CRIMINOLOGY

RESTRICTIVE DETERRENT EFFECTS OF A WARNING BANNER IN AN ATTACKED COMPUTER SYSTEM*

DAVID MAIMON,[1] MARIEL ALPER,[1] BERTRAND SOBESTO,[2] and MICHEL CUKIER[2]

[1]Department of Criminology and Criminal Justice, University of Maryland
[2]Department of Reliable Engineering, University of Maryland

KEYWORDS: cybercrime, deterrence, restrictive deterrence, honeypots, experiments

System trespassing by computer intruders is a growing concern among millions of Internet users. However, little research has employed criminological insights to explore the effectiveness of security means to deter unauthorized access to computer systems. Drawing on the deterrence perspective, we employ a large set of target computers built for the sole purpose of being attacked and conduct two independent experiments to investigate the influence of a warning banner on the progression, frequency, and duration of system trespassing incidents. In both experiments, the target computers (86 computers in the first experiment and 502 computers in the second) were set either to display or not to display a warning banner once intruders had successfully infiltrated the systems; 1,058 trespassing incidents were observed in the first experiment and 3,768 incidents in the second. The findings reveal that although a warning banner does not lead to an immediate termination or a reduction in the frequency of trespassing incidents, it significantly reduces their duration. Moreover, we find that the effect of a warning message on the duration of repeated trespassing incidents is attenuated in computers with a large bandwidth capacity. These findings emphasize the relevance of restrictive deterrence constructs in the study of system trespassing.

System trespassing, which is defined as "illegally gaining access to one or more computer systems after exploiting security vulnerabilities or defeating a security barrier" (McQuade, 2006: 83), is one of the fastest growing areas of cybercrime (Furnell, 2002). According to a recent survey of more than 580 information technology (IT) practitioners employed by large organizations and governmental agencies, 90 percent of U.S. corporations, both private and public, experienced multiple incidents of system trespassing during 2010 (Ponemon Institute, 2011; Whitman, 2003). These breaches are estimated to result

* Additional supporting information can be found in the listing for this article in the Wiley Online Library at http://onlinelibrary.wiley.com/doi/10.1111/crim.2014.52.issue-1/issuetoc.

This research was conducted with the support of the SANS Institute, the National Consortium for the Study of Terrorism and Responses to Terrorism in the University of Maryland, and the National Science Foundation Award 1223634. We thank Lawrence Sherman, Jean McGloin, Ray Paternoster, and Theodore Wilson for their helpful comments throughout the project. We also wish to thank Gerry Sneeringer and the Security Team of the Office of Information Technology at the University of Maryland for their insights on this research. Finally, we thank Wayne Osgood and the four anonymous reviewers for their helpful comments on this paper. Direct correspondence to David Maimon, Department of Criminology and Criminal Justice, University of Maryland 2220 LeFrak Hall, College Park, MD 20742 (email: dmaimon@umd.edu).

 doi: 10.1111/1745-9125.12028

in billions of dollars of financial losses annually, as well as in serious invasion of privacy for both customers and employees (Whitman, 2003). Nevertheless, despite the growing public and legal awareness of system trespassing and its consequences for commercial, governmental (Rantala, 2008), and individual computer users (Bossler and Holt, 2009), only scant attention has been given to this phenomenon in the criminological literature (Skinner and Fream, 1997).

Addressing this challenge, this work explores the effectiveness of sanction threats in attacked computer systems in preventing the progression, reducing the frequency, and shortening the duration of system trespassing incidents. Specifically, focusing on recent extensions of deterrence theory (Gibbs, 1975; Jacobs, 2010), we seek to answer four research questions. First, does a warning banner, displayed when a system trespasser intrudes on an information system for the first time, result in immediate termination of the system trespassing session? Second, does this warning banner reduce the frequency of repeated system trespassing incidents on the target computer? Third, does a warning banner affect the duration of first and repeated system trespassing incidents? And last, do varying computer configurations condition the effect of the warning banner on the duration of system trespassing incidents? To answer these questions, we designed a randomized trial using a large set of target computers built for the sole purpose of being attacked. This research design allows experimental investigation of the role of deterring cues in the development of first and repeated system trespassing incidents.

THEORETICAL BACKGROUND

SYSTEM TRESPASSING

Similar to trespassing in the physical world, system trespassing involves the violation of a use restriction on property by someone who has no right to access the property (Brenner, 2010). Overall, unauthorized users can access a computer either locally, by gaining physical access to it, or remotely, by logging in via the Internet (Anderson, 1980; Stallings, 2005). Depending on the motivation of the intruder (e.g., revenge, monetary gain, ideology, thrill, status, or addiction [McQuade, 2006; Wall, 2007; Yar, 2006]), the attacks could be harmless (e.g., exploring the Internet) or dangerous (e.g., reading and modifying privileged data, disrupting the system, using the system to attack other computers, or all of the above) for the target systems and their users (Stallings, 2005).

In an effort to gain remote unauthorized access to a system, system trespassers, who also are referred to as hackers or crackers (Furnell, 2002; Wall, 2007), randomly scan the Internet and look for open networked computer ports (Gadge and Patil, 2008). Once they have identified open ports, trespassers may use special software cracking tools—available for purchase and as open source software on the Internet—that systematically check all possible keys to a system until the correct one is found and access to the system is granted.[1] Once unauthorized access to a system is obtained, system trespassers may log

1. These powerful tools can generate millions of passwords in a short period of time using dictionary wordlists and smart rule sets in an effort to guess the right password to an account (Florêncio, Herley, and Coskun, 2007; Knudsen and Robshaw, 2011). Several tools even try different combinations of user names and passwords in their attempts to access a system, sidestepping the countermeasure of locking out a single account for failed password attempts.

in and out of the compromised system at any time, access and corrupt private information and files, and interrupt the ability of legitimate users to use the system. In addition, depending on the system configuration, intruders may use the compromised system to send spam e-mail, set up fake websites, launch denial-of-service (DoS) attacks against various network targets to deny legitimate users access to network resources (Garfinkel, Spafford, and Schwartz, 2003), or employ the system for launching subsequent system trespassing incidents to intrude on other computers (Brenner, 2010; McQuade, 2006).

To mitigate system trespassing and allow more secure and protected computing environments, extensive efforts have been made during the last 20 years to develop technical solutions for detecting and preventing unauthorized access to computers (Allen and Stoner, 2000; Mackey, 2003). Moreover, responding to the pressing need to generate deterrence against the operations of system trespassers, Congress enacted the Computer Fraud and Abuse Act in 1986, which allows for up to 10 years of imprisonment for computer misuse offenses (Kerr, 2009). The belief that credible threats of apprehension will deter computer crime offenders is based on the general assumption that the behaviors of individuals can be altered by the threat and imposition of punishments (Paternoster, 1987; Tittle, 1980).

DETERRENCE THEORY

Embedded in utilitarian philosophical principles (Beccaria, 1963 [1764]; Bentham, 1970 [1785]), deterrence theory advances the view that fear of sanctions and punishments inhibit the involvement of individuals in deviance and crime (Cusson, 1993; Gibbs, 1975; Paternoster, 1987). Focusing on the cost aspect of the cost–benefit equation, this classic theory emphasizes the influence of the perceptions of sanctions severity, certainty, and celerity in generating effective deterrence. However, during the last three decades, this theory has been advanced in three important ways. First, deterrence scholars now distinguish between the effect of punishments and sanction threats on punished offenders (i.e., specific deterrence) and the effect on the general public (i.e., general deterrence) (Paternoster and Piquero, 1995). Second, prompted by the seminal work of Becker (1968), contemporary scholars differentiate between the unique influence of objective (i.e., actual risks of punishment and apprehension) and subjective sanctions (i.e., individual perceptions of the certainty, severity, and celerity of punishment) on the probability of an individual to offend (Paternoster, 1987). Finally, contemporary theoreticians distinguish between the ability of sanction threats to prevent criminal involvement completely (absolute deterrence) and the effect of punishment threats in reducing the frequency and severity of individual offending (restrictive deterrence) (Gibbs, 1975; Jacobs, 2010). Surprisingly, although extensive research has evaluated how the former two dimensions of deterrence shape involvement in crime by an individual (Pratt et al., 2006), only meager research has examined aspects of restrictive deterrence (Paternoster, 1987) and its impact on the expression of crime (Gallupe, Bouchard, and Caulkins, 2011; Jacobs, 1993, 1996a).

Restrictive deterrence was described by Gibbs (1975) as "the curtailment of a certain type of criminal activity by an individual during some period because in whole or in part the curtailment is perceived by the individual as reducing the risk that someone will be punished as a response to the activity" (1975: 33). According to Gibbs (1975), restrictive deterrence applies to offenders who committed an act of crime at least once and is concerned primarily with reduction in the frequency of offending. Specifically, because

offenders believe that they will be punished for their offending at some point in time, they reduce the frequency of their offending to delay this point and exacerbate detection efforts (probabilistic deterrence).

More recent works refined these insights by Gibbs (1975) in two important ways. First, several scholars (Jacobs, 1996a; Tittle, 1980) proposed that restrictive deterrence could be a function of both specific and general deterrence. Similarly to ex-offenders, individuals who had never previously committed an act of crime also could employ restrictive deterrence strategies while committing their first offense. Second, Jacobs (2010) proposed that along with reducing the frequency of their offending, offenders restrict the scope of their deviance using other strategies that are consistent with restrictive deterrence, including reducing the seriousness of criminal acts, engaging in situational measures that reduce the risk of detection, and switching the locations and timing of criminal events (particularistic deterrence).

