REGULATING CYBERSECURITY

Nathan Alexander Sales,
George Mason University School of Law

Northwestern University Law Review, Forthcoming

George Mason University Law and Economics Research Paper Series

12-35
REGULATING CYBERSECURITY


Nathan Alexander Sales
George Mason University School of Law

ABSTRACT

The conventional wisdom is that this country’s privately owned critical infrastructure – banks, telecommunications networks, the power grid, and so on – is vulnerable to catastrophic cyberattacks. The existing academic literature does not adequately grapple with this problem, however, because it conceives of cybersecurity in unduly narrow terms: Most scholars understand cyberattacks as a problem of either the criminal law or the law of armed conflict. Cybersecurity scholarship need not run in such established channels. This article argues that, rather than thinking of private companies merely as potential victims of cyber crimes or as possible targets in cyber conflicts, we should think of them in administrative law terms. Firms that operate critical infrastructure tend to underinvest in cyberdefense because of problems associated with negative externalities, positive externalities, free riding, and public goods – the same sorts of challenges the modern administrative state faces in fields like environmental law, antitrust law, products liability law, and public health law. These disciplines do not just yield a richer analytical framework for thinking about cybersecurity, they also expand the range of possible responses. Understanding the problem in regulatory terms allows us to adapt various regulatory solutions for the cybersecurity context, such as monitoring and surveillance to detect malicious code, hardening vulnerable targets, and building resilient and recoverable systems. In short, an entirely new conceptual approach to cybersecurity is needed.
INTRODUCTION

The Red Army had been gone for years, but it still had the power to inspire controversy – and destruction.¹ In April 2007, the government of Estonia announced plans to relocate a contentious Soviet-era memorial in its capital city of Tallinn. Known as the Bronze Soldier, the Soviets erected the statue in 1947 to commemorate their sacrifices in the Great Patriotic War and their “liberation” of their Baltic neighbors. The local population, which suffered under the Bolshevik boot for decades, understandably saw the monument in a rather different light. Not long after the announcement, the tiny nation was hit with a massive cyberattack. Estonia, sometimes nicknamed “E-stonia,” is one of the most networked countries in the world – its citizens bank, vote, and pay taxes online² – and it ground to a halt for weeks. The country’s largest bank was paralyzed. Credit card companies took their systems down to keep them from

¹ The events in this paragraph are described in JOEL BRENNER, AMERICA THE VULNERABLE 127-30 (2011); RICHARD A. CLARKE & ROBERT K. KNACE, CYBERWAR 11-16 (2010); and Ian Traynor, Russia Accused of Unleashing Cyberwar to Disable Estonia, GUARDIAN (U.K.), May 16, 2007.

being attacked. The telephone network went dark. Newspapers and television stations were knocked offline. Who was responsible for launching what has come to be known as Web War I?3 The smart money is on Russia, though no one can say for sure.

It could happen here. Government officials like Richard Clarke, the former White House cybersecurity czar, have been warning of an “electronic Pearl Harbor” for years.4 Others lament the “gaping vulnerabilities”5 in America’s cyberdefenses and speculate that the economic effect of a major assault could be “an order of magnitude” greater than the September 11, 2001 terrorist attacks.6 Academic commentators generally agree. Some see the danger as “monumental”7 and the country’s “most pervasive and pernicious threat.”8 Others predict that America’s failure to secure its cyber assets could “take down the nation’s entire security and economic infrastructure”9 and “bring this country to its knees.”10 It has even been suggested that “[t]he very future of the Republic” depends on “protect[ing] ourselves from enemies armed with cyber weapons.”11 There are some naysayers,12 but the consensus that we stand on the brink of a cyber calamity is both broad and deep.

A large scale cyberattack on this country, as in Estonia, likely would target privately held critical infrastructure – banks, telecommunications carriers, power companies, and other firms whose compromise would cause widespread harm.13 Indeed, America’s critical infrastructure,
approximately 85 percent of which is owned by private firms, already faces constant intrusions. Yet the private sector’s defenses are widely regarded as inadequate. Companies are essentially on their own when it comes to protecting their computer systems, with the government neither imposing security requirements nor bearing a share of the resulting costs.16 According to Bruce Smith, the United States follows a “bifurcated approach to network security” that “relie[s] predominantly on private investment in prevention and public investment in prosecution.”17 Christopher Coyne and Peter Leeson likewise stress that our defensive strategy “is simply the sum of dispersed decisions of individual users and businesses.”18 Regular firms that operate in competitive markets (such as online retailers) may be more likely to effectively protect their systems against ordinary intruders (such as recreational hackers). But strategically significant firms in uncompetitive markets (such as power companies and other public utilities) seem especially unlikely to maintain defenses capable of protecting their systems against skilled and determined adversaries (such as foreign intelligence services).

The poor state of America’s cyberdefenses is partly due to the fact that the analytical framework used to understand the problem is incomplete. The law and policy of cybersecurity are undertheorized. Virtually all legal scholarship approaches cybersecurity from the standpoint of the criminal law or the law of armed conflict.19 Given these analytical commitments, it is

---


inevitable that academics and lawmakers will tend to favor law enforcement and military solutions to cybersecurity problems. These are important perspectives, but cybersecurity scholarship need not run in such narrow channels. An entirely new approach is needed. Rather than conceiving of private firms merely as possible victims of cyber crimes, or as potential targets in cyber conflicts, we should think of them in regulatory terms. Many companies that operate critical infrastructure tend to underinvest in cyberdefense because of problems associated with negative externalities, positive externalities, free riding, and public goods – the same sorts of challenges the modern administrative state encounters in a variety of other contexts, such as environmental law, antitrust law, products liability law, and public health law.

For instance, cybersecurity resembles environmental law in that both fields are primarily concerned with negative externalities. Just as firms tend to underinvest in pollution controls because some costs of their emissions are borne by those who are downwind, they also tend to underinvest in cyberdefenses because some costs of intrusions are externalized onto others. An attack on a power company will not just harm the intended target; it will also harm the target’s customers and those with whom the power company has no relationship. Because firms do not bear the full costs of their vulnerabilities, they have weaker incentives to secure their systems. Cybersecurity also resembles an antitrust problem. Antitrust law seeks to prevent anticompetitive behavior, and it traditionally has been skeptical of coordination among competitors. Some inter-firm cooperation could improve cybersecurity – sharing information about vulnerabilities and threats, for example, or developing industry wide security standards. Yet firms are reluctant to do so because they fear liability under the antitrust laws. Next, cybersecurity raises tort problems. Products liability law uses the threat of money damages to incentivize firms to take reasonable precautions when designing their products. This threat is almost entirely absent in the cybersecurity context, as companies face little risk of liability to those who are harmed by attacks on their systems or products. The incentive to patch vulnerabilities thus is weaker than it would be under a meaningful liability regime. Finally, cybersecurity resembles public health. A key goal of public health law is prevention – keeping those who have contracted a disease from spreading it to the healthy, a form of negative externality. Public health law uses vaccinations to promote immunity, biosurveillance to detect outbreaks, and quarantines to contain infectious diseases. Cybersecurity has similar goals – ensuring that critical systems are immune to malware, quickly detecting outbreaks of malicious code, and preventing contaminated computers from infecting clean systems.

Approaching cybersecurity from a regulatory vantage point does not just yield a richer analytical framework. It also expands the range of possible responses. The available solutions are determined by the threshold choice of analytical models; the more frameworks, the longer the menu of policy choices. If cyber insecurity resembles problems that arise in other regulatory contexts, then perhaps some of their solutions can be adapted here. Taken together, these
disciplines suggest four groups of responses: (1) monitoring and surveillance to detect malicious code; (2) hardening vulnerable targets and enabling them to defeat intrusions; (3) building resilient systems that can function during attacks and recover quickly; and (4) responding in the aftermath of attacks.

In particular, public health law’s distributed biosurveillance network might be used as a model for detecting cyber intrusions. Rather than empowering a single regulator to monitor internet traffic for outbreaks of malicious code, private firms could be tasked with reporting information about the vulnerabilities and threats they experience in much the same way hospitals report to public health authorities. To incentivize participation in this distributed surveillance network, firms might be offered various subsidies (on the theory that cybersecurity data is a public good that the market will tend to underproduce) and liability protections (such as an exemption from the antitrust laws). As for hardening targets, we might adopt industrywide security standards for companies that operate critical infrastructure. These protocols need not be issued in the form of traditional regulatory commands. Instead, as is sometimes the case in environmental law and other fields, the private sector should actively participate in formulating the standards. Tort law has a role to play as well: Threats of liability and offers of immunity might be used to incentivize firms to implement the protocols. Next, because it is inevitable that some cyberattacks will succeed, it is important that critical systems are able to survive and recover. Public health law offers several strategies for improving resilience. Systems that are infected with malware might be temporarily isolated to prevent them from spreading the contagion. Or firms might build excess capacity into their systems that can be deployed in emergencies – the equivalent of stockpiling vaccines and medicines. Finally, retaliation is thoroughly addressed in the existing criminal law and armed conflict literatures, but there is one response that deserves a brief mention here: “hackbacks,” in which a victim counterattacks the attacker. Because the counterattack might fall on a third party whose system unwittingly is being used by the assailant, hackbacks can incentivize firms to prevent their systems from being so commandeered. Hackbacks also might weaken attackers’ incentives. If assailants know that counterattacks can render their intrusions ineffective, they are less likely to commit them in the first place.

This article proceeds in three parts. Part I considers whether private companies are investing socially optimal amounts in cyberdefenses. Part II describes four regulatory frameworks – environmental law, antitrust law, products liability law, and public health law – and explains their relevance to cybersecurity. Part III surveys solutions used by these regulatory disciplines and considers how to adapt them for the cybersecurity context.

Several preliminary observations are needed. First, I use the terms “cyberattack” and “cyber intrusion” interchangeably to denote any effort by an unauthorized user to affect the data on, or to take control of, a computer system. As used here, the terms include all of the following: “viruses” (a piece of code that “infects a software program and then ensures that the infected program reproduces the virus”21), “worms” (“a standalone program that replicates itself”22),

21 O’Neill, supra note 19, at 246; see also Katyal, Criminal Law, supra note 10, at 1023; Sklerov, supra note 19, at 14-15.

22 Katyal, Criminal Law, supra note 10, at 1024; see also Sklerov, supra note 19, at 15. Viruses and worms are similar. A principal difference is that viruses require human action to propagate – such as clicking on a link or
“logic bombs” (malware that “tells a computer to execute a set of instructions at a certain time or under certain specified conditions”\textsuperscript{23}), and “distributed denial of service” (“DDOS”) attacks (in which a “master” computer conscripts “zombies” and orders them to disable a victim by flooding it with traffic\textsuperscript{24}). Second, this article emphatically is not a paean to traditional, command and control regulation. The conventional wisdom is to avoid cybersecurity regulation,\textsuperscript{25} in part because of doubts about the government’s ability to manage such a dynamic field. But as I hope to show in the following pages, cybersecurity need not, and in many cases should not, be pursued with heavy handed regulatory tools. It is possible to promote better cyberdefenses with private law, such as by modifying traditional tort law doctrines. As for public law, regulation need not take the form of rigid legal commands backed by threat of sanction; regulatory objectives often can be attained by appealing to private firms’ self interest – by offering positive incentives to improve their defenses, not just by punishing them when they fall short. The private sector’s poor defenses may represent a market failure, as some have argued,\textsuperscript{26} but “[t]here’s not much point in replacing a predictable market failure with an equally predictable government failure.”\textsuperscript{27}

I. AN EFFICIENT LEVEL OF CYBERSECURITY

Our national security “depends heavily on privately owned critical infrastructure.”\textsuperscript{28} A cyberattack on these private assets could be devastating: With a few keystrokes, adversaries could hack into banks and corrupt customer data, take control of power plants and bring down the electricity grid, open the floodgates of dams, and take telecommunications networks offline.\textsuperscript{29} Or worse. Despite the magnitude of the threat, the conventional wisdom is that the private sector is not adequately protecting itself. This section surveys the available evidence on the extent of private cybersecurity expenditures. It then predicts that ordinary firms in competitive markets (like online retailers) are more likely to be investing socially optimal amounts in cyberdefense, while strategically significant firms in uncompetitive markets (like public utilities) are more likely to be underinvesting.

The optimal level of cyber intrusions is not zero, and the optimal level of cybersecurity expenditures is not infinity. From an economic perspective, the goal is to achieve an efficient

\begin{flushleft}
\textsuperscript{23} Katyal, Criminal Law, supra note 10, at 1025; see also O’Neill, supra note 19, at 248.
\end{flushleft}

\begin{flushleft}
\textsuperscript{24} STEWART A. BAKER, SKATING ON STILTS 202-03 (2010); BRENNER, supra note 1, at 38-39; CLARKE & KNAKE, supra note 1, at 13-14; Lin, supra note 19, at 70; Yamada et al., supra note 16, at 226.
\end{flushleft}

\begin{flushleft}
\textsuperscript{25} CLARKE & KNAKE, supra note 1, at 108-09.
\end{flushleft}

\begin{flushleft}
\textsuperscript{26} ABA, supra note 18, at 8; BAKER, supra note 24, at 237; CSIS, supra note 8, at 50, Katyal, Digital Architecture, supra note 15, at 2285.
\end{flushleft}

\begin{flushleft}
\textsuperscript{27} BAKER, supra note 24, at 237; see also Coyne & Leeson, supra note 18, at 490; Powell, supra note 14, at 507.
\end{flushleft}

\begin{flushleft}
\textsuperscript{28} BRENNER, supra note 1, at 223; cf. ABA, supra note 18, at 8 (“[P]rivate sector security is often governmental security.”).
\end{flushleft}

\begin{flushleft}
\textsuperscript{29} Stewart Baker, Denial of Service, FOREIGN POL’Y 2 (Sept. 30, 2011); BRENNER, supra note 1, at 137-54; CLARKE & KNAKE, supra note 1, at 64-68;
\end{flushleft}
level of attacks, not to prevent all attacks.\textsuperscript{30} Suppose that the expected cost to society of a given cyberattack – its cost discounted by the probability that it will occur – is $5 billion. It would be efficient for society to invest up to $5 billion in countermeasures to prevent the attack. If the necessary countermeasures cost more than $5 billion, the cost of preventing the attack would exceed the resulting security gains. In short, it is worthwhile to invest in cyberdefenses whose marginal costs are less than the marginal benefits of preventing the attacks.\textsuperscript{31} Relatedly, some intrusions are more problematic than others. Cybersecurity is a form of risk management, where risk is a function of three variables: vulnerabilities, threats, and consequences. A company with easily hacked systems, that faces a high probability of attacks from sophisticated foreign intelligence services, and whose compromise would cause severe social harm, raises very different problems than a company with relatively robust defenses, that is unlikely to face skilled intruders, and whose compromise would have few consequences for society.

Are individual firms, and society as a whole, investing the right amount in cyberdefense? Most observers believe that firms are underinvesting – and are missing the mark by a wide margin. Richard Clarke proclaims the private sector response an “unmitigated failure,”\textsuperscript{32} and scholars generally agree.\textsuperscript{33} Very little empirical data is available, but the consensus view has at least some anecdotal support. Studies conducted by McAfee (a computer security firm) in 2010 and 2011 revealed low levels of investment in cyberdefense. The studies found that many firms regard cybersecurity as little more than “a last box they have to check,”\textsuperscript{34} and that they neglect network security because they find it too expensive.\textsuperscript{35} In particular, McAfee found that companies often have weak authentication requirements\textsuperscript{36} – tools that can verify that the person who is accessing a system is who he says he is, and is authorized to access the system. Even fewer have systems that can monitor network activity and identify anomalies.\textsuperscript{37} Other studies

\textsuperscript{30} Coyne & Leeson, \textit{supra} note 18, at 477-78.
\textsuperscript{31} Coyne & Leeson, \textit{supra} note 18, at 478.
\textsuperscript{32} Clarke & Knake, \textit{supra} note 1, at 104.
\textsuperscript{33} ABA, \textit{supra} note 18, at 8; Banks & Parker, \textit{supra} note 7, at 9; Katyal, \textit{Criminal Law}, \textit{supra} note 10, at 1019; Bruce K. Kobayashi, \textit{Private Versus Social Incentives in Cybersecurity, in Grady & Parisi}, \textit{supra} note 19, at 14; Sarnikar & Johnsen, \textit{supra} note 19, at 3, 16; Bruce Schneier, \textit{Computer Security: It’s the Economics, Stupid} 1 (May 16, 2002). \textit{But see} Coldebella & White, \textit{supra} note 14, at 240; Smith, \textit{supra} note 17, at 173 n.12. Some scholars argue that companies are providing a suboptimally high level of cybersecurity. Benjamin Powell reports that a 2000 study found that firms would invest in cyberdefenses if they were expected to produce a 20 percent return on investment, which was considerably lower than the 30 percent ROI typically required for information technology investments. Powell, \textit{supra} note 14, at 504. What mechanism could account for a tendency to overinvest? A firm’s IT department has incentives to overstater the vulnerabilities the company faces, as cybersecurity fears translate into a larger share of the company’s budget; for outside security vendors, such fears mean brisier business. Ross Anderson, \textit{Unsettling Parallels Between Security and the Environment} 2 (May 16, 2002); Bambauer, \textit{supra} note 12, at 604-06; Calkins, \textit{supra} note 19, at 198-99.
\textsuperscript{34} McAfee 2011, \textit{supra} note 15, at 1.
\textsuperscript{35} McAfee, \textit{In the Crossfire} at 14 (2010).
\textsuperscript{36} McAfee 2011, \textit{supra} note 15, at 14.
\textsuperscript{37} McAfee 2011, \textit{supra} note 15, at 15. It would be a mistake to read too much into these findings. The study’s methodology was to survey business executives in about a dozen countries, McAfee 2010, \textit{supra} note 35, at 1, 41; McAfee 2011, \textit{supra} note 15, at 3, and it “was not designed to be a statistically valid opinion poll with sampling and
reveal that some companies’ defenses are so poor they don’t even know when they have suffered an attack. Verizon reported that “fully 75 percent of the intrusions they investigated were discovered by people other than the victims and 66 percent of victims did not even know an intrusion occurred on the system.” 38 Finally, a 2011 study by the Ponemon Institute found that “73 percent of companies surveyed had been hacked, but 88 percent of them spent more money on coffee than on securing their Web applications.” 39

Are these levels of investment efficient? Whether a particular firm is making socially optimal investments in cybersecurity – and the related issue of who should pay for that company’s cyberdefenses – is a function of two intersecting questions. First, what is the defending firm? Is it a regular company in a competitive market, an operator of critical infrastructure in an uncompetitive market, or something in between? Second, who is the anticipated attacker? Is it a recreational hacker, a foreign intelligence service, or someone in between? The range of possibilities can be depicted in a simple graph:

---

38 Rattray et al., supra note 8, at 155; see also Jensen, Cyber Warfare, supra note 15, at 1536.
39 BRENNER, supra note 1, at 239.
The x axis depicts the firms that might be subject to a cyberattack. They are arranged from left to right in order of increasing strategic significance. (A strategically significant company is one whose compromise would result in substantial social harms.) On the far left are relatively insignificant firms in competitive markets – i.e., markets in which many companies offer the same good or service, and where disappointed consumers therefore may defect from one to another. An example would be online retailers, such as Amazon.com. To the right are financial institutions. These firms rate high on the strategic significance scale; a former CIA director predicted that an attack on a single bank “would have an order-of-magnitude greater impact on the global economy” than 9/11.40 Banks operate in fairly competitive markets, as consumers can easily move their accounts from one to another. Another step to the right are ISPs and telecommunications carriers. They, too, are strategically significant. When Russia crippled Georgia’s communications systems during their 2008 war, citizens “could not connect to any outside news or information sources and could not send e-mail out of the country.”41 These markets are less competitive; consumers typically have only a handful of internet providers or telephone companies to choose from. At the far right are power companies and other public utilities. These firms rate high on the strategic significance scale. A cyberattack on the power grid would be truly catastrophic. The industrial control, or SCADA,42 systems used by power plants and other utilities are increasingly connected to the internet.43 Hackers could exploit this connectivity to disrupt power generation and leave tens of millions of people in the dark for months44; they could even destroy key system components like turbines.45 (In 2009, the Stuxnet worm – “the most sophisticated cyberweapon ever deployed”,46 – caused similar physical damage to Iran’s nuclear program.47) Utility markets are uncompetitive. Municipalities typically have only one power company or natural gas supplier, and there is no meaningful prospect that disappointed consumers will switch to a competitor.

