How Sudden Censorship Can Increase Access to Information

WILLIAM R. HOBBS   *Northeastern University*
MARGARET E. ROBERTS   *University of California, San Diego*

*C*onventional wisdom assumes that increased censorship will strictly decrease access to information. We delineate circumstances when increases in censorship expand access to information for a substantial subset of the population. When governments suddenly impose censorship on previously uncensored information, citizens accustomed to acquiring this information will be incentivized to learn methods of censorship evasion. These evasion tools provide continued access to the newly blocked information—and also extend users' ability to access information that has long been censored. We illustrate this phenomenon using millions of individual-level actions of social media users in China before and after the block of Instagram. We show that the block inspired millions of Chinese users to acquire virtual private networks, and that these users subsequently joined censored websites like Twitter and Facebook. Despite initially being apolitical, these new users began browsing blocked political pages on Wikipedia, following Chinese political activists on Twitter, and discussing highly politicized topics such as opposition protests in Hong Kong.

Scholars have studied censorship by assuming it has an exogenous impact on a passive citizenry. Conventional wisdom posits that censorship should lower the probability that citizens access information (Morozov 2011; Lessig 1999). Not surprisingly, then, authoritarian governments appear to be tightening their grip on information environments, increasing their use of search filtering, content removal, and website blocking (Deibert et al. 2010; Kelly, Cook, and Truong 2012; Shirk 2011), along with rapid online censorship during large-scale collective action events (King, Pan, and Roberts 2013, 2014).

Here, we show that information environments more realistically function like ecosystems. They are sufficiently complex that censorship can inadvertently increase information access for some while reducing it for others. Because some citizens quickly adapt to censorship, the imposition of restrictions can have unexpected consequences by creating incentives for censorship circumvention. In certain circumstances, sudden censorship can even result in the opposite of the intended effect: an *increase* in access to off-limits information among people motivated by the new censorship to seek out avenues for evasion.

We show that censorship can increase access to information when a government extends censorship to previously uncensored media, especially when individuals have no ready alternatives to replace it. When favorite and difficult-to-replace media are suddenly blocked, those who would otherwise not take the time to evade censorship restrictions will have a stronger motive to learn how to continue to access the newly censored information. Because censorship evading technology (and, more broadly, social networks) that assist in evasion are rarely specific to particular information sources, acquisition of these tools and networks can give users unhindered access to *many* country-specific blocked sources. Although many users might learn censorship evasion only to maintain access to a preferred source, censorship evasion incidentally expands the set of information these users can easily access. We name this phenomenon a "gateway effect" as the motivation to access the newly censored information provides a gateway into access to information that has long been censored or blocked.[1] While many others have made the point that repression leads to strategic responses from dissenters (Tilly 1978; Ritter 2014; Francisco 2005; Pierskalla 2010), this article shows a mechanism through which repression can backfire inadvertently, without political or strategic motivation.

We illustrate this phenomenon by bringing to bear one of the first large, detailed, and individual-level datasets of the real-time censorship evasion behavior of internet users under policies of increased censorship. Using social media data from websites that are blocked by the Great Firewall in China, we measure the number and types of people who use virtual private networks (VPNs) to "jump" the Great Firewall to access information and networks blocked by the Chinese government. We show how increases in censorship through

[1] A formal model of this mechanism appears in the Supplementary Material to this paper.

blocking of the popular social networking website Instagram in China disrupted the habits of millions of individuals accustomed to visiting that site and increased evasion of the Great Firewall. In doing so, censorship of Instagram created a sudden increase in the number of people from China accessing websites and networks such as Twitter and Facebook that the Chinese government has blocked for many years. We present evidence that previously apathetic Chinese social media users began to engage in topics related to ongoing protest events and with people in Chinese activist circles shortly after the Instagram block.

Our evidence suggests a mechanism that could explain previous work that found a positive correlation between censorship and the likelihood of resistance against the government (Hassanpour 2014; Nabi 2014; Moore 1998). However, importantly, the "gateway effect" we explore does not rely on backlash against censorship, which has commonly explained this phenomenon in the past. Our mechanism also does not require pre-existing or nascent political interest on the part of the censored individuals. As governments expand the types of information that are off limits, they can simply create more incentives for individuals to engage in networks and technologies that allow them to access the media they are accustomed to consuming, and these technologies, in turn, allow them to access to information they did not have access to before.

To be clear, our findings do not suggest that censorship is ineffective in general or that increases in censorship will *always* backfire against governments. Our evidence suggests that, through the gateway effect, there are circumstances where censorship *can* increase access to information for a large subset of individuals. As we explore in our discussion, our findings suggest that gateway effects are most likely when increases in censorship are sudden and blunt—when applied during emerging protests or crises, for example.[2] By motivating more people to acquire the ability to evade censorship, a sudden increase in censorship can erode its own effectiveness, can politicize previously apolitical citizens, and can accumulate collective action potential that it often seeks to suppress.

The paper proceeds as follows. First, we describe the previous literature on the effects of censorship and introduce the logic behind how censorship can increase access to information and politicize previously apolitical users. Second, we describe our research design and the case we use to test our theory—the sudden block of Instagram in China in September 2014. Using our data, we describe the types of users who evade censorship before the sudden block of Instagram. Then, we show how the Instagram block increased access to information for the majority of Instagram users. We end with implications of this research, including a description of the scope conditions and circumstances under which censorship might lead to the gateway effects that we describe.

---

[2] Censorship might be more sudden and thus less effective when collective action events are decentralized and spontaneous, and therefore not anticipated by the government (Steinert-Threlkeld 2017).

CENSORSHIP AND ACCESS TO INFORMATION

A large literature has contributed to our knowledge of how censorship influences the political views and behavior of its intended targets. Most of the literature finds that censorship generally decreases access to information among citizens. Scholars have shown that by reducing connectivity to information sources, censorship and repression can successfully restrict information and demobilize individuals from engaging in the behavior the government deems objectionable (Enikolopov, Petrova, and Zhuravskaya 2011; Pierskalla 2013; Edmond 2013). Others have found that when media is restricted in authoritarian governments, citizens access what is readily available to them (Stockmann 2012; Geddes and Zaller 1989; Chen and Yang 2017) and do not discuss alternative information with each other because they do not know what others believe (Kuran 1997). Roberts (2018) shows that because consumers of media are impatient, even small increases in the price of information imposed by censorship can have large negative effects on information consumption, particularly when censorship goes unnoticed.

Government control of the media can also provide signals to citizens to follow the government line. Huang (2015) finds that propaganda in China can signal government strength and Stern and Hassid (2012) find that the ambiguous nature of off-limits information can induce risk-averse journalists to self-censor instead of spreading information. Censorship may even have adverse effects on information outside of a country by reducing incentives for citizens of other countries to engage in social media without a wider audience (Zhang and Zhu 2011).

Even the internet, which many scholars thought would be a form of "liberation technology" and would help with coordinating protests (Lynch 2011; Diamond 2010; Howard 2010; Shirky 2008; Tufekci and Wilson 2012; González-Bailón et al. 2011), is thought be be highly affected by government censorship. "Repression technology" on the internet, in the form of search filtering, content filtering, or blocks from firewalls, are thought to effectively decrease access to information in authoritarian regimes (Kalathil and Boas 2010; Morozov 2011; MacKinnon 2012; Rød and Weidmann 2015). Even when autocrats facilitate access to information, scholars have found that they do so strategically, for information-gathering purposes, indicating that many autocrats engage in sophisticated and calculated censorship (Lorentzen 2014; Egorov, Guriev, and Sonin 2009).

A few empirical studies have provided evidence to the contrary, suggesting that sometimes disruption of the media seems to have the opposite effect, stimulating increased public engagement or providing opportunities for the opposition (Rasler 1996; Hassanpour 2014). Explanations for why censorship may backfire often concentrate on political and social evaluations of increased censorship, including signaling and psychological models of reactance. In popular culture, this type of backlash against censorship has

commonly been known as the "Streisand effect,"[3] which occurs when censored information increases in perceived value because the censorship efforts unintentionally publicize it. For example, when Wikimedia, a Wikipedia-like website in France, was allegedly asked by France's intelligence agency to remove part of an article, the report of the attempted censorship launched an international campaign against it, and the page ended up becoming the most-viewed page on the website.[4] In this case, censorship itself caused political backlash that attracted attention to the information the government had hoped to make less accessible.

In this vein, scholars have argued that increases in censorship or sudden media disruptions may at times anger a broader population that opposes censorship. Evidence of Streisand effects around the world have led scholars to categorically call censorship futile (Nabi 2014) and are a potential explanation for some instances of censorship backfire (Jansen and Martin 2003). In some cases, observable censorship might signal regime weakness (Roberts 2018), potentially creating common knowledge that has been shown to co-ordinate collective action (Kuran 1991; Lohmann 1994; Chwe 2001). In other cases, media disruption removes regime-solidifying distractions, decentralizing information sharing to face-to-face interactions that can accelerate collective action (Hassanpour 2014; Kern and Hainmueller 2009). Scholars studying contentious politics have long noted that government repression can be countered by strategic calculations of dissidents, who can use such repression as a rallying cry for increased organization (Tilly 1978; Ritter 2014; Francisco 2005; Pierskalla 2010).[5]

THEORY: GATEWAY EFFECTS IN INFORMATION ACCESS

For the most part, the literature has posited that censorship acts to restrict information access and that citizens are unlikely to have any recourse. In the cases where evidence of censorship backfire has existed, it has relied on emotional and political backlash where citizens access the information because they are aware of government efforts to suppress it.[6] This often requires a well-organized opposition that can make the public aware of government censorship to bolster support for the cause or sabotage censorship efforts (Jansen and Martin 2003; Hess and Martin 2006; Yang 2014). Here,

we suggest that there may be more circumstances outside of backlash where censorship does not produce its intended effects. In particular, we posit that censorship has the capability to politicize and empower previously apathetic citizens without a well-organized effort. In this section, we describe the logic behind these circumstances and refer readers to a more explicit model of the mechanism in the Supplementary Material.

