3-2019

See no evil, hear no evil? Dissecting the impact of online hacker forums

Wei T. YUE
City University of Hong Kong

Qiu-Hong WANG
Singapore Management University, qiuhongwang@smu.edu.sg

Kai-Lung HUI
Hong Kong University of Science and Technology

See No Evil, Hear No Evil?

Dissecting the Impact of Online Hacker Forums

Wei T. Yue, Qiu-Hong Wang, and Kai-Lung Hui

Forthcoming in MIS Quarterly

Abstract

Online hacker forums offer a prominent avenue for sharing hacking knowledge. Using a field dataset culled from multiple sources, we find that online discussion of distributed denial of service (DDOS) attacks in hackforums.net decreases the number of DDOS-attack victims. A 1% increase in discussion decreases DDOS attacks by 0.032%–0.122%. This means that two DDOS-attack posts per day could reduce the number of victims by 700–2,600 per day. We find that discussion topics with similar keywords can variously increase or decrease DDOS attacks, meaning we cannot ascertain the impact of the discussion just by the post nature. Mentioning botnets, especially new botnets, increases the attacks, but the follow-up discussion decreases the attacks. Our results suggest that online-hacker-forum discussion may exhibit the *dual-use* characteristic. That is, it can be used for both good and bad purposes. We draw related managerial implications.

Keywords: Hacker forum, distributed denial of service attack, backscatter data, dual use, panel regression, content analysis.

Author names are list in reverse alphabetical order. Yue: College of Business, City University of Hong Kong. Wang: School of Information Systems, Singapore Management University. Hui: School of Business and Management, Hong Kong University of Science and Technology. Corresponding author: Qiu-Hong Wang (qiuhongwang@smu.edu.sg). We thank the seminar participants at HEC Paris, University of Texas at Dallas, Vytautas Magnus University, Warwick University and ESSEC Business School (Asia Pacific campus), for their helpful comments. We also thank Johannes Ullrich of the SANS Technology Institute for providing the backscatter and DShield data needed for this study. We thank Wlad Fomin, Wilson Weixun Li, Jiali Zhou, and their teams for the helpful research assistance. This research is supported in part by the Hong Kong SAR General Research Fund projects 16500715 and 152611, the Singapore National Research Foundation, NRF-TAU Collaboration Grant NRF2016NCR-NCR001-009, and the Singapore MOE Tier-1 Grant, 16-C220-SMU-002.

1. Introduction

The Internet brings unprecedented impacts to the society. One noteworthy change is the ease with which individuals share and discuss sensitive topics in online channels, including crime-related knowledge such as how to attack other people or new attack tools that can increase victims' damage. Such sharing and discussion may affect information security. In particular, Imperva, a cybersecurity-solution provider, has argued that hacker forums serve as a convenient venue for hackers to share hacking knowledge and collaborate on attacks. They suggest that hacker forums have become "the cornerstone of hacking":

> They are used by hackers for training, communications, collaboration, recruitment, commerce and even social interaction. Forums contain tutorials to help curious neophytes mature their skills. Chat rooms are filled with technical subjects ranging from advice on attack planning and solicitations for help with specific campaigns. Commercially, forums are a marketplace for selling of stolen data and attack software. (Imperva 2011, page 1)

However, there is an important difference between hacking and physical crimes. Because hacking involves using computing devices and networks to launch an attack, the hacker must acquire the related computing knowledge. Such knowledge, however, may also help potential victims defend against the attack. For example, the discussion of how to penetrate a firewall can help security managers improve firewall configuration. The spread of botnet data may help law enforcement agencies trace and neutralize the botnets. This is different from the knowledge on certain physical crimes, such as how to set off a bomb or spread a deadly virus, which inevitably contributes to damage and offers little benefit.

Accordingly, hacking tools and knowledge exhibit the *dual-use* characteristic (Katyal 2001) and can be used for both good and bad purposes. Because of dual use, it is unclear whether we should take action against the sharing and discussion of hacking knowledge. On one hand, such discussion may expose more people to hacking and hence promote aggression. It may also help like-minded hackers collaborate on attacking other people. On the other hand, hacking discussion may contribute to developing and spreading protection knowledge. Understanding hacker assets in online forums may educate users about their functions and characteristics (Samtani et al. 2015). Open discussion of hacking may remove its novelty for unskilled or amateur hackers such as *script kiddies*. It may also contribute to establishing a proper social norm, which could be one practical means of curbing cybercrimes (Katyal 2001). With these opposing influences, the net impact of hacking discussion on cyber attacks is an intriguing empirical question.

Here, using a unique dataset culled from multiple sources, we study the impact of online-hacker-forum discussion on the extent of distributed denial of service (DDOS) attacks, which is one of the most popular cyber attacks on the Internet. DDOS attacks cripple online services by flooding the servers with dummy requests. It affects many global enterprises, with some suffering revenue losses exceeding one million dollars per hour (Neustar 2017). The threat of DDOS attacks has reached an unprecedented scale due to rapid growth of unsecure devices on the Internet (Constantin 2016). However, most knowledge and tools related to DDOS attacks carry the dual-use characteristic, making it very difficult to prevent and deter. For example, firms often perform penetration and stress tests that use port scanning and traffic generators, both being commonly used for launching DDOS attacks. We focus on hacker forums because it is the major channel for hacking discussion on the Internet (Imperva 2011).

We compiled DDOS-attack discussions from hackforums.net, one of the most visited hacker forums on the Internet. Because all DDOS attacks target specific ports associated with different software applications, we connect the forum discussion to the DDOS attacks observed from 2007 to 2011 via the port numbers mentioned in the discussion. We identify the forum-discussion effect by regressing the number of DDOS attacks on the scattered forum posts over time and across the ports. We supplement this identification strategy with an instrumental-variable estimation and several validation and falsification exercises.

We find that discussion in hackforums.net generally *decreases* DDOS attacks. A 1% increase in DDOS-attack posts decreases the number of DDOS-attack victims by 0.032% to 0.122%. The size of this effect is economically significant as it implies two posts per day would reduce the number of DDOS-attack victims by 700 to 2,600 per day. Discussion in antichat.ru, a prominent Russian forum, also decreases DDOS attacks, but its effect is considerably smaller. Discussion in other hacker forums is not statistically correlated with DDOS attacks.

We buttress our estimation with several empirical strategies and find that our results are robust to the exclusion of outliers and variations in model specifications. We then scrutinize the content of the discussion. We find that topics with overlapping DDOS-attack keywords could have opposite influences on actual DDOS attacks. This seems consistent with the dual-use theory, which suggests that similar content or tools can have both good and bad impacts depending on the context. Nevertheless, the content analysis points to one interesting mechanism. Mentioning botnets, particularly new botnets, increases the number of DDOS attacks, but the follow-up discussion has an opposite effect: It tends to decrease the attacks.

This study makes three important contributions. First, it shows that encouraging more discussion need not be bad when hacking knowledge and discussion is openly accessible on the Internet. It provides alternative evidence countering recent findings that focus on the adverse consequences of online information exchange and the Internet (see, e.g., Kaplan and Moss 2003; Hunton 2009; Banks 2010; Chan and Ghose 2014; Chan et al. 2016).

Second, it highlights an intriguing challenge to regulating dual-use technologies. The knowledge and tools around DDOS attacks can be put into both good and bad uses, as reflected in our hacker-forum-post analysis. Although most posts are ostensibly malicious, developing the discussion actually led to fewer DDOS attacks. Our study suggests that we need more-focused identification strategies in studying the empirical impacts of dual-use technologies.

Third, this study provides novel evidence on the mechanism that underlies the discussion's impact. In particular, popular keywords may not help us predict its influence. Instead, the sequence matters – first mentioning an attack increases the number of attacks observed, but subsequent discussion decreases attacks. This finding contributes an important new perspective to public policy: We should pay closer attention to the development of public discussion instead of focusing on disclosure of malicious information per se.

The rest of this paper is organized as follows. Section 2 reviews related literature. Section 3 describes our setting and data. Section 4 presents the empirical model. Section 5 reports the results including the robustness and falsification tests and content analysis. Section 6 discusses the implications of this research and concludes the paper.

2. Related Literature

This study is related to the growing stream of research on hacker behavior. In an early work, Jordan and Taylor (1998) suggest that, similar to the computer-security community, the online hacker community may potentially enhance system protection through hacking. Hackers are interested in learning about computing technologies (O'Neil 2006; Auray and Kaminsky 2007). They perceive themselves as positive deviants who follow the greater cause of rectifying injustice (Olson 2012; Coleman 2013; Steinmetz and Gerber 2014) and whose expertise empowers them to challenge social conventions (Turgeman-Goldschmidt 2008).[1]

Recent research, however, has found sinister behaviors in online channels such as forums, chat rooms and social media. Holt and Lampkeb (2010) find that some people use online forums to trade stolen financial data. By scrutinizing the transactions of hacking tools in online forums, Holt (2012) finds that the hacker community supports cybercrimes. Such findings underscore the importance of identifying potential threats from the online hacker community. Benjamin et al. (2015) develop an automated content-analysis methodology that can detect the emerging threats from hacker forums, Internet relay-chat channels, and carding shops. Benjamin et al. (2016) develop an approach that can identify key cyber criminals based on social-network analytics. Using content-analysis techniques, Abbasi et al. (2014) identify and characterize expert hackers who may pose threats to society. Instead of scrutinizing specific hacker behavior and drawing inferences on their impacts from community activities per se, this study connects online hacker activities to real-world events.

[1] For a detailed discussion of the characteristics of highly skilled malware writers and hackers in an underground hacker social-networking group, please refer to Holt (2012).

With the proliferation of electronic commerce and social media, the impacts of online channels on offline outcomes have received great attention. For example, Godes and Mayzlin (2004) find that the dispersion of discussion across different Usenet forums can help predict new television programs. Antweiler and Frank (2004) show that the discussion in online message boards can help predict stock volatility. Chen et al. (2014) also find that peer opinions in social media help predict stock returns. Geva et al. (2015) find that online forum data and Google search data complement social media data to predict automotive sales. Rui et al. (2013) find that online word of mouth affects the box-office revenues of movies.

However, other studies have also found negative consequences of the Internet. Bhuller et al. (2013) find that broadband Internet penetration has promoted sex crimes, possibly due to easier access to pornography. Chan and Ghose (2014) find that the introduction of Craigslist has facilitated HIV transmission because of nonmarket casual hookups (in contrast to paid sexual transactions). Chan et al. (2016) find evidence that broadband Internet access leads to more racial hate crimes. The use of social media may also correlate with suicide (Dunlop et al. 2011; Luxton et al. 2012).

In general, this literature suggests that the activities in online channels tend to have the expected impacts – stock and movie promotion can increase stock returns and movie sales. Easier access to sex may increase sex crimes and HIV transmissions. The impacts of the Internet on other social phenomena may be more nuanced. For example, the proliferation of the Internet may decrease offline social participation but increase online social participation (Bauernschuster et al. 2014). The availability of online content should encourage a wider exposure to different contents, but increased customizability of online content could also lead to selective exposure – the so-

called *echo chamber* effect (Hosanagar et al. 2014; Flaxman et al. 2016). In situations like these, where theoretical analysis does not give unequivocal guidance, we must seek empirical insights. This is especially the case for hacker forum discussion because of the dual-use nature and moral ambiguity of hacking (Thomas 2005).

Ascertaining the impact of hacker-forum discussion is important because it informs public policy about the need for intervention. Prior research has considered regulating selected Internet activities. For example, prosecuting online transactions of dangerous exploits may keep the exploits from creating damage before security developers can find a solution (Stockton and Golabek-Goldman 2013). Subject to a similar set of law and regulation that govern newspaper and television, restricting the supply of harmful information online should help curb cyber attacks (Neumann 2013). For these regulations to work, we need a clear orientation of the online activities, viz. whether they increase or decrease the harm on other people. It is not easy to determine such orientations for online hacking discussion.

Accordingly, this study establishes the net empirical impact of hacker-forum discussion. Similar to the literature reviewed above, we exploit the rich discussion data in a representative hacker forum over five years. The forum contains millions of posts and comprises visitors from major economies in the world. We match its discussion to worldwide DDOS-attack data obtained from another source independent of the forum. Hence, we utilize the granular forum discussion data and the massive real-world cyberattack data to estimate the net impact of online-hacker-forum discussion. This impact is nontrivial because of the dual-use characteristic.

