

The effect of online ad campaigns on DDoS-attacks: A cross-national difference-in-differences quasi-experiment*

Asier Moneva^{1,2}

E. Rutger Leukfeldt^{1,2}

¹ Netherlands Institute for the Study of Crime and Law Enforcement (NSCR), Amsterdam, Netherlands; ² Center of Expertise Cyber Security at The Hague University of Applied Sciences, The Hague, Netherlands

Abstract

Research summary: European law enforcement agencies have begun to use targeted online ad campaigns to raise cybercrime awareness among at-risk populations. Despite their rapid proliferation, there is little research to support their efficacy and effectiveness. This study uses a quasi-experimental difference-in-differences design to evaluate the effect of seven campaigns deployed in 2021 and 2022 on the volume of DDoS-attacks recorded in six European countries: Denmark, Finland, Netherlands, Norway, Sweden, and Portugal. The results show mixed effects, suggesting that the campaigns are not clearly effective in reducing DDoS-attacks in the short term.

Policy implications: Law enforcement has partly justified the use of targeted online ad campaigns on the premise that they reduce DDoS-attacks. However, this study shows that the evidence in this regard is inconclusive. If public support for the use of such campaigns is

*This is the peer reviewed version of the following article: 'Moneva, A., & Leukfeldt, R. (2023). The effect of online ad campaigns on DDoS-attacks: A cross-national difference-in-differences quasi-experiment. *Criminology & Public Policy*, 1–26', which has been published in final form at <https://doi.org/10.1111/1745-9133.12649>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This article may not be enhanced, enriched or otherwise transformed into a derivative work, without express permission from Wiley or by statutory rights under applicable legislation. Copyright notices must not be removed, obscured or modified. The article must be linked to Wiley's version of record on Wiley Online Library and any embedding, framing or otherwise making available the article or pages thereof by third parties from platforms, services and websites other than Wiley Online Library must be prohibited.

to be secured in the long term, law enforcement will likely need to rely on stronger arguments. The effect of this type of campaigns needs to be further investigated.

Keywords: Cybercrime, DDoS-attacks, difference-in-differences, online ad campaigns, quasi-experiment

Introduction

Faced with the growing threat of cybercrime, law enforcement agencies around the world are stepping up their prevention efforts (e.g., Europol & Dutch National Police, 2021; Interpol, 2022; National Crime Agency, 2022). However, due to the scarcity of rigorous evaluations, it is unclear whether currently used cybercrime prevention strategies actually work to reduce cybercrime (e.g., Brewer et al., 2019). The lack of knowledge about the effect that these strategies may have on the target population is dangerous because it can lead to the oversight of possible negative consequences, such as criminogenic effects, and the wastage of resources. On the contrary, knowledge about the effect of the strategies contributes to generating an evidence base to determine what works to reduce cybercrime and to more efficient use of resources. Some police strategies have focused on reducing specific forms of cybercrime, like online ad awareness campaigns against (distributed) Denial-of-Service or (D)DoS-attacks.

DDoS-attacks are a type of cyber-attack that uses multiple compromised computer systems to flood a targeted server or network with Internet traffic, making it unavailable to legitimate users. Law enforcement agencies specifically focus on this type of cybercrime because it is considered to be *entry-level* (National Assessments Centre, 2022; National Cyber Crime Unit, 2017); in other words, a low-level illicit online activity—often perpetrated by individuals with limited technical expertise and resources—that may constitute a pathway to a serious criminal career. In their paper on the effects of police interventions on DoS attack markets, Collier and colleagues (2019) reported that the online ad campaigns designed by the UK National Crime Agency (NCA) to raise awareness against cybercrime and deter potential cybercriminals “led to a clear and lasting reduction in the number of attacks” (p. 58). Building on the work of Collier and colleagues (2019), this paper uses a cross-national quasi-experimental design to evaluate the effect of seven online ad awareness campaigns, coordinated by the Netherlands Police and implemented nationwide in six European countries, that aimed to divert and deter potential cybercriminals from committing DDoS-attacks and other cybercrimes.

Cybercrime prevention strategies are part of the national security agenda of an increasing number of countries around the world. Many of them are aimed at enhancing the cyber resilience of organizations and individuals, so that they are able to withstand, recover from, and adapt to cyber-attacks (for a definition of cyber resilience see Dupont, 2019). For example, at the international level, Interpol has launched several awareness campaigns (most recently “#YouMayBeNext”) to encourage online self-protection of users and organizations against cyber extortion threats (Interpol, 2022). Europol, the Netherlands Police, Kaspersky, and McAfee have created the public-private “No More Ransom” initiative to provide ransomware

victims with free tools to decrypt their files (Europol & Dutch National Police, 2021; see also Filiz et al., 2021). In the United Kingdom, the National Cyber Security Centre has developed the “Cyber Essentials” scheme, a toolkit to certify organizations that adopt basic cyber security measures (BritainThinks, 2020; see also Kemp, 2023). These campaigns emphasize the self-protection role that targets of cyber-attacks can play, although this is not the only way of prevention.

Another type of strategy, inspired by focused deterrence (Kennedy, 2012), focuses on changing the behavior of individual cyber offenders. Behavioral change would be achieved by combining an appropriate threat of sanction with the provision of viable pro-social alternatives. In this vein, the Federal Bureau of Investigation (FBI) of the United States published their “Cyber Strategy”, as a statement of intent to deter cybercriminals by increasing both the perceived risk of their actions and the punishment inflicted on them (Federal Bureau of Investigation, 2020). In Europe, together with the Dutch Ministry of Justice, the Netherlands Police implemented the “Hack_Right” program to cut pathways into cybercrime for young hackers and present them with pro-social alternatives (Schiks et al., 2021). The National Crime Agency (NCA) of the United Kingdom developed the “Cyber Choices” program, which aims to increase awareness of the illegality of cybercrime while promoting legal alternatives using online ad campaigns (Collier et al., 2021; National Crime Agency, 2022). These campaigns are sometimes combined with *knock-and-talk* police visits and educational workshops for potential cybercriminals who have acquired cybercrime tools or services (Collier et al., 2022). Online ad campaigns were later adopted by the Netherlands Police to redirect people who have not yet developed a serious cybercriminal career to cybersecurity (Moneva et al., 2022). At present, online ad campaigns continue to be developed and implemented in several European countries.

Cybercrime prevention through online ad campaigns

Since 2018, the NCA has pioneered the use of targeted advertising campaigns to deter youth from cybercrime and divert them toward prosocial behaviors, such as cyber security. Researchers have termed this strategy *influence policing* (Collier et al., 2021). In collaboration with behavioral psychologists, the NCA designed a series of messages and deployed them through Google Adwords (now Google Ads)—a popular online platform to display brief ads—so that potential cybercriminals could make an informed choice (Collier et al., 2022). In the original 2018/2019 campaign, users who entered into the Google search engine terms related to DDoS-attacks—such as “how do I DDoS”, “booter services”, “is hacking wrong”—were exposed to the ads (Figure 1). By clicking on the ads, users were redirected to a web page that displayed information about what DDoS-attacks are, the consequences that victims and perpetrators suffer, and a list of alternatives to promote pro-social behavior in cyber security. In January 2018, the first month the campaign was deployed, the ads generated about 12,700 impressions and 517 clicks (National Crime Agency, 2022). Since then, the NCA has been adjusting its keywords to increase the engagement generated by the ads, generating “some 5.32

million impressions and more than 57,000 clicks” in just over two weeks in 2020, according to a quote from the press (Krebs, 2020).

[DDoS is illegal - National Crime Agency - www.nationalcrimeagency.gov.uk](https://www.nationalcrimeagency.gov.uk)

[Ad www.nationalcrimeagency.gov.uk](https://www.nationalcrimeagency.gov.uk)

DDoS attacks are illegal in the UK under the Computer Misuse Act 1990

Figure 1: Original online ad by the NCA

In 2021, the Netherlands Police decided to carry out similar online ad campaigns. Lessons learned in the targeted online ad campaigns deployed in the United Kingdom informed the design and implementation of new campaigns in the Netherlands. There, the first campaign was deployed in 2021 and, after a short pilot, different ads were tested for 14 weeks to see which type of message generated the most engagement among the target population: users interested in DDoS-attacks and the gaming industry (Moneva et al., 2022). The ads generated 71,475 impressions and 4,457 clicks, and the results of the study showed that social messages (i.e., those that “use social consequences of behavior to invoke potential negative reinforcement from peers”) were the most engaging. Note that the most engaging ads are not necessarily also the most effective in changing behavior. Again, findings informed subsequent campaigns that were implemented in more countries. Currently, the campaigns have been adapted to advertise against several cybercriminal behaviors [launching DDoS-attacks, hacking, using Remote Access Tools (RATs)] in different countries, and have therefore been translated from English and Dutch into several languages (Danish, Finnish, Norwegian, Portuguese, and Swedish).

Prevention mechanisms underlying online ad campaigns

Criminologists have long sought to prevent crime by influencing criminal decision making (Clarke & Cornish, 1985). When provided with appropriate alternatives, criminal behavior would even be replaced by prosocial behavior (Kennedy, 2012). Broadly speaking, there are two main ways to achieve this change: by focusing on individuals and changing criminal predispositions, or by focusing on events and changing crime contexts (Hirschi & Gottfredson, 1986; Wortley & Townsley, 2017). The second option is characteristic of situational approaches to crime prevention and is considered to be immediate, inexpensive, and arguably more effective (Clarke, 1997). In criminology, the practical means of altering the situational determinants of crime and influencing the decision-making process of offenders to make crime less likely is situational crime prevention (Clarke, 1980, 2017). In a similar vein, recent advances in behavioral economics and social psychology propose to change behavior by manipulating the choice architecture, understood as the design of environments or systems that influence people’s decisions and behaviors (Thaler & Sunstein, 2009). This approach, known as *nudging*, has been applied in criminology to research offender decision-making and provide insights to develop crime prevention policies (for a review, see Pogarsky et al., 2018). Since nudges aim to influence

behavior without limiting freedom of choice (Thaler & Sunstein, 2009), they align with the strategy of the NCA and the Netherlands Police of using online ad campaigns to promote informed choices among offenders.

