The Effect of a Surveillance Banner in an Attacked Computer System: Additional Evidence for the Relevance of Restrictive Deterrence in Cyberspace

Theodore Wilson1, David Maimon1, Bertrand Sobesto2, and Michel Cukier2

Abstract

Objectives: Test whether the presence of a surveillance message on an attacked computer system influences system trespassers’ active engagement with the compromised system (i.e., entering computer commands). The hypothesized restrictive deterrent effect is tested both in the context of a first system trespassing incident and in the progression of repeated trespassing incidents in an attacked computer system. Methods: We designed a randomized controlled trial and deployed a series of virtual target computers with known vulnerabilities into the computer network of a large public

1 Department of Criminology and Criminal Justice, University of Maryland, College Park, MD, USA
2 A. James Clark School of Engineering, University of Maryland, College Park, MD, USA

Corresponding Author:
Theodore Wilson, Department of Criminology and Criminal Justice, University of Maryland, 2220 LeFrak Hall, College Park, MD 20742, USA.
Email: twilson5@umd.edu
university in the United States. The target computers were set to either display or not display a surveillance banner once system trespassers infiltrated them. Results: We find that the presence of a surveillance banner in the attacked computer systems reduced the probability of commands being typed in the system during longer first system trespassing incidents. Further, we find that the probability of commands being typed during subsequent system trespassing incidents (on the same target computer) is conditioned by the presence of a surveillance banner and by whether commands have been entered during previous trespassing incidents. Conclusions: These findings offer modest support for the application of restrictive deterrence in the study of system trespassing.

Keywords
classical theories, criminological theory, prevention, crime, deviance

Introduction

Fueled by an expansion of new surveillance technology and tools, an increasing number of international police departments and governmental agencies have adopted a range of surveillance technologies in an effort to prevent crime (Gill and Spriggs 2005; Sarno, Hough, and Bulos 1999; Winge and Knutsson 2003). While some of these technologies are used in the physical world, such as closed-circuit television (CCTV) cameras, others are employed to monitor online user behaviors, for example, computer-aided supervisory monitoring software that enables managers in large organizations to monitor their employees (Aiello 1993). Nevertheless, the effectiveness of these tools in reducing crime is still unclear (Welsh and Farrington 2004, 2009a, 2009b). Moreover, although prior research crafts an important framework regarding the relationship between formal surveillance and crime in the physical world, the effectiveness of surveillance cues in reducing online deviance has been grossly understudied. Several of these studies of online deviance focus upon the victimization experiences of organizations (D’Arcy, Hovav, and Galleta 2009; Straub 1990) as opposed to the offending or victimization experiences of individuals.

Addressing these issues, the present study draws on an extension of the deterrence perspective (Gibbs 1975; Jacobs 2010) to assess the impact of a surveillance banner upon the behavior of system intruders during system trespassing incidents (i.e., the unauthorized use of a computer system; Furnell 2002). Specifically, our research has three key goals. First, we
explore whether a surveillance banner displayed to system trespassers upon entry to a computer system influences the seriousness of a trespassing event, by reducing the probability of computer commands being entered into the compromised system during the first system trespassing incident (Jacobs 2010). Second, we test the effectiveness of a surveillance banner in reducing the volume and probability of repeated system trespassing incidents on the target computer (Gibbs 1975). Finally, we investigate whether the effect of a surveillance banner on an intruder’s decision to actively engage with the system (i.e., enter commands in the system) decays during subsequent system trespassing incidents (Sherman 1990). To achieve these goals, we designed and implemented a randomized controlled trial using a large set of target computers built for the sole purpose of being infiltrated and collected trespassing data over a seven-month period.

Theoretical Background

System Trespassing

System trespassing is defined by McQuade (2006:83) as the act of “illegally gaining access to one or more computer systems after exploiting security vulnerabilities or defeating a security barrier.” Unfortunately, the true prevalence of computer trespassing incidents is still difficult to estimate due to the inability to adequately detect computer trespassing. Nevertheless, a recent report by the Online Trust Alliance (2014) indicates that in the United States alone, over 740 million records were exposed during 2013 as a result of numerous system trespassing incidents. Moreover, the average cost of a single data breach to large organizations in the United States was estimated to be around US$5.4 million in 2012 (Ponemon Institute 2013).

Since detecting system trespassing events is not an easy task, many security standards and practices have been developed over the years by Information Technology (IT) managers to detect these events. One important security practice adopted by the Council on Cyber Security (2013) draws on formal surveillance practices and calls for the monitoring of log files (i.e., the files to which a computer system writes a record of its activities) that are produced by networked computers (Paquet 2012). Based on this policy, organizations should retain the log files of all of their networked computers and allow system administrators to go through these files and identify anomalies in the logs. While this policy is designed to enable prompt detection of trespassing incidents on the networked computers,
surveillance cues may also deter system trespassers during the criminal act and influence their respective developments (Hinduja and Kooi 2013).

**Surveillance and Monitoring for Crime Prevention**

Formal surveillance practices are designed to prevent crime by increasing offenders’ perceived threat of detection and punishment for initiating a criminal event. According to Clarke and Homel (1997), formal surveillance is achieved by the deployment of individuals whose primary responsibility is security (e.g., security guards or IT managers), or through the introduction of some form of technology (e.g., CCTV) that increases the likelihood of detecting offenders and criminal events (Welsh, Mudge, and Farrington 2010). Overall, an extensive literature has investigated the effect of formal surveillance in preventing crime in public spaces (for recent reviews of these studies, see Ratcliffe, Taniguchi, and Taylor 2009; Welsh and Farrington 2004, 2008, 2009a, 2009b; Welsh, Farrington, and O’Dell 2010). However, findings from these studies reveal mixed evidence for the effectiveness of formal surveillance means in reducing crime. Welsh and Farrington (2009b), for instance, report that although CCTV technology is effective in reducing car thefts from parking lots, it is ineffective in reducing violent and property crimes in city centers, public housing communities, and public transportation facilities. Similarly, while several studies find that electronic article surveillance (i.e., hard tags placed on clothing) reduces the magnitude of shoplifting and theft from retail shops (Bamfield 2004; DiLonardo and Clarke 1996), other studies fail to observe such a relationship (Hayes and Blackwood 2006).