We propose a "gateway effect" mechanism to explain how sudden censorship can increase information access. This gateway effect occurs when, because of censorship, citizens seek out the *same* information that they previously had access to by acquiring tools for censorship evasion—networks or technologies that enable them to bypass government restrictions on information. These citizens then gradually become exposed to other off-limits information, as this information is suddenly easy for them to access.

Individuals might not intend to seek out political information—and they do not necessarily begin to seek out off-limits information to arm themselves against the government. The gateway effect is driven by habit: citizens are accustomed to accessing the newly censored information and would like to continue to do so. Information that is indispensable or addictive will produce stronger gateway effects when censored than that which can be easily substituted.

The gateway effect will impact a portion, but typically not all of the population. There will be some citizens who are not sufficiently motivated to spend time and energy finding the tools to evade censorship to continue their habit. For citizens who do not seek out evasion tools, censorship will strictly decrease their access to information. However, those who are motivated enough by habit and capable enough to find ways to evade censorship will be driven by sudden censorship to find tools to continue to consume the newly censored information. This, in turn, will incidentally expose them to a range of other off-limits information, increasing the range of information that they have access to.

This access will accumulate over time, unless the government continually adjusts its behavior to counter that cumulative effect. Sudden censorship will arm citizens with tools for bypassing government restrictions as they seek these tools out to access newly censored information. Once censorship evasion tools are acquired and learned, they can be used more easily in all situations. Censoring more information sources further increases the total benefits of evasion. Evasion technology will grant access to a broader array of off-limits media with each censorship event.

Consider a few examples of this gateway effect. Citizens might be habitual readers of a particular author, whose books are suddenly banned by the government. Loyal readers of the author may then seek out black market bookstores to continue following the author. These bookstores, in turn, may become a gateway to a list of other off-limits titles, including those on subjects the individual may have heard about, but never had the opportunity to buy in the past. Or, citizens might participate in a religious organization that is suddenly

---

[3] Named for Barbra Streisand, who, in an attempt to have pictures of her home in Malibu, California, removed from the internet, instead brought more attention to them.

[4] "What Is the Streisand Effect?" *The Economist*, April 16, 2013, http://www.economist.com/blogs/economist-explains/2013/04/economist-explains-what-streisand-effect.

[5] Streisand-effect arguments are similar to that in some of the repression literature, where government violence can cause political backlash. See Davenport (2007) for an overview.

[6] There is also a literature about how censorship may not be as effective as authoritarians would like it to be because authoritarian governments struggle to implement it, see for example Yang (2013); Miller (2017); Lorentzen (2014).

banned by the government, forcing the organization underground. To continue worshiping, the religious organization may then meet in time periods or places that are less likely to be detected by the government, and in doing so, may be exposed to individuals from other organizations who are using the same strategies to evade government censure. In the case that we describe in the empirical section, the popular social media website Instagram was suddenly blocked by the Chinese government. Habitual users of Instagram then sought out software to evade the Great Firewall to continue to access Instagram. This, in turn, allowed them easy access to long-blocked websites such as Facebook and Twitter and political information on Wikipedia, which they signed up for and viewed soon after acquiring evasion software.

Gateway effects, of course, are not specific to government censorship. For example, changes in alcohol consumption, drug use, and prostitution during the Prohibition Era in the United States may have displayed this dynamic. During the Prohibition era, alcohol consumption was approximately 60% to 70% of pre-Prohibition consumption (Miron and Zwiebel 1991). In other words, many Americans kept drinking. Because underground alcohol distributors were also engaged in other black markets, Prohibition may have given habitual drinkers ready access to many other illicit goods and services. More recently, the Netherlands decided to permit the sale of cannabis in coffee shops precisely because policymakers were worried that gateway effects might occur when users of cannabis seek out markets and networks that also sell harder drugs (MacCoun 2010).

The logic of this proposed mechanism contrasts with that of a Streisand effect because it does not require backlash against the censorship itself, whether that backlash is the product of anger or mere curiosity to see what was blocked. In a Streisand effect, overall access to the newly censored information should increase, as citizens become more interested in the information *because of* censorship. In contrast, in the gateway effect, overall access to the newly censored information decreases since it has become more difficult to access—users do not rally around the censored information and gateway effects do not create new interest in the newly censored information. Instead, access to *other* information that has long been banned increases incidentally and even without political motivation among those who seek out new technologies or networks to access the information they are in the habit of consuming. The logic of the gateway effect is a more general explanation of what has been known as the "cute cat" theory of censorship (Zuckerman 2015), which posits that entertainment, and not politics, drives information consumption on the internet and the pairing between politics and entertainment makes censorship more difficult for governments. With the gateway effect, any type of newly censored information (including entertainment) can motivate evasion, and the increase in access is driven by the individual's subsequent exposure to other information that is facilitated by acquiring evasion technology and know-how.

THE GREAT FIREWALL AND GEOLOCATION

In this paper, we study the gateway effect in detail in the context of China's Great Firewall. The Great Firewall of China blocks foreign websites from mainland Chinese IP addresses, preventing Chinese citizens from accessing websites that the government deems objectionable. Mainland Chinese users—importantly, not including Hong Kong users—cannot access a whole host of foreign social media platforms from Twitter to Facebook to Google.[7]

To access these websites from China, a user must first log in to a computer outside of China using a VPN or a proxy. VPNs are not terribly difficult to acquire for a sophisticated user and as of writing citizens are not sanctioned for using them. However, the small costs in terms of time and money for Chinese users in combination with the availability of attractive alternative social media websites like Sina Weibo (the Chinese version of Twitter) and WeChat (the Chinese version of WhatsApp) mean that relatively few users regularly "jump" (evade) the Wall.

Those who do use a VPN to evade censorship can obtain full access to blocked social media websites and can use the internet without restrictions. Like their counterparts around the world, these users will sometimes use geolocation to "tag" themselves at a particular location. When users tag themselves in mainland China on blocked websites like Instagram or Twitter to specific locations, we know that they are using evasion technology to access these social media websites. We can estimate the dynamics of evasion of the Great Firewall by examining how the number of Instagram and Twitter posts geotagged to China changes over time and comparing these rates to locations that are not subject to the firewall, like Hong Kong. Of course, geolocated users are not a random sample of the entire population of users. To ensure that our findings are generalizable, we supplement the geolocated social media data with data from sources that do not rely on geolocation, such as application download statistics, Chinese-language Wikipedia page views, and accumulation of followers of Chinese-language Twitter accounts, which we describe in detail in the next section.

DATA

We estimate the dynamics of the gateway effect in response to the sudden censorship of Instagram in China, which occurred on September 29, 2014. Instagram was allegedly blocked in China in reaction to the protests in Hong Kong, which began on September 26, 2014.[8] We estimate the gateway effect by relying on a variety of

---

[7] Greatfire.org, which monitors the constantly updating group of blocked sites from China, estimates that 131 of 937 Alexa top 1000 domains are blocked in China, as of May 16, 2016, including media sites such as the *New York Times*, the *Wall Street Journal*, *Reuters*, and *Bloomberg* to social media sites like Blogspot, Twitter, Facebook, and YouTube.

[8] Park, Madison, "China's Internet Firewall Censors Hong Kong Protest News," CNN, September 30, 2014, http://www.cnn.com/2014/09/29/world/asia/china-censorship-hong-kong/

datasets. While some of these datasets necessarily rely on found and geotagged data that do not represent all users of any particular platform (Salganik 2017), we are more confident in our results because we observe the gateway effect in such a wide variety of data obtained from different sources.[9]

We first collect a sample of 419,279 geolocated Instagram posts from mainland China, between September 1, 2014 and October 30, 2014. To obtain a representative sample, we randomly sampled coordinate grids in mainland China proportional to their population and downloaded all geolocated Instagram posts for the time period with each sampled grid.[10] Instagram was not blocked by the Great Firewall until September 29, 2014, when it was suddenly blocked during the third day of large-scale protests in Hong Kong.[11] This allows for a convenient discontinuity to estimate what proportion of users persisted in accessing Instagram even after it was blocked.

Second, we use application download statistics from the app tracker App Annie to directly measure how much the Instagram block spurred new downloads of censorship-evasion technology. App Annie tracks the relative rank of application downloads on iPhones by country. We use this data source to measure the popularity of new downloads of VPN applications in China around the time period of the Instagram block.

Third, we study how the Instagram block influenced the popularity of websites that have been blocked in China for an extended period of time. We use App Annie to measure new application downloads that direct to blocked websites, such as Facebook and Twitter. In addition, we collect 2,321,861 geolocated Twitter posts from mainland China, beginning in March of 2014, using the Twitter garden hose, approximately 1% of all Tweets.[12] Twitter has been blocked by the Great Firewall since 2009, therefore users whose posts are geolocated to mainland China are using evasion software to access the website. We use this dataset to study new users who sign up for Twitter from China around the time period of the Instagram block. For comparison, we use 1,773,678 geolocated Twitter posts from Hong Kong over the same time period.[13] Hong Kong is not subject to the constraints of the Great Firewall, but is similar to China culturally and linguistically.

