Spring 2017

Authoritarianism and the Internet

Bryan H. Druzin
Gregory S. Gordon, *Chinese University of Hong Kong*

Law & Social Inquiry
Journal of the American Bar Foundation

Law & Social Inquiry
Volume 43, Issue 4, 1427–1457, Fall 2018

Authoritarianism and the Internet

Bryan Druzin and Gregory S. Gordon

This article argues that Internet censorship is more fragile than is generally supposed and is, in fact, vulnerable to abrupt collapse. The volume and rapidity of online communication renders perfect policing of the Internet technologically impossible. Authoritarian governments are thus forced to rely on Internet users to police themselves in the form of self-censorship. This strategy has proven largely successful—legal ambiguity regarding what constitutes impermissible speech fosters norms of self-censorship. This reliance on self-censorship, however, renders these censorial systems susceptible to shocks. We set out a model that explains sudden breakdowns in Internet censorship that we term "cyberspeech cascades." A cyberspeech cascade occurs when small expressions of online dissent produce large shifts in public perception regarding the acceptable limits of online expression that are, in fact, inaccurate. Online bandwagons of progressively more brazen speech proliferate into large-scale torrents of uncensored expression, triggering the temporary collapse of self-censorship norms online.

INTRODUCTION

Control over the flow of information is critical to authoritarian rule. In the Information Age, command of online communication has become indispensable for dictatorial regimes to realize this goal. It is thus no surprise that the contemporary repression apparatus devotes considerable energy and resources to snuffing out online dissent. Modern strategies of cyber censorship have been largely successful. However, we argue that Internet censorship is in fact more fragile than it appears. Cyber censorship, this article posits, is structurally circumscribed. The sheer volume of electronic communications that proliferate in present-day, wired societies represents a serious challenge to the successful control of information. In reality, effective surveillance of only a fraction of online communications is ever possible—the quantity and rapidity of online behavior renders it logistically impossible to police every post and every page on the Internet. To achieve large-scale suppression of online dissent, therefore, authoritarian and semi-authoritarian governments are forced to rely substantially on Internet users to police themselves in the form of self-censorship (Herold and Marolt 2011, 55).[1]

Bryan Druzin is an Assistant Professor of Law at the Chinese University of Hong Kong. He may be contacted at bryandruzin@cuhk.edu.hk.

Gregory S. Gordon is an Associate Professor of Law at the Chinese University of Hong Kong.

The conceptual foundations of this article build on earlier work by the lead author published by Cornell University (Druzin and Li 2016). For their insightful comments, the authors are deeply grateful to Eric Posner at the University of Chicago and former Permanent Judge of Hong Kong's Court of Final Appeal, the Hon. Mr. Justice Kemal Bokhary. Any errors are the authors' own.

1. Referring specifically to China, Herold and Marolt cite self-censorship as the government's "second weapon" to control the Internet.

This strategy gets results—fear of government scrutiny discourages the open expression of dissent (Deibert et al. 2008, 180, 197–98, 211, 216, 260–61, 325; Hachigian 2001; Yu 2009, 115; Hayhtio and Jarmo 2008, 196).[2] However, this reliance on self-censorship, we argue, renders cyber censorship implicitly susceptible to abrupt collapse. It is constructed upon the shaky edifice of collective paranoia—a fragile foundation that can crumble as the result of small, individual acts of defiance observed sequentially by other Internet users until a point is reached where the norms that underpin self-censorship suddenly collapse. Utilizing the theory of *informational cascades* in the economics literature, and the related concepts of *availability* and *norm cascades* in the legal academic literature, this article advances a model that explains sudden spontaneous collapses in censorial control unique to online communication. We refer to this phenomenon as a *cyberspeech cascade*.

To encourage Internet users to self-censor, authoritarian governments deliberately thrust their populations into a state of informational uncertainty regarding the precise boundaries of permissible speech. Because the legal boundaries of officially tolerated speech are left unclear, even modest, unintentional expressions of contentious online speech can trigger small shifts in perception regarding the acceptable limits of public expression that may then grow into larger ones. The basic model is as such: because the precise line between permissible and impermissible expression is unclear, small acts of uncensored online speech slightly alter collective perceptions of what constitutes acceptable public expression. With the perceived boundaries of permissible online speech slightly expanded, the process repeats itself: building on what they mistakenly believe are the new expanded limits of acceptable speech, fresh waves of Internet users engage in slightly more contentious speech, further pushing back the perceived boundaries of tolerated expression.

With sufficient iterations, this process creates a snowball effect as mass perceptions regarding the acceptable limits of public expression shift significantly. Reputational concerns will also arise, powerfully strengthening this process—a desire for social approval and a wish to avoid disapproval will further incentivize people, intentionally or not, to push the boundaries of public expression. Thus, trickles of progressively more brazen speech can proliferate into large-scale torrents of uncensored expression, triggering the temporary (or permanent) collapse of self-censorship norms.

This process is, we argue, facilitated by the unique character of online interaction—constraints on information, a linear sequentiality, the hyper-connectivity and speed of online communication, and, to the degree that it is present, anonymity between users—render such explosions of expression more likely to occur online than in the offline world. A cyberspeech cascade may vary in terms of initial speed and scale. However, the highly networked character and exponential potential of Internet communication allows collective shifts in perceptions as to the permissible boundaries of public expression to gain momentum and eventually to reach a tipping point. At that juncture, previously stable norms of self-censorship give way as online discourse is overwhelmed by the sheer enormity and widespread presence of uncensored speech. While informational cascades represent the focus of our model,

2. The authors cited are referring, respectively, to Latin America, the Middle East, Uzbekistan, Malawi, Belarus, Malaysia, and China.

availability and norm cascades, which incorporate reputational considerations, play a decisive role in producing cyberspeech cascades and, as such, form an integral part of our model.

The usefulness of the model we propose is in explaining initial breakdowns in online censorship that often precede eruptions of offline social dissent (existing cascade models already capture quite well the growth of political protest offline). The model is both descriptive and predictive. It describes the sudden collapses in Internet censorship that have been observed across the world in recent years. The model is predictive as it suggests that as Internet usage grows in repressive societies, governments will likely need to rely increasingly on self-censorship, rendering such systems of control even more vulnerable to outbreaks of cyberspeech cascades with potentially massive sociopolitical consequences. We present our argument not as established fact, but as a hypothesis that, ultimately, can be tested only by the unfolding of real-world events as we progress further into the Internet age. We are currently experiencing "the greatest information and communications revolution in human history" (World Bank 2016, xxi) and the full political impact of this transformation remains to be seen. It may be that future events will prove our model extraordinarily accurate, extraordinarily inaccurate, or somewhere in between.

In any event, this article provides a much-needed update to cascade theory in light of the transformative effect of twenty-first-century Internet technology. In an age in which a single individual can upload a video of a panda sneezing that is then viewed by a quarter of a billion people across the planet (Oloffson 2010), the concept of cascades needs to be reexamined. The Internet has become an essential form of communication for billions of people, performed with stunning rapidity. It is a medium for transmitting information at speeds and on a scale unprecedented in human history that has transformed the conduits of social connection upon which cascade theory was first understood. The world has changed, and we need to understand this change. Historically, strongmen have been able to dismantle printing presses with hammers or silence political opponents in prison cells. Internet communication, however, has created new opportunities to disrupt authoritarian control. This article seeks to expose that new reality.

The Vital Role of Self-Censorship in Maintaining Control

The initial enthusiasm with which many scholars proclaimed that the Internet was destined to democratize authoritarian societies—common in the early days of the technology—has largely dissipated (Morozov 2011). As is often the case, reality intruded. Although authoritarian and semi-authoritarian states have embraced the Internet as a tool to help strengthen their domestic economies, these governments have been careful to maintain a tight leash on online communication. Indeed, authoritarian governments the world over have shown remarkable skill at stifling online speech that may upset their hold on power (World Bank 2016, 178). While these governments employ various technology-based strategies of control to achieve this, self-censorship represents a key component in their ability to maintain control. This reliance on self-censorship is born out of necessity. Directly policing the flow

of information that streams daily across the Internet is a technological impossibility (Boas 2004, 438). Self-censorship is therefore instrumental in suppressing online speech critical of the state's practices. Indeed, authoritarian governments have shown themselves quite adept at fostering cultures of self-censorship that smother online chatter capable of generating offline dissent.

Virtually all states engage in some form of cyber censorship to some degree (e.g., censoring child pornography, hate speech, sites that facilitate illegal activity, etc.). However, some notably extreme state cyber censors are Bahrain, Belarus, China, Cuba, Ethiopia, Iran, North Korea, Pakistan, Russia, Saudi Arabia, Sudan, Syria, Turkmenistan, the United Arab Emirates, Uzbekistan, and Vietnam (Reporters Without Borders 2014). Although they vary in terms of the extent to which they censor the Internet, these states rank as the most hard-line cyber censors. Although it is difficult to say with certainty, online self-censorship appears to be commonplace in these societies. This should not be surprising as norms of self-censorship are already well entrenched in these countries—offline habits of guarded speech appear to have, for the most part, reflexively carried over to online expression.

While it may often be the case that societies that lack strong democratic institutions and rule of law drift naturally toward self-censorship in cyberspace, to ensure the migration of self-censorship norms to the cyber realm, these governments proactively encourage "self-censorship through the use of intensive surveillance of online forums" (Brown and Marsden 2013, 163). Many of these governments "use registration, licensing, and identity requirements to control what people do online and to create a climate of self-censorship" (Deibert et al. 2011, 11). A lack of user anonymity combined with the threat of serious legal repercussions for transgressing the accepted limits of public expression has helped establish robust norms of cyber self-censorship.[3]

Legal vagueness as to what constitutes impermissible speech plays a critical role in inducing self-censorship by cultivating a sense of collective paranoia that causes citizens to limit the scope of their individual expression. Broadly worded laws designed to prompt self-censorship and self-monitoring are characteristic of what has been dubbed "second-generation" cyber control (Deibert 2010, 571). Law is used to "restrict permissible communications and to create a climate of fear, intimidation and ultimately self-censorship" (Glanville 2010, 81). The precise ambit of permissible speech is left unclear so as to maximize the range within which people voluntarily restrain their behavior online (Link, Madsen, and Pickowicz 2013, 89–90),[4] creating a chilling effect on public speech (Glanville 2010, 81; Allen-Ebrahimian 2015; Gibson and Lowes 2007; IBP 2013, 46).[5] The subtle yet potent impact of self-censorship has been analogized to a giant snake resting over the heads of those who sit self-consciously beneath it: "Normally the great snake doesn't move. It doesn't have to. It feels no need to be clear about its prohibitions. Its constant silent message is 'You yourself decide,' after which, more often than not, everyone in its shadow makes his or her large and small adjustments—all quite 'naturally'" (Link 2002).