Unfortunately, this prior research is primarily relevant to crime occurring in the physical world, and it does not necessarily contribute to our understanding of surveillance in cyberspace. Moreover, these studies do not investigate the influence of security personnel and surveillance systems on the progression of a criminal incident. Addressing these issues, our study seeks to extend the scope of surveillance research by assessing the role of surveillance cues in the development of illegal computer system trespassing incidents. In line with previous surveillance research, we suspect that surveillance cues aimed at increasing offenders’ perceptions of the risks associated with detection of their criminal behavior may be effective in cyberspace and deter system trespassers’ behaviors during a system trespassing event. This rationale is also consistent with the deterrence perspective.
Restrictive Deterrence

Deterrence theory can be traced back to the utility-based principles presented in the works of Bentham ([1785] 1970) and Beccaria ([1764] 1963). Both Bentham and Beccaria conceived of individuals as rational actors who are susceptible to the influence of sanctions, as they weigh the potential costs and benefits of committing a potential criminal act, in what is termed a hedonistic calculus (Bentham [1785] 1970). Deterrence theory extends this tradition to focus more exclusively on the influence of the severity, celerity, and certainty of sanctions in prompting an individual to refrain from crime (Paternoster 1987). Emphasizing the need to communicate the sanction threat in an effective manner, Geerken and Gove (1975) suggest that a successful system of deterrence needs to inform offenders that (1) the probability that a criminal act will be detected by authorities is high, (2) after being detected, the probability of receiving sanctions is high, and (3) the cost associated with the sanction is great enough to offset the potential reward of the criminal act.

Contemporary branches of deterrence theory attempt to account for different theoretical aspects of deterrence, including the impact of punishment avoidance on an individual’s decision to initiate a criminal event (Stafford and Warr 1993) and the distinction between objective and subjective sanctions (Becker 1968). Gibbs’ (1975) differentiation between absolute and restrictive deterrence is an additional elaboration of the theory.

According to Gibbs (1975), absolute deterrence is a process whereby an offender is wholly deterred from engaging in any criminal conduct due to fear of potential sanctions. This abstention of criminal behavior is “absolute” in that potential offenders completely refrain from engaging in crime. In contrast, restrictive deterrence is a process whereby an offender is not wholly deterred from engaging in crime, but instead modifies his or her behavior in a specific manner to reduce the probability of detection and punishment (Gibbs 1975; Jacobs 2010). Gibbs’ initial conception of restrictive deterrence focused primarily upon a reduction in the frequency of crime, which necessarily limits the applicability of his constructs to those offenders who are already engaged in crime. Recognizing this limit, Jacobs (1993) later extended Gibbs’ platform to allow the analysis of an individual modifying the initial act of crime, including actions that reduce the seriousness of the criminal act, modification behaviors or measures that reduce the risk of detection, and those that alter the spatial-temporal component of the criminal event (Jacobs 2010).

Empirical investigations of different aspects of restrictive deterrence theory are still preliminary within the criminological field, with only a handful
of authors attempting to test the theory in real-world settings (Jacobs 1993, 1996; Jacobs and Cherbonneau 2012; Maimon et al. 2014). Jacobs (1996), for instance, observes that crack dealers frequently change location to avoid detection by both formal and informal agents within a neighborhood. Similarly, Beauregard and Bouchard (2010) identify concealment measures and the use of gloves by rapists, techniques that reduce both the evidence left at the crime scene and the eventual likelihood of apprehension. These studies, while not clearly tying surveillance cues in the environment to offenders’ behavioral changes during a crime, do demonstrate offenders modifying their offending behavior to reduce the likelihood that they will be detected and/or arrested. Nevertheless, while these studies describe the progression of criminal events in the physical world, their ability to facilitate understanding of the relevance of deterrence in the progression of criminal cyber events is limited.

**Deterrence in Cyberspace**

Overall, previous empirical work on deterrence in cyberspace is relatively scarce in the criminological literature (D’Arcy et al. 2009; Maimon et al. 2014; Straub 1990). Nevertheless, extensive theoretical literature has debated the effectiveness of deterrence strategies (including threats of retaliation) in preventing and mitigating the consequences of cybercrimes (Goodman 2010; Harknett 1996). Some of these works suggest that the main problem associated with a cyber deterrence strategy is the limited capability and intent demonstrated by the legal system to punish cyber offenders, especially small-scale offenders. Furthermore, Harknett (1996) proclaims that the inherently anonymous nature of cyberspace drastically increases individuals’ ability to avoid detection and escape penalties for their criminal online behaviors, and in turn reduces the effectiveness of deterrence strategies in cyber space (also see Denning and Baugh 2000).

In contrast, other scholars argue that it is not necessary to identify specific individuals for deterrence to take effect (Goodman 2010; Maimon et al. 2014). This claim is supported by Png and Wang’s (2009) findings which indicate that system trespassers attempt to maximize the expected net benefit from their cyber operations by directing efforts against targets that provide the maximum return and minimum risks. Further, Maimon and colleagues (2014) have recently found that a warning banner displayed to system trespassers upon entry to an attacked computer system reduced the duration of both first and repeated system trespassing incidents. However, while Maimon and colleagues focus on the effect of a sanction threat in their
work, they do not test the effect of surveillance and monitoring cues in an attacked computer system on the progression of system trespassing events.