Fourth, we collect a sample of 238,918 geolocated social media posts from Sina Weibo, the Chinese version of Twitter, as an additional comparison set. We use these data in the subsequent section to compare the types of users who are savvy and interested enough to participate in Twitter by evasion to those who participate in the more accessible Sina Weibo.

Fifth, we compare page views of blocked Chinese-language Wikipedia pages before and after the Instagram block using Wikipedia page-view data.[14] We include an analysis of which of the blocked Chinese language Wikipedia pages had surges in popularity directly following the block.

Last, we estimate how the Instagram block affected the accumulation of new followers of Chinese-language accounts on Twitter. To assess the block's long-term impact, we also explore the amount that sensitive political conversations occurring years after the block are discussed by those who signed up for Twitter on the day of the Instagram block. Both the Wikipedia and Twitter data do not rely on geolocation and directly measure whether the block increased access to off-limits information through page views and Twitter follows.

RESEARCH DESIGN

In this section, we describe the empirical setup of our argument that increased censorship can lead to expanded access to information by disrupting users' habitual behavior. Our empirics will demonstrate that Chinese users will begin to access blocked information (and social networks) because a website that they habitually access becomes suddenly censored, inspiring them to acquire evasion technology. We demonstrate this finding with four sequential tests, summarized in Table 1. First, we estimate the proportion of users who stay on Instagram after the block by examining the geolocated Instagram data. Because the block completely prevented Chinese internet users from accessing Instagram, we can be confident that anyone accessing Instagram from China is doing so through a VPN. Second, we test whether the block inspired the acquisition of censorship technology by examining download ranks for popular VPNs from Chinese iPhone users. Third, we test whether the block inspired the use of websites that have long been blocked by the Chinese government by measuring Facebook and Twitter application downloads and directly measuring activity of Chinese social media users on Twitter. Last, we estimate whether the block resulted in users engaging more in political information by estimating how the participation of new users affected the popularity of blocked media and activists

---

[9] In the Supplementary Material, we provide a bounding analysis to account for potential bias resulting from geolocation.

[10] We sampled approximately 25% of the grid by population (i.e., the sampled grid coordinates accounted for approximately 25% of the population of China). Our estimates suggest we successfully sampled a smaller fraction—15% of the population. A user's Instagram post can appear in our data if it was posted publicly with geolocation in mainland China during September or October 2014.

[11] To our knowledge, no other websites were suddenly blocked that day if they had not been blocked before.

[12] This was part of a broader collection of worldwide tweets. Using the Twitter API, all worldwide geolocated tweets were requested every hour starting in March 2014. Because the garden hose only allows for a total of 1% of worldwide (geolocated or not geolocated) tweets to be collected, if geolocated tweets were greater than 1% of all tweets that hour, the request returns a random sample of all geolocated tweets up to 1%. Scholars have estimated that geolocated tweets are approximately 2–3% of the whole Twitter sample (Leetaru et al. 2013). This means that our collection contains approximately one-half to one-third of all geolocated Tweets, and approximately 1% of tweets in China.

[13] These posts were collected in the same worldwide API call as the mainland China tweets.

[14] Page-view data at https://dumps.wikimedia.org/other/pagecounts-raw.

TABLE 1. Empirical Tests

Question	Test
1. Do users evade censorship to access Instagram?	Number/proportion of Instagram users who *continue* to access site after block
2. Does total censorship evasion increase?	VPN application download ranks
3. Do users access other, already blocked information?	Facebook/Twitter application download ranks, active Twitter users, Twitter sign-ups, Wikipedia page views
4. Do users absorb and engage with blocked political information?	Mentions of Hong Kong protests on Twitter, number of followers of Twitter activists and media, lasting discussions of political topics

---

on Twitter, measuring changes in access to off-limits information on Wikipedia, and testing whether the Instagram users began engaging in political conversation.

WHO TYPICALLY EVADES CENSORSHIP?

Before we analyze how user behavior is affected by the sudden increase in censorship, we use a comparison of geolocated users on Twitter and geolocated users on Sina Weibo, the Chinese version of Twitter, to describe the profile of the typical Chinese user who evades censorship before the Instagram block. Those who are already using a VPN to jump the firewall at the time of the block would not be affected by the block as they already have access to off-limits information. We include this section to 1) describe the static efficacy of the Great Firewall and 2) describe the profiles of those who are already evading censorship to show what types of social networks new individuals would be introduced to.

Consistent with previous research, the data on Twitter users in China before the Instagram block reveals that very few individuals regularly evade the Great Firewall of China and the firewall practically prevents many users from accessing information blocked by the wall. Previous studies of censorship evasion in China have estimated that 3%–15% of Chinese internet users engage in censorship evasion.[15] Consistent with this research, we find that very few users post from China on Twitter, particularly in comparison to similar locations where Twitter is not censored. In our sample of geolocated Twitter posts, on average about 1,600 unique Twitter users geolocate to mainland China every day. Since we believe our sample covers approximately 1% of all Twitter posts from China, we expect that there are around 160,000 Twitter users who tweet each day from mainland China, or about 0.026% of all internet users in China.[16] In comparison, on average, there are 900

unique Twitter users every day in our sample geolocated to Hong Kong. This suggests that around 90,000 Twitter users tweet each day from Hong Kong, or about 1.7% of internet users in Hong Kong.[17] This rough, back-of-the-envelope calculation suggests that Hong Kong has around 65 times more Twitter users per internet user than China does, despite having similar language, culture, alternative social media sites like Sina Weibo and Wechat, and political reasons to join Twitter. The Hong Kong-China Twitter comparison suggests that the small costs of evasion of the firewall are generally effective in keeping mainland users off of the blocked website, a finding that is largely consistent with previous research (Song, Faris, and Kelly 2015; Chen and Yang 2017).

What types of people in China evade censorship to get on Twitter? Not surprisingly, those using Twitter from mainland China are much more likely to speak a language other than Chinese than those who are using Sina Weibo. When users sign up for a social media account, they can indicate the language they would like their social media account to appear in. Of the geolocated Twitter users before the Instagram block, only 39% of them indicated that their primary language was Chinese when they signed up for Twitter, and 48% indicated that their primary language was English. In comparison, 99% of Sina Weibo users indicate that their primary language is Chinese, and only 0.05% English. Many of the users of Twitter in China are either using Twitter as ex-pats or have extensive foreign language training and experience abroad. These users are likely to have more resources and reasons to evade censorship by jumping the firewall. In the Supplementary Material, we show that Chinese language Twitter users before the Instagram block are highly clustered in urban areas in comparison to Sina Weibo users.

Twitter and Weibo users are also quite different in terms of the content they post. The differences in content suggest that Chinese Twitter users are more technologically savvy and more interested in politics than the average Chinese user on Weibo. In Table 2, we use

---

[15] See Roberts (2018); Faris, Robert, John Palfrey, Ethan Zuckerman, Hal Roberts, and Jillian York, "2010 Circumvention Tool Usage Report," October 14, 210, https://cyber.law.harvard.edu/publications/2010/Circumvention_Tool_Usage; Mander, Jason. "90 Million VPN Users in China Have Accessed Restricted Social Networks," *Global Web Index*, November 24, 2014, http://www.globalwebindex.net/blog/vpn-in-china

[16] We estimate that there are 610,650,000 internet users in China, approximately 45% of the Chinese population (http://www.internetlivestats.com/internet-users/china). If the late 2014 Twitter population was close to 2016 estimates (Russell, Jon, "Twitter Estimates

it has 10 million users in China," *Tech Crunch*, July 5, 2016, https://techcrunch.com/2016/07/05/twitter-estimates-that-it-has-10-million-users-in-china/), then this corresponds to 2% of Chinese Twitter users tweeting any given day.

[17] We estimate that there are 5,254,000 internet users in Hong Kong or about 74% of the population (http://www.internetlivestats.com/internet-users/china-hong-kong-sar).

TABLE 2. Words Most Associated with Twitter and Weibo Users, Mutual Information. This figure shows that Chinese language Twitter users were more likely to use political words than Weibo users.

	Words Associated with Twitter	Words Associated with Weibo
1	im: i'm	自己: myself
2	发布: released	知道: know
3	香港: Hong Kong	今天: today
4	刚刚: just	播放: play
5	照片: photos	生活: life
6	北京: Beijing	努力: work hard
7	good: good	别人: others
8	the: the	看看: watch
9	night: night	mv: mv
10	中心: center	时候: time
11	和平: peace	什么: what
12	apple: Apple	永远: forever
13	day: day	通过: through
14	my: my	快乐: happy
15	国际: international	开心: happy
16	will: will	坚持: persistence
17	公民: citizens	喜欢: like
18	政府: government	没有: none
19	中国: China	事情: situation
20	happy: happy	我们: we
21	人民: the people	明星: stars
22	see: see	生日: birthday
23	自由: freedom	觉得: think
24	国家: country	现在: now
25	like: like	心里: in my heart

mutual information to calculate the words most associated with Twitter posts and most associated with Weibo posts. Mutual information measures the amount of information a word contains about whether the document that the word is in is part of a class, in this case Twitter and Weibo (Manning, Raghavan, and Schütze 2008). We only consider Twitter and Weibo users where the account is registered in Chinese. Even so, Twitter users are much more likely to inject English into their Tweets than those on Sina Weibo—15 out of the 35 words most associated with the Twitter posts are English rather than Chinese. In addition, Twitter users are much more likely to talk about technology—Apple appears in the list of words associated with Twitter. Last, Twitter users are much more likely to talk about politics—political words like international, citizens, government, China, country, freedom, and the politically restive province Xinjiang (not shown, 35th most distinctive) are included in the words associated with Twitter, whereas there are no explicitly political words in the list associated with Weibo.

