Title
Networks, deregulation, and risk: the politics of critical infrastructure protection

Permalink
https://escholarship.org/uc/item/5b93r8tr

Author
Ellis, Ryan Nelson

Publication Date
2011

Peer reviewed|Thesis/dissertation
Networks, Deregulation, and Risk: The Politics of Critical Infrastructure Protection

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Communication by

Ryan Nelson Ellis

Committee in charge:
Professor Robert Horwitz, Chair
Professor Lewis Branscomb
Professor Gary Fields
Professor Akos Rona-Tas
Professor Michael Schudson

2011
The dissertation of Ryan Nelson Ellis is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

____________________________________________________________________

____________________________________________________________________

____________________________________________________________________

____________________________________________________________________

____________________________________________________________________

Chair

University of California, San Diego

2011
DEDICATION

For my father, mother, and brother.

For Erin Dunne and B. Ellis.
# TABLE OF CONTENTS

Signature Page ........................................................................................................................................................................ iii

Dedication ........................................................................................................................................................................................................ iv

Table of Contents ........................................................................................................................................................................................................ v

List of Figures ..................................................................................................................................................................................................... vii

List of Tables ................................................................................................................................................................................................... viii

List of Charts .................................................................................................................................................................................................. x

List of Images ................................................................................................................................................................................................... xi

Acknowledgments .................................................................................................................................................................................................. xii

Vita ....................................................................................................................................................................................................................... xiv

Abstract of the Dissertation ........................................................................................................................................................................... xvi

Foreword: The Place of Communication within the Dissertation ........................................ 1

Chapter 1: Introduction: Post-9/11 Anxiety and the Problem of Critical Infrastructure Protection ........................................................................................................................................................................... 9

Chapter 2: Stumbling Toward Resilience: Revisiting and Rethinking Infrastructure Regulation ..................................................................................................................................................................................................... 32

Chapter 3: The Political Origins of Infrastructure Vulnerability: Regulatory Reform, Tight Coupling, and Complex Systems .............................................................................................................................................................. 113

Chapter 4: Risk, Politics, and Culture: Risk Society and Reflexive Modernization ............ 206

Chapter 5: Delivering the Mail: Terrorism, Efficiency, and the Politics of Post-Anthrax Bioterrorism Security ........................................................................................................................................................................ 231

Chapter 6: Green Security? The Environmental Movement, the Transportation of Hazardous Materials, and the War on Terror .................................................................................................................................................. 311

Chapter 7: The Creation of Mandatory Cybersecurity Reliability Standards: Autonomy, Interdependency, and Electric Power ...................................................................................................................................................... 349

Conclusion: The Politics of Critical Infrastructure Protection ................................................. 394
LIST OF FIGURES

Figure 3-1: ISO and RTO North American Operating Map ........................................ 204

Figure 3-2: Typical Elements of a Control System ..................................................... 205

Figure 5-2: Features of Intelligent Mail Barcode ...................................................... 308

Figure 5-3: Sample Intelligent Mail Barcode ............................................................ 309
LIST OF TABLES

Table 2-1: Post Office Mail Volume and Clerks Employed, 1910-1968 ...................... 104
Table 2-2: Postal Employment, 1930-1968 .............................................................. 104
Table 2-3: Postal Operating Deficit, 1926-1969 ....................................................... 105
Table 2-4: Postal Deficit as Percentage of Total Annual Postal Expenditures, 1926-1969 ........................................................................................................ 106
Table 2-5: Post Office Department Appropriation Categories .................................. 107
Table 2-6: Post Office Research and Development Spending, 1950-1969 ............... 108
Table 2-7: Percentage of Total Postal Obligations Devoted to Research and Development, 1950-1970 .......................................................................................... 108
Table 2-8: Annual Postal Investment in Large Capital Projects, 1955-1967 ............ 109
Table 2-9: Percentage of Postal Obligations Devoted to Large Capital Projects, 1955-1967 ........................................................................................................ 110
Table 2-10: U.S. Rail Mileage, 1840-1920 .................................................................. 111
Table 2-11: Percentage of Inter-City Freight Carried by Rail, 1930-1980 .............. 111
Table 2-12: Percentage of Industrial Chemicals by Mode of Transportation, 1963-1977 ........................................................................................................ 112
Table 3-1: Postal Operating Deficit, 1926-1969 ....................................................... 199
Table 3-2: Percentage of USPS Mail Volume Sorted Mechanically, 1968-1980 ...... 200
Table 3-3: Average Return on Investment Per Decade, Railroad Industry, 1940s-2000s ........................................................................................................ 201
Table 3-4: Percentage Utility and Non-Utility U.S. Electricity Generation, 1990-2009 ........................................................................................................ 202
Table 3-5: Total U.S. Electricity Generation and Resale, 1998-2009 ..................... 203
Table 5-1: Cross-Contamination of U.S. Postal Facilities ........................................ 304
Table 5-2: Characteristics of Five Processing and Distribution Centers Involved in 2001
Anthrax Attacks .................................................................................................................. 306