From a theoretical perspective, it is worth noting that some consider nudging to be “simply an arm of what is widely known as ‘situational crime prevention’” (Roach et al., 2017, p. 32). In this sense, online ad campaigns have been considered a situational crime prevention measure capable of removing excuses for noncompliance with the law by alerting consciences and controlling disinhibitors (Moneva et al., 2022). However, there may be a subtle but important difference between situational crime prevention and nudging that lies in the timing of implementation. Whereas situational crime prevention would require that the measures be applied immediately prior to the criminal behavior (Felson & Eckert, 2018; see also Clarke, 1980, 2017), nudges could be implemented earlier, to produce their effects in the short term (Thaler & Sunstein, 2009). In the case of online ad campaigns, potential cybercriminals may see the ads at different temporal distances from the crime, depending on their intent. When the criminal intent is still forming and potential offenders are simply looking for information, the distance is greater. This means that advertisements can be seen hours, days, weeks, or months before the crime. In contrast, when the criminal intent is fully formed and potential offenders are looking for cybercrime services, the distance to the criminal behavior is minimal. In these cases the ads may be seen minutes or even seconds before the crime. The greater the distance between the viewing of the ad and the criminal behavior, the further away the situational crime prevention mechanism is. In such a case, it may be more appropriate to speak of nudging.

It is unclear, however, whether nudges are effective in changing the behavior of potential cybercriminals. Experimental evidence using honeypots and deterring warning banners yields mixed results. Experiments on hacking behavior show that deterrent banners do not have an effect on the volume of initial or repeat intrusions on computer systems (Maimon et al., 2014; Vetterl, 2020; Wilson et al., 2015), but do appear to reduce the duration of intrusions (Maimon et al., 2014; Stockman et al., 2015). Note that these studies have been criticized for not being able to distinguish humans from bots (Holt, 2017; Vetterl, 2020). Another experiment on Internet users shows that some warning banners halved the number of visits to a barely legal pornography website (Prichard et al., 2021). Two possible explanations for the limited effect of deterring warning banners are the values and perceptions of hackers, such as peers’ support and curiosity, and the use of tools that hinder actor attribution (Holt, 2017). Yet a study examining the relationship between targeted online ads by the police and cybercrime (DDoS-attacks) suggests that the campaigns have a reduction effect (Collier et al., 2019).

The Present Study

After developing and implementing two online ads campaigns to cut pathways into cybercrime in the Netherlands in 2021 (Moneva et al., 2022), the Team High Tech Crime of the Dutch Po-

lice coordinated a third international campaign under the European Multidisciplinary Platform Against Criminal Threats (EMPACT) and the scrutiny of Europol and Interpol in 2021-2022. This campaign was joined by representatives of the national police forces of five other European countries: Denmark, Finland, Norway, Portugal, and Sweden. All three campaigns share the focused deterrence philosophy of the original campaign proposed by the NCA’s National Cyber Crime Unit as part of the Cyber Choices project, represented by two “Ds”: Deter and Divert. This paper evaluates the effectiveness of these campaigns in reducing the volume of DDoS-attacks in the six countries participating in the consortium using data collected by the Cambridge Cybercrime Center (Thomas et al., 2017).

Depending on the design of the ads, there are arguably two overlapping nudging mechanisms that can influence user behavior: deterrence, through *informational nudges*, and socialization, through *social comparison nudges* (Pogarsky & Herman, 2019). In Criminology, deterrence refers to the use of punishment or threat of punishment—directly or indirectly—to discourage criminal behavior (e.g., Nagin, 2013). On the other hand, socialization refers to the process by which individuals acquire the attitudes, values and behaviors of their peers towards crime (e.g., Akers & Jennings, 2015). Law enforcement often resorts to informational nudges to convey deterrent messages to potential offenders and thus increase their perception of risk in order to change their behavior, as in the case of online ad campaigns (Collier et al., 2021, 2019; Moneva et al., 2022). An example of such messages would be “A DDoS attack is illegal. Carrying out a DDoS is illegal in the Netherlands under the Criminal Code”. Sometimes law enforcement resorts to social comparison nudges to show social disapproval towards criminal behavior, in what have been called social messages (Moneva et al., 2022). An example of such messages would be “DDoS ruins it for everyone. So you want your friends to be unable to play games because you play a prank?”. Both deterrent and socializing mechanisms in campaigns could therefore jointly influence the decision making process of potential cybercriminals.

There are two challenges to evaluate the effect of the campaigns on DDoS-attacks. The first is to reach the *right population* with the ads. The literature suggests that DDoS-attacks are an entry-level cybercrime often launched or hired as-a-service by young people interested in video games (Noroozian et al., 2016; Thomas et al., 2017). Given that the target population of the campaigns are young people under the age of 34, we expect the campaigns to reach potential cybercriminals (Moneva et al., 2022). The second is attribution: identifying the country of *origin and target* of the attacks (Brenner, 2007). If the vast majority of DDoS-attacks originate in a country that does not participate in the campaigns, the effect of the online ads will be minimal. A recent analysis of data from the `stresser.gg` booter reveals that a large portion of the attacks that originate in the same country they are targeted (Santanna, n.d.). For example, of the 1521 attacks recorded against the Netherlands, 731 (48.1%) originated in the Netherlands. Given this ratio, we expect that the effect of the campaigns, if any, would be statistically significant ($\alpha = 0.05$). A precedent supports these assumptions. In an evaluation of the effect of law enforcement interventions in DDoS markets, Collier and colleagues (2019) found that the original NCA campaign “may have had the effect of dissuading new users from becoming involved, halting the rising demand for attacks for a period of seven or eight months” (p. 61). Since we use the same data source, we hypothesize that:

The 2021-2022 online ad campaigns against cybercrime reduced the volume of DDoS-attacks in the short term in the countries where the campaigns were implemented compared to similar countries where the campaigns were not implemented.

Methods

Data

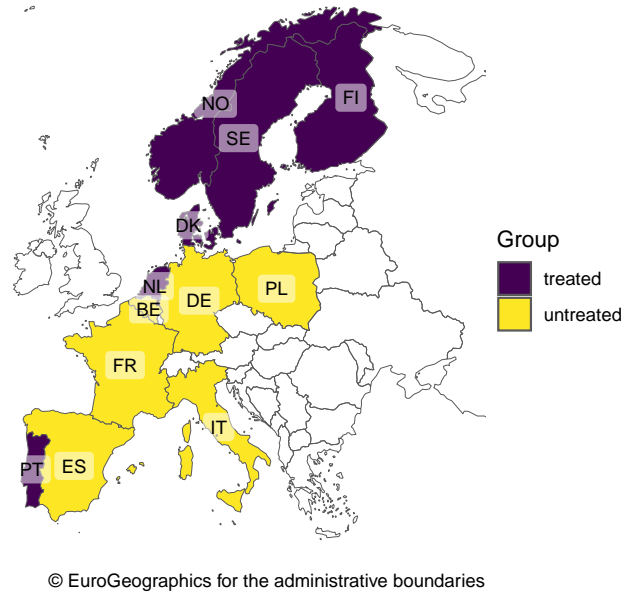


Figure 2: European countries included in the treated and untreated groups

To evaluate the effect of the campaigns on DDoS-attacks, we used User Datagram Protocol (UDP) packet victim data collected by researchers from the Cambridge Cybercrime Centre. To collect the data, the researchers deployed honeypots that mimic the reflector servers commonly used in DDoS-attacks. Through about 100 sensors, the honeypots record incoming packets, most of which refer to websites offering DDoS-as-a-service, often referred to as *booters* or *stressers*. Each UDP packet observation contains the timestamp when it was recorded, its duration in seconds, a count of the number of upcoming packets, the destination port number, and the IP address or prefix of the target. For details on the data, data collection, and ethical issues, see Thomas and colleagues (2017).

We used data for 12 countries: the six that implemented the campaigns or *treated group*—Denmark, Finland, Netherlands, Norway, Portugal, Sweden—and six other similar countries that did not or *untreated group*—Belgium, France, Germany, Italy, Poland, Spain (Figure 2). We followed the three criteria described in Huntington-Klein (2022, p. 433) to choose the untreated group:

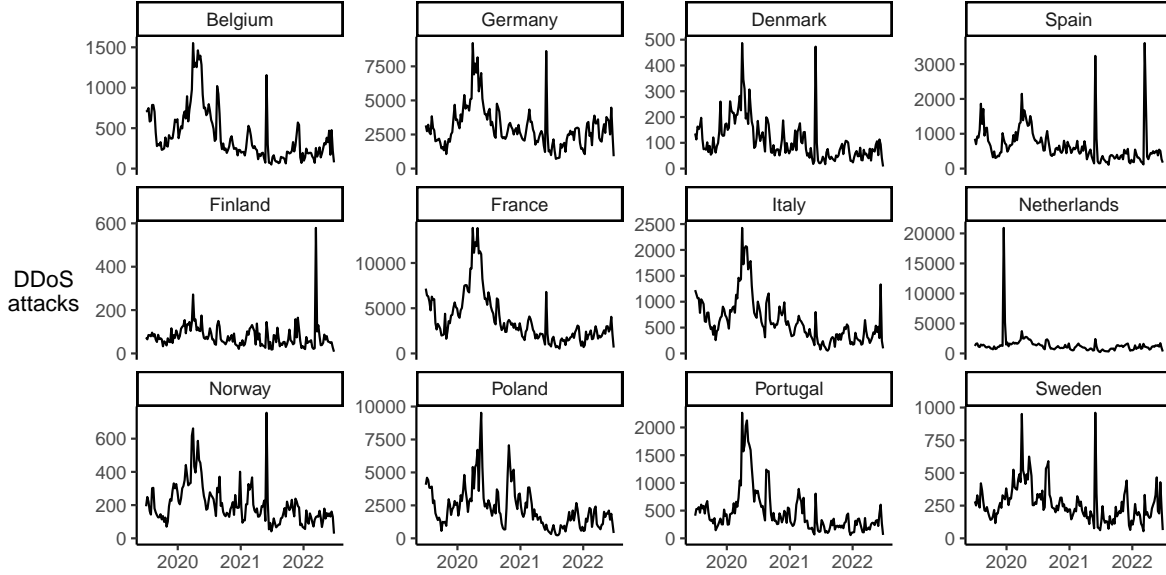


Figure 3: DDoS-attacks trends by country

- We had no reason to believe that the trend of DDoS-attacks would suddenly change in the untreated countries around the implementation of the campaigns. This could occur, for example, in the event of a large-scale police intervention. Note that we were regularly in contact with the Netherlands Police, who coordinated the campaigns.
- The untreated countries are similar in many ways, in the sense that they are all European and developed countries.
- The untreated countries had similar DDoS-attack trajectories to those of the treated countries prior to the implementation of the campaigns. (Appendix A provides empirical support for this claim, and Section 5.4 provides more details on the selection of countries and the analytical strategy.)