Consistent with previous research, this comparison suggests that the existence of the firewall has created two different social media communities in China—one for ordinary individuals and one for individuals who find it worth their while to spend time and money to participate in websites that are censored. While the typical citizen participates in domestic social media and might find participating in blocked social media sites inconvenient, those in China who jump the firewall to participate on Twitter live in urban areas, are more technologically savvy, seem to have international language experience, and are relatively likely to discuss politics.

DIRECT EFFECTS OF THE INSTAGRAM BLOCK

What happens to the size and composition of Instagram users in China when censorship increases? On September 26, 2014, Hong Kong protesters began sit-in protests in response to reforms being considered by the Hong Kong government to the electoral system. On September 29, 2014, the Chinese government suddenly disrupted social media by blocking Instagram, a popular social media website used for posting photos.

Instagram was widely used in China before it was suddenly blocked. In 2011, when Instagram was first taking hold in China, one of the founders of the site indicated that Instagram had around 100,000 downloads each week from China.[18] In the month before Instagram was blocked in China, it was in the top 100 Apple application downloads for users in China and in the top 10 for photo and video sharing.[19]

The Instagram block had a large, negative impact on the number of people using Instagram from China. Our data indicate that a little less than half of the users who were previously geolocating to China using Instagram discontinued their use of Instagram. As shown in Figure 1, before the block, our sample contains, on average, 6,368 unique users geolocating to China every day. After the block, the number of unique users within our sample geolocating with Instagram decreases to a mean of 3,376. If, similar to Twitter, only 2–3% of Instagram users geotag their posts (Leetaru et al. 2013), and if we collected one out of four geotagged posts,[20] then there were potentially one million unique users posting to Instagram from China every day before the block, and half a million unique users posting to Instagram from China every day after the block.

The evidence provided here shows that the Instagram block did not create a Streisand effect as censorship of Instagram did not create *more* interest in or attract new users to Instagram. Yet, given that experts have long estimated that very few people in China use or have ever used VPNs, the persistence of Instagram users should be surprising. A full 53% of Instagram users continued using Instagram, despite the block. While there is no direct data available on the total number of Instagram users in China, survey data

---

[18] Millward, Steven. "Instagram Has 100,000 Downloads in China Each Week, Looking at Integration With Weibo," *Tech Asia*, November 1, 2011, https://www.techinasia.com/instagram-china-weibo
[19] Data from application analytics firm App Annie.
[20] We randomly sampled and searched grid coordinates that covered the residences of approximately one fourth the population of China.

FIGURE 1. **The Instagram block's effect on the number of unique Instagram users geolocating from mainland China.**
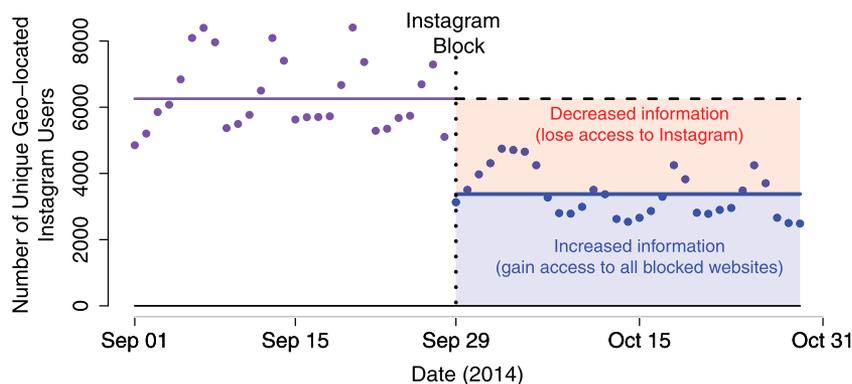
The Instagram block led to an approximately 50% decrease in active, geolocated Chinese Instagram users. The *x*-axis in this figure is the date and the *y*-axis is the number of unique users on Instagram who added a geotagged post to the site that day (a small fraction of all users). We add horizontal lines for the before (purple) and after (blue) Instagram block means of the number of daily, unique geotagging users, as well as the number of these users we would have expected had there been no block (black dotted line). The blue shaded area highlights that 50% of active Chinese Instagram users were accessing an uncensored version of the internet after Instagram was blocked, while the red shaded area highlights that 50% of Chinese Instagram users were no longer active on Instagram after it was blocked.

suggests that approximately 5% of internet users use Facebook.[21] Facebook—which is blocked in China—was downloaded with less frequency than Instagram before the Instagram block according to App Annie. To give a very rough estimate of the extent of Instagram use, if 5% of internet users in China used Instagram before the block, there would be 30,532,500 total Instagram users in China before the block. Roughly, this would suggest that around 16.1 million people—three times the number of all internet users in Hong Kong—were motivated enough to evade the firewall in the days following the block, to access a social media site that they were accustomed to using.[22]

This back-of-the-envelope calculation is a rough approximation, and we explore different parameters for the estimate in the Supplementary Material. Here, we note that the 30 million Instagram users estimate before the block and 16 million Instagram users after roughly matches our data in Figure 1 if approximately 1 in 30 people post any given day and if Instagram has approximately the same geolocation rate as Twitter.[23] The 1 in 30 posting rate is similar to an independent

activity estimate obtained by comparing posting rates to Instagram's published monthly active users.[24]

If the geolocation rates are higher on Instagram than on Twitter, then the 30 and 16 million estimates should be lower. For example, if the geolocation rate is 5% and geolocating and nongeolocating users are equally affected, then 8 million people continued to use Instagram after the block.[25] This matches our observation that Instagram use in Mainland China was close to six times higher after the block than Hong Kong Instagram use, which we show in the Supplementary Material. Instagram had 2.3 million users in Hong Kong by late 2016, according to the Hong Kong Police Force,[26] so the estimate that 8 to 16 million people jumped the firewall to continue to access Instagram after the block appears well-grounded.

EFFECTS OF THE INSTAGRAM BLOCK ON VPN ACQUISITION

It could be that the users who persisted in using Instagram were already using VPNs and therefore the block did not increase their access to information. However, the application analytics website, App Annie, suggests that the Instagram block caused a large, sudden increase in new VPN application downloads, suggesting

---

[21] Kemp, Simon, "Digital in 2016," We Are Social Singapore, http://www.slideshare.net/wearesocialsg/digital-in-2016

[22] Note that, typically, a small fraction of social media users post content on a daily basis and the "lurkers" (people who view the site but do not post) do not show up in our daily activity data. Also, some of the observed, posting users may have already had VPNs. If 3–15% of these users already had VPNs, then approximately 13.7–15.7 million people may have downloaded a VPN on the day of the block to get on Instagram. We provide a detailed accounting of the assumptions made for this back-of-the-envelope calculation and bounds for this estimate in the Supplementary Material.

[23] 16 million ≈ 3,376 times 4 (1/4 sampling rate) times 30 (1/30 daily activity) times 40 (1/40 geolocation).

[24] Of people who post at any time over a year (and many people never post—even if they often view the site), about 5% post on any given day. See Gaffney, Devin. 2016. Medium: https://medium.com/dgaff/estimating-instagram-s-actual-population-statistics-2ee32dd219f4 . Note, while this article uses Instagram posts, lurking is just as important for our purposes—since both lurking and posting imply VPN access.

[25] (16 times 0.5) ≈ (3,376 times 4 times 30 times 20)

[26] http://www.police.gov.hk/ppp_en/03_police_message/instagram.html

FIGURE 2. Left: Proportion of Tweets from China mentioning "ins" by day. Right: The Instagram block's effect on the rank of VPN applications on iPhones from mainland China, from AppAnnie.com.
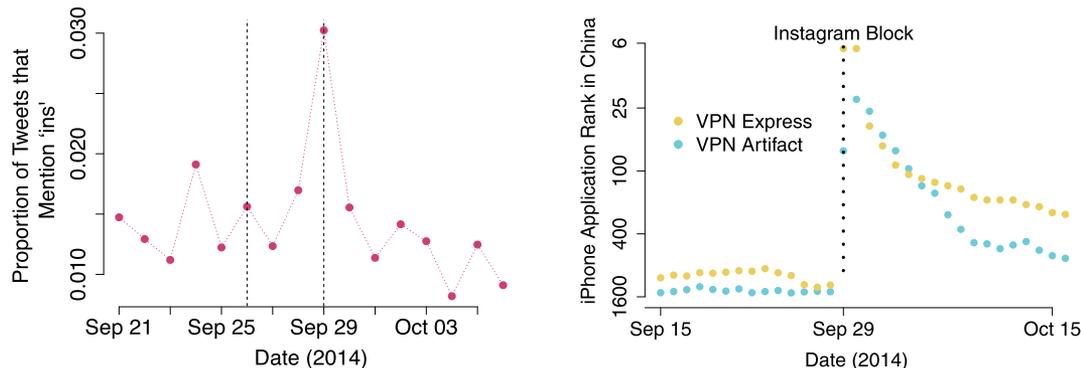
In the left panel of this figure we show that 3% of tweets in China mentioned Instagram on the day of the Instagram block. The right panel of this figure shows that the download ranks of VPN Express and VPN Artifact increased from ranks lower than 1,000 to the top 10 most popular applications in China on the day of the Instagram block.

that the Instagram block created millions of *new* VPN users. On September 28, 2014, VPNs did not make it on the top 10 downloaded productivity applications for iPhones. But on September 29, 2014, four of the top 10 downloaded free productivity applications in China were VPNs: VPN Express, GreenVPN, VPNArtifact, VPN in Touch. To take one stark example, on September 28, 2014, VPN Express was the 1,229th most downloaded iPhone app in China. By the next day, September 29, 2014, it was the sixth most downloaded. The right panel of Figure 2 shows the time series rank of two of these VPN applications during the time period surrounding the Instagram block.[27]

Consistent with our hypothesis that habits were driving censorship evasion, those who continued to use Instagram by downloading a VPN were those who were most accustomed to using Instagram. We use the numbers of likes and comments for users on Instagram as a proxy for the amount of interaction each user experiences on Instagram. We include in the Supplementary Material the difference between the number of likes and comments before the Instagram block between the users who stayed on Instagram in comparison to those who stopped using Instagram. Those who continued to use Instagram had substantially more interactions with other users before the block, including more comments and likes.