3. Online anonymity is examined again in the section entitled "Cyberspeech Cascades from Reputation." It should be noted, however, that this is a discussion of anonymity between Internet users, not between individual Internet users and government monitors.

4. The cited authors specifically reference China's cyber censorship system.

5. The sources cited are referencing, respectively, China, Singapore, and Myanmar.

US jurisprudence has long recognized this potential under the *chilling effect doctrine*, which holds that vaguely defined statutes can produce a "chilling effect" on citizens' willingness to engage in public speech out of fear of potential legal repercussions, real or not (Youn 2013, 1481–94). This chilling effect has been noted with reference to online speech (Deibert et al. 2008, 232). Legal opaqueness "naturally creates powerful incentives for cautious self-censorship and may, in fact, be the only way [to] exert control over the huge volumes of information circulating [online]" (Ford 2015, 24). Uncertainty as to the precise boundaries of legal permissibility and fear of possible legal reprisal is a very effective way to cultivate behavioral norms that encourage Internet users to self-censor.

Yet, we argue that this reliance on self-censorship represents an inbuilt vulnerability. While authoritarian states have, on the whole, successfully managed the Internet, this vulnerability is latent in online communication and has periodically expressed itself in episodic eruptions of uncensored online speech. This has occurred, we argue, precisely because the exact boundaries of permissible speech are strategically left fuzzy. This absence of clarity has a curious potential in cyberspace to produce the opposite of its intended effect—it may open the door to cyberspeech cascades. Before exploring this in more detail, we need to first unpack key concepts in the cascade literature.

Cascade Theory: The Power of Information and Reputation

Cascade theory seeks to explain social imitation, in which individuals adopt the behavior of one another en masse. The literature on cascades presents several models of this herd-like behavior.

A Brief Overview of the Cascade Literature

The literature describes two kinds of foundational cascades: *informational cascades* and *reputational cascade* (Ellickson 2001, 51). Informational cascades, a concept developed by Siushil Bikhchandani, David Hirshleifer, and Ivo Welch (1992), occur when, faced with informational uncertainty, individuals mimic the behavior of others. In an informational cascade, individuals are motivated solely by informational needs. Other motivating factors such as reputational concerns are not incorporated into this model. Reputational cascades are similar; however, in the case of reputational cascades, individuals emulate the decisions of others not because of information uncertainty but because they seek social approval and wish to avoid disapproval (Sunstein 1996a, 2033; 1996b, 903; Kuran 1998, 623). The result is that people engage in *preference falsification* (expressing preferences that differ from what they genuinely believe) in order outwardly to conform to the views of others.

Timur Kuran developed the early literature on reputational cascades.[6] In "Sparks and Prairie Fires," "Now Out of Never," and more fully in *Private Truths, Public Lies*, Kuran discusses how people's willingness to speak is driven both by

6. Kuran draws from the twin theories of conformity in social psychology: *information social influence* and *normative social influence*.

reputational concerns and by their informational needs (Kuran 1989, 1991, 7; 1997). He argues that both factors jointly determine individual "thresholds" and whether and when a cascade occurs and its subsequent impact (Kuran 1998, 623). Building on Kuran's work, Sunstein introduces the concept of *norm cascades* (Sunstein 1996b). Norm cascades are similarly driven by a combination of motives: reputational concerns and information acquisition. Norm cascades model sudden collective swings in societal norms. Synthesizing much of this earlier work, Kuran and Sunstein (1999) introduce the concept of *availability cascades*, which combines informational and reputational mechanisms.

The model we propose—cyberspeech cascades—weaves together these three cascade types: informational cascades, availability cascades, and norm cascades. These three theories are essential to our argument, and so they are explained in detail below.

A Closer Look: Informational Cascades

Informational cascades occur when an agent's actual beliefs change in response to the public conduct of others. The public information obtained from the behavior of previous actors overrides an agent's private signals, irrespective of accuracy. As each agent changes her behavior, this is observed by other actors, which in turn causes a change in the beliefs of other actors and their subsequent behavior. The process then repeats itself through multiple sequences of observation and belief modification, resulting in an informational cascade. The phenomenon is rooted in the volatility and fragility of mass behavior. Experts "call upon ideas like informational cascades to describe how peripheral knowledge can pervade a society, changing the perceptions and realities of political power" (Joyce 2010, 211). Bikhchandani, Hirshleifer, and Welch describe the concept as follows: "An informational cascade occurs when it is optimal for an individual, having observed the actions of those ahead of him, to follow the behaviour of the preceding individual without regard to his own information" (1992, 994). The agent alters her behavior on the assumption that the publicly available information is reliable.[7]

Informational cascades are in fact quite common. For instance, in attempting to exit an airport after disembarking from an airplane, a passenger might join a small cluster of other passengers ahead of her moving in a particular direction, assuming that these individuals must be going toward the airport exit. That assumption, however, could be entirely erroneous—the fellow passengers may, for example, be simply heading to the restroom. As other disembarking passengers likewise act on this mistaken assumption and the crowd grows larger, an erroneous informational cascade forms.

Consider the following example of an informational cascade. A consumer must decide between shopping at two different stores—Store X and Store Y (Banerjee 1992, 798–99).[8] The consumer has very little private information with which to

7. In social psychology, this tendency is known as *informational social influence*.

8. Banerjee employs a comparable consumer-choice model that involves choosing between two restaurants.

make her decision. However, the consumer perceives that Store X is empty while Store Y has several shoppers milling about inside. Working on the assumption that the presence of shoppers is a reliable indication that Store Y is the superior choice, the consumer chooses Store Y. In doing so, however, the consumer adds to the informational appeal of Store Y for the following consumer faced with the same choice. Given similar thought processes, subsequent consumers faced with the same choice will replicate the actions of the first consumer, who through her initial choice unwittingly amplified the informational appeal of Store Y. In this fashion, Store Y attracts an increasing number of shoppers. This pattern is self-reinforcing: the impression that Store Y is the superior choice grows with each new shopper. Even a shopper who arrives late in the informational cascade and would prefer shopping at Store X may be persuaded to choose Store Y as a result of the large crowd of shoppers in Store Y.

However, Store Y—and this is an important point—may not necessarily be superior to Store X. It may in fact be inferior. The initial shoppers in Store Y may have simply preferred the color of the awning in front of Store Y to that of Store X. Or they might have perceived that Store X looked slightly better organized than Store Y in the way its products were stacked on the shelves. In either case, their choice is not a reliable indication of the superiority of Store Y. Erroneous informational cascades based on misplaced assumptions may be triggered "where initial misrepresentative signals start a chain of incorrect decisions that is not broken by more representative signals received later" (Anderson and Holt 1997, 859; Sterelny et al. 2013, 30–31). The information is inaccurate yet an informational cascade nevertheless takes hold.

Under conditions of sequential observation and limited private information, an informational cascade is difficult to avoid: it will form either with the third or fourth agent. Because the model is sequential, after one consumer makes a choice, the next consumer follows such that each agent is able to see the choices of all the other agents who decided before her. Each consumer has a degree of private information concerning the decision (i.e., which store is better), but, operating in a state of informational uncertainty, assigns equal value to her private and public knowledge. The first consumer, utilizing strictly her private information (as this is the only information she has at her disposal at this juncture), decides to shop in Store Y (as her personal knowledge mandates). Assuming the next consumer's private information similarly suggests that Store Y is the superior option, an informational cascade will initiate immediately. This is because the third agent will choose Store Y even if her private information indicates that Store X is the better option. The third agent gives equal weight to her private and public information and her public information now exceeds her private information. If the second consumer's private information indicates that Store X is the better choice, the decisions of the first and second consumers will negate each other in terms of the available public information. As such, the third consumer will be placed in an informational void and, as a result, she will be directed by her private information. However, regardless of her choice, her decision will initiate an informational cascade because the fourth agent's public information will now exceed her private information (i.e., a store with two customers as opposed to another with only one). Thus, provided there is

sequentiality of choice and sufficient limits on private information, an informational cascade will form either with the third or the fourth adopter.

A Closer Look: Availability Cascades

Availability cascades extend informational cascade theory by incorporating the idea of availability heuristics and adding a second component: reputational concerns (Kuran and Sunstein 1999). Kuran and Sunstein assume actors utilize heuristics to assess the current state of affairs in which they find themselves. This "involves estimating the probability of an event on the basis of how easily instances of it can be brought to mind" (1999, 706). Thus, they argue, actors imitate each other not only because they believe that others have more reliable information (as is the case with informational cascades), but also "because they take the simple availability of information as an indication of its reliability" (Lemieux 2004, 20). This in turn stimulates further adoption of the belief, triggering "a chain reaction that gives the perception increasing plausibility through its rising availability in public discourse" (Kuran and Sunstein 1999, 683). This self-reinforcing process of collective belief is, critically, strengthened by reputational concerns, as individuals change "their public responses in the interest of maintaining social acceptance" (Kuran and Sunstein 1999, 683). Actors adjust their behavior, and in doing so, they reinforce the cascade. Sunstein and Kuran posit the idea of *availability entrepreneurs*, social agents who "attempt to trigger availability cascades likely to advance their own agendas" (Kuran and Sunstein 1999, 687).

An availability cascade can alter people's assessments of risk (the model has found extensive application in the field of regulatory research and finance theory). Availability "shapes, on the one hand, judgments about the magnitudes of various risks and, on the other, the acceptability of these risks" (Kuran and Sunstein 1999, 712). For instance, the availability cascade model informs us that because of its elevated profile in the public conversation, people are likely to overestimate the risk of being wounded in a terrorist attack compared with being injured in the bathroom (in the United States, approximately 235,000 people are injured in bathrooms every year; see Bakalar 2011). This shift in risk perception caused by availability cascades, as we will see, also contributes to breakdowns in online censorial control.

