary rewards for reporting vulnerabilities. The vulnerability market provides an advantage to the "protection side" by increasing the incentives for benign security professionals to report vulnerabilities. Compared with nonmarket disclosure, Ransbotham, Mitra, and Ramsey (2012) demonstrate that market-based disclosure provides security advantages, for it restricts the diffusion of vulnerability exploitation, reduces the risk of exploitation, and decreases the volume of exploitation attempts.

Another example of the dual use problem is the dissemination of hacking knowledge. On the one hand, discussion about hacking techniques exposes more people to hacking and helps like-minded hackers collaborate on cyber-attacks, which may promote aggression. On the other hand, such discussion can help develop and spread protection knowledge, which may help curb cyber-attacks. Using the data from hackforums.net, one of the largest hacking technology forums, Yue, Wang, and Hui (2019) find that a one percent increase in DDoS attack discussion decreases DDoS attacks by 0.032% to 0.122%. This finding suggests that, surprisingly, allowing hacking discussion actually helps reduce cybercrimes.

The last dual use example is the hacker group. Though most existing studies regard hackers as offensive and defenders as defensive, realistically hackers' knowledge and techniques about attacks can be used to protect the systems (Moore and Anderson 2012). At the national security level, Moore, Friedman, and Procaccia (2010) find that when two parties can either stockpile vulnerabilities for offensive advantages or share vulnerabilities to secure the systems, the offensive behavior is likely to emerge endogenously, even though defensive behavior is more preferred.

At the micro level, dual use hackers may open new possibilities for reducing cybercrimes when hackers are appropriately incentivized. One noteworthy practice that leverages dual use hackers is the bug bounty program, which has been pursued by many large organizations, including Google, Facebook, Tencent, and even the US Department

of Defense. In bug bounty programs, hackers will be given monetary reward, namely "bug bounty", for reporting "bugs" in organizations' systems. Zhou and Hui (2019) study the impact of bug bounty programs. Their analysis shows that using bug bounty programs to divert hackers from exploitation provides economic (and sometimes security) advantages over using only self-defense and law enforcement. Until now, the dual use feature of hackers is relatively less explored in the traditional crime and even the cybercrime) literature. Future research should explicate its potential and examine its practical and policy implications in tackling hacking and cybercrimes.

Conclusions

This review analyzes hacking behaviors using a conventional demand and supply framework. It synthesizes the relevant cybercrime and hacking studies using six key influencing factors. In reviewing these studies, the key differentiations of cybercrimes from conventional crimes are contrasted along these six factors. Table 1 summarizes the six factors, their key features, how they influence cybercrimes, their impacts on the equilibrium crime level and return, and the examples drawn from the relevant studies.

Understanding hacking behaviors not only increases our knowledge about information security. It also helps us devise effective ways to curb cybercrimes. Existing studies on hacking behaviors are scarce and scattered (Mahmood et al. 2010), yet a message is clear from this essay: the classical economic theory of crime can help analyze many new features of cybercrimes. The contribution of this essay is to systematically investigate these new features and, accordingly, how the classical theory applies. We hope this review provides a useful framework for researchers and practitioners to understand hacking behaviors via an economic perspective.

Table 1. Summary of results

Influencing factor	Influence channel	Key differentiation	Impacts on equilibrium	Example
Law enforcement	Supply curve	**Enforcement difficulty:** the law cannot be enforced effectively in the cyberspace. Hackers can often strategically escape from punishment	Tends to shift the supply curve rightward, leading to higher equilibrium quantity of cybercrimes but lower crime returns.	US CAN-SPAM Act (e.g., Lee 2005, Grimes 2007),
Hacking benefits		**Non-financial incentives:** Many hackers are motivated by non-financial incentives (e.g., they seek psychological or other non-monetary gains)	Tends to shift the supply curve rightward, leading to higher equilibrium quantity of cybercrimes but lower crime returns.	Hackers motivated by variety seeking (Ooi et al. 2002); Fame-driven hackers (Leeson and Coyne 2005)
Hacking costs		**Economies of scale:** mass cyber-attacks exhibit high economies of scale	Economies of scale would shift the supply curve rightward (cf. diseconomies of scale), leading to a higher equilibrium quantity of cybercrimes but lower crime returns.	Economies of scale vs. diseconomies of scale in cyber-attacks (Png and Wang 2009)
Legal alternatives		**White collar crime:** Hackings is often commissioned by educated, "white-collar" workers	Increasing the attractiveness of white collar jobs or IT jobs shifts the supply curve more than marginal jobs or non-IT jobs, which tends to decrease the equilibrium quantity of cybercrimes.	Unemployment and cybercrimes (Png et al. 2008)
Private defense	Demand curve	**Inefficiencies:** Spillover, misaligned incentives, and information asymmetry in private defense	Inefficiencies can shift the demand curve to either the left or right, which could lead to either too much or too little cybercrimes when compared with the social optimum.	Positive spillover vs. negative spillover effects (Varian 2004; Grossklags et al. 2008)
Dual use problem	Both	**Dual use:** policy choice is ambiguous when a technology faces the "dual use" problem	Affects both supply and demand. The net impact on cybercrimes is an empirical question.	Vulnerability disclosure (Arora et al. 2006, Mitra and Ransbotham 2015, Ransbotham et al. 2012); hacking discussion (Yue et al. 2019)