3. The Data

We compiled our data from multiple sources. To measure the extent of DDOS attacks over time, we obtained backscatter data from the Internet Storm Center (ISC) of the SANS Institute. The ISC maintains a worldwide collection of network security sensor logs from its voluntary Internet subscriber base. These sensors report abnormal traffic to the ISC. Hence, they provide a good and comprehensive overview of all malicious activities on the Internet.

The backscatter data record malicious attacks generating SYN-ACK packets in the ISC's sensor networks. In a SYN-ACK DDOS attack, the attacker exploits transmission control protocol's (TCP's) three-way handshake process and floods a victim with SYN packets from forged senders. The victim responds to each of these SYN requests with a SYN-ACK packet – the *backscatter* packet – and then waits for the forged senders' final confirmations. These confirmations will never come, however, which causes the victim's system to create open sessions. With too many open sessions, the victim will have fewer resources for legitimate requests.

The SYN attack and backscatter packets go through a certain port in the victim's computer system. The ISC aggregates these backscatter packets by port and counts the number of unique source Internet Protocol (IP) addresses, corresponding to DDOS-attack *victims*, on a daily basis.[2] The use of backscatter data to study cyber attacks is common in the literature (see, e.g., Moore et al. 2006; Zhang and Parashar 2010; Kim et al. 2012; Hui et al. 2017).

The port number is a good variable for linking DDOS attacks to forum discussion. Open ports provide an access point to a victim's computer system. Probing open ports and exploiting

[2] Details of the ISC and backscatter data are available at https://isc.sans.edu/ [accessed November 20, 2017]. The destination IP address in the backscatter data are mostly forged. Hence, they are not usable for our purposes.

the vulnerabilities of Internet applications using those open ports are common preliminary actions before hackers launch cyberattacks (Panjwani et al. 2005). We choose port and day as the units of analysis because the ISC groups the backscatter data by port and day and because we cannot specifically associate each observed DDOS attack to posts in the forum discussion. The port number ranges from 0 to 65,535. We have a total of 1,826 days of data in our sample.

Because port usage varies by Internet application and some ports, such as 80 and 21, are used more often than other ports, it is important to control for the frequency of attacks on the ports.[3] We compiled the number of vulnerabilities associated with each port over time from the National Vulnerability Database and Open Source Vulnerability Database. We also downloaded the number of threats and risks associated with each port over time from Symantec's Enterprise Security Response Unit. Vulnerabilities and threats affect the ease of compromising a computer. Hence, the number of vulnerabilities and number of threats may correlate with the extent to which a port is attacked, making them pertinent control variables.

For the main analysis, we obtained the data from hackforums.net (Hackforums), which is one of the largest English forums dedicated to hacking discussion on the Internet. Hackforums ranked third in the *Hacking* subcategory and first in the *Chats and Forums* subcategory under the *Hacking* subcategory in Alexa.com (Alexa).[4] The Anti-Security Movement (Anti-Sec) recognizes Hackforums as being "*notable within the hacking underground and the computer security world*" and "*one of the largest communities of hackers and script-kiddies alike currently at large in cyber*

[3] For example, most Web traffic goes through port 80 or 8080. Most email services use port 25 (SMTP), 110 (POP), 143 (IMAP), 465 (SSL/TLS encrypted SMTP), or 993 (SSL/TLS encrypted IMAP).

[4] Alexa classifies websites into 17 categories. Hacking is one of the subcategories under *Computers.* For more details, please refer to http://www.alexa.com/topsites/category [accessed January 16, 2017].

space." (Anti-Sec 2009). Users need to seek approval from an administrator to create an account and must log in to view and post messages on Hackforums.

The discussion in Hackforums was not active until 2007. For this study, we downloaded all posts in the *hacking* section of Hackforums from 2007 to 2011 (total 1,826 days), comprising 2,960,893 posts distributed across 23 subforums and 355,222 threads. With these posts, we conducted multiple rounds of text extraction and verification to identify the posts discussing DDOS attacks and the corresponding port numbers. We further scrutinized the DDOS-attack posts using various text-mining techniques to explore the content of the discussion in Hackforums. We report the details of text extraction and processing in Sections 3.1–3.3.

To assess the boundary of our findings, we collected additional discussion from another prominent English hacker forum, Hellboundhackers.org (HBH), the popular Chinese hacker forums hackbase.com (Hackbase) and 2cto.com (referred as HHLM from its Chinese acronym), and the popular Russian hacker forums antichat.ru (Antichat) and xaker.name (Xaker). Table 1 presents the ranking and the total numbers of posts, threads, and subforums in each of the six forums in 2007–2011. Although these forums do not have the highest ranks in the *hacking* categories in Alexa, we select them because the other higher-ranked forums are not focused on hacking or were started much later and hence do not cover our data window, 2007–2011. Table 2 presents the distribution of forum visitors. Evidently, Hackforums has more diverse visitors. The Chinese forums have the most concentrated visitors from China.

[Insert Tables 1 and 2 here]

3.1. Port and DDOS Post Extraction

As is evident in Table 1, the forums contain millions of posts. It is practically infeasible for us to read all of these posts manually. Accordingly, we conducted multiple rounds of text extraction supplemented by manual screening to identify posts mentioning a port or DDOS attacks. We report the detailed procedures and statistics in the online appendix.

In particular, we followed three steps to identify port numbers. First, we removed posts containing irrelevant numbers such as date or IP address. Second, we separated the remaining posts into two sets, the candidate set and the irrelevant set. The candidate set contains all posts that either have the keyword *port* and a number, or other keywords related to common protocols and the corresponding port numbers (e.g., *TCP* with port 80, *telnet* with port 23, *SMTP* with port 25, etc.). Third, two research assistants (RA) independently read all posts in the candidate set to confirm whether they indeed contain a port number. The RAs then compared their results to resolve any inconsistency in the screening.[5]

To test the performance of our procedure, we generated three test samples for each forum. The RAs read *all* posts in these test samples to establish a benchmark. We then applied the three steps above to each test sample. The results show that the recall rates, defined as the fraction of extracted posts mentioning a port over all posts mentioning a port, mostly exceed 90% after the second step.[6] We provide the details of this assessment and the full results of the

[5] Because identifying a port number does not require any subjective or strong judgment, we asked the two RAs to discuss and resolve any inconsistencies in the independent screening. Most inconsistencies arose because of human errors, such as typos or overlooking a port number. We engaged different RAs who are familiar with English, Chinese, and Russian to process the corresponding posts.

[6] The key purpose of this assessment is to estimate the extent to which our procedure would miss posts mentioning a port in the second step when we classify some posts as irrelevant without further screening. The third step does not apply here because the RAs read all posts in the test samples. We randomly selected 1,000 threads in each test sample for each forum except HBH, which had relatively little discussion. We randomly selected 500 threads

test sample screening in the online appendix, Section A1. In view of the high recall rates and the significant saving in labor (the first two steps help us remove more than 90% of the posts; the third step of manually screening further helps us remove 50%–80% of the candidate posts), we applied the same procedure to process all forum posts. The fourth column in Table 1 reports the number of extracted posts mentioning a port in each forum.

Next, we followed a four-step procedure to identify discussions of DDOS attacks. First, we obtained a large number of articles from the Internet related to DDOS attacks, such as the techniques and tools involved. Second, we removed common stop words such as *the*, *is*, *at*, and *on* (and similar stop words for posts of other languages) from these articles and ranked their keywords by frequency. Third, we separated the posts into two sets, the candidate set and the irrelevant set. The candidate set contained all posts that have a high score in terms of DDOS-attack keyword ranks and frequencies. Fourth, two RAs independently read all posts in the candidate set to decide whether they were indeed discussing DDOS attacks. We repeated the first three steps multiple times to fine-tune the keyword lists.

Similar to the port-number extraction, we evaluated the accuracy of our DDOS-attack post extraction using three test samples for each forum. The RAs read all posts in these test samples. We then applied the four-step procedure described above and crosschecked the results with the manual screening. The results show that the recall rates, defined as the fraction of extracted

in each HBH test sample. The total number of posts used in this assessment varies across the test samples and forums because the sampled threads contain different numbers of posts.

DDOS-attack posts over all DDOS-attack posts, mostly exceed 90% after the third step.[7] We provide the details of this assessment in the online appendix, Section A2.

Because we use the port number to connect forum discussions with the observed DDOS attacks, we extracted DDOS-attack posts only from all *threads* that contain a port number in at least one of their posts. We extracted DDOS-attack posts from the entire thread instead of specific posts mentioning the port numbers because DDOS-attack discussion may span multiple posts, but not all of these posts mention a port number.[8] The last column in Table 1 reports the number of DDOS-attack posts in each forum. Overall, the keyword extraction in the first three steps helps us remove 60%–90% of irrelevant posts across the different forums. The fourth step of manual screening further helps us remove 40%–90% of the candidate posts.

In our main analysis, we measure port-related DDOS-attack discussion by counting the number of posts that mentioned a port or replied to an earlier post that mentioned a port in a thread that contains at least one DDOS-attack post. We call them *DDOS-thread–port-effective* posts.[9] We report robustness tests using other measures in the online appendix. Figures 1 and 2 plot the daily average numbers of DDOS-attack victims and DDOS-thread–port-effective posts in 2007–2011 across forums and the five most commonly discussed ports, 80 (HTTP), 21 (FTP), 82 (xB browser), 8080 (alternative HTTP), and 443 (TLS/SSL) in Hackforums.[10] Evidently, the DDOS

[7] Here again, the fourth step does not apply because the RAs read all posts in the text samples. We randomly selected 1,000 threads in each test sample for each forum except HBH, which had relatively little discussion. We randomly selected 500 threads in each HBH test sample.

[8] As we shall see in Section 4, the effect of DDOS-attack discussion in other threads without a port number is captured by the day fixed effects in the empirical model. Hence, it will not affect the significance of our estimates of the port-related DDOS-attack discussion effect.

[9] Hereafter, we use the convention "*X* effective" to refer to all posts that either mentioned *X* or replied to an earlier post that mentioned *X*, and "*Y* thread" to refer to all posts in a thread that contains a post mentioning *Y*.

[10] The brackets contain common protocols or Internet applications using the corresponding ports.

attack and forum discussion often trend in the opposite direction especially when they are connected by port number. Figures 1 and 2 present model-free evidence that hacker-forum discussion might be negatively correlated with the observed DDOS attacks.

[Insert Figures 1 and 2 here]

Note that the magnitude of the DDOS-thread–port-effective posts in Figures 1 and 2 may seem disproportionately large when compared with the total number of extracted posts reported in Table 1. This is because we count the *effective* posts by including all the follow-ups to the original posts mentioning the port number. Furthermore, a post can mention multiple port numbers. Because we organize the data by port, we count a post multiple times if it mentions more than one port.

3.2. Content Analysis

To gain a deeper understanding on the content discussed in the port-related DDOS-attack posts, we conducted two sets of unsupervised and supervised text processing. In the first analysis, we applied the latent Dirichlet allocation (LDA) method, an unsupervised modeling technique, to explore the topics discussed in the DDOS-thread–port-effective posts extracted in Section 3.1.[11] To ensure robustness, we repeated the LDA analysis by generating different sets of topics and

[11] The LDA method models each document as a finite mixture of latent topics, with each topic being a mixture of keywords with some probability distribution (Blei et al. 2003). Because we do not know the topics, we cannot use any ground truth to assess a LDA model. Hence, it extracts different topics and keyword distributions depending on the number of topics specified by the researcher. We use the port-effective DDOS thread as the unit of a "document" in the LDA analysis. It is more likely to extract meaningful topics from an elaborate discussion in a thread of posts instead of individual posts, which tend to be too granular and often contain incomplete discussion.

testing whether these discussion topics correlate with the observed DDOS attacks. We report the detailed LDA modeling results and topic keywords in the online appendix, Section A3.

Furthermore, DDOS attacks often involve using coordinated compromised computers (the *botnet*). As reported in Section 5.3, the LDA modeling results indicate that a botnet is indeed a conspicuous discussion topic in Hackforums. Hence, in the second analysis, we applied term-frequency–inverse-document-frequency (tf-idf) weighting, a supervised classification technique, to identify botnet discussion from all DDOS-thread–port-effective posts. To enhance the specificity of our analysis, we further conducted keyword extraction to identify posts discussing two new botnet techniques, Mariposa botnet and Zbot, that prevailed during our data window of 2007–2011.[12] In the empirical analysis, we test whether the discussion of these botnets correlate with the observed DDOS attacks in the ISC backscatter data. We report the detailed keyword extraction steps and results in the online appendix, Section A4.