The Current Study

We seek to directly address the theoretical debate regarding the effectiveness of deterrence strategies in cyberspace, by assessing the influence of a surveillance banner in a compromised computer system upon the seriousness (Jacobs 2010) of intruders’ engagement with the system. Indeed, all system trespassing involves unauthorized access to a computer system. However, once they’ve infiltrated the system, trespassers can choose whether to actively engage with the compromised system. Regardless of the intruder’s intent, any action performed on the system beyond the initial unauthorized access is a form of active manipulation of the system and could be considered an escalation of the trespassing event. One way of initiating an active engagement with the target system requires trespassers to enter commands directly into the console of the compromised system. By entering commands, trespassers may gather intelligence on the system and its legitimate users, manipulate content, and/or install a backdoor that will allow them easy access to the target computer in the future (McQuade 2006). However, such engagement with the system increases the probability of detection by legitimate users of the system: Designated log files record all of the commands typed on the system, and there is always a danger that someone monitors the log files in an effort to identify intruders on the system. Thus, we presume that system trespassers who attempt to avoid detection will abstain from escalating their illegal online behaviors and avoid entering commands on the system.

Integrating the rationale afforded by Geerken and Gove (1975) with past research demonstrating that employees’ awareness of surveillance tools on their networked computers reduces their intention to misuse their computers (D’Arey et al. 2009), we suspect that displaying a targeted and specific surveillance message in an attacked computer system will raise trespassers’ awareness of the presence of surveillance on the system and increase the trespassers’ perceived risk of detection. Since one means by which system trespassers can reduce the potential for detection is through reducing the seriousness of the trespassing event by abstaining from entering commands in the system, we suggest that a surveillance banner displayed to trespassers upon entry to a computer system reduces the probability of computer commands being entered into the compromised system during a first system trespassing incident.
In addition to the possible influence of a surveillance banner on the seriousness of a trespassing incident, we test its effectiveness in reducing the frequency and probability (Gibbs 1975) of repeated trespassing incidents on the target computer. Indeed, a surveillance banner on a target system would only be displayed to a system trespasser upon successful entry to the computer system. Nevertheless, the surveillance message might influence trespassers’ willingness to initiate repeated trespassing events. Specifically, Gibbs (1975) suggests that since repeat offenders believe that their criminal behavior will eventually be detected and punished, they reduce the frequency of their offending in an effort to delay this point of detection and frustrate detection efforts. Thus, consistent with Gibbs’ (1975) assertion, our second research hypothesis proposes that the presence of a surveillance banner reduces the probability and frequency of repeated system trespassing incidents on the compromised system.

Finally, our third research goal explores whether the effect of a surveillance banner on the probability of commands being entered on the system decays beyond the first system trespassing incident to subsequent incidents (Sherman 1990). Indeed, Sherman proclaims that the unwillingness of potential offenders to initiate a criminal event once encountering a deterring cue in the environment for the first time, such as police crackdowns, might decay with time once they learn through trial and error that they had overestimated the certainty of being caught. Drawing on Sherman’s (1990) theoretical claims, as well as emerging empirical evidence suggesting that individuals update their risk of punishment based on past experiences of punishment and punishment avoidance (Anwar and Loughran 2011), we believe that the effect of the surveillance banner on system trespassers’ engagement with the system during subsequent trespassing events may vary based on whether commands had been entered during the first trespassing incident. Specifically, we expect that for those system trespassers who were deterred by the surveillance banner and did not enter commands in the system during the first system trespassing incident, the restrictive deterrent effect of the surveillance banner should extend to the second system trespassing incident. Adopting the rationale afforded by Sherman (1990) and Anwar and Loughran’s (2011), we suspect that this group of offenders should still be capable of being deterred during the second system trespassing incident, as these trespassers’ perception of the certainty of punishment should theoretically remain unchanged since they did not receive additional information with which to update their perceptions of the risk of detection and punishment. Thus, we propose that a surveillance banner reduces the probability of computer commands being entered into the compromised
system during a second system trespassing event if the intruder did not engage with the system previously.

In contrast, we believe that surveillance banners that did not deter trespassers from engaging with the system during the first trespassing event will not deter trespassers’ engagement with the system in subsequent trespassing events. Specifically, we suspect that intruders who were not deterred from actively engaging with the system during the first trespassing event will likewise not be deterred during their second system trespassing incident. If there were no sanctions levied as a result of their active engagement with the system during the first system trespassing incident, this should result in a lower updated perception of the certainty of detection (Loughran et al. 2012). Therefore, we hypothesize that a surveillance banner will have a null effect on the probability of commands being entered on the system during the second system trespassing incident if commands were already typed into the system during the first system trespassing event.

Data and Methods

Experimental Design

To test our research hypotheses, we designed and implemented a randomized controlled trial that allows the collection of data on system trespassing events while preserving the unique context in which they exist. In line with Maimon and colleagues’ (2014) research design, we employed 300 public Internet protocol (IP) addresses of a large public American university and deployed experimentally controlled target computers onto the university’s network.1 The target computers were virtually deployed and maintained from a set of physical servers housed within the university’s server farm; they were designed to emulate real computer systems with vulnerable entry points and Linux-based operating systems (CentOS).2 We did not advertise the existence of the target computers or their assigned IP addresses. Instead, we deployed our target computers over a seven-month period (from April 4, 2013, until November 3, 2013) and waited for trespassers to find the computers and attempt to compromise them.

