Letters, Power Lines, and Other Dangerous Things

The Politics of Infrastructure Security

The title-level DOI for this work is:

doi:10.7551/mitpress/10541.001.0001

6

Regulating Cybersecurity: The Unexpected Remaking of Electric Power

Nothing seems to be updating on the computers. … We've had people calling and reporting trips and nothing seems to be updating in the event summary. … I think we've got something seriously sick.
—FirstEnergy transmission operator to IT department, August 14, 2003[1]

On August 14, 2003, the evening commute for millions of New Yorkers did not go as planned. Subways stopped running, elevators stalled in mid-descent, and traffic congestion made driving nearly impossible. The power was out. Thousands abandoned their typical commute and set out on foot, streaming across the Brooklyn Bridge on their way home. Others decided not to brave the hectic conditions and bunked wherever they could.[2] The largest blackout in the nation's history stretched across eight states and parts of Canada, leaving over 50 million people without power.[3]

The blackout was a classic normal accident. It was set in motion by a mix of mundane and decidedly unspectacular failures: an employee at the Midwest Independent System Operator (MISO) turned off MISO's state estimator, a computer system that monitors system contingencies, in order to troubleshoot an unrelated problem and then forgot to turn it back on. At around this time, Eastlake5, a power generator in northern Ohio, tripped due to operator error. This normally would not have been too significant. FirstEnergy should have been able to import enough power to cover the loss of Eastlake5. But in quick succession, falling trees snapped a number of high-voltage lines.[4] The MISO employee's failure to switch the state estimator back on and unrelated computer failures at FirstEnergy now took on much more significance. They made both the line outages and the failure of Eastlake5 difficult to account for, leading to further problems. The comedy of errors—each unremarkable on its own and largely unrelated—continued to compound. A series of seemingly small and local failures cascaded through the system, throwing transmission lines and

generators off-line. The risks of a tightly coupled and complex electric power system became stark: these apparently minor failures turned out to be fatal. The deregulation of electric power had replaced the staid utility system with vast networks of tightly integrated players—generators, transmission operators, distribution systems, marketing companies, and others—lashed together through technical and nontechnical linkages. The fragility of this new constellation now came rushing into view. A few tree branches fall, a computer system is accidentally turned off, an operator makes an unfortunate but not entirely unexpected mistake and then, suddenly, the lights go dark and 50 million people across the U.S. and Canada cannot turn them back on.

The outage dwarfed the 1977 New York blackout. It had significant regional—rather than local—impacts. Shortly after 4:00 p.m. on August 14, 2003, New York, Toronto, Cleveland, Detroit, and many other cities came face-to-face with life without electricity as a cascading power failure moved from Ohio to parts of the Northeast, to the Midwest, and to Ontario.[5] Speaking in San Diego, President Bush promised a thorough review of the nation's electrical system.[6]

* * * *

This chapter examines the reorganization of electric power that followed the passage of the Energy Policy Act of 2005. The new law transformed federal regulatory powers governing the nation's electrical system and called for the creation of the first set of mandatory cybersecurity regulations—known as the Critical Infrastructure Protection (CIP) Standards—for the electric power grid. The law was initially driven by—and on paper very much appeared to be indebted to—industry interests. The *Toronto Star* derisively called the new law the "Leave No Energy Company Behind Act" and dubbed it a "porkfest of subsidies" for the energy sector.[7] However, what looked on its face to be an attempt by the electric power industry to further consolidate its power morphed into something much different—new forms of accountability and public oversight of the electric power industry. Regulators from the Federal Energy Regulatory Commission (FERC) took a new set of thin administrative powers and transformed them into a robust system of checks and balances. Industry had drafted the federal legislation to leave regulators with a marginal role. But federal regulators had different ideas. They turned a rubber stamp into a hammer.

The development of the new cybersecurity regulations for the electric power grid reveals an important story about the possibilities of security

politics. The transformation of infrastructures into dangerous things made the 2003 U.S.-Canada blackout legible in a new way. The entwined cultural, institutional, and organizational changes that unspooled after 9/11 made it possible to view the power grid as a target of malicious subversion. The blackout became not an aberration or freak accident but something closer to a warning: a sign of the fragility of the grid—a sign of what *might* happen if enterprising terrorists or other malicious actors set their minds to attacking the power system. This framing was vital to the Energy Policy Act of 2005. It was used to justify layering new forms of security protections and a new system of regulatory power onto the power system. These changes created a space and an opportunity for federal regulators to seize an important oversight role. Ultimately, FERC leveraged its limited powers to implement cybersecurity regulations that were significantly tighter than the standards proposed by the industry. FERC bucked industry preference for limited administrative regulations containing ample loopholes and instead secured comprehensive cybersecurity controls. These new regulations directly confronted the social and technical challenges unleashed by deregulation (see chapters 1 and 2). Now, the unintended consequences of decades of technological and institutional change were called to account. New robust cybersecurity regulations created an avenue for ongoing public review and participation in the oversight of electric power. Like the case of freight rail (see chapter 5), concerns over terrorism led to the creation of new forms of accountability that unexpectedly shuffled and reordered the power between industry and other publics. To the surprise of industry insiders and many skeptics, security concerns were leveraged to reinvigorate public oversight of the industry. These changes were, as we will see, anything but a foregone conclusion.

Reforming Electric Power: The Electric Power Industry Makes Its Case—and Gets Nowhere

By the late 1990s, electric power had entered a brave new world. Deregulation had radically altered the industry. Vertical monopolies were replaced by new centrally managed regional markets. This shake-up (detailed in chapter 2) unfolded over decades, but by the end of the 1990s, industry incumbents were getting nervous. This new setup seemed to teeter on the edge of failure. Where before electric power had been defined by a fairly limited club of traditional utilities supervising and controlling their own systems, now thousands of new and different organizations worked to keep the lights on. This transition gave industry players pause. So much

has to go *right* for electric power to work, and this new, complex ecosystem appeared to be ripe (and indeed was) for accidents or unscrupulous manipulation and gaming of the deregulated system for profit. The Enron scandal and the 2003 blackout would prove these were anything but idle fears.

In the late 1990s, key industry players drafted model federal legislation to help manage some of the unintended consequences of this competitive new world.[8] The North American Electric Reliability Council (NERC) oversaw the development of the proposed legislation in what came to be called the *NERC consensus language*. NERC had long worked as a voluntary industry association and forum for power companies. It had been created by the electric power industry after the 1965 Northeast blackout to promote industry best practices and oversee the development of voluntary standards.[9] In the late 1990s, NERC worked with all corners of the industry—including the Edison Electric Institute (the industry trade association that represents private electric companies), the American Public Power Association (which represents publicly owned utilities), the Electric Consumers Resource Council (a trade association representing large industrial power customers), and state and federal regulators—to develop model legislation.[10] The need for reform was clear. As T. J. Glauthier, deputy secretary of Energy, Department of Energy, noted in testimony before Congress in 1999: "As we move to a more competitive environment, the reliability of our bulk power systems can no longer be entrusted to voluntary standards."[11] This changing landscape made some sort of mandatory reliability rules necessary. As backers of the NERC proposal repeatedly noted in congressional testimony, electric power had become too complex and too reliant on thousands of different interdependent players with different agendas and interests to be left to voluntary standards alone.