Note that LDA modeling and tf-idf weighting require good understanding of the language used in the forums and significant processing resources. As reported in Section 5.4, we find that except in Hackforums, the DDOS-attack posts did not have a sizeable impact on the DDOS attacks observed in our data. Therefore, in view of the difficulty in scrutinizing posts in other languages, we conduct these two sets of analysis only for the discussion in Hackforums. We defer studying the contents in other forums to future research.

4. Empirical Model

Our basic specification is a dynamic panel fixed-effects model,

[12] For details of the Mariposa botnet and Zbot, please refer to https://en.wikipedia.org/wiki/Mariposa_botnet and https://en.wikipedia.org/wiki/Zeus_(malware) [accessed January 30, 2017].

$$r_{it} = \alpha_1 r_{i,t-1} + \alpha_2 f_{i,t-1} + \alpha_3 x_{it} + p_i' + d_t' + \varepsilon_{it}, \quad (1)$$

where r_{it} denotes the number of victim IPs attacked via port *i* in day *t*, $f_{i,t-1}$ denotes the number of DDOS-attack posts related to port *i* in day $t-1$, x_{it} includes the control variables including the number of threats issued and number of vulnerabilities on port *i* in day *t*, p_i' denotes port fixed effects, d_t' denotes day fixed effects, and ε_{it} captures idiosyncratic random errors.

We use the forum discussion lagged by one day, $f_{i,t-1}$, instead of the contemporaneous discussion to allow for the possibility that it may take time for the discussion to diffuse into the hacker community. Furthermore, hackers need not take action immediately after participating in the discussion. Parameterizing the discussion using a lagged variable allows us to capture these delayed effects. We report variations of this specification in the online appendix by including the contemporaneous discussion and the discussion lagged by more days. The estimation results of these variations are highly consistent with the results reported below.

We also include the number of DDOS-attack victims lagged by one day, $r_{i,t-1}$, in the model. In general, r_{it} and $r_{i,t-1}$ may be correlated if DDOS attacks exhibit intertemporal substitution. Such intertemporal substitution could occur when the attack trends are cyclical or encompass novelty or recency effects (i.e., recent attacks removing the novelty of launching further attacks, causing the near-term attack rate to decrease). Omitting such intertemporal correlations may bias the estimation of the forum-discussion effect.

It is well known, however, that including the lagged dependent variable in a within-group estimator produces biased estimates if the number of observations per cross-sectional unit is small (Blundell and Bond 1998). This is not the case in our setting because our dataset contains

1,826 days of observations per port. Hence, the bias due to the inclusion of $r_{i,t-1}$ is negligible (Bond 2002).[13]

The port fixed effects, p_i', help capture any variations in DDOS attacks due to application design. For example, many SYN attacks target ports 80 and 3389, which are the default ports used for World Wide Web and Remote Desktop services. Similarly, the day fixed effects, d_t', capture variations that are generic across all ports. For example, the propensity of DDOS attacks may vary because of holidays, noteworthy world events such as the 2008 financial crisis, the release of new DDOS-attack or -protection tools in the market or, simply, general DDOS-attack discussion that does not mention a specific port. We cluster the standard errors, ε_{it}, by port to allow for flexible correlations in DDOS attacks over time.

Furthermore, as shown in Figures 1 and 2, the number of DDOS-attack victims and forum discussion vary widely over time and across ports and forums, and their trends are somewhat skewed. Accordingly, we use the double-log specification, which fits skewed data better. Where necessary, we add one before taking logarithm to avoid logarithm of zero. With this specification, we can interpret the estimated coefficients as elasticities.

With the model in equation (1), we utilize a specific piece of content – the port number – to associate the discussion with the DDOS attacks. This association is highly focused and provides a powerful tool to scrutinize the discussion effect. We identify the impact of the discussion by

[13] A common treatment for models with a lagged dependent variable is to use the dynamic generalized method of moments (GMM) estimator (Arellano and Bond 1991). However, we cannot obtain the GMM estimator because our dataset contains many ports and days, which generate too many observations and lagged instruments in the GMM model. The GMM estimator does not converge to give any estimate. Because the bias due to the lagged dependent variable is negligible in our setting (with many observations over time for each port), the fixed-effects model should give reasonably accurate estimates of the forum-discussion effect.

exploiting the lagged and scattered distribution of DDOS-attack posts involving different ports over time. We strengthen the identification and test the robustness of our empirical model with instrumental variables (IV) and several validation tests.

5. Results

The six hacker forums and the vulnerability and threat databases mentioned 35,450 ports. With 1,826 days in 2007–2011, we have a panel of 35,450 ports × 1,826 days = 64,731,700 observations (64,696,250 after removing one day of observations because we use lagged variables). Table 3 presents the descriptive statistics and correlations of the variables.

[Insert Table 3 here]

Our main analysis includes only the discussion data from Hackforums, which has the highest global traffic rank among the hacker forums. Referring to Table 2, Hackforums also has the widest diversity and lowest concentration in terms of the visitors' countries of origin. Focusing on Hackforums allows us to make a better and proper comparison with the content analysis reported in Section 5.3. We explore the effects of the DDOS-attack discussion in the other five hacker forums in Section 5.4.

Table 4, Column (1) reports the result of estimating Model (1). Both the number of threats and number of vulnerabilities have positive and significant correlations with the number of DDOS-attack victims, which is well expected because threats and vulnerabilities make the port a more attractive target for cyber attacks. The lagged number of victim IPs has a positive impact, which does not support the presence of intertemporal substitution in the attacks.

[Insert Table 4 here]

More importantly, the number of DDOS-attack posts has a significant *negative* impact on the number of DDOS-attack victims. A 1% increase in DDOS-attack posts *decreases* the number of DDOS-attack victims by 0.032%. Our dataset contains an average of 196 DDOS-thread–port-effective posts and 2.18 million victim IPs per day. This estimate implies that increasing the discussion by two posts per day would decrease the number of victim IPs by around 700 per day.[14] This impact is economically significant.

5.1. Identification

The estimation in Table 4, Column (1) does not account for endogeneity. For example, there could be reverse causality, in the sense that the attackers or victims may share their experience with Hackforums participants immediately before or after the attacks. Omitted variables, such as underground "blackhat" discussion or the sharing of hacking knowledge in the "dark" web, may also bias the coefficient of forum discussion.

Our use of lagged DDOS-attack discussion as an independent variable should ameliorate the threat from reverse causality. Besides, such reverse causality, if it exists, would cause the estimated effect to bias *upward*. As we find a *negative* impact of forum discussion, the reverse-causality argument does not refute our finding that the forum discussion has reduced the number

[14] From Table 3, the number of DDOS-thread–port-effective posts in Hackforums = 5.541 ÷ 1,000 × 35,450 ≅ 196 per day. The number of DDOS-attack victims = 61.417 × 35,450 ≅ 2.18 million per day. One percent of 196 is around two. 0.032% of 2.18 million is around 700. Note that we count a post multiple times if it mentions more than one port number. We also count a victim IP address multiple times if it was attacked via multiple ports. Hence, the average numbers of DDOS-thread–port-effective posts and victim IPs in Table 3 contain duplicated entries. We calculate the effect size in the same way in all the estimates reported in this paper.

of DDOS-attack victims. The inclusion of lagged number of DDOS-attack victims as an independent variable also helps control for the effects of omitted variables.

Nevertheless, it is instructive to devise an identification strategy that is robust to omitted variables, particularly when we cannot observe real *blackhat* hackers or underground hacking activities (cf. Olson 2012; Coleman 2014). We first use an IV identification strategy. Following the procedures in Section 3.1, we constructed an IV by counting the number of posts that mentioned or replied to an earlier post that mentioned a port in a thread that *does not* contain any DDOS-attack post. In other words, we use the non-DDOS-thread–port-effective posts as IV. The assumption is that the tendency to post a hacking discussion is correlated across topics (the relevant condition for IV tests), but it is unlikely for the non-DDOS-attack posts to correlate with the DDOS attacks in our data (the exclusion restriction for IV tests).

Table 4, Column (2) reports the IV estimator obtained by two-stage least squares (2SLS) regression. The coefficient of lagged non-DDOS-attack posts in the first-stage regression, 0.084, is positive and statistically significant ($p < 0.01$). The Kleibergen–Paap Wald statistic (Kleibergen and Paap 2006) is 26.92, which exceeds the critical value of 16.38 for a maximal size of 10% for the Wald test in 2SLS (Stock and Yogo 2005). Hence, our IV passes the weak-instrument test.[15] The *C* statistic (Hiyashi 2000) is 9.317 ($p < 0.01$), implying the number of DDOS-attack posts is indeed endogenous. Hence, the IV estimator is preferred over the uninstrumented estimator. For brevity, we omit the first-stage regression and the detailed IV diagnostics.

[15] The Anderson-Rubin Wald statistic is 12.56 ($p < 0.01$), suggesting the 2SLS estimator is robust even if the instrument is weak.

The effect of the number of DDOS-attack posts, −0.122, is almost three times bigger than that reported in Table 4, Column (1), and continues to be statistically significant ($p < 0.01$). As discussed above, if reverse causality is present or some omitted variables have caused the DDOS attacks and discussion to co-move, then we would expect the true effect of the discussion to be *more negative*, which is the case with the IV estimator. The IV estimator implies that a 1% increase in the number of DDOS-attack posts, or two posts per day, would decrease the number of DDOS-attack victims by 0.122%, or more than 2,600 per day.

Despite the strong IV estimation results, we prefer to use the uninstrumented estimator for two reasons. First, our goal is to estimate the *sign* of the impact of DDOS-attack discussion with confidence. Using a more conservative estimate runs a lower risk of overstating the impact. Second, we cannot construct good IVs for some of the following tests. For better comparison, we use the uninstrumented estimator for Model (1) as the benchmark.

We next assess the validity of our finding via the concept of falsification. The idea is that the treatment effect should not exist in a setting where it should not apply. If it is also confirmed in a falsification test, then the treatment effect found in the key research of interest may be spurious instead of causal (Prasad and Jena 2013). The falsification test is particularly helpful in our setting because our dataset is large. With close to 65 million observations, we face a high risk of identifying spurious or erroneous associations.

To conduct such a test, we obtained a set of intrusion data reported by the firewalls in the ISC DShield sensors.[16] These intrusion attempts exclude DDOS attacks. Hence, the discussion

[16] For details of the DShield project, please refer to ISC's website, http://www.dshield.org/reports.html [accessed February 3, 2017]. Note that the DShield intrusion data contains only 1,674 days, not 1,826 days as in the backscatter (DDOS-attack) data.

of DDOS attacks *should not* affect the number of victims in the DShield intrusion data. Table 4, Columns (3) and (4) report two tests related to this strategy. In the first test, we use the number of IPs detecting intrusions as the dependent variable and *all* port-effective posts discussing *all* security attacks to measure forum discussion. Consistent with Table 4, Column (1), the number of port-effective posts has a negative and statistically significant impact, meaning general port discussion in Hackforums causes fewer intrusions. In the second test, we use the number of DDOS-thread–port-effective posts as the key independent variable. Because the DShield data does not include DDOS attacks, the discussion of DDOS attacks should not have an impact. As shown in Table 4, Column (4), this is indeed the finding. Collectively, these two tests suggest that the forum discussion effect is highly specific to DDOS attacks.

Finally, we conduct another falsification exercise to test the validity of our use of the port number to link the forum discussion and DDOS attacks. Table 4, Column (5) reports a variation where we randomize the forum discussion variable on the right-hand-side of Model (1); we deliberately mismatch the ports between the forum discussion and DDOS attacks. Supporting our use of the port as the linking variable, the effect of forum discussion with the randomized port numbers (placebo) is not statistically significant (p = 0.924). Once again, this test suggests that our finding is not caused by some general trends that move both the discussion in Hackforums and the observed DDOS attacks from the backscatter data.