Although Gibbs contended that the "deterrent effect of punishment is largely restrictive" (1975: 34), systematic empirical investigations of this concept are still preliminary and relatively scarce (Jacobs, 1993, 1996a, 1996b; Jacobs and Cherbonneau, 2012; Jacobs and Miller, 1998; Gallupe, Bouchard, and Caulkins, 2011; Paternoster, 1989; Wright and Decker, 1994). Specifically, most of the prior research that has dealt with this concept has been qualitative in nature (Jacobs, 1993, 1996a, 1996b; Jacobs and Cherbonneau, 2012) and has drawn on relatively small samples (Beauregard and Bouchard, 2010; Jacobs, 1996b; Jacobs and Cherbonneau, 2012). Thus, whereas these studies have been crucial to our understating of the adoption of situational measures by offenders that reduce the risk of detection during the course of criminal incidents, they have not generated a direct link between the presence of sanction threats in the environment and the restriction of the scope of criminal activities by the offenders. Moreover, these studies have revealed little regarding the relationships between the presence of deterring cues in the environment and the progression of a criminal event. We attempt to bridge this empirical gap by examining the effect of deterring messages (i.e., warnings) on the progression and duration of system trespassing incidents.

DETERRENCE AND WARNINGS

Advancing the view that a system of deterrence is a communication mechanism that informs potential offenders about the probability of someone detecting their criminal act and their likelihood of being punished for their criminal behavior, Geerken and Gove (1975) proposed that a successful deterrence process is determined by the degree to which a deterring message is conveyed to a target audience. Consistent with the assertions of Geerken and Gove (1975), we contend that posting warnings represents one important avenue for transmitting coherent deterring messages to potential offenders (Clarke, 1997; Cusson, 1993). Indeed, extensive research has assessed the effectiveness of warning signs in preventing the occurrence of criminal incidents (Coleman, 2007; Decker, 1972; Eck and Wartell, 1998; Goldstein, Cialdini, and Griskevicius, 2008; Green, 1985; Schultz and Tabanico, 2009; Schwartz and Orleans, 1967; Slemrod, Blumenthal, and Christian, 2001; Wenzel and Taylor, 2004). However, results from these studies have been mixed. A few of these works have suggested that warnings are effective in deterring illegal behavior like claim padding of insured persons (Blais and Bacher, 2007), thefts of Jobseekers' Allowance unemployment checks (Tilley, 2005), and unsafe driving

(Rama and Kulmala, 2000). In contrast, other studies have indicated that warning and deterring signs carry no effect on prostitution (Lowman, 1992), theft of cable television signals (Green, 1985), and illicit parking meter use (Decker, 1972). Still other research has demonstrated that the presence of warning signs may encourage the development of petty crimes like pickpocketing (Ekblom, 1991; Grabosky, 1996). In an effort to explain this perverse effect, Grabosky (1996) suggested that in some cases, warnings may advertise the unacceptable behavior, excite potential offender curiosity, and entice the rebellious nature of the offenders.

Importantly, although most of these studies have assessed the effectiveness of warnings in preventing the *occurrence* of a criminal event, almost no studies have investigated the effect of warning signs on the progression and *duration* of a criminal incident (for an exception, see Green, 1985). In line with the claims by both Gibbs (1975) and Jacobs (2010), we believe that differentiating between the occurrence and progression of a criminal event is of theoretical importance because although sanction threats may not prevent the emergence of a criminal incident, they may still alter its characteristics. Specifically, it could be the case that the presence of sanction threats in the environment may not cause offenders to abort the criminal event (Cusson, 1993) but simply cause offenders to change their course of criminal action (Green, 1985; Guerette and Bowers, 2009; Jacobs and Cherbonneau, 2012).

Moreover, whereas the relationships between deterring cues and crime had been examined extensively in the physical world (for instance, Sherman and Weisburd, 1995), no prior study has investigated the impact of warnings on deviant behaviors that take place in cyberspace. However, extensive theoretical literature has debated the application of denial (i.e., physically preventing an attacker from obtaining a threatening technology), prevention (i.e., the use of defensive measures to disrupt an attack), and futility (i.e., using means to minimize the effect of a successful attack in the system) strategies for deterring the development of cyber attacks (Elliott, 2011; Geers, 2012; Goodman, 2010; Harknett, 1996). Although some theoreticians have proposed that because of the anonymity inherent in the Internet and the difficulty associated with tying cyber criminals to their crimes, the role of deterrence is minimal in the context of cyberspace (Blank, 2001; Harknett, 1996); others have proposed that attributing cyber attacks to specific individuals is not necessary for deterring the activities of cyber criminals (Goodman, 2010). Unfortunately, despite this theoretical scholarship, no empirical initiatives have been launched to assess the effectiveness of punishment threats in cyberspace. Thus, drawing on the assumption that exposure to deterrent messages is the first necessary condition to deter cyber aggression (Geerken and Gove, 1975; Goodman, 2010), we explore the impact of warning banners posted on computer systems on the progression and duration of system trespassing incidents.

THIS STUDY

One important recommendation, proposed by the National Institute for Standards and Technology for minimum-security controls in governmental, industrial, and private agencies (NIST, 2009), encourages information technology (IT) managers in governmental and proximate organizations to display an approved system use notification message before granting users access to the system. According to the guidelines, this notification should be implemented in the form of a warning banner that includes 1) the

organizational policies regarding unauthorized access and use of the system and 2) the criminal and civil penalties that are associated with trespassing. Our first goal in this work is to assess the effectiveness of this type of warning in determining the progression of a first system trespassing incident. Specifically, we explore whether the presentation of a warning banner during the first time a system trespasser accesses the system results in immediate cessation of the trespassing incident. However, because no prior research has studied the effect of warnings in cyberspace (Goodman, 2010) and because of the mixed results reported in the criminological literature regarding the effectiveness of warnings in preventing illegitimate behaviors in the physical world (Ariel, 2012; Cusson, 1993; Green, 1985), we hypothesize that *a warning banner in an attacked computer system could discourage, encourage, or have no effect on the progress of a first system trespassing incident.*

Drawing on the restrictive deterrence literature (Gibbs, 1975; Jacobs, 2010), we also assess the impact of deterring messages in influencing the *frequency* of repeated system trespassing incidents on the target computer. Consistent with the discussion of "probabilistic deterrence" proposed by Gibbs (1975), we suspect that once encountering a deterring message in an attacked computer system, system trespassers will attempt to avoid detection and punishment for their illegal acts by reducing the frequency of repeated trespassing events on the compromised system. Therefore, we hypothesize that *the presence of a sanction threat in an attacked computer system reduces the frequency of repeated trespassing incidents on the target computer.*

Similarly, adopting the assumption that offenders respond to sanction threats by restricting the scope of their criminal behaviors (Jacobs, 2010), we hypothesize that *the presence of a warning banner in an attacked computer system shortens the duration of both first and repeated system trespassing incidents.* Indeed, prior research has indicated that limiting the time of possessing a stolen vehicle is a common strategy that is used by automobile thieves to avoid apprehension by the police (Jacobs and Cherbonneau, 2012). However, given the absence of relevant data, no previous study has investigated the relationship between the presence of deterring cues in the environment and the duration of a criminal incident. Consistent with prior findings indicating that computer attackers use various strategies, including limiting exposure on the system, to evade detection by IT personnel and intrusion detection systems (Wagner and Soto, 2002), we suspect that the presence of sanction threats on the target computer triggers conscious/unconscious wariness among perpetrators, which in turn attenuates their willingness to expose themselves on the attacked system and shortens the duration of a trespassing incident.

Finally, because system trespassing incidents take place in heterogeneous computing environments, we suspect that attributes of these environments convey different opportunities and risks for detection, which in turn condition the effect of a warning on the duration of a system trespassing incident. Indeed, prior research has indicated that property offenders pick up important cues from the environment and rely on these cues before deciding to initiate a criminal act; Weaver and Carroll (1985), for instance, reported that shoplifters observe potential "facilitators of shoplifting," such as store layout and accessibility of items, before making judgments about criminal opportunities in a store setting. Moreover, several studies have indicated that the effectiveness of posted instructions and warning signs in shaping the actions of individuals depends on the context in which these warnings are advertised (Ariel, 2012; Keizer, Lindenberg, and Steg, 2008; Slemrod, Blumenthal, and Christian, 2001). However, no prior research has tested the interactive effect

between deterring cues and characteristics of the environment on the duration of a criminal incident.

Acknowledging the role of computing environment in determining the functionality of a system, we suspect that the random access memory (RAM) size (i.e., the ability of a computer to process data quickly) and bandwidth capacity (i.e., the amount of data that can be carried from one computer to another per unit of time) of an attacked computer are important characteristics that facilitate varying opportunities for the development of a system trespassing incident. Specifically, systems with small RAM size access and process information more slowly than systems with large RAM size. Similarly, computers with low bandwidth capacity offer a slower communication and data transfer rate than computers with high-bandwidth-capacity connections (Tanenbaum, 2006). As a result, the execution of commands and transfer of information on small-RAM/low-bandwidth-capacity computers is slow and requires spending longer periods of time when working with these systems. Because limited functionality and longer processing time on a system introduce a greater probability of detection and fewer opportunities for subsequent operations, it is possible that when system trespassers encounter a deterring message on such computers, they will be more likely to comply with it. Accordingly, our final research hypothesis suggests that *large RAM size and high bandwidth capacity of the target computer system attenuate the effect of a warning banner on the duration of a system trespassing incident in such a way that the effect of a warning banner will be less pronounced on such computers than on low-RAM and low-bandwidth-capacity computers.*

EXPERIMENT 1

The goal of our first experiment was to determine the impact of a warning message in the target computer system on the progression, frequency, and duration of system trespassing incidents. In line with the legal definition of system trespassing (McQuade, 2006) and prior conceptualizations of system trespassing incidents (Alata et al., 2006; Berthier and Cukier, 2009), we operationalize a system trespassing incident as any event in which an unauthorized person accesses and logs in to a computer system. To achieve our research goal, we designed a randomized experiment employing a series of targeted computers called "honeypots."