The centerpiece of the NERC proposal was self-regulation. The electric power industry would police itself. The model legislation sought to impose binding rules on these new players in order to make sure there were no "weak links" that could undermine or jeopardize the reliability of the larger interconnected power system. The proposed model legislation would transform NERC. Under the plan, it would be designated as an electric reliability organization (ERO). NERC would then be able to legally enforce the standards developed by industry. In effect, NERC would keep doing what it had been doing for years—serving as a venue for industry to hash out its own rules—only now the standards it developed would be legally binding. The industry trade association would now have real teeth. Under the proposed terms, NERC would oversee a process in which industry members would draft, propose, and vote on

possible mandatory and legally binding reliability standards. NERC, and not federal regulators, would then enforce these standards.[12]

This was a deeply conservative response to the challenges unleashed by deregulation. Federal regulators—FERC—would have little power and little say in shaping reliability regulations. They would review the standards that NERC eventually submitted against very narrow criteria, making sure they were developed according to the guidelines that the industry put in place and, at most, looking for procedural missteps. But the draft text was explicit: federal regulators would defer to industry on the *substance* of regulations. Regulators would wield a rubber stamp. Their approval would transform the standards that industry developed into legally binding rules, but FERC would have little opportunity to actively shape the content of standards. Deregulation had sloughed off many of the regulations and controls that had long governed electric power. Industry was not about to advocate for tighter federal regulations now. If anything, the NERC consensus language would give industry *even more power* to control and shape electric power as they saw fit, with little meaningful public oversight or accountability.

Industry heavyweights, state public utility commissioners, and federal regulators all lined up in support for the NERC consensus language.[13] During the 2000 election, candidate George W. Bush made energy reform a key priority.[14] After the election, President Bush quickly created the National Energy Policy Development Group (NEPDG), chaired by Vice President Dick Cheney, and put together a comprehensive plan for reform.[15] The group included cabinet-level participants and a cadre of industry representatives and lobbyists.[16] The deliberations of the group were secret and controversial. The industry representatives and lobbyists were not technically members of the NEPDG, a formal designation designed to prevent public disclosure of their participation.[17] A years-long legal battle over public disclosure of the NEPDG's records and the scope of executive privilege culminated in a ruling in favor of the vice president and the nondisclosure of the NEPDG documents.[18] Nonetheless, information about the group trickled out. The picture that emerged showed industry and government working hand in glove. The NEPDG met extensively with titans of the energy industry. Exxon Mobile, Enron, Duke Energy, Florida Power & Light, the American Petroleum Institute, the Interstate Natural Gas Association of America, and others representing the power industry had significant input into the group's deliberations.[19] Environmental groups were, on the contrary, given a cursory meeting (Vice President Cheney did not attend, and half of the allotted meeting time was taken up by

introductions).[20] The NEPDG produced a sweeping report titled the *National Energy Policy*. It substantially reflected the priorities of industry, including recommending opening the Arctic National Wildlife Refuge to drilling, offering subsidies for coal, and ramping up support for oil and gas production.[21] Unsurprisingly, the report also supported the adoption of the NERC consensus language that industry had been flogging for a number of years. In its litany of recommendations, the report called for legislation based on the principles outlined in the draft legislation.[22]

Despite the powerful backing of the executive branch, industry, and federal and state regulators, the NERC consensus language stalled. It was added to bill after bill in Congress. But the language could not get passed. Over and over again, the NERC consensus language died in committee. By as early as 1999, FERC chairman James Hoecker was starting to sound weary, remarking in congressional testimony that he had been pushing this same proposal before both the House of Representatives and the Senate for the better part of two years.[23] By 2003 the model language was starting to look like it might be a dead end … and then the blackout happened.

The August 2003 Blackout: Revisiting Electric Power

The 2003 blackout was the worst power outage in the nation's history. Over 50 million people in the U.S. and Canada lost power, and 508 generators at 265 plants fell off-line, affecting 61,800 MW of electric load.[24] The blackout's economic impact was staggering; estimates place losses at between $4 and $10 billion dollars.[25] On August 15, the day after the blackout, President Bush and Canadian prime minister Jean Chrétien announced the formation of the U.S.-Canada Power System Outage Task Force. The task force was charged with investigating the causes of the blackout and offering recommendations on how to improve the reliability of the electric power system.[26] As part of its work, it would investigate the blackout and conduct a broad review of the electric power system.[27] U.S. Secretary of Energy Spencer Abraham and Herb Dhaliwal, the Canadian minister of natural resources, were appointed to lead the task force.

The task force divided its work across three working groups: the Electric System Working Group, Nuclear Working Group, and Security Working Group.[28] Public-sector experts drawn from the Department of Energy, Nuclear Regulatory Commission, FERC, Federal Bureau of Investigation, Department of Homeland Security, U.S. national labs, state utility boards, and their Canadian counterparts staffed the working groups and worked in close collaboration with industry.[29]

The task force released its interim report in November 2003 and its final report in April 2004.[30] The final report identified a mix of technical failures, human errors, and institutional factors as contributing causes.[31] The task force found both low- and high-tech failures. It placed the blame on both faulty vegetation practices—tree trimming—and various computer glitches and errors.[32] The failure was not merely technical: at critical moments, human error—failure to turn critical equipment back on, failure of an IT team to recognize that their system reboot had failed to fix a stalled alarm, and other decisions or moments of inaction contributed to the outage.[33] The failings ran deep. The task force found pervasive institutional challenges. It echoed a by-now-familiar complaint: the lax and ambiguous regime of voluntary regulations that oversaw electric power was ill-suited to confront the challenges of ensuring that a competitive electric power system remained reliable. In examining the practices of FirstEnergy, the MISO, and PJM Interconnection—all key players in the outage—the task force found multiple violations of the industry's voluntary reliability standards. In other cases, the standards on the books were so ambiguous that they made meaningful compliance difficult.[34]

Terrorism, Cybersecurity, and the U.S.-Canada Power System Outage Task Force

The task force examined another possible source behind the massive outage: terrorism. The post-9/11 cultural and institutional recalibration toward questions of homeland security and fears over terrorism hung over the task force's work. The threat of terrorism now provided a powerful backdrop for nearly *all* discussions of infrastructure policy. After 9/11, major infrastructure failures were greeted with foreboding: Could this be another terrorist attack?[35] The power outage initially sparked such fears. On the day of the blackout, the New York Police Department activated its terrorism-response plans and deployed armed teams to protect high-profile targets, such as the New York Stock Exchange.[36] It was a tense moment. Only a few weeks earlier in July, U.S. and Canadian government officials had issued warnings to the energy sector based on collected intelligence indicating that Al-Qaeda might attack power plants.[37] However, in remarks delivered in San Diego several hours after the blackout, President Bush attempted to quell fears of terrorism in a brief statement to reporters, noting: "One thing I think I can say for certain is that this was not a terrorist act."[38] Yet concerns lingered. On August 18, *Al-Hayat*, an Egyptian news outlet, reprinted a communiqué attributed to Al-Qaeda claiming credit for the power outage.[39] Similar claims appeared in various news outlets in the following weeks and months.[40]

The possibility that the blackout *could* have been caused by a terrorist act informed the task force's work and ultimately shaped its recommendations. The task force examined seriously charges that the outage could have been caused by terrorism. In creating the task force, President Bush and Prime Minister Chrétien appointed three U.S. and three Canadian public officials to serve directly under the leadership of Secretary Abraham and Minister Dhaliwal. Among the high-profile U.S. appointees was Secretary Tom Ridge of the Department of Homeland Security. The selection of Secretary Ridge highlighted the prioritization of security and terrorism within the task force.[41] Similarly, the creation and designation of the Security Working Group signaled the prominence that security concerns would hold within the inquiry—and it underlined the broader prominence that security now held for discussions of infrastructure.[42] The Security Working Group ultimately confirmed President Bush's initial assertion that terrorism played no role in the outage. The Working Group found no evidence supporting published claims of responsibility and discounted any role of malicious activity.[43] But, despite concluding that this was not a terrorist attack, the task force nonetheless placed security concerns front and center in its eventual report and recommendations.