In a first phase of data processing, we transformed the IP addresses or prefixes of the packet targets into country codes using the `iptools` R package (Rudis et al., 2021). We then filtered the data by the protocols commonly abused for DDoS-attacks: CHARGEN, DNS, LDAP, MDNS, NTP, PORTMAP, QOTD, SNMP, SQLMon, SSDP (Collier et al., 2019; Thomas et al., 2017). To distinguish actual attacks from mere scans, we also use the threshold of “5 packets per sensor with no gap of more than 900 seconds” (Thomas et al., 2017, p. 4; see also Collier et al., 2019). The final dataset contains 2,061,759 DDoS-attacks. Figure 3 shows the weekly distribution of DDoS-attacks by country.

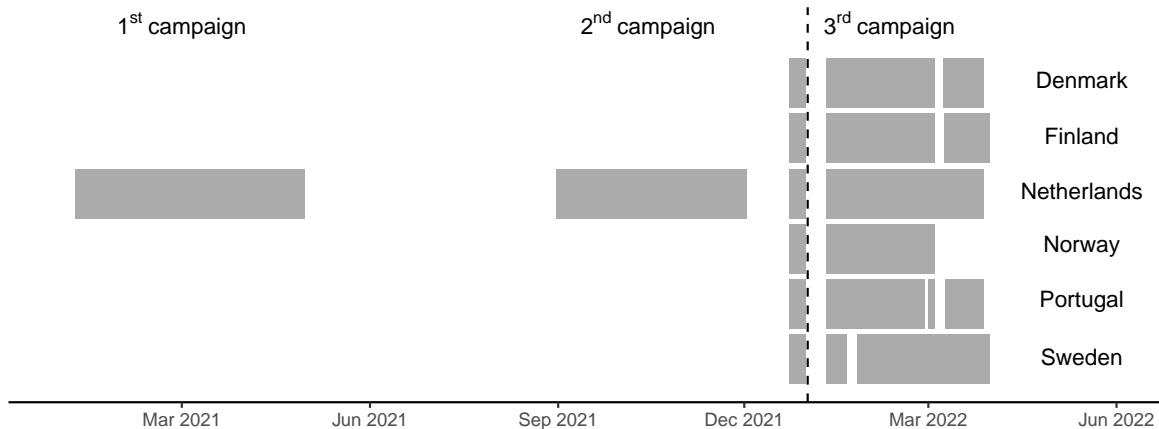


Figure 4: Campaign deployment timeline

Online ad campaigns

European law enforcement agencies deployed three online ads campaigns throughout 2021 and the first quarter of 2022 to reduce cybercrime (Figure 4). The first and second campaigns were deployed by the Netherlands Police in the Netherlands. Including the pilot study, the first campaign ran from January 8 to April 30, 2021 (see Moneva et al., 2022). The second covered the period from August 31 to December 2, 2021, although its effect was fading for about a week by then due to the limited remaining budget. (This campaign was in its full potential from September 1 to November 23.) The third campaign was implemented simultaneously, albeit with slight variations, from December 23, 2021 to the end of March 2022 in the six countries of the treatment group. In Norway, Google prematurely terminated the campaign due to an alleged violation of the circumventing systems policy ¹. The first two campaigns in the Netherlands were specifically aimed at reducing DDoS-attacks in the gaming industry, while the third had a more general purpose and aimed to also reduce acts of hacking and the use of RATs.

Stimuli: online ads and landing pages

The stimulus used in the treatment group was exposure to online ads. In particular, we used Google’s online advertising program, Google Ads, to display short messages in the search results (Google, 2023). The ads appear in the search results when users enter certain cybercrime-related keywords into the search engine from any of the treated countries. Examples of search queries that trigger ads are “free booter”, “how to ddos”, “botnet for hire”, and “stresser service”. The first two campaigns were specifically targeted at DDoS-attacks and their ads showed four possible texts; the third campaign had a more general purpose and its ads showed three

¹“Your account is suspended. Your account violated the Circumventing systems policy.”

Online ads for the first and second campaign:

A DDoS attack is illegal | Gamechangers

[Ad publicaties.politie.nl/DDoS/attack](https://publicaties.politie.nl/DDoS/attack)

Carrying out a DDoS is illegal in the Netherlands under the Criminal Code

DDoS ruins it for everyone | Gamechangers

[Ad publicaties.politie.nl/DDoS/attack](https://publicaties.politie.nl/DDoS/attack)

So you want your friends to be unable to play games because you play a prank?

Do you want to DDoS a game? | Gamechangers

[Ad publicaties.politie.nl/DDoS/attack](https://publicaties.politie.nl/DDoS/attack)

Learn more about DDoS attacks and their impact on gaming

Play fair; losers do DDoS | Gamechangers

[Ad publicaties.politie.nl/DDoS/attack](https://publicaties.politie.nl/DDoS/attack)

Win fairly in an e-sports competition by training your gaming skills

Figure 5: Example online ads translated into English

Online ads for the third campaign:

DDoS ruins it for everyone | DDoS is punishable by law

[Ad https://www.politie.nl/ddos](https://www.politie.nl/ddos)

DDoS-attacks cause serious damage to organizations, businesses and individuals

Deploying a RAT is burglary | RATs may be illegal

[Ad https://www.politie.nl/algemeen/remote-access-tools.html](https://www.politie.nl/algemeen/remote-access-tools.html)

The police are cracking down on cybercrime. In any form and anywhere.

Illegal hacking is breaking and entering | Illegal hacking is punishable

[Ad https://www.politie.nl/algemeen/hacken.html](https://www.politie.nl/algemeen/hacken.html)

Illegal hacking can cause serious harm to organizations, businesses and individuals

Figure 6: Example online ads translated into English

possible texts. Since ads that use native language seem to generate slightly more engagement among users at a slightly lower cost, campaigns that employ them may be more sustainable (Moneva et al., 2022). For this reason, the ads were displayed in the primary official language of the treated countries. Table 5 in Appendix B shows the original texts for each country, along with the target cybercrime. Figure 5 and 6 show the seven type of ads translated into English. These text designs were rotating among the users who triggered them.

When users clicked on the ads, they were redirected to a landing page on the police domain. The page contained information about the specific type of cybercrime mentioned in the ad, the legal consequences of its execution, as well as pro-social alternatives in the field of cyber security. The text is presented in an informal tone adapted to the gamer culture using jargon and memes. For the landing page designs of the first and second campaigns, see Moneva and colleagues (2022). The designs for the third campaign are shown in Appendix C.

Analytic strategy: difference-in-differences

To determine whether, in the countries that implemented the campaigns, online ads had any effect on the volume of DDoS-attacks, we use a difference-in-differences design. Difference-in-differences is a “quasi-experimental research design that researchers often use to study causal relationships in public health settings where randomized controlled trials are infeasible or unethical” (Wing et al., 2018, p. 453). In our design, the six European countries that implemented campaigns constitute the treatment group, while six other European countries that did not implement campaigns constitute the control or *untreated* group. In this way, we can compare, within the treated countries, the periods before and after a campaign was implemented—differences—and compare these periods, in turn, to the same ones in the untreated countries—in differences. There is one exception. Since the third campaign in the Netherlands was implemented shortly after the end of the second, due to possible delayed effects, we do not consider the short period of time preceding it as untreated. Therefore, we did not evaluate that campaign.

To calculate the differences-in-differences, we used two-way fixed effects models, and long-term effect models. We estimated two-way fixed effects models to calculate single estimates of the overall effect of online ad campaigns on DDoS-attacks while controlling for differences between treatment groups and periods with the following formula (Huntington-Klein, 2022, p. 446):

$$Y = \alpha_g + \alpha_t + \beta_1 treated + \epsilon$$

where Y are logged DDoS-attacks; α_g are the fixed effects between treated and untreated countries; α_t are the fixed effects between treated and untreated periods; $treated$ is a binary variable indicating when and where a campaign was active; β_1 is the difference-in-differences estimate of the effect of the campaigns on the DDoS-attacks; and ϵ is the error term.

Table 1: Treated and untreated groups per campaign

Treated group	Untreated group
1st campaign	
Netherlands	Belgium, France, Germany, Italy
2nd campaign	
Netherlands	Belgium, France, Germany, Italy
3rd campaign	
Denmark	Belgium, France, Germany, Spain
Finland	France, Germany, Italy
Norway	France, Germany, Italy
Portugal	France, Germany, Poland, Spain
Sweden	France, Germany, Italy, Poland

We also estimated long-term effect models to examine the effect of the treatment over the course of several weeks using the following formula (Huntington-Klein, 2022, p. 454):

$$Y = \alpha_g + \alpha_t + \beta_{-(T_1-1)}treated + \dots + \beta_{-1}treated + \beta_1treated + \dots + \beta_{T_2}treated + \epsilon$$

where T_1 and T_2 are the weeks before and after the treatment. We chose the week as the unit of analysis because, unlike the month or the quarter, it allows to observe and compare the treatment effect over equal periods, while also taking into account the effect of weekends on criminal activity.