A Closer Look: Norm Cascades

The third cascade type, norm cascades, explains abrupt transformations in societal norms (Sunstein 1996b). Sunstein argues that, for reputational reasons, a certain number of actors conceal their privately held views but may grow more emboldened to express them as others do likewise. Sunstein argues that norms may be far more fragile than is generally believed. He distinguishes between norm bandwagons and norm cascades. According to Sunstein: "Norm bandwagons occur when small shifts lead to large ones, as people join the 'bandwagon'; norm cascades occur when there are rapid shifts in norms" (Sunstein 1996b, 909). A norm bandwagon will emerge when "the lowered cost of expressing new norms encourages an ever-increasing number of people to reject previously popular norms, to a 'tipping point' where it is adherence to the

old norms that produces social disapproval" (Sunstein 1996b, 912).[9] At that point, a norm cascade will erupt, bringing about a sudden collapse of the norm. Norm cascade theory is rather broad in scope, focusing on abrupt wholesale cultural or political reconfigurations, such as the collapse of communism in Eastern Europe and the dismantling of apartheid in South Africa (Sunstein 1996b, 912). Robert Ellickson describes the norm cascade process as follows: "[I]n the paradigm case an upstart norm starts slowly, gains momentum, and culminates in a triumphant rush. Various authors refer to a tipping point being passed, to an equilibrium changing not gradually but in punctuated fashion, or to a cascade being triggered" (Ellickson 2001, 51).

Similar to availability entrepreneurs, Sunstein argues that norm cascades are jumpstarted by *norm entrepreneurs* who attempt to rally support for their views by exploiting "private dissatisfaction with existing norms in order to bring about large-scale social change" (Sunstein 1996b, 929). Sunstein cites political actors such as Martin Luther King, Jr., Ronald Reagan, and Jerry Falwell as examples of norm entrepreneurs (Sunstein 1996b, 929). Certainly, figures such as Adolf Hitler and Mahatma Gandhi would also fall into this camp. For our model, norm cascades are highly relevant because they may play a key role in producing the sometimes sudden and punctuated character of a cyberspeech cascade.

A cyberspeech cascade, which occurs entirely in the context of online behavior, comprises informational cascades, availability cascades, and norm cascades. The two sections that follow explain how these three cascade types combine to produce cyberspeech cascades.

Cyberspeech Cascades from Information

The strategic use of legal vagueness to induce self-censorship creates a curious dynamic. Entire populations are deliberately thrust into a state of informational uncertainty regarding the acceptable limits of online speech. This uncertainty creates a crack through which cyberspeech cascades may enter. The cyberspeech cascade model we propose holds that the following set of assumptions generally pertain across closed-speech societies: (1) a significant percentage of Internet users, out of fear, self-censor and engage in preference falsification; (2) nevertheless, a minority of Internet users will—deliberately or inadvertently—push gently against the perceived limits of open expression (we shall refer to these actors as "cyberspeech entrepreneurs"); (3) these agents engage in extremely limited degrees of non-self-censored speech, crossing ever so slightly the perceived boundaries between nonpermissible and permissible speech (the model requires only very modest acts of contentious speech). If these three assumptions hold, cyberspeech cascades will intermittently occur given the structural properties of Internet communication and cyber censorship's reliance on self-censorship. In this and the following section, we describe how cyberspeech cascades work. We first examine the impact of informational cascades; we then look at the role of reputational motivations.

9. The concept of tipping points was first popularized by Thomas Schelling in his well-known work *Micromotives and Macrobehaviour* (1978).

Informational Cascades

In the previous section, we described an informational cascade using the example of choosing between two stores. Incorrect beliefs about the acceptable boundaries of open online expression can, we argue, cascade in a similar fashion. To apply the model of informational cascades to online speech, we need to substitute certain variables (Bikhchandani, Hirshleifer, and Welch 1992).[10] Rather than a choice between two stores, the competing choices in our model are self-censorship versus non-self-censorship. For expository convenience, we define these two choices as "closed speech" (i.e., self-censored speech) and "open speech" (i.e., non-self-censored speech).[11] Agents possess both private and public information. A speaker's private information consists of the understanding that certain speech is impermissible; her public information is gleaned from the online conduct of other speakers. For the purposes of our model, a distinction must also be made between "cyberspeech bandwagons" and "cyberspeech cascades."[12] The former involves a modest uptick in speakers communicating openly. The latter occurs when multiple bandwagons of open speech converge in a synergistic manner. This takes the form of a large-scale eruption of open speech that ripples across social media and other online speech platforms, causing vast numbers of Internet users to suddenly doubt their previously held perceptions of the permissibility of certain forms of speech.

The essential component for an informational cascade (uncertainty) is built into the strategy of using legal vagueness to induce self-censorship. Because the dividing line between permissible and impermissible speech is left unclear in authoritarian societies (Ford 2015, 24), Internet users are stuck in a state of informational uncertainty. Although legal ambiguity as to permissive speech is very effective in encouraging self-censorship, this lack of clarity can cut both ways. This state of uncertainty may keep a lid on dissent in general, but public perceptions of what constitutes permissible speech are vulnerable to misinterpretation at the perimeters of collective understanding.

Lacking clear demarcations of what constitutes impermissible speech, the average Internet user, on the (unwarranted) assumption that numbers convey reliability, will look to her peers to determine what exactly acceptable speech is. Like the shopper in our above example who, while inclined to shop at Store X, is nevertheless persuaded by the behavior of other consumers to shop at Store Y, an Internet user may be persuaded by a cyberspeech bandwagon's growing numbers to likewise engage in open speech. Bikhchandani, Hirshleifer, and Welch have shown that in a "general setting with sequential choices . . . at some stage a decision maker will ignore his private information and act only on the information obtained from previous decisions. . . . In the absence of external disturbances, so do all later individuals" (1992, 994). This same dynamic may occur with respect to open online speech. Although extreme forms of open speech are clearly understood as impermissible, a great deal of online discourse cannot be easily classified by Internet users, and so on

10. Bikhchandani, Hirshleifer, and Welch (1992) provide a formal model of informational cascades.
11. Henceforth, the terms open speech and closed speech refer specifically to online speech.
12. This mirrors Sunstein's "norm bandwagons" and "norm cascades."

the periphery of what is understood as acceptable expression, the Internet user's private information is extremely weak. It is here, along the margins of public discourse, that the envelope may be incrementally and even inadvertently pushed by the behavior of cyberspeech entrepreneurs.

The model does not require cyberspeech entrepreneurs to engage in extreme forms of political speech, nor does it even have to be performed intentionally. In many instances, cyberspeech entrepreneurs may not even be conscious that they are crossing the acceptable line of expression; their conduct may simply flow from a natural lack of caution and a penchant for audacious speech. This is a unique aspect to our model that distinguishes it from the existing cascade literature.[13] Unlike norm entrepreneurs and availability entrepreneurs, cyberspeech entrepreneurs are not envisioned here as having to push so vigorously or in such an organized manner against prevailing norms and beliefs. Cyberspeech entrepreneurs may act in a highly decentralized, spontaneous, and, most crucially, unintentional fashion. This is due to the unique structural character of online communication[14]—the highly interconnected character of online communication can rapidly amplify small shifts in speech. The interesting component of our model is that the entire process may unfold as an unintended consequence of very small acts. No one may wish for a cyberspeech cascade to emerge, yet one may emerge nevertheless.

In any case, intentional or not, the model assumes that cyberspeech entrepreneurs will only push very gently against the boundaries of permissible speech, and will not do so if it represents a significant departure from their understanding of what is deemed acceptable by the censors. They will not engage in political speech that may threaten their own physical safety and liberty. Thus, each cyberspeech bandwagon will shift perceptions of acceptability only to a very minor, almost negligible degree. The loosening of closed speech is initially slight and proceeds at first incrementally. After building on itself in an almost imperceptible manner, however, the proliferation of cyberspeech bandwagons may eventually reach a point such that a cascade emerges, triggering a large-scale shift in the character of online speech.

A lone cyberspeech bandwagon will produce little more than a slight uptick in open speech. Most of these bandwagons will remain isolated and ultimately peter out. However, because the interlinked structure of the Internet allows for exponential growth, a cyberspeech bandwagon has the potential to spawn additional bandwagons. If this occurs, the process may begin to snowball, and as is the nature of exponential growth, the perceived boundaries of censorship may begin to move more quickly. This process of exponential increase is illustrated in Figure 1.

Figure 1 visually depicts a cyberspeech cascade, describing it on four levels. The first level shows a cyberspeech bandwagon yielding new ones that, in turn, trigger others. This may exceed the rate of one to two, as illustrated in the diagram's fourth level, which depicts a rate of one to three. The important point is that the

13. This feature of the model is discussed in more detail in the section entitled "Interaction Between Cascade Types."

14. The unique structural composition of online behavior is explored in greater detail in the section entitled "The Unique Conditions of Online Interaction."

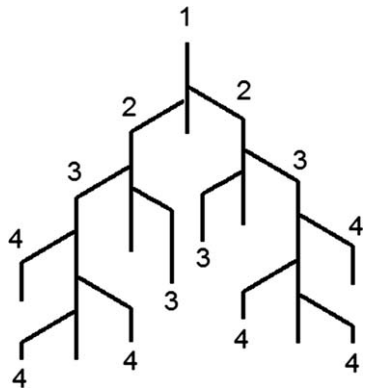

FIGURE 1.
Depiction of a Cyberspeech Cascade

cascade can explode rapidly into multiple layers of growth, and from only an initial trickle, it can eventually capture within its sweep millions of Internet users. At each successive level, there are lesser degrees of self-censorship.

Of course, Figure 1 does not capture the complex reality of cyberspeech cascades. We have simplified it for purposes of illustration. Cyberspeech cascades are rarely so linear and precise in how they unfold—in fact, cyberspeech cascades may result in a matrix of complexity. Such may be the dizzying intricacy of the chain reaction that charting its exact pathways of interconnection would be an exercise in futility.

Regardless of the precise circuitry of growth, however, a cyberspeech cascade will conform to a general pattern. Each successive bandwagon further widens the borders of permissible speech ever so slightly but repeatedly so that the boundaries of what is perceived as acceptable quickly recede. Each bandwagon builds on the one before it, further pushing back the tacitly understood boundaries of permissible speech as people misinterpret the previous bandwagon as reliably signaling that a wider ambit of public expression is now officially tolerated. The scale of each bandwagon may range in number—a mere handful to millions of Internet users could be implicated at various stages.

The process can be explained more technically as follows: Bandwagon 1 yields Speech X, where X signifies a slight increase of open speech. Subsequently, cyberspeech entrepreneurs in Bandwagon 2, misinterpreting that Speech X reliably indicates a greater tolerance of uncensored speech, produce Speech X+1, where 1 denotes an even higher degree of open speech. Bandwagon 3 then builds on Bandwagon 2, with cyberspeech entrepreneurs delivering Speech X+2, which bears even fewer indicia of censorship. In sequential fashion, Bandwagon 4 employs speech X+3. This cycle repeats itself in an iterated fashion until a cyberspeech cascade occurs. The effect of the cyberspeech cascade will be self-reinforcing as it reorders public perceptions regarding what constitutes permissible speech. Previously stable norms of self-censorship collapse as online discourse is overwhelmed by the large-scale eruption of open speech.