EXPANDED ACCESS TO BLOCKED WEBSITES

Users who downloaded a VPN to continue to access Instagram did not just remain on Instagram, but also appeared to begin interacting with a number of websites that had long been of limits in China. Websites already blocked in China saw a surge in use from China. The left panel of Figure 3 shows the number of application downloads of Twitter and Facebook—both blocked from the mainland—in China. Both applications peaked on the day of the Instagram block, indicating that the new users evading censorship expanded their presence on social media to websites that had long been blocked and that they had not previously interacted with.

In the sample of Twitter data we collected, geolocated to China, we see unprecedented numbers of new users joining Twitter. The right panel of Figure 3 shows the effect of the Instagram block on Twitter use. The increase in Twitter use was driven largely by new account creations. In the Supplementary Material, we include a figure that shows new account creations by day. On the day of the Instagram block, the rate of new user account creation jumped more than 600%.

Of course, it could be that the new users of Twitter joined not because of the Instagram block but instead to follow developments in the Hong Kong protests. However, our data do not support this interpretation of events. First, we show in Figure 3 that the timing of the user increase corresponds with the Instagram block and not the protests. The Hong Kong protests began on September 26, and the first two days of the protest did not seem to generate enough interest among mainland users to significantly increase the number of users on Twitter. The increase in new users also corresponds to a spike in conversation about Instagram (left panel of Figure 2). A full 3% of tweets within our sample use the word "ins" on September 29,[28] three times the amount that it is typically used within the sample.

Second, if new Twitter use was inspired by the Hong Kong protests, we would expect a similar increase in

---

[27] We include a long-term plot on VPN Express (we don't have long-term data for VPN Artifact) in the Supplementary Material to show that this jump is very significant even on a longer time scale.

[28] We use "ins" because it will pick up both those using the word 'Instagram' and also its nickname, "Ins."

FIGURE 3.   Left: The Instagram block's effect on the rank of Facebook and Twitter on iPhones from mainland China, from AppAnnie.com. Right: Comparison of tweets per day from Mainland China and Hong Kong before and after the Instagram block.
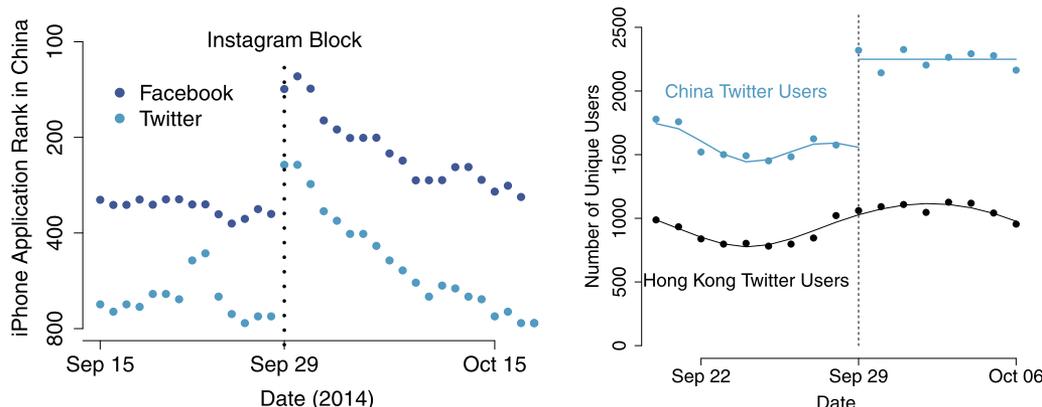
The left panel of this figure shows the change in download ranks for Facebook and Twitter before and after Instagram was blocked. The right panel of this figure shows that the Chinese Twitter users in our sample increased 30% the same day that we observe a spike in Instagram mentions and several days after the beginning of the Hong Kong protests. This increase only occurred in China and not in Hong Kong. The lines in this panel were fit using a smoothing spline.

the area most affected by the protest—Hong Kong. The right panel of Figure 3 shows that the jump in Twitter use only affects mainland China. This suggests that the increase in Twitter use is not due to events within Hong Kong, but indeed due to the Instagram block, which only affected mainland China.

To further understand the sudden increase in Twitter users on September 29, we take a closer look at the new geolocating users within our sample who joined Twitter on September 29, and compare them to existing Twitter users from China. Overall, these new users look very different from the typical Twitter user, indicating that the Instagram block encouraged less political, more typical Chinese users to jump the Firewall. First, the new users are much less international than the typical Chinese Twitter user. Eighty percent of the new Twitter users indicated that their preferred language was Chinese, in comparison to only 39% of regular Twitter users. Only 18% of new users indicated that they preferred English, compared to 49% of regular Twitter users. New users are also much less likely to reside in China's most international cities. Whereas approximately 30% of regular Twitter users geolocate to Beijing and Shanghai, only 15% of the new users were geolocated to these two major cities.

In addition to basic user characteristics, we also coded the first ten people that each of the new users followed. We find that almost 80% of the users that the new users followed fall into the category of arts and entertainment or sports. One new Twitter user best summarizes the new users' motivation to join Twitter, suggesting that the motivation for joining Twitter was not political, but rather incidental:

> "Today Chinese not on INS, want to download something to use, then try the twitter, the results can, can pay attention to Justin Bieber proud" [sic]

POLITICIZATION

The Instagram block precipitated a large increase in followers for Chinese-language Twitter users, enhancing the network of the blocked website among mainland citizens. Not only did new users follow celebrities, they also began following political accounts, including accounts of news sites such as *New York Times* Chinese and *Apple Daily* (a Hong Kong news source highly critical of the mainland Chinese government). We show in the right panel of Figure 4 the total additional followers to these accounts, by day, compared to their average in the 30 days preceding the block. Overall, by November 1, there were 33,750 more followers of Chinese Twitter users than we would have expected based on data from the month before the block.

Even though the new Twitter users were not originally interested in politics, preliminary evidence suggests that they started becoming politically interested a few days after joining Twitter. We code the first ten tweets of each of the new users starting September 29 into categories of whether or not these Tweets are political mentions of Hong Kong. Only 1% of the first ten tweets of new users mention politics in Hong Kong, in comparison to the first ten tweets of old users on Twitter on September 29, where 3% of them mention politics in Hong Kong.

The left panel of Figure 5 displays the proportion of Tweets that mention politics in Hong Kong for new users and old users. Although new users show relatively little interest in Hong Kong their first day on the site, by the second day, they begin to look like the other Twitter users. As one new Twitter user, whose Twitter activity mostly revolves around a Thai pop star, puts it: "Something have happened in HK, but people on mainland do not know because of the government of mainland, pray for you."

FIGURE 4.   Left: Daily new followers to *New York Times* Chinese and *Apple Daily* Twitter accounts (based on new user sign-up dates). Right: Cumulative increase in followers, compared to preblock trend, of any Chinese language user (based on new user sign-up dates) compared to expected increase in followers.
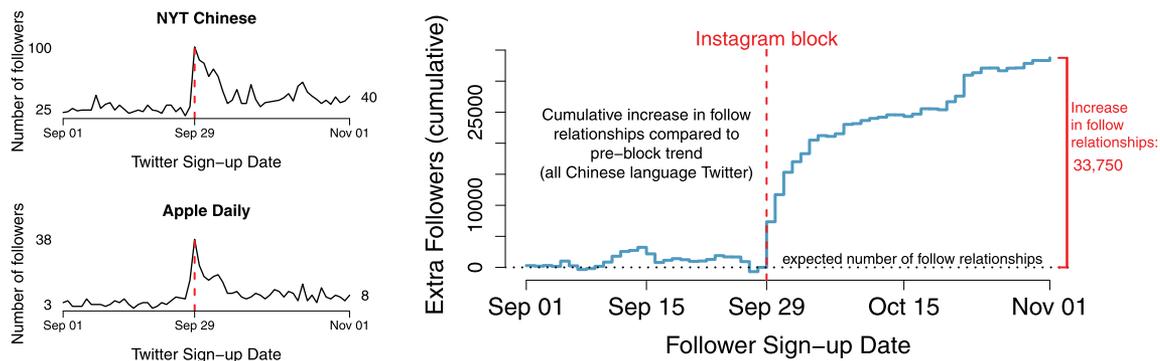
The left panel of this figure shows the sign-up dates of followers of the *New York Times* Chinese and *Apple Daily* Twitter accounts. Many followers of these accounts signed up for Twitter immediately following the Instagram block. This increase in sign-ups—users who eventually followed *NYT* Chinese and *Apple Daily*—continues long after the Instagram block. The right panel of this figure shows that all Chinese language Twitter users accumulated approximately 33,750 more followers from new Twitter sign-ups than what we would expect based on pre-block trends. This cumulative increase was calculated using a cumulative sum of the number of new followers minus the number of expected followers, where the expected followers was the mean daily number of new followers prior to the Instagram block.