5.2. Robustness

We conduct several sets of robustness tests. First, as discussed in Section 3.1, we use the number of DDOS-thread–port-effective posts to measure forum discussion. Technically, for each type (port or DDOS attack) of post, we can measure the discussion in three ways – the posts

mentioning the port number or DDOS attack, the effective posts including the posts mentioning the port number or DDOS attack and all subsequent follow-ups, and the entire thread with the port or DDOS-attack posts. This implies there are 3 × 3 = 9 ways of measuring DDOS-attack discussion. We report estimations using the other eight measures in the online appendix. The results are qualitatively similar to those reported in Table 4, Column (1).

Next, recall that we use the number of DDOS-attack posts lagged by one day to measure the discussion. We report variations in the online appendix using the contemporaneous forum discussion and other orders of lagged discussion. Regardless of the number of lags, the effect of the discussion on DDOS attacks, if significant, is always negative. We also test the robustness of our empirical model by omitting the threat and vulnerability variables, omitting the lagged attack variable, and including port–month fixed effects. The port–month fixed effects should control for seasonality if the DDOS attacks to specific ports follow different trends over time. As reported in the online appendix, none of these variations changes our conclusion.

We then test whether our result is robust to the exclusion of outliers. The ISC data records an attack only if a backscatter packet is received. This means that the number of attacks to a port is missing (instead of zero) if no backscatter packet is received from that port. In our analysis, we imputed the missing data by assigning a value of zero if some other ports have recorded DDOS attacks on the same day (hence suggesting that the ISC sensors were in operation). In another test reported in the online appendix, we exclude all imputed data, which essentially removes all observations with no DDOS attacks.[17] Similarly, in two other estimations, we omit port 0, which

[17] This test also addresses the concern that our dataset contains many observations with zero DDOS attacks. One way to explain such a large number of zeros is to model a separate data generation process for them using a zero-inflated negative binomial regression. However, it is computationally infeasible to conduct such a regression here

is not used for serious purposes in practice, and the five most attacked ports, 6881, 80, 53, 4672, and 137. We also prune the data before May 2, 2007, the date when Hackforums was officially launched (though it started operation earlier than that day). As reported in the online appendix, our conclusion is robust to the exclusion of these outliers.

Note that the estimation in Table 4, Column (1) includes some ports mentioned in other forums but not Hackforums. If the discussion in other forums has caused more DDOS attacks via the mentioned ports, then the inclusion of these ports for which no DDOS-attack discussion is observed in Hackforums could bias the effect of the DDOS-attack discussion *downwards*. In the online appendix, we report a test that includes only the 28,860 ports specifically mentioned in Hackforums. Our conclusion remains unchanged.[18]

Finally, our dataset contains millions of observations. With such a large sample, the *p*-value would quickly converge to zero even when the effect size is negligible for practical purposes. To avoid exaggerating the significance of the impact of DDOS-attack discussion, we report the elasticities, which range from 0.032 (the uninstrumented estimator) to 0.122 (the IV estimator). These estimates are economically significant as they imply a few posts in Hackforums could have reduced hundreds, or even thousands, of victims per day.

Following the advice of Lin et al. (2013), we report two coefficient–p-value–sample-size (CPS) charts in Figure 3. The first chart plots the coefficient of the number of DDOS-attack posts

because our model is dynamic and richly parameterized with tens of thousands of fixed effects. In any case, we do not have any strong reason to expect that the zero attacks follow a different process.

[18] As reported in Section 5.4, we do not find a positive correlation between the DDOS-attack discussion in the other forums with the number of DDOS-attack victims.

in 30 regressions that progressively increase the number of ports in the sample. The second chart plots a similar graph with sample size progressively increasing by the number of days. Evidently, the effect is robust in smaller samples. The 95% confidence intervals mostly lie below, and are often quite distant from, zero. Therefore, the negative effect of the discussion is robust even if we use the most conservative bounds for the confidence intervals.

[Insert Figure 3 here]

5.3. Content of the Discussion

We scrutinize the DDOS-thread–port-effective posts in Hackforums to examine the role of content on DDOS attacks. We conduct two analyses. The first applies topic modeling using LDA analysis. We generate multiple topic models that differ in the number of topics and assign the DDOS-thread–port-effective posts into these topics. We follow Quinn et al. (2010) and give more weight to words specific to a topic to provide stronger distinctive power by downplaying common words, such as *hack* and *port*, which tend to appear in many topics.

We then use the number of posts in each of these topics as an independent variable in the regression. Specifically, we estimate a variant of Model (1):

$$r_{it} = \alpha_1 r_{i,t-1} + \alpha_2' T_{i,t-1} + \alpha_3 x_{it} + p_i' + d_t' + \varepsilon_{it}. \quad (2)$$

Model (2) is similar to Model (1) except that we replace the number of DDOS-attack posts, $f_{i,t-1}$, by the constituent topics, $T_{i,t-1}$. For brevity, we report the LDA procedure, topic keywords, and regression results in the online appendix, Section A3. As one example, Figure 4 presents the keywords in the 4-topic model, which has the lowest perplexity score (Brown et al. 1992) and

best predicts the sample of DDOS-thread–port-effective posts in Hackforums. Table 5, Column (1) reports the estimation of Model (2) using the four topic variables.

[Insert Figure 4 and Table 5 here]

In general, across all LDA models with different numbers of topics, the informative topics and their follow-up discussion tend to have negative influences on DDOS attacks. These topics often include such keywords as *http, error, file, short, work, nice, great, link, post, tutori* or *tut, click, includ, plea* or *plz, view, spoiler, result, packet, program, use, messag, hack* and *site.* However, some topics *increase* DDOS attacks. These topics often include such keywords as *link, download, bot, ip, viru, server, host, plea, know, file, run, click* and *password.* Evidently, some keywords appear in multiple topics that have distinct impacts on DDOS attacks.[19]

Nevertheless, the LDA modeling results provide some indicative insights. Keywords such as *bot, viru* and *irc* tend to belong to topics that increase DDOS attacks.[20] Guided by this observation, our second analysis applies tf-idf weighting to extract botnet posts. We focus on botnet instead of computer virus because it is often involved in DDOS attacks. To increase the specificity of our inference, we further conduct another round of keyword extraction to identify posts discussing new botnets that prevailed during our sampling window. We have identified two

[19] We applied stemming and lemmatization (Manning et al. 2008, pp. 32–34) to preprocess the posts and transform derivatively related forms of a word into common base forms before applying the LDA analysis. For example, we transform the variants *include, including, included* and *includes* into the common base form *includ.* This explains why some of the keywords such as *includ, plea* and *viru* are truncated. Although some of these top keywords appear in multiple topics in the same model, the cosine similarity (Singhal 2001) between the topics that significantly correlate with DDOS attacks in the regressions mostly lies below 0.2. This implies that the LDA topics are quite distinctive in content. We also checked the cosine similarity between the topics *across* the different LDA models. Similar topics have qualitatively similar impacts in the regressions.

[20] Referring to the online appendix, Tables A3 and B4, the effect of topics containing the keyword "bot" is either insignificant or positive and statistically significant.

such botnets, Mariposa botnet and Zbot. We report the details of the tf-idf weighting procedure and the botnet (and new botnets) post extraction in the online appendix, Section A4.

Furthermore, because the LDA results hint that some topics with keywords such as *nice, great,* etc., may decrease attacks, in the next estimation, we separately consider the influences of botnet mentions and follow-ups using a variant of Model (1),

$$r_{it} = \alpha_1 r_{i,t-1} + \alpha_{21} f_{i,t-1} + \alpha_{22} m_{i,t-1} + \alpha_{23} w_{i,t-1} + \alpha_3 x_{it} + p_i' + d_t' + \varepsilon_{it}, \quad (3)$$

where $m_{i,t-1}$ is the number of botnet mentions related to port *i* and day $t-1$, and $w_{i,t-1}$ is the number of follow-up posts weighted by either the number of botnet posts in the same thread or the duration, in terms of number of days, since the first botnet mention in the same thread.

Table 5, Columns (2)–(5) report the estimation results. In Columns (2) and (4), we weight the follow-ups by the total number of botnet posts in the same thread. In Columns (3) and (5), we weight the follow-ups by the number of days since the first botnet mention appeared. The regressions in Columns (2) and (3) use general botnet variables. The regressions in Columns (4) and (5) focus on two new botnets, Mariposa botnet and Zbot.

The result is illuminating. In general, botnet mentions increase DDOS attacks. However, as more posts contribute to discussing the botnets, the number of attacks decreases. We obtain this result regardless of whether we use the number of posts or duration as the weight for the follow-up discussion. Note that this result is obtained after we have controlled for the DDOS-attack trends using the fixed effects and the lagged attack variable, and the forum discussion trends using the lagged number of DDOS-attack posts. Hence, the botnet variables should capture the incremental impact of the forum due to the botnet discussion. The botnet posts are particularly damaging. During 2007–2011, there were around $0.856 \div 1{,}000 \times 35{,}450 \cong 30$ botnet

posts per day. The estimates in Table 5, Columns (2) and (3) suggest that increasing botnet mentions by one post, 3%, would *increase* the number of victim IPs by around 970 to 2,600. Similarly, the estimates in Table 5, Columns (4) and (5) suggest that Mariposa botnet or Zbot mentions in Hackforums could potentially increase the number of DDOS-attack victims by millions per day!

Overall, the content analysis suggests that the discussion consists of heterogeneous topics. These topics could variously increase or decrease DDOS attacks even though they contain some common keywords. Mentioning botnets, especially new botnets, tends to increase DDOS attacks. The follow-up discussion tends to decrease DDOS attacks.

5.4. Additional Forums

To test the boundary of our findings, we obtained data from five other forums and extracted their port and DDOS-attack posts using the same procedures as in Section 3.1. In Table 6, Column (1), we report the estimation by including the DDOS-thread–port-effective posts from the other five forums as additional independent variables. The effect of Hackforums discussion remains negative and statistically significant. Among the other forums, only the discussion in Antichat has a significant influence on DDOS attacks, and the effect is also negative.

[Insert Table 6 here]

However, the Antichat effect is smaller than that of Hackforums – the elasticity is only −0.005. Referring to Table 3, Antichat had $121.9 \div 1{,}000 \times 35{,}450 \cong 4{,}300$ posts per day. The estimate in Table 5, column (1) implies increasing DDOS-attack discussion in Antichat by 1% – around 43 posts per day – would decrease the number of DDOS-attack victims by 130 per day.

This effect is less than (130 / 43) ÷ (700 / 2) = 0.9% of that of Hackforums, the discussion in which could spare 700 victims with merely two posts per day.

Table 6, Column (2) reports another estimation that groups the six forums by language – English for Hackforums and HBH, Chinese for Hackbase and HHLM, and Russian for Antichat and Xaker. The result is similar. Only the discussion in the English and Russian forums mattered, which is likely due to Hackforums and Antichat. The estimate in Table 6, Column (3) groups all the forum posts together. Here again, the conclusion is similar except that the effect size becomes smaller because the grouping dilutes the influence of Hackforums' posts.[21]

Taken together, these results suggest that only the discussion in Hackforums and Antichat had material impacts on the observed DDOS attacks. This could be due to differences in user profile and the scope of our data. Referring to Tables 2 and 3, the visitors to the Chinese forums, Hackbase and HHLM, are mostly from China, whereas the volume of posts in HBH and Xaker is relatively small. By contrast, the traffic of Hackforums and Antichat is more globalized and hence aligns better with the global nature of our DDOS-attack data.[22]

6. Discussion and Conclusions

By compiling a comprehensive dataset from the field, we find that hacker-forum discussion of DDOS attacks *decreases* the number of DDOS-attack victims. Content analysis using LDA topic

[21] In the online appendix, we report another set of estimates that enter the forum discussion variables one by one. The results are qualitatively similar.

[22] Referring to Table 2, the Antichat visitors were quite concentrated in Russia. However, the prior literature has suggested that Russian hackers often target victims in foreign countries instead of Russia (Howard 2009, pp. 172). Although the discussion in Antichat adds to the negative impact of Hackforums, as shown in Figure 1, the discussion in Antichat exhibited some discrete bursts around July 2008 and died down subsequently. Hence, its impact may not be robust and persistent. In any case, the size of the Antichat effect is less than 0.9% of that of Hackforums. It is less significant for practical considerations.

modeling and tf-idf classification shows that discussion topics with many overlapping keywords can variously increase or decrease DDOS attacks. Mentioning botnets, especially new botnets, tends to increase the attacks, but the follow-up discussion tends to decrease DDOS attacks. The size of the Hackforums discussion effect is large and economically significant. The discussion effect is considerably smaller and often insignificant among the other hacker forums.