The task force moved beyond the particulars of the 2003 blackout to offer a more general assessment of the power system. The unbound imagination that now characterized much post-9/11 thinking was on full display: rather than narrowing its exploration to what had happened, the task force started to consider what was simply possible. It found reasons to worry. The task force found no concrete evidence of foul play, but it did find significant opportunities for real harm. Hacking the computer systems crucial to the electric power system could result in great damage. The task force, in an analysis that would have been familiar to staffers working nearly a decade earlier on the President's Commission on Critical Infrastructure Protection (see chapter 3), saw that the use of information and communication technologies (ICTs) created openings for enterprising terrorists to manipulate the power grid. The Security Working Group seized on the presence of such vulnerabilities as a significant area needing improvement.[44] The final report noted that "the increased reliance on IT by critical infrastructure sectors, including the energy sector, has increased the vulnerability of these systems to disruption via cyber means."[45] It noted that supervisory control and data acquisition (SCADA) systems are particularly susceptible to corruption. For the task force, the transformation of ICTs within electric power from standalone proprietary systems to generative, multipurpose machines operating

across corporate networks and the Internet created new cybersecurity challenges.[46] This new technological ecosystem, in the task force's view, could be exploited to undermine the functionality of the grid. The report readily accepted that electric power was a target, noting: "The generation and delivery of electricity has been, and continues to be, a target of malicious groups and individuals intent on disrupting this system."[47] In addition to targeted attacks, the spread of malware designed to exploit flaws in common operating systems and software applications could incidentally threaten electric power systems, despite not specifically being tailored for this sort of purpose.[48] The task force, both implicitly and explicitly, wrestled with the challenges presented by the reorganization of electric power. It grappled with a difficult question: Given the broader structural changes in the industry and the ongoing merger of operational technology (OT) with a more conventional information technology (IT) environment, how could electric power be secured?

The task force's discussion of cybersecurity was important. It reworked traditional notions of what security means within the electric power sector.[49] Historically, *security* had a fairly clear meaning in the context of electric power. It was defined as "the ability of the electric system to withstand sudden disturbances such as electric short circuits or unanticipated loss of system elements."[50] Security, in this sense, was taken to relate to system stability and reliability. A secure system within this framework meant planning for and having on hand adequate resources, particularly generating capacity, to meet fluctuating demand. Security practices comprised the difficult work of keeping the system in a state of balance that matched generation and consumption. System security confronted a range of possible disruptions, including operator error, spikes in demand, poor weather, and accidents—but it did not focus primarily on intentional disruptions. Under this rubric, a secure system is one that has adequate capacity to meet demand.

In a different vein, the NEPDG—Vice President Cheney's energy advisory group—had made energy security a key focus of its report back in May 2001. In that document, *energy security* was taken to mean securing access to cheap and reliable sources of power. As a practical matter, the NEPDG used energy security to argue for a reduction in oil imports in favor of increasing domestic production, exploring untapped and underdeveloped oil reserves across the globe, and strengthening trade with oil-producing nations.[51] That document had a few boilerplate sentences about infrastructure security and the perils of intentional disruption.[52] But the bulk of the discussion of energy security in this earlier report was

tethered to ideas about market volatility and resource scarcity. Energy *insecurity*, in this telling, was about spikes in prices for oil and possible production shortfalls. Energy security, as of May 2001, was a question of ramping up domestic production and opening up untapped resources.

The task force transformed the meaning of *security* within discussions of electric power: it split from both the more familiar use of security within electric power operations and the recently put forward notion of energy security found in Vice President Cheney's NEPDG report. The U.S.-Canada Power System Outage Task Force adopted a different notion of security. It aligned security with malicious activity, contemporary fears about terrorism, and attacks targeting infrastructure. In a conceptual remapping, it employed in its report what had become a more colloquial notion of security that means something closer to "defense against intentional attack." This was quietly revolutionary. Here, the idea of a "secure" system became something quite different: rather than emphasizing resource adequacy, security now became linked with authorized access, defensive postures designed to repel or discourage saboteurs, and the broader interests of geostrategic security. This was an important shift. By explicitly and implicitly emphasizing the threat of terrorism and malicious intentional disruption, the task force expanded the meaning of security to include a larger set of contingences that blurred conversations about electric power reliability with the larger ongoing discussion about terrorism and infrastructure security. Now, the task force argued, security planning must confront the possibility of enterprising intentional disruption. For decades, operators worried about the impact of storms, demand fluctuations, and accidents on the electric power system; now, they began to worry about terrorism.

The conceptual transformation mapped out in the blackout report was radical. It led the task force to question the standard approach to contingency management that the industry had long used, and it led to an intense focus on questions about cybersecurity. Historically, power operators relied on what is known as the *N-1 Criterion* to guide reliable operation. Long used by operators and written into the industry's voluntary guidelines as NERC "Operating Policy 2.A—Transmission Operations," N-1 directs operators to ensure that at any moment their system can withstand the loss of the *single* most critical transmission or generation asset without significant disruption of service.[53] N-1, or normal operating conditions minus one, provides a margin of error: it is intended to ensure that systems can absorb the single worst outage or contingency at any given moment without lapsing into a large-scale cascading failure.[54] This rubric became encoded

within standard operating procedures across the industry. During day-to-day operation, controllers continually reevaluate and adjust system performance as system conditions change (demand shifts, assets come online or move off-line, and so on) in order to ensure that adequate resources are always available. As conditions shift, operators continually adapt to a new N-1 state—that is, they make adjustments to survive the current single largest contingency within the reconfigured system conditions. Ideally, these changes are supposed to occur within 30 minutes, giving operators a short window to accommodate changing conditions on the ground and make sure their system is robust.[55]

The task force's redefinition of security chipped away at the utility of N-1. This approach to contingency management works well to confront random failures. N-1 as a planning guide assumes that when the mundane failures that have traditionally occupied the worries of operators occur—high winds, storms, operator error, and the like—they will only trip a single asset at a time.[56] Malicious attacks, however, do not always follow a random distribution. Storms do not coordinate and plan; terrorists might. Targeted assaults can occur at multiple locations and hit many assets simultaneously to maximize harm. In confronting terrorism, multiple simultaneous contingences, rather than just the single worst contingency, start to appear as legitimate possibilities.[57] What's worse, as NERC's chief security officer would soon note, the diffusion of common ICTs across the industry created possibilities for horizontal disruptions that could ripple across and within connected systems.[58] Malware might target and manipulate not just a single controller within a single utility; it could potentially target *all* of the power companies that use a similar hardware and software mix. Rather than disabling or disrupting a single isolated system, this could cause failures at multiple points within the larger electric power system at the same time. These types of failures are not well accounted for under N-1.[59] The task force argued that terrorism and the new ICT landscape created different challenges. The possibility of targeted simultaneous attack or horizontal failures of standard computer equipment go beyond the scope of N-1.