References

Akerlof, G. A. (1970). Quality Uncertainty and the Market Mechanism. The Quarterly Journal of Economics, 84(3), 488-500.

Allodi, L. (2015, March). The heavy tails of vulnerability exploitation. In International Symposium on Engineering Secure Software and Systems (pp. 133-148). Springer, Cham.

Allodi, L., Massacci, F., & Williams, J. M. (2017). The work-averse cyber attacker model: Theory and evidence from two million attack signatures. Available at SSRN 2862299.

Anderson, R., & Moore, T. (2006). The economics of information security. Science, 314(5799), 610-613.

Arora, A., Nandkumar, A., & Telang, R. (2006). Does information security attack frequency increase with vulnerability disclosure? An empirical analysis. Information Systems Frontiers, 8(5), 350-362.

Ayres, I., & Levitt, S. D. (1998). Measuring positive externalities from unobservable victim precaution: an empirical analysis of Lojack. The Quarterly Journal of Economics, 113(1), 43-77.

Bandyopadhyay, T., & Mookerjee, V. (2019). A model to analyze the challenge of using cyber insurance. Information systems frontiers, 1-25.

Barber, R. (2001). Hackers profiled—who are they and what are their motivations?. Computer Fraud & Security, 2001(2), 14-17.

Becker, G. S. (1968). Crime and punishment: An economic approach. In The economic dimensions of crime (pp. 13-68). Palgrave Macmillan, London.

Böhme, R., & Schwartz, G. (2010, June). Modeling Cyber-Insurance: Towards a Unifying Framework. In Workshop on the Economics of Information Security.

Brunt, R., Pandey, P., & McCoy, D. (2017, June). Booted: An analysis of a payment intervention on a ddos-for-hire service. In Workshop on the Economics of Information Security.

Cartwright, E., Castro, J.H., & Cartwright, A. (2019). To pay or not: game theoretic models of ransomware. Journal of Cybersecurity, 5(1), tyz009.

Cárdenas, A., Radosavac, S., Grossklags, J., Chuang, J., & Hoofnagle, C. J. (2009, August). An economic map of cybercrime. TPRC.Chalfin, A., & McCrary, J. (2017). Criminal deterrence: A review of the literature. Journal of Economic Literature, 55(1), 5-48.

Coleman, E. G. (2012). Coding freedom: The ethics and aesthetics of hacking. Princeton

University Press.

Coleman, E. G., & Golub, A. (2008). Hacker practice: Moral genres and the cultural articulation of liberalism. Anthropological Theory, 8(3), 255-277.

Cook, P. J., & MacDonald, J. (2011). Public safety through private action: an economic assessment of BIDS. The Economic Journal, 121(552), 445-462.

Ehrlich, I. (1996). Crime, punishment, and the market for offenses. Journal of Economic Perspectives, 10(1), 43-67.

Florêncio, D., & Herley, C. (2013). Where do all the attacks go?. In Economics of information security and privacy III (pp. 13-33). Springer, New York, NY.

Freeman, R. B. (1999). The economics of crime. Handbook of labor economics, 3, 3529-3571.

Fultz, N., & Grossklags, J. (2009, February). Blue versus red: Towards a model of distributed security attacks. In International Conference on Financial Cryptography and Data Security (pp. 167-183). Springer, Berlin, Heidelberg.