Since the outcome variable is right skewed—as is usual with crime counts—we take its logarithm for the analyses. This also minimizes the impact of outliers on the results. The resulting model estimates can be more easily interpreted by applying the formula $(exp(\beta) - 1) \times 100$ (Wooldridge, 2020), which calculates the percentage increase in DDoS-attacks in the treated country for each one-unit increase in DDoS in the untreated countries.

To check whether the untreated group was appropriate, we evaluated the parallel trends assumption. This assumption implies that the treated and untreated groups would have described a similar variation over time if the treatment had not been implemented. To evaluate the parallel trends assumption, we initially compared the DDoS-attack trends between each treated country and all the untreated countries for each campaign, both visually and statistically (Appendix A). Note that these tests are not final, but indicators of how plausible the assumption is (Huntington-Klein, 2022). For the statistical tests we estimated the same two-way fixed effects models in a fake treatment period of the same duration as the real one immediately prior to the treatment. The results of this test of *prior trends* test should reflect that there are no differences between groups to support that the DDoS attack trends are parallel and—therefore—that the untreated group is appropriate (Huntington-Klein, 2022). Although visual inspection showed similar DDoS attack trends across countries, suggesting

Table 2: Difference-in-differences (two-way fixed effects) estimates of the effect of online ad campaigns on logged DDoS-attacks

Treatment	Estimate	Std. Error	p-value	Sig.	Num. Obs.	Adj. R-squared
1st campaign						
Netherlands	0.056	0.117	0.654		490	0.936
2nd campaign						
Netherlands	0.259	0.078	0.029	*	160	0.968
3rd campaign						
Denmark	-0.313	0.153	0.110		735	0.953
Finland	0.132	0.208	0.571		588	0.959
Norway	-0.077	0.210	0.737		572	0.955
Portugal	-0.448	0.124	0.023	*	735	0.903
Sweden	0.031	0.147	0.843		735	0.934
FE: Week						
FE: Country						

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

that parallel trends were plausible, statistical tests did not do so in all cases. To further support the parallel trends assumption, we adjusted the untreated group for each treated country by eliminating countries with discordant DDoS trends until the difference-in-differences were null. This resulted in different control groups for each country (Table 1).

Software

We used R version 4.2.0 (R Core Team, 2022) and RStudio version 2022.02.3 (RStudio Team, 2020) to process and analyze the data. We used the tidyverse R package version 1.3.1 (Wickham et al., 2019) to manipulate and visualize the data, and the fixest R package version 0.10.4 (Bergé, 2018) to estimate the statistical models.

Results

We examined the effect of online ad campaigns on DDoS-attacks in two ways: first we fitted two-way fixed effects models to obtain single estimates summarizing the overall effect of the campaigns; and then we fitted long-term effect models to obtain a series of estimates over several weeks to gain insight into the effect on DDoS trends.

The two-way fixed effect model results show mixed effects of the campaigns on DDoS-attacks (Table 2). In the Netherlands, it appears that the second campaign had a significant positive

effect, increasing DDoS-attacks by 29.6% ($est = 0.259$; $SE = 0.078$; $p = 0.029$). In contrast, in Portugal, it appears that the third campaign had a significant negative effect, reducing DDoS-attacks by 36.1% ($est = -0.448$; $SE = 0.124$; $p = 0.023$). In the other campaigns and countries, there were non-significant increases and decreases ranging from -26.9% in Denmark to 14.1% in Finland. Adjusted R-squared figures, all greater than 0.903, show that the model fits the data well and thus adequately represents the relationship between treatment and outcome.

Table 3: Difference-in-differences (Long-Term Effects) estimates of the effect of online ad campaigns on logged DDoS-attacks

Treatment	1st campaign (Netherlands)				2nd campaign (Netherlands)				3rd campaign (Denmark)				3rd campaign (Finland)				3rd campaign (Norway)				3rd campaign (Portugal)				3rd campaign (Sweden)			
	Estimate	Std. Error	p-value	Sig.	Estimate	Std. Error	p-value	Sig.	Estimate	Std. Error	p-value	Sig.	Estimate	Std. Error	p-value	Sig.	Estimate	Std. Error	p-value	Sig.	Estimate	Std. Error	p-value	Sig.	Estimate	Std. Error	p-value	Sig.
Week 1	0.028	0.065	0.690		0.176	0.133	0.256		-0.335	0.265	0.276		0.053	0.177	0.785		-0.589	0.177	0.045	*	0.330	0.143	0.083		0.196	0.134	0.219	
Week 2	-0.011	0.166	0.950		0.153	0.150	0.365		-1.647	0.215	0.002	**	0.206	0.200	0.378		-0.108	0.200	0.627		-0.414	0.154	0.055		1.256	0.134	0.001	***
Week 3	-0.146	0.186	0.475		0.212	0.144	0.213		-0.288	0.177	0.178		0.139	0.055	0.088		-0.196	0.055	0.039	*	0.644	0.136	0.009	**	0.612	0.046	0.000	***
Week 4	0.002	0.186	0.991		0.269	0.119	0.088		0.222	0.255	0.433		-0.143	0.156	0.427		-0.372	0.156	0.097		0.639	0.137	0.010	**	0.895	0.110	0.001	**
Week 5	0.224	0.155	0.224		0.108	0.117	0.407		-0.349	0.206	0.165		0.307	0.178	0.183		-0.159	0.178	0.438		0.735	0.185	0.016	*	0.573	0.126	0.010	*
Week 6	0.155	0.245	0.562		0.455	0.153	0.041	*	-0.053	0.127	0.699		0.001	0.125	0.993		-0.513	0.125	0.026	*	0.606	0.116	0.006	**	0.427	0.092	0.010	**
Week 7	0.162	0.255	0.560		0.553	0.075	0.002	**	-0.396	0.061	0.003	**	0.075	0.200	0.734		-0.714	0.200	0.038	*	0.882	0.100	0.001	***	0.200	0.185	0.340	
Week 8	0.217	0.190	0.316		0.271	0.108	0.066		-0.330	0.144	0.083		0.201	0.234	0.453		-0.054	0.234	0.833		0.930	0.231	0.016	*	0.359	0.292	0.285	
Week 9	0.144	0.166	0.436		0.369	0.050	0.002	**	0.169	0.168	0.371		0.220	0.298	0.513		-0.305	0.298	0.381		0.646	0.232	0.050	*	0.186	0.286	0.552	
Week 10	0.195	0.113	0.161		0.159	0.123	0.265		-0.359	0.171	0.104		-0.125	0.262	0.666		-0.151	0.262	0.604		0.828	0.117	0.002	**	0.282	0.181	0.193	
Week 11	0.223	0.143	0.195		0.192	0.140	0.243		-0.483	0.196	0.069		-0.683	0.323	0.125		-0.543	0.406	0.273		0.020	0.230	0.936		-0.239	0.251	0.395	
Week 12	0.167	0.185	0.418		0.046	0.123	0.727		0.449	0.490	0.324		-0.937	0.122	0.005	**					-0.149	0.429	0.746		-1.176	0.159	0.002	**
Week 13	0.463	0.246	0.133		0.059	0.158	0.729		-1.034	0.654	0.189		2.021	0.195	0.002	**					0.172	0.646	0.803		0.449	0.160	0.049	*
Week 14	0.216	0.163	0.256		0.053	0.257	0.848		-0.636	0.520	0.288		0.468	0.099	0.018	*					-0.319	0.539	0.585		0.183	0.117	0.192	
Week 15	0.102	0.111	0.410						-0.742	0.147	0.007	**	0.417	0.248	0.190						0.147	0.203	0.510		0.545	0.204	0.056	
Week 16	-0.134	0.152	0.426																									
Week 17	0.104	0.084	0.281																									
Num. Obs.	490.000				160.000				735.000				588.000				572.000				735.000				735.000			
Adj. R-squared	0.964				0.966				0.964				0.979				0.969				0.918				0.948			
FE: Week																												
FE: Country																												

Note: * p < 0.05, ** p < 0.01, *** p < 0.001

The long-term effect model results also show mixed effects of the campaigns on DDoS-attacks (Table 3). Adjusted R-squared, for all models above 0.918, again indicate that the fixed effects and the treatment explain a high proportion of the variation in DDoS-attacks. Except for the first campaign, there were occasional significant increases and decreases in all countries compared to the week prior to implementation. This may be because weekly estimates are more responsive to the temporal concentration of DDoS-attacks than estimates over longer periods, although there may be another reason. For the estimates of these models to be reliable, both the reference point and the estimates corresponding to the untreated period must fluctuate around 0. This occurs at almost all times in almost all the campaigns, except at very specific times (e.g., about 20 weeks before the deployment of the campaigns in the Nordic countries) and consistently in the case of Portugal. Since the reference point for the Portuguese campaign is too low with respect to the average trend described by the estimates in the untreated period, the model shows positive weekly coefficients during the campaign. However, compared to the general trend, these weekly estimates exhibit a null or negative effect. These nuances are not reflected in the results of the two-way fixed effect model, which results in a more straightforward interpretation. Therefore, in cases like this, looking at the trends that the estimates describe over time, instead of the coefficients, may be more insightful.

Figure 7 shows the weekly estimates of DDoS-attacks over time and the trends they describe. Each plot in the figure shows the evaluation results for each campaign and country. The vertical axis shows the logged estimates of DDoS-attacks with 95% confidence intervals. The horizontal axis shows the number of weeks since the reference point—the week prior to the deployment of the campaigns—represented in the plots as the only estimate without a confidence interval. The shaded area corresponds to the period when the campaigns were active. To examine the effect of campaigns on DDoS attack trends, we added a purple trend line based on a linear regression model for the period when campaigns were and were not active and compared them. The slope of the line reflects the strength and direction of the relationship between the passing of time and DDoS-attacks; or, in other words, how DDoS-attacks vary over time.