The tightly bound and hyper-connected nature of the Internet allows even tiny acts of dissent (even unintentional ones) to reverberate and rapidly gain momentum as they ripple out across the Internet. Like a light touch on a spider's web, this impact is amplified and may be felt throughout the entire system. What began as small shifts in speech patterns eventually turns into a flood of open discourse—eddies become ripples and ripples crest into tidal waves. Ironically, the informational uncertainty that is so effective in inducing self-censorship sets the very conditions for cyberspeech cascades.

Cyberspeech Cascades from Reputation

Informational cascades, however, are just one piece in a cyberspeech cascade. A cyberspeech cascade will invariably consist of a mix of informational cascades and reputational cascades. To the extent that a speaker possesses a strong online identity, reputational motives will come into play—Internet users will jump on cyberspeech bandwagons for reputational reasons in addition to informational reasons. Anonymity between users is thus a crucial variable in our model. If online speech occurs under conditions of anonymity, reputational motives cannot emerge; to the degree that there is a genuine absence of anonymity, however, reputational concerns will kick in.

Online Anonymity Between Users

Anonymity between users varies considerably across Internet platforms.[15] For instance, speech that occurs on social networking services, such as Facebook, is far from anonymous and undoubtedly gives rise to powerful reputational concerns. These reputational concerns are especially robust because these connections usually comprise a speaker's close social ties with users who engage in repeated interactions. Online speech in group e-mail trails and chat groups on instant messaging clients—such as Whatsapp, WeChat, and Tencent QQ—that comprise personal (or semi-personal) ties may likewise produce strong reputational concerns. Although the connections are less intimate, professional networking services, such as LinkedIn and ResearchGate, will also generate powerful reputational motives. Users on micro-blogging platforms such as Twitter and Sina Weibo (China's Twitter) also lack anonymity. Compared with other forms of social media, micro blogging is generally more open to the public and typically consists less of personal connections (in the form of followers). The lack of anonymity on these online speech platforms may give rise to powerful reputational concerns.

However, we must be cautious here. Although social networking platforms like Facebook are anything but anonymous, a great deal of online communication in fact occurs in what is the practical equivalent of anonymity. We give the term *functional anonymity* to the situation in which an Internet user employs a pseudonym (online handles or screen names) that reveals sparse to no information regarding

15. The reader should note that we are referring here to anonymity in relation to other Internet users, not government monitors. A cyberspeech cascade does not occur because users believe monitoring has changed; rather, the cascade erupts because users believe that official tolerance of open speech has changed.

their personal identity. It would be difficult to argue that the use of online aliases generates strong reputational motives. This is especially true when online communication involves predominately one-shot interactions (as opposed to repeated interactions). Moreover, many of these screen names are site specific, with Internet users employing multiple online identities in which they have very little invested.

Indeed, the lack of reputational concerns is clear from the level of vitriol, undisguised racism, sexism, and verbal abuse that saturates online discourse. For instance, the comment threads on YouTube, news sites, blog sites, Internet forums, online chats, multiplayer online games, and other online fora are replete with speech clearly liberated from the constraints of basic civility. It is often the case that with "commenters able to hide behind a cloak of anonymity, the blog and chatroom have become forums for hatred and bile" (Adams 2011). Social norms and reputational concerns that would otherwise keep such speech in check offline are routinely ignored by people when they interact online. This fact speaks to the lack of strong reputational motives in these fora—the result of functional anonymity. It is telling that the caustic discourse so commonplace in the comment threads of YouTube videos is rarely seen on personal Facebook pages. The use of screen names is insufficient to trigger reputational concerns. Reputational concerns will not arise, or if they do, they will be comparatively weak. Although star bloggers or YouTube celebrities with millions of followers may be heavily invested in their online identities, this is not true for their millions of followers operating under random screen names posting the odd comment here and there. Even where stable online identities are used (e.g., Facebook profiles) on speech platforms such as the comment threads beneath news stories, a lack of repeated interaction and the one-shot nature of these interactions inhibit reputational concerns of any significant nature.

It should be noted that the point here is not that Internet users communicating in functional anonymity do so with respect to government monitors—potential surveillance does not change with the use of screen names. The point here is merely that there is no, or very weak, reputational concerns with respect to other users. This is important because it means that fear of social disapproval or a desire for social approval will not play a role under such conditions.

It is important that we really appreciate the varied character of online communication. In some cases, speakers possess well-established online social identities and repeatedly interact; in other cases, speakers operate in functional anonymity. The first will give rise to reputational concerns; the second will not. As a result, online communication produces a mix of cascade types. Because online speech operates across a broad range of speech platforms—some that provide anonymity between users and some that do not—a full-blown cyberspeech cascade will involve informational and reputation-related cascades as it sweeps across the Internet. Where reputational motives are actuated, cascades that involve reputational motives, such as availability and norm cascades, will emerge. Let us now examine, in turn, the impact of each of these.

Availability Cascades

To the extent that they involve a reputational aspect and not simply informational issues, cyberspeech cascades are better captured by the concept of availability

cascades. Where the speaker possesses a robust online identity, availability cascades may emerge because reputational factors will come into play. Unlike simple informational cascades, availability cascades are also driven by powerful reputational motives. Reputational motives such as earning social approval and avoiding disapproval come to the fore (Kuran 1998, 686). Reluctant to deviate from the script, many Internet users join the growing bandwagon because they wish to be seen as adopting the correct position (i.e., engaging in online speech critical of the government). An Internet user asks: "How will my community view me if I fail to endorse prevailing opinion?" This provides incentive to parrot these views in public (Kuran 1998, 729). Reputational concerns can be very powerful when conditions allow. If an actor believes that the acceptable limits of public expression have really changed,[16] otherwise politically indifferent Internet users may engage in online speech that is critical of the government driven solely by a desire to gain social approval and avoid social disapproval. The degree of concern about reputation will vary between people: "people differ from one another along these dimensions; some care enormously about their reputations, while others care very little" (Kuran and Sunstein 1999, 729). However, on the macro level of vast numbers of agents interacting, reputational motives will exert a discernable impact. One need only look at the tidal waves of ideological conformity that wash across Facebook feeds to begin to see how powerful reputational concerns can be in patterning collective online opinion, genuine or not.

Availability cascades also have an informational component. Like informational cascades, Internet users' uncertainty regarding the precise boundaries of permissible speech facilitates the emergence of availability cascades. However, the availability cascade model invokes an informational mechanism that differs from informational cascades. This mechanism, the *availability heuristic*, was briefly discussed earlier.[17] The availability heuristic "involves estimating the probability of an event on the basis of how easily instances of it can be brought to mind. This heuristic can produce substantial distortions whenever certain alternatives are easier to imagine than others" (Kuran 1998, 706). Kuran and Sunstein give the example of a person overestimating the prevalence of AIDS "simply because many of his acquaintances have the disease and he can easily think of AIDS cases" (Kuran 1998, 706). Applied to online speech, the availability heuristic suggests that an Internet user's understanding of what constitutes impermissible online expression will be distorted by how easily she can recall encountering examples of open speech online (this tendency is compounded by the fact that instances of self-censorship are, by definition, not easily discerned). Thus, a rise in the availability of uncensored speech, particularly if this is concentrated in a localized network of Internet users (e.g., between a cluster of interlinked Facebook accounts), can produce the incorrect impression that the acceptable limits of permissible speech have changed.

16. Note that the state of informational uncertainty in which agents operate remains a crucial element.

17. It should be noted that the availability heuristic may manifest through platforms marked by functional anonymity. However, the informational component to availability cascades is treated here together with its reputational component for clarity of exposition and structure.

Availability cascades have been traditionally used to explain misperceptions of increased risk conflated by media attention (e.g., disease threats, airplane crashes, the threat of terrorism). However, availability cascades may also in principle explain a mistaken belief in decreased risk—that is, the incorrect perception, bolstered by its simple availability, that engaging in open speech online has become less risky when in fact it has not. With their perception of risk now slightly altered, Internet users become more inclined to engage in open speech themselves, and in so doing, further increase the availability of open speech. This triggers a chain reaction of individual responses that make the perception that a wider range of open speech is now being tolerated by the censors appear increasingly plausible simply by virtue of its growing online availability, when in fact a wider range of open speech is not being officially tolerated by the censors (Kuran 1998, 685).

Reputational concerns will have an impact *only* if actors believe it is safe to engage in open speech. This needs to be emphasized. The average Internet user will not risk personal safety, life, and liberty merely for reputational gains. The perception that a wider ambit of liberal expression is actually being tolerated by the censors remains a necessary condition. When this condition is satisfied, however, reputational motives can powerfully reinforce shifts in collective perception, adding further fuel to the fire. As people join growing bandwagons of open speech motivated by reputational concerns, other users will misinterpret this increase in availability as an indication that the censors have grown more tolerant of open speech.

Norm Cascades

Reputational motives are also very useful in accounting for the sudden, punctuated shifts in Internet self-censorship patterns that have been observed in many instances.[18] Sunstein's norm cascade model describes in a very detailed fashion how this may come about. Norm cascades are fascinating in how they play upon collective fear and distort mass perceptions. Again, uncertainty regarding the precise boundaries of permissible speech provides the initial space for norm cascades to arise; however, the reputational component operates in a distinctive manner. The norm cascade model holds that "social conditions are often more fragile than might be supposed, because they depend on social norms to which—and this is the key point—people may not have much allegiance" (Sunstein 1996b, 909). Many individuals, driven by reputational fear, engage in preference falsification, and so wait in the wings for an opportunity to defy the prevailing norm. Like norm cascades, the present model assumes that a significant percentage of Internet users self-censor out of fear, engaging in preference falsification. An uptick in open speech, however, lowers the perceived danger in engaging in open speech. This shift in perception emboldens Internet users previously reluctant to engage in open speech to join the growing bandwagon.