40 Quoted in David E. Sanger et al., U.S. Steps Up Effort on Digital Defenses, N.Y. TIMES, Apr. 27, 2009, at A1; see also Sarnikar & Johnsen, supra note 19, at 1; Sklerov, supra note 19, at 19-20.

41 CLARKE & KNAKE, supra note 1, at 19; see also BRENNER, supra note 1, at 39-40; Jensen, Cyber Warfare, supra note 15, at 1540.

42 The acronym stands for “supervisory control and data acquisition.” CLARKE & KNAKE, supra note 1, at 98; CSIS, supra note 8, at 54; Randal C. Picker, Cybersecurity: Of Heterogeneity and Autarky, in Grady & Parisi, supra note 19, at 126.

43 BRENNER, supra note 1, at 97; Steven R. Chabinsky, Cybersecurity Strategy, 4 J. Nat’l Sec. L & Pol’y 27, 28 n.1 (2010); Condon, supra note 19, at 407; Coyne & Leeson, supra note 18, at 474; CSIS, supra note 8, at 54; Sklerov, supra note 19, at 18.


45 BRENNER, supra note 1, at 110; CLARKE & KNAKE, supra note 1, at 100, 107; ECONOMIST, supra note 3, at 4; Gable, supra note 2, at 59-60.

46 William J. Broad et al., Israeli Test on Worm Called Crucial in Iran Nuclear Delay, N.Y. TIMES, Jan. 15, 2011; see also BRENNER, supra note 1, at 102; Ellen Nakashima, Homeland Security Tries to Shore up Nation’s Cyber Defenses, WASH POST., Oct. 1, 2011; Kim Zetter, How Digital Detectives Deciphered Stuxnet, the Most Menacing Malware in History, WIRED, Jul. 11, 2011.

47 Bambauer, supra note 12, at 585-86; BRENNER, supra note 1, at 103; John Markoff, A Silent Attack, but not a Subtle One, N.Y. TIMES, Sept. 27, 2010.
The y axis depicts the assailants that might commit a cyberattack. They are arranged from bottom to top in order of increasing sophistication. (A sophisticated attacker is one who is capable of compromising the most secure systems; unsophisticated attackers are only capable of compromising relatively unsecured systems.) At the bottom are recreational hackers – the stereotypical teenagers out for “a digital joy ride.”48 One step above are “hacktivists.” Hacktivists are relatively skilled hackers who use cyber intrusions to advance a political agenda; they typically do not group themselves into formal organizations.49 An example is “Anonymous,” a group that launched DDOS attacks on financial institutions that stopped allowing customers to send money to WikiLeaks, an anti-secrecy group that had published a number of classified documents.50 Next are organized crime syndicates, such as those operating out of Russia.51 They, too, are fairly sophisticated; they engage in cyber intrusions primarily for financial gain; and by definition they are structured organizations.52 (International terrorists might be placed here as well, though they have shown little enthusiasm or aptitude for cyberattacks thus far.53 On the other hand, al Qaeda reportedly established an “academy of cyberterrorism” in Afghanistan,54 and computers taken from members contained information about SCADA systems in the United States.55) At the top are foreign governments’ militaries and intelligence services. These are the most sophisticated adversaries of all and they are capable of breaking into even highly secure systems. Internet giant Google recently saw its Gmail service penetrated by Chinese spies who wanted to eavesdrop on the Dalai Lama.56 Similarly, RSA – a software firm that issues online security credentials for the Pentagon, defense contractors, and other sensitive enterprises – was compromised so badly (probably by China) that it had to offer new credentials to all its customers.57

The curve roughly predicts the combinations of victims and attackers that are likely to occur. Quadrant (4) involves high frequency, low severity attacks. Retailers and other relatively insignificant firms can expect to be targeted fairly often by relatively unsophisticated recreational hackers and by more sophisticated hacktivists. Quadrant (2) involves attacks that are low frequency and high severity. More strategically significant firms like ISPs and public utilities will face attacks from sophisticated militaries and intelligence services, and perhaps from organized crime syndicates. These attacks will only occur rarely, but they are likely to be

48 Dunlap, supra note 12, at 358.
49 Byron Acohido, Cyberattacks Likely to Escalate this Year, USA TODAY, Jan. 10, 2012.
50 Somini Sengupta, 16 People Arrested in Wave of Attacks on Web Sites, N.Y. TIMES, Jul. 20, 2011.
52 BRENNER, supra note 1, at 7, 25.
53 Condron, supra note 19, at 405; Dunlap, supra note 12, at 359-60.
54 Joel P. Trachtman, Global Cyberterrorism, Jurisdiction, and International Organization, in Grady & Parisi, supra note 19, at 260-61.
55 BRENNER, supra note 1, at 106.
56 BAKER, supra note 24, at 208-13; BRENNER, supra note 1, at 46-47; Ellen Nakashima, Google to Enlist NSA to Help It Ward off Cyberattacks, WASH POST., Feb. 4, 2010; Rosenzweig, supra note 14, at 6.
devastating. In quadrant (3), recreational hackers and hacktivists might launch attacks against utilities and similarly significant enterprises, but these targets are probably less attractive to them than they are to foreign militaries or intelligence services.\footnote{Zetter, supra note 46 (“[C]ontrol systems aren’t a traditional hacker target, because there’s no obvious financial gain in hacking them . . . .”).} In quadrant (1), foreign governments are unlikely to target insignificant firms like retailers, because they gain little by compromising them, though organized crime may do so.

We are now in a position to make predictions about various companies’ cybersecurity expenditures. The closer we are on the curve to the lower left corner, the higher the probability that the firm is investing a socially optimal amount in cyberdefense. This is so in part because the expected social cost of an attack on an ordinary company is fairly low. Society will not grind to a halt if Amazon.com is knocked offline; bookworms might experience minor annoyance but they will still be able to buy a copy of \textit{Macbeth} from Barnes & Noble. In addition, these companies are unlikely to face attacks by skilled and determined foreign governments, so it is not necessary for them to spend huge sums of money on the very best and most impregnable defenses. The efficient level of cybersecurity investment for them thus is fairly low. Importantly, market forces may provide these firms with meaningful incentives to protect their systems against cyberattacks. Retailers, banks, and similar companies operate in competitive markets. The risk of customer exit provides them with strong incentives to cater to customer demand. If consumers want the companies with which they do business to provide better security against cyberattacks – the jury is out on that question, incidentally\footnote{Compare \textsc{Brenner, supra} note 1, at 225, 226; \textit{and} Paul M. Schwartz \& Edward J. Janger, \textit{Notification of Data Security Breaches}, 105 \textsc{Mich. L. Rev.} 913, 946-47 (2007); \textit{with} Dunlap, \textit{supra} note 12, at 361; \textit{and} Doug Lichtman \& Eric P. Posner, \textit{Holding Internet Service Providers Accountable}, \textit{in} Grady \& Parisi, \textit{supra} note 19, at 256.} – they will have good reason do so. (Note that current liability rules both diminish and augment these incentives. The federal wiretap act makes it a crime to intercept electronic communications, and some ISPs fear that this prohibition prevents them from filtering botnet traffic or other malware; the threat of liability undermines their incentives to improve the security of their systems.\footnote{See infra notes 194 to 201 and accompanying text.} By contrast, the Gramm-Leach-Bliley Act requires banks, on pain of significant money damages, to protect customer data against unauthorized access; the threat of liability amplifies their incentives to improve the security of their systems.\footnote{See \textit{infra} notes 202 to 210 and accompanying text.})

The closer we are on the curve to the upper right corner, the lower the probability that the firm is adequately investing in cybersecurity.\footnote{\textsc{Sarnikar \& Johnsen, supra} note 19, at 17, 22-23.} Quadrant (2) – low frequency, high severity – is the opposite of quadrant (4). First, the expected social cost of a cyberattack is monumental. The consequences of an attack on, say, the power grid, would reverberate throughout the economy, causing harm to the utility and its customers and also to third parties with which the company has no contractual relationships. Because the expected cost of an attack on these firms is so high, it is efficient to invest greater sums in securing them against intruders. In addition, the modest (and low cost) defenses that are usually capable of thwarting recreational hackers will do nothing to prevent intrusions by foreign governments; more expensive countermeasures are
needed to protect against these exceptionally sophisticated adversaries. The socially optimal level of cybersecurity investment for these firms thus is fairly high. Second, power companies and other utilities are not subject to market forces that might incentivize them to improve their cyberdefenses. Utilities face little if any competition; a given customer typically will be served by only one power company. Customer exit is essentially impossible, and the utility therefore has weaker incentives to supply what its customers are demanding. (This absence of favorable market forces may help explain why public utilities often fail to implement even relatively costless security measures. Many electric companies use vendor default passwords to protect their SCADA systems, and a recent study found that they take an average of 331 days to implement security patches for these systems. Perhaps not coincidentally, hackers – most likely Chinese and Russian spies – have been able to insert logic bombs into the power grid.)

If this analysis is correct, then strategically significant firms in uncompetitive markets are less likely to adequately invest in cybersecurity than ordinary firms in competitive markets. The question then becomes who should be responsible for securing these most sensitive companies against the most dangerous adversaries. Economists often argue that risk should be allocated to the low cost avoider. If the government can reduce a vulnerability more efficiently than a firm, it should pay; if the firm can reduce the vulnerability more efficiently, it should pay. There is no single low cost avoider in this context. Defending critical infrastructure against sophisticated cyberattackers is a task that features dueling comparative advantages. Private firms typically know more than outsiders, including the government, about the architecture of their systems, so they often are in a better position to know about weaknesses that intruders might exploit. The private sector thus has a comparative advantage at identifying cyber vulnerabilities. On the other hand, the government’s highly skilled intelligence agencies typically know more than the private sector about malware used by foreign governments and how to defeat it. The government thus has a comparative advantage at detecting sophisticated attacks and developing countermeasures. This suggests that responsibility for defending the most sensitive systems against the most sophisticated adversaries should be shared.

What might such a partnership look like? All private firms might be asked to provide a baseline level of cybersecurity – modestly effective (and modestly expensive) defenses that are capable of thwarting intrusions by adversaries of low to medium sophistication. The government then would assume responsibility for defending public utilities and other sensitive enterprises.

---

63 Availability bias is another reason why firms might tend to underinvest in cyberdefense. The United States has not experienced a major cyber incident that has captured the public’s imagination, so firms might irrationally discount the probability that they will suffer a catastrophic attack. John Grant, Will There Be Cybersecurity Legislation?, 4 J. NAT’L SEC. L & POL’Y 103, 111 (2010); McAfee 2010, supra note 35, at 14.

64 McAfee 2011, supra note 15, at 8.

65 BRENNER, supra note 1, at 98.


67 Katyal, Criminal Law, supra note 10, at 1095-96; LAWRENCE LESSIG, CODE 2.0 at 169-17 (2006).

68 See infra notes 263 to 265 and accompanying text.

69 See infra notes 266 to 268 and accompanying text.
against catastrophic attacks by foreign militaries and other highly sophisticated adversaries.\textsuperscript{70} This arrangement – basic security provided by firms, supplemental security provided by the government – is in a sense the opposite of what we see in realspace criminal law. In realspace, the government offers all citizens a baseline level of protection against criminals (in the form of police officers, prosecutors, and courts). Individuals may supplement these protections at their own expense, such as by installing alarm systems in their homes or hiring private security guards.\textsuperscript{71} This arrangement also is consistent with our intuitions about the respective roles of government and the private sector. Consider another realspace analogy. In World War II, factories weren’t expected to install antiaircraft batteries to defend themselves against Luftwaffe bombers.\textsuperscript{72} Nor would we expect power plants to defend themselves against foreign governments’ cyberattacks. Protecting vital national assets from destruction by foreign militaries is a quintessential, perhaps the quintessential, government function.\textsuperscript{73}

The division of labor I am suggesting also seems sound from an economic standpoint. If a firm invested in extraordinarily expensive cyberdefenses capable of thwarting doomsday attacks by China’s intelligence service and Russia’s military, it would effectively be subsidizing the rest of the population. The company would capture some benefits of increased security, but a large portion of the benefits would be in the form of a positive externality conferred on others.\textsuperscript{74} In other words, the firm would be providing a public good (a good that is both non-rivalrous and non-excludable).\textsuperscript{75} Economic theory predicts that public goods will be underprovided on the market; a standard response is to subsidize them. So the government might provide a sensitive enterprise with a subsidy equal in value to its costs of defending against the most sophisticated cyberattackers.\textsuperscript{76} This subsidy could take many forms. The government could either pay for the firm’s defenses directly or reimburse it for its cybersecurity expenditures. Or the company could be offered various tax credits, deductions, and other benefits. Or it could be granted immunity from certain forms of legal liability. (In that case, the subsidy would not run from society as a whole, but from those who were injured by the firm’s otherwise unlawful conduct and whose entitlement to redress has been extinguished. This sort of subsidy is potentially regressive.) Or the government might provide the company with intelligence about the types of attacks it is likely to face. (This sort of subsidy appears to be occurring already. The NSA reportedly is providing malware signature files to Google and certain banks to help them detect sophisticated intrusions into their systems.\textsuperscript{77})

\textsuperscript{70} Rabkin & Rabkin, at 4; Trachtman, \textit{supra} note 54, at 272.
\textsuperscript{71} Rosenzweig, \textit{supra} note 14, at 20.
\textsuperscript{72} \textit{Clarke & Knaake}, \textit{supra} note 1, at 144; Rosenzweig, \textit{supra} note 14, at 5-26.
\textsuperscript{73} Brenner, \textit{supra} note 1, at 223; CSIS, \textit{supra} note 8, at 15; Katyal, \textit{Digital Architecture}, \textit{supra} note 15, at 2282.
\textsuperscript{74} Sarnikar & Johnsen, \textit{supra} note 19, at 17.
\textsuperscript{75} \textit{See infra} notes 129 to 135 and accompanying text.
\textsuperscript{76} Amitai Aviram, \textit{Network Responses to Network Threats}, in Grady & Parisi, \textit{supra} note 19, at 149, 156; Bambauer, \textit{supra} note 12, at 658; \textit{Clarke & Knaake}, \textit{supra} note 1, at 113-14; Rosenzweig, \textit{supra} note 14, at 10; Sarnikar & Johnsen, \textit{supra} note 19, at 22-23.
\textsuperscript{77} \textit{See infra} notes 267 to 268 and accompanying text.
II. CYBERSECURITY FRAMEWORKS, CONVENTIONAL AND UNCONVENTIONAL

The vast majority of academic commentary regards cybersecurity as a problem of the criminal law or the law of armed conflict.\(^{78}\) The problem is not that these conventional approaches are mistaken. The problem is that they are incomplete. Treating cybersecurity as a matter for law enforcement or the military brings certain challenges into sharper focus. But it tends to obscure others.

Cybersecurity is beset by externalities; it’s “externalities galore.”\(^{79}\) An externality is “an effect on the market the source of which is external to the market”\(^{80}\); it occurs when an actor’s conduct results in the imposition of a cost or benefit on a nonconsenting third party. Externalities can be either positive or negative. “Positive externalities occur whenever an activity generates benefits that the actor is unable to internalize,” such as through prices; “[n]egative externalities occur when one’s activity imposes costs on others” that likewise are not transmitted through prices.\(^{81}\) Economic theory predicts that the market will oversupply negative externalities, relative to socially optimal levels, “because the producer will internalize all the benefits of the activity but not all of the costs.”\(^{82}\) It also predicts that the market will undersupply positive externalities because third parties will free ride. Externalities thus represent a form of market failure.\(^{83}\) The standard government response to a negative externality is to discourage the responsible conduct (as with taxation or regulation); the standard response to a positive externality is to encourage the responsible conduct (as with a subsidy).\(^{84}\)

Cybersecurity can be understood in these terms. If a company suffers an intrusion, much of the harm will fall on remote third parties; the attack results in a negative externality.\(^{85}\) It can be extraordinarily difficult to internalize these costs. The class of persons affected by the intrusion is likely to be so large that it would be prohibitively expensive to use market exchanges to internalize the resulting externalities; the transaction costs are simply too great. Nor can tort law internalize the costs, as firms generally do not face liability for harms that result from cyberattacks on their systems or products. Because many companies do not bear these externalized costs, they ignore them when deciding how much to spend on cyberdefense. They therefore tend to underinvest relative to socially optimal levels. (This is true both of companies that produce computer products, such as software manufacturers, and companies that use them,\(^{86}\)

\(^{78}\) See sources cited supra note 19.

\(^{79}\) Picker, supra note 42, at 115.


\(^{81}\) Elkin-Koren & Salzberger, supra note 80, at 563.

\(^{82}\) Coyne & Leeson, supra note 18, at 479.

\(^{83}\) Coyne & Leeson, supra note 18, at 479; Timothy F. Malloy, Regulating by Incentives, 80 TEX. L. REV. 531 534-35 n.13 (2002).

\(^{84}\) Coyne & Leeson, supra note 18, at 479; Rosenzweig, supra note 14, at 10.

\(^{85}\) See infra notes 118 to 124 and accompanying text.
such as ISPs and electric companies.) Cyberattacks also involve positive externalities. A company that secures itself against intruders makes it harder for assailants to use its systems to attack others. Investments in cyberdefense thus effectively subsidize other firms. Because the investing company doesn’t capture the full benefit of its expenditures, it has weaker incentives to secure its systems. And because other companies are able to free ride on the investing firm’s expenditures, they have weaker incentives to adopt defenses of their own.

These externality and free rider problems are largely overlooked by the conventional approaches to cybersecurity, but they can be illuminated if we consult alternative regulatory frameworks – frameworks like environmental law, antitrust law, products liability law, and public health law. In short, a wider selection of analytical lenses is needed to fully comprehend cybersecurity challenges in all their complexity.

A. The Conventional Approaches: Law Enforcement and Armed Conflict

Scholars typically use two analytical frameworks to understand cyberattacks: criminal law and the law of armed conflict. Consider the former first. Broadly speaking, the criminal law seeks to protect members of society from unjustified acts of violence against their persons or property. The criminal law pursues this objective by imposing sanctions, such as incarceration, on those adjudged to have violated the law. These penalties, it is alternatively said, will either punish those who have transgressed society’s moral code (retribution), or dissuade the perpetrator or others from committing similar offenses in the future (specific or general deterrence), or isolate the dangerous perpetrator from society (incapacitation), or teach the misguided perpetrator the error of his ways (rehabilitation). Cyberattacks fit into this conceptual framework fairly comfortably. A person who hacks into another’s computer may have thereby violated any number of laws, such as the federal Computer Fraud and Abuse Act. Society regards this sort of conduct as sufficiently blameworthy that it proscribes it and subjects those who engage in it to criminal penalties of varying severity.