Again, the trend analysis shows mixed results. It appears that the first campaign in Netherlands did not have an effect on DDoS-attacks, as the initial flat trend is also followed by another flat trend. However, in the second campaign, also in Netherlands, we observe a downward trend change despite the overall increase—although the untreated period is shorter in the second campaign than in the first campaign and is possibly influenced by the latter. The third campaign is particularly interesting because it allows cross-national comparisons over the same period. In Denmark, the trend is maintained, although the volume of DDoS-attacks reduced. In Finland, as in Denmark, the trend is maintained but again reduced. Norway also shows a downward trend change and a reduction in the volume of attacks. In Portugal, the only southern European country in the sample, we observe a downward trend change after a few weeks. In Sweden the increasing trend reverses to a decreasing one.

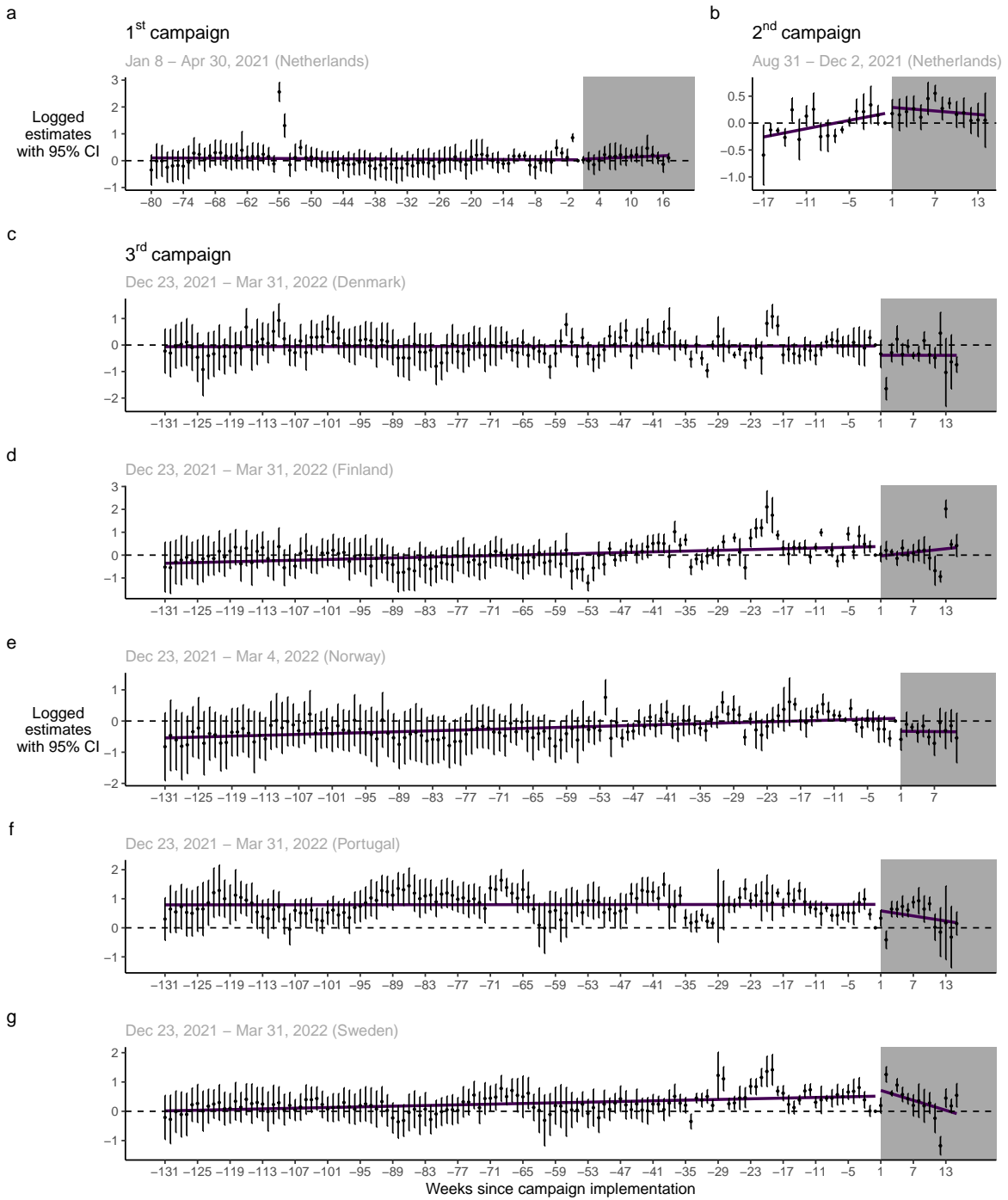


Figure 7: Dynamic effects of online ad campaigns on logged DDoS-attacks per country

Discussion

This article presented the results of a quasi-experimental, cross-national evaluation of seven online ad awareness campaigns aimed at cutting off pathways into cybercrime for Internet users interested in DDoS-attacks and other cybercrime. Using the volume of DDoS-attacks recorded by an infrastructure of sensors deployed by the Cambridge Cybercrime Centre as an objective outcome measure (see Collier et al., 2019; Thomas et al., 2017), the evaluation compared the variation in attacks before and while the campaigns were active in twelve countries, six of which implemented them. The difference-in-differences estimates suggest the campaigns produced mixed effects on DDoS-attacks. Two-way fixed effects model results were non-significant in five of the seven campaigns, significant and positive in one campaign (increase), and significant and negative in another (decrease). Further visual inspection of the long-term effect weekly estimates suggests that the linear trends of DDoS-attacks remain the same or decrease when the campaigns were active compared to the previous period of inactivity. Overall, the results do not provide clear support for the hypothesis that online ad campaigns reduce DDoS-attacks.

There are several possible explanations for the contrast between these results and those of Collier and colleagues (2019). First, it should be noted that during the period for which Collier and colleagues reported the effect of the campaigns, the NCA was also conducting knock-and-talk visits and subsequent workshop interventions with potential cybercriminals (Collier et al., 2022). It is therefore possible that the reduction effect reported was not due solely to the campaigns, but to the knock-and-talk visits, or rather the combination of the two. This explanation would be in line with other crime reduction findings from focused deterrence strategies that use similar outreach strategies to knock-and-talk visits to warn prolific (traditional) offenders (Abt, 2019; Kennedy, 2012). Another possible explanation for our results is that, as argued by Holt (2017), the effect of deterrence in cyberspace is limited, which would render the campaigns ineffective. In line with this, some authors argue that potential cybercriminals may perceive a low risk of detection and therefore low certainty of punishment, which could be amplified by the influence of deviant peers (Brewer et al., 2019). Although not all ad texts used in the evaluated campaigns contain deterrent informational nudges, but also—for example—social comparison nudges (Moneva et al., 2022), it could be argued that by including the name or logo of the police as an authority figure in the ads or landing pages could have an inherent deterrent meaning. However, this would not explain why the results of Collier and colleagues (2019) suggest a reducing effect of the campaigns, since all of the ad texts used by the NCA were deterrent. In that case, the interpretation would be the opposite: that the ads of the NCA, having a higher proportion of deterrent texts, would have had a greater reduction effect.

Certain events (e.g., booter takedowns, cyber operations, national stay-at-home orders) as well as seasonality can have an impact on the temporal distribution of DDoS-attacks (Collier et al., 2019). It is therefore important that seasonality is consistent for the treated and untreated groups. Visual and statistical analyses supporting the parallel trends assumption suggest that temporal discrepancies in the distribution of DDoS-attacks between countries was not a serious

threat to the reliability of the results. It could also be debated whether the discrepancy is due to the use of different analysis techniques. However, we believe that if the results do not replicate in a different segment of the same data source, it is possible that the original findings were not entirely robust. Findings from different evaluations should replicate to inform public crime prevention strategies with confidence. We believe that our findings are persuasive in this regard, as they refer to evaluations of seven campaigns, in three different time periods, and six different countries.

Assessing the key research assumptions

Our research design relied on three important assumptions that, if not supported, are likely to affect the findings. Although we already addressed them previously, it is worth revisiting them in light of the results. First, in order to observe a significant reduction in DDoS-attacks, the ads had to reach the right population. The campaigns delivered the ads to young males searching for information about cybercrime in general and also about specific cybercrime services. However, not everyone who meets these conditions will commit cybercrime in the short term. In fact, many never will. In addition, it should be noted that some users use ad blocking software. This type of software can block most online ads, especially the most annoying ones, such as pop-ups, or flashing animations (Coalition for Better Ads, n.d.), but some ad blocking software does not filter Google Ads because they are considered non-intrusive. As a member of the [Coalition for Better Ads](#), Google works to make its ads more acceptable to users, which may reduce blocking rates.

The second assumption has to do with the problem of attribution in cyberspace. It refers to the difficulty of accurately identifying the actors responsible for cyber-attacks (Brenner, 2007). In this study, attribution could be problematic because we do not know the country of origin of the DDoS-attacks. Assuming the campaigns are effective, this implies that only when a significant portion of attacks originates from the targeted country, we can observe a significant reduction effect in attacks. Since it is difficult to determine the geolocation of individuals who hire DDoS services (beyond the location of the booter itself), it is difficult to determine whether potential cybercriminals are exposed to the campaigns. However, a recent unpublished analysis of data from the `stresser.gg` booter revealed that 48.1% of DDoS-attacks targeting the Netherlands originate in the Netherlands (Santanna, n.d.). While this is anecdotal evidence, if the percentage of attacks that are hired from one country to attack the same country is similar on other booters, such figures would reinforce the idea that the problem of attribution would be less relevant to territorially restricted online ad campaigns such as Google Ads.