A honeypot is a "security resource whose value lies in being probed, attacked or compromised" (Spitzner, 2002: 40). This technical tool is a real computer that serves as a flexible decoy and permits the collection of information on intruders and "live" attacks. Because honeypots have no production value, any network activity that is sent their way or initiated by them means that system trespassers have successfully infiltrated the system, the system has been compromised, and the target computer is used for the malicious operations of intruders. Although information technology managers use *production honeypots* for detecting and mitigating attacks against their networks, cybersecurity scholars employ *research honeypots* to explore who the attackers are, what they are doing on the compromised systems, and what kind of tools they use (Spitzner, 2002).

Indeed, several previous studies have employed research honeypots to generate a better understanding of the etiology of system trespassing. Alata et al. (2006), for instance, collected 38 attack sessions (i.e., unique system trespassing incidents from start to end) over a period of 131 days and provided preliminary results about the skills and attack patterns of system trespassers. A more extensive study was conducted by Berthier and

Cukier (2009), where the authors studied 1,171 attack sessions and analyzed 250 examples of rogue software collected over a period of 8 months. Finally, Salles-Loustau et al. (2011) recorded a total of 211 system trespassing incidents over a period of 167 days and examined evidence at each stage of the trespassing sequence, from discovery to intrusion and exploitation of software. Like these studies, we use honeypots (which will be identified as target computers in this work) to study the progress and development of system trespassing incidents. However, in contrast to past research that has focused solely on technical aspects of system trespassing, we draw on deterrence theory and design a randomized trial to assess the impact of an intervention (i.e., a warning) on the progress and development of system trespassing incidents.

DESIGN

In our first experiment, we used 80 public Internet Protocol (IP) addresses that were provided to us by the information technology team of a large American university and deployed identical target computers on the university network.[2] These target computers were set up as computer systems with the Linux Ubuntu 10.04 operating system (Canonical Group Limited, London, U.K.). To gain access to the target computers, system trespassers had to break into these systems successfully through frequently scanned and vulnerable entry points. After infiltrating the target computers, trespassers were assigned to either a treatment or a control target computer, and a system trespassing incident was initiated. To allow the collection of meaningful data on system trespassing incidents, we monitored the different components of the system trespassing incident using specialized software (Sebek, gateway, and OpenVZ hosts and containers) that records the system trespassing sessions for later analysis. Our focus in this work is on the first and repeated system trespassing incidents recorded on the target computers.

PROCEDURES

Unlike common experimental designs that require active subject recruitment, we did not recruit subjects to participate in our experiment. Instead, we deployed our target computers on the university network for a period of 2 months (April 1 to May 30, 2011) and waited for system trespassers to find our systems and employ special software cracking tools (McQuade, 2006) to break into them. To simulate a genuine environment, the target computers were modified to reject the login attempts by system trespassers on its public IP addresses until a predefined number of attempts. The predefined threshold was a random number between 150 and 200. When this threshold was reached, the target computer was "successfully" infiltrated and allowed the intruder access to the system by creating a new user with the latest credentials attempted by the system trespasser.[3]

Once access to our target computer had been granted, system trespassers were randomly assigned to either a warning (treatment) or a no-warning (control) computer and initiated a system trespassing incident. When assigned to a warning target computer, the

2. An IP address is an identifier for a computer or device on the Internet network. Networks that use the standard Internet protocol (i.e., TCP/IP protocol) route Internet traffic based on the IP address of the destination.

3. To limit the number of deployed target computers per attacker IP address, the system rejected any login attempt from an IP address that had already deployed target computers.

following message appeared on the screen of the intruder immediately after he or she broke into the system successfully:

> The actual or attempted unauthorized access, use, or modification of this system is strictly prohibited. Unauthorized users are subject to institutional disciplinary proceedings and/or criminal and civil penalties under state, federal, or other applicable domestic and foreign laws. The use of this system is monitored and recorded for administrative and security reasons. Anyone accessing this system expressly consents to such monitoring and is advised that if monitoring reveals possible evidence of criminal activity, the Institution may provide the evidence of such activity to law enforcement officials.

In contrast, when assigned to a no-warning target computer, no message appeared on the screen of the intruder.

To assess the effect of warnings, we allowed system trespassers to employ the target computers and initiate repeated system trespassing incidents for a period of 30 days—including sharing the target computer with friends, allowing access to other intruders, renting it, using it to attack other computer systems, and so on. To ensure that trespassers do not engage in activities that jeopardize our computer networks, and those of others, we constantly monitored their activity. Using special software (Sebek keylogger), we then recorded each trespassing incident. At the end of a 30-day period, we blocked access to the target computer by the system trespasser, cleaned it, and redeployed it on the network, so if a system trespasser wanted to regain access to the system, then she or he had to break into the system again before initiating a system trespassing incident.

During the 2 months of the experimental period, 86 target computers were deployed and infiltrated by system trespassers (42 of the computers had a warning banner installed), and 971 system trespassing incidents were recorded; 451 of the system trespassing incidents were recorded on the no-warning computers, and 520 sessions were recorded on the warning treatment computers. Importantly, most of the target computers experienced repeated system trespassing incidents. Information regarding the actual number of target computers deployed for the treatment and control conditions, as well as the number of system trespassing incidents recorded on these computers, is presented in appendix A in the online supporting information.[4] In an effort to address our list of research hypotheses, we first run our analyses using data on the first system trespassing incidents only ($n = 86$ trespassing sessions), and then we employ data on the entire poll of trespassing incidents recorded during the experimental period ($N = 971$ sessions).

OUTCOME MEASURES

Because system trespassing incidents are similar to other criminal events in the sense that they have a beginning and an end, we timed the start and termination points of each trespassing session and calculated the duration each incident lasted. We then created two dependent measures. The first measure, *immediate incident cessation*, is a dummy measure (1 = immediate incident cessation) indicating the termination of a trespassing incident

4. Additional supporting information can be found in the listing for this article in the Wiley Online Library at http://onlinelibrary.wiley.com/doi/10.1111/crim.2014.52.issue-1/issuetoc.

after a period of 5 seconds from its start.[5] The second measure, *incident duration*, is a continuous measure that taps the elapsed time (in seconds) between the beginning and the end of a system trespassing incident.

RESULTS

First Trespassing Incidents

We begin with analyzing the *first* trespassing incidents recorded on each target computer, and we test for a significant difference between the proportions of immediate incident cessation on warning and no-warning target computers. To achieve this goal, we perform a t test for comparing two proportions, with immediate incident cessation as a dependent variable. The results from this test revealed a nonsignificant main effect for warning ($Z = -1.46$, $p > .05$). Specifically, although the proportion of immediate incident termination is larger on the warning than on the no-warning computers (40 percent vs. 25 percent), the main effect of the warning is insignificant.

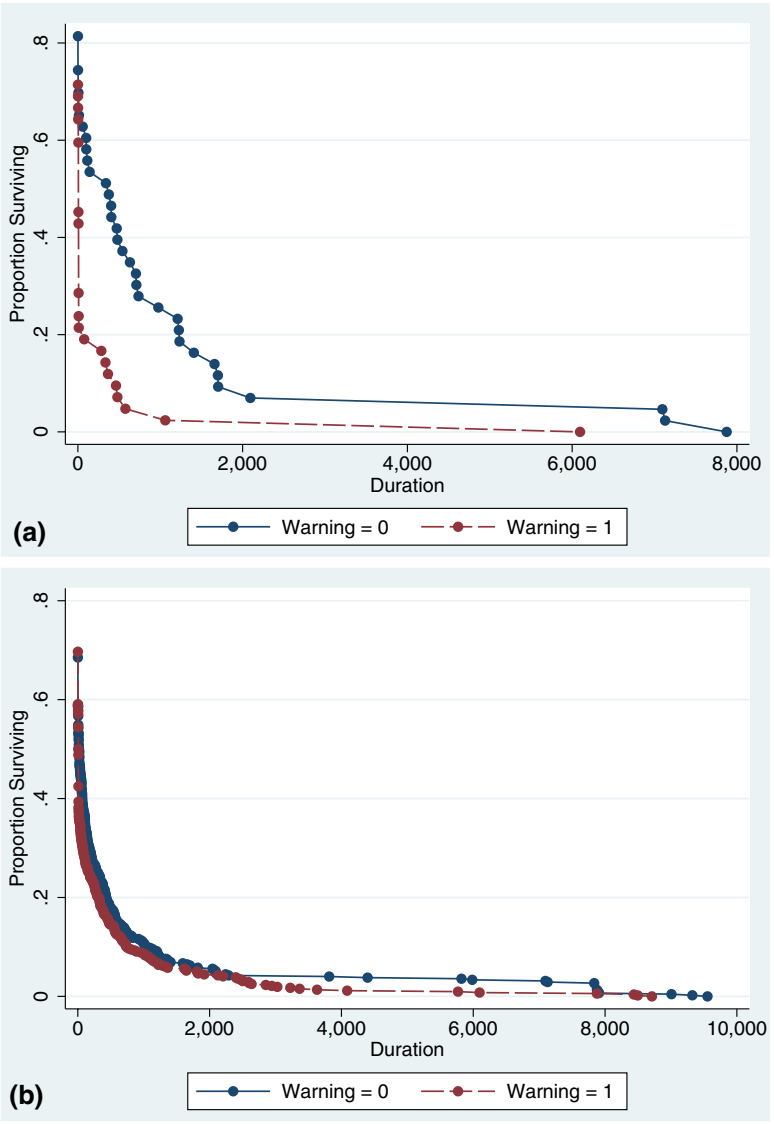
Turning to an exploration of the effect of a sanction threat on the duration of trespassing incidents, we investigate the influence of a warning banner on the survival time of system trespassing incidents. However, because of the right-skewed distribution of the survival time of trespassing incidents, we cannot simply compare the average durations of system trespassing incidents on the warning and no-warning target computers. Instead, we employ event history analysis techniques that allow for estimating and comparing the proportion of units surviving an event [i.e., survival function $S(t)$], as well as a prediction of the rate at which durations end [i.e., hazard rate $h(t)$] (Box-Steffensmeier and Jones, 2004).