Scholars who approach cybersecurity from a law enforcement perspective focus on the “whodunit” questions. Who was responsible for launching this particular attack? Was it an individual hacker or a larger criminal enterprise? The law enforcement framework also emphasizes jurisdictional questions. Which court (or courts) properly may exercise subject matter jurisdiction over a given cyberattack? State courts, federal courts, or perhaps an international tribunal? Should jurisdiction be determined by the location of the target? By the location of the attacker? By the location in which the effects of the attack are felt? Should cyberattacks be subject to universal jurisdiction – the notion that a court may try certain crimes

---

86 See infra notes 126 to 135 and accompanying text.
87 See sources cited supra note 19.
89 Gable, supra note 2, at 99-117
regardless of where in the world they occurred? How might courts gain personal jurisdiction over those suspected of committing the attack, especially if they are overseas? Do existing extradition treaties cover the range of offenses that cybercriminals might commit? Should the United States negotiate new bilateral agreements with key international partners (such as our European allies), or with countries in which cyberattacks are likely to originate (such as China and Russia)? Or should there be a multilateral global convention on cybercrime, one that will facilitate extradition of suspects from their home countries to the states in which they will stand trial for their alleged crimes?

The law enforcement framework also emphasizes punishment and deterrence. Certain economic theories of criminal law posit that a person’s willingness to commit crimes is a function of the expected penalty for that activity – i.e., the sanction for particular offense discounted by the probability he will get caught. The greater the sanction, and the greater the likelihood of detection and punishment, the less likely a person will choose to commit that crime. The question then becomes what should be done to increase the deterrent effect of laws that proscribe various cyber intrusions? Should the penalties for violating these statutes be increased? To what level? Should society invest more resources in detecting cyber crime, thereby increasing the probability that perpetrators will be caught and punished? Or should lawmakers pursue “cost deterrence,” the objective of which is to increase the costs one must incur to perpetrate cybercrime?

The second conventional approach is to regard cyberattacks from the standpoint of the law of armed conflict (LOAC). The LOAC, also known as international humanitarian law (IHL), is a body of international law that regulates a state’s ability to use force in several ways. First, it sets forth the circumstances in which a state lawfully may engage in armed conflict – the \textit{jus ad bellum} regulations. For instance, the United Nations Charter forbids signatories “from the threat or use of force against the territorial integrity or political independence of any state” but also recognizes an inherent right to use force in self defense: “Nothing in the present Charter shall impair the inherent right of individual or collective self-defense if an armed attack occurs . . .” Second, the LOAC regulates what kinds of force may be used during an authorized armed conflict – the \textit{jus in bello} regulations. For instance, a state may not deliberately kill civilians or destroy civilian infrastructure (the \textit{distinction or discrimination} principle), may not inadvertently inflict harm on civilian populations and structures that is disproportionate to the importance of

\begin{itemize}
\item \textbf{92} ABA, \textit{supra} note 18, at 13; Gable, \textit{supra} note 2, at 65; Katyal, \textit{Criminal Law}, \textit{supra} note 10, at 1006, 1011, 1040; O’Neill, \textit{supra} note 19, at 265-68; K.A. Taipale, \textit{Cyber-Deterrence} 18 (Apr. 2010).
\item \textbf{94} Katyal, \textit{Criminal Law}, \textit{supra} note 10, at 1012; \textit{see also} O’Neill, \textit{supra} note 19, at 265-88.
\item \textbf{95} U.N. Charter Art. 2.
\item \textbf{96} U.N. Charter Art. 51.
\end{itemize}
the military objective (proportionality), and may not cause more harm to legitimate targets than is needed to achieve the military objective (necessity).\(^97\)

Cybersecurity is often described in LOAC terms. Scholars who see cybersecurity as an armed conflict problem focus on determining who was responsible for a particular attack.\(^98\) Was this attack launched by a state or an international terrorist organization, in which case the LOAC would probably permit some form of military retaliation? Or was it carried out by criminals, in which case the distinction principle likely would rule out a military response? If the attacker was in fact a state or terrorist group, which one? Was it China, or maybe Russia, or perhaps North Korea? Or was it al Qaeda, or al Qaeda in the Arabian Peninsula, or Hezbollah? Until the identity of the assailant it known, it will be unclear against whom to retaliate – or even whether retaliation is lawful at all.\(^99\)

Another set of important questions concerns how to characterize a cyber incident. Is a given intrusion an act of espionage or an attack? It can be quite difficult to answer that question, because the steps an intruder would take to steal information often are identical to the steps it would take to bring down a system. If the intrusion is properly understood as an attack, does it rise to the level of an “armed attack” that triggers the right of self defense?\(^100\) Should these questions be resolved with an “instrument based” test (a cyber intrusion counts as an armed attack when it causes harms that previously could have been caused only by a kinetic attack\(^101\))? Or a less demanding “effects” or “consequence based” test (a cyber intrusion counts as an armed attack when it has a sufficiently harmful effect on the targeted state?\(^102\))? Or an even less demanding “intent” test (a cyber intrusion counts as an armed attack whenever it evinces a hostile intent, regardless of whether it causes actual damage\(^103\)). The LOAC approach also emphasizes possible responses. When a nation suffers a cyberattack, is it limited to responding with a cyber intrusion of its own?\(^104\) Or may a victim retaliate by launching a kinetic attack?\(^105\) How severe must the cyberattack be before a kinetic response would be justified?


\(^98\) Graham, *supra* note 19, at 92; Lin, *supra* note 19, at 77.

\(^99\) Condron, *supra* note 19, at 414.


\(^101\) Graham, *supra* note 19, at 91; Sklerov, *supra* note 19, at 54.


\(^103\) WALTER GARY SHARP, SR., CYBERSPACE AND THE USE OF FORCE 129-31 (1999). Some scholars describe the intent test as a form of “strict liability.” See, e.g., Graham, *supra* note 19, at 91; Sklerov, *supra* note 19, at 55. This seems incorrect. A strict liability regime imposes liability solely on the basis of the social harm produced by the actor’s conduct, without reference to his mens rea. WAYNE R. LAFAVE, CRIMINAL LAW § 5.5, at 288-89 (5th ed. 2010). It would be more accurate to say that the intent test imposes liability solely on the basis of mens rea, without any requirement that the actor’s conduct result in social harm.

\(^104\) Condron, *supra* note 19, at 415-16; Graham, *supra* note 19, at 89-90.

\(^105\) Jensen, *Computer Attacks, supra* note 19, at 229-30.
Other problems arise from the fact that much of the world’s critical infrastructure is “dual use” – it serves a state’s civilian population but also is relied upon by the state’s political leadership and armed forces.\(^{106}\) (In the United States, civilian networks carry up to 98 percent of the federal government’s communications traffic, including 95 percent of defense-related traffic.\(^{107}\)) When, if ever, may a combatant direct a cyberattack at an adversary’s dual use infrastructure?\(^{108}\) Finally, the LOAC approach to cybersecurity focuses on deterrence. Given the differences between cyber conflicts and kinetic ones, how can a state dissuade its adversaries from committing cyberattacks? Key differences include the fact that it is difficult to determine who was responsible for a given intrusion, the possibility that a retaliatory cyber strike might end up harming innocent third parties more than the actual assailant, and the fact that different nations are more (or less) dependent on cyber infrastructure and therefore have more (or less) to lose from an exchange of cyber weapons.\(^{109}\)

A central problem for both the law enforcement and armed conflict approaches to cybersecurity is determining the identity of the assailant. Yet attribution is extraordinarily difficult; the challenges are “staggering”\(^{110}\) and “[n]o one has come close to solving” them.\(^{111}\) This is so because of the basic architecture of the internet. The internet’s TCP/IP protocol was designed to move packets of data as efficiently as possible; it is utterly unconcerned with who sent them.\(^{112}\) As such, it is fairly easy for attackers to obscure their true identities by routing their intrusions through a series of dispersed intermediary computers.\(^{113}\) These attribution difficulties can severely frustrate the law enforcement and armed conflict approaches to cybersecurity.

**B. Cybersecurity as an Environmental Law Problem**

A principal goal of environmental law is to regulate externalities. Various forms of environmental degradation can be described as negative externalities – i.e., spillover costs that are imposed on third parties and that are not transmitted through prices.\(^{114}\) A coal fired power plant imposes negative externalities on those who live downwind when, as a byproduct of


\(^{110}\) Jensen, *Computer Attacks, supra* note 19, at 234.


\(^{112}\) Bambauer, *supra* note 12, at 595-96; LESSIG, *supra* note 67, at 44.


\(^{114}\) *See supra* notes 79 to 84 and accompanying text.
productive activity, it emits pollutants that increase the incidence of asthma. Sometimes these
externalities are geographic; pollutants emitted by a factory in Ohio might affect residents of
New York.\textsuperscript{115} Sometimes they are temporal; carbon emissions today might affect the planet’s
climate for future generations.\textsuperscript{116} The critical point is that these costs are borne by people other
than those who are responsible for the pollution, and it is usually impossible to use market
transactions to internalize the costs onto the polluter. Many scholars therefore believe that
regulatory controls are necessary. For instance, Richard Lazarus cites a “need for government
regulation because of the spatial and temporal spillovers caused by unrestricted resource
exploitation.”\textsuperscript{117} These controls often take the form of strict limits on the regulated activity,
backed by the threat of civil damages or criminal sanctions, though less coercive forms of
regulation exist.

Cybersecurity can be understood in these terms. First, consider negative externalities.\textsuperscript{118}
A given firm – whether it is a company that produces computer products (such as a software
manufacturer) or a company that uses them (such as an ISP or electric company) – will not bear
the full costs of its cyber insecurities. (By “cyber insecurity,” I refer to a firm that suffers a
cyberattack after failing to implement defenses capable of defeating the attack.) Instead, some of
these costs are borne by third parties; they are partially externalized.\textsuperscript{119} Imagine a cyberattack
that disables a power plant. The intrusion would harm the utility as well as consumers who buy
electricity from it\textsuperscript{120} – hospitals, manufacturers, and others. The attack also would harm a
number of third parties who have no relationship with the power company – hospital patients,
downstream manufacturers in the supply chain, and so on. These “indirect effects of a
cyberattack are almost always more important to the attacker than the direct effects.”\textsuperscript{121} And it
would be prohibitively expensive to internalize them through market exchanges; the transaction
costs would be staggering, as it is extraordinarily difficult to identify the universe of third parties
affected by the intrusion.

The fact that many costs of cyberattack are externalized onto third parties is enormously
significant. Some commentators have argued that firms have strong “financial incentives to

\textsuperscript{115} See, e.g., Massachusetts v. EPA.


\textsuperscript{117} Lazarus, \textit{supra} note 116, at 1005-06.

\textsuperscript{118} Anderson, \textit{supra} note 33, at 1. One potential difference between pollution and cyber insecurity is that pollution is
a harmful byproduct of socially beneficial activity (such as manufacturing) whereas cyberattacks involve intentionnally malicious conduct. Rattray et al., \textit{supra} note 8, at 171. Yet cyber intrusions likewise may be seen as a
harmful byproduct of beneficial activity. A cyberattack on a computer is a byproduct of the computer being
connected to the internet. And connecting a computer to the internet is socially beneficial because it produces
network effects; by joining the network, the user increases its value to all users. \textit{Post}, \textit{supra} note 90, at 47-49.

\textsuperscript{119} ABA, \textit{supra} note 18, at 8; Anderson, \textit{supra} note 33, at 1; Jim Harper, \textit{Government-Run Cyber Security? No
Thanks} 1 (Mar. 13, 2009); Rosenzweig, \textit{supra} note 14, at 9-10; Schwartz & Janger, \textit{supra} note 59, at 928.

\textsuperscript{120} Aviram, \textit{supra} note 76, at 155; Lin, \textit{supra} note 19, at 68.

\textsuperscript{121} Lin, \textit{supra} note 19, at 68.
Those incentives are weaker than might be supposed. A firm that is deciding how much to invest in securing its systems will not account for the costs that an attack will impose on third parties. Firms tend to oversupply pollution, since they capture all the benefits of the associated productive activity but not all of the resulting costs. In a similar way, firms tend to oversupply cyber insecurity – or, to say the same thing, they tend to undersupply cyberdefense – because they internalize all of the benefits but only some of the costs. Many firms thus tend to invest less in cyberdefense than would be optimal from a societal standpoint.

The point can be illustrated with a simple hypothetical. Imagine a cyberattack that will result in $1 million in expected costs for the target firm and $10 million in expected costs for third parties. From a societal standpoint, it would be worthwhile to invest up to $11 million to prevent the attack – the sum of the expected harms to the firm and third party victims. But from the company’s standpoint, it would only be worthwhile to invest up to $1 million to prevent the attack. If the firm spent more than that, the cost to it of the precautions would exceed the benefit to it, and the firm would be conferring uncompensated benefits on third parties. The firm effectively would be providing a security subsidy. Thus, there is a gap between the welfare of the company and the welfare of society as a whole. Levels of cybersecurity investment that are efficient for particular firms may turn out to be inefficient for society at large.

Cybersecurity also can be understood as a positive externality. When a firm expends resources to defend itself against intruders, that investment makes other users’ systems marginally more secure as well. This is so because the defenses not only help prevent harm to the company’s system, they also help prevent the firm’s system from being used to inflict harm on others’ systems. If Pepsi’s network is well defended, it is less likely to be infected by a worm and thus less likely to transmit the malware to Coke. The effect is to decrease the overall incidence of infection on the internet, but the investing firm does not capture the full benefit. A classic positive externality. Cyberdefenses can differ from realspace defenses in this respect. If I install an alarm in my home, that might prevent burglars from breaking into my house, but it won’t necessarily decrease the overall incidence of burglary. The alarm might simply displace the burglar who would have targeted me onto my neighbor – a form of negative externality. By contrast, cyberdefenses can make my system more secure at the same time they increase the overall security of the internet.

---

122 Nojeim, supra note 14, at 134; see also Coldebell & White, supra note 14, at 236, 241; Dunlap, supra note 12, at 361; Yang & Hoffstadt, supra note 15, at 203.
123 ABA, supra note 18, at 8; Coyne & Leeson, supra note 18, at 479; Rosenzweig, supra note 14, at 9-10.
124 Coyne & Leeson, supra note 18, at 480.
125 Sarnikar & Johnsen, supra note 19, at ___.
126 Katyal, Criminal Law, supra note 10, at 1081-82; O’Neill, supra note 19, at 278; Rosenzweig, supra note 14, at 9; Sarnikar & Johnsen, supra note 19, at 15-16.
127 Neal Kumar Katyal, Community Self-Help, 1 J.L. ECON. & POL’Y 33, 46 (2005); Katyal, Criminal Law, supra note 10, at 1081; O’Neill, supra note 19, at 278.
128 But see Kobayashi, supra note 33, at 16; Rosenzweig, supra note 14, at 9.
Relatedly, some aspects of cybersecurity resemble public goods. A public good is both non-rivalrous (one person’s use of the good doesn’t reduce its availability for use by others) and non-excludable (the owner of the good can’t prevent particular persons from using it). A classic example of a public good is a large municipal park – open to all comers, and one person enjoying a crisp fall afternoon on a park bench (generally) doesn’t prevent anyone else from doing the same. Some scholars argue that cybersecurity information is a public good – e.g., information about the vulnerability of particular system, or the most effective way to counter a particular cyberthreat – that the market will tend to underproduce. There is also a sense in which defensive measures themselves are public goods. Like a municipal park, cyberdefenses can be non-rivalrous. When Pepsi expends resources to secure its computer network, that doesn’t decrease the amount of security available for Coke; it actually can increase security for third parties. Cyberdefenses also can be non-excludable. When Pepsi secures its system against conscription into a botnet, it isn’t possible to specify which third parties will enjoy the benefit of Pepsi’s immunity – Coke, but not Snapple. All such users are thereby protected from a DDOS attack launched from Pepsi’s system.

Understanding cybersecurity in terms of positive externalities and public goods can help explain why many firms underinvest in defense. It’s a free rider problem. A company that decides to better secure its computer systems thereby produces benefits that accrue to a number of third parties, and it is impossible to exclude them from receiving those benefits. The investing firm therefore has a weaker incentive to expend resources on cyberdefenses because such expenditures effectively subsidize other firms, including its competitors. Third parties likewise have weaker incentives to secure their own systems; they would prefer to free ride on a rival firm’s investment in, say, anti-spyware software than to purchase the product on their own. The overall effect is to weaken the incentive of all firms to invest in protecting their networks. “The individual undertaking the security precautions does not internalize all the benefits, and will seek to free-ride off the efforts taken by others”; as a result, “theory predicts that security will be undersupplied on the market.”

129 CSIS, supra note 8, at 50; Kobayashi, supra note 33, at 15; Powell, supra note 14, at 498.

130 Elkin-Koren & Salzberger, supra note 80, at 559; James Grimmelman, The Internet Is a Semicommons, 78 FORD. L. REV. 2799, 2806 (2010); Rosenzweig, supra note 14, at 8-9; see also Harold Demsetz, The Private Production of Public Goods, 13 J.L. & ECON. 293, 295 (1970) (distinguishing between non-rivalrous goods, which are properly characterized as public goods, and non-exclusive goods, which are properly characterized as “collective goods”).

131 Kobayashi, supra note 33, at 16; Rosenzweig, supra note 14, at 9. But see Amitai Aviram & Avishalom Tor, Overcoming Impediments to Information Sharing, 55 ALA. L. REV. 231, 234-35 (2004) (arguing that information can be a rivalrous good, insofar as sharing it can cause a firm to “los[e] a competitive edge over rivals that benefit from the information”).

132 Kobayashi, supra note 33, at 20-21; Trachtman, supra note 54, at 270.

133 Trachtman, supra note 54, at 270.

134 Aviram & Tor, supra note 131, at 238; CSIS, supra note 8, at 50; Elkin-Koren & Salzberger, supra note 80, at 559; Sarnikar & Johnsen, supra note 19, at 16; Trachtman, supra note 54, at 281. But see Powell, supra note 14, at 504-05.

135 Coyne & Leeson, supra note 18, at 480; see also Sarnikar & Johnsen, supra note 19, at 16.
Environmental law and the underlying economic principles it reflects thus provide an important framework through which we might better understand the problem of cybersecurity. Firms tend to underinvest in cyberdefenses for the same reason they tend to underinvest in pollution control technologies – because insecurities that result in successful cyberattacks produce negative externalities that are borne by third parties. Firms also tend to underinvest in cyberdefenses because such expenditures create positive externalities and provide opportunities for free riding. Understood in these terms, the challenge for a cybersecurity regime is to internalize the externalities – to ensure that firms that impose negative externalities by failing to secure their systems are made to bear the resulting costs.

C. . . . as an Antitrust Problem

The ultimate goal of antitrust law, promoting consumer welfare, is achieved by restraining businesses from engaging in anticompetitive conduct. Antitrust law is especially concerned about the possibility that firms will take coordinated action that undermines competition – an agreement by firms to divide a market among themselves, for instance. Antitrust also is apprehensive about information sharing among competitor firms; such exchanges, it is feared, can “facilitate anti-competitive collusion or unilateral oligopolistic behavior.”\(^\text{136}\) Hence section 1 of the Sherman Act sweepingly prohibits “[e]very contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States.”\(^\text{137}\)

Antitrust law often subjects coordinated conduct by multiple firms to stricter scrutiny than isolated conduct by a single firm; “the legal presumption against” joint arrangements “is sometimes thought to be strong.”\(^\text{138}\) Many such arrangements – namely, coordinated action that can be characterized as a “naked” restraint, or a restraint that “is formed with the objectively intended purpose or likely effect of increasing price or decreasing output in the short run”\(^\text{139}\) – are condemned under a “per se rule.”\(^\text{140}\) With a per se rule, there is no need to inquire whether a particular arrangement actually has anticompetitive effects. Instead, antitrust law takes a shortcut and presumes that the conduct is harmful.\(^\text{141}\) This approach may lead to the occasional false positive – coordinated action that is actually beneficial to consumers but that nevertheless is condemned as unlawful. But the conventional wisdom is that the costs of these false positives would be dwarfed by the decision costs of distinguishing the small number of naked restraints that are procompetitive from the much larger number that are anticompetitive.