Our findings highlight the importance of scrutinizing the discussion rather than merely the disclosure of sensitive information in an online community. Prior research suggests that properly designed vulnerability disclosure mechanisms can help reduce cyberattacks (Arora et al. 2010; Ransbotham et al. 2012). In our setting, hacker forums are often used to disclose attack-related information. We asked two RAs to classify the nature of the discussion by manually reading the titles and leading posts of all port-related DDOS-attack threads. Table A5.1 in the online appendix reports their classification. Similar to the findings in previous research (e.g., Holt and Lampkeb 2010; Holt 2012), many posts in Hackforums are indeed malicious in nature. Among the 2,781 threads in Hackforums, 2,458 (88%) are ill intentioned. Paradoxically, although these posts are seemingly malicious, development in their discussion decreases DDOS attacks.

How does the discussion reduce instead of increase DDOS attacks? Further analysis of the communication patterns and content provides some hints. Recall from Section 5.4 that only the discussion in Hackforums and Antichat had statistically significant impacts on DDOS attacks. As shown in Table A5.1 in the online appendix, Hackforums and Antichat also had most replies. In Hackforums, DDOS-attack threads average 40.6 posts, with some threads exceeding 1,000 posts. Incredibly, Antichat's DDOS-attack threads averaged 403 posts! Such elaborate discussion may create a rich knowledge base on cyberattacks (T. Wang et al. 2013; Kim and Kim 2014; Samtani

et al. 2015; J. G. Wang et al. 2015). As hacking knowledge carries the dual-use characteristic, the discussion may be put into good (protection) use.

We further scrutinize the top ten keywords and find that Hackforums and Antichat seem to share some similarities. Specifically, as shown in Table A5.2 in the online appendix, within the post titles, both bot- and DDOS-related keywords are ranked highly in Hackforums and Antichat. When we look at the top keywords in the leading and reply posts in Tables A5.3 and A5.4, we see that both Hackforums and Antichat contain more technical terms such as *server, port, file* and *http,* capturing the snippets of the technical depth of the discussion. Overall, we find that the nature of the posts may not be directly indicative of their impacts. The discussion development could bring surprising results in real-world outcomes.

Our content analysis shows that the topics discussed in hacker forums are heterogeneous even though they contain many common and highly related keywords. These topics can variously increase or decrease DDOS attacks, meaning we cannot rely on keyword inspection to analyze the real-world impact of online forums or, more broadly, social networking websites (cf. Abbasi et al. 2014; Benjamin et al. 2015; Benjamin et al. 2016). We need delicate empirical strategies to dissect the mechanisms and dynamics that underscore the impact of online discussion.

Our analysis in Section 5.3 suggests one interesting mechanism – mentioning new botnets causes more attacks, but the follow-up discussion reduces the attacks. This means that when online disclosure of cyberattacks has happened, facilitating its continuous discussion instead of sanctioning it may be advisable. This implication is novel, and it extends prior research that has largely omitted the evolution of online discussion and its impacts (e.g., Holt and Lampkeb 2010;

Holt 2012). This evolution may be especially important for discussion that possesses the dual-use characteristic, such as hacking knowledge and tools.

Practically, this research adds a new perspective to the continuing debate on censorship of malicious public information. The prevailing argument for such censorship is that it prevents dangerous information from falling into the "wrong" hands, which could adversely influence society. Some lawmakers indeed consider such censorship necessary in protecting the public. For example, the Computer Misuse and Cybersecurity Act of Singapore censors content that could potentially cause public mischief. Recently, some legislators have considered bills restricting the use of end-to-end encryption (Martin 2016; McLaughlin 2016).

However, before we adopt these measures, we must consider if it is practicable to stop the disclosure of malicious information. People who are determined to share such information may turn to underground channels. Our research shows that once the initial (malicious) post is published, the subsequent discussion contributes to decreasing the attacks. When the Internet promotes user-generated content and when it is difficult to ascertain the real identity of Internet publishers, censoring the discussion of harmful content need not help protect the public. It may simply deprive the public of the knowledge needed to protect itself.

Note that censorship can carry a cost too. It may impede freedom of expression and user innovation (von Hippel and Paradiso 2008). It may also remove online knowledge and strategic intelligence that can be used against perpetrated crimes in fighting against terrorism (Bambauer 2009; Holt and Lampkeb 2010). If the concerned information exhibits dual use and does not empirically cause more crimes, society should not bear the censorship costs. At least in the case of hacking discussion, our findings do not support a blanket censorship policy.

Nonetheless, a recent incident highlights the intricate dilemma in regulating online hacker forums. On October 21, 2016, a malware named Mirai initiated a massive DDOS attack in the United States, causing a large-scale Internet outage. The Mirai source code was first made publicly available on Hackforums on September 30, 2016. Subsequent to the Mirai outbreak, Hackforums permanently shut down its Server Stress Testing section because the section has become a top destination for people to buy DDOS-for-hire services. Although the forum administrator believes that there are legal and legitimate uses of website stress-testing tools (which are also the same tools used for DDOS attacks), he eventually shut down the section (Kan 2016).

The Mirai incident is consistent with our content analysis – the first mention of a botnet (in this case, publishing the source codes of Mirai) increases DDOS attacks, and the impact of the mention could be remarkable. However, what we cannot observe is the subsequent effect of the forum discussion. Had Hackforums not shut down the section containing the Mirai discussion, our research would predict that the attacks would decrease as the discussion of Mirai continued. This raises the question – was the shutdown necessary? Evidently, legitimate users of the Server Stress Testing section of Hackforums have suffered a collateral damage.

Our findings provide specific guidance on hacker-forum regulation. Posts mentioning new botnets and, by extension, new hacking techniques may cause a big harm to other users. We may want to closely monitor the development of such posts. For example, the forum administration can try to promote more whitehat discussion that provides more knowledge and perspectives to help the community tackle the new threats. Because the discussion in hacker forum correlates with real cyber attacks, hacker forums can be a novel and effective avenue for regulators and law

enforcement agencies to analyze global cyberattack trends. It may also serve as a practical channel for communication between hackers and security managers.

In fact, given the potential benefit of elaborate discussion, firms could potentially harness hacker forums as a security training ground. To be proficient in information security, a person needs a wide skill set from sustained training and real cyberattack or defense experience. This explains the proliferation of professional programs, such as the Certified Ethical Hacker training program offered by the International Council of E-Commerce Consultants (EC-Council). Some companies even wrestle with the dilemma of hiring criminal hackers to advance their knowledge (Armerding 2012). Instead of seeking such help, perhaps firms could simply venture into online hacker forums to train in-house experts.

Overall, our empirical findings and the arguments above point to a clear conclusion – hacker-forum discussion can help offset the harm caused by malicious sharing. It can also serve as a good knowledge platform for both firms and regulators. Hence, blanket censorship or shut down of hacker forums is ill advised. However, selective moderation of the forum content may be necessary, especially for novel but impactful new malicious tools.

There are several limitations in this research. We cannot identify whether whitehat or blackhat hackers have contributed the posts in the hacker forums. Without intimate access to individual hackers, studying hacker motivation is always challenging especially in large-scale empirical studies. We also lack data on underground hacking groups. All six forums studied here are public and so need not be popular among true blackhat hackers. Future work should study secret underground channels such as closed or private IRC chat rooms.

Furthermore, we study only one type of cyberattack, DDOS attacks. We argue it is an appropriate context because its related knowledge possesses the dual-use characteristic and, as Figure 2 shows, the top-discussed ports in Hackforums are often top targets of DDOS attacks. The ISC backscatter dataset has a broad coverage in time and geographical regions, and it allows us to trace each attacked port. These features help us unambiguously link the DDOS attacks to the forum discussion. Nonetheless, extending the analysis to other cyberattacks such as computer virus, Trojan horse, worm, and phishing can enrich our understanding of the boundary of the forum discussion impact.

Although our dataset is comprehensive, it does not allow us to study the impact of less general forums that have specific geographic-region, country, or language scope. To specifically test the effect of such forums, we need more granular attack data. The challenge is that we need to match the forum discussion to the attack via the port. This is a demanding requirement.

Empirically, because we want to identify the forum discussion effect using the mentioned port numbers, we use the day fixed effects to control for general DDOS-attack discussion without any port number. Such discussion may also affect DDOS attacks. Our empirical framework cannot provide a separate identification of this general discussion effect because it mingles with other day-specific port-invariant influences, such as the holiday effect.

We conclude the paper by suggesting several future research directions. First, we should continue to explore the mechanisms underlying the impact of forum discussion. Does the forum discussion decrease DDOS attacks because it educates security managers or removes the novelty of the attacks? Perhaps the discussion has established an ethical social norm?

Second, future research should test the dual-use theory empirically. It will be a challenging test because, by this theory, the same piece of content can have opposite impacts depending on the context. One possibility to overcome this challenge is to attach the visitors to the posts and infer the positive or negative impacts of the discussion through the roles played by the visitors (e.g., whether they are hackers or security managers).

Finally, it will be meaningful to extend this analysis to other online-community or social-networking websites with sensitive discussion, such as the forums dedicated to political issues or indecent affairs. We need better empirical evidence to inform public policy before taking action to regulate the exchange in these emerging online platforms.

References

Abbasi, A., Li, W., Benjamin, V., Hu, S., and Chen, H. (2014). "Descriptive Analytics: Examining Expert Hackers in Web Forums," *2014 IEEE Joint Intelligence and Security Informatics Conference (JISIC)*. The Hague, Netherlands: IEEE.

Armerding, T. (2012). "Should Companies Hire Criminal Hackers?" *CSO Online*, July 27, 2012. Available http://www.csoonline.com/article/2132036/malware-cybercrime/should-companies-hire-criminal-hackers-.html [accessed February 15, 2017].

Anti-Sec (2009). "Ant-Sec – We are going to terminate Hackforums.net and Milw0rm.com - New Apache 0-day exploit uncovered," *Full Disclosure Mailing List Archives*, July 15, 2009. http://seclists.org/fulldisclosure/2009/Jul/164 [Accessed January 12, 2017]

Antweiler, W. and Frank, M. Z. (2004). "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards," *Journal of Finance* (59:3), pp. 1259-1294.

Arellano, M. and Bond, S. (1991). "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies* (58:2), pp. 277-297.

Arora, A., Krishnan, R., Telang, R., and Yang, Y. (2010). "An Empirical Analysis of Software Vendors' Patch Release Behavior: Impact of Vulnerability Disclosure," *Information Systems Research* (21:1), pp. 115-132.

Auray, N. and Kaminsky, D. (2007). “The Professionalisation Paths of Hackers in IT Security: The Sociology of a Divided Identity,” *Annales Des Télécommunications* (62:11-12), pp. 1312-1326.

Banks, J. (2010). "Regulating Hate Speech Online," *International Review of Law, Computers & Technology* (24:3), pp. 233-239.

Bambauer, D. E. (2009). “Filtering in OZ: Australia's Foray into Internet Censorship,” *University of Pennsylvania Journal of International Law* (31:2), pp. 493-530.

Bauernschuster, S., Falck, O., and Woessmann, L. (2014). “Surfing Alone? The Internet and Social Capital: Evidence from an Unforeseeable Technological Mistake,” Journal of Public Economics (117), pp. 73-89.

Benjamin, V., Li, W., Holt, T., and Chen, H. (2015). “Exploring Threats and Vulnerabilities in Hacker Web: Forums, IRC and Carding Shops,” *2015 IEEE International Conference on Intelligence and Security Informatics (ISI)*, Baltimore, United States, pp. 85-90.

Benjamin, V., Zhang, B., Nunamaker Jr., J. F., and Chen, H. (2016). “Examining Hacker Participation Lenght in Cybercriminal Internet-Relay-Chat Communities,” *Journal of Management Information Systems* (33:2), pp. 482-510.

Bhuller, M., Havnes, T., Leuven, E., and Mogstad, M. (2013). “Broadbank Internet: An Information Superhighway to Sex Crime?” *Review of Economic Studies* (80:4), pp. 1237-1266.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). “Latent Dirichlet Allocation,” *Journal of Machine Learning Research* (3), pp. 993-1022.