The task force followed the by-now-familiar testimony of industry insiders and regulators and concluded that an overhaul of the governance of electric power was long overdue. The traditional approaches to managing system reliability no longer seemed to fit. This was not a new insight. But the task force did offer a new pressing rationale for mandatory standards: security. Voluntary standards and N-1 might have made sense once upon a time, but as the task force concluded, fears over

terrorism—wherein malicious actors are assumed to always be lying in wait—combined with a technological system rife with exploitable vulnerabilities seemed to call for a drastically new approach. If the industry's standard way of managing contingencies was no longer useful, what might take its place? The task force had a suggestion: mandatory standards and new specific cybersecurity regulations.

Mandatory Regulations and Cybersecurity: The Task Force's Recommendations

The U.S.-Canada Power System Outage Task Force's final report provided a path for reform. It offered fundamental recommendations for overhauling the governance of electric power and presented specific recommendations for confronting the particular challenges of cybersecurity. The final report contained 46 recommendations.[60] At the highest level, the recommendations called for an overhaul of the process of developing and enforcing reliability standards. Voluntary measures and reliability standards had governed electric power for decades.[61] The task force recommended scrapping this approach. It noted that voluntary standards were routinely ignored with little punishment. NERC—the industry trade group—developed these standards, but they did not carry the force of law, and NERC had few tools to meaningfully address noncompliance.[62] The task force turned and endorsed the model legislation developed (to no avail) years before and recommended that NERC be given new legally binding power to develop and enforce reliability standards for the industry.

It was unsurprising to see the blackout task force yet again dust off the NERC consensus language. This language had become de rigueur for nearly *all* proposals to overhaul electric power. The task force's recommendations closely followed the NERC consensus language. In keeping with what by now had become a standard set of recommendations, the task force called for a staggered regulatory structure: NERC would internally create and enforce mandatory reliability standards in consultation with industry, and federal regulators, FERC, would have limited power to accept or reject the NERC-developed standards.[63] Under the task force's model, federal regulators would certify a single entity as the Electric Reliability Organization (ERO). The ERO would represent the industry and develop and enforce mandatory reliability regulations. The task force made plain in its final report that the ERO concept was designed and proposed with the understanding that NERC would fill this role.[64] The power industry would self-police. The industry would work through the ERO to develop reliability standards and enforce compliance. The role for

federal regulators was narrowly drawn: FERC would have the authority to accept or deny proposed reliability standards but not draft or develop regulations; it would be called to defer to industry in defining the content of the standards.[65] Within this scheme, FERC's approval would give industry-developed regulations legal standing. The proposed scheme would enhance—and not diminish—the power of industry: it would govern itself.

The task force did, however, move beyond the developed model legislation. Thirteen of the 46 recommendations in the report specifically addressed a new topic: cybersecurity.[66] Just as it had with the more general recommendations, the task force once again followed industry's lead. NERC had already been working on some modest voluntary cybersecurity guidelines before the blackout. In a coincidence, on August 13, 2003, *the day before the blackout*, NERC released its first voluntary cybersecurity standards, Urgent Action Standard 1200 (UA 1200).[67] UA 1200 contained 16 different sections concerning access controls, critical asset identification, and training. It was nonspecific. It asked relevant players to create and document plans, but it said little about what those plans should actually look like. Having a plan on file was enough.[68] The task force did not provide a review of UA 1200, but it nonetheless advocated adopting the mandatory reliability standards as a first step in addressing the cybersecurity issues now plaguing the industry.[69]

The NERC consensus language was not dead after all—post-9/11 fears and the work of the blackout task force gave it a new urgency, a new rationale, and a new life.

The Energy Policy Act of 2005: Reforms—Intended and Otherwise

On August 8, 2005, President George W. Bush signed the Energy Policy Act of 2005 into law. It was a significant piece of legislation.[70] In over 1,700 pages of text, it promised substantial changes to U.S. energy policy. In many ways the law appeared to be the apotheosis of deregulatory zeal. It continued the now long-standing trend of removing or relaxing administrative controls over electric power and serving up generous benefits to powerful industry players. The law offered over $14 billion in tax breaks for the oil, gas, and coal industries (it provided little support for renewable energy).[71] It loosened or cut environmental protections and created exemptions to clean water laws for the oil and gas industries.[72] The new law repealed the Public Utilities Holding Company Act of 1935—the key Depression-era law that had limited consolidation within electric power (see chapter 1). The new law was largely shaped by—and in some cases

directly drafted by—the energy industry. The report originally produced by Vice President Cheney's NEPDG provided a rough blueprint for the new legislation. The law contained most—but not all—of the recommendations included in that report.[73] The revolving door of energy executives who had worked with Vice President Cheney's group—including representatives from Constellation Energy Group, British Petroleum, and many others—to draft the *National Energy Policy* report could not have been happier.[74] Four years after the fact, their vision had been written into law.

Critics saw a disaster. Just as they had criticized the Cheney report years earlier, they attacked the law. Anna Aurilio, legislative director for the Public Interest Research Group, a consumer-focused nonprofit formed by Ralph Nader in the 1970s, tartly noted that "it's Christmas in August for big energy, and consumers get lumps of coal."[75] Environmental groups were equally furious.

But the new law had not only been shaped by wheeling-and-dealing industry executives. The 2003 blackout and the larger post-9/11 recasting of infrastructures as dangerous things played an important role. At a ceremony at the Department of Energy's Sandia National Laboratories in Albuquerque, New Mexico, President Bush touted the economic impact of the bill. Bush noted that the bill would help lower energy costs and support a growing economy. But the rhetoric of security was not lost in the mix. Bush was clear, stating: "It's an economic bill, but … it's also a national security bill."[76]

The NERC consensus language *finally* became law. Title XII of the act, known as the Electricity Modernization Act of 2005, for the first time granted authority to the industry to create and enforce their own legally binding mandatory reliability standards.[77] For nearly a decade, the industry had been trying to get some version of this legislation passed. The task force's work and the new post-9/11 context were crucial. It freighted the proposal with a new pressing rationale: protection against terrorism. As in the case of freight rail and the postal system, security concerns amplified and legitimized ideas that had already been circulating.