In order to successfully infiltrate our target computers, intruders had to “guess” the credentials of “legitimate users” of the system—typically with toolkits3 designed expressly for the purpose of cracking the legitimate login credentials to user accounts (McQuade 2006). To emulate a genuine computing environment, the target computers were set to reject the login attempts by system trespassers until a predefined threshold of attempts
(a generated random number between 100 and 200) was reached. When this threshold was met, the login credentials used during the \( n \)th attempt were treated as the legitimate credentials for the system. System trespassers then had to input these credentials into the target computer in order to allow further access to the attacked system.

**Procedure**

As soon as the intruders reached the predetermined \( n \)th attempt and gained access to the system, they were randomly assigned to one of four conditions. In the first condition (treatment 1), the target computers were configured to display the following surveillance banner upon each entry to the system: “This system is under continuous surveillance. All user activity is being monitored and recorded.” Figure 1 presents a screen shot of the manner in which the surveillance banner was presented to system trespassers. In the second condition (treatment 2), no banner was displayed, but we installed two surveillance-based processes: an open-source surveillance software (Zabbix) that emulated the process of monitoring the target computers and a customized agent called “monitor” that checked system configurations. The surveillance software did not impact the system performance and was only visible to trespassers upon calling forth a display depicting all of the processes running on the target computer. These processes would appear in the list of all of the processes running on the target computer as “zabbix_agentd” and “monitor,” respectively. The Linux command to call forth this information is “ps.”

In the third condition (treatment 3), target computers were configured to present the surveillance banner (upon each entry to the system) and run the surveillance software described earlier. In the last condition (control), the target computer had neither the surveillance banner nor the surveillance software installed. We allowed trespassers to work with their assigned computer for a period of 30 days, which allowed for the initiation of repeated trespassing events against our target computers. At the end of the 30-day period, we blocked trespassers’ access to the target computer, cleaned it, and redeployed it.

Over the seven-month experimental period, 660 target computers were successfully compromised and retained at least one system trespassing event. As depicted in Table 1, the number of computers in each condition ranged from a minimum of 155 (treatment 2) to a maximum of 172 (control). These target computers experienced a total of 2,942 trespassing incidents during the experimental period with computer commands entered on the attacked system in 1,318 of these incidents.
In order to assess whether trespassers had indeed been exposed to the second treatment (i.e., surveillance software), we investigated the frequency with which the “ps” command was entered in our target computers. These

Figure 1. Screenshot of surveillance banner.

Table 1. Descriptive Statistics by Experiment Condition.

<table>
<thead>
<tr>
<th>Target Computer Type</th>
<th>Target Computers with System Trespassing Incidents</th>
<th>System Trespassing Incidents</th>
<th>Trespassing Incidents with Commands</th>
<th>Average Trespassing Incidents per Target Computer</th>
<th>Average Trespassing Incidents with Commands per Target Computer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>172</td>
<td>805</td>
<td>385</td>
<td>4.68 (5.66)</td>
<td>2.24 (3.23)</td>
</tr>
<tr>
<td>Surveillance banner</td>
<td>155</td>
<td>641</td>
<td>255</td>
<td>4.13 (3.76)</td>
<td>1.65 (1.65)</td>
</tr>
<tr>
<td>Surveillance process</td>
<td>164</td>
<td>686</td>
<td>311</td>
<td>4.18 (3.50)</td>
<td>1.90 (2.00)</td>
</tr>
<tr>
<td>Surveillance banner and process</td>
<td>169</td>
<td>810</td>
<td>367</td>
<td>4.79 (4.50)</td>
<td>2.17 (2.41)</td>
</tr>
<tr>
<td>Total</td>
<td>660</td>
<td>2,942</td>
<td>1,318</td>
<td>4.46 (4.47)</td>
<td>2.00 (2.42)</td>
</tr>
</tbody>
</table>

Note: Standard deviations are presented in parentheses where appropriate.

In order to assess whether trespassers had indeed been exposed to the second treatment (i.e., surveillance software), we investigated the frequency with which the “ps” command was entered in our target computers. These
investigations yielded that the “ps” command was entered on less than 33 percent of the target computers, and approximately 25 percent of those computers assigned to receive the surveillance software condition. Further, displaying the running processes on the system does not guarantee that a trespasser was able to notice the entries for “zabbix_agentd” and “monitor” or to identify them as surveillance software. Because this uncertainty would make the surveillance software have a diminished or even negligible effect upon trespassers’ subsequent behavioral outcomes, we consolidated the four conditions from the $2 \times 2$ factor design. Concordantly, the control group and the software-only group (treatment 2) were consolidated into a single no banner group, while the surveillance banner only and the surveillance banner and software groups (treatments 1 and 3) were consolidated into a banner group. This consolidation led to the no banner group retaining 336 target computers that produced a total of 1,491 system trespassing incidents, while the banner group retained 324 target computers that produced a total of 1,451 system trespassing incidents. Notably, this consolidation did not detrimentally invoke bias according to any of the measures reported here. Nevertheless, sensitivity analyses were conducted to ensure that this methodological decision did not bias the ultimate results.6