This continues until a tipping point from predominately closed speech to predominately open speech is reached. At this stage, the cost structure flips so that it

18. We discuss several examples of this punctuated change in self-censorship norms in the section entitled "The Unique Conditions of Online Interaction."

is reputationally more costly to appear out of step with the growing chorus of speech critical of the government than remain silent. Misperceiving that the official boundaries of permissible speech have now shifted, large numbers of Internet users who previously self-censored will cease doing so. Even more crucially, however, Internet users otherwise indifferent, motivated by the reputational consequences of not conforming, become more likely to engage in such speech. Alexis de Tocqueville described the dynamic well: "More frightened of isolation than of committing an error, they [join] the masses even though they [do] not agree with them" (de Tocqueville 1856, 259). When this stage is reached, a communal shift in self-censorship norms may occur (Roe 1996, 663–65; Picker 1997, 1250–51) in the punctuated, sudden "triumphant rush" described by Ellickson (2001, 51).

This inversion in collective fear has close parallels to the spiral of silence theory in the political science literature and mass communication theory (Noelle-Neumann 1974, 43–51). The theory holds that as a particular view begins to gain ascendency in a social environment, individuals adjust their behavior accordingly. The result is that those holding the view become more vocal and those who do not become less willing to speak out, triggering "a spiraling process which increasingly establishes one opinion as the prevailing one" (Noelle-Neumann 1974, 44). Spiral of silence theory explains the same en masse shift in social behavior described by the norm cascade model. Fear of voicing a minority opinion cows people into silence, and those holding a majority view grow more emboldened to voice their view.

Interactions Between Cascade Types

Which cascade type emerges through which online platform ultimately depends on the degree of anonymity between users. However, once arisen, informational cascades, availability cascades, and norm cascades will interact, creating "cascade synergies" that can accelerate the wholesale collapse of Internet self-censorship systems.

Mutually Reinforcing

Given the character of online communication, the potential for informational cascades is the greatest. This is because informational cascades are not driven by reputational motives, and as such, may emerge across any speech platform. Availability cascades and norm cascades, on the other hand, may only emerge through speech platforms where reputational motives come into play—that is, where speakers possess a strong online identity. Nevertheless, as cascades of open speech gain momentum and flood across a mixture of online speech platforms, all three cascade types will invariably be ignited. Once this occurs, informational cascades, availability cascades, and norm cascades will feed off one another.

As changes in speech patterns move into areas of the Internet that give rise to reputational concerns, informational cascades will help produce availability and norm cascades by providing the initial uptick in open speech critical to triggering these cascade types. In the cascade literature, availability cascade and norm cascade

models require availability entrepreneurs and norm entrepreneurs to achieve this. The present model does not require such committed campaigners for social change. Shifts in speech patterns may occur from the unintended actions of Internet users and extremely trivial incidences of open speech that quickly build on one another in a sequential manner.

In the other direction, availability and norm cascades can in turn powerfully reinforce cyberspeech bandwagons that begin as simple informational cascades. As cyberspeech bandwagons flood into areas of Internet communication that generate reputational concerns, such as Facebook pages or other forms of social media, fear of social disapproval and a desire for social approval will further incentivize actors to engage in open speech. In this manner, cascades in one domain can trigger or reinforce cascades in another. Although informational cascades and cascades that involve reputational concerns may initially emerge in different sectors of the Internet, once arisen, these cascade models may reinforce each other like the three legs of a tripod.

This is an important point because simple informational cascades, the literature tells us, are extremely fragile—they can be reversed by new public information (Bikhchandani, Hirshleifer, and Welch 1992, 994). Conformity "is brittle. The arrival of a little information . . . can shatter an informational cascade" (Bikhchandani, Hirshleifer, and Welch 1992, 1004). It is thus not difficult to dam up an informational cascade. The new public information "needs only to offset the information conveyed by the action of the last individual before the start of the cascade, even if millions subsequently imitated" (Bikhchandani, Hirshleifer, and Welch 1992, 1006). As such, cyberspeech cascades are not a perfectly stable equilibrium. Cascades of open speech are themselves just as vulnerable to changes in counterinformational cascades.

Reputational Concerns Make Informational Cascades More Robust and Harder to Reverse

This can be discerned in the government control strategies employed by state censors. For example, the use of "Internet commentators" by Chinese censors (some estimates put the number of these Internet commentators as high as 300,000; see D'Hooghe 2014, 55) to guide and redirect online discourse by writing responses to online postings (So 2012, 449) is designed precisely to achieve this end—reversing growing cyberspeech bandwagons by introducing counterbandwagons. Russian censors have similar programs, employing an army of trolls in an effort to shape online speech.[19] The Kremlin's Internet Research Agency systematically spreads false information, posting "pro-Kremlin propaganda online under fake identities . . . in order to create the illusion of a massive army of supporters," smear political opposition, and neutralize the impact of online speech of an antigovernment flavor.

This administrative arm of the Russian government—informally referred to as the Troll Farm (Chen 2015)—strategically shapes public opinion by creating content across various social networks, such as Twitter, Facebook, VKontakte (Russia's

19. A troll is Internet slang for an Internet user who deliberately posts provocative messages.

Facebook), and the comment threads of Russian news outlets (Chen 2015). While it is unlikely that the Chinese or Russian censors conceptualize their efforts in terms of cyberspeech bandwagons and cascades, the goal of such soft control methods is to redirect the flow of online discourse so as to thwart the emergence of cyberspeech cascades (Rawnsley and Rawnsley 2015). These methods of curating online discourse constitute the external disturbances identified by Bikhchandani, Hirshleifer, and Welch (1992) that may disrupt an informational cascade.

Cascades that possess a reputational component (i.e., availability and norm cascades), however, are far more resilient than lone informational cascades (Lemieux 2004, 20).[20] This is because the addition of "reputational factors to informational motivations implies that individuals will ignore their private signals even more often, and that the information made available by everyone's choices will be even more biased toward following the crowd" (Lemieux 2004, 20). As such, when informational cascades combine with availability and norm cascades, an eruption of open speech on the Internet may be quite difficult to stop once begun.[21]

The Unique Conditions of Online Interaction

Cyberspeech cascades are unique because they occur online. We posit that the structure of online discourse renders online communication uniquely susceptible to cascades.[22] The informational cascade model has been criticized in the same vein as most economic models—as being oversimplified, crude representations of reality (Lemieux 2004, 17). There is, of course, merit to this position. Human behavior is not always as systematic as economists like to portray it. However, we argue that both the informational cascade model and the other cascade models discussed here grow even more robust in the context of online behavior. There are basic structural reasons for this.

Sequentiality and a Clearer Awareness of Others' Behavior

Informational cascades are open to the criticism that they assume perfect sequentiality with each agent able to observe the behaviors of all the agents before her (Shiller 1995, 183). Actors make decisions sequentially, "with later people watching the actions of earlier people, and from these actions inferring something about what the earlier people know" (Easley and Kleinberg 2010, 484). Yet, situations that exhibit perfect sequentiality are in reality more the exception than the norm (Shiller 1995, 182–83). The Internet, however, better simulates this with users behaving in a highly structured, usually sequential fashion. Indeed, it is often perfectly sequential.

20. Lemieux is referring specifically to availability cascades, but the reasoning also extends to norm cascades.

21. This would suggest that censors would be best served by concentrating their efforts in fora that allow for reputational concerns, such as social media.

22. The section that follows draws from an earlier article by the lead author. See Druzin and Li (2016).

One need only glance at the standard layout of chat rooms, message boards, micro-blog posts, social media such as Facebook, e-mail threads, and the comments to online news articles to appreciate their sequential structure: commenters respond in a successive fashion, able to view previous comments (Holt 2004, 15–16; Adams 2007).[23] The number of likes on a Facebook page or post proceeds in a sequential fashion and, as such, is susceptible to a bandwagon effect as the number of likes reaches a tipping point. This sequential nature is true for Internet chatter as a whole: each user has access, either in its entirety or partially, to the past online behavior of other users—literally at their fingertips.[24]

In the real world, agents do not always act in such a well-formatted sequential pattern and, in general, are not perfectly able to view the behavior of preceding agents. Thus, the informational cascade model's assumption that the behavior of all the previous agents is known is better approximated in an online setting and more readily maps onto online behavior than it does onto other informational environments. This heightened availability of information also increases the likelihood of availability and norm cascades. Belief patterns can proliferate and cluster in semi-insulated corners of the Internet, producing an increased availability of other actors' behavior. The structure of Internet communication allows for clear and immediate access to these availability swells.

Connectivity and Speed

Behavioral cascades apply more robustly to Internet-based communities also in terms of the networked character and speed of online communication. During the 2009 antigovernment protests in Iran (referenced again later), Iranians began taking to city rooftops in a nightly ritual to shout "Allahu Akbar!" ("God is Great!") as a form of protest (Mackey 2009). With the practice spreading and in danger of cascading, authorities promptly cracked down on the practice, successfully extinguishing its growth. The speed and hyper-connectivity of online communication, however, renders similar methods of cascade control more challenging to pull off in the cyber world.

The hyper-degree of connectivity the Internet affords is nowhere approximated outside of the cyber realm (Friedman 2007, 516–17). The Internet connects vast numbers of people, transcending geographical, class, and social boundaries. This connectivity is currently experiencing explosive growth with the increased usage of smartphones and other Internet-enabled electronic devices (OECD 2012, 21). The Internet is now in its second stage of development, "evolving from a data network connecting PCs with wires to a much broader network reaching a wide range of new portable devices such as mobile phones and tablet computers" (OECD 2012, 21). The highly interconnected nature of the Internet makes cyberspeech cascades even more likely to occur in that it accelerates the transmission of information, and

23. Holt notes this sequential structure. Likewise, P. C. Adams argues that the structure of online discourse is best conceptualized as a tree diagram.

24. Online tools such as hashtags—a type of online tagging system—increase the ease with which users can locate messages with a specific theme or content.

where communication is not anonymous, will produce powerful reputational concerns. What was once referred to as a global village has now arguably become a global parlor: the collective chatter and discussion of large masses of people is occurring on a scale hitherto unimagined in human history (Ren and Zhai 2013, 112–13). The astonishing speed at which information can now be transmitted complements this hyper-connectivity. This speed of communication amplifies the effects of informational availability, as well as norm cascades, making sudden perceptual shifts regarding what constitutes acceptable online speech far more likely than in the offline world.

Bounded Rationality

Another reason suggesting that the informational cascade model and the availability cascade model better track online behavior is the limited knowledge that participants have of their fellow netizens. Human beings are not engines of perfect rationality. This is captured by the concept of bounded rationality. Bounded rationality is simply the idea that actors' decisions are limited by the information they have, the cognitive limitations of their minds, and the time available to make decisions. The informational cascade model assumes that actors are boundedly rational—that is, they make their decisions within a field of incomplete information (Bikhchandani, Hirshleifer, and Welch 1992). A key component of the model is that persons are unable to observe others' private information directly, but they may make inferences about this private information based on what these other persons do.