The third assumption relates to the reliability of difference-in-differences models. There is an ongoing debate on how best to estimate these models and, as usual, it depends on the context (e.g., Huntington-Klein, 2022; Wing et al., 2018). We follow Huntington-Klein (2022) to estimate two-way fixed effect and long-term effect models, as well as to assess the parallel trends assumption with visual inspection and the test of *prior trends*. One of the advantages

of these models is that, by controlling for time, they account for seasonality. This is important in the case of outcome measures that are the result of human activity, such as crime in general and DDoS-attacks in particular, since they have a strong seasonal component (e.g. can be heavily influenced by holidays). We define our outcome variable following the Cambridge Cybercrime Centre’s formula (Collier et al., 2019; Thomas et al., 2017), although we are aware that Internet Service Providers (ISPs) have their own formulas for measuring DDoS-attacks. This type of cybercrime is highly variable and therefore difficult to predict. The presence of outliers—disproportionately high (or low) numbers of attacks recorded at a given time—makes DDoS-attacks challenging to model. To mitigate this issue, we checked the parallel trends assumption and estimated the models on a logged version of the DDoS, a common approach for highly varying outcome measures (Wooldridge, 2020).

Implementation and cost of the intervention

The EMMIE framework proposes five dimensions to inform stakeholders about the performance of crime prevention initiatives (Johnson et al., 2015). Its five dimensions are Effect, Mechanisms, Moderators, Implementation, and Economic costs. Above we already reported the size and direction of the effect of the campaigns on DDoS-attacks, as well as the possible mechanisms behind the effect, and the contexts that may have influenced the outcome (the EMM). We have yet to discuss the implementation of the campaigns and their cost (the IE).

The implementation of the campaigns was the result of a close collaboration between law enforcement and academia, with occasional input from industry partners, both online marketing experts, ISPs, and the gaming industry. The field knowledge of the Cyber Offender Prevention Squad of the Dutch National Police was key to targeting potential cybercriminals, as well as coordinating the implementation of the campaigns at national and international level. Academics contributed guidance on measures, design, and evaluation of the campaigns. In the first campaign, input from the gaming industry was essential to tailor the content of the ads and landing pages and improve their engagement. ISPs corroborated police suspicions regarding the usual targets of DDoS-attacks, which supported the proposed research design. Marketing experts were tasked with configuring the campaigns according to law enforcement and academia specifications, and deploy them. Although initially outsourcing certain tasks may be useful due to lack of expertise or time, in the long run it may facilitate the coordination and monitoring of the campaigns if law enforcement develops its own resources to devote to the campaigns. This would be especially relevant for long-term *influence policing* strategies. Given the growing momentum of the campaigns, and the duty of law enforcement to police cyberspace, global players such as Interpol and Europol can play a pivotal role in coordinating these campaigns on an international scale.

For a prevention campaign to be sustainable in the long-term, and obtain public support, it is important that it be cost-effective. Although we are not able to provide an estimate of the cost-effectiveness of the campaigns based on the cost of possibly averted cybercrime (if any), we can elaborate on their financial aspects. Rather than estimating the total cost

of a campaign, we suggest that the cost be measured based on the cost per click (CPC) of the ads. Ads are triggered based on specific keywords that multiple advertisers may bid on. Typical keyword categories include terms and actions around DDoS-attacks, botnets, booters and stressers (including active booter names), and gaming (see Appendix A in Moneva and colleagues (2022)). For example, the keyword “ddos attack” may be linked to preventive ads deployed by law enforcement, but also to commercial ads from cyber security companies. It is often the case that the more generic the keywords are, the stiffer the competition, which increases the CPC. It is possible that keywords in English are more expensive than those in other languages with fewer speakers because the former have the potential to reach more people (Moneva et al., 2022). CPCs for DDoS-attack-related keywords typically range from \$1-3. Advertisers should look for that optimal point between cost and engagement. It is also important to optimize the performance of campaigns over time based on these parameters. If the cost is too high, it may be wise to look for alternative keywords that are most specific to the target group. Google Ads offers the option to optimize campaigns automatically. We recommend that campaigns are supervised by human staff, as keyword selection is crucial to keep campaigns minimally intrusive (Collier et al., 2022), but automatic optimization may be an alternative if resources are scarce. Although the campaigns evaluated here were financed by the national law enforcement agencies of the countries involved, some countries also involved third parties to manage the campaigns. In these cases, it is advisable to anticipate the expense and set aside a budget to hire external companies.

Future research directions

Cambridge data is a valuable source that allows estimating the volume of DDoS-attacks suffered by countries with high temporal accuracy and the analysis of weekly and even daily cybercrime trends (see Thomas et al., 2017). This study evaluates weekly variations in the volume of DDoS-attacks over time and between countries, but future research could investigate whether the campaigns affect other outcome measures such as the duration of DDoS-attacks. Perhaps potential cybercriminals exposed to the ads will launch smaller attacks hoping to avoid detection. Since Cambridge data records IP addresses and prefixes, it would also be possible, in theory, to resolve them and identify different types of targets for DDoS-attacks, such as education, gaming industry, or government. Since the campaigns target potential cybercriminals, who in turn appear to be frequently targeting the gaming industry (Noroozian et al., 2016; Thomas et al., 2017), distinguishing between types targets would allow a more detailed analysis of the effect of the campaigns. Put another way, there could be an effect that we are not detecting.

Future research could also examine who is actually reached by the online ad campaigns. As pointed by Collier and colleagues (2022), “messages may be seen by the wrong people, or not seen by the right people” (p. 11), so it is important to minimize targeting inaccuracies. The primary target of the campaigns are potential cybercriminals, who tend to be young males interested in gaming and specific forms of cybercrime (Collier et al., 2019), and there is some

evidence suggesting that the campaigns do indeed reach this population (Moneva et al., 2022). Still, a detailed analysis of the socio-demographic characteristics of the users reached by the campaigns would help to identify more targeted and potentially more cost-effective keywords that would make the campaigns less intrusive and more sustainable. Google Ads and Google Trends data would be useful to examine this topic.

Finally, this study has not taken into account country-level factors in the analyses. It would be interesting to examine, for example, differences in compliance, cybercrime and cyber security awareness, gaming population, and involvement in cyber operations. For example, it is possible that the gaming community using booters in Portugal is small compared to that in the Netherlands, which would make it easier to reach a larger proportion of potential cyber-criminals in Portugal with the campaigns (not only through the ads themselves, but also, for example, through word of mouth), thus increasing their potential effectiveness. This evaluation of the effect of online ad campaigns against DDoS-attacks in Europe can provide some general insights for cybercrime prevention, but it remains crucial to take into account the unique characteristics and context of any region seeking to implement this strategy. Tailoring campaigns to the regional cybercrime landscape, cultural and linguistic factors, legal and regulatory frameworks, and collaboration and partnerships will likely increase their effectiveness.

Conclusion

This study has quantitatively and quasi-experimentally evaluated the effect of seven online advertising campaigns on the volume of DDoS-attacks recorded in six European countries. The results show mixed effects, so the hypothesis that the campaigns reduce DDoS-attacks in the short term did not receive clear empirical support. For the time being, this means that the use of such campaigns should not be justified solely on the basis of their effectiveness in reducing cybercrime, but also on other grounds such as their capacity to inform the population-at-risk about the illegality of DDoS-attacks and booter services. We need to improve our research designs and cybercrime measures to more accurately determine the effect of online ad campaigns on cybercrime.

References

- Abt, T. (2019). *Bleeding out: The devastating consequences of urban violence—and a bold new plan for peace in the streets* (First edition). Basic Books.
- Akers, R. L. & Jennings, W. G. (2015). *Social Learning Theory* (A. R. Piquero, Ed.; pp. 230–240). John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118512449.ch12>
- Bergé, L. (2018). *Efficient estimation of maximum likelihood models with multiple fixed-effects: The {r} package {FENmlm}*.

- Brenner, S. W. (2007). At light speed: Attribution and response to cybercrime/terrorism/warfare. *Journal of Criminal Law and Criminology*, 97(2), 379–475. <https://scholarlycommons.law.northwestern.edu/jclc/vol97/iss2/2/>
- Brewer, R., Vel-Palumbo, M. de, Hutchings, A., Holt, T., Goldsmith, A. & Maimon, D. (2019). *Cybercrime Prevention: Theory and Applications*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-31069-1>
- BritainThinks. (2020). *Review of cyber essentials influence on cyber security attitudes and behaviours in UK organisations* (pp. 1–6). <https://www.ncsc.gov.uk/information/setting-baseline-ce-prior-to-iasme>
- Clarke, R. V. (1980). “Situational” Crime Prevention: Theory and practice. *The British Journal of Criminology*, 20(2), 136147. <https://doi.org/10.1093/oxfordjournals.bjc.a047153>
- Clarke, R. V. (Ed.). (1997). *Situational crime prevention: successful case studies* (2nd ed.). Harrow; Heston.
- Clarke, R. V. (2017). *Situational crime prevention* (R. Wortley & M. Townsley, Eds.; 2nd ed., p. 125). Routledge, Taylor & Francis Group.
- Clarke, R. V. & Cornish, D. B. (1985). Modeling offenders’ decisions: A framework for research and policy. *Crime and Justice: A Review of Research*, 6, 147185.
- Coalition for Better Ads. (n.d.). *Determining a better ads standard based on user experience data*.
- Collier, B., Flynn, G., Stewart, J. & Thomas, D. R. (2022). Influence government: Exploring practices, ethics, and power in the use of targeted advertising by the UK state. *Big Data & Society*, 9(1), 205395172210787. <https://doi.org/10.1177/20539517221078756>
- Collier, B., Thomas, D. R., Clayton, R. & Hutchings, A. (2019). *IMC '19: ACM Internet Measurement Conference*. 50–64. <https://doi.org/10.1145/3355369.3355592>
- Collier, B., Thomas, D. R., Clayton, R., Hutchings, A. & Chua, Y. T. (2021). Influence, infrastructure, and recentering cybercrime policing: evaluating emerging approaches to online law enforcement through a market for cybercrime services. *Policing and Society*, 1–22. <https://doi.org/10.1080/10439463.2021.1883608>
- Dupont, B. (2019). The cyber-resilience of financial institutions: significance and applicability. *Journal of Cybersecurity*, 5(1), tyz013. <https://doi.org/10.1093/cybsec/tyz013>
- Europol & Dutch National Police. (2021). *No more ransom*. <https://www.nomoreransom.org/en/about-the-project.html>
- Federal Bureau of Investigation. (2020). *FBI cyber strategy*. <https://www.ic3.gov/Media/PDF/Y2020/PSA201008.pdf>
- Felson, M. & Eckert, M. (2018). *Introductory criminology: The study of risky situations* (1 Edition). Routledge, Taylor & Francis Group.
- Filiz, B., Arief, B., Cetin, O. & Hernandez-Castro, J. (2021). On the Effectiveness of Ransomware Decryption Tools. *Computers & Security*, 111, 102469. <https://doi.org/10.1016/j.cose.2021.102469>
- Google. (2023). *How to be successful with google ads*. <https://support.google.com/google-ads/answer/6080949>
- Hirschi, T. & Gottfredson, M. R. (1986). *The distinction between crime and criminality* (T. F. Hartnagel, R. A. Silverman, & G. Nettler, Eds.; p. 4469). Transaction Books.