Garg, V., Camp, L. J., & Kanich, C. (2013). Analysis of ecrime in crowd-sourced labor markets: Mechanical turk vs. freelancer. In The economics of information security and privacy (pp. 301-321). Springer, Berlin, Heidelberg.Grimes, G. A. (2007). Compliance with the CAN-SPAM Act of 2003. Communications of the ACM, 50(2), 56-62.

Hui, K. L., Hui, W., & Yue, W. T. (2012). Information security outsourcing with system interdependency and mandatory security requirement. Journal of Management Information Systems, 29(3), 117-156.

Hui, K. L., Kim, S. H., & Wang, Q. H. (2017). Cybercrime deterrence and international legislation: Evidence from distributed denial of service attacks. MIS Quarterly, 41(2), 497.

Hui, K. L., Ke, P. F., Yao, Y., & Yue, W. T. (2019). Bilateral Liability-Based Contracts in Information Security Outsourcing. Information Systems Research, 30(2), 411-429.

Jordan, T., & Taylor, P. (1998). A sociology of hackers. The Sociological Review, 46(4), 757-780.

Jordan, T., & Taylor, P. (2004). Hacktivism and cyberwars: Rebels with a cause?. Routledge.

Katyal, N. K. (2001). Criminal law in cyberspace. University of Pennsylvania Law Review, 149(4), 1003-1114.

Kigerl, A. C. (2016). Deterring spammers: impact assessment of the CAN SPAM act on email spam rates. Criminal Justice Policy Review, 27(8), 791-811.

Kshetri, N. (2006). The simple economics of cybercrimes. IEEE Security & Privacy, 4(1), 33-39.

Kunreuther, H., & Heal, G. (2003). Interdependent security. Journal of risk and uncertainty,

26(2-3), 231-249.

Kwon, J., & Johnson, M. E. (2011, June). An Organizational Learning Perspective on Proactive vs. Reactive investment in Information Security. In WEIS.

Kwon, J., & Johnson, M. E. (2018). Meaningful Healthcare Security:: Does Meaningful-Use Attestation Improve Information Security Performance?. MIS Quarterly, 42(4), 1043-1067.

Lee, C. H., Geng, X., & Raghunathan, S. (2013). Contracting information security in the presence of double moral hazard. Information Systems Research, 24(2), 295-311.

Lee, Y. (2005). The CAN-SPAM Act: a silver bullet solution?. Communications of the ACM, 48(6), 131-132.

Leeson, P. T., & Coyne, C. J. (2005). The economics of computer hacking. JL Econ. & Pol'y, 1, 511.

Leontiadis, N., Moore, T., & Christin, N. (2014, November). A nearly four-year longitudinal study of search-engine poisoning. In Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (pp. 930-941). ACM.

Lott, J. R. (2013). More guns, less crime: Understanding crime and gun control laws. University of Chicago Press.

Mahmood, M. A., Siponen, M., Straub, D., Rao, H. R., & Raghu, T. S. (2010). Moving toward black hat research in information systems security: an editorial introduction to the special issue. MIS quarterly, 34(3), 431-433.

Majoras, D. P., Leary, T. B., Harbour, P. J., & Leibowitz, J. (2005). Effectiveness and enforcement of the CAN-SPAM Act: A report to Congress. Federal Trade Commission.

McCoy, D., Dharmdasani, H., Kreibich, C., Voelker, G. M., & Savage, S. (2012, October). Priceless: The role of payments in abuse-advertised goods. In Proceedings of the 2012 ACM conference on Computer and communications security (pp. 845-856). ACM.

Miller, A. R., & Tucker, C. E. (2011). Encryption and the loss of patient data. Journal of Policy Analysis and Management, 30(3), 534-556.

Mitra, S., & Ransbotham, S. (2015). Information disclosure and the diffusion of information security attacks. Information Systems Research, 26(3), 565-584.

Moore, T., & Anderson, R. (2012). Internet security. The Oxford Handbook of the Digital Economy'(Oxford University Press 2011).

Moore, T., Clayton, R., & Anderson, R. (2009). The economics of online crime. Journal of Economic Perspectives, 23(3), 3-20.

Moore, T., Friedman, A., & Procaccia, A. D. (2010, September). Would a'cyber warrior'protect us: exploring trade-offs between attack and defense of information systems. In Proceedings

of the 2010 New Security Paradigms Workshop (pp. 85-94). ACM.