**Outcome Measures**

Our study’s main unit of analysis is the target computer; the constructed variables and subsequent analyses were all conducted at the level of the target computers. In order to test our three research hypotheses, we constructed a list of dependent measures. The first measure taps the presence of any commands having been entered in the target computer during the first system trespassing incident. This measure is coded as a binary variable, with (1) indicating the presence of at least one entered computer command on the first system trespassing incident and (0) indicating no such commands entered during the first system trespassing incident. A second binary measure is employed for testing our second research hypothesis and distinguishes between computers with more than one recorded trespassing event (1) and target computers with only one recorded trespassing incident (0). Finally, our third binary measure taps the presence of any commands entered in the target computer during the second system trespassing incident: (1) indicates the presence of at least one computer command entered into the system during the second system trespassing incident and (0) indicates the absence of any such commands entered into the system during the second system trespassing incident.
**Results**

We begin by investigating the effect of the surveillance banner on the seriousness of a trespassing event and examining the probability of commands being typed in the target computers during the first system trespassing event. Results from our analyses are presented in Panel A of Table 2. Since some of the trespassing incidents lasted for a duration of 0 seconds, we first present our analyses for the total pool of first trespassing incidents recorded on our target computers and then restrict subsequent analyses to first trespassing incidents that retained a duration greater than 5, 25, and 50 seconds, respectively. Beginning with the left column of Panel A, one may observe that 36 percent of the first system trespassing incidents recorded on the surveillance banner computers included computer commands being entered into the system. In contrast, 39 percent of the first system trespassing incidents recorded on the no surveillance banner computers included computer commands being entered into the system. Although the volume of system trespassing incidents involving computer commands is lower on the surveillance banner computers than on the no surveillance banner computers, consistent with our first research hypothesis, this difference is not statistically significant. However, since 79 of the first trespassing events recorded on our target computers lasted less than 5 seconds, and were thus too short for trespassers to either read the banner or input commands into the system, we suspect that these estimates may be biased downward. Therefore, we reran our analysis while removing these potentially downward biasing observations; results from these analyses are depicted in columns 3, 4, and 5 of Panel A.

Upon limiting the sample to those trespassing events that lasted longer than 5 seconds, the estimated difference between the proportion of first system trespassing incidents with commands recorded on the surveillance banner computers and the proportion of first system trespassing incidents with commands recorded on the no surveillance banner computers increases to 5.4 percent ($p = .09$). This difference increases further to 7 percent ($p = .08$) and 8 percent ($p = .05$) when assessing those first system trespassing incidents that lasted longer than 25 seconds and 50 seconds, respectively. The unique pattern revealed by these findings would seem to indicate that system trespassers who study the banner for longer periods of time opt to exit the system without entering commands. However, this finding may also indicate that system trespassers enter commands into the system more rapidly in the presence of a surveillance banner. To investigate which of these interpretations is empirically supported, we reran our analysis for
Table 2. Commands Entered during the First and Second Trespassing Events.

Panel A: Probability of any commands being typed during the first trespassing event, by event duration

<table>
<thead>
<tr>
<th>Target Computer Type</th>
<th>All First Trespassing Incidents</th>
<th>First Trespassing Incidents Longer Than 5 Seconds</th>
<th>First Trespassing Incidents Longer Than 25 Seconds</th>
<th>First Trespassing Incidents Longer Than 50 Seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>No surveillance banner</td>
<td>0.39 (0.03)</td>
<td>0.44 (0.03)</td>
<td>0.46 (0.03)</td>
<td>0.45 (0.03)</td>
</tr>
<tr>
<td>Surveillance banner</td>
<td>0.36 (0.03)</td>
<td>0.39 (0.03)</td>
<td>0.39 (0.03)</td>
<td>0.37 (0.03)</td>
</tr>
<tr>
<td>Difference</td>
<td>0.03 (0.04)</td>
<td>0.05 (0.04)</td>
<td>0.07 (0.05)</td>
<td>0.08 (0.05)</td>
</tr>
<tr>
<td>z Statistic</td>
<td>0.60</td>
<td>1.31</td>
<td>1.37</td>
<td>1.61</td>
</tr>
<tr>
<td>p Value</td>
<td>0.27</td>
<td>0.09</td>
<td>0.08</td>
<td>0.05</td>
</tr>
<tr>
<td>n</td>
<td>660</td>
<td>581</td>
<td>481</td>
<td>423</td>
</tr>
</tbody>
</table>

Panel B: Frequency and probability of repeated trespassing incidents

<table>
<thead>
<tr>
<th>Target Computer Type</th>
<th>Average System Trespassing Incidents per Target Computer</th>
<th>Probability of Repeated Trespassing Incidents</th>
</tr>
</thead>
<tbody>
<tr>
<td>No surveillance banner</td>
<td>4.44 (4.73)</td>
<td>0.24 (0.02)</td>
</tr>
<tr>
<td>Surveillance banner</td>
<td>4.48 (4.17)</td>
<td>0.22 (0.02)</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.04 (0.35)</td>
<td>0.02 (0.03)</td>
</tr>
<tr>
<td>Test statistic</td>
<td>-0.12</td>
<td>0.57</td>
</tr>
<tr>
<td>p Value</td>
<td>0.55</td>
<td>0.31</td>
</tr>
</tbody>
</table>
Table 2. (continued)

Panel C: Probability of any commands being typed during the second trespassing event, by presence of commands typed during the first trespassing incident

<table>
<thead>
<tr>
<th>Target Computer Type</th>
<th>Unconditioned Probability</th>
<th>No Prior Commands Entered During the First Trespassing Incident</th>
<th>At Least One Prior Command Entered during the First Trespassing Incident</th>
</tr>
</thead>
<tbody>
<tr>
<td>No surveillance banner</td>
<td>0.52 (0.03)</td>
<td>0.47 (0.04)</td>
<td>0.63 (0.05)</td>
</tr>
<tr>
<td>Surveillance banner</td>
<td>0.48 (0.03)</td>
<td>0.38 (0.04)</td>
<td>0.67 (0.05)</td>
</tr>
<tr>
<td>Difference</td>
<td>0.04 (0.04)</td>
<td>0.09 (0.05)</td>
<td>−0.04 (0.07)</td>
</tr>
<tr>
<td>z Statistic</td>
<td>1.02</td>
<td>1.62</td>
<td>−0.48</td>
</tr>
<tr>
<td>p Value</td>
<td>0.15</td>
<td>0.05</td>
<td>0.68</td>
</tr>
<tr>
<td>n</td>
<td>507</td>
<td>339</td>
<td>168</td>
</tr>
</tbody>
</table>