- Holt, T. J. (2017). On the Value of Honey Pots to Produce Policy Recommendations: Sanction Threats on Online Behaviors. *Criminology & Public Policy*, 16(3), 739–747. <https://doi.org/10.1111/1745-9133.12315>
- Huntington-Klein, N. (2022). *The effect: An introduction to research design and causality*. CRC Press, Taylor & Francis Group.
- Interpol. (2022). *Awareness campaigns*. <https://www.interpol.int/Crimes/Cybercrime/Awareness-campaigns>
- Johnson, S. D., Tilley, N. & Bowers, K. J. (2015). Introducing EMMIE: an evidence rating scale to encourage mixed-method crime prevention synthesis reviews. *Journal of Experimental Criminology*, 11(3), 459–473. <https://doi.org/10.1007/s11292-015-9238-7>
- Kemp, S. (2023). Exploring public cybercrime prevention campaigns and victimization of businesses: A Bayesian model averaging approach. *Computers & Security*, 127, 103089. <https://doi.org/10.1016/j.cose.2022.103089>
- Kennedy, D. M. (2012). *Deterrence and Crime Prevention: Reconsidering the Prospect of Sanction* (1st ed.). Routledge. <https://doi.org/10.4324/9780203892022>
- Krebs, B. (2020). UK ad campaign seeks to deter cybercrime. *Krebs on Security*. <https://krebsonsecurity.com/2020/05/uk-ad-campaign-seeks-to-deter-cybercrime/>
- Maimon, D., Alper, M., Sobesto, B. & Cukier, M. (2014). Restrictive deterrent effects of a warning banner in an attacked computer system. *Criminology*, 52(1), 3359. <https://doi.org/10.1111/1745-9125.12028>
- Moneva, A., Leukfeldt, E. R. & Klijnssoon, W. (2022). Alerting consciences to reduce cybercrime: a quasi-experimental design using warning banners. *Journal of Experimental Criminology*. <https://doi.org/10.1007/s11292-022-09504-2>
- Nagin, D. S. (2013). Deterrence in the Twenty-First Century. *Crime and Justice*, 42(1), 199–263. <https://doi.org/10.1086/670398>
- National Assessments Centre. (2022). *Youth pathways into cyber crime in the UK* (pp. 1–8). <https://nationalcrimeagency.gov.uk/who-we-are/publications/596-nac-youth-pathways-into-cyber-crime/file>
- National Crime Agency. (2022). *Cyber choices: Helping you choose the right and legal path*. <https://nationalcrimeagency.gov.uk/what-we-do/crime-threats/cyber-crime/cyberchoices>
- National Cyber Crime Unit. (2017). *Pathways into cyber crime* (pp. 1–18). <https://www.nationalcrimeagency.gov.uk/who-we-are/publications/6-pathways-into-cyber-crime-1/file>
- Noroozian, A., Korczyński, M., Gañan, C. H., Makita, D., Yoshioka, K. & Eeten, M. van. (2016). *Who gets the boot? Analyzing victimization by DDoS-as-a-service* (F. Monrose, M. Dacier, G. Blanc, & J. Garcia-Alfaro, Eds.; Vol. 9854, pp. 368–389). Springer International Publishing. http://link.springer.com/10.1007/978-3-319-45719-2_17
- Pogarsky, G. & Herman, S. (2019). Nudging and the choice architecture of offending decisions. *Criminology & Public Policy*, 18(4), 823–839. <https://doi.org/10.1111/1745-9133.12470>
- Pogarsky, G., Roche, S. P. & Pickett, J. T. (2018). Offender Decision-Making in Criminology: Contributions from Behavioral Economics. *Annual Review of Criminology*, 1(1), 379–400. <https://doi.org/10.1146/annurev-criminol-032317-092036>

- Prichard, J., Wortley, R., Watters, P. A., Spiranovic, C., Hunn, C. & Krone, T. (2021). Effects of Automated Messages on Internet Users Attempting to Access “Barely Legal” Pornography. *Sexual Abuse*, 107906322110138. <https://doi.org/10.1177/10790632211013809>
- R Core Team. (2022). *R: A language and environment for statistical computing*. <https://www.R-project.org/>
- Roach, J., Weir, K., Phillips, P., Gaskell, K. & Walton, M. (2017). Nudging down theft from insecure vehicles. A pilot study. *International Journal of Police Science & Management*, 19(1), 31–38. <https://doi.org/10.1177/1461355716677876>
- RStudio Team. (2020). *RStudio: Integrated development environment for r*. <http://www.rstudio.com/>
- Rudis, B., Keyes, O. & Smith, T. (2021). *Iptools: Manipulate, validate and resolve 'IP' addresses*. <https://CRAN.R-project.org/package=iptools>
- Santanna, J. J. (n.d.). *Analysis of stresser.gg booter website leaked database*.
- Schiks, J. A. M., van 't Hoff-de Goede, M. S. & Leukfeldt, E. R. (2021). *Een alternatief voor jeugdige hackers? Plan- en procesevaluatie van Hack_Right* (pp. 1–170). <https://www.politieenwetenschap.nl/publicatie/politiewetenschap/2021/een-alternatief-voor-jeugdige-hackers-361/>
- Stockman, M., Heile, R. & Rein, A. (2015). *An Open-Source HoneyNet System to Study System Banner Message Effects on Hackers*. 1922. <https://doi.org/10.1145/2808062.2808069>
- Thaler, R. H. & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness* (Rev. and expanded ed). Penguin Books.
- Thomas, D. R., Clayton, R. & Beresford, A. R. (2017). *2017 APWG symposium on electronic crime research (eCrime)*. 79–84. <https://doi.org/10.1109/ECRIME.2017.7945057>
- Vetterl, A. (2020). *Honeypots in the age of universal attacks and the internet of things* (pp. 1–115). <https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-944.pdf>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). *Welcome to the {tidyverse}*. 4, 1686. <https://doi.org/10.21105/joss.01686>
- Wilson, T., Maimon, D., Sobesto, B. & Cukier, M. (2015). The Effect of a Surveillance Banner in an Attacked Computer System: Additional Evidence for the Relevance of Restrictive Deterrence in Cyberspace. *Journal of Research in Crime and Delinquency*, 52(6), 829–855. <https://doi.org/10.1177/0022427815587761>
- Wing, C., Simon, K. & Bello-Gomez, R. A. (2018). Designing Difference in Difference Studies: Best Practices for Public Health Policy Research. *Annual Review of Public Health*, 39(1), 453–469. <https://doi.org/10.1146/annurev-publhealth-040617-013507>
- Wooldridge, J. M. (2020). *Introductory econometrics: A modern approach* (Seventh edition). Cengage Learning.
- Wortley, R. & Townsley, M. (2017). *Environmental criminology and crime analysis: Situating the theory, analytic approach and application* (R. Wortley & M. Townsley, Eds.; 2nd ed., p. 125). Routledge, Taylor & Francis Group.

Acknowledgements

We would like to thank Richard Clayton, Cambridge Cybercrime Center, for facilitating the data for the study and helping us to understand it. Also for sharing with us, together with Ben Collier, from The University of Edinburgh, their thoughts on the early stages of the research design and commenting on an earlier draft of the manuscript. We would also like to thank all the members of the Cyber Offender Prevention Squad, High Tech Crime Team, Netherlands Police—and especially Wouter Klijnsoorn—for their invaluable and diligent liaison work between academics and police during the design and implementation of the campaigns.

Authors' bio

Asier Moneva is a Postdoctoral Researcher at the Netherlands Institute for the Study of Crime and Law Enforcement (NSCR) and at the Center of Expertise Cyber Security of The Hague University of Applied Sciences. Asier is interested in where, when, and how cyber offenders commit cybercrime. With his research, he aims to generate knowledge to better understand cybercrime, and to find solutions to reduce it or mitigate its impact.

Rutger Leukfeldt is a Senior Researcher at the Netherlands Institute for the Study of Crime and Law Enforcement (NSCR) and the Director of the Center of Expertise Cyber Security of The Hague University of Applied Sciences. Rutger has conducted research into the human factor of cybercrime for 15 years. During that period, he has been involved in both fundamental academic research and applied research for companies and governments. Rutger carries out both quantitative and qualitative studies, but his expertise lies in qualitative methods. Over the years, he analyzed numerous large scale police investigation and interviewed both cybercriminals and victims.

Appendix A. Support for the parallel trends assumption

The two-way fixed effects models in the fake treatment periods indicate that there were no statistically significant differences ($\alpha = 0.05$) between the countries in the treated and untreated groups prior to treatment (Table 4), suggesting that the parallel trends assumption is plausible and that the comparison groups are therefore appropriate. Figure 8 plots the trends of DDoS-attacks before the treatment. Overall, the trends increase and decrease in parallel over time, which further suggests that the parallel trends assumption is plausible.