The Energy Policy Act amended Section 215 of the Federal Power Act. The law called for the development of mandatory reliability standards for the bulk power system and, in particular, required new cybersecurity standards.[78] It specified that reliability standards cover all the facilities and control systems necessary for the reliable operation of the bulk power system or portions of the bulk power system. Under the terms of the new law, generation and transmission were included within the "bulk power system," while distribution to end users was not covered. The text of the

law made it clear that reliability standards should be designed to support the operation of the interconnected power system such that sudden disturbances will not lead to cascading failure.[79]

The Energy Policy Act followed the template set by NERC and its industry partners (and the U.S.-Canada Power System Outage Task Force): FERC is given limited powers; industry controls—or is supposed to control—the drafting and enforcement of reliability standards. The act called for FERC to certify the ERO. Any entity could apply to become the ERO, but it was clear that NERC was going to be the ERO.[80] Once certified, the ERO can file reliability standards with FERC for approval. In examining the standard, FERC can approve the standard if it finds that it is just, reasonable, and in the public interest or remand the standard back to the ERO for further consideration. FERC's power is limited: it cannot draft reliability standards but rather only approve or remand. Additionally, FERC's review criteria are narrow. FERC is bound to give "due weight to the technical expertise" of the ERO when examining a proposed standard.[81] FERC is specifically instructed to defer on technical matters of content to the ERO. FERC can, however, upon its own initiative or in response to a received complaint, direct the ERO to develop a reliability standard.[82] But again, regulators cannot dictate the content or terms of the standard: they must defer to the technical competence of the ERO. The act also specified that the ERO is responsible for enforcement and can levy penalties on any covered entity found to be in noncompliance, subject to FERC review.[83]

The Energy Policy Act explicitly embraced cybersecurity within its purview. This was a new wrinkle not found in the earlier model legislation. The new law required the ERO to develop reliability standards that apply to intentional cyberattacks or what it termed a "cybersecurity incident."[84] It fully adopted the reworked notion of security discussed by the blackout task force. It defined a cybersecurity incident as "a malicious act or suspicious event that disrupts, or was an attempt to disrupt, the operation of these programmable electronic devices and communication networks including hardware, software and data that are essential to the reliable operation of the bulk-power system."[85] Though the act generally left the content of reliability standards to be determined, it singled out cybersecurity as an area where the ERO *must* develop standards.

The legislation sought to preserve the power of the industry. The Energy Policy Act supported a well-worn path: giving the electric power industry more power. Similar to the discussion of hazardous materials transportation in chapter 5, the threat of terrorism was used to justify an expansion of regulatory powers into a new domain: security. Unlike the expansion of

the Department of Transportation's power, which included a broad security mandate, the Energy Policy Act of 2005 took pains to limit the role of federal regulators in devising new security regulations. The law created a regulatory structure that supported the centrality of industry in the regulatory process. This was firmly in line with the broader promarket restructuring that occurred as part of deregulation. Federal regulators, it seemed, would have little to do but approve whatever NERC passed along. The NERC consensus language, the NEPDG's report, the blackout report, and the Energy Policy Act were each industry-friendly initiatives that sought to increase the power of industry over the terms of the bulk electric power system.

Yet, surprisingly, federal regulators did not sit idly by and rubber-stamp industry-approved reliability standards. On the contrary, in the coming years they leveraged their limited power to substantially redefine the terms of cybersecurity regulations. In doing so, they created greater public involvement and accountability over the operation of the electric power industry and successfully challenged the postderegulation status quo.

Implementing the Energy Policy Act: Regulatory Activism and Transformation

NERC quickly sought recognition as an ERO under the new terms of the Energy Policy Act and moved to convert its existing voluntary guidelines into mandatory reliability standards.[86] But it soon became apparent that the marriage between NERC and FERC was going to be a good deal more contentious than expected. Through the initial approval process, NERC and FERC clashed over the degree of industry autonomy, the specificity of new regulations, and the penalties that would be associated with noncompliance. This was the first test of the freshly minted regulatory process. Both industry and federal regulators attempted to set the tenor and contours of the new relationship.

FERC quickly moved to secure its turf. FERC Order 672, issued February 3, 2006, implemented key portions of the Energy Policy Act by defining the criteria for the certification of an ERO and outlining procedures for the establishment, approval, and enforcement of reliability standards.[87] In the order, FERC asserted a prominent role in defining reliability standards and rejected calls to defer to industry. FERC interpreted the Energy Policy Act broadly. It saw a role for itself that went well beyond simply affirming industry reliability standards. FERC's interpretation cut against the intent behind the long-in-the-works NERC consensus language, the task force recommendations, and, to a large degree, the apparent purpose of the Energy Policy Act. The interpretation was plausible within the boundaries

set forth by the act, but it was a generous interpretation. Most importantly, in defining the meaning of *due weight*, the importance of flexibility within mandatory standards, and the scope of remand authority, FERC's order consistently interpreted the Energy Policy Act in ways that favored an enlarged federal role and limited the autonomy and power of industry. This was *not* what industry had been fighting for. For nearly a decade, they had been trying to pass legislation that would allow them to enforce their own rules on their own terms. On paper, it looked like they had gotten everything they wanted. But FERC seemed to have other ideas.

FERC's order reiterated the basic criteria that the Energy Policy Act had set out for the ERO. Order 672 specifies that for an applicant to be certified as the ERO it must be independent, impartial, and open to all industry stakeholders.[88] Under the outlined terms, the ERO will develop reliability standards through open, deliberative processes involving industry stakeholders and then submit them for approval by FERC.[89] But FERC used Order 672 to make clear that the review process was a serious matter that could potentially shape the direction and content of standards. It would not be a rubber stamp. FERC cautioned that it would closely review proposed standards to ensure they did not amount to bland "lowest common denominator" rules that appeased all corners of industry.[90] The Energy Policy Act charges FERC to review proposed reliability standards while giving due weight to the technical expertise of the ERO. There is a tension between the call for FERC to defer to the expertise and the prerogatives of industry, on the one hand, and FERC's duty to perform oversight, on the other. In Order 672, FERC made clear how it intended to resolve these competing values. *Due weight* would not translate into a carte blanche grant of authority to the ERO. Industry hoped that FERC would adopt a limited frame of review and only scrutinize standards for procedural violations related to the development process (i.e., instances when the ERO deviates from its own guidelines). FERC rejected these calls. FERC refused to accept the claim that due weight establishes what is known as a *rebuttable presumption*—that is, a presumption that a finding by the ERO is to be taken as true unless a specific objection against it is offered—in favor of all ERO proposals. Such a position would have tied FERC's hands and allowed it to intervene only when a specific complaint was lodged. FERC rejected the contention that a standard that passes through the internal ERO development and vetting process automatically passes muster as just, reasonable, and nondiscriminatory.[91] Members of industry had wanted just such a presumption. Accepting it would have effectively shut FERC out of *all* substantive decisions, with little to do but certify that

industry standards were drafted according to the procedures laid out by industry.[92] FERC refused to be boxed in. Instead, Order 672 staked a strong claim in favor of expanded federal regulatory authority.

FERC's initial order regarding the ERO was a sobering response to the power industry and NERC. It rejected the notion that reliability standards should simply offer general guidelines (as UA 1200 and other voluntary standards had), as opposed to detailed requirements. Industry hoped that FERC would only review the ultimate aims of a proposed standard, rather than conduct an analysis of the precise mechanisms through which such aims are supported. FERC rejected this notion. In some cases, it noted, FERC would have to weigh in on the particularities of standards—not just the end goals.[93] FERC asserted that to ensure that reliability standards are technically sound, uniform, and unambiguous, an examination of implementation would in some cases be indispensable.[94] While FERC accepted the general contention that reliability standards would, in some instances, admit flexibility, it did not accept that this need rendered scrutiny or review of specific operations as beyond the scope of its duties.[95] FERC, here and throughout Order 672, signaled that regulations would in some cases define larger aims *and* in other cases define specific mandates—FERC review could and would take both under examination.[96] Here again, FERC enlarged the scope of its authority to conduct review: it could and would involve itself in substantive issues involving the implementation of reliability standards.