Each Internet user is operating from a position of imperfect information regarding other agents. Users only see the online behavior of other agents, but they have no way of knowing if the actions of these other actors are predicated on authoritative information. As a result, users often assume (wrongly) that other actors are working with more perfect information and are therefore inclined to interpret their behavior as conveying reliable information. In our example of the shopper choosing between two stores, if the latecomer shopper does not directly inquire about the decision process of all the other shoppers, she has no way of knowing that their decision to shop at that store was merely the result of an informational cascade. This is precisely the case with online behavior. The behavior of others is highly decontextualized. There is seldom any way of discovering the motivation of other Internet users and, as result, a general impression that other agents are acting on bona fide information will tend to arise, particularly if we see the same behavior in very large numbers.

Bounded rationality likewise enhances the strength of availability cascades. The cognitive biases that give rise to an availability heuristic are the result of bounded rationality. Essential to Kuran and Sunstein's model is that a society is "composed of boundedly rational individuals who benefit immensely from using cognitive rules of thumb" (Kuran and Sunstein 1999, 690). This cognitive limitation forces actors to rely on heuristics to assess "the magnitude of a risk or the seriousness of an alleged social problem" (Kuran and Sunstein 1999, 689). Because Internet users are operating under conditions of limited information, they are frequently

more boundedly rational than offline actors. Hence, we argue, both informational and availability cascades may emerge more robustly through online communication, as the conditions are ripe for their emergence.

Better "Rational" Actors

One final point should be made. Informational, availability, and norm cascades assume a rational-choice model: a cascade occurs because each actor rationally makes her decision based on the behavior of preceding actors or seizes the opportunity to enhance her reputational standing (Sunstein 1996b, 935; Kuran and Sunstein 1999, 689). Given the cognitive limitations under which actors operate, it is rational to look to the informational signals of others (Kuran and Sunstein 1999, 689). Likewise, it is rational to be concerned about one's reputation (Kuran and Sunstein 1999, 689). Internet users are arguably better rational actors. In the non-cyber world, a host of variables may reframe the information upon which a player is acting. For instance, an actor may attribute the fact that Store Y is full of shoppers not to the fact that Store Y's products or prices are better, but to the fact that its awning is a more attractive color. This additional information may blunt the impact of an informational cascade. On the reputational side of the ledger, knowledge that other actors are merely jumping on a bandwagon in an attempt to merely gain social approval may deflate the reputational appeal of also joining the bandwagon.

However, with respect to online speech that occurs on online bulletin boards or is conveyed through brief tweets via social media, there is a relative paucity of information (particularly where functional anonymity predominates). A vast amount of context is stripped away. The behavior of others is often reduced to a string of utterances, a like on a Facebook page, or a reposting of a message or image. Actors thus form their impressions based on the very narrow slice of information provided them. Such a myopically focused set of information is, arguably, more fertile soil for informational and availability cascades given that in cyberspace a great deal of context—other actor's motivations, external considerations, even tone—is absent. Behavior reinforced by the credibility of large numbers will exert a more powerful impact because there is less information that may correct the mistaken inferences drawn. Under such conditions, agents are better "rational" actors, even if the conclusions they draw are in error.

For these reasons, we argue that online speech is especially susceptible to sudden breakdowns in self-censorship norms—more so than traditional conduits of communication. Many authoritarian governments have not adjusted to this new reality.[25] Sudden upticks in open speech online, especially in the initial stages, if unchecked, with the passage of time can become increasingly difficult to dam. By dissipating the sense of collective paranoia that undergirds self-censorship, an unexpected explosion of open speech online in the form of a cyberspeech cascade can seriously destabilize, and ultimately bring down, an authoritarian government's control over the Internet.

25. Perhaps China, the world's foremost Internet censor, is an exception here, demonstrating exceptional skill in suppressing contentious online speech.

Examples of Cyberspeech Cascades and the Growing Significance of Online Communication

The aim of this last section is to highlight examples of cyberspeech cascades in order to provide empirical support for our model. We list several examples below. It should be noted, however, that cyberspeech cascades provide only a partial explanation of these events. Our model only relates to these political upheavals in the limited sense that these events were facilitated by the initial breakdown of online self-censorship. To be sure, a myriad of exogenous factors contributed to the cases cited below. The Arab Spring, for instance, was a highly dynamic process, and cyberspeech cascades cannot solely account for the massive sociopolitical transformation that occurred. Cyberspeech cascades are better understood as a catalyst than as a sole cause of these events. Indeed, in many cases, government responses to the sudden shifts in societal speech patterns themselves played a pivotal role in the spread and impact of changing norms regarding acceptable speech. Indeed, in many instances, the overreaction and strong-arm tactics of authorities precipitated the political dissent these governments were trying desperately to contain.

However, a necessary condition in all of the cases listed below was an initial, sudden disintegration of a collective fear regarding the permissible boundaries of online expression. With this constraint cleared away, a surge of open speech online facilitated the emergence of collective action offline (in the cases where such action followed). Cyberspeech cascades' usefulness is in accounting for this initial, decisive breakdown in online self-censorship norms. Cyberspeech cascades only relate to the misperception that the censors have grown more tolerant of open online speech. Critically, this misunderstanding breaks down self-censorship patterns in online speech, which then allows for the possibility of more significant offline dissent. During the Arab Spring, for example, the use of social media such as Twitter and Facebook played a significant role in facilitating the protests that flared up across North Africa and the Middle East (Howard and Hussain 2012, 110; Adi 2014, 23–28; Druzin and Li 2015). We argue that online cyberspeech cascades (consisting of informational, availability, and norm cascades) assisted these events by first collapsing norms of online self-censorship within these countries—a crucial initial step to what then followed.

Examples of cyberspeech cascades, however, need not be as dramatic as in the case of the events that occurred during the Arab Spring. Online cyberspeech bandwagons are undoubtedly a common occurrence in closed-speech societies; it is simply that these bandwagons emerge and dissipate without ever evolving into cyberspeech cascades and are thus quickly forgotten. Numerous large-scale cyberspeech cascades of a distinctly political nature have erupted just over the last decade. Table 1 lists several major cyberspeech cascades by country in chronological order. These examples of public anger rapidly gathered momentum through online discussion, eventually growing into cyberspeech cascades. Although the list is by no means exhaustive, these examples reach back ten years and appear to be growing more numerous, although a more rigorous empirical study would be required to confirm that such events are becoming more frequent, as Table 1 suggests.

TABLE 1.
Examples of Notable Large-Scale Cyberspeech Cascades Across Various States

Belarus	2006: Antigovernment chatter emerged online in the wake of presidential elections widely considered rigged. Calls for action began spontaneously flaring up on Internet blogs and discussion forums as online traffic surged to key sites, culminating in "unconventional protests activities, such as flash mobs, in the Belarusian capital following the election" (Bulhakau 2006, 82–83).
China	2008: A China Central Television (CCTV) report on the need for Internet restrictions featured a story involving an elementary school student who allegedly stumbled upon Internet pornography. Accusations that the story was staged began to proliferate online. Condemnation eventually audaciously decried all CCTV reporting as propaganda (Herold and Marolt 2011, 138).
Egypt	2008: A Facebook group in support of protesting millworkers gradually "attracted 70,000 followers and eventually coalesced into a political movement known as the 'April 6 Youth Movement'" (Burnell, Randall, and Rakner 2014, 168).
Iran	2009: Video recording of a female Iranian student, Neda Agha-Soltan, shot and killed by Iranian police, began circulating online. The forty-second video was reposted with increasing frequency with netizens "sharing links on various blogs on the internet, discussing the case via social media platforms etc." (Basten 2011, 4). The sharing of the image became an expression of solidarity with those protesting the legitimacy of Iran's presidential elections, and the incident featured prominently in the street demonstrations that followed, with protesters displaying portraits of the assassinated woman.
Iran	2009: The use of Twitter escalated quickly throughout the period in and around the Iranian presidential elections. It reached an unprecedented 200,000 tweets per hour during the elections and was utilized to coordinate a massive protest three days later. As the use of social media surged by protesters, photos, videos, blog posts, tweets, and SMS messages flowed between protesters (Featherman 2015, 23).
China	2009: An Internet meme and image of a fictional animal called a "grass-mud horse," a parody of government Internet censorship, swept China's Internet, going viral (Wines 2009). The name, when pronounced in Mandarin, is a double entendre for a popular obscenity, and became a powerful semiotic device to protest Internet censorship. Interestingly, the incident illustrated how a cascade may not technically even come in the form of speech—expression may come in the form of the posting and sharing of images.
Egypt	2010: Images of a young Egyptian man, Khaled Said, who was beaten to death by police was posted online. An anonymous human rights activist created a Facebook page "We are All Khaled Said." Within days of Said's death, 130,000 Egyptians joined the page to "get and share updates about the case." The page became the biggest dissident Facebook page during the toppling of the Mubarak regime with close to a half a million users (Preston 2011). By February 2011, the Facebook page garnered over 600,000 likes and was a key organizational conduit for the protests that eventually toppled the Mubarak regime (Alexander 2011).
Tunisia	2010–2011: Video of the self-immolation of a Tunisian protester was posted online. Facebook events spreading information on upcoming protests began popping up. Within days, a well-known Tunisian rap artist uploaded an anti-regime anthem on his Facebook page. These actions of online dissent were initial waves that eventually culminated in the toppling of the Tunisian government in January 2011 (Cammett et al. 2015, 426).

Table 1. *Continued*

Bahrain	2011: As demonstrations against the ruling political establishment began in February 2011, the number of tweets of an antigovernmental character escalated rapidly (Youngs 2013, 166).
Syria	2011: After a Syrian man soaked himself in gasoline and set himself on fire, opposition groups took to Facebook and Twitter in a failed call for demonstrations (Taylor 2014, 87). However, netizens were emboldened and a Facebook page called "The Syrian Revolution 2011 Syrian revolt against Bashar al-Assad" quickly garnered over 41,000 fans, video of the protests emerged on online, and Syrian Twitter chatter escalated (Flock 2011).
China	2011: Online discussion surrounding the collision of two high-speed trains in Wenzhou unreported by state media quickly escalated into widespread indignation on China's micro blogs. Within two days, 5.3 million posts emerged on Sina Weibo (Xu 2011). The story eventually attracted over 500 million posts (Denyer 2016). The event illustrated that "the speed with which information can be circulated defies the keyword-based filtering that Chinese authorities use to censor the Internet" (Tkacheva et al. 2013, 105–07).
Russia	2011: Prompted by online protest, first thousands and then tens of thousands of Russians took to the streets to protest against electoral fraud in the December parliamentary elections (White, Sakwa, and Hale 2014, 140). Key catalysts to these demonstrations were an "explosion of internet use that eroded the dominance of state television . . . online political networks," and the role of key cyberspeech entrepreneurs (White, Sakwa, and Hale 2014, 140).
China	2011: Demands to relocate a damaged chemical plant were "enabled by the speed of the Internet and the number of people using it" (Tkacheva et al. 2013, 108). One day after discussions surged online, 12,000 demonstrators converged in Dalian. Censors were unable to delete posts quickly enough to prevent the offline protests (Tkacheva et al. 2013, 108).