These effects extended beyond geolocated social media and application downloads. We were able to measure off-limits information access directly by using page views of Chinese language Wikipedia pages (zh.wikipedia.org) that were blocked in China before and after the Instagram block.[29] We recorded which Chinese-language Wikipedia pages were blocked by the Great Firewall on September 29, 2014 using data obtained from Great Fire (greatfire.org). The pages we study are political, as only *specific*, political pages were blocked at the time in mainland China, including pages such as "human rights" (人权) and Chinese dissident Liu Xiaobo (刘晓波).

In Figure 5, we show that there was a large spike in views of Chinese language Wikipedia pages blocked in China on the days after the Instagram block. There were approximately 160,000 more page views of blocked Wikipedia pages on September 29 than in the week preceding the Instagram block. Unlike other analyses, we were not able to distinguish worldwide page views of Chinese language Wikipedia pages from page views coming from Mainland China; however, the patterns of information access very closely reflected those in our geotagged results.

In Table 5, we show the censored Wikipedia pages that had the largest increases in views on the days following the Instagram block. If those viewing the blocked Wikipedia pages were long-time users of VPNs, we should only see spikes in page views of pages related to ongoing events, like the Hong Kong protests.

Of course, we do see some interest in Wikipedia views related to Hong Kong, particularly on the first day of the Instagram block from September 28–29. However, we also see large increases in page views about sensitive historical topics in China and mainland political leaders, including interest in the June 4, 1989 crackdown in Tiananmen Square, its leaders, and aftermath, and mainland Chinese political leaders such as Mao Zedong, Deng Xiaoping, Jiang Zemin, Xi Jinping, and Hu Jintao. We also see rising interest in the list of websites that the PRC blocks. These are viewing patterns that would be more typical of new users who had just jumped the firewall, rather than of old VPN users who had presumably consumed this information long ago. These changes are particularly pronounced on the 30th, after new users have had a day to adjust to the unfamiliar and more open information environment.

IMPLICATIONS

The Instagram block resulted in the unintended consequence of giving a large number of "normal" Chinese citizens incentives to evade censorship and therefore access to websites and information that many had likely never previously encountered. In what situations do we expect that the gateway effect will be relevant? And what are the implications of this expanded access to information? In this section, we describe the situations in which we expect that gateway effect to be strongest.

The increased activity in political discussion on the part of new users could shed light on the mechanisms that underlie existing empirical evidence that media

---

[29] Page view data at https://dumps.wikimedia.org/other/pagecounts-raw.

FIGURE 5.   Left: Tweets that mention politics in Hong Kong, comparison of new users and old users. Right: Page views for Chinese language Wikipedia pages blocked in China. Bottom: Changes in Wikipedia views.

Largest increases in views **September 28th to 29th** *First day of Instagram block*	Largest increases in views **September 29th to 30th** *Second day of Instagram block*
Occupy Central	**People's Republic of China blocked websites list**
Hong Kong	**Jiang Zemin**
June 4th Incident	Radio Australia
Hong Kong Independence Movement	**Hu Jintao**
pan-Democrats	**Zeng Qing**
Xi Jinping	**Wang Weilin (Tank Man)**
Sunflower student movement	**Li Peng**
Guo Baixiong	**Ling Jihua**
Article 23 Hong Kong	**Tiananmen Square Incident**
Zhang Xuan (celebrity who supported HK movement)	**Zhou Yongkang**
Alliance for True Democracy	**Wu'erkaixi (June 4 leader)**
Bo News (dissident Chinese news network, located in U.S.)	**Zhang Dejiang**
Chai Ling (June 4 leader)	YouTube
Jasmine Revolution	**Wen Jiabao**
Hu Yaobang	AV Actress (Japanese porn stars)
Taiwan	**Mao Zedong**
Jiang Zemin	**Deng Xiaoping**
People's Republic of China blocked websites list	**Ling Gu (son of Ling Jihua)**
Democracy	**Wu Bangguo**
Wu'erkaixi (June 4 leader)	**Hua Guofeng**

The left panel of this figure shows that users who signed up for Twitter after the Instagram block began mentioning protest events in Hong Kong about a day after their arrival on the site. The right panel of this figure shows page views of Chinese language Wikipedia pages that were blocked in China before and after Instagram was blocked. The increase in Wikipedia page views shows that the Instagram block facilitated increased access to information for some users. In the bottom table, the change in Wikipedia views shows that new viewers accessed pages that had long been censored including those related to the 1989 Tiananmen Square protests (general political and historical events are in bold).

disruption can increase the potential for collective action. However, we do not expect that the gateway effect will be important in all instances of increased censorship. We hypothesize that the increase in access to information will occur primarily in three situations: 1) when censorship is sudden enough to disrupt citizen habits, 2) when the newly censored information is difficult to substitute with uncensored alternatives, and 3) when there is low probability for punishment for evasion.

Sudden censorship is more likely to cause the habit disruption that creates gateway effects. When users have full access to a website one day, which is completely blocked the next, they may be in the middle of conversations or projects that create short-term incentives for them to seek out evasion tools.[30] The gate-

way effect that sudden censorship creates may explain why the Chinese government has opted for more gradual blocks of websites in the past. In 2010, following a conflict with Google, the Chinese government throttled Google for years, choosing not to outright block the website until 2014 (Millward 2011). In retrospect, this may have been an effort to wean users off of Google without creating immediate incentives to jump the firewall. Similar to government efforts to mitigate Streisand effects (Jansen and Martin 2015), sudden censorship may leave the government less time to blame the block on other causes or hide the existence of censorship.

Unlike the Google block, which was precipitated by a conflict between the tech giant and the Chinese government, censorship in reaction to collective action events may necessarily be sudden. In the case of the Instagram block, the Chinese government was purportedly worried about images of police using tear gas on

---

[30] Note that Streisand effects and other forms of backlash against repression may also be aggravated with suddenness, see Martin (2007).

protesters in Hong Kong circulating to mainland citizens on Instagram.[31] If these images were the most immediate, dangerous piece of information that could accelerate the protest, a sudden block that decreases use of Instagram could be strategic for a government, even if it increases access to evasion technology and other blocked websites for millions of people in the longer term.

Second, gateway effects are more likely when the newly censored information has fewer substitutes. It is increasingly clear that one of the primary information-control strategies of the Chinese government is to create distractions to off-limits information (King, Pan, and Roberts 2017). Chen and Yang (2017) show that even when given evasion software, Chinese citizens often do not have incentives to evade the firewall because they have low demand for information across the firewall. The government has actively encouraged the development of Chinese social media websites to provide substitutes for those that are blocked by the firewall. Instagram, one of the most popular social networking websites in the world, not only has no clear substitute in China, but has been shown to be more addicting than websites (Roberts, Yaya and Manolis 2014). Social networks formed on Instagram may be difficult for consumers to replace. The results described in this paper therefore may not hold for other websites.

Third, we expect that gateway effects will be most important when the probability for punishment for evasion is low. In the China case, those who access blocked websites through VPNs are not punished. In other more totalitarian contexts, censorship evasion could be accompanied with punishment. In these cases, the gateway effect will likely be muted, as citizens will have other reasons not to acquire evasion tools. In these cases, cycles of repression and dissent might be more important (Moore 1998; Davenport 2005, 2007). We include this possibility in a model of the gateway effect in the Supplementary Material.

Is the Instagram block a unique case of the gateway effect, or are there other cases like it? We believe that, while it is difficult to document the existence of gateway effects because they are often associated with illicit or illegal behavior, they are common both today and throughout history. One example of a recent gateway effect comes from media reports about the crackdown on bitcoin and the Chinese social media website WeChat in 2017.[32] Although there were no estimates of the size of the gateway effect from that crackdown in September 2017, like we provide for Instagram in 2014 here, the restrictions on digital currencies and WeChat appeared to have led users to explore overseas havens for digital currency and to encrypted platforms banned in China, including encrypted messaging service Telegram. Beyond China, sudden blocks of Twitter, Facebook, Google, or entertainment such as Netflix that have become common in countries such as Turkey, Iran, and Egypt may motivate people to download VPNs, which may, in turn, introduce them to other information that has long been blocked in their country. Banned books, TV, or newspapers may motivate people to seek out underground bookstores, or get antennas that can access off-limits radiowaves, and this technology may introduce them to other books, TV, and newspapers that have content more dangerous to the government (Kern and Hainmueller 2009). Gateway effects also expand outside of censorship. Newly banned drugs might encourage users to seek out drug dealers and be exposed to other more dangerous substances; this is one explanation for why heroin use has increased with measures to crack down on opioids.[33] We hope that future research can explore examples of gateway effects outside of that which we have outlined here.