Yet some inter-firm cooperation is beneficial to consumers,\(^\text{142}\) and antitrust law can struggle to determine whether a given instance of joint action is pro- or anticompetitive.\(^\text{143}\) In the

\(^{136}\) Aviram & Tor, supra note 131, at 236.


\(^{138}\) HOVENKAMP § 5.1, at 211; see also id. § 5.1b, at 214-16.

\(^{139}\) HOVENKAMP § 5.1a, at 212.

\(^{140}\) HOVENKAMP § 5.1, at 211.

\(^{141}\) HOVENKAMP § 5.1, at 211-12.

\(^{142}\) Aviram & Tor, supra note 131, at 231; HOVENKAMP § 5.1, at 211.
cybersecurity context, various forms of coordination and information sharing can help firms better defend themselves against intrusions and thus prevent consumers from incurring losses. Firms in a particular industry might agree to exchange threat information. An ISP that discovers it has been victimized by a particular form of malware could alert others to be on the lookout for the same threat. Or firms could share vulnerability information. A power plant that learns that its SCADA system can be compromised by a particular type of intrusion could tell other companies about the vulnerability. Firms also might share countermeasure information. A company might discover that a particular security solution is an especially effective way to defend against, say, a DDOS attack, and the company might notify other firms to use the same technique. Finally, firms might agree to establish a uniform set of cybersecurity standards for their industry, along with a monitoring and enforcement mechanism to ensure that all members are implementing the agreed upon measures. They might, in other words, form something like a cartel.

Which brings us to the problem. Coordinating on cyberdefense could give rise to antitrust liability, and firms therefore are reluctant to share information with their competitors or to adopt common security standards. These liability fears appear to be fairly widespread. A 2002 analysis found that, among the private sector’s “major concerns about fully communicating cybervulnerabilities,” one of the most important is “the potential for antitrust action against cooperating companies.” In a 2009 report, the American Bar Association likewise recounted the concerns of several firms that “antitrust laws created a barrier to some forms of sharing” cybersecurity information. Government officials have reported the same fears. The White House’s 2009 Cyberspace Policy Review acknowledged that some inter-firm coordination takes place, but went on to report that “some in industry are concerned that the information sharing and collective planning that occurs among members of the same sector under existing partnership models might be viewed as ‘collusive’ or contrary to laws forbidding restraints on trade.”

143 Aviram & Tor, supra note 131, at 236; HOVENKAMP § 5.1, at 211.
144 Emily Frye, The Tragedy of the Cybercommons, 58 BUS. LAW. 349, 369 (2002); Lichtman & Posner, supra note 59, at 236.
145 Aviram & Tor, supra note 131, at 263.
146 Kobayashi, supra note 33, at 23.
148 Frye, supra note 144, at 374. The other two concerns are “an increased risk of liability” and the “loss of proprietary information.” Id.
149 ABA, supra note 18, at 10.
150 Cyberspace Policy Review 18-19 (2009). But see BRENNER, supra note 1, at 228 (dismissing the possibility that cybersecurity coordination might give rise to antitrust liability); Rosenzweig, supra note 14, at 16 (same). Cybersecurity experts sometimes exchange information about threats and vulnerabilities notwithstanding the antitrust laws. For instance, an informal collaboration between researchers at Symantec, the computer security company, and several freelance computer experts in Europe revealed that Stuxnet, originally thought to be a “routine and unambitious” piece of malware, was in fact a sophisticated cyberweapon aimed at Iran’s nuclear program. Zetter, supra note 46. This episode is important for two reasons. First, it confirms that information sharing can produce significant cybersecurity gains. Second, it suggests that information sharing is more likely to take place
These concerns seem well founded. There are a number of scenarios in which cybersecurity coordination conceivably could trigger liability under federal antitrust statutes. For instance, suppose that firms in a particular industry agree to implement a uniform set of cybersecurity practices. It is improbable that these new standards will be costless. Whether the companies have agreed to purchase and install new firewall software, or to transition from vulnerable commercial off the shelf (“COTS”) operating systems to more expensive proprietary operating systems, the measures are likely to affect their bottom lines. Industry members might decide to absorb these increased costs, depending on the elasticity of consumer demand for the goods or services they offer. But they might further decide to pass on these costs to consumers, either in the form of a general price hike or as a free standing surcharge.

Would the arrangement be lawful? This sort of venture may well amount to price fixing in violation of the Sherman Act. Even if the participating firms are not setting a specific price for their products (everyone will now charge $50 for widgets instead of $45), they are still establishing a premium that will be assessed for their products (everyone will increase the price they charge for their widgets by $5). The economic effect is the same. Indeed, the arrangement may even amount to a “naked” restraint that triggers review – and therefore probably reflexive condemnation – under the per se rule.\textsuperscript{151} The venture also might stand condemned as an unlawful tying arrangement. Tying occurs when a firm requires a consumer to purchase one product as a condition of purchasing another\textsuperscript{152}; Canon refuses to sell you a camera unless you also buy a flash. Like naked restraints, tying arrangements typically are reviewed under a per se rule.\textsuperscript{153} Transferring the increased costs of cybersecurity to consumers might be seen as an effort to force them to buy a new security product in addition to the firm’s basic product. Imagine a bank that previously would have offered financial services, such as the ability to use a credit card, for $45 a year. After the agreement, it now sells financial services plus enhanced security for $50 a year. That additional $5 represents the price for a separate product, cybersecurity, which consumers may or may not independently wish to purchase.\textsuperscript{154}

As a second example of how shared cybersecurity standards might violate the antitrust laws, consider an arrangement that imposes no new costs on consumers – at least not directly. Suppose firms in a particular industry agree to install intrusion-detection or -prevention capabilities to scan for malware on their networks.\textsuperscript{155} These systems rely on a technique known as “deep packet inspection,” in which all data traversing the network is scanned and checked against signature files of known malware.\textsuperscript{156} The effect is often to slow down the network’s where there is little risk of antitrust liability. Symantec and European researchers could freely exchange information because they did not offer competing goods or services, so the arrangement was unlikely to be condemned as a contract, combination, or conspiracy in restraint of trade. 15 U.S.C. § 1.

\textsuperscript{151} HOVENKAMP § 5.1a, at 212.
\textsuperscript{152} HOVENKAMP § 10.1, at 435.
\textsuperscript{153} But see Jefferson Parish Hosp. Dist. No. 2 v. Hyde, 466 U.S. 2, 40 (1984) (O’Connor, J., concurring) (arguing that tying arrangements should be reviewed under a rule of reason).
\textsuperscript{154} See sources cited supra note 59.
\textsuperscript{155} POST, supra note 90, at 85.
\textsuperscript{156} CLARKE & KNake, supra note 1, at 161-62; LESSIG, supra note 67, at 55-56; Lynn, supra note 19, at 103.
performance, sometimes dramatically.\textsuperscript{157} Suppose further that the firms decide to absorb the costs of the monitoring or detection system rather than pass them on to their consumers. Would that forbearance save the arrangement from antitrust liability? Not necessarily. The shared security standards still plausibly could be described as an unlawful price fixing agreement. While the participating companies have not agreed to raise prices directly, they have indirectly accomplished something similar; instead of requiring consumers to pay a higher price for the same product, the firms have agreed to require consumers to pay the same price for a lesser product (where speed is an important component of the product’s value).

Notice that clear and unambiguous prohibitions on inter-firm coordination may not be necessary to deter businesses from cooperating with one another. Mere uncertainty about the applicability of the antitrust laws – and the corresponding risk of liability – may be enough to dissuade firms from participating in joint cybersecurity ventures. The deterrent effect of legal ambiguity is likely to be especially strong in this context because of the sanctions that may be imposed on antitrust defendants. Firms that are alleged to have violated federal antitrust laws face criminal prosecutions\textsuperscript{158} as well as federal civil actions,\textsuperscript{159} state civil actions,\textsuperscript{160} and lawsuits by aggrieved private parties,\textsuperscript{161} and each type of civil litigation carries the prospect of treble damages payouts to the successful plaintiffs.\textsuperscript{162} In light of these potential sanctions, private firms will have good reasons to avoid coordinating their efforts to improve cybersecurity.

One final observation about cybersecurity and antitrust. Beyond liability concerns, there are other serious impediments to coordination and information sharing. The difficulties of forming and maintaining cartels are well known. Among other problems, individual cartel members have strong incentives to cheat, such as by offering a greater quantity of product or by charging a higher price than allotted by the cartel.\textsuperscript{163} In the cybersecurity context, businesses will have comparable incentives to shirk their responsibilities to implement any agreed upon (and likely costly) security standards. In addition, firms may be especially reluctant to share information with their competitors.\textsuperscript{164} If a firm discovers an effective way to defend its systems against a particular form of cyber intrusion, that information gives it a comparative advantage over rivals that may not be as adept at protecting their own networks. Sharing the information with competitors enables them to free ride and thereby eliminates the firm’s comparative advantage. As such, even if fears of liability under the antitrust laws were eliminated completely, it is doubtful that firms would fully cooperate with one another. Nevertheless, liability concerns appear to be a significant impediment to cybersecurity coordination and

\textsuperscript{157}\textsc{Clarke & Knaek, supra} note 1, at 81; Smith,\textsc{ supra} note 17, at 180.

\textsuperscript{158}\textsc{CITE}


\textsuperscript{160}15 U.S.C. § 15c.


\textsuperscript{162}Compare 15 U.S.C. § 15(a) (treble damages in private lawsuits), with\textsc{ id.} § 15a (treble damages in lawsuits by United States), with\textsc{ id.} § 15c(a)(2) (treble damages in lawsuits by state attorneys general).

\textsuperscript{163}Hovenkamp § 4.1a, at 161-68.

\textsuperscript{164}Aviram & Tor,\textsc { supra} note 131, at 234; see also Nathan Alexander Sales,\textsc{ Share and Share Alike}, 78 Geo. Wash. L. Rev. 279, 319-20 (2010).
information sharing. Reducing these fears would not by itself ensure cooperation, but doing so would make it more likely at the margin.

D. . . . as a Products Liability Problem

Private investment in cybersecurity also resembles a tort problem – more precisely, a products liability problem. Broadly speaking, the law of products liability has two complementary goals. First, from an ex post perspective, the law seeks to compensate consumers injured by products that did not perform as expected. Second, from an ex ante perspective, products liability law uses the risk of money damages to incentivize firms to take reasonable precautions when designing and manufacturing products.165

The branch of products liability law that is most relevant to cybersecurity concerns design defects. In a design defect case, the theory is that “the intended design of the product line itself is inadequate and needlessly dangerous.”166 (By contrast, a manufacturing defect occurs when a product suffers from “a random failing or imperfection,”167 such as a crack in a Coke bottle that causes it to explode,168 and a marketing defect occurs when an otherwise safe product “become[s] unreasonably dangerous and defective if no information explains [its] use or warns of [its] dangers.”169) In its infancy, products liability law typically assigned blame on a theory of strict liability.170 A plaintiff could recover damages by establishing that a given product had a defective design and that he was injured by that defect; it wasn’t necessary to show that the manufacturer was negligent, or otherwise blameworthy, in producing the defect.171 The modern approach abandons strict liability in favor of a negligence standard.172 How do courts determine whether a manufacturer was at fault when it produced a product with a design defect? One common approach is the risk-utility test.173 The test, which has its roots in Learned Hand’s negligence formula,174 compares “the risks of the product as designed against the costs of making the product safer.”175 If the risks associated with a product can be reduced by a significant amount at a relatively low cost, a manufacturer that declines to do so is negligent. If

166 DOBBS § 355, at 980; MICHAEL I. KRAUSS, PRINCIPLES OF PRODUCTS LIABILITY 81 (2011).
167 DOBBS § 355, at 979.
168 Lee v. Crookson Coca-Cola Bottling Co., 188 N.W.2d 426 (Minn. 1971).
169 DOBBS § 355, at 981.
171 DOBBS § 353, at 974-75.
172 RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § ___ (1998); see also DOBBS § 353, at 977; KRAUSS, supra note 166, at 40; LANDES & POSNER, supra note 165, at 292.
173 RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 2, cmts. a & f (1998); see also DOBBS § 357, at 985-87; LANDES & POSNER, supra note 165, at 291-92.
174 United States v. Carroll Towing Co. 159 F.2d 169 (2d. Cir. 1947).
175 DOBBS § 357, at 985.
the risks can be reduced only by a small amount at a relatively high cost, a manufacturer that declines to do so is not negligent.

This system of tort liability creates important incentives for manufacturers to prevent or eliminate design defects. Imagine a company that makes residential furnaces; it is trying to decide whether to remedy a design defect that increases the probability the furnaces will explode. The company will do so if the expected benefits of reducing the risk of explosion exceed the expected costs of making the fix. Without tort liability, the benefit of making defect-free furnaces is lower than it otherwise would be. Furnaces that occasionally explode would damage the firm’s reputation, and some consumers likely would buy competitors’ products instead. The manufacturer benefits to the extent it reduces these harms. But it doesn’t face the prospect of paying money damages to homeowners whose houses burned down. The cost-benefit calculus looks very different once a products liability regime is in place. Now, a decision to eliminate the design defect will be more beneficial. The company thereby reduces its exposure to potentially ruinous money damages awards, at least where the defect can be cured in a sufficiently low-cost way that the failure to do so would be deemed negligent. In short, a products liability regime increases a firm’s expected benefit of remedying design defects – namely, the benefit of foregone money damages, discounted by the probability that they would be awarded. It thus increases the number of circumstances in which firms will find it welfare maximizing (benefit > cost) to improve the safety of their products. The result is that, at the margin, products will be safer than they otherwise would be.

Internet-related goods and services sometimes suffer from design defects that increase their vulnerability to cyberattacks. Perhaps the best known example is Microsoft Windows. The operating system software, which accounts for more than 90 percent of the PC market, is notoriously riddled with vulnerabilities. These vulnerabilities stem in part from the software’s size. Windows Vista, released in 2007, featured some 50 million lines of code, compared to 35 million for Windows XP (released in 2001) and just 15 million for Windows 95 (released in 1995). It is more or less inevitable that the programmers who write these millions of lines will make mistakes, and it can be quite difficult to detect and repair them. (Given that it probably would cost a great deal to eliminate all of these vulnerabilities, the failure to do so may not be negligent under the risk-utility test.) Other examples abound. Indeed, many of the vulnerabilities described in Part I can be understood as the results of design defects. Consider the decision by power companies to connect generators and other elements of the electrical grid to the internet. This can be seen as a form of defective system design, in that internet

---

177 Lichtman & Posner, supra note 59, at 255.
179 See, e.g., __.
180 Lohr & Markoff, supra note 178.
181 Dorothy Denning, Information Warfare and Security 12 (1999); Katyal, Digital Architecture, supra note 15, at 2264-65; Sarnikar & Johnsen, supra note 19, at 18.
182 But see Lichtman & Posner, supra note 59, at 255 (arguing that improving the security of Windows is “simply a matter of investing more resources in product design as well as testing”).
connectivity exposes the nation’s power grid to potentially catastrophic cyberattacks in exchange for relatively minor benefits.\textsuperscript{183} The same can be said of companies that continue to protect their SCADA systems with vendor supplied default passwords\textsuperscript{184} – a defect, incidentally, that could be remedied at a negligible cost.

The incentives to cure these design defects are fairly weak, because poor cyber security generally does not trigger civil liability. “[L]iability has played virtually no role in achieving greater Internet security.”\textsuperscript{185} One reason for this is a venerable chestnut of tort law known as the economic loss doctrine. The economic loss doctrine provides that, while a defendant who causes physical injuries is also liable for any resulting economic harms, he generally is not liable for freestanding economic harms. “When commercial or economic harm stands alone, divorced from injury to person or property, courts have not imposed a general duty of reasonable care.”\textsuperscript{186} Many of the harms that would result from a cyberattack on, say, the power grid or the financial sector would be purely economic in nature. An automobile manufacturer might be unable to run its assembly line because the power is out, or a consumer might default on a loan because he can’t make a payment online. Few of these harms would derive from a physical injury, and they therefore would not be actionable under the economic loss doctrine. For instance, in 2009, the Massachusetts Supreme Judicial Court dismissed a lawsuit brought by credit unions against a retailer after hackers accessed the retailer’s computer systems and stole customer credit card data. The court emphasized that, because “the plaintiffs suffered only economic harm due to the theft of the credit card account information,” the “economic loss doctrine barred recovery on their negligence claims.”\textsuperscript{187} (Cyberattacks that cause injuries to person or property presumably would remain actionable, as would any resulting economic harms. So, for instance, if an attacker exploited a design defect in a dam’s control system and opened the floodgates,\textsuperscript{188} the dam operator might be held liable for the deaths of the downstream landowners and any corresponding economic losses.)

The problem can be understood in Coasean terms.\textsuperscript{189} Consider the famous example of a train that emits sparks that burn the wheat in neighboring fields. Regardless of whether the legal entitlement is initially assigned to the railroad (a right to emit sparks) or the farmers (a right to be free from incinerated crops), the parties will bargain to reallocate the entitlement to its socially most efficient use (assuming that the transaction costs are sufficiently small). In the

\textsuperscript{183} See supra notes 42 to 47 and accompanying text.

\textsuperscript{184} See supra notes 63 to 65 and accompanying text.

\textsuperscript{185} BRENNER, supra note 1, at 224; see also Schnieier, supra note 33, at 2.

\textsuperscript{186} DOBBS § 452, at 1282; see also id., § 452, at 1285-87 (discussing exceptions to the economic loss doctrine); LANDES & POSNER, supra note 165, at 251. The rule’s familiar rationales are, first, the fact that “financial harm tends to generate other financial harm endlessly and often in many directions” and the corresponding recognition that liability “would be onerous for defendants and burdensome for courts”; and, second, the notion that “contract law is adequate to deal with the problem and also usually more appropriate.” DOBBS § 452, at 1283.


\textsuperscript{188} Frye, supra note 144, at 350; see also Sklerov, supra note 19, at 20.

cybersecurity context, the absence of tort liability essentially grants firms a legal right to refrain from taking precautions that would protect third parties from attacks on their systems or products. This may be an efficient allocation of the legal entitlement in some contexts, but not always. In these latter circumstances, companies and third parties theoretically should negotiate and establish a new legal right to be free from harm due to cyber intrusions. But Coasean bargaining over cybersecurity seems unlikely to occur because of the staggering transaction costs: It would be prohibitively expensive, if not impossible, for companies to bargain with everyone who conceivably could be injured by cyberattacks on their systems or products.

Beyond tort, it is doubtful that other sources of law credibly will threaten cybersecurity shirkers with liability. Contract law does not seem well suited to the task. Software manufacturers typically do not offer warranties that their products are secure. Indeed, some do not “sell” software at all. They merely grant a license, and users cannot install the software unless they click a button to accept the terms and conditions of the license – which very often include a limit on the manufacturer’s liability. Likewise, federal law extends broad immunity to internet service providers. Section 230 of the Communications Decency Act provides that an ISP will not be “treated as the publisher or speaker of any information” that transits its network. This statute has been interpreted by at least one federal appellate court to foreclose a lawsuit alleging that an ISP negligently failed to prevent malware from being sent over its network. From the standpoint of a profit-maximizing firm, then, the expected benefits of remediing a cyber vulnerability often will be lower than the expected costs. Without the prospect of civil liability, firms have weaker incentives to invest in measures to secure their systems and products against cyberattacks.