Note: Standard deviations are presented in parentheses where appropriate.
trespassing events that were recorded on surveillance and no surveillance banner computers and that only lasted between 0 and 25 seconds. Results from this analysis indicate that the proportion of first system trespassing incidents that lasted between 0 and 25 seconds that included computer commands is .33 for both surveillance banner and no surveillance banner computers. This finding refutes the assumption that system trespassers simply enter commands at a faster rate in the presence of a surveillance banner, and instead, supports the idea that system trespassers who take the time to reflect upon the banner refrain from actively engaging with the system.

Next, we tested the effect of the surveillance banner in reducing the frequency and probability of repeated trespassing events on the target computer. We first counted the number of repeated trespassing events recorded on banner and no banner computers across the 30-day period of access. Results from this analysis are presented in Panel B of Table 2. Note that the average number of repeated trespassing events on the banner and no banner computers are similar at 4.48 and 4.44, respectively. Results from a t-test comparing the difference between these two means reveal that this difference is statistically insignificant ($t = -.12, p = .55$). A comparison of the proportion of computers retaining repeated trespassing incidents between banner and no banner computers yields similar values at .22 and .24, respectively, as indicated in the last column of Panel B. This small difference of .02 is not found to be statistically significant ($p = .31$) and in conjunction with the t-test results reported earlier, does not support our second research hypothesis.

Moving to our third research question, we investigated whether the marginal effect of the surveillance banner on the probability of commands being entered on the system during the first system trespassing event persists beyond the first trespassing event to the second and subsequent incidents. The analysis pertaining to the second incident (as well as subsequent trespassing incidents) is not as straightforward as the analyses discussed for the first research hypothesis, as there is a potential history effect wherein trespassers have some experience working with our target computers during the first trespassing event. Specifically, when analyzing data pertaining to the second trespassing events, trespassers could differ according to two important dimensions: (1) exposure to a surveillance banner and (2) having entered commands during the first trespassing event. Trespassers’ reactions to the surveillance banner during the first trespassing event and their willingness to enter commands in the system can be reasoned to provoke biases in assessing the effect of the surveillance banner on the probability of entering any command on the second
trespassing event. As such, our analysis assesses the effect of the surveillance banner on the probability of commands being typed on the system during the second trespassing event for (1) the whole pool of target computers experiencing at least two trespassing events, (2) the subsample of target computers on which no commands were entered during the first trespassing incident, and (3) the subsample of target computers on which commands were entered during the first trespassing event. These results are depicted in Panel C of Table 2.

Starting with the left column of panel C, we first examine the unconditioned effect of the presence of a surveillance banner on the probability of commands being entered in the system during the second trespassing event. As depicted in the table, the unconditioned difference in the probability of any commands being typed on banner and no banner target computers during the second trespassing event is 4.5 percent ($p = .15$). However, applying the third research hypothesis to the subsample of target computers that experienced at least two trespassing events and on which no commands were entered during the first trespassing incident reveals support for our theoretical rationale. Specifically, we find that commands were entered into 47 percent of the no surveillance banner computers during the second trespassing incidents when no engagement was made with the system during the first trespassing event. By contrast, commands were entered into only 38 percent of the surveillance banner computers during the second trespassing incidents if no commands were entered into the system during the first trespassing event. The difference between these proportions is statistically significant at $p = .05$ and suggests that a surveillance banner reduces the probability of commands being entered in the target computer during the second trespassing event if no commands were previously entered during the first event.

In analyses not shown here, we tested whether the effect of a surveillance banner on the compromised system persisted beyond the second trespassing incident to subsequent incidents. These analyses proceeded in a similar fashion to those of the second system trespassing incident, by running multiple analyses conditioned on whether any commands had been entered at any point during any of the previous system trespassing incidents. Each of these subsequent assessments produced null results and suggested that the effect of a surveillance banner decays beyond the second trespassing incident for this sample.

Investigation of the effect of a surveillance banner on target computers recording at least two trespassing events, and on which commands were entered during the first trespassing event, further supports our expectations.
As may be observed in the last column of Panel C, the difference between the proportion of surveillance and no surveillance target computers with at least two trespassing events, and on which commands were entered during the first trespassing event, is not statistically significant and runs counter to theoretical expectations according to deterrence theory. This result reaffirms our theoretical claim and indicates that such a surveillance banner does not have an effect on the probability of commands being entered on the system during the second system trespassing incident if commands were already typed into the system during the first system trespassing event.

Discussion

System trespassing is a growing public concern and is estimated to cost governments, businesses, and individuals millions of dollars annually (Ponemon Institute 2013). Drawing on Gibbs’ restrictive deterrence perspective (Gibbs 1975; Jacobs 2010), we designed a randomized controlled trial in an effort to investigate the role of a surveillance banner in an attacked computer system in determining system trespassers’ engagement with the system. First, we asked whether a surveillance banner displayed to system trespassers upon infiltrating a computer system would curtail the probability of computer commands being entered into the attacked system during a first system trespassing event, and thus result in a less serious trespassing event. Second, we investigated whether the presence of a surveillance banner in an attacked computer system would decrease the frequency and probability of repeated system trespassing incidents on that particular computer. Finally, we explored whether the effect of a surveillance banner on the probability of commands being entered in the first trespassing incident persists in subsequent incidents. Findings from our unique field experiment afford several insights.