We first use standard life table methods to examine the effect of a warning on the time until termination of the system trespassing incidents (Namboodiri and Suchindran, 1987). To determine whether warning banners influence the time until termination, we compare the survival distribution of *first* trespassing incidents observed on target computers with a warning banner with the corresponding survival distribution of *first* trespassing incidents recorded on computers with no warning message. The results from this analysis are presented in figure 1a. As indicated in the figure, across all time points, the proportion of first trespassing incidents that survived longer periods of time is smaller on the treatment (warning) than on the control (no-warning) target computers.

To test whether the effect of a warning on the duration of system trespassing incidents is significant, we assess the impact of a warning banner on the hazard of incident termination by generating a dummy variable indicating whether a system trespassing incident was recorded on a warning or a no-warning target computer (1 = warning) and estimating a Cox proportional-hazard regression (Box-Steffensmeier and Jones, 2004; Walters, 2009). Similar to a simple regression, the Cox model aims to explore the relationships between dependent and independent variables. However, in contrast to a simple ordinary least-squares regression, a Cox model allows investigations of the relationships

5. For both experiments discussed in this article, we chose 5 seconds as a cutoff threshold because we wanted to ensure sufficient time for attackers to see and read the banner. However, in analyses not shown, we also tested the effect of a warning banner on immediate session cessation while using different cut points (0, 3, and 7 seconds). The results from these analyses were identical to those reported in this article.

Figure 1. Time to System Trespassing Incident Termination—Experiment 1



(a) First trespassing incidents ($n = 86$)
(b) All trespassing incidents ($N = 971$)

between the survival of the event and independent measures of interest (Box-Steffensmeier and Jones, 2004). The results from the estimated Cox model are presented in table 1, model 1. Consistent with our research hypothesis, model 1 confirms that a warning banner in the target computer is positively associated with the hazard of first system trespassing incident termination. Specifically, the hazard ratio estimate of our warning

Table 1. System Trespassing Incident Duration Regressed over Warning Configuration (Experiment 1)

Variables	Model 1		Model 2	
	First Observed Incidents		All Observed Incidents	
	(Cox Regression, $n = 86$)		(Frailty Model, $N = 971$)	
	Coefficient (SE)	Hazard Ratio	Coefficient (SE)	Hazard Ratio
Warning	.97***	2.62	.26*	1.29
	(.26)		(.13)	
Theta	—	—	.23***	—
Log likelihood	−233.08***		−4,601.05**	

ABBREVIATION SE = standard error.
*$p < .05$ (two-tailed); **$p < .01$ (two-tailed); ***$p < .001$ (two-tailed).

measure indicates that a warning banner more than doubles the rate of first system trespassing incident termination, and results in shorter duration of first trespassing incidents.[6]

All Trespassing Incidents Recorded

Turning to an investigation of the effect of a warning banner on the volume of repeated trespassing incidents, we employ information from the entire poll of trespassing incidents recorded against our target computers (i.e., 971 incidents), and we estimate whether the mean number of repeated trespassing incidents recorded on the warning computers is significantly different than the mean number of repeated trespassing incidents observed on the no-warning computers. The results from a *t* test for comparing the means of two groups reveal an insignificant difference between the averages of these groups ($t = –1.11$, $p > .05$). Accordingly, although the average number of trespassing incidents is higher on the warning than on the no-warning computers (12 on the warning target computers vs. 10 on the no-warning computers), this difference is insignificant. This finding reveals no support to our assumption that a warning banner reduces the frequency of repeated system trespassing incidents on the target computers.

Next, we compare the survival distributions of *all* system trespassing incidents recorded on the warning and no-warning computers.[7] Figure 1b presents results from this comparison. Similar to the pattern observed for the first trespassing incidents, this comparison reveals that the proportion of trespassing incidents that survived longer periods of time is smaller on the treatment (warning) than on the control (no-warning) target computers. To estimate the effect of a warning banner on the hazard of system trespassing incident cessation, we employ shared-frailty models (or random-effect models). Overall, these models

6. This finding is consistent with results obtained from a log-rank test for comparing the difference between the survival curves of two groups: log-rank chi square = 15.42, $p < .001$.
7. In an analysis not shown, we estimated whether the duration of first trespassing incidents is related to the number of repeated attacks against the system, computing a Pearson correlation coefficient. The finding from this test indicates an insignificant relationship between the first trespassing session duration and the number of repeated attacks against the system. A similar pattern was observed in experiment 2.

are unique extensions of the classic Cox model that account for the heterogeneity and dependence issues generated by repeated observations (Box-Steffensmeier, De Boef, and Joyce, 2007; Liu, Wolfe, and Huang, 2004).[8] Shared-frailty models are particularly useful in the context of our work because they allow us to estimate the effect of a warning on the hazard of trespassing incident termination while accounting for the frailty shared among all repeated trespassing incidents that are observed for the same target computer.

The results from our estimated shared-frailty model are reported in table 1, model 2. As indicated in the model, the effect of a warning banner in the target computers is positive and significant on the hazard of trespassing session termination. Accordingly, the hazard ratio estimate of our warning measure indicates that a warning banner increases the probability of trespassing session termination by 29 percent. This finding further confirms our second research hypothesis and demonstrates that a warning banner reduces the duration of system trespassing incidents on the attacked system.

EXPERIMENT 2

Our goal in the second experiment was to replicate the findings from experiment 1 while accounting for different system configurations that might moderate the effect of a warning banner on the duration of system trespassing incidents. Based on previous research that explored the influence of the interactive relationships between deterrence and opportunity on the occurrence of crime, we hypothesized that the RAM size and bandwidth capacity of a computer system condition the effect of a warning banner on the duration of system trespassing incidents. Specifically, we predicted that because smaller RAM size and lower bandwidth capacity require users to spend longer periods of time on the system, it is possible that when system trespassers encounter a deterring message on slower computers, they will be more likely to terminate the system trespassing incident earlier.

DESIGN AND PROCEDURE

In experiment 2, we used 300 public IP addresses that were provided to us by the IT team at a large American university, and we deployed our target computers on the university network. In line with our design in experiment 1, system trespassers had to infiltrate these target computers through a frequently scanned and vulnerable entry point. However, in contrast to experiment 1, in experiment 2, we employed a 2 [warning banner, no banner (control)] $\times$ 2 [low (512 Mbytes) RAM, high (2.25 Gbytes) RAM] $\times$ 2 [low (128 Kbits/s) bandwidth, high (512 Kbits/s) bandwidth] $\times$ 2 [low (5 Gbytes) disk space, high (30 Gbytes) disk space] factorial design. The advantage of this experimental setting is that it allows examination of the responses of system trespassers to a deterring stimulus in different computing environments.

We deployed our target computers on the university network for a period of 6 months (October 4, 2011 to April 3, 2012), and we waited for system trespassers

8. The underlying premise of the Cox model assumes that event times are independent. In the presence of correlated events and heterogeneity, the independence assumption is violated and leads to incorrect estimates of the standard errors of the model. To correct for this issue, the shared-frailty models assume that the unobserved effects across subjects/observations are commonly and randomly distributed across groups of observations.

to find our systems and employ special software cracking tools to break into them successfully. We built a genuine computer network environment by programming the target computers to deny login attempts by intruders on its public IP addresses until a predefined threshold was reached (the predefined threshold was a random number between 150 and 200). When this threshold was reached, the target computer was "successfully" infiltrated, intruders were assigned randomly to one of 16 target computer configurations, and intruders were allowed to initiate a system trespassing incident. System trespassers were allowed to employ the system for a period of 30 days; yet their activities were monitored closely. At the end of the 30 days, we prevented access to the target computer by the system trespasser, cleaned the computer, and redeployed it.

Overall, a total of 502 target computers (of which 259 had a warning banner installed) were deployed and infiltrated, and more than 3,700 system trespassing incidents were recorded (2,041 on target computers with a warning banner) during the 6 months of the experimental period. Similar to the pattern observed in experiment 1, most of the target computers experienced repeated system trespassing incidents. Data on the number of target computers deployed with each configuration, as well as the number of system trespassing incidents recorded on them, are presented in appendix B in the online supporting information. To answer our research questions, we take a similar approach to that we employed when analyzing data from our first experiment: The first round of analysis uses information on the first trespassing incidents only (i.e., $n = 502$ incidents), and the second round analyzes information from the entire poll of trespassing incidents (i.e., $N = 3{,}768$ incidents).