Not only are liability fears failing to incentivize firms to take better precautions against cyberattacks, they are actually discouraging them from doing so. Companies sometimes are reluctant to better secure their systems because of concerns that these steps could expose them to civil liability. For instance, ISPs typically do not offer assistance if they discover that their customers’ PCs have been infected by malware. ISPs often are able to tell, through routine traffic analysis, that a particular machine on the network is part of a botnet or has been infected by a worm. “[B]ut they don’t dare inform the customer (much less cut off access) out of fear that customers would . . . try to sue them for violating their privacy.” Doing so might even be a crime. The federal wiretap act makes it unlawful to “intentionally intercept[] . . . any wire, oral, or electronic communication,” and some companies fear that filtering botnet traffic or

190 Frye, supra note 144, at 367.
191 BRENNER, supra note 1, at 224.
193 Green v. America Online, 318 F.3d 465, 471 (3d Cir. 2003); see generally Lichtman & Posner, supra note 59, at 247-52.
194 BRENNER, supra note 1, at 229; CLARKE & KNAKE, supra note 1, at 164.
195 CLARKE & KNAKE, supra note 1, at 164-65; see also BRENNER, supra note 1, at 229; Coldebelia & White, supra note 14, at 236-37.
196 18 U.S.C. § 2511(1)(a)
other malware might fall within this prohibition.\(^{197}\) And while federal law makes an exception for ISPs that intercept communications to protect their own property,\(^{198}\) there is no parallel exception for intercepts intended to protect the property of subscribers. Likewise, some ISPs are using deep packet inspection to examine the data streams on their networks for malicious code (this is probably lawful under the exception mentioned above, or a separate exception for “mechanical or service quality control checks”\(^{199}\)). But even when they uncover malware, ISPs “have been reluctant to ‘black hole’ (or kill) malicious traffic because of the risk that they might be sued by customers whose service is interrupted.”\(^{200}\) Again, as in the antitrust context,\(^{201}\) even if the applicable service contracts or state and federal laws do not clearly forbid these measures, the mere risk of liability may be enough to dissuade firms from undertaking them.

While firms with poor cyberdefenses generally do not face the prospect of civil lawsuits, there is one context in which a credible liability threat exists: data breaches in the financial services sector. The Gramm-Leach-Bliley Act of 1999 directs a group of federal agencies, such as the Federal Trade Commission and the Federal Deposit Insurance Corporation, to issue data security regulations for financial institutions.\(^{202}\) In particular, the act mandates the adoption of “administrative, technical, and physical safeguards” that will, among other things, “insure the security and confidentiality of customer records and information” and “protect against unauthorized access to or use of such records.”\(^{203}\) The sanctions for violating these data security requirements can be severe. The GLB Act does not enumerate specific penalties, but rather directs the enforcing agencies to apply the act’s requirements according to their respective enabling statutes.\(^{204}\) Thus, for example, a bank subject to FTC jurisdiction would face a civil penalty of up to $16,000 for each violation.\(^{205}\) If the FTC treated every customer affected by a cyber intrusion as a separate violation, the penalties very quickly would become staggering. If 100 customers have their data compromised the bank would face up to a $1.6 million penalty, 10,000 customers would mean up to a $160 million penalty, and so on.

Not coincidentally, financial institutions are widely believed to do a better job of protecting customer data than members of other industries.\(^{206}\) Unlike other firms, which typically spend only modest sums on cybersecurity, most banks make large investments,

\(^{197}\) Brenner, supra note 1, at 229-30.


\(^{200}\) Clarke & Knaake, supra note 1, at 163; see also McAfee 2010, supra note 35, at 5.

\(^{201}\) See supra notes 158 to 162 and accompanying text.


\(^{204}\) 15 U.S.C. § 6805(b).

\(^{205}\) 16 C.F.R. § 1.98.

\(^{206}\) ABA, supra note 18, at 21; Frye, supra note 144, at 367-68; Powell, supra note 14, at 501-05. But see Gable, supra note 2, at 84 (emphasizing that banks remain vulnerable to cyberattack).
“between 6 and 7 percent of their entire information technology budgets.”207 Financial institutions also are more likely to adopt specific security measures like intrusion detection and prevention systems, antivirus software, smart cards, and biometrics.208 The unique risk of liability that banks face may be responsible, at least in part, for that record. The GLB Act has the effect of increasing the expected benefit of cybersecurity – namely, avoiding potentially crippling civil penalties – and thus creates strong incentives for banks to invest in defenses. (Another explanation is the risk of customer exit. It is relatively easy for a customer who fears cyber intrusions to switch banks, so the bank has an incentive to maintain data integrity.209) Of course, the GLB Act’s emphasis on protecting consumer data might distort firms’ cybersecurity investments. Rather than expending resources on defenses against the attacks they regard as the most dangerous, or the most likely to occur, financial institutions will tend to prioritize defenses against the one form of intrusion singled out by their regulators – the compromise of customer data.210 The effect may be to ensure that firms are well defended against one threat at the expense of increased exposure to many other threats.211 Even so, Gramm-Leach-Bliley remains an example of how the risk of civil liability might be used to incentivize firms to improve (at least some of) their cyberdefenses.

E. . . . as a Public Health Problem

As several scholars have noted, in more or less detail, cybersecurity can be thought of in terms of public health.212 A critically important goal for any cybersecurity regime is to keep attacks from happening and to contain their ill effects.213 The same is true of public health, the ultimate goal of which is prevention.214 Unlike medical practice, which typically has an ex post orientation (treating illnesses that have already occurred), public health is primarily oriented toward ex ante solutions – preventing people from contracting infectious diseases, preventing pathogens from spreading, and so on. Broadly summarized, public health law – including the subset known as public health emergency law – involves government efforts “to persuade, create incentives, or even compel individuals and businesses to conform to health and safety standards for the collective good.”215 These interventions are sometimes defended, controversially, on paternalistic grounds. The notion is that the state may curtail individuals’ freedoms to promote their own interests, which in this context means their physical health and safety.216 By far the

207 Powell, supra note 14, at 502
208 Powell, supra note 14, at 503.
209 See supra notes 40 to 47 and accompanying text.
210 Similar distortions may arise at the state level, as a number of states have enacted laws requiring designated companies to disclose breaches of customer data. Vincent R. Johnson, Cybersecurity, Identity Theft, and the Limits of Tort Liability, 57 S.C. L. REV. 255, ___ (2005); Schwartz & Janger, supra note 59, at 917.
211 Cf. BAKER, supra note 24, at 238-39; McAfee 2010, supra note 35, at 29.
212 See generally Rattray et al., supra note 8; IBM, supra note 19; see also Coyne & Leeson, supra note 18, at 480; Hunker, supra note 19, at 202-03; Katyal, Criminal Law, supra note 10, at 1081; Rosenzweig, supra note 14, at 19.
213 Katyal, Community, supra note 127, at 34; Katyal, Criminal Law, supra note 10, at 1078-79.
214 LAWRENCE O. GOSTIN, PUBLIC HEALTH LAW 19 (2d ed. 2008).
215 GOSTIN, supra note 214, at xxii.
216 GOSTIN, supra note 214, at 50-54.
more common, and widely accepted, justification for public health law is the risk of harm to others. The state may coerce persons who have contracted an infectious disease or are at risk of doing so, the theory goes, to prevent them from transmitting the disease to, and thereby harming, others.\footnote{GOSTIN, supra note 214, at 49.} Seen in this light, a principal objective of public health law is to internalize negative externalities – in particular, the costs associated with spreading infections to others.

Public health law contemplates three specific measures that are relevant here: mandatory inoculations to reduce susceptibility to infectious diseases, biosurveillance to monitor for epidemics and other outbreaks, and isolation and quarantine to treat those who have been infected and prevent them from spreading the pathogen.\footnote{GOSTIN, supra note 214, at 11, 39.} We will consider each in turn along with their potential relevance to cybersecurity.

Inoculation, in which a healthy subject is exposed to a pathogen, helps prevent disease both directly (a person who is inoculated against a disease is thereby rendered immune) and indirectly (the person’s immunity reduces the risk that he will transmit the disease to others). Inoculation mandates can take several forms. In the nineteenth and early twentieth centuries, state and local governments sometimes opted for direct regulation – a firm legal requirement that citizens must receive a particular vaccine, backed by the threat of sanctions.\footnote{GOSTIN, supra note 214, at 379.} (In the 1905 case of Jacobson v. Massachusetts, the Supreme Court upheld such a requirement against a lawsuit invoking Fourteenth Amendment’s privileges or immunities, due process, and equal protection clauses.\footnote{197 U.S. at 34.} According to the Court, mandatory inoculation is a permissible exercise of the states’ police powers.) The modern approach usually involves a lighter touch. Now, state and local governments typically create incentives for citizens to undergo inoculation by making it a condition of eligibility for certain valuable benefits. The best known example is to deny children access to public schools unless they have been vaccinated.\footnote{GOSTIN, supra note 214, at 379-80, 382; Hunker, supra note 19, at 203; Lichtman & Posner, supra note 59, at 255.} (The Supreme Court upheld such a scheme in 1922 in Zucht v. King.\footnote{260 U.S. 174 (1922).})

It isn’t necessary to inoculate all members of a population to frustrate the transmission of a given disease. This is due to “herd immunity.” Herd immunity theory proposes that, for a contagious disease that is transmitted from person to person, chains of infection are likely to be disrupted when large numbers of a population are immune or less susceptible to the disease.\footnote{Katyal, Criminal Law, supra note 10, at 1081.} The critical number is typically around 85 percent of the population, but it can be as low as 75 percent for some diseases (such as mumps) and as high as 95 percent for others (such as pertussis – i.e., whooping cough).\footnote{http://www.bt.cdc.gov/agent/smallpox/training/overview/pdf/eradicationhistory.pdf.} Herd immunity is a form of positive externality – those who undergo

\footnotetext[217]{GOSTIN, supra note 214, at 49.}
\footnotetext[218]{GOSTIN, supra note 214, at 11, 39.}
\footnotetext[219]{GOSTIN, supra note 214, at 379.}
\footnotetext[220]{197 U.S. 11 (1905).}
\footnotetext[221]{197 U.S. at 34.}
\footnotetext[222]{GOSTIN, supra note 214, at 379-80, 382; Hunker, supra note 19, at 203; Lichtman & Posner, supra note 59, at 255.}
\footnotetext[223]{260 U.S. 174 (1922).}
\footnotetext[224]{Katyal, Criminal Law, supra note 10, at 1081.}
\footnotetext[225]{http://www.bt.cdc.gov/agent/smallpox/training/overview/pdf/eradicationhistory.pdf.}
vaccination provide an uncompensated benefit to those who do not – and, as a result, there is a free rider problem.\(^\text{226}\) Many people would prefer to enjoy the benefits of herd immunity without themselves undergoing vaccination, which is costly (money, discomfort, risk of reaction, etc.); they would rather be part of the 15 percent than the 85 percent. The effect is to weaken each person’s incentive to undergo vaccination, and overall vaccinations may drop below the levels needed to support herd immunity. The need to overcome this free rider problem helps explain why state and local governments sometimes use their coercive powers to require inoculation. (Another approach would be to provide subsidies to those who have been inoculated. Public school vaccination requirements can be understood in these terms; the government is subsidizing the education of children who are inoculated.)

Ensuring widespread immunity – not to disease, but to malicious code – is also an important goal of cybersecurity. The average internet-connected computer may be even more susceptible to infection by malware than the average person is to infection by a pathogen, because malicious code can propagate more efficiently than disease. Many pathogens are transmitted by person-to-person contact; you are unlikely to contract polio unless you come into close proximity with someone who is already infected. But one can contract malware from virtually any (networked) computer in the world. In effect, the internet brings dispersed systems into direct contact with one another; alternatively, the internet is a disease vector that, like mosquitoes and malaria, can transmit a contagion between dispersed systems. It is therefore essential for the elements at the edge of the network (such as the SCADA system that runs the local power plant) to maintain effective defenses against cyber intrusions (such as disconnecting the power plant’s controls from the public internet). And there’s the rub. As with herd immunity, cybersecurity raises free rider problems.\(^\text{227}\) A user who takes steps to prevent his computer from being infected by a worm or impressed into a botnet thereby makes other systems more secure; if the user’s machine is not infected, it cannot transmit the malware to others. But the user receives no compensation from those who receive this benefit; he does not internalize the positive externality. He therefore has weaker incentives to secure his system, as he – like everyone else – would prefer to free ride on others’ investments. A critical challenge for any cybersecurity regime, then, is to reverse these incentives.

The second key element of public health law is biosurveillance. “Biosurveillance is the systematic monitoring of a wide range of health data of potential value in detecting emerging health threats.”\(^\text{228}\) Public health officials collect and analyze data to determine a given disease’s incidence, or “the rate at which new cases occur in a population during a specified period,” as well as its prevalence, or “the proportion of a population that are cases at a point in time.”\(^\text{229}\) Effective biosurveillance is a vital first step in managing an epidemic or other outbreak.\(^\text{230}\) Biosurveillance takes place through a partnership among the U.S. Centers for Disease Control

\(^{226}\) Coyne & Leeson, supra note 18, at 480; GOSTIN, supra note 214, at 378-79. See generally supra notes 134 to 135 and accompanying text.

\(^{227}\) See supra notes 126 to 135 and accompanying text.

\(^{228}\) GOSTIN, supra, at 291.

\(^{229}\) Rattray et al., supra note 8, at 152.

\(^{230}\) IBM, supra note 19, at 11.
and Prevention, the CDC’s state level counterparts, and front line health care providers (such as hospitals, clinics, and individual medical practitioners). Many, if not all, states have enacted legislation requiring specified health care professionals to notify state authorities if their patients have contracted any number of infectious diseases, such as smallpox or polio. These reports typically include the patient’s name, the type of disease, his medical history, and other personal information. State authorities then share the data with the CDC. These reports are not required by law, but most states appear to be fairly conscientious about them. Public health law thus relies on a system of distributed surveillance. No central regulator is responsible for collecting all the data needed to detect and respond to infectious disease outbreaks. Instead, the system relies on individual nodes within a far flung network – from state agencies to hospitals to individual doctors – to gather the necessary information and route it to the CDC’s central storehouse. CDC then analyzes the data and issues alerts advising state agencies and medical practitioners about disease trends and offering recommendations about how to respond.

The third public health intervention involves containing infectious diseases once an outbreak has occurred, and preventing them from spreading further. Two key measures are isolation and quarantine. Isolation and quarantine differ in subtle ways, though in colloquial usage the terms are essentially synonymous. The goal of each is to segregate from the population those who have contracted or been exposed to an infectious disease and thus prevent them from transmitting it to those who are well. Isolation and quarantine are often coupled with mandatory treatment, which helps reduce the risk of further contagion; a person who has been cured of an infectious disease cannot transmit it to others. The rationale for these interventions is the familiar harm principle – i.e., the risk that a person who has contracted or been exposed to a pathogen will infect others. Isolation and quarantine thus seek to reduce negative externalities.

At the federal level, isolation and quarantine are accomplished under the Public Health Service Act of 1944. The Secretary of Health and Human Services has authority under the act to

---

231 GOSTIN, supra note 214, at 295-96.
232 http://www.cdc.gov/mmwr/pdf/wk/mm5853.pdf
233 GOSTIN, supra note 214, at 297.
234 GOSTIN, supra note 214, at 296; Hunker, supra note 19, at 202-03.
235 This reporting scheme is permissible under the Health Insurance Portability and Accountability Act privacy rule, which generally limits the use and disclosure of protected health information, 45 C.F.R. § 164.502(a), but which contains an exception for disclosures to public health authorities, 45 C.F.R. § 164.512(b). The reporting is probably constitutional as well. The Supreme Court in Whalen v. Roe, 429 U.S. 589 (1977), upheld, against a Fourth Amendment challenge, a similar New York law requiring physicians to report information about drug prescriptions.
236 Rattray et al., supra note 8, at 154.
237 Isolation involves separating persons who are known to be infected with a disease, for as long as the disease remains communicable. GOSTIN, supra note 214, at 429. Quarantine involves separating persons who, though asymptomatic, may have been exposed to a disease, for the period of communicability. GOSTIN, supra note 214, at 429.
238 GOSTIN, supra note 214, at 411-12.
239 GOSTIN, supra note 214, at 414-15.
“make and enforce such regulations as in his judgment are necessary to prevent the introduction, transmission, or spread of communicable diseases” into or within the United States.\textsuperscript{240} The law further provides for the “apprehension, detention, or conditional release” of persons who may have been exposed to any one of several communicable diseases that the president has specified by executive order.\textsuperscript{241} The list, which was updated most recently in 2005,\textsuperscript{242} includes cholera, tuberculosis, plague, smallpox, SARS, and several other diseases.\textsuperscript{243} Large scale isolation and quarantine are rarely used; the most recent example is from the 1918 Spanish flu pandemic (and it was carried out under different legal authorities). However, isolation and quarantine are sometimes used for particular individuals. In May 2007, HHS issued an isolation order for an American with multi-drug resistant tuberculosis who flew from the Czech Republic to Canada and then crossed the land border into the United States.\textsuperscript{244} Violations of the quarantine regulations carry criminal penalties – a fine of up to $1000 and incarceration for up to a year.\textsuperscript{245}

Both biosurveillance and isolation/quarantine have important lessons for cybersecurity. Like the public health system, effective cyberdefenses depend on information about the incidence and prevalence of various kinds of malware. Users – both individuals on their PCs and firms that operate extensive networks – need to know what new forms of malicious code are circulating on the internet in order to secure their systems against them. As for isolation and quarantine, a critical cybersecurity challenge is to ensure that systems infected with malicious code do not spread the contagion to other, healthy computers. In both cases, cybersecurity faces the same problems that arise in realspace disease control, and public health solutions therefore might be adapted for the cyber context.

There is, of course, a significant difference between infectious diseases and malicious computer code: Diseases (typically) develop and spread on their own, whereas malware is created by human beings and (sometimes) requires human intervention to propagate. This is true as far as it goes, but the differences between cyberspace and realspace pathogens can be overstated. Infectious diseases can be engineered (e.g., biological weapons), and sometimes malware is able to spread on its own (e.g., a worm that is programmed to search for other computers on which to replicate itself\textsuperscript{246}). Another potential obstacle is the antiquity of public health statutes.\textsuperscript{247} Many of these laws have been on the books for decades, even a century, and they do not necessarily reflect contemporary scientific understandings of disease.\textsuperscript{248} Plus they often restrict civil liberties and privacy to a degree rarely seen today.\textsuperscript{249} The judicial precedents

\begin{thebibliography}{99}
\bibitem{240} 42 U.S.C. § 264(a).
\bibitem{241} 42 U.S.C. § 264(b).
\bibitem{242} Exec. Order No. 13375 (Apr. 1, 2005).
\bibitem{243} Exec. Order No. 13295 (Apr. 4, 2003).
\bibitem{244} http://www.hhs.gov/asl/testify/2007/06/t20070606b.html.
\bibitem{245} 42 U.S.C. § 271(a).
\bibitem{246} \textit{See supra} note 22 and accompanying text.
\bibitem{247} \textit{Institute of Medicine, The Future of the Public’s Health in the 21st Century} 4 (2002).
\bibitem{248} \textit{Gostin, supra} note 214, at 24.
\bibitem{249} \textit{Gostin, supra} note 214, at 24.
\end{thebibliography}
upholding these statutes against various constitutional challenges typically antedate the Supreme Court’s modern civil rights and liberties jurisprudence. It is not clear that today’s Court would uphold, say, mandatory vaccination of adults as readily as it did in 1905. Yet even if public health law fits uneasily into contemporary constitutional law, it can still be a useful framework for cybersecurity. This is so because, as explained below, the cyber versions of public health interventions can be friendlier to civil liberties and privacy than their realspace counterparts.