Finally, in Order 672 FERC asserted that the terms of the Energy Policy Act allowed it to not only remand proposed standards but to also include as part of this action a deadline for the ERO to address FERC's concerns and submit a revised standard. This, in time, would prove to be a very useful tool. Though FERC could not directly draft reliability standards, by asserting the right to remand with a deadline for revision and resubmission (and assess penalties if its directions were not followed), FERC carved out a proactive role for itself in the drafting process.[97] By remanding a proposal with a detailed explanation, FERC's hand could now weigh heavily on how industry would develop new standards. Through this process, FERC could become a de facto author of new standards.

Taken together, Order 672 was and is bracing: It offered a broad interpretation of the powers afforded to FERC under the Energy Policy Act, and it significantly qualified industry autonomy. It upset the postderegulation status quo. The Energy Policy Act and the recommendations of the task force intended to ensure that industry would have a decisive say over mandatory reliability standards while setting aside a minor role for

federal regulators to approve such standards in order to make enforcement legally binding. Yet in implementing the act, FERC expanded its role. Rather than being content to sit on the sidelines and offer rote approval of standards developed by industry, FERC interjected itself into the substantive work of defining new standards. FERC at each turn consistently offered readings of its authority that supported a greater federal role at the expense of industry and the ERO. Order 672 signaled to industry that the development of mandatory regulations—all regulations, not just cybersecurity regulations—would not be a pro forma exercise. Indeed, as would soon become clear, FERC intended to leave its mark on reliability standards.

Critical Infrastructure Protection Reliability Standards: Security, Autonomy, and the Challenges of Tightly Coupled Complex Systems

The scope and importance of FERC's power grab quickly became clear. Shortly after filing for certification as an ERO, NERC submitted its first batch of proposed mandatory cybersecurity standards, known as the Critical Infrastructure Protection (CIP) Reliability Standards, for FERC review.[98] The proposed CIP standards had already been approved by NERC through industry ballot.[99] This was a key test of FERC's power. Regulators made it clear that the proposed standards were deeply flawed. Subsequent iterations of the standards (version 2.0 and beyond) would have to look far different.[100] FERC had three key complaints. First, it took aim at two particular clauses within the standards—the "reasonable business judgment" clause and the provision that allowed regulated parties to "accept risk," in lieu of mitigating an identified risk, and opt out of the standards. Second, the regulators challenged the vagueness of the standards. Third, they criticized the weak penalties associated with the standards.[101] FERC used its power to push for successive versions of the mandatory reliability standards to be more comprehensive and binding.[102] On each issue, FERC pushed against industry preference. Underlying FERC's arguments on each of the above three issues was the recognition of the unique challenges of ensuring security within complex and tightly coupled networks. For FERC, security concerns were not merely rhetorical gloss: the regulators took concerns about malicious attacks seriously.

NERC's first set of proposed mandatory CIP reliability standards (CIP version 1) created a framework for the identification and protection of critical cyber assets to support the reliable operation of the bulk electric system.[103] In its initial filing, NERC outlined the intended scope of CIP reliability standards by defining the key terms *cyber assets*, *critical*

assets, and *critical cyber assets*. It defined cyber assets as "programmable electronic devices and communication networks including hardware, software, and data."[104] Critical assets were defined as "facilities, systems, and equipment which, if destroyed, degraded, or otherwise rendered unavailable, would affect the reliability or operability of the Bulk Electric System."[105] Critical cyber assets were described as "Cyber Assets essential to the reliable operation of Critical Assets.[106] Each standard would apply to the entities playing a role in the generation and transmission of electric power."[107] As drafted, CIP reliability standards were designed to prevent and respond to the malicious disruption or the attempted disruption of critical cyber assets.[108] Version 1 of the CIP standards offered a mix of general and specific requirements. In FERC's summary, the standards "require, among other things, that the responsible entities establish plans, protocols, and controls to safeguard physical and electronic access, to train personnel on security matters, to report security incidents, and to be prepared for recovery actions."[109] The CIP standards required responsible entities to produce security plans, access control policies, and incident response and recovery plans. More specifically, regulations also called for the use of antimalware software for all critical cyber assets (what it referred to as "Malicious Software Prevention"), the creation of a secure electronic perimeter around all critical cyber assets, and electronic and physical access controls (for an overview of each reliability standard, see table 6.1).[110]

FERC noted that the standards were much improved from the voluntary standards that had previously been in effect. The voluntary standard, UA 1200, covered much of the same ground but relied on self-certification—essentially requiring covered entities to be responsible for determining whether or not they were in compliance—and explicitly disallowed any economic penalties for noncompliance.[111] The proposed mandatory reliability standards moved away somewhat from these limitations, but FERC raised important and significant objections.

Rewriting the Standards: Reasonable Business Judgment and Acceptance of Risk

FERC took issue with language inserted into the proposed standards that provided industry with ample opportunity to, in effect, opt out of meaningful regulation. Each proposed standard contained text indicating that regulated entities may use "reasonable business judgment" in applying the standard.[112] To federal regulators, this was an enormous loophole. NERC outlined the meaning of *reasonable business judgment*

Table 6.1
Proposed cybersecurity standards: CIP-002–CIP-009.

Reliability standard	Title	Purpose	Key requirements
CIP-002–1	Critical Cyber Asset Identification	Identification and documentation of the critical cyber assets associated with the reliable operation of the bulk electric system.	1) Identification of critical cyber assets through "risk-based" methodology. 2) Documentation of critical cyber assets. 3) Annual review.
CIP-003–1	Security Management Control	Development and implementation of minimum security management controls to protect critical cyber assets.	1) Implement and annually review security plan covering standards CIP-002—CIP-009. 2) Implement and document security plan covering access and protection of information relating to critical cyber assets. 3) Implement and document plan for altering critical cyber assets.
CIP-004–1	Personnel and Training	Provide training and risk assessment for all personnel having access to critical cyber assets.	1) All employees, including contractors, with access must receive training. 2.) All personnel with access must receive a risk assessment (background examination). 3) Maintenance of list of personnel with access.
CIP-005–1	Electronic Security Perimeter(s)	Identification and protection of an electronic security perimeter(s), within which all critical cyber assets reside.	1) All critical cyber assets to be included within an electronic security perimeter. 2) Implement and document access control for perimeter. 3) Monitor and log access to perimeter.

(continued)

Table 6.1 (continued)
Proposed cybersecurity standards: CIP-002–CIP-009.

Reliability standard	Title	Purpose	Key requirements
CIP-006–1	Physical Security of Critical Cyber Assets	Creation of physical security program for critical cyber assets.	1) Implement physical security plan. 2) Control, monitor, and log physical access to critical cyber assets. 3) Test and maintain physical security systems at least once every three years.
CIP-007–1	Systems Security Management	Definition of methods, processes, and procedures for securing systems determined to be critical cyber assets.	1) Test all changes involving changes to critical cyber assets. 2) Disable all unused ports. 3) Use malicious software prevention programs and patch-management program. 4) Limit administrator access to critical cyber access. 5) Monitor systems for security breach. 6) Conduct annual cyber vulnerability assessment.
CIP-008–1	Incident Reporting and Response Planning	Identification, classification, response, and reporting of cybersecurity incidents.	1) Develop and implement incident response plan. 2) Document cybersecurity incidents.
CIP-009–1	Recovery Plans for Critical Cyber Assets	Development and implementation of recovery plans for critical cyber assets.	1) Create and annually review recovery plans. 2) Exercise plan annually. 3) Maintain backup information for successful restore of critical cyber assets.