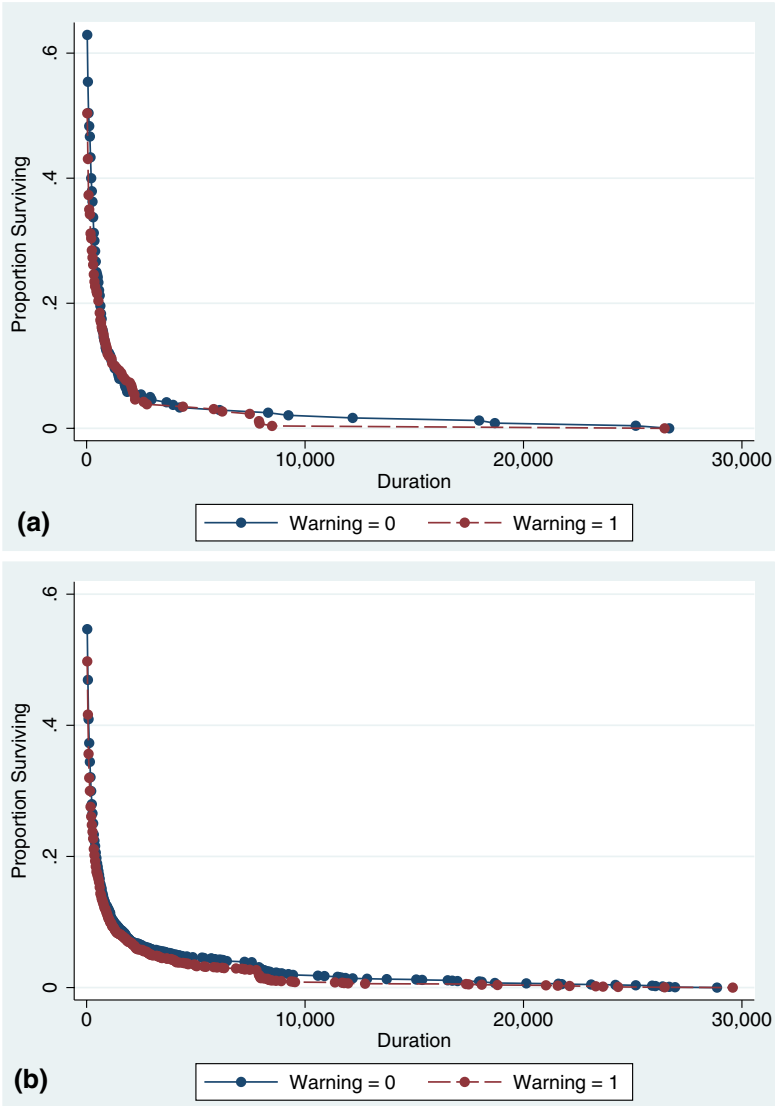
RESULTS

First Trespassing Incidents

Consistent with our first experiment, we generated two dependent measures. The first, *immediate incident cessation*, is a binary measure (1 = immediate incident cessation) indicating a trespassing incident terminated after a period of 5 seconds from its start. The second, *incident duration*, is a continuous measure that taps the elapsed time (in seconds) between the beginning and the end of a system trespassing incident. We then employed information from the *first* trespassing incident recorded on each target computer to explore the effect of a warning banner on immediate incident termination. Consistent with the findings reported for experiment 1, the results of a t test for comparing two proportions, with immediate incident cessation as a dependent variable, revealed an insignificant main effect for the warning ($Z = .57$, $p > .05$). Specifically, the proportion of first system trespassing incidents that were terminated on the warning target computers up to 5 seconds after a trespassing incident had started is almost identical to the proportion of incidents that were terminated in the same period on the no-warning computers (18 percent on the no-warning vs. 16.6 percent on the warning target computers). This finding corroborates evidence from our first experiment and indicates that a warning banner does not lead to immediate termination of a system trespassing incident.

Next, we compare the survival distribution of *first* trespassing incidents on warning and no-warning target computers. The results from this analysis are presented in figure 2a. As indicated in the figure, the proportion of first trespassing incidents that survived longer periods of time is smaller on the treatment (warning) than on the control (no-warning)

Figure 2. Time to System Trespassing Incident Termination—Experiment 2



(a) First trespassing incidents ($n = 502$)
(b) All trespassing incidents ($N = 3{,}768$)

target computers. To explore the impact of a warning banner on the hazard of trespassing incident termination, we generated a dummy variable indicating whether a system trespassing incident was recorded on a warning or no-warning target computer (1 = warning). We then used this measure to estimate its effect on incident survival time using a Cox proportional-hazard regression. The results from this analysis are presented in table 2,

Table 2. First System Trespassing Incidents Duration Regressed over Computer System Configurations (Experiment 2, $N = 502$)

Variables	Model 1		Model 2		Model 3	
	Coefficient (SE)	Hazard Ratio	Coefficient (SE)	Hazard Ratio	Coefficient (SE)	Hazard Ratio
Warning	.20* (.10)	1.22	.28† (.14)	1.32	.32* (.15)	1.38
Memory size	—	—	.26 (.14)	1.30	—	—
Bandwidth size	—	—	—	—	.14 (.15)	1.14
Warning × Memory	—	—	−.19 (.20)	.82	—	—
Warning × Bandwidth	—	—	—	—	−.25 (.20)	.78
Log likelihood	−1,928.62**		−1,926.93**		−1,927.88**	

ABBREVIATION SE = standard error.
†$p < .10$ (two-tailed); *$p < .05$ (two-tailed); **$p < .01$ (two-tailed).

model 1. As indicated in model 1, our findings confirm that a warning banner in the target computer is positively associated with system trespassing incident termination. Specifically, the hazard ratio estimate of the warning indicator suggests that a warning banner increases the rate of first trespassing incident termination by 22 percent. These findings reaffirm our second research hypothesis and demonstrate that a warning banner reduces the duration of system trespassing incidents on the target computer. Nevertheless, note that the size of the warning effect shrinks considerably in this model and only just reaches statistical significance ($p = .05$).

Turning to our final research hypothesis, we explore whether computing environments moderate the effect of a warning banner on the hazard of system trespassing incident termination. We first create two dummy measures indicating whether a trespassing incident was recorded on a low- or high-memory target computer (high memory = 1) and whether an incident was recorded on low- or high-bandwidth-connectivity computer (high bandwidth = 1). We then run two separate Cox models to test for the presence of significant interactive effects between our deterrence (i.e., warning) and computing environment (i.e., memory size and bandwidth capacity) measures on the hazard rate of system trespassing termination. The results from these analyses are reported in table 2, models 2 and 3. As indicated in model 2, the Cox model estimates reveal an insignificant interactive effect between the warning and memory size on the hazard of first system trespassing incidents. Similarly, model 3 reveals that the interaction between the warning banner and the bandwidth size is insignificant on the hazard rate of first trespassing incident termination. These findings stand in contrast to our theoretical expectations and suggest that the effect of a warning banner on the hazard of first system trespassing incident termination is not conditioned by the computing configurations of the target computer.

All Trespassing Incidents

In an effort to estimate the effect of a warning banner on the frequency of repeated trespassing incidents, as well as assess whether the patterns observed in our

Table 3. All System Trespassing Incidents Duration Regressed over Computer System Configurations (Experiment 2, $N = 3{,}768$)

Variables	Model 1		Model 2		Model 3	
	Coefficient (SE)	Hazard Ratio	Coefficient (SE)	Hazard Ratio	Coefficient (SE)	Hazard Ratio
Warning	.13† (.07)	1.14	.12 (.10)	1.13	.27** (.10)	1.30
Memory size	—	—	.11 (.10)	1.11	—	—
Bandwidth size	—	—	—	—	.20* (.10)	1.22
Warning × Memory	—	—	.01 (.14)	1.01	—	—
Warning × Bandwidth	—	—	—	—	−.29* (.14)	1.33
Theta	.23***		.23**		.23**	
Log Likelihood	−16,876.10**		−16,874.60**		−16,873.72**	

ABBREVIATION SE = standard error.
$^{\dagger}p < .10$ (two-tailed); $^{*}p < .05$ (two-tailed); $^{**}p < .01$ (two-tailed).

analyses of first trespassing incidents hold across the entire poll of trespassing incidents, we rerun our analyses using information from the full sample of trespassing incidents collected in the second experiment. Again, we start by determining any significant differences between the number of repeated trespassing incidents recorded on the warning and no-warning computers. The results from a *t* test for comparing the means of two groups reveal an insignificant difference between the averages of these groups ($t = -1.11, p > .05$). Specifically, the average number of trespassing incidents that was recorded on the warning and no-warning target computers is almost identical (7.79 on the warning target computers vs. 7.19 on the no-warning computers). Consistent with the findings reported in the first experiment, we find no evidence that the presence of a warning banner reduces the frequency of repeated system trespassing incidents on the target computer.

Next, we examine whether warning banners influence the duration of trespassing incidents by comparing the survival distributions of *all* trespassing incidents recorded on the warning and no-warning target computers throughout the experimental period. Figure 2b presents results from this comparison. At first glance, it seems as though the survival distributions of system trespassing incidents observed on the warning and no-warning target computers are indistinguishable. Nevertheless, across all time points, the proportion of trespassing incidents that survived longer periods of time is smaller on the treatment (warning) than on the control (no-warning) target computers.

To assess the magnitude of a warning banner on the hazard rate of trespassing incident termination, we next estimate a shared-frailty model that accounts for the frailty shared among all repeated trespassing incidents originating on the same target computer. The findings from this analysis are presented in table 3, model 1. As noted in model 1, the presence of a warning banner in the target computer is positively associated with trespassing incident termination. Specifically, the hazard ratio estimate of the warning indicator suggests that a warning banner increases the rate of system trespassing incident

termination by 14 percent. Nevertheless, the effect of warning is only marginally significant ($p < .10$) in this model.

Finally, we employ shared-frailty models to test whether different system configurations condition the effect of a warning banner on the hazard rate of system trespassing incident termination. The results from these analyses are presented in table 3, models 2 and 3.[9] In model 2, we test the hypothesis that RAM size conditions the effect of a warning banner on the hazard of system trespassing incident survival. As shown in table 3, the interaction between these two measures is insignificant in the model. This finding stands in contrast to our theoretical expectations and suggests that the effect of a warning banner on the hazard of system trespassing incident termination is not conditioned by the RAM size of the target computer.