First, we find mixed support for the hypothesis suggesting that the presence of a surveillance banner in an attacked computer system reduces the probability of commands being entered in the system during a first system trespassing event. Specifically, the deterrent effect was only observed for those system trespassing events that lasted longer than 50 seconds and not for shorter system trespassing events. Thus, this finding indicates that surveillance cue is somewhat effective in eliciting deterrence and raising detection concerns among system trespassers who take an adequate amount of time toward considering both the content of the banner as well as the prospects for actively engaging with the target computer. This conditioned effect, in addition to substantive reasons justifying this conditional analysis,
lends modest support to Jacobs’ (1993) claims regarding an offender’s reduction of the seriousness of the criminal act in response to a deterring cue and suggests that system trespassers are less likely to escalate their offending in the presence of a surveillance cue when they take an adequate amount of time to consider the content of the cue. The observed effect of a surveillance banner for trespassers’ engagement with the system during first trespassing incidents is promising from a policy point of view, as delaying the entering of commands to later trespassing events allows more time for IT practitioners to respond to a trespassing incident and ameliorate the compromise on the system.

Second, we find no effect of a surveillance banner on the frequency and probability of repeated trespassing events against our target computers. Although inconsistent with Gibbs’ (1975) and Jacobs’ (1993) assertions regarding the effect of restrictive deterrence on the frequency of repeated offending in the physical world, our finding joins prior empirical evidence revealing that a warning banner in an attacked computer system is ineffective in reducing the volume of repeated trespassing events against a target system (Maimon et al. 2014). We suspect that the reason for the statistically insignificant effect of a surveillance banner on the volume of repeated trespassing incidents is embedded in system trespassers’ tendency to keep access to multiple compromised computer systems at any given time. Specifically, according to Spitzner (2002), system trespassers infiltrate as many computer systems as they can and keep access to these computers for their future operations. Therefore, it is possible that the surveillance banner did not influence the volume of repeated trespassing incidents on our systems because trespassers did not recall the potential presence of surveillance in the target computer before initiating a repeated trespassing incident (Maimon et al. 2014).

Finally, we find that those system trespassers who did not enter any commands during their first trespassing incident on a given target system are subsequently deterred from escalating their offending during the second trespassing event. In contrast, those intruders who previously entered commands during the first trespassing incident are not deterred from entering subsequent commands into the system during the second trespassing event. This finding is consistent with parallel investigations into the potential for deterrence decay within hot-spot policing (Sherman 1990). However, while prior research focuses on the decaying effect of deterrence practices on the overall level of crime in a community over time (Ross 1984), our study is focused on the progression and sequencing of system trespassing incidents against target computers. Moreover, this finding lends credence to our
theoretical expectations that those trespassers who entered commands during the first trespassing event and subsequently receive no sanctions for their illegal activities may update their perception of the certainty of detection to a lower level (Anwar and Loughran 2011) and are thus more likely to engage in crime (Loughran et al. 2012). Although we do not directly observe this updating process in our work, our observations match a priori expectations afforded by these principles. Specifically, our observations yielded that individuals who entered system commands during the first trespassing event were also far more likely to engage the system with additional commands than those trespassers who did not enter commands during the first session, regardless of the assigned treatment condition. This heterogeneity within the population of trespassers offers a rich area of inquiry for further analysis of deterrence and cybercrime.

These findings retain implications for future theoretical development within the deterrence tradition and the continued application of deterrence principles to cyberspace. Specifically, a few scholars believe that it is impossible to deter system trespassers from their online behaviors due to their ability to disguise their identity and location in cyberspace (Denning and Baugh 2000). In contrast, other scholars argue that system trespassers’ actions on attacked computer systems consistently demonstrate their rational behavior and that it is not necessary to identify individuals in cyberspace in order for deterrence to take effect (Goodman 2010). Our findings support the latter position. Coupled with Maimon and colleagues’ (2014) findings, we suspect that our work clears the ground for the beginning of an empirical body of research that supports the continued assessment and application of deterrence concepts in cyberspace. While most prior work in the context of deterrence attempted to predict the influence of deterring cues on the occurrence of a criminal event, the unique context of cyberspace allows investigations of deterrence concepts and their impact on the progression and development of sequences of criminal events. Thus, we call for further criminological research around the online constructs that could alter the pathways of online criminal behavior.

Our results also contribute to the ongoing discussion in support of recognizing the human component inherent to cybercrime, as an additional modicum with which to prevent and modify cybercriminal activity. Specifically, we believe that our findings support the implementation of similar surveillance banners on organizations’ networked computers. This straightforward policy could be applied with minimal cost and effort on legitimate users’ computers. We do not find any potential negative repercussions to altering an entry banner to a computer system in this manner,
as the system conditions are not physically modified to otherwise prevent or modify legitimate user activity. However, replication should still be sought before these results contribute directly to the enactment of policy protocols. Additionally, experimentation with alternative wordings of the presented banner should be conducted to separate the effect of the content of the banner from the presentation of a banner upon entry to a computer system. Future research is necessary to elucidate the explicit mechanisms underlying the effect of such banners, as well as comparable deterring cues introduced into cyber environments.