III. REGULATORY PROBLEMS, REGULATORY SOLUTIONS

This concluding section examines the responses of environmental, antitrust, products liability, and public health law to the various challenges that arise in those fields, and it considers how those solutions might be adapted to the field of cybersecurity. The range of possible responses to cyber insecurity depends on our understanding of that problem. The security measures we choose are determined by our antecedent choice of how to describe the problem in the first place. If we regard cybersecurity from the standpoint of law enforcement and armed conflict, we will tend to favor the responses of law enforcement and armed conflict – stronger penalties for cyber intrusions, say, or retaliating with kinetic attacks, and so on. Those are plausible frameworks and equally plausible solutions. But they are not the only ones. A wider angle lens is needed. Going beyond the conventional approaches, and conceiving of cybersecurity in terms of the regulatory disciplines surveyed above, brings into focus certain aspects of the problem that otherwise might have gone unnoticed. A more comprehensive set of analytical frameworks also enlarges the menu of legal and policy responses available to decisionmakers.

Taken together, the frameworks described in Part II suggest that an effective cybersecurity regime should include four components: (1) monitoring and surveillance to detect malicious code; (2) hardening vulnerable targets and enabling them to defeat intrusions; (3) building resilient systems that can function during an attack and recover quickly; and (4) responding in the aftermath of an attack. There are two complementary objectives here: preventing intrusions from happening at all, and enabling firms to withstand the intrusions that do take place. Stronger defenses would provide an obvious, first order level of protection: Better defense means less damage. They also would provide an important second order level of protection: Stronger defenses can help achieve deterrence. By enabling victims to defeat, survive, and recover from cyberattacks, these measures increase the expected costs of an intrusion to an attacker (i.e., the costs one must bear to overcome the defenses) and also decrease its expected benefits. And that means weaker incentives to attack in the first place; why expend your scarce resources to try to take down the power grid if the effort is likely to fail?

250 Jacobson v. Massachusetts, 197 U.S. 11 (1905). But see GOSTIN, supra note 214, at 130 (proposing that the Court “indisputably” would reach the same result if it decided Jacobson today).

251 See infra notes 272 to 273 and accompanying text.

252 Cf. Nojeim, supra note 14, at 131; Trachtman, supra note 54, at 265.

253 Bambauer, supra note 12, at 673; Yochai Benkler, Peer Production of Survivable Critical Infrastructures, in Grady & Parisi, supra note 19, at 76-77; BRENNER, supra note 1, at 214; CLARKE & KNAKE, supra note 1, at 159.

254 CSIS, supra note 8, at 26; Lynn, supra note 19, at 99-100; Taipale, supra note 92, at 36.
Of course, it is inevitable that some attacks will succeed. Some intrusions can be prevented or mitigated but others cannot, and any defensive scheme necessarily will be imperfect.\textsuperscript{255} (This is so because, in cyberspace, offense is much less costly than defense. “Defending a modern information system” is like “defending a large, thinly populated territory like the nineteenth century Wild West: the men in black hats can strike anywhere, while the men in white hats have to defend everywhere.”\textsuperscript{256}) The goal therefore is not to develop impregnable defenses. Doing so may be impossible from a technological standpoint, and even if they were feasible they may be inefficiently costly.\textsuperscript{257} Instead, the goal is to adopt efficient levels of investment in defenses that are better at protecting society’s critical systems than current defenses are.\textsuperscript{258} Another important point is that cyberdefense is not a one-size-fits-all proposition. Security measures should be tailored to the unique risks faced by specific firms or industries - their combinations of vulnerabilities, threats, and consequences.\textsuperscript{259} The strongest (and presumably most costly) defenses should be reserved for the firms that are most vulnerable to cyberattacks, that face the most severe threats (e.g., from foreign intelligence services as opposed to recreational hackers), and whose compromise would have the most devastating consequences for society. Strategically unimportant firms might get by with modest defenses, whereas robust defenses may be needed for critical industries.\textsuperscript{260} Finally, what follows is by no means an exhaustive list of possible responses to cyber insecurity. It is merely a list of responses that are implied if we conceive of cybersecurity in environmental, antitrust, products liability, and public health terms. Other solutions, suggested by other analytical frameworks, may be just as promising.

A. Monitoring and Surveillance

Effective cybersecurity depends on the generation and exchange of information about cyber threats.\textsuperscript{261} An ideal system would create and distribute vulnerability data (the holes intruders might exploit to gain access to computer systems), threat data (the types of malware circulating on the internet and the types of attacks firms have suffered), and countermeasure data (steps that can be taken to prevent or combat infection by a particular piece of malicious code).\textsuperscript{262} Perhaps the best way to collect this information is through a distributed surveillance network akin to the biosurveillance system at the heart of public health law. Companies are unlikely to participate in this sort of arrangement due to fears of liability under antitrust and other laws. A suite of measures is therefore needed to help foster favorable incentives, including

\begin{itemize}
  \item Bambauer, supra note 12, at 673; CSIS, supra note 8, at 51; Gable, supra note 2, at 65; IBM, supra note 19, at 12; Lynn, supra note 19, at 99; Sklerov, supra note 19, at 8; Taipale, supra note 92, at 9.
  \item See supra notes 30 to 31 and accompanying text.
  \item DOROTHY DENNING, \textit{INFORMATION WARFARE AND SECURITY} 12 (1999).
  \item ABA, supra note 18, at 21; Katyal, \textit{Criminal Law}, supra note 10, at 1080; Nojeim, supra note 14, at 119.
  \item See supra notes 59 to 66 and accompanying text.
  \item \textit{But see} CSIS, supra note 8, at 45 (information sharing should not be “a primary goal”).
  \item See supra notes 144 to 146 and accompanying text.
\end{itemize}
subsidies, threats of liability, and offers of immunity. These steps won’t guarantee that firms will collect and share cybersecurity data, but they will make such arrangements more viable than they are at present.

Public health law’s system of distributed biosurveillance seems well suited to the challenge of gathering and disseminating data about a vast range of cyber threats. Like health care providers who diagnose and then report their patients’ infectious diseases, firms could be tasked with monitoring their systems for vulnerabilities and intrusions, then reporting their findings (as well as the countermeasures they have implemented) to designated recipients. Such a system would take advantage of important information asymmetries. Individual companies often know more than outsiders about the vulnerabilities in their systems and the types of intrusions they have faced; they have a comparative advantage in compiling this data.263 The principal alternative – surveillance by a single, central regulator – is unlikely to be as effective. As Hayek emphasized, “the knowledge of the [economic] circumstances of which we must make use never exists in concentrated or integrated form but solely as the dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess.”264 The same is true of cybersecurity data. A central regulator lacks the capacity to examine each device that is connected to the internet to determine its vulnerabilities, nor can it inspect every data packet transiting the internet to determine whether it contains malicious code. And even if the scope of the project wasn’t prohibitively vast, the privacy costs associated with a central monitor – especially a government monitor – likely would be intolerable. Instead, the more efficient course would be to rely on individual firms to gather the relevant information.265

While firms would be responsible for the lion’s share of monitoring, there is still an important role for the government: providing especially sensitive companies (such as power companies and ISPs) with information about especially sophisticated forms of malware. Here, the comparative advantage is reversed; the government’s highly resourceful intelligence agencies are simply better than the private sector at detecting intrusions by sophisticated adversaries like foreign militaries and developing countermeasures.266 The government can provide these firms with the signatures of malware used in previous attacks, and firms can use the signature files to detect future intrusions. In 2010, for instance, the National Security Agency began assisting Google in detecting intrusions into its systems. The partnership was announced in the wake of reports that sophisticated hackers, most likely affiliated with China’s intelligence service, had broken into Google’s systems and collected data about users, including a number of human rights activists.267 The NSA reportedly has entered a similar partnership with a number of large banks.268

263 Bamberger, supra note 203, at 391-92; CSIS, supra note 8, at 53; Katyal, Criminal Law, supra note 10, at 1091. See generally Bamberger, supra note 203, at 399 (emphasizing “the information asymmetries between regulated firms and administrative agencies”).


265 CLARKE & KN AKE, supra note 1, at 162.

266 Condron, supra note 19, at 407; Coldebella & White, supra note 14, at 240. But see O’Neill, supra note 19, at 265, 27; Taipale, supra note 92, at 9.

267 Nakashima, Google, supra note 56.

What should be the architecture of the system used to disseminate the vulnerability, threat, and countermeasure information compiled by private firms? At least two possibilities exist. Some commentators have called for the creation of a central repository of cybersecurity data—a “cyber-CDC,”\textsuperscript{269} as it were. Under such a system, an individual firm would notify the clearinghouse if it discovers a new vulnerability in its systems, or a new type of malicious code, or a particular countermeasure that is effective against a particular kind of threat. The repository would analyze the information, looking for broader trends in vulnerabilities and threats, then issue alerts and recommendations to other firms. This clearinghouse might be a government entity, as in public health law, but it need not be. An alternative architecture would be for firms to exchange cybersecurity information with one another directly, on a peer-to-peer basis, rather than first routing it through a central CDC-type storehouse. One advantage of the peer-to-peer approach is that it may be more resilient. A central storehouse would be an attractive target for cyber adversaries, and the entire system would fail if it were compromised.

Distributed surveillance may be an even better fit for cybersecurity than for public health, for several reasons. First, malicious computer code often can be detected more quickly than biological pathogens,\textsuperscript{270} which means that countermeasures can be put in place rapidly. Biosurveillance can be slow because the incubation period for certain diseases—i.e., the amount of time between when a disease is contracted and when its symptoms first manifest—can be days or weeks. By contrast, it is possible to detect known malware in real time, as the code is passing through a company’s system (assuming, of course, that the firm’s intrusion detection systems know what to look for, which is not always the case\textsuperscript{271}). Second, cyberthreat monitoring has the potential to raise fewer privacy concerns than biosurveillance.\textsuperscript{272} Health care providers often give authorities intensely sensitive information about individual patients, such as their names, Social Security numbers, and other personally identifiable information, as well the diseases they have contracted.\textsuperscript{273} A properly designed cyber monitoring system need not compile and disseminate information of the same sensitivity. Collection and sharing could be limited to information about the incidence and prevalence of known malware. The fact that a particular system has been infected by the “ILoveYou” worm exposes a great deal less personal information, and thus raises weaker privacy concerns, than the fact that a particular patient suffers from HIV or breast cancer.

This framework has important limitations. Malware detection is an inexact science.\textsuperscript{274} One challenge is that deep packet inspection and other forms of network monitoring typically work by comparing streams of data against signature files of known malicious code.\textsuperscript{275} These systems are only as good as their underlying signatures. If there is no signature for a particular

\textsuperscript{269} IBM, supra note 19, at 13-14; see also Sharp, supra note 8, at 25.
\textsuperscript{270} Rattray et al., supra note 8, at 152.
\textsuperscript{271} See infra notes 274 to 277 and accompanying text.
\textsuperscript{272} But see Nojeim, supra note 14, at 126.
\textsuperscript{273} GOSTIN, supra note 214, at 297.
\textsuperscript{274} CLARKE & KNAKE, supra note 1, at 126; Sklerov, supra note 19, at 74.
\textsuperscript{275} See supra note 156 and accompanying text.
type of malware, chances are it will not be detected. As a result, sophisticated “zero day” attacks – so called because they occur before the first day on which security personnel become aware of them and begin to develop countermeasures – may well go unnoticed.\footnote{Rosenzweig, supra note 14, at 28 n.23; Zetter, supra note 46.} Former CIA director Jim Woolsey emphasizes that “[i]f you can’t deal with a zero-day attack coming from a thumb drive, you have nothing.”\footnote{Quoted in McAfee 2011, supra note 15, at 1.} Of course, these are the very sorts of attacks likely to be launched by sophisticated adversaries like foreign intelligence services. Public health law’s biosurveillance framework thus is probably better at detecting intrusions of low to modest complexity than those undertaken by foreign governments.

The challenge, then, is to provide firms with incentives to collect and disseminate information about cyber vulnerabilities, threats, and countermeasures.\footnote{Nojeim, supra note 14, at 128.} At present companies have strong disincentives to do so, partly due to fears of legal liability,\footnote{See supra notes 148 to 150, 194 to 201 and accompanying text.} but also because of concerns about compromising trade secrets, losing customer goodwill, reputational harms, and so on.\footnote{ABA, supra note 18, at 10; Aviram, supra note 76, at 154; Aviram & Tor, supra note 131, at 240; Bambauer, supra note 12, at 611; Todd Brown, supra note 14, at 232; Coldebbala & White, supra note 14, at 236; Frye, supra note 144, at 369; Katyal, Digital Architecture, supra note 15, at 2278; Nojeim, supra note 14, at 135; Powell, supra note 14, at 501; Rosenzweig, supra note 14, at 9; Sarnikar & Johnsen, supra note 19, at 19; Schwartz & Janger, supra note 59, at 931; Smith, supra note 17, at 172-73. But see O’Neill, supra note 19, at 281.} Public health law facilitates collection and sharing through direct regulation, such as state statutes requiring health care providers to notify authorities about patients who have contracted various infectious diseases.\footnote{See supra notes 228 to 235 and accompanying text.} A similar arrangement might be adopted for cyberspace. The government could require firms to gather information about the vulnerabilities in their systems, the types of attacks they have suffered, and the countermeasures they have used to combat malware, and then to disseminate the data to designated recipients.\footnote{Frye, supra note 144, at 370-71.} (Imposing such an obligation would not eliminate companies’ incentives to withhold cybersecurity data. It would simply make it more costly for them to do so, where cost is equal to the sanctions for hoarding discounted by the probability of punishment. Increased costs mean that firms are more likely to collect and share cybersecurity data, but some will still find it advantageous to hoard.) There is also a less coercive, and probably more effective, alternative. Cybersecurity data is a sort of public good, and economic theory therefore predicts that it will be underproduced.\footnote{See supra notes 129 to 131 and accompanying text. But see Aviram & Tor, supra note 131, at 234-35 (arguing that information can be a rivalrous good).} One way to encourage the provision of public goods is to subsidize them, so firms might be offered bounties to compile and exchange the needed information.\footnote{Nojeim, supra note 14, at 128.} These bounties could be direct payments from the government or, more probably, tax credits or deductions. The subsidies also could take the form of enhanced intellectual property protections for the cybersecurity information firms generate. If the subsidies are large enough, firms will have an incentive not
just to report the data they have already compiled, but to invest in discovering previously unknown vulnerabilities, threats, and countermeasures.  \(^{285}\)

Antitrust law also can help recalibrate firms’ incentives.  \(^{286}\) Antitrust is often skeptical of information sharing and other forms of cooperation among companies.  \(^{287}\) But in the cybersecurity context, such arrangements can enhance consumer welfare: Agreements to exchange vulnerability, threat, and countermeasure information can help prevent cyberattacks from taking place (or at least mitigate their effects) and thereby minimize the resulting consumer losses.  \(^{288}\) One way to incentivize companies to cooperate is to alleviate their apparently widespread fears of antitrust liability. This could be accomplished through judicial, administrative, or legislative action. Federal courts could expressly discard the per se approach and substitute a rule of reason when reviewing private sector agreements to share cybersecurity data or to adopt common security protocols. This doctrinal shift would reduce the risk of false positives – i.e., the danger that the coarse grained per se rule might invalidate a cybersecurity initiative that is actually welfare enhancing. Instead, arrangements would be judged on a case by case basis, and would stand or fall based on the degree to which they actually advance or hinder consumer welfare. While this approach shows promise, it also has some significant drawbacks. The judicial route may not do enough to remove legal uncertainty. At the time companies are deciding whether to enter a cybersecurity venture, it will not always be possible to predict whether reviewing courts would sustain or invalidate it. A wrong guess could have severe consequences: Firms that are found to have violated federal antitrust law must pay treble damages to the successful plaintiffs.  \(^{289}\) This continued risk of legal liability will dissuade companies from entering arrangements that ultimately may withstand judicial scrutiny but, then again, possibly might not. In short, the uncertain prospects of ex post judicial approval may not provide firms with enough assurance ex ante.

A more promising approach would be for administrative agencies to sponsor cybersecurity exchanges, as some in Congress have proposed.  \(^{290}\) Agencies with special expertise in cybersecurity (such as the NSA and DHS) could partner with the agencies that are responsible for enforcing federal antitrust laws (the Federal Trade Commission and the Justice Department’s antitrust division) to establish fora in which companies could establish common security standards and exchange information. The government’s participation in these fora would offer assurances that they are being used for legitimate purposes and not as vehicles for anticompetitive conduct. From the standpoint of participating firms, this approach is advantageous because it offers them de facto antitrust immunity.  \(^{291}\) It is unlikely that an FTC or DOJ that sponsored a cooperative cybersecurity arrangement later would go to court to have it

\(^{285}\) But see Malloy, supra note 83, at 572-73 (predicting that firms will tend to neglect “regulatory investments” – i.e., expending scarce resources to obtain benefits offered to those who comply with government regulations).

\(^{286}\) Cf. Adler, supra note 147, at __.

\(^{287}\) See supra notes 136 to 141 and accompanying text.

\(^{288}\) See supra notes 144 to 146 and accompanying text.

\(^{289}\) 15 U.S.C. § 15(a); id. § 15a; id. § 15c(a)(2).


\(^{291}\) BRENNER, supra note 1, at 228.
invalidated. And while the blessing of these agencies does not formally bind other potential plaintiffs, such as state attorneys general or private parties, their determination that a proposed venture is permissible under federal antitrust laws probably would receive a healthy dose of judicial deference. Government sponsorship has another advantage: It can help solve the coordination and free rider problems associated with collective action. Left to their own devices, companies would prefer for their competitors to bear the expense of implementing new security standards and then free ride on the resulting security gains; they also would prefer to gain a competitive advantage by using other firms’ information and then withholding their own data from their competitors. A regulator can mitigate these tendencies. It can coerce firms into participating in the forum and complying with its requirements; it also can withhold the forum’s benefits from firms that shirk.

A third measure would be for Congress to enact a cybersecurity exception to the antitrust laws. The upside of a legislative carve out is that it would eliminate virtually all risk of liability and thus remove one powerful disincentive for companies to cooperate on cybersecurity initiatives. Ideally, such a measure would be narrowly tailored to the precise sort of inter-firm cooperation that is desired – the exchange of vulnerability, threat, and countermeasure information and the development of common security protocols. In other words, the exemption would be pegged to specific conduct, and would not extend antitrust immunity to entire industries (as used to be the case with major league baseball). A broader exception would offer few additional cybersecurity gains and could open the door to anticompetitive conduct.

We also might consult products liability law for ideas on how to incentivize companies to exchange cybersecurity data. Firms have weaker incentives to search for vulnerabilities in the systems they operate or the products they offer, because they typically cannot be sued by those injured by any resulting breaches. At the same time, other companies, such as ISPs, are reluctant to monitor network traffic for malicious code because of fears that doing so could expose them to legal liability. Adjusting the liability rules could help recalibrate these incentives. Lawmakers might use a combination of carrots and sticks. Offers of immunity would increase companies’ expected benefits of compiling and sharing cybersecurity data; threats of liability would increase their expected costs of failing to do so.

Consider the carrots first. Firms could be offered immunity from various laws that presently inhibit them from collecting and exchanging certain information about cyber vulnerabilities and threats. One way to accomplish this would be for Congress to expand the

---

292 Kobayashi, supra note 33, at 23; Sarnikar & Johnsen, supra note 19, at 22-23.
293 Katyal, Community, supra note 127, at 52.
295 See supra notes 185 to 193 and accompanying text.
296 See supra notes 194 to 201 and accompanying text.
297 Sarnikar & Johnsen, supra note 19, at 22.
298 Malloy, supra note 83, at 531-32. But see id. at 573 (predicting that firms will tend to neglect “regulatory investments” – i.e., complying with regulations to receive the benefits they offer).
service provider exception to the federal wiretap act’s general ban on intercepting electronic communications.299 The exception could be broadened to authorize ISPs to monitor network traffic for malicious code that threatens their subscribers’ systems, not just their own systems. Congress also could authorize (or perhaps even require) ISPs to notify customers whose systems are found to be infected by malware.300 It further could expressly preempt any state laws to the contrary. This would foreclose any claims that monitoring for malware violates a given state’s privacy law or breaches the terms of service between an ISP and its subscribers. In all cases, eligibility for these forms of immunity could be conditioned on information sharing: A company would not be able to take advantage of the safe harbor unless it shared the information it discovered with other firms. The result would be to create strong incentives to exchange data about threats and vulnerabilities.