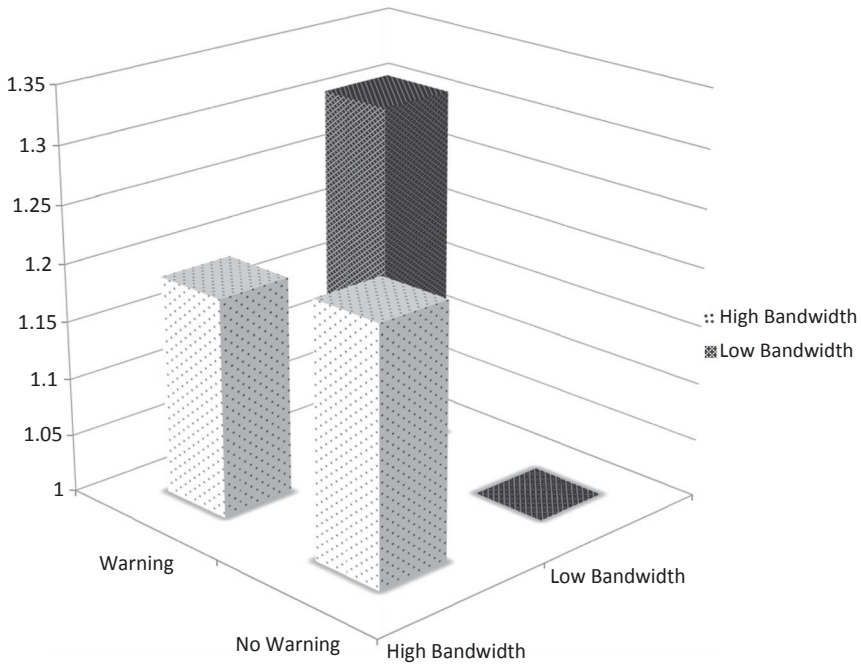
In model 3, we examine the interactive effect of a warning and bandwidth capacity on the hazard rate of system trespassing incident termination. In line with our research hypothesis, the results from this analysis reveal a significant interaction between these two measures ($b = –.29, p < .05$). Specifically, this finding suggests that the presence of a warning banner in the target computer leads to a slower decay of system trespassing incidents on high-bandwidth-capacity computers than on low-bandwidth-capacity computers. To illustrate this point, we calculate the hazard rate of system trespassing incident termination for each possible combination (target computers with no warning and a low bandwidth capacity serve as a reference group), and we plot our results in figure 3. As indicated in the figure, the hazard rate of trespassing incident termination on a *low*-bandwidth target computer with a warning banner is 31 percent. In contrast, the hazard rate of system trespassing incident termination on a *high*-bandwidth-capacity target computer with a warning banner is only 19 percent. These findings indicate that a warning banner produces more deterrence and shorter duration of system trespassing incidents on target computers with a low bandwidth capacity.

DISCUSSION

The heavy reliance of contemporary societies on computers and the Internet has facilitated fertile ground for the development of computer-focused crimes like system trespassing (Furnell, 2002). Unfortunately, although extensive research has explored technological aspects of system trespassing (Alata et al., 2006; Berthier and Cukier, 2009; Salles-Loustau et al., 2011), until now, only scant attention has been given in the criminological field to the effectiveness of existing cybersecurity measures in deterring unauthorized access to computer systems by system trespassers. Addressing these theoretical and empirical gaps, we drew on the restrictive deterrence literature (Gibbs, 1975; Jacobs, 2010) and raised hypotheses regarding the effectiveness of a warning message (Geerken and Gove, 1975) in preventing the progression and duration of system trespassing incidents. Four research hypotheses were examined.

9. We also constructed a binary measure indicating whether a trespassing incident was recorded on a low- or high-disk-space target computer (high disk space = 1), and we tested for the presence of significant interactive effects between warning and disk space size on the hazard of system trespassing incident termination. The first test used information from the first incident only, and the second used info from all trespassing incidents recorded on our systems. The results from these analyses indicated an insignificant interactive effect between these measures on the hazard rate of system trespassing incident termination.

Figure 3. Proportional Differences in Hazard Rate of Trespassing Session Termination Relative to No-Warning and Low-Bandwidth Target Computers



First, we hypothesized that a warning banner in the target computer could prevent, encourage, or have no effect on the progress of the first system trespassing incident (from a given trespasser) recorded on a target computer. Second, we suspected that the presence of a sanction threat in a target computer would reduce the frequency of repeated system trespassing incidents against the computer. Third, we predicted that the presence of a warning banner in the target computer would decrease the duration of both first and repeated system trespassing incidents. Finally, we hypothesized that large RAM size and high bandwidth capacity on the target computer attenuate the effect of a warning banner on the duration of system trespassing incidents. To test these hypotheses, we employed 1) a large set of target computers built for the sole purpose of being attacked and 2) two randomized experimental designs. The findings from these two independent experiments revealed several consistent and important insights.

First, a warning banner in the attacked computer system does not cause the immediate prevention of a first system trespassing incident. The findings from the two experiments indicate that in contrast to the classic deterrence model, the proportion of system trespassing incidents that were terminated in the first 5 seconds after a warning banner appeared (in the beginning of an incident) is similar to the proportion of incidents that ended after the same period of time when a warning banner was not displayed. We suspect that the inability of a sanction threat in the compromised system to prevent the progression

of a trespassing incident relates to the substantial time and efforts the system trespassers invested before obtaining access to the target computer. Specifically, to break into the target computers successfully, system trespassers need to scan the network, identify system vulnerabilities, discover open computer ports, and "guess" the correct combination of username and password to the target computer. This process is time consuming and demanding, and it might not always end in success. It could be the case that when "successfully" breaking into the system, system trespassers like to harvest the fruits of their success and engage in an exploratory first trespassing incident, independent of the presence of a warning banner in the system. Moreover, because hacking is considered by many attackers to be a mundane and easy routine (Spitzner, 2002; Wall, 2007), this finding lines up with the Wikström (2006) assumption that once an act of crime becomes a habit, deterring cues and messages have no effect on the decision of an offender to commit the act. System trespassers in this sense seem to follow the Pogarsky definition of incorrigible offenders, that is, "offenders who are impervious to dissuasion" (2002: 433) and to be insensitive to threats of legal sanctions.

Similarly, we find no evidence that a warning banner reduces the volume of repeated system trespassing incidents on the target computer. This finding seems to stand in contrast to the claim by Gibbs (1975) that offenders reduce the frequency of their criminal involvement in response to deterring cues. However, we suspect that we do not detect a restrictive deterrent effect on the frequency of repeated system trespassing incidents in our experiments because system trespassers are likely to keep access to multiple compromised computer systems at any given time and do not recall the unique configuration on each of these systems. Specifically, Spitzner (2002) claimed that system trespassers infiltrate as many computer systems as they can and then keep access to these computers for their future operations. Therefore, it could be the case that the warning banner has no effect on the frequency of repeated trespassing incidents because trespassers simply do not remember the presence of a sanction threat in the target computer prior to initiating a repeated trespassing incident.

In contrast, and in line with the third hypothesis, our findings reveal that the presence of a warning banner in the target computer significantly shortens the duration of first and repeated system trespassing incidents. Specifically, findings from our two independent experiments suggest that a warning banner in the target computers significantly increases the hazard of system trespassing incident termination (although this effect is less pronounced in the second experiment). These findings offer important support to the ideas of restrictive deterrence proposed by Gibbs (1975) and Jacobs (2010), and it is suggested that the presence of a sanction threat in the target computer results in system trespassers restricting the scope of their criminal activity. Specifically, we suspect that the warning messages presented to system trespassers during the trespassing incident 1) raised the concerns of trespassers regarding the possible protection measures taken by the victim in an effort to defend the system and 2) made them expose themselves on the target computer for shorter periods of time. Alternatively, it could be the case that a warning banner automatically activates cautious behavior from system trespassers and reduces their willingness to expose themselves for longer periods of time on the target computer (Bargh, Chen, and Burrows, 1996). In either case, although considerable attention had been given in the criminological literature to the notion of punishment avoidance (Stafford and Warr, 1993), this study is among few to focus on the actions taken by offenders in efforts to avoid punishment.

This finding joins previous empirical evidence that demonstrates offender responsiveness to sanction threats in the environment (Jacobs, 1996a, 1996b; Jacobs and Cherbonneau, 2012; Weaver and Carroll, 1985). Nevertheless, future studies should further explore the influence of sanction threats on other dimensions of restrictive deterrence. For instance, subsequent analyses should investigate the influence of punishment threats on the seriousness of system trespassing incidents and on the willingness of system trespassers to engage in risky online behaviors while using the target computer. Future work also should investigate how increasing the risk of being detected influences the actions of system trespassers on the target computer.

Fourth, we find mixed support for the assumption that the bandwidth capacity of a computer system conditions the effect of a warning banner on the duration of system trespassing incidents. Specifically, an analysis of the first trespassing incidents recorded on our target computers reveals no significant interactive effect between warning and bandwidth capacity on the hazard rate of system trespassing incident termination. In contrast, when analyzing the entire poll of system trespassing incidents (i.e., both first and repeated trespassing incidents), we find that the effect of warnings on the hazard rate of incident termination is conditioned by the bandwidth capacity of the target computer. The latter finding is consistent with our theoretical expectation: Low-bandwidth-capacity computers offer fewer opportunities for subsequent online operations and a greater probability of detection, and in turn, these computers encourage compliance by system trespassers with the deterring message and restriction of their criminal activity (Jacobs, 2010).

Importantly, system trespassers do not need to check the bandwidth capacity of the target computer to determine its functionality. Signs like delays in the appearance of commands on the intruder screen, as well as a low rate of data transfer between the target computer and that of the intruder (for instance, the download time of a 250-Kbytes file on the 128-Kbytes/s target computer is 15 seconds vs. 3 seconds on the 512-Kbytes/s computers), serve as important indicators of the bandwidth connectivity and functionality of the target system for the system trespasser. Thus, if the system trespasser cannot communicate with the target computer in an effective and rapid way, then he or she has no reason to remain on the system and expose himself or herself for long periods of time. This issue is particularly relevant for explaining the mixed findings regarding the interactive effects of warning and bandwidth capacity during first and subsequent system trespassing incidents. Specifically, we suspect that when encountering a warning during the first trespassing incident on the system, intruders explore and experiment with the system cautiously, while limiting their exposure on the system regardless of its bandwidth configuration. Using the intelligence gained during the first system trespassing incident, repeated trespassers would then have some understanding of the capabilities of the system, and the next time they trespass, they will limit their activities only to necessary operations, and then leave. Indeed, additional experiments are required to confirm this finding.