Source: Proposed NERC CIP reliability standards collected in NERC, "Cyber Security: Standards CIP-002–1–CIP-009–1."

in a "Frequently Asked Questions" document submitted as part of its filing. In it, NERC noted that the phrase had a 200-year history in business and corporate common law and here was meant "to reflect—and to inform—any regulatory body or ultimate judicial arbiter of disputes regarding interpretation of these Standards—that Responsible Entities have a significant degree of flexibility in implementing these Standards."[113] Historically, the principle protects corporate officers and board members from judicial review. Under the accepted legal meaning of this phrase, corporate board members and directors are granted wide discretion in their official capacity to act on behalf of the business, and they are obligated to act in what they perceive to be the best financial interests of the company.[114] Their decisions are protected from judicial review as long as they do not act with gross negligence. As NERC described:

> Courts generally hold that the phrase indicates reviewing tribunals should not substitute their own judgment for that of the entity under review other than in extreme circumstances. A common formulation indicates the business judgment of an entity—even if incorrect in hindsight—should not be overturned as long as it was made (1) in good faith (not an abuse or indiscretion), (2) without improper favor or bias, (3) using reasonably complete (if imperfect) information as available at the time of the decision, (4) based on a rational belief that the decision is in the entity's business interest.[115]

The principle of reasonable business judgment would, if included, plainly protect the autonomy of industry players to operate with wide discretion.[116]

The power industry argued that reasonable business judgment was an important part of the CIP standards. For NERC, Pacific Gas and Electric, San Diego Gas and Electric, and others the inclusion of this language would protect the right of regulated entities to weigh the presumed costs and benefits associated with any new security measure. In their estimation, security should be carefully considered within the broader scope of the costs and benefits to each individual company. Without such a qualification, they noted, new forms of security could be prohibitively expensive.[117] Likewise, Tampa Energy and Progress Energy and others asserted that the reasonable business judgment qualification was necessary to provide responsible entities the flexibility to implement CIP standards in ways best suited for their operating profile.[118]

FERC disagreed. The reasonable business judgment qualification was a significant problem. It would—if included—make a mockery of the very idea of mandatory standards. This language would allow regulated parties to disregard (so-called) mandatory reliability standards if, in their estimation, doing so conflicted or was unsupported by the financial best

interests of the company. Including this qualification would protect regulated parties from effective oversight; it would narrow the circumstances in which FERC and the courts could step in and review the actions of regulated parties.[119] For FERC, allowing individual firms wide discretion ignored an obvious tension: the pursuit of individual financial well-being and the overall security of the bulk power system do not always align. The CIP reliability standards sought (and seek) to promote the security of the interconnected electric power grid, not the best interests of any *single* company's shareholders.[120] As FERC remarked in reviewing NERC's proposal, "business convenience cannot excuse compliance with mandatory Reliability Standards."[121] Indeed, mandatory reliability standards were initially called for precisely because these two interests—parochial financial interests and security—often diverged. The interdependent nature of the power grid, and more specifically the ICTs and networks connecting different parties, rendered the reasonable business judgment standard particularly troubling.[122] Weak cybersecurity by one party involved in the bulk power system could undermine the security of other interconnected parties and the larger grid. As FERC noted, "Cyber standards are essential to protecting the Bulk-Power System against attacks by terrorists and others seeking to damage the grid. Because of the interconnected nature of the grid, an attack on one party can affect the entire grid."[123] In other words, the security of a system is only as strong as its weakest link.

FERC took aim at another provision in the CIP standards that would allow regulated players to effectively opt out of the regulations. A number of the initially proposed CIP standards included clauses that would have allowed regulated entities to issue a statement "accepting risk" in lieu of complying with the outlined requirements.[124] Multiple proposed standards included this "opt out" language.[125] For example, CIP-007 would have allowed parties to opt out of the requirements to close all unused ports (connections), adopt regular patches for known vulnerabilities involving critical cyber assets, or use antimalicious software protection.[126] If a party issued a statement accepting the risk in these cases, it would no longer be bound by the specific requirement of the standard.[127] In NERC's view, this provision allowed responsible parties to acknowledge the inevitable limitations of particular security strategies and investments.[128] Tampa Electric and Idaho Power joined NERC on similar grounds, supporting the provision. In their view, the elimination of all risk was neither possible nor, when costs were taken into account, necessarily desirable.[129] Industry sought these loopholes or exceptions as a way of maintaining their independence. FERC rejected these claims and strengthened its oversight

power. Regulators argued that the effectiveness of mandatory reliability standards could not be squared with allowing entities to opt out of key requirements at their own discretion.[130] As FERC remarked, there were "no controls or limits on a responsible entities use" of the acceptance of risk exception.[131] As it noted, FERC would have little power to review these sorts of enormously consequential decisions:

> A responsible entity may invoke the "acceptance of risk" exception without any explanation, mitigation efforts, evaluation of the potential ramifications of accepting the risk, or other accountability. In essence, the phrase "or an acceptance of risk" allows a responsible Entity to opt out of certain provisions of a mandatory Reliability Standard at its discretion.[132]

The acceptance of risks was not tempered by even a modest requirement calling on entities to justify or explain why they had chosen this decision. Deciding to accept a risk could be done for essentially any reason. Echoing its concerns regarding reasonable business judgment, FERC remarked that due to the "interconnected control systems of various entities, an acceptance of a cyber risk by one entity is actually an acceptance of risk for all of those connected entities because the entity that initially accepted the risk is now the weak link in the chain."[133] The acceptance of risk by one party, then, would have an impact on others. It would force all other interconnected parties to also accept this risk regardless of the wisdom of doing so.[134]

FERC's objections to the seemingly bland language that dotted the standards were important. Here, in the trenches of bureaucratic minutiae, the basic terms between the industry and the public and its proxies were being sorted out. For FERC, both the reasonable business judgement and the acceptance of risk provisions were traps: they turned mandatory standards into mere suggestions; they turned federal regulators into a staff of pliant observers; and they would make industry, essentially, beyond review. FERC fought back. It directed that these clauses must be struck from the next iteration of the CIP standards. Regulators offered a stern warning: if NERC did not comply, it would be fined and stripped of its status as an ERO, and FERC would find a different organization to serve as the industry steward.

Rewriting the Standards: Specificity and Penalties for Noncompliance

FERC worked to give the CIP standards real bite. During the initial review of the proposed standards, it became clear that federal regulators were not going to rubber-stamp whatever came across their desks. In addition to closing the various loopholes identified above, they also worked to make

the standards more specific and the penalties for noncompliance meaningful. The first set of proposed CIP standards focused almost exclusively on general, vague outcomes at the expense of offering specific guidance concerning how to meet goals.[135] The regulations often amounted to little more than bookkeeping exercises. At the same time, the first slate of proposed penalties for violating the standards amounted to little more than a slap on the wrist and did little to encourage compliance. FERC flexed its power, and the standards were remade.