As for the sticks, below I propose modifying tort law’s traditional economic loss doctrine (under which a defendant generally is not liable for freestanding economic harms, only economic harms that result from physical injuries) in the cybersecurity context.301 Firms that implement approved industry wide security standards would enjoy immunity from lawsuits seeking redress for injuries sustained from an intrusion; companies that disregard the protocols would be subject to lawsuits for any resulting damages. Under such a scheme, a company that has implemented the standards might have its immunity stripped if it fails to share information about known weaknesses in its systems or products. As for firms that fail to adopt the security standards, the lack of information sharing could be treated as an aggravating factor; extra damages could be imposed on firms that are aware of vulnerabilities or threats but fail to share that information with other companies. This series of tiered penalties would result in marginal deterrence; firms would have good reason not only to implement the approved security standards, but to exchange the threat and vulnerability information on which those protocols depend.

B. Hardening Targets

A second objective for a cybersecurity regime is to harden critical systems against attack by developing effective security protocols.302 The goal of such measures is to prevent cyber intruders from infecting these systems at all, as opposed to limiting the amount of damage intrusions can do; the objective is to increase the impregnability of critical systems, as opposed to their survivability.303 (Of course, some cyberattacks inevitably will succeed, so enhancing survivability – discussed below304 – is an essential goal as well.) The regulatory disciplines surveyed above suggest various techniques for encouraging companies to adequately secure their networks. Environmental law suggests the need for industry wide security standards; these rules should be developed through collaborative partnerships between regulatory agencies and private firms, rather than imposed via direct regulation. Products liability law suggests that pairing

300 BRENNER, supra note 1, at 229-31; CLARKE & KNAKE, supra note 1, at 164-65.
301 See infra notes 332 to 338 and accompanying text.
302 CLARKE & KNAKE, supra note 1, at 159.
303 See supra note 253 and accompanying text.
304 See infra Part III.C.
threats of liability with offers of immunity can incentivize firms to implement the security standards. And public health law’s use of mandatory vaccinations might be adapted by incentivizing firms to take certain minimum steps to secure their systems. Again, different firms and industries face different vulnerabilities, threats, and consequences, so the resulting security standards should be calibrated to the particular conditions in individual industries.

Regulators could improve critical systems’ defenses by establishing and enforcing new cybersecurity protocols, akin to the environmental regulations that restrict, say, the amount of sulfur dioxide a given source may emit into the atmosphere. A cyberattack on critical infrastructure will not just harm the targeted company, it also will impose negative externalities on a number of remote third parties. It is usually impossible to internalize the resulting costs through market exchanges, because the transaction costs would be prohibitive. Regulatory standards can help manage these spillovers. It should be emphasized at the outset that the specific content of any cybersecurity standards is well beyond the scope of this article. My focus here is not on the technical feasibility or policy advantages of any particular defensive measure. Instead, the focus of this article is establishing regulatory mechanisms by which new cybersecurity standards – whatever their content – may be adopted.

Turning to that question, one obvious option would be for administrative agencies to use traditional “command and control” regulation – i.e., issue a set of mandatory standards and incentivize firms to comply with them by threatening civil or criminal penalties. This is a fairly common approach in environmental law, and some scholars have urged the government to adopt it here. Neal Katyal argues that “direct government regulation” of cybersecurity “is the best solution,” and calls for regulatory agencies to issue “the equivalent of building codes to

---

305 It is also possible to develop new cybersecurity standards through litigation. Harper, supra note 119, at 2; Johnson, supra note 210, at 275-76; Rosenzweig, supra note 14, at 23. A court might hold, for instance, that a given firm’s failure to adopt a particular security measure breaches a general duty of care. This option seems less promising than the regulatory approach for several reasons. First, courts may not have the technical expertise to fashion detailed security protocols for complicated systems and products. Second, there is the problem of legal uncertainty. A regulation is likely to be more comprehensive than a series of incremental judicial opinions, especially in the context of a highly complex subject matter like cybersecurity, and relying on litigation thus runs the risk that firms will not know what is expected of them. There is, of course, an important role for litigation – the prospect of civil liability creates incentives for firms to comply with the regulatory standards. See infra notes 332 to 344 and accompanying text. But litigation should be limited to enforcing the standards, not formulating them in the first place.

306 Just within the law review literature – to say nothing of computer science, economics, and other fields – authors have debated relatively modest regulations such as mandating that firms use encryption, firewalls, and intrusion detection systems, Condron, supra note 19, at 410; Gable, supra note 2, at 94-95, requiring companies that operate certain sensitive systems to authenticate users before granting them access, Nojeim, supra note 14, at 131-33; Sklerov, supra note 19, at 22-24, and disconnecting vulnerable SCADA systems from the internet, CLARKE & KNAKE, supra note 1, at ___; McAfee 2010, supra note 35, at 34. Others have debated even more dramatic proposals, such as requiring ISPs to monitor the traffic that flows over their networks for malicious code, Lichtman & Posner, supra note 59, at 222; Katyal, Criminal Law, supra note 10, at 1007, 1095-1101; Taipale, supra note 92, at 34, or moving to an entirely new internet architecture (such as IPv6) in which anonymity is reduced and user activity is capable of being traced, Bambauer, supra note 12, at 590, 601; BAKER, supra note 24, at 231-32; Frye, supra note 144, at 354; Katyal, Digital Architecture, supra note 15, at 2269-70; LESSIG, supra note 67, at 45, 54; POST, supra note 90, at 84; Taipale, supra note 92, at 31.

307 Malloy, supra note 83, at 531.
require proper design and performance standards for software.”

Likewise, a prominent think tank argues that “the federal government bears primary responsibility” for cybersecurity and that “it is completely inadequate” to leave the matter “to the private sector and the market.” Some have even called for the federal government to take over certain sectors of the economy in the name of cybersecurity. According to an ABA task force, “government may also need to ‘semi-nationalize’ some sectors (like the electricity grid) where isolation is not an option and the adverse consequences of certain low probability events are likely to be very high.” It isn’t steel mills, but Harry Truman would have admired the proposal.

Traditional command and control regulation seems ill suited to the task of securing the nation’s critical cyber infrastructure. The better course would be to involve the firms that operate these assets in establishing and implementing new security protocols. Private sector participation – an approach sometimes seen in environmental law – is desirable for several familiar reasons. First, information asymmetries: Companies often (though not always) know more than regulators about the vulnerabilities in their systems, the types of intrusions they have faced, and the most effective countermeasures for dealing with those threats. A related concern is that regulators probably lack the knowledge necessary to determine the socially optimal level of cyber breaches and set the security standards accordingly. The market, through the price system, is capable of aggregating and processing this information in a way that central planners cannot. Third, rapid technological change makes it difficult for regulators to formulate durable security rules. Vulnerabilities, threats, and countermeasures are in a constant state of flux, and regulatory standards cannot keep pace with these developments. Notice and comment rulemaking rarely takes less than 18 months, sometimes much longer, and the rules likely would be obsolete before the ink in the Federal Register was dry. Fourth, there is a risk that government protocols will stifle innovation. If regulatory agencies promulgate a set of mandatory standards, regulated firms will have less reason to search for newer and more efficient countermeasures; they will simply implement the government’s directives.

What specific role should private firms have in developing and implementing cybersecurity standards? At least two possibilities come to mind. First, regulators could

---


309 CSIS, *supra* note 8, at 15; *see also* Frye, *supra* note 144, at __.

310 ABA, *supra* note 18, at 27.


312 *See supra* notes 263 to 265 and accompanying text.

313 *Coyne & Leeson, supra* note 18, at 488-89; *Kobayashi, supra* note 33, at 26-27; *Powell, supra* note 14, at 502, 505.

314 *BAKER, supra* note 24, at 235, 237; *CSIS, supra* note 8, at 51; *Rosenzweig, supra* note 14, at 10.

315

316 CSIS, *supra* note 8, at 51; *Kobayashi, supra* note 33, at 26.

317 *See also* Kobayashi, *supra* note 33, at 27 (calling for the creation of private “security cooperatives” as an “alternative to government standards”). *But see* Aviram, *supra* note 76, at 150, 182 (arguing that private security standard setting is unlikely to be effective due to high enforcement costs).
practice a form of “delegated regulation” in which they mandate broad security goals and establish the penalties for falling short, then leave it up to companies to achieve those goals in whatever manner they deem most effective. Regulation by delegation is said to be appropriate where administrative agencies have the capacity to “identify specific outcomes but cannot easily codify in generally-applicable rules the means for achieving them.” It is also desirable where, as is often the case, private firms have a comparative advantage at developing innovative and efficient ways to achieve regulatory goals. Environmental law sometimes follows this approach (as do other fields such as food safety and securities regulation). For instance, the EPA’s “bubble” approach to the Clean Air Act allowed polluters to offset increased emissions from one source with decreased emissions from other sources, providing them with an incentive to experiment with new technologies that could reduce emissions at lower cost. Delegated regulation seems a good fit for cybersecurity, though not perfect one. Giving companies discretion to implement the government’s security standards achieves three of the four benefits of private action mentioned above: It avoids (some) problems with information asymmetries, allows for flexibility in reacting to fast changing technologies, and promotes rather than stifles private sector innovation. However, difficulties would remain with formulating the standards. As is true of much command and control regulation, agencies lack the knowledge needed to determine the socially optimal level of cyber breaches and set the security standards accordingly.

An alternative would be a form of “enforced self-regulation” in which private companies develop the new cybersecurity protocols in tandem with the government. These requirements would not be handed down by administrative agencies, but rather would be developed through a collaborative partnership in which both regulators and regulated would play a role. In particular, firms might prepare sets of industrywide security standards. (The National Industrial Recovery Act, famously invalidated by the Supreme Court in 1935, contained such a mechanism, and today the energy sector develops reliability standards in the same way.) Or agencies could sponsor something like a negotiated rulemaking in which regulators, firms, and

---

319 Bamberger, supra note 203, at 380; see also ABA, supra note 18, at 9; CLARKE & KNAKE, supra note 1, at 134; Jensen, Cyber Warfare, supra note 15, at 1565.
320 Bamberger, supra note 203, at 389.
322 Bamberger, supra note 203, at 390-91.
323 Chevron U.S.A. Inc. v. Natural Resources Defense Council, Inc., 467 U.S. 837 (1984); Malloy, supra note 83, at 536, 541; see also Malloy, supra note 83, at 547-49 (discussing conflicting accounts of whether bubble approach actually promoted innovation).
324 Bamberger, supra note 203, at 461 (citing IAN AYRES & JOHN BRAITHWAITE, RESPONSIVE REGULATION: TRANSCENDING THE DeregULATION DEBATE 101-32 (1995)).
325 ABA, supra note 18, at 9; Coldebella & White, supra note 14, at 241-42; Katyal, Criminal Law, supra note 10, at 1099.
327 CSIS, supra note 8, at 52-53.
other stakeholders forge a consensus on new security protocols. 328 In either case, agencies then would ensure compliance through standard administrative techniques like audits, investigations, and enforcement actions. 329 This approach would achieve all four of the benefits of private action mentioned above: It avoids (some) problems with information asymmetries, takes advantage of distributed private sector knowledge about vulnerabilities and threats, accommodates rapid technological change, and promotes innovation. On the other hand, allowing firms to help set the standards that will be enforced against them may increase the risk of regulatory capture – the danger that agencies will come to promote the interests of the companies they regulate instead of the public’s interests. 330 The risk of capture is always present in regulatory action, but it is probably even more acute when regulated entities are expressly invited to the decisionmaking table. 331

Products liability law likewise offers several strategies for hardening critical infrastructure against cyberattacks. The prospect that a company might be required to pay money damages to those who have been injured by an attack on the systems they operate or the products they offer would internalize costs that are now externalized onto others. Liability thus would incentivize firms to offer goods (such as computer software) and services (such as online banking) that are more secure. 332 At present, companies face little risk of liability for the injuries that result from their failure to prevent cyber intrusions, because the economic loss doctrine generally limits them to paying damages for physical injuries and associated economic harms, not for freestanding economic injuries. 333 Modifying this default rule of de facto immunity could help foster incentives for firms to improve their cyberdefenses.

What could a recalibrated liability regime for cybersecurity look like? Again, a combination of carrots and sticks could be used. Congress might abolish the economic loss doctrine for injuries that result from a given company’s (wrongful) failure to prevent a cyberattack. In its place, lawmakers could substitute a regime that imposes liability or offers immunity based on what steps a company has taken to secure its products or systems. As for the

---

329 CSIS, supra note 8, at 52.
330 George Stigler, The Theory of Economic Regulation, 2 BELL J. ECON. & MANAGEMENT SCI. 3 (1971). A related problem is that, because of information asymmetries, agencies often depend on the companies they regulate to provide the data they need to formulate rules. Yet firms will have an incentive to underestimate vulnerabilities and threats to persuade regulators to approve lenient (and less costly) security protocols. Coyne & Leeson, supra note 18, at 489. (Of course, that concern is also present in traditional regulation.) There are also doctrinal difficulties. Depending on how the public-private partnership is structured, it conceivably could violate what remains of the nondelegation doctrine. See, e.g., Carter v. Carter Coal Co., 298 U.S. 238 (1936) (striking down a statute that authorized coal producers to establish minimum prices in certain geographic regions on the ground that it was an unconstitutional delegation of legislative power to private companies).
331 USA Group Loan Servs., Inc. v. Riley, 82 F.3d 708, __ (7th Cir. 1996) (Posner, C.J.) (describing negotiated rulemaking as “an abdication of regulatory authority to the regulated, the full burgeoning of the interest-group state, and the final confirmation of the ‘capture’ theory of administrative regulation”).
332 Coyne & Leeson, supra note 18, at 492; Lichtman & Posner, supra note 59, at 232-39; Hunker, supra note 19, at 211; Johnson, supra note 210, at 260; Rosenzweig, supra note 14, at 23; Schnieier, supra note 33, at 2; Yang & Hoffstadt, supra note 15, at 207-10.
333 See supra notes 185 to 187 and accompanying text.
carrots, firms that implement the security standards that are developed in tandem with regulators, but nevertheless suffer cyberattacks, could be offered immunity from lawsuits seeking redress for the resulting damages. This cyber safe harbor could extend not just to purely economic injuries (for which firms currently enjoy de facto immunity) but also to physical injuries and the associated economic harms (for which firms presently may be held liable). The scope of immunity thus would be broader than under current law, but it would only be available to companies that take the desired steps to improve their cyberdefenses. (Lawmakers might use the SAFETY Act as a model. The Support Anti-terrorism by Fostering Effective Technologies Act of 2002 grants immunity to firms that sell certain anti-terrorism goods and services, so long as they comply with various standards, including a requirement that they carry liability insurance.)

As for the sticks, firms that fail to implement the agreed security measures and then suffer cyberattacks could be exposed to liability for the full range of injuries that result from the intrusions. The severity of the damages could be pegged to the severity of their misconduct, thereby achieving marginal deterrence. So, for instance, a company that fails to adopt the approved security standards might be made to pay compensatory damages or even a smaller fixed sum set by statute, but a company whose conduct is more egregious – one that fails to share information about known vulnerabilities or threats, for instance – might be eligible for exemplary damages. (For inspiration, lawmakers might look to the Gramm-Leach-Bliley Act, which imposes liability on banks that fail to protect consumer data, resulting in relatively robust defenses in the financial services sector against cyber intrusions.) Such a liability regime would increase both a firm’s expected benefits of implementing the security protocols that are developed in tandem with regulators, as well as the expected costs of defying them.

Civil liability also would help promote a more robust market for cybersecurity insurance. Insurers can have a profound effect on the steps firms take to secure their systems and products against cyber intrusions, because they can insist that companies implement various security measures as a condition of coverage or charge higher premiums of those that do not. In effect, insurance companies provide a sort of second order regulation, enforcing cybersecurity standards by refusing to bear the losses of firms with poor records or engaging in price discrimination against them. The result is to provide insureds with financial incentives to implement the defenses their insurers are calling for. These incentives have already borne fruit. According to Bruce Schneier, “[f]irewalls are ubiquitous because auditors started demanding firewalls. This changed the cost equation for businesses. The cost of adding a firewall was expense and user annoyance, but the cost of not having a firewall was failing an audit.” Enforcement by

334 Coldebella & White, supra note 14, at 235.
335 BAKER, supra note 24, at 234-35.
336 6 U.S.C. §§ 441-44.
337 See supra note 301 and accompanying text.
338 15 U.S.C. § 6801(b); see supra notes 202 to 210 and accompanying text.
339 Bamberger, supra note 203, at 456; BRENNER, supra note 1, at 225; Coyne & Leeson, supra note 18, at 491-92; Rosenzweig, supra note 14, at 23-24.
340 Schneier, supra note 33, at 1.
insurers also can decrease the government’s enforcement costs; there is less need for regulators to verify that firms are complying with the agreed security standards if insurers, pursuing their own financial interests, are already doing so.

At present, the market for cybersecurity insurance is fairly underdeveloped (though some insurance companies have begun to offer coverage\(^{341}\)). This is so in large part because firms currently face very little risk of liability for injuries resulting from cyberattacks on their systems or products; why insure when one is effectively immune?\(^ {342}\) The prospect of civil liability is a critical first step in creating a viable market for cybersecurity insurance.\(^ {343}\) Lawmakers might further stimulate the market by offering various kinds of subsidies. For instance, the government might provide insurers with more information (including, perhaps, classified information) about the incidence, prevalence, and consequences of various sorts of malicious code. Insurers could use this data to assess more accurately the probability of cyber intrusions and their potential costs, which would help in setting insurance premiums.\(^ {344}\) Or the government might offer tax benefits to insurers that offer cybersecurity policies. Or it might require certain companies (such as strategically important firms like public utilities or companies that supply goods or services to the government) to carry cybersecurity insurance.

Public health law suggests a final approach to hardening critical infrastructure against cyberattacks. Most states have enacted laws requiring schoolchildren to be vaccinated against various diseases, and lawmakers might adopt similar measures for cyberspace. In both contexts, compulsory inoculation helps reduce negative externalities. In the same way that an unvaccinated child might infect his classmates with a pathogen, a computer system that lacks effective cyberdefenses might be commandeered into a botnet. Compulsory inoculation also helps create positive externalities. A child who has been vaccinated contributes to herd immunity and thereby decreases the probability that other, unvaccinated students will contract the disease. In the same way, companies that adopt effective cyberdefenses make it less likely that their systems will be used to transmit malware to other users.

What would mandatory vaccination look like in cyberspace? Several variants exist. The most coercive approaches involve direct regulation, akin to a requirement that all citizens receive a particular vaccine. One option would be for lawmakers to mandate that every computer user (or, less dramatically, firms in particularly sensitive industries such as the telecommunications sector) install certain security products on their systems (such as antivirus software or firewalls). Think of it as a digital equivalent of the Patient Protection and Affordable Care Act’s “individual mandate” to purchase health insurance.\(^ {345}\) An alternative would be for the government to require ISPs to provide their customers with a specified security software package.\(^ {346}\) ISPs presumably

\(^{341}\) Coyne & Leeson, supra note 18, at 491; Yang & Hoffstadt, supra note 15, at 208-09.