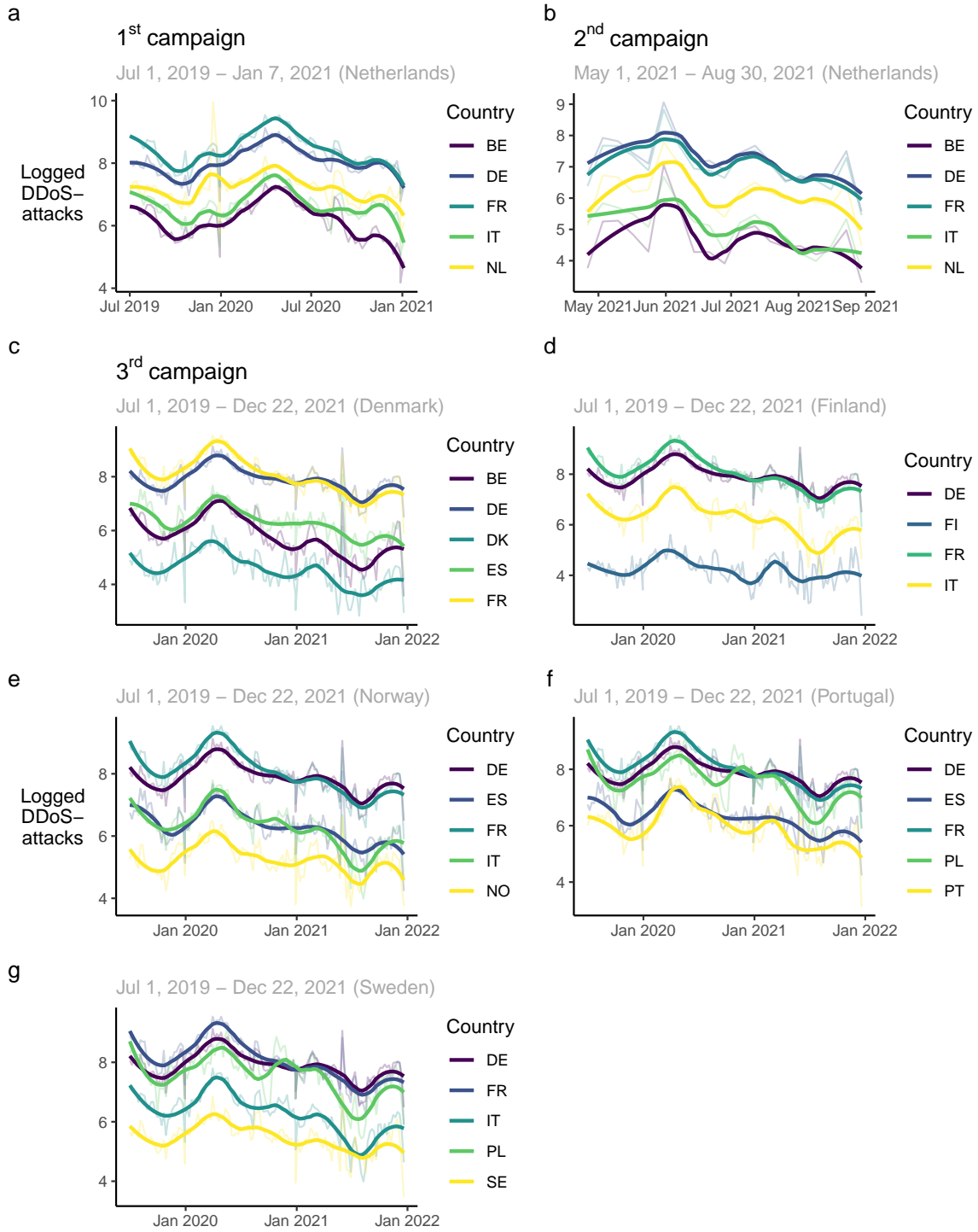


Figure 8: Parallel trends per country prior to the online ad campaigns

Table 4: Difference-in-differences (two-way fixed effects) estimates using a fake treatment period

Treatment	Estimate	Std. Error	p-value	Sig.	Num. Obs.	Adj. R-squared
1st campaign						
Netherlands	0.028	0.150	0.861		410	0.932
2nd campaign						
Netherlands	0.130	0.091	0.228		95	0.971
3rd campaign						
Denmark	0.025	0.117	0.838		660	0.963
Finland	0.350	0.156	0.111		528	0.966
Norway	0.307	0.156	0.145		528	0.962
Portugal	-0.204	0.119	0.160		660	0.916
Sweden	0.228	0.111	0.110		660	0.942
FE: Week						
FE: Country						

Table 5: Original ad texts of the first and second campaigns

Crime	Headline	Description
1st and 2nd campaign (Netherlands)		
DDoS	Een DDoS-aanval is strafbaar	Het plegen van een DDoS is strafbaar in Nederland onder het Wetboek van Strafrecht.
DDoS	Play fair; verliezers DDoS-en	Win eerlijk in een e-sport competitie door je gaming skills te trainen.
DDoS	DDoS verpest het voor iedereen	Dus je wilt dat je vrienden niet meer kunnen gamen omdat jij een grapje uithaalt?
DDoS	Wil je een game DDoS-en?	Leer meer over DDoS-aanvallen en de impact hiervan op gamen.

Appendix B. Original ad texts

Appendix C. Landing page design for the third online ad campaign

Hé jij white-hat! Interesse in de digitale wereld? Goed bezig. Zorg er alleen wel voor dat je de online grenzen kent ;-)



Als je het schoolsysteem hackt, kunnen er strengere regels komen op school. Hierdoor hebben jij en je vrienden minder vrijheid.

Illegaal hacken is cybercrime en strafbaar in Nederland onder artikel 138ab in het Wetboek van Strafrecht.

Sam weet het wachtwoord van de social media account van zijn vriend. Hij gebruikt dit om in het social media account van zijn vriend te hacken, omdat hij zijn berichten wil lezen.

De politie treedt hard op tegen cybercrime; in welke vorm en waar dan ook. Pleeg je cybercrime? Houd dan rekening met de gevolgen:




	Veroordeling
Je krijgt een boete, taakstraf of een gevangenisstraf.	
	Strafblad
Je krijgt een strafblad. Dit maakt het vinden van een stage of baan lastig, maar ook het reizen naar andere landen.	
	Tools
Je raakt je computer en/of telefoon kwijt.	



This is Bill
Bill is on the internet
Bill checks responsible disclosures before hacking
Bill is smart
Be like Bill

Schade

Daarnaast kan een DDoS serieuze schade toebrengen aan mensen en bedrijven:

		
Jouw game-account wordt geband	Telecomproviders zijn offline	112 is onbereikbaar

Meld hier een hack anoniem

[\(\(aangifte-of-melding-doen/ik-wil-anoniem-iets-melden.html\)\)](#)

Wil je écht winnen?

Heb je interesse in IT en wil je nu direct iets nieuws leren? Check de onderstaande initiatieven!

 <p>#CRIMEDIGGERS TEST JE DIGITALE SKILLS https://www.crimediggers.nl</p>	 <p>Gamechangers https://publicaties.politie.nl/changeyourgame/</p>	 <p>DIVD https://www.divd.nl/</p>
---	--	--

Samenwerking

De bestrijding van cybercrime doen we als politie niet alleen. In deze strijd werken wij samen met Europol en partners.



Figure 9: Landing page for the hacking online ads of the third campaign

Table 6: Original ad texts of the third campaign

Crime	Headline	Description
3rd campaign (Denmark)		
DDoS	DDoS ødelægger det for alle	Det er strafbart i Danmark at udføre et Ddos angreb
Hacking	Ulovlig hacking er indbrud	Du vil få en bøde, samfundstjeneste eller en fængselstraf
RAT	Brug af RAT tools er tyveri	Du vil få en plettet straffeattest. Det vil gøre det sværere at finde elevplads eller job
3rd campaign (Finland)		
DDoS	DDoS pilaa kaikkien ilon	Palvelunestohyökkäyksen tekeminen on Suomen lakien mukaan rangaistava teko
Hacking	Laiton hakkerointi on rikos	Laiton hakkerointi on Suomen lain mukaan rangaistavaa
RAT	RAT = kyberrikos = ei vitsi	Haittaohjelman käyttö on Suomen lain mukaan rangaistavaa
3rd campaign (Netherlands)		
DDoS	DDoS verpest het voor iedereen	DDoS-aanvallen veroorzaken ernstige schade aan organisaties, bedrijven en individuen
Hacking	Illegaal hacken is inbreken	Illegaal hacken kan ernstige schade toebrengen aan organisaties, bedrijven en individuen
RAT	Een RAT inzetten is inbraak	De politie pakt cybercriminaliteit aan. In welke vorm en waar dan ook.
3rd campaign (Norway)		
DDoS	ddos ødelegger for alle	DDoS-angrep er straffbart i Norge etter straffeloven, og kan medføre fengselsstraff
Hacking	Hacking er datainnbrudd	Ulovlig hacking (datainnbrudd) er straffbart i Norge, og kan medføre fengselsstraff
RAT	Bruk av RAT er et datainnbrudd	Politiet bekjemper alle former for datakriminalitet
3rd campaign (Portugal)		
DDoS	DDoS prejudica toda a gente	Ficas com um registo criminal e vai ser mais difícil conseguires um estágio ou um emprego.
Hacking	Hacking é ilegal	A Polícia investiga qualquer tipo de cibercrime sob qualquer forma
RAT	Instalar um RAT é como roubar	A Polícia investiga qualquer tipo de cibercrime sob qualquer forma!
3rd campaign (Sweden)		
DDoS	DDoS förstör för alla	DDoS-attacker är straffbara enligt svensk lag
Hacking	Olovlig hacking = dataintrång	Att hacka någon en enda gång kan göra att du hamnar i straffregistret
RAT	Använda en RAT = dataintrång	Så du vill hindra dina kompisar från att spela på grund av ditt "skämt"?