The Government Reaction and Long-term Impacts

The longer term impact of the gateway effect will be conditional of the government's strategic reaction to increased citizen evasion. Not long after the Instagram block, the Chinese government began increasing censorship in a different way—by raising the Great Firewall. Instead of blocking more websites, the government instead increased the difficulty of successfully jumping the firewall. Though the crackdown on VPNs does not appear to have happened in one day, reports of difficulties of accessing VPNs begin in November 2014, culminating in January 2015 when Beijing is said to have "upgraded" the firewall. During this time period, VPN providers began reporting disruptions in their services and users reported being increasingly frustrated with VPN access.[34]

While we cannot tell for certain whether this crackdown in the months following the Instagram block was in reaction to the increased number of people evading censorship, the Chinese government's upgrade of the firewall after the Instagram block was conceivably an effort to reduce a ballooning censorship-evading population. This implies that autocrats will have incentives to follow expansions in censorship like the Instagram block with crackdowns on censorship evasion. The VPN crackdown following the block is consistent with other empirical evidence that media disruptions are accompanied by larger crackdowns, but suggests an alternative mechanism (Gohdes 2015).

---

[31] Park, Madison, "China's Internet Firewall Censors Hong Kong Protest News," CNN, September 30, 2014, http://www.cnn.com/2014/09/29/world/asia/china-censorship-hong-kong/
[32] "China's WeChat Crackdown Drives Bitcoin Enthusiasts to Telegram," *South China Morning Post*, September 15, 2017, http://www.scmp.com/news/china/policies-politics/article/2111322/chinas-wechat-crackdown-drives-bitcoin-enthusiasts

[33] Friedersdorf, Conor, "How Drug Warriors Helped to Fuel the Opioid Epidemic," *The Atlantic*, April 4, 2016. https://www.theatlantic.com/politics/archive/2016/04/how-drug-warriors-helped-to-fuel-the-heroin-epidemic/476679/
[34] "China Blocks Websites in 'Censorship Campaign' Ahead of Major Internet Conference," *The Guardian*, November 18, 2014, http://www.theguardian.com/world/2014/nov/18/china-blocks-access-websites-before-hosting-world-internet-conference; "In China, VPN Internet Access Tools Suffer Further Disruptions," *Reuters*, January 23, 2015, http://www.reuters.com/article/us-china-internet-vpn-idUSKBN0KW0WS20150123.

FIGURE 6. Instagram block post-mortem: Did the effects persist?

Signup dates of people who mentioned Liu Xiabo after his death in July 2017
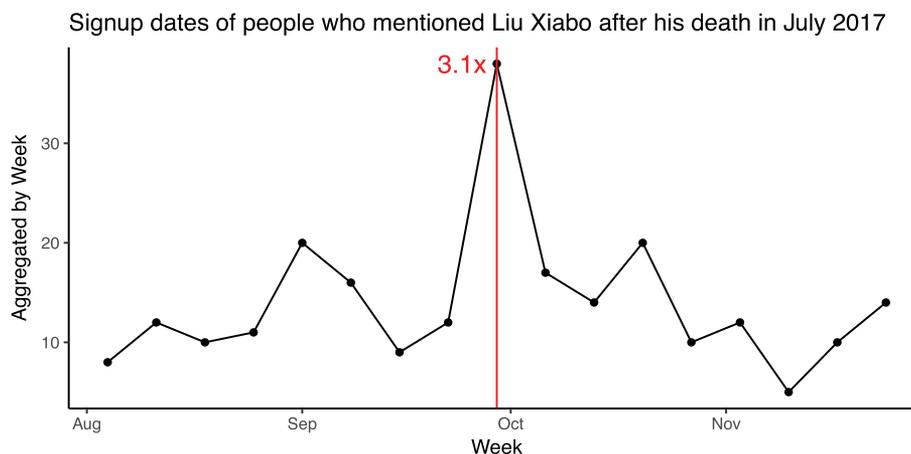
This figure shows the number of Chinese language Twitter users by sign-up date who mentioned democracy activist and Nobel Peace Prize Laureate Liu Xiaobo on Twitter after his death in July 2017. The number of people who discussed his death and signed up for Twitter after the Instagram block on the week of September 29 was about three times higher compared to the weekly average of users who signed up from August through September 2014.

If governments frequently make evasion more difficult following an expansion of censorship, this likely churns networks of those engaged on blocked social media sites creating unknown implications for collective action. Some individuals who already access blocked information may drop out once evasion becomes more costly. This removes individuals for whom evasion is difficult or who have limited motivation to pay the costs of evasion. At the same time, as governments block more websites, it will create incentives for new individuals to jump the firewall if they are accustomed to using the censored sites. This introduces newly censured individuals to existing, politically interested social networks and removes relatively disinterested and/or resource-constrained individuals.

The effects of this churn are currently unclear. On one hand, the replacement of peripheral members in a social network could disrupt collective action potential by reducing continuity in a network (Barberá et al. 2015). New users might only gradually build new connections to replace old ones in the networks. On the other hand, new members could alter social dynamics in a network by introducing new ideas and behaviors to homogeneous and stale social groups, similar to dynamics observed in successful creative networks (Uzzi and Spiro 2005). By disrupting old relationships and introducing unfamiliar ones to replace them, churn potentially increases the likelihood of emergent collective behaviors.

To assess whether the effects of the Instagram block continued despite increased efforts by the Chinese government to limit access to VPNs, we looked at the sign-up dates of Twitter users who discussed democracy activist and Nobel Peace Prize Laureate Liu Xiaobo after his death in July 2017. This was a simple test to see if we could still detect an effect of the 2014 block even three years later in 2017. We chose Liu Xiaobo because

a mention of him was unambiguously political and explicitly censored in Mainland China, so we were unlikely to pick up false positives.[35] Liu Xiaobo was also a useful case because he died just before we revisited this analysis in late July 2017 and it was convenient for us to collect posts about him on Twitter as Twitter users were discussing his death.

We show in Figure 6 that while relatively few Chinese language Twitter users who signed up in 2014 discussed Liu Xiaobo's death, there were around three times as many users discussing the topic who signed up in the week following the Instagram block as we would expect based on average sign-up numbers before September 29, 2014. Our previous evidence showed that six times the average number of people signed up for Twitter following the Instagram block, suggesting that the Instagram-inspired Twitter users were less likely than average to become politicized in the long-term than the average new Twitter user, a trend we would expect given that they were initially motivated to jump the firewall to access entertainment. Yet, because of the sheer number of people who were introduced to Twitter because of the Instagram block, even with a lower politicization rate, this evidence suggests that the block increased the number of people who engage with political information on Twitter and that the Instagram block has had long-term political ramifications for China.

CONCLUSION

Despite conventional wisdom that increases in censorship strictly decreases access to information, we

---

[35] For example, if we look for the similarly sensitive word "harmonious," we will detect a mix of political and ordinary language.

find that expansions in censorship to previously uncensored websites can incentivize citizens to invest in censorship-evasion technology. Upon learning how to evade censorship, these individuals, in turn, have more, not less, access to information and begin engaging in conversations, social media sites, and networks that have long been off-limits to them. Censorship, unexpectedly, increases access to information for the individuals for whom it provides a reason to search out off-limits information. Our findings suggest a potential mechanism through which users actually have access to more information despite increased censorship. If censorship is applied to largely apolitical social media sites like Instagram, it can pull apolitical users into censorship-evasion networks and engage them in political discussion and coordination.

Importantly, these patterns of behavior were not forms of psychological reactance—such as a Streisand effect, where individuals explicitly seek out information that is being hidden. Instead, users gained access to more information simply by maintaining their current social media consumption (and there were fewer users of the newly blocked media itself). In blocking Instagram, the Chinese government reduced the number of users on Instagram, but increased access to websites that are typically more political such as Twitter, Wikipedia, and Facebook.

SUPPLEMENTARY MATERIAL

To view Supplementary Material for this article, please visit https://doi.org/10.1017/S0003055418000084.

Replication materials can be found on Dataverse at: https://doi.org/10.7910/DVN/Q8NRTS.