Blundell, R. and Bond, S. (1998). “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models,” *Journal of Econometrics* (87:1), pp. 115-143.

Bond, S. (2002). “Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice,” *Cemmap Working Paper*, The Institute for Fiscal Studies, Department of Economics, University College London, April 3, 2002.

Brown, P. F., Della Pietra, V. J., Mercer, R. L., Della Pietra, S. A., and Lai, J. C. (1992). “An Estimate of an Upper Bound for the Entropy of English,” *Computational Linguistics* (18:1), pp. 31-40.

Chan, J., Ghose, A., and Seamans, R. (2016). " The Internet and Racial Hate Crime: Offline Spillovers from Online Access," *MIS Quarterly* (40:2), pp. 381-403.

Chan, J., and Ghose, A. (2014). "Internet's Dirty Secret: Assessing the Impact of Online Intermediaries on HIV Transmission," *MIS Quarterly* (38:4), pp. 955-975.

Chen, H., De, P., Hu, Y., and Hwang, B.H. (2014). “Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media,” *The Review of Financial Studies*, (27:5), pp. 1367-1403.

Coleman, E.G. (2013). *Coding Freedom: The Ethics and Aesthetics of Hacking*. New Jersey: Princeton University Press.

Coleman, E.G. (2014). *Hacker, Hoaxer, Whistleblower, Spy: The Many Faces of Anonymous.* London, NY: Verso.

Constantin, L. (2016). “Armies of Hacked IoT Devices Launch Unprecedented DDoS Attacks,” *PC World*, September 27, 2016. https://www.pcworld.idg.com.au/article/607509/armies-hacked-iot-devices-launch-unprecedented-ddos-attacks/ [Accessed November 6, 2017]

Dunlop, S. M., More, E., and Romer, D. (2011). “Where do youth learn about suicides on the Internet, and what influence does this have onsuic idal ideation?” *Journal of Child Psychology and Psychiatry* (52:10), pp. 1073-1080.

Flaxman, S., Goel, S., and Rao, J. (2016). “Filter Bubbles, Echo Chambers, and Online News Consumption,” Public Opinion Quarterly (80:S1), pp. 298-320.

Geva, T., Oestreicher-Singer, G., Efron, N. and Shimshoni, Y. (2015). “Using Forum and Search Data for Sales Prediction of High-Involvement Products,” *MIS Quarterly*, forthcoming, available at https://ssrn.com/abstract=2294609 [accessed February 15, 2017].

Godes, D., and Mayzlin, D. (2004). “Using Online Conversations to Study Word-of-Mouth Communication,” *Marketing Science* (23:4), pp. 545-560.

Hiyashi, F. (2000). *Econometrics*. Princeton: Princeton University Press.

Holt, T. J. (2012). “Examining the Forces Shaping Cybercrime Markets Online,” *Social Science Computer Review*, (31:2), pp. 165-177.

Holt, T. J. and Lampkeb, E. (2010). “Exploring stolen data markets online: products and market forces,” *Criminal Justice Studies* (23:1), pp. 33-50.

Hosanagar, K., Fleder, D., Lee, D., and Buja, A. (2014). “Will the Global Village Fracture Into Tribes? Recommender Systems and Their Effects on Consumer Fragmentation,” *Management Science*, (60:4), pp. 805-823.

Howard, R. (2009). *Cyber Fraud: Tactics, Techniques, and Procedures*. Boca Raton, FL: Auerbach Publications.

Hunton, P. (2009). "The Growing Phenomenon of Crime and the Internet: A Cybercrime Execution and Analysis Model," *Computer Law and Security Review* (25:6), pp. 528-535.

Hui, K.L., Kim, S.H., and Wang, Q.H. (2017) "Cybercrime Deterrence and International Legislation: Evidence from Distributed Denial of Service Attacks," *MIS Quarterly* (41: 2), pp. 497-523.

Imperva (2011). *Hacker Intelligence Initiative, Monthly Trend Report #5*. https://www.imperva.com/docs/HII_Monitoring_Hacker_Forums.pdf [accessed February 10, 2017]

Jordan, T. and Taylor, P. (1998). A Sociology of Hackers. *The Sociological Review* (46:4), pp. 757-780.

Kan, M. (2016). "Hacking forum cuts section allegedly linked to DDoS attacks," *CIO.com*, October 28, 2016. Available at http://www.cio.com/article/3136357/hacking/hacking-forum-cuts-section-allegedly-linked-to-ddos-attacks.html [accessed February 15, 2017].

Kaplan, J. E. and Moss, M. P. (2003). *Investigating Hate Crimes on the Internet.* Washington DC: Partners Against Hate, U.S. Department of Justice.

Katyal, N. K. (2001). "Criminal Law in Cyberspace," University of Pennsylvania Law Review (149:4), pp. 1003-1114. Available at http://scholarship.law.upenn.edu/penn_law_review/vol149/iss4/2

Kim, S. H. and Kim, B. C. (2014). "Differential Effects of Prior Experience on the Malware Resolution Process," *MIS Quarterly* (38:3), pp. 655-678.

Kim, S.H., Wang, Q. H. and Ullrich, J. (2012). "A Comparative Study of Cyber Attacks," *Communications of the ACM* (55:3), pp. 66-73.

Kleibergen, F. and Paap, R. (2006). "Generalized Reduced Rank Tests using the Singular Value Decomposition," *Journal of Econometrics* (133:1), pp. 97-126.

Lin, M.F., Lucas Jr., H. C. and Shmueli, G. (2013). "Too Big to Fail: Large Samples and the p-Value Problem," *Information Systems Research* (24:4), pp. 906-917.

Luxton, D. D., June, J. D., and Fairall, J. M. (2012). "Social Media and Suicide: A Public Health Perspective," *American Journal of Public Health*, (102:S2), S195-S200.

Manning, C. D., Raghavan, P., and Schutze, H. (2008). *Introduction to Information Retrieval*, Cambridge University Press.

Martin, A. J. (2016). "UK gov says new Home Sec will have powers to ban end-to-end encryption," *The Register*, July 14, 2016. Available at https://www.theregister.co.uk/2016/07/14/gov_says_new_home_sec_iwilli_have_powers_to_ban_endtoend_encryption/ [accessed February 15, 2017].

McLaughlin, J. (2016). "Bill That Would Ban End-to-End Encryption Savaged by Critics," The Intercept.com, April 9, 2016. Available at https://theintercept.com/2016/04/08/bill-that-would-ban-end-to-end-encryption-savaged-by-critics/ [accessed February 15, 2017].

Moore, D., Shannon, C., Brown, D. J., Voelker, G. M., and Savage, S. (2006). “Inferring Internet Denial-of-Service Activity,” *ACM Transactions on Computer Systems* (TOCS) (24:2), pp. 115-139.

Neumann, P. R. (2013). “Options and Strategies for Countering Online Radicalization in the United States,” *Studies in Conflict & Terrorism*, (36:6), pp. 431-459.

Neustar. (2017). *Worldwide DDoS Attacks & Cyber Insights Research Report: Taking Back the Upper Hand from Attackers*. May 2017.Olson, P. (2012). *We Are Anonymous*. New York, NY: Hachette Book Group, Inc.

O’Neil, M. (2006). “Rebels for the System? Virus Writers, General Intellect, Cyberpunk and Criminal Capitalism,” *Journal of Media and Cultural Studies* (20:2), pp. 225-241.

Panjwani, S., Tan, S., Jarrin, K. M. and Cukier, M. (2005). “An experimental evaluation to determine if port scans are precursors to an attack,” *Proceedings of the 2005 International Conference on Dependable Systems and Networks (DSN’05)*.

Prasad, V. and Jena, A. B. (2013). “Prespecified Falsification End Points: Can They Validate True Observational Associations?” *JAMA* (309:3), pp. 241-242.

Quinn, K. M., Monroe, B. L., Colaresi, M., Crespin, M. H. and Radev, D. R. (2010). “How to Analyze Political Attention with Minimal Assumptions and Costs,” *American Journal of Political Science*, (54:1), 209–228.

Ransbotham, S., Mitra, S. and Ramsey, J. (2012). “Are Markets for Vulnerabilties Effective?” *MIS Quarterly* (36:1), pp. 43-64.

Rui, H.X., Liu, Y.Z., and Whinston, A. (2013). “Whose and what chatter matters? The effect of tweets on movie sales,” *Decision Support Systems* (55:4), pp. 863-870.

Samtani, S., Chinn, R., and Chen, H. (2015). “Exploring hacker assets in underground forums,” *IEEE International Conference on Intelligence and Security Informatics (ISI)*, pp. 31-36.

Singhal, A.(2001). “Modern Information Retrieval: A Brief Overview,” *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering* (24:4), pp. 35-43.

Steinmetz, K. f. and Gerber, J. (2014). "The Greatest Crime Syndicate since the Gambinos: A Hacker Critique of Government, Law and Law Enforcement,” *Deviant Behavior* (35:3), pp. 243-261.

Stock, J. H. and Yogo, M. (2005). “Testing for Weak Instruments in Linear IV Regression,” In Andrews DWK *Identification and Inference for Econometric Models*. New York: Cambridge University Press. pp. 80-108.

Stockton, P. N. and Golabek-Goldman, M. (2013). “Curbing the Market for Cyber Weapons,” *Yale Law & Policy Review*, (32:1), pp. 239-266.

Thomas, J. (2005). “The moral ambiguity of social control in cyberspace: a retro-assessment of the ‘golden age’ of hacking,” *New Media & Society*, (7:5), pp. 599-624.

Turgeman-Goldschmidt, O. (2008). “Meanings that Hackers Assign to their Being a Hacker,” *International Journal of Cyber Criminology*, (2:2), pp. 382-396.

von Hippel, E. and Paradiso, J. A. (2008) “User Innovation and Hacking,” *IEEE Pervasive Computing* (7:3), pp. 66-69.

Wang, J.G., Xiao, N., and Rao, H. R. (2015). “An Exploration of Risk Characteristics of Information Security Threats and Related Public Information Search Behavior,” *Information Systems Research* (26:3), pp. 619-633.

Wang, T., Kannan, K. N., and Ulmer, J. R. (2013). “The Association between the Disclosure and the Realization of Information Security Risk Factors,” *Information Systems Research* (24:2), pp. 201-218.

Zhang, G.S., and Parashar, M. (2010). “Cooperative Detection and Protection against Network Attacks using Decentralized Information Sharing,” *Cluster Computing* (13:1), pp. 67-86.

Table 1. Hacker Forums and Post Distributions

	Traffic rank[+]	Posts[++]	Threads[++]	Subforums[++]	Port-related posts[++]	DDOS-attack posts[++]
Hackforums	Third in the *Hacking* subcategory under *Computers* in Alexa	2,960,893	355,222	23	24,610	13,410
HBH	Seventeenth in the *Hacking* subcategory under *Computers* in Alexa	63,300	8,058	39	302	69
Hackbase	Fifth in the *Hacker* subcategory under *Computers/Security* in Chinese in Alexa	1,733,924	175,021	9	5,884	430
HHLM	First in the *Hacker* subcategory under *Computers/Security* in Chinese in Alexa	388,938	52,154	11	4,194	1,284
Antichat	Not categorized in Alexa, but has higher ranking than most of the sites in the *Hacking* subcategory under *Computers* in Russian in Alexa	1,356,780	145,512	68	9,588	626
Xaker	Eighth in the *Hacking* subcategory under *Computers* in Russian in Alexa	55,127	9,830	35	744	124

[+]We obtained all ranking information from Alexa on January 22, 2017. Because Alexa does not publish historical statistics, we cannot obtain the ranking information in 2007–2011. [++]2007–2011 numbers.