FERC and the industry had vastly different ideas about regulatory philosophy. For NERC, regulations would ideally specify outcomes and goals while deferring to the regulated parties on how to best meet those goals.[136] The CIP standards initially put forward general goals and left each individual entity to figure out *how* it might meet those goals.[137] For example, CIP-002 called for the use of "risk-based" methodology in developing cybersecurity plans. Yet the reliability standard offered no interpretation as to what such a methodology might entail.[138] FERC was leery. The lack of specificity made the standards little more than requirements to keep documents on file.[139] Of the 41 total requirements included within the initial slate of proposed CIP standards (each standard contained multiple requirements), *38* related to the mere possession of documents.[140] If these standards were adopted, FERC would have nothing to do except ensure that various documents were filed. It would lack any real power to review the substance of how regulated parties were confronting cybersecurity challenges.[141] FERC called for more concrete and binding standards. Without specific guidance, each entity could interpret the requirements as it wished and, in the process, undermine the interconnected bulk power system.[142]

FERC took issue with the proposed penalties attached to the draft standards. It was bad enough that the standards would, if adopted, allow regulated parties to pick and choose which standards to comply with. NERC also proposed paltry penalties for noncompliance. The penalties were set against a scale based on "violation risk factors." Each individual requirement and subrequirement of a mandatory reliability standard was assigned a corresponding violation risk factor. These factors were divided into three categories: lower, medium, or high.[143] These categories affixed the upper and lower limits of monetary penalties that could be assessed for infractions.[144] A high violation risk factor indicated that a violation of the requirement could directly contribute to or cause system instability and a cascading failure. A medium violation risk factor was reserved for requirements that, if violated, would affect the interconnected power

system but be unlikely to lead to cascading failure or instability. Lower violation risk factors were assigned to requirements that were only administrative in nature: these types of violations were not expected to affect the capability of the bulk power system.[145] The fixed range of penalties differed significantly: infractions of lower violation risk factor requirements carried fines between $1,000 and $25,000, violations of medium-level requirements carried fines between $2,000 and $335,000, and violations of high-level requirements ran from $4,000 to $1 million.[146]

NERC initially labeled the overwhelming majority of the requirements in the CIP standards—85%—as "lower" violation risk factors.[147] It attempted to classify mandatory cybersecurity standards as little more than administrative requirements. NERC proposed a $25,000 cap for most penalties. Federal regulators balked. This, in their view, failed to recognize the seriousness of the challenges of cybersecurity.[148] For example, CIP-002 called for the identification and recording of all critical assets and critical cyber assets. This carried lower and medium violation risk factors, respectively.[149] Yet, as FERC noted, these requirements were not merely administrative: they controlled how, *or even if*, the larger family of reliability standards would apply. Failure to identify a critical asset or critical cyber asset would, in effect, fail to trigger *any* of the other CIP standards.[150] FERC pushed back and sought real penalties. The regulators called for increasing the severity of 43 of the 162 proposed violation risk factors.[151]

The Surprising Success of FERC

FERC changed the CIP standards in a meaningful way. The Energy Policy Act was supposed to help codify industry control. Indeed, the initial version of the standards proposed by industry was incredibly lax—it allowed industry to opt out of regulation as they saw fit, imposed few specific mandates, and carried only modest penalties. Outside observers viewed the standards as a significant failure.[152] Yet something surprising happened. The standards improved. FERC pushed back and chiseled out a space for the real, meaningful public review of electric power. The regulators did not accept the watered-down standards initially proposed, and they did not accept sitting on the sidelines approving whatever the industry cooked up. FERC approved the initial proposed standards but, at the same time, used its authority to direct NERC to make significant changes and resubmit a new version of the standards for review. It was not an idle request. If NERC failed to develop standards that responded to FERC's specific concerns, FERC would strip NERC of its power. It was a serious penalty: removal of NERC as the ERO. FERC's demands were met. In the next

iteration of standards proposed by NERC—CIP version 2.0—the reasonable business judgment and accept the risk provisions were eliminated, the CIP standards became more specific, and the penalties were increased for 43 different types of violations.[153] This pattern has now repeated. Through multiple revisions—FERC approved the sixth version of the CIP standards in 2016—FERC has used its power to greatly alter and improve the CIP standards.[154] Although it cannot draft new regulations, FERC found a way to hold the pen. Taken together, FERC transformed the proposed mandatory reliability standards from vague administrative standards containing multiple avenues for opting out into something far closer to comprehensive regulations. The regulators have held industry to account.

Public Accountability and Security

The development of the first mandatory cybersecurity regulations for the electric power system is an interesting case. These new interventions confront the contradictions of deregulation. The move toward competitive markets inadvertently introduced new forms of vulnerability into the bulk power system. After the blackout of 2003, experts reviewing the bulk power system underscored the need to create mandatory cybersecurity standards to confront the threat of terrorism and intentional disruption. The tension between security and the market became an acute problem.

The Energy Policy Act of 2005 outlined what were basically conservative responses to these problems. It appeared, at first blush, to be little more than the retrenchment of the ethos of deregulation. It put in place measures that would—or at least seemed to—enhance the power of industry. The complex standards process that followed the act was designed to allow industry to self-police with little oversight from federal regulators or the public. Industry, it seemed, would have far greater power to control the terms upon which electric power operated, while other infrastructure publics would be shut out. Yet the development of mandatory regulations took an unexpected turn. FERC transformed the seemingly thin powers it was left with and carved out a significant role for expanded federal—that is, public—oversight of electric power. FERC expanded its power during the initial approval of NERC as the ERO—giving it more power to oversee and define *all reliability* standards moving forward. In time, this affected not just those standards related to cybersecurity but a larger set of standards related to emergency preparedness, facility design, interchange scheduling and coordination, and other topics. FERC put its power to use during the approval and review of the CIP reliability

standards and successfully transformed what were lax regulations into tight controls. Federal regulators took seriously the prioritization of security and followed through to a reasonable conclusion: tough standards were needed. FERC seized on the ambiguity within the Energy Policy Act's call for federal review in the public interest, on the one hand, and its stated preference for deference to the desires and judgment of industry, on the other. It offered an aggressive interpretation of the scope of federal power. Rather than simply enforcing industry-approved provisions, federal regulators took their role to review proposed standards seriously and directly challenged industry preferences in a meaningful way. New legislation designed to give industry broad power was reinterpreted by federal regulators to support robust public accountability. There is an important, if familiar, lesson here: once on the books, agencies and the public can spin regulatory authority in surprising ways.

The creation and implementation of mandatory cybersecurity reliability standards was a battle to determine authorship over new regulatory controls. How the uncertain threat of terrorism and an expanded notion of security would be translated into routine and standard practices was an open question that pitted industry and federal regulators against each other. FERC's success ensures that there will be meaningful public review of how electric power companies operate. New forms of public accountability have been created. The establishment of a substantial role for FERC in overseeing regulatory standards is a significant victory for the public. Robust FERC review ensures that electric power will be held to account by public institutions—electric power companies must justify and balance their interests against a larger set of concerns. Likewise, FERC's process of regulatory review ensures transparent decision-making and ample opportunities for comment from diverse stakeholders: it creates a durable forum through which different infrastructure publics can both access information about electric power and have their voices heard. The regulatory process opens the workings of electric power to public discussion and review. This is not trivial: the work of FERC's regulators to reassert themselves within electric power leads both to better security controls and, importantly, a recalibration of the social relations that define the electric power system. Deregulation, at its core, attempted to close infrastructures to public scrutiny and accountability. It greatly diminished the role that public institutions play in overseeing infrastructures in deference to the market. Here, through the guise of security, the workings of the infrastructure are pried open with a snap, and the public is let back in.