All these events involved a sudden eruption of open expression online. In some cases, this loosening of online self-censorship spilled offline, taking the form of mass demonstrations and civil disobedience. This was certainly the case with the events that unfolded in Egypt and Tunisia over 2010–2011, which culminated in regime change in both those countries. Other cyberspeech cascades outlined in Table 1, however, were only momentary breakdowns in online self-censorship norms. Some, for instance, the events of 2009 in Iran, generated significant social unrest yet fell short of regime change. Other events, such as the 2009 online mass condemnation of CCTV as propaganda and the 2011 Wenzhou high-speed rail event in China, were merely temporary eruptions of open speech that bore no offline consequences. They were merely online tremors that saw no earthquake.

In fact, although it is difficult to say with certainty, the vast majority of cyberspeech bandwagons likely never even come close to escalating into full-blown cyberspeech cascades. However, cyberspeech bandwagons on any scale pose a threat to online censorship. If a cyberspeech bandwagon is not promptly quelled, it may spawn multiple cyberspeech bandwagons, with each igniting more bandwagons, and each further pressing back the perceived boundaries of officially sanctioned speech. This poses an enormous challenge for government censors. To prevent the emergence of large-scale cyberspeech cascades, censors must diligently extinguish them

TABLE 2.
A Tabulated List of Popular Internet Speech Platforms Showing the Approximate Number of Registered Users and a Description of the Platform

Speech Platform	Number of Users	Description
Facebook	1.65 billion	General social networking
YouTube	1 billion	Video sharing
WhatsApp	1 billion	Instant messaging
Tencent QQ	829 million	Instant messaging
WeChat	762 million	Instant messaging
LinkedIn	400 million	Business networking
VKontakte	360 million	General social networking
Twitter	310 million	Social micro blogging
Sina Weibo	222 million	Social micro blogging

Sources: Freier (2015); Caruso (2016); Chen (2016); Globe and Mail (2016); Isaac (2016a, 2016b); Sisario (2016); Stewart (2016); Wingfield (2016).

as they begin to form. As Internet penetration grows in closed-speech societies, strategies of direct censorship such as deleting contentious speech will become increasingly difficult. The situation may devolve into a game of whack-a-mole, where censors snuff out cyberspeech bandwagons only to have many more pop up. Small cyberspeech bandwagons are not themselves so dangerous; however, their potential to spark cyberspeech cascades and reshape collective perceptions represents a potential challenge to the authoritarian state's control over the flow of online information and the regimentation of public opinion more generally.

The massive numbers under discussion here should be properly appreciated. Gone are the days when Internet communication was a novelty reserved for academics and economic elites in the developed world with Internet connections. It now represents a substantial channel of communication for humanity—close to half the world's population is now online (World Bank 2016, 8). The approximate number of users for just the few Internet speech platforms referenced in this article is jaw-dropping (Table 2).

Social networking platforms Qzone (China's Facebook) and Google+, not mentioned here, boast 645 million and a stunning 1.6 billion users, respectively (Bischoff 2014). The Internet speech platforms listed in Table 2 are growing so quickly that these numbers will almost certainly be outdated by the time these words are read. Given the growing importance of Internet communication, cascade theory needs to be reexamined. Trends in mass social behavior driven by this stunning degree of hyper-connectivity are uniquely susceptible to cascades. This has deep sociopolitical and cultural implications that are, however, mostly beyond the scope of this short discussion.

CONCLUSION

This article has argued that, while outwardly robust, Internet censorship is in fact more fragile than is generally supposed. The article sets out a behavioral

economics model that explains sudden breakdowns in cyber censorship, weaving together separate strands of cascade theory—informational cascades, availability cascades, and norm cascades—to create the concept of a cyberspeech cascade. A cyberspeech cascade occurs when expressions of contentious online speech (deliberate or even unintentional) trigger small shifts in public perception regarding the permissible limits of online expression that then proliferate into large-scale torrents of uncensored speech. Changes in perceptions build on prior shifts, becoming the catalyst for further change. By triggering online chain reactions of open speech, each progressively more brazen, these communication swells can cause the temporary (or permanent) collapse of self-censorship norms online.[26]

The model suggests that as Internet penetration grows across the planet, absent breakthrough leaps in technology, censorial regimes will need to rely increasingly on self-censorship, which will render such systems of control more vulnerable to unexpected outbreaks of cyberspeech cascades. The fact that self-censorship is a major component of the authoritarian state's ability to control the Internet is problematic for such governments in that a susceptibility to cyberspeech cascades is structurally built into online interaction. The implications of this for the state's ability to maintain control over online speech are far-reaching. Aung San Suu Kyi once observed that for citizens in authoritarian regimes, "fear tends to be the order of the day . . . yet even under the most crushing state machinery courage rises up" (2010, 185).

As this article has demonstrated, in the context of cyber censorship, courage may be little more than timid, often unintended acts of slightly bolder expression. Yet such discourse, amplified through the interconnected and rapid nature of online communication, may be quite powerful. Recent history offers several examples of cyberspeech cascades piercing the gossamer veil of self-censorship in illiberal states. China, Iran, Egypt, Belarus, Bahrain, Syria, Russia, and Tunisia—each of these governments has had to deal with this threat to varying degrees over the past decade. Most of these states have managed to keep a lid on dissent; however, given finite human resources and the limits of monitoring technology, these governments remain vulnerable to future cyberspeech cascades with, potentially, more permanent consequences. Perhaps technology will eventually be equal to the task of snuffing out each and every expression of online dissent. Until then, however, the potential for cyberspeech cascades will remain a threat to the authoritarian state.

REFERENCES

Adams, Paul C. 2007. *Atlantic Reverberations: French Representations of an American Presidential Election*. Burlington, VT: Ashgate.

26. It should be noted that the model may be applied to any taboo view provided there is sufficient collective uncertainty regarding the acceptability of its online expression (racism, sexism, etc.). It is, however, especially relevant to regimes of cyber censorship because users within such systems are deliberately placed in a state of informational uncertainty and preference falsification is rampant. A broader application of the model to online speech more generally forms the basis of related research currently being explored by the lead author (Druzin).

Adams, Tim. 2011. How the Internet Created an Age of Rage. *Guardian*, July 24. https://www.theguardian.com/technology/2011/jul/24/internet-anonymity-trolling-tim-adams (accessed February 3, 2017).

Adi, Mohammad-Munir. 2014. *The Usage of Social Media in the Arab Spring*. Munster: Lit Verlag.

Alexander, Anne. 2011. Internet Role in Egypt's Protests. *BBC*, February 9. http://www.bbc.com/news/world-middle-east-12400319 (accessed March 1, 2017).

Allen-Ebrahimian, Bethany. 2015. The "Chilling Effect" of China's New Cybersecurity Regime. *Foreign Policy*, July 10. http://foreignpolicy.com/2015/07/10/china-new-cybersecurity-law-internet-security/ (accessed February 3, 2017).

Anderson, Lisa R., and Charles A. Holt. 1997. Information Cascades in the Laboratory. *American Economic Review* 87:847–62.

Bakalar, Nickolas. 2011. Watch Your Step While Washing Up. *New York Times*, August 15. http://www.nytimes.com/2011/08/16/health/research/16stats.html?_r=0 (accessed April 2, 2016).

Banerjee, Abhijit V. 1992. A Simple Model of Herd Behavior. *Quarterly Journal of Economics* 107:797–817.

Basten, Nathali. 2011. *(In)Visibility of Death: Neda—Constructions of Gender and Identity in the Media*. Munich: GRIN Verlag.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades. *Journal of Political Economy* 100:992—1026.

Bischoff, Paul. 2014. Tencent Owns 3 of the World's 5 Biggest Social Networks. *Tech in Asia*, November 3. https://www.techinasia.com/tencent-owns-3-worlds-5-biggest-social-networks (accessed February 3, 2017).

Boas, T. 2004. Weaving the Authoritarian Web. *Current History* 103:438–43.

Brown, Ian, and Christopher T. Marsden. 2013. *Regulating Code: Good Governance and Better Regulation in the Information Age*. Cambridge, MA: MIT Press.

Bulhakau, Valer. 2006. *The Geopolitical Place of Belarus in Europe and the World*. Warsaw: Fundacja Przestrzeni Obywat.

Burnell, Peter, Vicky Randall, and Lise Rakner. 2014. *Politics in the Developing World*. Oxford: Oxford University Press.

Cammett, Melani, Ishac Diwan, Alan Richards, and John Waterbury. 2015. *A Political Economy of the Middle East*. Boulder, CO: Westview Press.

Caruso, Catherine. 2016. Can a Social-Media Algorithm Predict a Terror Attack? *MIT Technology Review*, June 16. https://www.technologyreview.com/s/601700/can-a-social-media-algorithm-predict-a-terror-attack/ (accessed February 4, 2017).

Chen, Adrian. 2015. The Agency. *New York Times*, June 7, MM57.

Chen, Lulu Yilun. 2016. $7,000-a-Month "Shameless China" Blogger Loses All with One Post. *Bloomberg News*, July 18. http://www.bloomberg.com/news/articles/2016-07-17/-7-000-a-month-shameless-china-blogger-loses-all-with-one-post (accessed February 3, 2017).

de Tocqueville, Alexis. 1856. *L'Ancien Régime et la Révolution*. Paris: Michel Levy Freres.

Deibert, Ronald. 2010. *Access Controlled: The Shaping of Power, Rights, and Rule in Cyberspace*. Cambridge, MA: MIT Press.

Deibert, Ronald, John Palfrey, Rafal Rohozinski, and Jonathan Zittrain 2011. *Access Contested: Security, Identity, and Resistance in Asian Cyberspace*. Cambridge, MA: MIT Press.