Finally, we find no support for the assumption that computer RAM size moderates the effect of a warning banner on the duration of system trespassing incidents. It could be that computer RAM size does not condition the effect of a warning because at the end of the day, the minimum RAM size required for the execution of commands and for running the tools of the intruder on the target computer is not very high, and therefore it does not affect the exposure of the trespasser on the target computer system. Alternatively, it could be that the target computer RAM size determines the specific uses trespassers find for the system. In that case, future research should investigate whether RAM size conditions the

effect of warnings on the probability of storing files on the target computer, setting up fake websites, or initiating subsequent attacks on other computer systems.

These findings support the view suggesting that criminological theories, particularly the deterrence perspective, should be implemented in the study of computer-focused crimes. Specifically, we suspect that the unique cyberspace realm allows investigations of theoretical constructs in a way that brings scholars closer to the antecedents of human behavior. For instance, in the current study, we did not actively recruit subjects to participate in our study, nor did we generate an unnatural environment for them to work in (Wright and Decker, 1994). As a consequence, the unique setting and tools that are common in the cyberspace realm brought us closer to the offenders and their reactions to a punishment threat and refined our understanding regarding the effect of a deterring message on the development of a criminal event. Moreover, it enabled empirical investigation of the effect of a sanction threat on the *occurrence* of criminal events and on the *progression and duration* of criminal incidents (Gibbs, 1975). Indeed, our findings suggest that once encountering a warning banner in the attacked system, system trespassers are willing to pursue the criminal act; yet they drop the connection with the target computer sooner than when they do not encounter such a banner. These findings may prove useful in contributing to the ongoing dispute regarding the application of deterrence strategies in cyberspace (Elliott, 2011; Geers, 2012). Future studies should further explore the "causal chain" of events that shape the decision making of offenders in cyberspace and shape their behavior in the presence of sanction threats.

In addition to its theoretical contributions, we believe that this study carries some policy implications for computer users and for IT managers who are in charge of protecting organizational networks and computer systems from data breaches and trespassing incidents. Indeed, the National Institute of Standards and Technology (NIST) recommends the display of a warning banner when all computer users (both legitimate and illegitimate) attempt to log in to the system (NIST, 2009). However, to date, no prior research has assessed the effectiveness of such warnings in influencing the behaviors of users and system trespassers. Our study is the first to show that displaying a warning banner in the attacked computer system does not prevent the occurrence of a trespassing incident and does not reduce the number of repeated trespassing incidents on the target computer, but it does reduce the duration of system trespassing incidents. These insights could be used by computer users and IT managers who debate whether implementing a warning banner in their own computer systems is beneficial for their needs. Future work should assess the effectiveness of other security measures to help IT managers tailor specific security solutions that are unique to their organizations.

Despite the important theoretical and practical implications that are derived from the findings presented in this work, it is essential to emphasize a few caveats about its design and results. First, we deployed our target computers on the Internet network infrastructure of a single educational institution. Although the two independent experiments indicate that our findings could be replicated and are valid over time, future research should further validate these findings by deploying target computers on a range of educational, industrial, and governmental networks. Second, it could be the case that the dosage of our treatment is simply not enough to obtain compliance with the warning by subjects (Piantadosi, 2005). Specifically, it could be that a more aggressive or more ambiguous warning (Sherman, 1990) would have produced different results for our immediate system trespassing cessation measure. Related to this, because we administered a similar warning message across the treatment groups, it is difficult for us to determine to which

of the warning banner components system trespassers are actually responding. Future research should address this issue and test the effectiveness of different warning forms on the progress and development of system trespassing events. Third, we have no way to tell whether system trespassers realized that they were not using actual systems but honeypots. Finally, we set up our target computers as computer systems with the Linux operating system. Although we have no reason to believe that system trespassers behave differently when intruding on computers with a Microsoft (Microsoft Corporation, Redmond, WA) or an Apple (Apple Inc., Cupertino, CA) environment installed, this point should be clear.

In sum, this work suggests that although a warning banner may be ineffective in preventing the occurrence of system trespassing incidents, it causes system trespassers to change their course of criminal action and stay on the target computer for shorter periods of time. This finding supports the restrictive deterrence perspective (Gibbs, 1975) and presents evidence that ties the administration of sanction threats to the engagement of offenders in detection avoidance strategies in response to such threats. We believe that these findings demonstrate the relevance of the deterrence perspective in the study of system trespassing events and provide clear ground for the development of an interdisciplinary explanation on the etiology of computer-focused crimes.

REFERENCES

Alata, Eric, Vincent Nicomette, Mohamed Kaâniche, Marc Dacier, and Matthieu Herrb. 2006. Lessons learned from the deployment of a high-interaction honeypot. *EDCC* 6:18–20.

Allen, Julia, and Ed Stoner. 2000. *Detecting Signs of Intrusions*. Pittsburgh, PA: Carnegie Mellon, Software Engineering Institute.

Anderson, James P. 1980. *Computer Security Threat Monitoring and Surveillance*. Technical Report. Fort Washington, MD: James P. Anderson Co.

Ariel, Barak. 2012. Deterrence and moral persuasion effects on corporate tax compliance: Findings from a randomized control trial. *Criminology* 50:27–69.

Bargh, John A., Mark Chen, and Lara Burrows. 1996. Automaticity of social behavior: Direct effects of traits construct and stereotype activation and action. *Journal of Personality and Social Psychology* 71:230–44.

Beauregard, Eric, and Martin Bouchard. 2010. Cleaning up your act: Forensic awareness as a detection avoidance strategy. *Journal of Criminal Justice* 38:1160–6.

Beccaria, Cesare. 1963 [1764]. *On Crimes and Punishments*. New York: Macmillan.

Becker, Gary S. 1968. Crime and punishment: An economic approach. *Journal of Political Economy* 76:169–217.

Bentham, Jeremy. 1970 [1785]. *An Introduction to the Principles of Morals and Legislation*. New York: Oxford University Press.

Berthier, Robin, and Michel Cukier. 2009. An evaluation of connection characteristics for separating network attacks. *International Journal of Security and Networks* 4:110–24.

Blais, Etienne, and Jean-Luc Bacher. 2007. Situational deterrence and claim padding: Results from a randomized field experiment. *Journal of Experimental Criminology* 3:337–52.

Blank, Stephen. 2001. Can information warfare be deterred? In *Information Age Anthology, Volume III: The Information Age Military*, eds. David S. Alberts and Daniel S. Papp. Washington, DC: Command and Control Research Program.

Bossler, A. M., and Thomas J. Holt. 2009. On-line activities, guardianship, and malware infection: An examination of routine activities theory. *International Journal of Cyber Criminology*, 3:400–20.

Box-Steffensmeier, Janet M., Suzanna De Boef, and Kyle A. Joyce. 2007. Event dependence and heterogeneity in duration models. *Political Analysis*, 15:237–56.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. Cambridge, U.K.: Cambridge University Press.

Brenner, Susan W. 2010. *Cybercrime; Criminal Threats from Cyberspace*. Westport, CT: Praeger.

Clarke, Ronald V. 1997. Introduction. In *Situational Crime Prevention: Successful Case Studies*, ed. Ronald V Clarke. Guilderland, NY: Harrow and Heston.

Coleman, Stephen. 2007. *The Minnesota Income Tax Compliance Experiment: Replication of the Social Norms Experiment*. http://ssrn.com.

Computer Fraud and Abuse Act of 1986, 18 U.S.C. § 1030 (1986).

Cusson, Maurice. 1993. Situational deterrence: Fear during the criminal event. In *Crime Prevention Studies*, Vol. 1, ed. Ronald V. Clarke. Monsey, NY: Criminal Justice Press.

Decker, John F. 1972. Curbside deterrence? An analysis of the effect of a slug-rejecter device, coin-view window and warning labels on slug usage in New York City parking meters. *Criminology* 10:127–42.

Eck, John E., and Julie Wartell. 1998. Improving the management of rental properties with drug problems: A randomized experiment. *Crime Prevention Studies* 9:161–85.

Ekblom, Paul. 1991. Talking to offenders: Practical lessons for local crime prevention. In *Urban Crime: Statistical Approaches and Analyses*, ed. Oriol Nello. Barcelona, Spain: Institut d'Estudis Metropolitans de Barcelona.

Elliott, David. 2011. Deterring strategic cyberattack. *IEEE Security and Privacy* 9:36–40.

Florêncio, Dinei, Cormac Herley, and Baris Coskun. 2007. Do strong web passwords accomplish anything? Paper presented at the 2nd USENIX Workshop on Hot Topics in Security, Boston, MA.

Furnell, Steven. 2002. *Cybercrime: Vandalizing the Information Society*. Boston, MA: Addison-Wesley.

Gadge, Jayant, and Anish Anand Patil. 2008. Port scan detection. Paper presented at the 16th IEEE International Conference on Networks, New Delhi, India.

Gallupe, Owen, Martin Bouchard, and Jonathan P. Caulkins. 2011. No change is good change? Restrictive deterrence in illegal drug markets. *Journal of Criminal Justice* 39:81–9.

Garfinkel, Simpson, Gene Spafford, and Alan Schwartz. 2003. *Practical UNIX and Internet Security*, 3rd ed. Sebastopol, CA: O'Reilly.