REFERENCES

Barberá, Pablo, Ning Wang, Richard Bonneau, John T. Jost, Jonathan Nagler, Joshua Tucker, and Sandra Gonález Bailón. 2015. "The Critical Periphery in the Growth of Social Protests." *PLOS ONE* 10 (11): e0143611.
Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (2): 169–217.
Brown, William W., and Morgan O. Reynolds. 1973. "Crime and "Punishment": Risk Implications." *Journal of Economic Theory* 6 (5): 508–14.
Chen, Yuyu, and David Y. Yang. 2017. "1984 or the Brave New World? Evidence from a Field Experiment on Media Censorship in China" (working paper).
Chwe, Michael Suk-Young. 2001. *Rational Ritual: Culture, Coordination, and Common Knowledge*. Princeton: Princeton University Press.
Davenport, Christian. 2005. "Repression and Mobilization: Insights from Political Science and Sociology." In *Repression and Mobilization*, eds. Christian Davenport, Hank Johnson, and Carol Mueller. Minneapolis: University of Minnesota Press, vi–xlii.
Davenport, Christian. 2007. "State Repression and Political Order." *Annual Review Political Science* 10: 1–23.
Deibert, Ronald, John Palfrey, Rafal Rohozinski, Jonathan Zittrain, and Miklos Haraszti. 2010. *Access Controlled: The Shaping of Power, Rights, and Rule in Cyberspace*. Cambridge: MIT Press.
Diamond, Larry. 2010. "Liberation Technology." *Journal of Democracy* 21 (3): 69–83.
Edmond, Chris. 2013. "Information Manipulation, Coordination, and Regime Change." *The Review of Economic Studies* 80 (4): 1422–58.
Egorov, Georgy, Sergei Guriev, and Konstantin Sonin. 2009. "Why Resource-Poor Dictators Allow Freer Media: A Theory and Ev-

idence from Panel Data." *American Political Science Review* 103 (4): 645–68.
Eide, Erling, Paul H. Rubin, and Joanna Mehlop Shepherd. 2006. *Economics of Crime*. Hanover: Now Publishers Inc.
Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya. 2011. "Media and Political Persuasion: Evidence from Russia." *American Economic Review* 101 (7): 3253–85.
Francisco, Ronald A. 2005. The Dictator's Dilemma. In *Repression and Mobilization*, eds. Christian Davenport, Hank Johnson, and Carol Mueller. Minneapolis: University of Minnesota Press, 58–81.
Geddes, Barbara, and John Zaller. 1989. "Sources of Popular Support for Authoritarian Regimes." *American Journal of Political Science* 33 (2): 319–47.
Gohdes, Anita R. 2015. "Pulling the Plug: Network Disruptions and Violence in Civil Conflict." *Journal of Peace Research* 52 (3): 352–67.
González-Bailón, Sandra, Javier Borge-Holthoefer, Alejandro Rivero, and Yamir Moreno. 2011. "The Dynamics of Protest Recruitment Through an Online Network." *Scientific Reports* 1 (197).
Hassanpour, Navid. 2014. "Media Disruption and Revolutionary Unrest: Evidence From Mubarak's Quasi-Experiment." *Political Communication* 31 (1): 1–24.
Hess, David, and Brian Martin. 2006. "Repression, Backfire, and the Theory of Transformative Events." *Mobilization: An International Quarterly* 11 (2): 249–67.
Howard, Philip N. 2010. *The Digital Origins of Dictatorship and Democracy: Information Technology and Political Islam*. Oxford: Oxford University Press.
Huang, Haifeng. 2015. "Propaganda as Signaling." *Comparative Politics* 47 (4): 419–44.
Jansen, Sue Curry, and Brian Martin. 2003. "Making Censorship Backfire." *Counterpoise* 7 (3): 5–15.
Jansen, Sue Curry, and Brian Martin. 2015. "The Streisand Effect and Censorship Backfire." *International Journal of Communication* 9: 656–71.
Kalathil, Shanthi, and Taylor C. Boas. 2010. *Open Networks, Closed Regimes: The Impact of the Internet on Authoritarian Rule*. Washington, D.C.: Carnegie Endowment for International Peace.
Kelly, Sanja, Sarah G. Cook, and Mai Truong. 2012. *Freedom on the Net 2012: A Global Assessment of Internet and Digital Media*. Washington D.C.: Freedom House.
Kern, Holger Lutz, and Jens Hainmueller. 2009. "Opium for the Masses: How Foreign Media Can Stabilize Authoritarian Regimes." *Political Analysis* 17 (4): 377–99.
King, Gary, Jennifer Pan, and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107 (2): 1–18.
King, Gary, Jennifer Pan, and Margaret E. Roberts. 2014. "Reverse-Engineering Censorship in China: Randomized Experimentation and Participant Observation." *Science* 345 (6199): 1–10.
King, Gary, Jennifer Pan, and Margaret E. Roberts. 2017. "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument." *American Political Science Review* 111 (3): 484–501.
Kuran, Timur. 1991. "Now Out of Never: The Element of Surprise in the East European Revolution of 1989." *World Politics* 44 (1): 7–48.
Kuran, Timur. 1997. *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge: Harvard University Press.
Leetaru, Kalev, Shaowen Wang, Guofeng Cao, Anand Padmanabhan, and Eric Shook. 2013. "Mapping the Global Twitter Heartbeat: The Geography of Twitter." *First Monday* 18 (5).
Lessig, Lawrence. 1999. *Code: And Other Laws of Cyberspace*. New York: Basic Books.
Lohmann, Susanne. 1994. "The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-1991." *World Politics* 47 (1): 42–101.
Lorentzen, Peter. 2014. "China's Strategic Censorship." *American Journal of Political Science* 58 (2): 402–14.

Lynch, Marc. 2011. "After Egypt: The Limits and Promise of Online Challenges to the Authoritarian Arab State." *Perspectives on Politics* 9 (2): 301–10.

MacCoun, Robert J. 2010. "What Can We Learn from the Dutch Cannabis Coffeeshop Experience?"*RAND (working paper),* 1–32. https://www.rand.org/content/dam/rand/pubs/working_papers/2010/RAND_WR768.pdf.

MacKinnon, Rebecca. 2012. *Consent of the Networked: The Worldwide Struggle For Internet Freedom*. New York: Basic Books.

Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to Information Retrieval*. Cambridge: Cambridge University Press.

Martin, Brian. 2007. "Slow Injustice." *Social Alternatives* 26 (4): 5–9.

Miller, Blake. 2017. "The Limits of Commercialized Censorship in China" (working paper).

Millward, Stephen. June 30, 2011. "Google+ Not Actually Blocked in China, Just Being Slowly Throttled." *TechAsia,* http://www.techinasia.com/google-plus-china.

Miron, Jeffrey A., and Jeffrey Zwiebel. 1991. "Alcohol Consumption During Prohibition." *The American Economic Review* 81 (2): 242–7.

Moore, Will H. 1998. "Repression and Dissent: Substitution, Context, and Timing." *American Journal of Political Science* 42 (3): 851–73.

Morozov, Evgeny. 2011. *The Net Delusion: The Dark Side of Internet Freedom*. New York: PublicAffairs.

Nabi, Zubair. 2014. "Resistence Censorship is Futile." *First Monday* 19 (11).

Pierskalla, Jan Henryk. 2010. "Protest, Deterrence, and Escalation: The Strategic Calculus of Government Repression." *Journal of Conflict Resolution* 54 (1): 117–45.

Pierskalla, Jan H., and Florian M. Hollenbach. 2013. "Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa." *American Political Science Review* 107 (2): 207–24.

Rasler, Karen. 1996. "Concessions, Repression, and Political Protest in the Iranian Revolution." *American Sociological Review* 61 (1): 132–52.

Ritter, Emily Hencken. 2014. "Policy Disputes, Political Survival, and the Onset and Severity of State Repression." *Journal of Conflict Resolution* 58 (1): 143–68.

Roberts, Margaret E. 2018. *Censored: Distraction and Diversion Inside China's Great Firewall*. Princeton: Princeton University Press.

Roberts, James A., Luc Yaya, and Chris Manolis. 2014. "The Invisible Addiction: Cell-Phone Activities and Addiction Among Male and Female College Students." *Journal of Behavioral Addictions* 3 (4): 254–65.

Rød, Espen Geelmuyden, and Nils B. Weidmann. 2015. "Empowering Activists or Autocrats? The Internet in Authoritarian Regimes." *Journal of Peace Research* 52 (3): 338–51.

Salganik, Matthew J. 2017. *Bit by Bit: Social Research in the Digital Age*. Princeton: Princeton University Press.

Shirk, Susan L. 2011. *Changing Media, Changing China*. Oxford: Oxford University Press.

Shirky, Clay. 2008. *Here Comes Everybody: The Power of Organizing Without Organizations*. New York: Penguin Press.

Song, Sonya Yan, Robert Faris, and John Kelly. 2015. "Beyond the Wall: Mapping Twitter in China." *Berkman Center Research Publication* (2015-14).

Stern, Rachel E., and Jonathan Hassid. 2012. "Amplifying Silence: Uncertainty and Control Parables in Contemporary China." *Comparative Political Studies* 45 (10): 1230–54.

Steinert-Threlkeld, Zachary C. 2017. "Spontaneous Collective Action: Peripheral Mobilization During the Arab Spring." *American Political Science Review* 111 (2): 379–403.

Stockmann, Daniela. 2012. *Media Commercialization and Authoritarian Rule in China*. Cambridge: Cambridge University Press.

Tilly, Charles. 1978. *From Mobilization to Revolution*. New York: McGraw-Hill.

Tufekci, Zeynep, and Christopher Wilson. 2012. "Social Media and the Decision to Participate in Political Protest: Observations from Tahrir Square." *Journal of Communication* 62 (2): 363–79.

Uzzi, Brian, and Jarrett Spiro. 2005. "Collaboration and Creativity: The Small World Problem." *American Journal of Sociology* 111 (2): 447–504.

Yang, Guobin. 2013. Social Dynamics in the Evolution of China's Internet Content Control Regime. In *Routledge Handbook of Media Law*, eds. Monroe E Price, Stefaan Verhulst, and Libby Morgan. New York: Routledge, 285–302.

Yang, Guobin. 2014. "Internet Activism & the Party-State in China." *Daedalus* 143 (2): 110–23.

Zhang, Xiaoquan, and Feng Zhu. 2011. "Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia." *The American Economic Review* 101 (4): 1601–15.

Zuckerman, Ethan. 2015. "Cute Cats to the Rescue? Participatory Media and Political Expression." In *From Voice to Influence: Understanding Citizenship in a Digital Age*, eds. Danielle Allen and Jennifer S. Light. Chicago: University of Chicago Press, 131–54.