Deibert, Ronald, John Palfrey, Rafal Rohozinski, Jonathan Zittrain, and Janice Gross Stein. 2008. *Access Denied: The Practice and Policy of Global Internet Filtering*. Cambridge, MA: MIT Press.

Denyer, Simon. 2016. The Internet Was Supposed to Foster Democracy. China Has Different Ideas. *Washington Post*, July 10. https://www.washingtonpost.com/world/asia_pacific/the-internet-was-supposed-to-foster-democracy-china-has-different-ideas/2016/07/10/42954bbc-1dd9-11e6-b6e0-c53b7ef63b45_story.html (accessed February 3, 2017).

D'Hooghe, Ingrid. 2014. *China's Public Diplomacy*. Leiden: Martinus Nijhoff.

Druzin, Bryan, and Jessica Li. 2015. The Art of Nailing Jello to the Wall: A Reassessment of the Political Power of the Internet. *Journal of Law & Policy* 40:1–39.

——. 2016. Censorship's Fragile Grip on the Internet: Can Online Speech be Controlled? *Cornell International Law Journal* 49:369–414.

Easley, David, and Jon Kleinberg. 2010. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge: Cambridge University Press.

Ellickson, Robert. 2001. The Evolution of Social Norms: A Perspective from the Legal Academy. In *Social Norms*, ed. Michael Hechter and Karl-Dieter Opp, 35–75. New York: Russell Sage Foundation.

Featherman, Chris. 2015. *Discourses of Ideology and Identity: Social Media and the Iranian Election Protests*. New York: Routledge.

Flock, Elizabeth. 2011. Syria Revolution: A Revolt Brews Against Bashar al-Assad's Regime. *Washington Post*, March 15. http://www.washingtonpost.com/blogs/blogpost/post/syria-revolution-revolt-against-bashar-al-assads-regime/2011/03/15/ABrwNEX_blog.html (accessed February 4, 2017).

Ford, Christopher A. 2015. *China Looks at the West: Identity, Global Ambitions, and the Future of Sino-American Relations*. Lexington: University Press of Kentucky.

Freier, Anne. 2015. Sina Weibo Revenue and Statistics. *Business of Apps*, December 7. http://www.businessofapps.com/sina-weibo-revenue-and-statistics/ (accessed March 2, 2017).

Friedman, Thomas L. 2007. *The World Is Flat 3.0: A Brief History of the Twenty-First Century*. New York: Picador.

Gibson, Timothy A., and Mark Douglas Lowes. 2007. *Urban Communication: Production, Text, Context*. New York: Rowman & Littlefield.

Glanville, Jo. 2010. *Brave New Words: Is Technology the Saviour of Free Speech?* London: Sage.

Globe and Mail. 2016. Tencent's QQ Web Browser May Put Users' Data at Risk: Report. *Globe & Mail*, March 29. http://www.theglobeandmail.com/technology/popular-web-browser-in-china-may-put-users-data-at-risk-report/article29425182/ (accessed February 3, 2017).

Hachigian, Nina. 2001. China's Cyber-Strategy. *Foreign Affairs*, March/April. https://www.foreignaffairs.com/articles/asia/2001-03-01/chinas-cyber-strategy (accessed February 3, 2017).

Hayhtio, Tapio, and Rinne Jarmo. 2008. *Net Working/Networking: Citizen Initiated Internet Politics*. Helsinki: University of Tampere.

Herold, David Kurt, and Peter Marolt. 2011. *Online Society in China: Creating, Celebrating, and Instrumentalising the Online Carnival*. New York: Routledge.

Holt, Richard. 2004. *Dialogue on the Internet: Language, Civic Identity, and Computer-Mediated Communication*. Westport, CT: Praeger.

Howard, Philip, and Muzammil Hussain. 2012. Egypt and Tunisia: The Role of Digital Media. In *Liberation Technology: Social Media and the Struggle for Democracy*, ed. Larry Diamon and Marc F. Plattner, 110–23. Baltimore, MD: Johns Hopkins University Press.

IBP. 2013. *Myanmar Internet and E-Commerce Investment and Business Guide—Regulations and Opportunities*. Washington, DC: International Business Publications.

Isaac, Mike. 2016a. Other Gains Don't Offset Twitter Miss on Revenue. *New York Times*, April 27, B1.

——. 2016b. WhatsApp Encryption Now Covers All Messages. *New York Times*, April 6, B3.

Joyce, Mary C. 2010. *Digital Activism Decoded: The New Mechanics of Change*. New York: International Debate Education Association.

Kuran, Timur. 1989. Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution. *Public Choice* 61:41–74.

——. 1991. Now Out of Never: The Element of Surprise in the East European Revolution of 1989. *World Politics* 44:7–48.

——. 1997. *Private Truths, Public Lies*. Cambridge, MA: Harvard University Press.

——. 1998. Ethnic Norms and Their Transformation Through Reputational Cascades. *Journal of Legal Studies* 27:623–59.

Kuran, Timur, and Cass Sunstein. 1999. Availability Cascades and Risk Regulation. *Stanford Law Review* 51:683–768.

Kyi, Aung San Suu. 2010. *Freedom from Fear and Other Writings*. London: Penguin.

Lemieux, Pierre. 2004. Following the Herd. *Regulation* 26:16–21.

Link, Perry. 2002. China: The Anaconda in the Chandelier. *New York Review of Books*, April 11. http://www.nybooks.com/articles/archives/2002/apr/11/china-the-anaconda-in-the-chandelier/ (accessed February 3, 2017).

Link, Perry, Richard P. Madsen, and Paul G. Pickowicz. 2013. *Restless China*. New York: Rowman & Littlefield.

Mackey, Robert. 2009. Updates on the Protests in Iran. *New York Times*, July 9. http://thelede.blogs.nytimes.com/2009/07/09/latest-updates-on-iran-election-protests/?_r=0 (accessed April 12, 2016).

Morozov, Evgeny. 2011. *The Net Delusion: How Not to Liberate the World*. London: Penguin.

Noelle-Neumann, Elisabeth. 1974. The Spiral of Silence: A Theory of Public Opinion. *Journal of Communication* 24:43–51.

OECD. 2012. *OECD Internet Economy Outlook 2012*. Paris: OECD.

Oloffson, Kristi. 2010. YouTube's 50 Best Videos. *Time*, March 29. http://content.time.com/time/specials/packages/article/0,28804,1974961_1974925_1974743,00.html (accessed February 3, 2017).

Picker, Randal. 1997. Simple Games in a Complex World: A Generative Approach to the Adoption of Norms. *University of Chicago Law Review* 64:1225–88.

Preston, Jennifer. 2011. Movement Began with Outrage and a Facebook Page that Gave It an Outlet. *New York Times*, February 5. http://www.nytimes.com/2011/02/06/world/middleeast/06face.html?_r=0 (accessed April 2, 2016).

Rawnsley, Gary D., and Ming-Yeh T. Rawnsley. 2015. *Routledge Handbook of Chinese Media*. New York: Routledge.

Ren, Fujun, and Jiequan Zhai. 2013. *Communication and Popularization of Science and Technology in China*. Heidelberg: Springer Science & Business Media.

Reporters Without Borders. 2014. Enemies of the Internet 2014: Entities at the Heart of Censorship and Surveillance. http://12mars.rsf.org/2014-en/enemies-of-the-internet-2014-entities-at-the-heart-of-censorship-and-surveillance/ (accessed May 17, 2016).

Roe, Mark J. 1996. A Chaos and Evolution in Law and Economics. *Harvard Law Review* 109: 641–68.

Schelling, Thomas C. 1978. *Micromotives and Macrobehavior*. New York: W.W. Norton & Company.

Shiller, Robert J. 1995. Conversation, Information, and Herd Behavior. *American Economic Review* 85:181–85.

Sisario, Ben. 2016. Music World Bands Together Against YouTube, Seeking Change to Law. *New York Times*, May 31. https://www.nytimes.com/2016/06/01/business/media/music-world-bands-together-against-youtube-seeking-change-to-law.html (accessed July 6, 2016).

So, Alvin Y. 2012. *Handbook of Contemporary China*. Singapore: World Scientific Publishing.

Sterelny, Kim, Richard Joyce, Brett Calcott, and Ben Fraser. 2013. *Cooperation and its Evolution*. Cambridge, MA: MIT Press.

Stewart, James B. 2016. Facebook Has 50 Minutes of Your Time Each Day. It Wants More. *New York Times*, May 5. https://www.nytimes.com/2016/05/06/business/facebook-bends-the-rules-of-audience-engagement-to-its-advantage.html (accessed May 17, 2016).

Sunstein, Cass R. 1996a. On the Expressive Function of Law. *University of Pennsylvania Law Review* 144:2021–53.

——. 1996b. Social Norms and Social Roles. *Columbia Law Review* 96:903–68.

Taylor, William C. 2014. *Military Responses to the Arab Uprisings and the Future of Civil-Military Relations in the Middle East: Analysis from Egypt, Tunisia, Libya, and Syria*. New York: Palgrave Macmillan.

Tkacheva, Olesya, Lowell H. Schwartz, Martin C. Libicki, Julie E. Taylor, Jeffrey Martini, and Caroline Baxter. 2013. *Internet Freedom and Political Space*. Santa Monica, CA: Rand Corporation.

White, Stephen, Richard Sakwa, and Henry E. Hale. 2014. *Developments in Russian Politics 8*. New York: Palgrave Macmillan.

Wines, Michael. 2009. A Dirty Pun Tweaks China's Online Censors. *New York Times*, March 11. http://www.nytimes.com/2009/03/12/world/asia/12beast.html?_r=0 (accessed May 17, 2016).

Wingfield, Nick. 2016. Microsoft Buys LinkedIn for $26.2 Billion, Reasserting Its Muscle. *New York Times*, June 14. https://www.nytimes.com/2016/06/14/business/dealbook/microsoft-to-buy-linkedin-for-26-2-billion.html (accessed April 9, 2016).

World Bank. 2016. *World Development Report 2016: Digital Dividends*. Washington, DC: World Bank.

Xu, Xiaowen. 2011. Internet Facilitated Civic Engagement in China's Context: A Case Study of the Internet Event of Wenzhou High-Speed Train Accident. Master's thesis, Columbia University, New York.

Youn, Monica. 2013. The Chilling Effect and the Problem of Private Action. *Vanderbilt Law Review* 66:1473–1539.

Youngs, Gillian. 2013. *Digital World: Connectivity, Creativity and Rights*. New York: Routledge.

Yu, Haiqing. 2009. *Media and Cultural Transformation in China*. Oxford: Taylor & Francis.