Geerken, Michael R., and Walter R. Gove. 1975. Deterrence: Some theoretical considerations. *Law & Society Review* 9:497–513.

Geers, Kenneth. 2012. The challenge of cyber attack deterrence. *Computer Law and Security Review* 26:298–303.

Gibbs, Jack. 1975. *Crime, Punishment, and Deterrence*. New York: Elsevier Scientific.

Goldstein, Noah J., Robert B. Cialdani, and Vladas Griskevicius. 2008. A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research* 35:472–82.

Goodman, Will. 2010. Cyber deterrence: Tougher in theory than in practice? *Strategic Studies Quarterly* (Fall):102–35.

Grabosky, P. N. 1996. Unintended consequences of crime prevention. In *The Politics and Practice of Situational Crime Prevention, Crime Prevention Studies*, Vol. 5, ed. Ross Homel. Monsey, NY: Criminal Justice Press.

Green, Gary S. 1985. General deterrence and television cable crime: A field experiment in social control. *Criminology* 23:629–45.

Guerette, Rob T., and Kate J. Bowers. 2009. Assessing the extent of crime displacement and diffusion of benefits: A review of situational crime prevention evaluations. *Criminology* 47:1331–68.

Harknett, Richard J. 1996. Information warfare and deterrence. *Parameters* (Autumn):93–107.

Jacobs, Bruce A. 1993. Undercover deception clues: A case of restrictive deterrence. *Criminology* 31:281–99.

Jacobs, Bruce A. 1996a. Crack dealers' apprehension avoidance techniques: A case of restrictive deterrence. *Justice Quarterly* 13:359–81.

Jacobs, Bruce A. 1996b. Crack dealers and restrictive deterrence: Identifying narcs. *Criminology* 34:409–431.

Jacobs, Bruce A. 2010. Deterrence and deterrability. *Criminology* 48:417–41.

Jacobs, Bruce A., and Michael Cherbonneau. 2012. Auto theft and restrictive deterrence. *Justice Quarterly*. E-pub ahead of print. doi: 10.1080/07418825.2012.660977.

Jacobs, Bruce A., and Jody Miller. 1998. Crack dealing, gender and arrest avoidance. *Social Problems* 45:550–69.

Keizer, Kees, Siegwart Lindenberg, and Linda Steg. 2008. The spreading of disorder. *Science* 322:1671–85.

Kerr, Orin S. 2009. *Computer Crime Law*, 2nd ed. St. Paul, MN: West.

Knudsen, Lars R., and Matthew J. B. Robshaw. 2011. Brute force attacks. *The Block Cipher Companion, Information Security and Cryptography* 3:95–108.

Liu, Lei, Robert A. Wolfe, and Xuelin Huang. 2004. Shared frailty models for recurrent events and terminal events. *Biometrics* 60:747–56.

Lowman, John. 1992. Street prostitution control: Some Canadian reflections on the Finsbury Park experience. *British Journal of Criminology* 32:1–17.

Mackey, David. 2003. *Web Security for Network and System Administrators*. Boston, MA: Cengage Learning.

McQuade, III, Samuel C. 2006. *Understanding and Managing Cybercrime*. Upper Saddle River, NJ: Pearson Education.

Namboodiri, Krishnan, and C. M. Suchindran. 1987. *Life Table Techniques and Their Applications*. New York: Academic Press.

NIST. 2009. *Recommended Security Controls for Federal Information Systems and Organization*. Washington, DC: U.S. Department of Commerce.

Paternoster, Raymond. 1987. The deterrent effect of the perceived certainty and severity of punishment: A review of the evidence and issues. *Justice Quarterly* 4:173–217.

Paternoster, Raymond. 1989. Absolute and restrictive deterrence in a panel of youth: Explaining the onset, persistence/desistence and frequency of delinquent offending. *Social Problems* 36:289–309.

Paternoster, Raymond, and Alex R. Piquero. 1995. Reconceptualizing deterrence: An empirical test of personal and vicarious experiences. *Journal of Research and Crime and Delinquency* 32:241–86.

Piantadosi, Steven. 2005. *Clinical Trials: A Methodologic Perspective*, 2nd ed. New York: Wiley.

Pogarsky, Greg. 2002. Identifying deterrable offenders: Implications for deterrence research. *Justice Quarterly* 19:431–52.

Ponemon Institute. 2011. *Second Annual Cost of Cyber Crime Study; Benchmark Study of U.S. Companies*. Research report. http://www.hpenterprisesecurity.com/collateral/report/2011_Cost_of_Cyber_Crime_Study_August.pdf.

Pratt, Travis C., Francis T. Cullen, Kristie R. Blevens, Leah E. Daigle, and Tamara D. Madensen. 2006. The empirical status of deterrence theory: A meta- analysis. In *Taking Stock: The Status of Criminological Theory*, eds. Francis T. Cullen, John Paul Wright, and Kristie R. Blevins. New Brunswick, NJ: Transaction.

Rama, Pirkko, and Risto Kulmala. 2000. Effects of variable message signs for slippery road conditions on driving speed and headways. *Transportation Research* 3:85–94.

Rantala, Ramona. 2008. *Bureau of Justice Statistics Special Report: Cybercrime Against Businesses, 2005*. Washington, DC: Bureau of Justice Statistics.

Salles-Loustau, Gabriel, Robin Berthier, Etienne Collange, Bertrand Sobesto, and Michel Cukier. 2011. Characterizing attackers and attacks: An empirical study. Paper presented at Proceedings of the 17th IEEE Pacific Rim International Symposium on Dependable Computing (PRDC 2011), Pasadena, CA.

Schultz, Wesley P., and Jennifer J. Tabanico. 2009. Criminal beware: A social norms perspective on posting public warning signs. *Criminology* 47:1201–22.

Schwartz, Richard, and Sonya Orleans. 1967. On legal sanctions. *University of Chicago Law Review* 34:282–300.

Sherman, Lawrence W. 1990. Police crackdowns: Initial and residual deterrence. *Crime and Justice* 12:1–48.

Sherman, Lawrence W., and David L. Weisburd. 1995. General deterrence effects of police patrol in crime hot spots: A randomized, controlled trial. *Justice Quarterly* 12:625–48.

Skinner, William F., and Anne M. Fream. 1997. A social learning theory analysis of computer crime among college students. *Journal of Research in Crime and Delinquency* 34:495–518.

Slemrod, Joel B., Marsha Blumenthal, and Charles W. Christian. 2001. Taxpayer response to an increased probability of audit: Evidence from a control experiment in Minnesota. *Journal of Public Economics* 79:455–83.

Spitzner, Lance. 2002. *Honeypots: Tracking Hackers*. Boston, MA: Addison-Wesley Longman.

Stafford, Mark C., and Mark Warr. 1993. A reconceptualization of general and specific deterrence. *Journal of Research in Crime and Delinquency* 30:123–35.

Stallings, William. 2005. *Wireless Communications and Networks*. Upper Saddle River, NJ: Pearson Prentice-Hall.

Tanenbaum, Andrew S. 2006. *Computer Networks*, 4th ed. Upper Saddle River, NJ: Pearson Prentice-Hall.

Tilley, Nick. 2005. *Handbook of Crime Prevention and Community Safety*. Devon, U.K.: Willan.

Tittle, Charles R. 1980. *Sanctions and Social Deviance*. Westport, CT: Praeger.

Wagner, David, and Paolo Soto. 2002. Mimicry attacks on host-based intrusion detection systems. In *Proceedings of the 9th ACM Conference on Computer and Communications Security* (pp. 255–64). New York: ACM.

Wall, David S. 2007. *Cybercrime: The Transformation of Crime in the Information Age*. Cambridge, U.K.: Polity.

Walters, Stephan J. 2009. *What Is a Cox Model?* London, U.K.: Hayward Medical Communication, Hayward Group.
Weaver, Frances M., and John S. Carroll. 1985. Crime perceptions in a natural setting by expert and novice shoplifters. *Social Psychology Quarterly* 48:349–59.
Wenzel, Michael, and Natalie Taylor. 2004. An experimental evaluation of tax-reporting schedules: A case of evidence-based tax administration. *Journal of Public Economics* 88:2785–99.
Whitman, Michael E. 2003. Enemy at the gate: Threats to information security. *Communication of the ACM* 46:91–5.
Wikström, Per-Olof H. 2006. Linking individual, setting, and acts of crime. Situational mechanisms and the explanation of crime. In *The Explanation of Crime: Contexts, Mechanisms, and Development*, eds. Per-Olof H. Wikström and Robert J. Sampson. Cambridge, U.K.: Cambridge University Press.
Wright, Richard, and Scott H. Decker. 1994. *Burglars on the Job*. Boston, MA: Northeastern University Press.
Yar, Majid. 2006. *Cybercrime and Society*. Thousand Oaks, CA: Sage.

David Maimon is an assistant professor of criminology and criminal justice at the University of Maryland—College Park. His research interests include cybercrime, experimental methods, and community and crime.

Mariel Alper is a criminology and criminal justice Ph.D. candidate at the University of Maryland. Her research interests include reactions to crime, reentry, and communities.

Bertrand Sobesto is a doctoral student in reliability engineering at the University of Maryland—College Park. He operates a honeypot network in collaboration with other universities and companies in the United States and abroad. His main research interests include cybercrime, malware, and network flow analysis.

Michel Cukier is the director for Advanced Cybersecurity Experience for Students (ACES) and the associate director for education for the Maryland Cybersecurity Center (MC2). He is an associate professor of reliability engineering with a joint appointment in the Department of Mechanical Engineering at the University of Maryland—College Park. His research covers dependability and security issues. His latest research focuses on the empirical quantification of cybersecurity.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Appendix A. Honeypots Deployed and Number of System Trespassing Incidents Recorded (Experiment 1)
Appendix B. Experiment 2