Table 2. Forum Visitors by Country

	Hackforums		HBH	Hackbase		HHLM		Antichat		Xaker
	5/2015	1/2017	5/2015	5/2015	1/2017	5/2015	1/2017	5/2015	1/2017	5/2015
Algeria		0.6%								
Australia	2.7%	1.5%								
Azerbaijan									3.8%	
Bangladesh	0.7%									
Belarus								2.4%	2.0%	2.4%
Belgium		0.7%								
Brazil	0.8%									
Canada	3.3%	4.8%							0.5%	
China		1.1%		92.6%	69.2%	88.8%	96.8%			
Croatia	0.7%	0.9%								
Czech Republic									1.0%	
Denmark	1.6%	1.8%								
Egypt	2.2%	1.1%								
Finland									0.9%	
France	2.2%	1.4%							1.2%	
Germany	1.4%	5.5%							5.0%	
Greece	1.2%	1.2%								
Hong Kong		0.6%			1.2%	0.6%				
Korea				4.5%	8.2%					
India	22.6%	5.1%	11.7%							
Indonesia	1.20%									
Iran		1.0%								
Israel									0.6%	
Italy	1.3%	1.4%							3.6%	
Japan				0.9%	19.9%		1.3%			
Kazakhstan								4.2%	3.0%	3.3%
Kuwait		1.5%								
Latvia									0.8%	1.5%
Mexico		0.8%								
Morocco		0.5%								
Netherlands	5.0%	3.8%						2.4%	2.0%	
Nigeria		0.8%								
Norway	3.3%	3.7%								
Pakistan	0.9%									
Philippines	1.1%									
Poland	1.0%	0.5%							1.2%	
Portugal	1.70%									
Romania	0.6%	1.4%								
Russia	0.9%							67.8%	46.5%	48.4%
Saudi Arabia	2.1%	0.6%								
Singapore		0.6%								
Slovenia	0.9%									
Spain	0.6%	2.0%							1.8%	
Sweden	5.5%	1.1%							1.6%	
Taiwan				1.1%	0.9%					
Turkey	1.8%	0.6%								
Ukraine								12.6%	5.9%	6.4%
United Kingdom	7.8%	10.6%							1.5%	
United States	14.7%	28.9%	28.2%	0.5%		8.8%	1.0%		5.6%	
Uzbekistan									2.3%	

Note: We obtained all visitor data from Alexa. Each entry is the percentage of visitors from the corresponding country. We have no visitor data for HBH and Xaker in May 2015. Because Alexa does not publish historical statistics, we cannot obtain the visitor data in 2007–2011.

Table 3. Descriptive Statistics and Correlations

Variable	Mean	Std. dev.	Min	Max	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
(1) Number of victim IPs	61.417	151.220	0	93,546		0.015	0.105	0.356	0.765	0.382	0.599	0.067	0.175	0.219	(11)									
(2) Number of vulnerabilities	0.004	2.234	0	5	0.008		0.014	0.232	0.007	0.013	0.007	0.001	0.001	0.001	(12)									
(3) Number of threats	0.001	1.040	0	3	0.010	−0.000		0.942	0.056	0.082	0.056	0.009	0.019	0.013	(13)									
(4) DDOS-thread–port-effective posts (Hackforums)	5.541	314.530	0	186	0.068	0.004	0.000		0.249	0.176	0.207	0.025	0.063	0.068	(14)									
(5) DDOS-thread–port-effective posts (HBH)	0.014	13.795	0	37	0.007	−0.000	0.026	0.003		0.127	0.506	0.066	0.107	0.087	(15)									
(6) DDOS-thread–port-effective posts (Hackbase)	0.168	36.283	0	43	0.009	0.001	−0.000	0.020	0.003		0.178	0.019	0.125	0.057	(16)									
(7) DDOS-thread–port-effective posts (HHLM)	1.182	260.762	0	187	0.008	0.000	0.000	0.013	0.001	0.020		0.044	0.113	0.190	(17)									
(8) DDOS-thread–port-effective posts (Antichat)	121.928	1,116.805	0	893	0.063	0.004	0.002	0.104	0.004	0.022	0.011		0.017	0.101	(18)									
(9) DDOS-thread–port-effective posts (Xaker)	0.063	21.550	0	33	0.004	−0.000	−0.000	0.041	−0.000	0.004	0.001	0.012		0.369	(19)									
(10) Non-DDOS-thread–port-effective posts (Hackforums)	8.402	260.051	0	165	0.067	0.002	0.000	0.305	0.004	0.022	0.015	0.106	0.065		(10)									
(11) DDOS-thread–port-effective posts (English)	5.555	314.877	0	186	0.069	0.004	0.001	0.999	0.047	0.020	0.013	0.104	0.041	0.305										
(12) DDOS-thread–port-effective posts (Chinese)	1.350	263.990	0	187	0.009	0.000	0.000	0.015	0.001	0.157	0.991	0.014	0.002	0.017										
(13) DDOS-thread–port-effective posts (Russian)	121.991	1,117.278	0	893	0.064	0.004	0.002	0.105	0.004	0.022	0.011	1.000	0.032	0.107										
(14) DDOS-thread–port-effective posts (All)	128.897	1,225.354	0	893	0.078	0.004	0.002	0.355	0.016	0.059	0.227	0.941	0.040	0.180										
(15) Botnet mention	0.856	79.376	0	70	0.051	0.003	−0.000	0.766	0.002	0.008	0.006	0.056	0.021	0.178										
(16) Botnet: Post duration	100.270	7,479.931	0	10,300	0.043	0.002	0.000	0.383	0.003	0.013	0.011	0.081	0.030	0.229										
(17) Botnet: Number of bot posts	364.187	30,905.800	0	21,600	0.048	0.002	0.000	0.599	0.002	0.006	0.006	0.056	0.017	0.166										
(18) New DDOS bot mention	0.002	1.873	0	6	0.006	−0.000	−0.000	0.068	−0.000	0.002	0.001	0.008	0.042	0.031										
(19) New bots: Duration	9.261	3,015.712	0	4,625	0.014	−0.000	−0.000	0.175	0.000	0.004	0.001	0.019	0.032	0.083										
(20) New bots: Number of new bots posts	0.406	176.404	0	695	0.006	−0.000	−0.000	0.220	0.000	0.002	0.000	0.012	0.045	0.048										

Note: Except for number of victim IPs, we multiply the means and standard deviations of all other variables by 1,000 to enhance the readability of the numbers. *N* = 64,731,700 observations for all variables.

Table 4. Panel Fixed-Effects Regression Results

Variables	(1) OLS	(2) IV Estimator	(3) DShield: Port posts	(4) DShield: DDOS posts	(5) Randomized Ports
Lagged number of victim IPs	0.579***	0.579***	0.106***	0.106***	0.579***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
Number of threats	0.331***	0.291***	−0.039	−0.000	0.345***
	(0.107)	(0.104)	(0.255)	(0.248)	(0.109)
Number of vulnerabilities	0.104***	0.094***	0.125	0.136	0.108***
	(0.027)	(0.026)	(0.145)	(0.144)	(0.028)
Number of DDOS-attack posts	−0.032***	−0.122***		−0.006	
	(0.006)	(0.029)		(0.026)	
Effective port posts			−0.064***		
			(0.018)		
Placebo					−0.000
					(0.002)
Port fixed effects	Yes	Yes	Yes	Yes	Yes
Day fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	64,696,250	64,696,250	59,343,300	59,343,300	64,696,250
Adjusted R^2	0.979	0.979	0.607	0.607	0.979
Number of port	35,450	35,450	35,450	35,450	35,450

Notes: Column (1): Baseline estimate. Column (2): 2SLS estimation with the number of non-DDOS-thread–port-effective posts as the IV. Column (3): Use the number of target IPs in the DShield intrusion data as the dependent variable and all port-effective posts to measure forum discussion. Column (4): Use the number of target IPs in the DShield intrusion data as the dependent variable and the number of DDOS-thread–port-effective posts to measure forum discussion. Column (5): Randomly match the forum discussion with the number of DDOS-attack victims. Robust standard errors clustered by port in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 5. Content Analysis

Variables	(1) 4-topic model	(2) Botnet: Number of posts	(3) Botnet: Duration	(4) New botnet: Number of posts	(5) New botnet: Duration
Lagged number of victim IPs	0.579***	0.579***	0.579***	0.579***	0.579***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number of threats	0.332***	0.334***	0.331***	0.331***	0.330***
	(0.107)	(0.108)	(0.107)	(0.107)	(0.107)
Number of vulnerabilities	0.105***	0.105***	0.104***	0.104***	0.104***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Number of DDOS-attack posts		−0.044***	−0.024***	−0.031***	−0.031***
		(0.009)	(0.006)	(0.006)	(0.006)
Topic 1	0.009***				
	(0.003)				
Topic 2	−0.008***				
	(0.002)				
Topic 3	−0.013***				
	(0.003)				
Topic 4	0.000				
	(0.001)				
Botnet mention		0.040***	0.015*		
		(0.010)	(0.009)		
Botnet follow-ups weighted by botnet posts		0.000			
		(0.003)			
Botnet follow-ups weighted by duration			−0.008***		
			(0.002)		
New botnet mention				0.132***	0.094**
				(0.040)	(0.039)
New botnet follow-ups weighted by new botnet posts				−0.030**	
				(0.014)	
New botnet follow-ups weighted by duration					−0.014**
					(0.006)
Port fixed effects	Yes	Yes	Yes	Yes	Yes
Day fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	64,696,250	64,696,250	64,696,250	64,696,250	64,696,250
Adjusted R^2	0.979	0.979	0.979	0.979	0.979
Number of port	35,450	35,450	35,450	35,450	35,450

Notes: Column (1): Include four LDA topics. Column (2): Add botnet mention and follow-ups weighted by total number of botnet posts in the same thread. Column (3): Add botnet mention and follow-ups weighted by duration since the first botnet mention. Column (4): Add Mariposa botnet and Zbot mention and follow-ups weighted by total number of Mariposa botnet and Zbot posts in the same thread. Column (5): Add Mariposa botnet and Zbot mention and follow-ups weighted by duration since the first Mariposa botnet or Zbot mention. Robust standard errors clustered by port in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 6. Results with the Discussion from Other Forums

Variables	(1) Add Other forums	(2) Group by language	(3) Group all forums
Lagged number of victim IPs	0.579***	0.579***	0.579***
	(0.002)	(0.002)	(0.002)
Number of threats	0.326***	0.332***	0.342***
	(0.105)	(0.107)	(0.109)
Number of vulnerabilities	0.104***	0.104***	0.108***
	(0.027)	(0.027)	(0.028)
Hackforums posts	−0.032***		
	(0.006)		
HBH posts	0.038		
	(0.047)		
Hackbase posts	0.009		
	(0.011)		
HHLM posts	0.003		
	(0.003)		
Antichat posts	−0.005***		
	(0.001)		
Xaker posts	−0.020		
	(0.024)		
All English posts		−0.032***	
		(0.006)	
All Chinese posts		0.003	
		(0.003)	
All Russian posts		−0.005***	
		(0.001)	
All Forum posts			−0.006***
			(0.001)
Port fixed effects	Yes	Yes	Yes
Day fixed effects	Yes	Yes	Yes
Observations	64,696,250	64,696,250	64,696,250
Adjusted R^2	0.979	0.979	0.979
Number of port	35,450	35,450	35,450

Notes: Column (1): All forum discussion included as different independent variables. Column (2): Forum discussion grouped by language. Column (3): Group all forum discussion into one variable. Robust standard errors clustered by port in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