Mukhopadhyay, A., Chatterjee, S., Bagchi, K. K., Kirs, P. J., & Shukla, G. K. (2019). Cyber risk assessment and mitigation (CRAM) framework using logit and probit models for cyber insurance. Information Systems Frontiers, 21(5), 997-1018.

Mukhopadhyay, A., Chatterjee, S., Saha, D., Mahanti, A., & Sadhukhan, S. K. (2013). Cyber-risk decision models: To insure IT or not?. Decision Support Systems, 56, 11-26.

Nagle, F., Ransbotham, S., & Westerman, G. (2017). The effects of security management on security events. In Annual Workshop on the Economics of Information Security.

Ooi, K. W., Kim, S. H., Wang, Q. H., & Hui, K. L. (2012). Do hackers seek variety? an empirical analysis of website defacements. AIS.

Png, I. P., Wang, C. Y., & Wang, Q. H. (2008). The deterrent and displacement effects of information security enforcement: International evidence. Journal of Management Information Systems, 25(2), 125-144.

Png, I. P., & Wang, Q. H. (2009). Information security: Facilitating user precautions vis-à-vis enforcement against attackers. Journal of Management Information Systems, 26(2), 97-121.

Police Executive Research Forum. 2014. "The Role of Local Law Enforcement Agencies in Preventing and Investigating Cybercrime." In. Police Executive Research Forum Washington, DC.

Police Executive Research Forum. 2018. "The Changing Nature of Crime And Criminal Investigations." In. Police Executive Research Forum Washington, DC.

Polinsky, A. M., & Shavell, S. (2007). The theory of public enforcement of law. Handbook of law and economics, 1, 403-454.

Posner, R. A. (1979). Optimal sentences for white-collar criminals. Am. Crim. L. Rev., 17, 409.

Ramzan, Z. (2010). Phishing attacks and countermeasures. In Handbook of information and communication security (pp. 433-448). Springer, Berlin, Heidelberg.

Ransbotham, S. (2010, June). An Empirical Analysis of Exploitation Attempts Based on Vulnerabilities in Open Source Software. In WEIS.

Ransbotham, S., & Mitra, S. (2009). Choice and chance: A conceptual model of paths to information security compromise. Information Systems Research, 20(1), 121-139.

Ransbotham, S., Mitra, S., & Ramsey, J. (2012). Are Markets for Vulnerabilities Effective?. MIS Quarterly, 43-64.

Rao, J. M., & Reiley, D. H. (2012). The economics of spam. Journal of Economic Perspectives, 26(3), 87-110.

Raphael, S., & Winter-Ebmer, R. (2001). Identifying the effect of unemployment on crime. The Journal of Law and Economics, 44(1), 259-283.

Romanosky, S., Telang, R., & Acquisti, A. (2011). Do data breach disclosure laws reduce identity theft?. Journal of Policy Analysis and Management, 30(2), 256-286.

Schell, B. H., & Dodge, J. L. (2002). The hacking of America: Who's doing it, why, and how. Greenwood Publishing Group Inc..

Stigler, G. J. (1970). The optimum enforcement of laws. Journal of Political Economy, 78(3), 526-536.

Taylor, P. (2012). Hackers: Crime and the digital sublime. Routledge.

Thomas, D. (2002). Hacker culture. U of Minnesota Press.

Turgeman-Goldschmidt, O. (2005). Hackers' accounts: Hacking as a social entertainment. Social Science Computer Review, 23(1), 8-23.

Varian, H. (2004). System reliability and free riding. In Economics of information security (pp. 1-15). Springer, Boston, MA.

Vasek, M., Wadleigh, J., & Moore, T. (2015). Hacking is not random: a case-control study of webserver-compromise risk. IEEE Transactions on Dependable and Secure Computing, 13(2), 206-219.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. Journal of the American Statistical Association, 60(309), 63-69.

Xu, Z., Hu, Q., & Zhang, C. (2013). Why computer talents become computer hackers. Communications of the ACM, 56(4), 64-74.

Yue, W.T., Wang, Q. H., & Hui, K. L. (2019). See no evil, hear no evil? Dissecting the impact of online hacker forums. MIS Quarterly, 43(1), 73.

Zhou, J., & Hui, K. L. (2019). Bug Bounty Programs, Security Investment and Law Enforcement: A Security Game Perspective.