While our work is the first to assess the influence of a surveillance banner on the progression and development of computer system trespassing events, it is not without limitations. First, the results are not necessarily generalizable beyond the setting of large academic institutions in the United States. Further research should investigate the effect of surveillance banners in alternate venues, including government and commercial entities, to assess whether our findings are externally valid. Related to this, it should be noted that our sampling procedure necessarily limits generalizable statements from being applied to the entire population of trespassers, as our target computers are targets of opportunity (i.e., random computers with vulnerabilities that are logged onto the Internet) as opposed to targets of choice (i.e., target computers that raise specific interest among trespassers, such as computers employed by the Federal Bureau of Investigation or Central Intelligence Agency). System trespassers who are more sophisticated and looking for specialized documents or information would not necessarily attempt to enter a Linux-based device that retains no inherent sensitive information or utility beyond its computing power. Nevertheless, there is reason to believe that the majority of system trespassers are looking for targets of opportunity as opposed to targets of choice, which should limit the impact of this sampling bias (McQuade 2006). Second, the target computers we used in this experiment ran a Linux-based operating system and retained a known vulnerability to allow offenders access to the system. Further research is necessary to ascertain whether our findings bear fruit toward reducing the severity of trespassing events targeting systems using alternate operating systems, system configurations, and/or means of access. Third, our experiment was unable to differentiate between potential human and bot-based attackers, and we are further unable to comment on the exact proportion of the sample that is human as opposed to autonomous. However, this would likewise only serve to downward bias any observed effects. Further, this produces a more accurate depiction of the environment in which an IT manager would deploy such a treatment in practice, as they would not
be able to immediately distinguish automated from human attackers in all cases. Fourth, our experimental design assumes that system trespassers are English literate and would be able to understand the presented surveillance banner. While this limitation would serve to downward bias the results obtained from our analysis, additional analyses identifying system trespassers’ English reading skills would be beneficial to the field. Fifth, we neither had nor employed measures to tap system trespassers’ perceptions regarding the probability of detection and the certainty of sanctions to their outcomes. In fact, in order to ensure that the Internet infrastructure of the large American University we worked with is not put at risk, we were not allowed to communicate with the system trespassers at all. Related to this, it is possible that the surveillance banner’s content did not present system trespassers with a sufficient dosage that would prevent their engagement with the attacked system during short trespassing events. Future research should employ stronger interventions that will increase system trespassers’ perception of detection and sanctions. For instance, future studies should employ more active signs of surveillance on the attacked system, such as presenting the last commands entered by the system trespasser on the computer screen and revealing the IP address from which the trespasser communicates with the attacked system. Finally, we did not differentiate commands entered into our systems on a basis of harm or malicious intent. Instead, we classified any commands entered into the system as constituting an escalation in the offense. Without having access to the intruder directly, it is impossible to assign intent of any type to the commands entered into our systems in all cases. Future research is needed toward exploring alternative conceptualizations of outcome measures for evaluating deterrent cues in cyberspace.

Despite these limitations, our study provides preliminary and modest evidence regarding the effectiveness of a surveillance banner in restricting system trespassers’ engagement with an attacked computer system. These findings should encourage social scientists to further apply soft science models in their quest to understand human behavior in cyberspace. At the same time, our findings should encourage IT practitioners to consider the application of human-based security solutions in the systems they are protecting. While adoption of human-based security solutions may not prevent the occurrence of trespassing events, it may delay trespasser behaviors on the attacked system and mitigate the resultant damage.

**Acknowledgment**

We thank Ray Paternoster and Tom Loughran 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 insight on this research.

**Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was conducted with the support of the National Science Foundation Award 1223634.

**Notes**

1. An Internet protocol (IP) address is the public identification number that a computer system uses to connect to the Internet and communicate with other systems. A single IP address consists of four numbers in sequence ranging from 1 to 256 (e.g., 111.11.111.11). The 300 IP addresses we used were from a single, continuous subnet devoted to this experiment. A subnet is a range of IP addresses (e.g., 111.11.111.1–111.11.111.256).

2. The target computers were deployed with default software for Linux-based operating systems. No special software or data were added to these systems, other than the surveillance software incorporated as a specific treatment.

3. These toolkits enable the prospective intruder to select a username to attempt to gain access to the system with.

4. Since the number of attempts limits the ability of human users to arbitrarily type in login credentials by hand, it prioritizes intruders’ use of brute force toolkits. Most brute force toolkits do not report the number of trials needed to gain legitimate access to the system, so whether the threshold was set at 100 or 1,000 attempts should not have an immediate bearing upon the resulting sample included in the study.

5. The randomization procedure took place before the intruders could access the system and can thus be considered to be exogenous to any and all differences observed on the target system over the course of a deployment.

6. These results did not substantively differ from those presented in this work and are available upon request.

7. Zero second duration and other short trespassing events could be prompted by connectivity issues, server issues, and/or human error in closing the window prematurely. The distribution of these connectivity issues theoretically should not be related to whether the trespasser was assigned to receive a surveillance banner. This logic is especially salient with regard to 0second duration sessions, as the
intruder would not have been able to see, let alone read, the surveillance banner in these instances.

8. These short trespassing events are only removed for the purposes of these analyses with the first system trespassing incident. They are retained for all of the subsequent analyses.

References


Author Biographies

Theodore Wilson is a PhD student in the Department of Criminology and Criminal Justice at the University of Maryland–College Park. His research interests include cybercrime, offender decision making, and quantitative methods.

David Maimon is an associate professor of Criminology and Criminal Justice at the University of Maryland–College Park. His research interests include cybercrime, experimental methods, and community and crime.

Bertrand Sobesto is a security engineer in the Division of Information Technology at the University of Maryland–College Park. His research interests include cybercrime, malware examination, and network flow analysis.

Michel Cukier 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 interests include system dependability and security issues.