\(^{342}\) ABA, supra note 18, at 8; BRENNER, supra note 1, at 225. Another challenge is that it is difficult for insurers to write policies when – as is often the case with cyberattacks – the probability and consequences of an incident are uncertain.

\(^{343}\) Rosenzweig, supra note 14, at 23.

\(^{344}\) Coyne & Leeson, supra note 18, at 491-92; Frye, supra note 144, at 366-67.

\(^{345}\) 26 U.S.C. § 5000A(a).

\(^{346}\) CLARKE & KNAKE, supra note 1, at 165; Sharp, supra note 8, at 25.
would pass on the costs of the software to their subscribers, so the effect would be the same as
the individual mandate approach – users would be made to pay a premium for a security product
they previously declined to purchase. Or, the government could compensate the ISPs for the
costs of making the security package available to their subscribers. In that event, the scheme
would represent a (likely regressive) wealth transfer from taxpayers who do not use computers to
those who do.

Another, less coercive, set of options would withhold or offer certain benefits to
incentivize security improvements; they are the equivalent of making vaccination a condition of
eligibility to attend public schools. The ability to access the internet (as opposed to local or
proprietary networks) is a valuable benefit of the service one receives from an ISP – for many
subscribers it is the most valuable benefit ISPs offer – and it might be conditioned on a
subscriber taking steps to improve cyber security. In particular, regulators could direct ISPs to
refuse to route users’ traffic to the public internet unless they are able to verify that the users
have installed specified security software on their systems. Alternatively, government web
sites could refuse any traffic sent from a system that has not adopted specified security measures.
Users thus would be unable to, for example, post comments in an online rulemaking docket or
check the status of a tax refund unless they adopted the security measures. (This sort of measure
depends on the ability to authenticate the identity of the sender, as well as the presence of various
cyberdefenses on its system. That capability does not presently exist, because the TCP/IP
routing protocol is unconcerned with the sender’s identity, though some scholars believe an
authenticated internet is inevitable.) Finally, the government could offer tax credits or
deductions to firms (or individual users) that install the specified security software on their
systems – another (likely regressive) wealth transfer.

C. Survivability and Recovery

The third thing an ideal cybersecurity regime would do is promote resilience, thus
limiting the amount of damage attackers can do to critical infrastructure. Here, the goals are
survivability and recovery, not impregnability. The need to build resilience into the nation’s
cyberdefenses is based on the hard reality that, no matter how good one’s defenses are, some
attackers will be able to breach them. As a result, it is not enough to try to prevent attacks
altogether. It is also necessary to minimize the amount of harm that the inevitably successful
intrusions can do, and to restore victims to the status quo ante as quickly as possible. Public
health law offers several strategies for improving resilience. Quarantine and isolation laws,
which help limit the spread of infectious diseases during outbreaks, might be adapted for
cyberspace. In the event of a cyberattack, systems that become infected with malware might be
temporarily disconnected from the internet. Or certain especially sensitive systems (such as the
power grid or financial networks) that have not been infected might nevertheless be isolated as a
preventive measure. Finally, firms might undertake the cyberspace equivalent of stockpiling

347 Rattray et al., supra note 8, at 160.
348 See supra notes 112 to 113 and accompanying text.
349 BAKER, supra note 24, at 231-32; LESSIG, supra note 67, at 45.
350 See supra note 253 and accompanying text.
vaccines and medicines – they might build excess capacity into their systems that can be called into service in emergencies.

In realspace, quarantine and isolation aim at minimizing the harm a pathogen can do; once an outbreak is underway, we want to contain the disease and limit the number of people to whom it can spread. In other words, the objective is to regulate negative externalities. Quarantine and isolation might be adapted for cyberspace – where the goal is to prevent malicious code from infecting more machines – in any number of ways. The most straightforward approach would be for authorities, in the event of a cyberattack, to order systems that are known or suspected to be infected with malware to temporarily disconnect from the internet. While in quarantine, the systems could be inspected to see if they are in fact carrying malicious code. If not, they could be reconnected; if so they could be repaired. The analogy to public health law is fairly exact: Separation of the infected, whether physical or virtual, prevents them spreading the contagion to others and presents an opportunity for treatment. While potentially effective, this approach has a significant drawback – legitimate users will be unable to access the infected system while it is offline. Putting a bank into cyber quarantine doesn’t just keep hackers from stealing money, it also keeps a customer from logging on to pay his credit card bill. A less drastic way of preventing the spread of malware would be to isolate traffic rather than systems. Infected systems would remain connected to the internet, but authorities could use (or require firms to use) deep packet inspection to determine if the data the systems are sending and receiving contains malware. If a given packet is found to be carrying malicious code, it could be interdicted; if not, it would be allowed to continue on its way. The public health analogy is allowing a man infected with SARS to leave an isolation facility and go about his business, while wearing a surgical mask that intercepts the respiratory droplets through which the virus is spread. The virtue of this finer grained variant is that it allows legitimate users to continue to access an infected system even as attackers are prevented from using it for their malign purposes; the hackers are thwarted, but customers can still access their accounts (although perhaps a bit more slowly than usual). On the other hand, traffic quarantines will only be as effective as the packet sniffers and malware signature files on which they rely, and sophisticated adversaries might be able to defeat both.

Another, more controversial set of options involves preventive quarantine – separating systems that have not been infected but that are vulnerable. This approach would turn public health law on its head; rather than isolating the sick, authorities would isolate the healthy. The most aggressive variant would be for regulators to require a select group of strategically significant firms, such as the power grid, financial institutions, telecommunications carriers, to temporarily disconnect from the internet if a cyberattack takes place.351 (Senator Nelson Rockefeller introduced legislation along these lines in 2009352; critics denounced it as an “internet kill switch.”353) Preventive quarantine would be a fairly effective way of preventing malware from spreading to critical infrastructure – a system that isn’t on the internet cannot contract a virus that spreads via the internet. But it wouldn’t be infallible. Even “air gapped”

351 Brenner, supra note 1, at 234; Clarke & Knaake, supra note 1, at 167; Picker, supra note 42, at 126.
systems – those that are physically separated from the internet\textsuperscript{354} – are vulnerable to infection via USB devices and other removable media.\textsuperscript{355} A disconnection requirement also could prove quite costly: The affected systems would be unavailable to legitimate users for as long as the order remained in effect. There is also a risk that regulators might pull the disconnection trigger too readily. As an alternative to a strict disconnection requirement, regulators might direct firms to implement security countermeasures of their own devising. (Senator Joseph Lieberman introduced legislation along these lines in 2010\textsuperscript{356}; it likewise was denounced as an internet kill switch.\textsuperscript{357}) Whatever the content of these security protocols – encrypting data to prevent its theft, for instance, or requiring users to authenticate themselves before gaining access to the system – they might be established through the collaborative regulatory partnership described above.\textsuperscript{358} An even more modest version of preventive quarantine would be, as above, to segregate traffic rather than entire systems. In the event of a cyberattack, packet sniffers might be used to inspect all traffic that is sent to and from designated systems. This would allow the systems to continue to operate more or less as usual, though perhaps at a cost of less security.

Another important goal is to ensure that critical systems are able to continue functioning during a cyberattack and recover quickly thereafter. One way to achieve this is to build systems with excess capacity – i.e., to include more capabilities than a firm needs for its day to day operations, but that can be held in reserve and called into service if an attack takes place.\textsuperscript{359} In particular, regulators might require certain companies to build their systems with excess bandwidth. A “strategic reserve of bandwidth” is an especially useful countermeasure for defending against denial of service attacks\textsuperscript{360}; if a company’s servers are being overwhelmed, the reserve bandwidth can be brought into service to process the requests. Regulators also might require certain companies to maintain redundant data storage capabilities. These firms might routinely back up their data to servers that are dispersed, both geographically and in network terms. If a cyberattack corrupted their systems, it would be relatively easy to wipe them clean and restore the data from an uncorrupted backup.\textsuperscript{361} An attacker thus might succeed in taking down one site “only to find that the same content continues to appear through other servers. This is like playing electronic Whac-A-Mole on a global scale . . . .”\textsuperscript{362} These sorts of measures can be thought of as akin to the public health practice of stockpiling medicines and vaccines for use in a crisis. The CDC may not need 300 million doses of smallpox vaccine in its everyday operations, but they would prove critical in the event of an outbreak.

\begin{itemize}
\item \textsuperscript{354} \textit{Brenner, supra} note 1, at 84; Ellen Nakashima, \textit{Cyber-Intruder Sparks Massive Federal Response}, \textit{Wash. Post}, Dec. 8, 2011.
\item \textsuperscript{355} \textit{Baker, supra} note 24, at 216; Baker, \textit{supra} note 29, at 3; \textit{Brenner, supra} note 1, at 61; \textit{Clarke & Knaake, supra} note 1, at 127; Nakashima, \textit{Cyber-Intruder, supra} note 354.
\item \textsuperscript{356} Protecting Cyberspace as a National Asset Act of 2010, S. 3480, 111th Cong. (2010).
\item \textsuperscript{357} Adam Cohen, \textit{What’s Missing in the Internet Kill-Switch Debate}, \textit{Time}, Aug. 11, 2010.
\item \textsuperscript{358} See \textit{supra} Part III.B.
\item \textsuperscript{359} Benkler, \textit{supra} note 253, at 75.
\item \textsuperscript{360} Taipale, \textit{supra} note 92, at 37.
\item \textsuperscript{361} Bambauer, \textit{supra} note 12, at 637; Sarnikar & Johnsen, \textit{supra} note 19, at 2, 21; Taipale, \textit{supra} note 92, at 38.
\item \textsuperscript{362} \textit{Brenner, supra} note 1, at 179.
\end{itemize}
Excess capacity can be expensive; requiring firms to keep reserves of (largely unused) bandwidth costs money, and “[h]aving information located in multiple places makes it more costly to maintain.” How to pay for these requirements? One obvious option would be for companies to pass their costs of complying with resilience mandates to their customers in the form of price increases, service decreases, or both. A difficulty with this approach is that improving a given company’s ability to withstand an attack does not just confer benefits on its customers. It also confers benefits on those with whom the company has no relationship; if Citibank can continue to operate notwithstanding a DDOS, its customers will still be able to pay their bills, and third party vendors will still be able to receive payments. Excess capacity thus creates positive externalities, and the customers who pay higher prices for excess capacity are effectively subsidizing the third parties. Another option would be for the government to offer various subsidies to firms that are subject to survivability mandates. This approach is based on a recognition that excess capacity is, in a sense, a public good that the market will tend to undersupply.

In part because excess capacity requirements can be costly, it wouldn’t be advisable for regulators to apply them to all firms in the marketplace.

D. Responding to Cyberattacks

The fourth and final component of an effective cybersecurity regime is responding to individuals, groups, and states that have committed cyberattacks. This topic is exhaustively covered in the existing literature, and it naturally lends itself to analysis under the law enforcement and armed conflict approaches to cybersecurity. For instance, scholars have proposed better international cooperation on cybercrime investigations, increasing the penalties for certain computer related offenses, increasing the costs that perpetrators must bear to commit cybercrimes, treating intrusions as “armed attacks” that trigger the right to self defense under the United Nations Charter, treating cyberattacks as acts of aggression that justify retaliating with conventional military force, and so on. This article does not seek to add to this already voluminous literature. There is, however, one type of response that deserves a brief mention: active self defense, or “hackbacks.” Like the regulatory solutions discussed above, hackbacks can incentivize firms to better secure their systems against cyber intruders.

A hackback is an in-kind response to a cyberattack. The victim essentially mounts a counterattack against the assailant, “shutting down the attack before it can do further harm and/or damaging the perpetrator’s system to stop it from launching future attacks.” This might be accomplished in several ways. If a victim detects that it is experiencing a cyberattack, it might direct a flood of traffic to the servers through which the attack is being routed, temporarily overwhelming them and preventing them from continuing the intrusion. Or it might hack into

363 Bambauer, supra note 12, at 637.
364 See supra notes 129 to 135 and accompanying text.
365 See supra Part II.A.
366 Sklerov, supra note 19, at 25; see also Condron, supra note 19, at 410-11; O’Neill, supra note 19, at 280.
367 Condron, supra note 19, at 410-11.
the responsible servers, taking control of them or damaging them. Some scholars believe that hackbacks are the most effective defense against cyberattacks. This is so in part because active self defense can avoid the attribution problem; a victim firm that is experiencing an intrusion could retaliate against any computer that is attacking it without knowing who is behind the incident or his purposes. (Needless to say, active self defense is only possible if the victim is aware that it is under attack. Hackback will not be an option if, as is sometimes the case, the intrusion goes undetected.)

Active self defense fits into the law enforcement framework fairly comfortably. Although hackbacks are probably illegal under the Computer Fraud and Abuse Act – the victims are, after all, perpetrating cyber intrusions of their own – fundamental principles of criminal law can explain why they might be acceptable. The basic idea is justification. Conduct that ordinarily is condemned can become permissible, or even desirable, in certain circumstances. Homicide is typically illegal, as society deems it blameworthy to take the life of another human being, but we are allowed to use deadly force against those who pose a threat to our lives or the lives of others. The same might be said of hackbacks. Society ordinarily condemns those who break into others’ computers, but one might be justified in hacking an intruder’s machine to protect one’s own system. Hackbacks also might be described in armed conflict terms, though perhaps not as neatly. Scholars have extensively debated whether the law of armed conflict allows a government to engage in active self defense against those who have launched cyberattacks against its computer systems. A related question is whether the LOAC allows private individuals to respond in kind when their systems are attacked. Initially one might think the answer is no; hackbacks by civilians against foreign adversaries can be seen as a use of force by noncombatants in violation of the LOAC. However, the LOAC allows noncombatants to participate in hostilities in certain circumstances. *Levee en masse* refers to the right of a civilian population to resist invaders by force, provided that they organize themselves into units, carry arms openly, and otherwise obey the laws of war. Active self defense against

---

368 Smith, *supra* note 17, at 177-78.
376 Third Geneva Convention art. 4A(6); see also Watts, *supra* note 44, at 435. Cf. Rabkin & Rabkin, at 11-12 (analogizing private citizens who conduct cyber intrusions with a state’s blessing to privateers who operate under letters of marque).
cyberattacks might be thought of as a digital (though far less organized) version of the *levee en masse*. Just as civilians are entitled to use pistols and rifles to repel invading armies, they also might be entitled to use hackers’ tools to stop attacks by foreign governments or groups – especially when those attacks are directed at their systems.377

Active self defense is controversial, but it offers one potential benefit that has been largely overlooked in the literature. Like the other regulatory solutions discussed in this article, hackbacks can incentivize firms to improve the security of their systems. Cyber perpetrators typically do not launch attacks directly; to obscure their responsibility, they usually route an attack through a chain of unsecured intermediary systems before reaching the ultimate target.378 If a victim firm responds to an intrusion with active self defense, it is likely that these third party systems will suffer harm.379 (The realspace equivalent would be a driver who leaves his car unlocked; the car is then stolen by bank robbers and destroyed when the thieves open fire and the bank’s security guards shoot back.) Many scholars regard this third party problem as a sufficient reason to forbid hackbacks.380 Yet the prospect of damage to third parties may actually be desirable. The threat of harm provides firms with incentives to better secure their systems and prevent them from being used as conduits for attacks on others. Suppose Citibank knows that, if attackers gain control of its computers and use them to conduct DDOS attacks, the victims will be allowed to retaliate against Citibank’s machines. Citibank will have a fairly strong incentive to ensure that its computers are not commandeered into botnets. Damage from hackbacks thus would internalize some of the costs that third parties impose on others by maintaining insecure systems.381 (Likewise in realspace. If drivers know that security guards are allowed to damage getaway cars even if they are stolen, they will lock their doors.) Active self defense also might weaken attackers’ incentives to commit cyberattacks. If assailants know that victims will be able to use hackbacks to render their attacks ineffective, or less effective, they will have less reason to undertake them in the first place. By increasing the futility of intrusions, hackbacks can help achieve deterrence.382 Active self defense thus can simultaneously foster favorable incentives to improve security and weaken unfavorable incentives to commit attacks.383

377 But see Davis Brown, supra note 13, at 191-92.
378 See supra note 113 and accompanying text.
379 Epstein, supra note 369, at 31; Katyal at 62-63; Kerr, supra note 373, at 205; Smith, supra note 17, at 180.
382 O’Neill, supra note 19, at 280; Sklerov, supra note 19, at 10. See generally supra note 254 and accompanying text.
383 To be sure, active self defense might foster negative incentives. As Orin Kerr points out, allowing hackbacks “would create an obvious incentive for attackers to be extra careful to disguise their location or use someone else’s computer to launch the attack.” Kerr, supra note 373, at 205. Allowing hackbacks also would “encourage foul play designed to harness the new privileges”; one example is the “bankshot attack,” in which an assailant who wants a computer to be attacked “can route attacks through that one computer toward a series of victims, and then wait for the victims to attack back at that computer.” Kerr, supra note 373, at 205-06; see also Katyal, *Community*, supra note 127, at 63. It cannot be predicted a priori whether the harmful conduct produced by these negative incentives would be greater or lesser than the beneficial conduct produced by the positive incentives.
CONCLUSION

Cyberthreats aren’t going away. As society increasingly comes to rely on networked critical infrastructure such as banks and the power grid, assailants will find that they have ever more to gain by attacking these digital assets. And we will find that we have ever more to lose.

It therefore becomes essential to think about cybersecurity using an analytical framework that is rich enough to account for the problem in all its complexity. Cybersecurity is too important, and too intricate, to leave to the criminal law and the law of armed conflict. Instead, as this article has proposed, an entirely new conceptual approach is needed – an approach that can account for the systematic tendency of many private firms to underinvest in cyberdefense. Companies sometimes fail to secure their systems against attackers because they do not bear the full costs of the resulting intrusions; the harms are partially externalized onto third parties. Firms also tend to neglect cybersecurity because by improving their own defenses they contribute to the security of others’ systems; the benefits are partially externalized, which creates opportunities for free riding. If these problems sound familiar, that’s because they are. These challenges of negative externalities, positive externalities, and free riding are similar to challenges that the modern administrative state encounters in a number of other settings. Cybersecurity thus resembles the problems that arise in environmental law, antitrust law, products liability law, and public health law. Scholars and lawmakers might look to these other fields for suggestions on how to incentivize private firms to improve their defenses; conceiving of cybersecurity in regulatory terms opens the door to regulatory solutions.

Of course, “regulatory solutions” need not mean “command and control solutions.” Often it will be possible to promote better cybersecurity by appealing to firms’ self interest – encouraging them to improve their defenses by immunizing them from liability or offering other subsidies, not just sanctioning them when they fail to do so. For instance, rather than empowering a central regulator to monitor the internet for outbreaks of malicious code, companies should use something like public health law’s distributed biosurveillance network to collect and share information about cyberthreats with one another. Similarly, the private sector should play an active role in establishing industrywide cybersecurity standards, as it frequently does in environmental law and other regulatory contexts. Offers of immunity and threats of liability then would be used to encourage companies to adopt the agreed upon standards. As for improving the ability of critical systems to survive intrusions, infected computers could be temporarily disconnected from the internet to keep them from spreading the malware, and companies should be encouraged to build their systems with excess capacity (such as reserve bandwidth and remote backups) that can be called into service during cyberattacks. Finally, lawmakers might loosen the restrictions on “hackbacks,” to incentivize firms to protect their systems from being commandeered into attacks on third parties.

Virtually no one is happy with the state of America’s cyberdefenses, and scholars have felled entire forests exploring how to prosecute cyber criminals more effectively or retaliate against countries that launch cyberattacks. Maybe we’ve been asking the wrong questions. Maybe what we need to secure cyberspace isn’t cops, spies, or soldiers. Maybe what we need is administrative law.