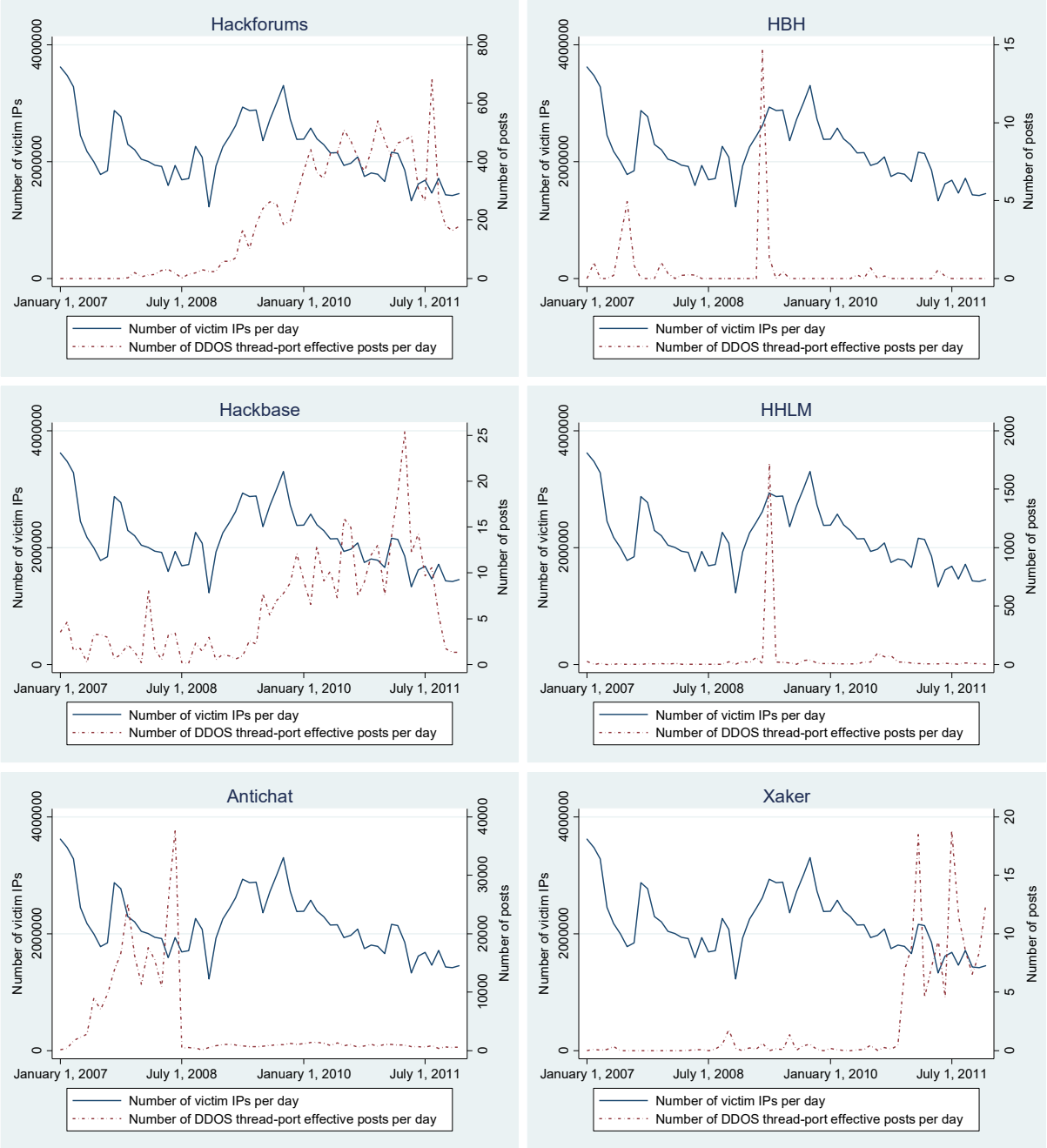


Figure 1. DDOS-Attack Victims and Forum Discussion over Time

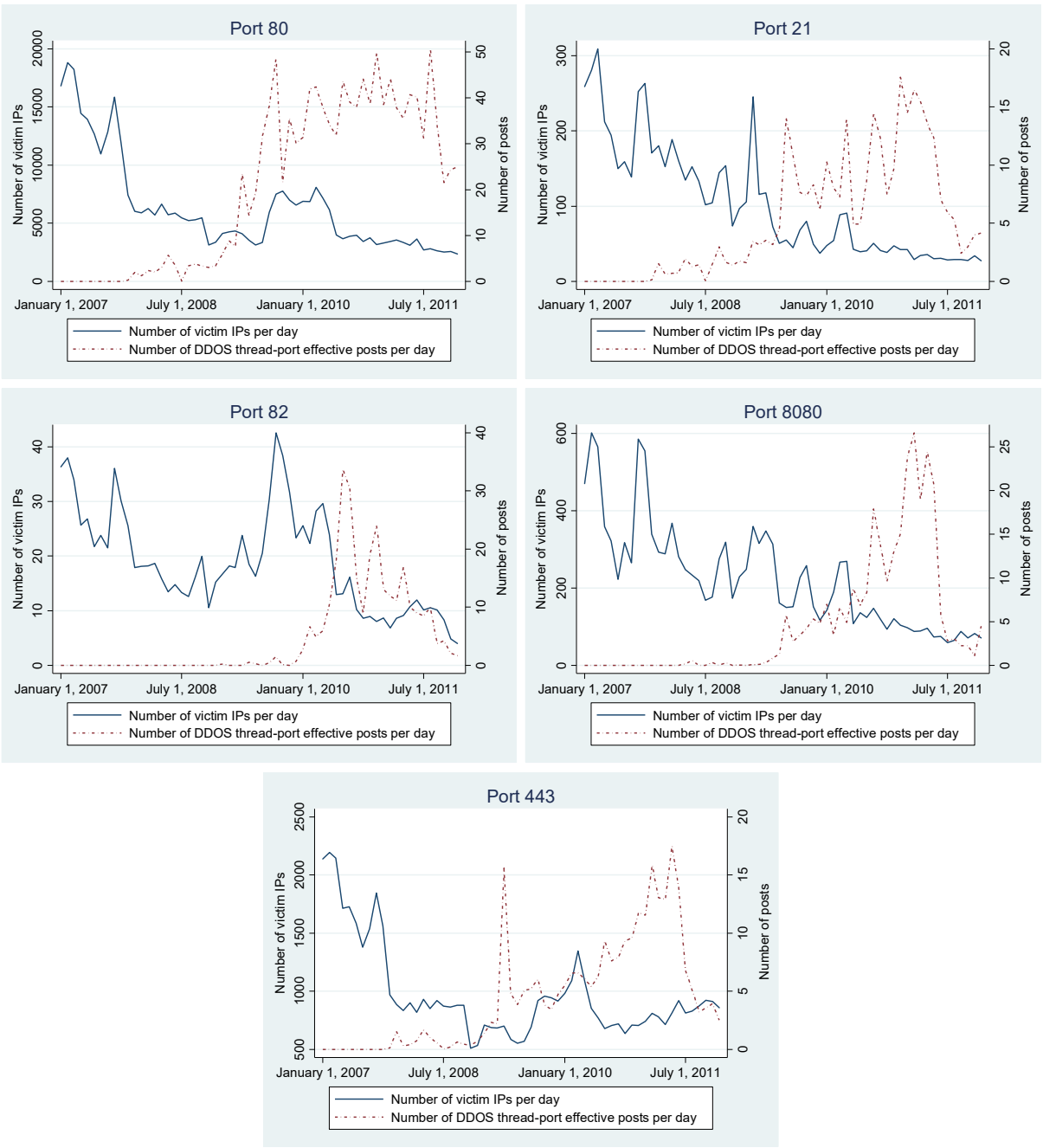
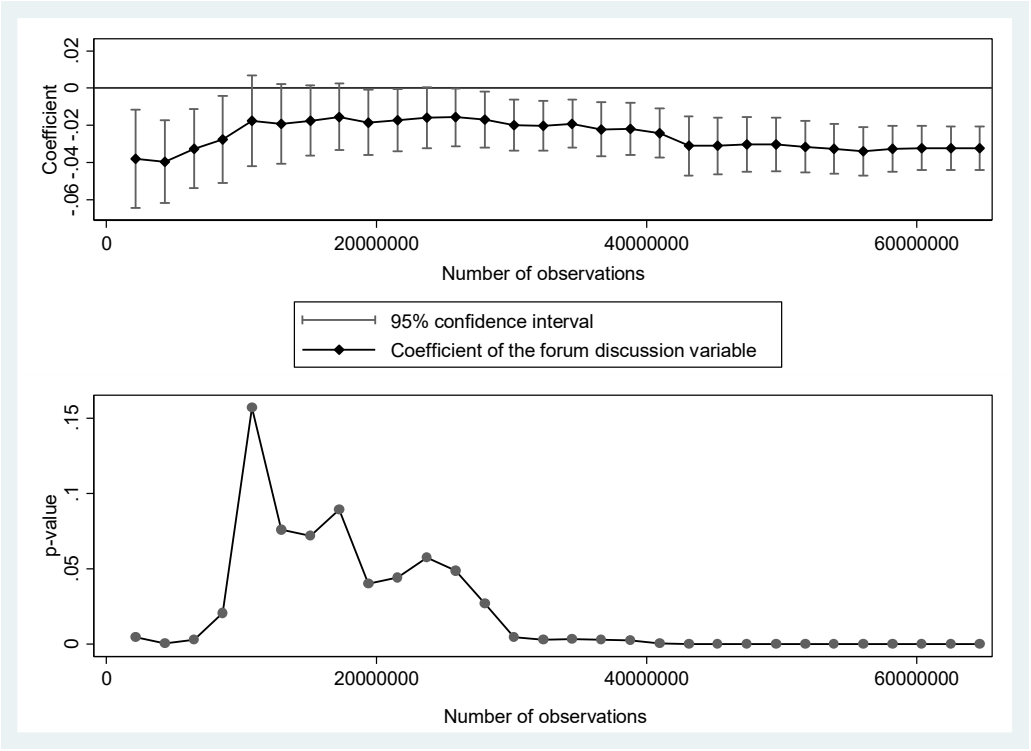
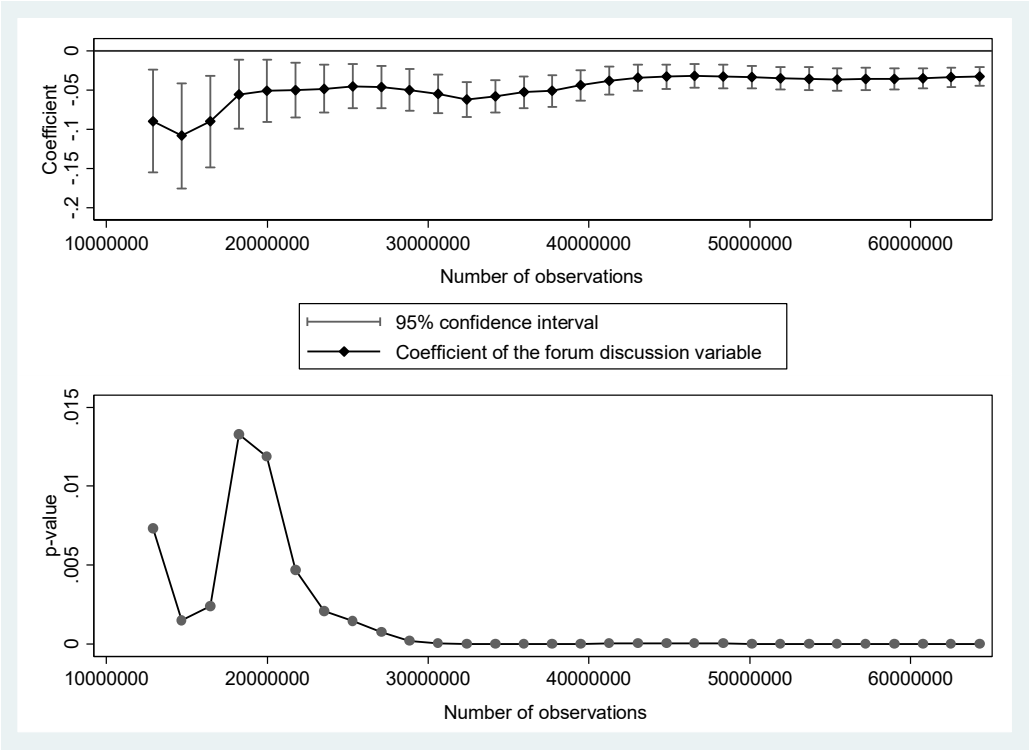


Figure 2. DDOS-Attack Victims and Forum Discussion by Port



(a) Plots with Subsamples Varying by Number of Ports



(b) Plots with Subsamples Varying by Number of Days

Figure 3. CPS Charts

Topic ID	Number of posts	Keywords (intensity of the shade reflects keyword weights)									
1	6,861	link	download	plea	file	send	password	version	viru	sourc	upload
		updat	code	plz	bot	backdoor	messag	remov	detect	add	compil
2	40,247	nice	work	great	post	tutori	tut	man	thank	thread	one
		hack	now	know	use	much	share	ing	keep	well	plea
3	2,679	http	error	file	includ	foundhttp	sql	result	warn	program	vulner
		found	inject	admin	open	invalid	php	commandsadd	platform	miss	print
4	33,047	use	port	server	bot	work	ip	open	connect	know	run
		one	want	make	host	comput	program	don	find	set	see

Figure 4. Top 20 Keywords in the Four-Topic LDA Model