Table 5-3: Labor Costs as Percent of Total USPS Costs .................................................. 310

Table 7-1: Proposed Cybersecurity Standards: CIP-002–CIP-009................................. 392
LIST OF CHARTS

Chart 5-1: Simplified Model of Mail Processing and Distribution Network ................... 305

Chart 5-2: Cross-Contamination of U.S. Postal Facilities ........................................... 307
LIST OF IMAGES

Image 6-1: Chemicals of Mass Destruction........................................................................ 348
ACKNOWLEDGMENTS

The dissertation could not have been written without the assistance, support, and encouragement of many important individuals and key institutions. In particular, I would like to thank my advisor, Robert Horwitz, for his unfailing support and incisive comments over the past several years. His willingness to make my dissertation a priority and engage with new drafts, often with little more than a moment’s notice, is much appreciated. I would also like to thank my committee members—Lewis Branscomb, Gary Fields, Akos Rona-Tas, Michael Schudson—for their feedback and encouragement. I would also like to thank Jessica Wallack for her involvement during the early stages of the dissertation. Without their assistance, the project would simply not have been possible.

While completing the dissertation, I was fortunate to be apart of a variety of fertile intellectual and collegial work environments. The staff, faculty members, and fellow graduate students of the Department of Communication, University of California, San Diego (UCSD) provided a rich and interesting workplace which I feel lucky to have experienced. In particular, I would like to thank current and former graduate colleagues Lyn Headley, Kinda Al-Fityani, Michael Esposito, J.R. Osborn, Zara Mirmalek, and Ray-Shyng Chou. Colleagues at the Muir College Writing Program, UCSD likewise offered support and fine examples of professional decorum while working on the project. Recently, the University of California Institute on Global Conflict and Cooperation (IGCC) has served as a stimulating and exciting environment in which to work. I would like to thank my IGCC colleagues for their assistance during the past year.
Material research support came from a number of different sources. A fellowship from the Department of Communication and grant from the Dean of Social Sciences, UCSD provided resources to conduct research into the changing institutional framework governing infrastructure services. A grant from the University of California, Los Angeles Institute for Research on Labor and Employment supported research on the role of labor in defining new forms of postal security. Finally, the Smithsonian Institution provided valuable support in the form of a scholarship from the Smithsonian National Postal Museum for travel to Washington, D.C. and research on the postal system. In concert with support form the Smithsonian, Thomas Lera, Winton M. Blount Chair in Research at the National Postal Museum, and William A. Sandrik graciously offered their time and expertise, while Michael Pergam and Sara Marques provided warm hospitality.

Finally, the dissertation owes a large debt to the support and patience of friends and family. My mother, father, and brother have provided encouragement at every turn. Erin Dunne and B. Ellis offered unlimited understanding, partnership, and love. To them, I offer my deepest thanks and appreciation. Words cannot truly express my gratitude or good fortune.
VITA

EDUCATION

PhD  Communication, 2011
     University of California, San Diego

MA  Communication, 2005
     University of California, San Diego

BA  Communication and Sociology, 2001
     Boston College

HONORS AND AWARDS

Smithsonian National Postal Museum/National Philatelic Exhibitions of Washington, DC
(NAPEX), Scholarship (2009-2010)
Dissertation Fellowship, University of California, San Diego (2009)
Grant, Institute for Research on Labor and Employment, University of California, Los Angeles (2009)
Dean of Social Sciences Travel Fund Award, University of California, San Diego (2009)
Rita Lloyd Moroney Award, for “Binding the Nation Together: The Politics of Postal Service” (2008)
Member, Phi Beta Kappa Honor Society (2001)

PUBLICATIONS, WORKING PAPERS, AND CONFERENCE PAPERS

Publications


Working Papers
Conference Papers


PROFESSIONAL EXPERIENCE (Selected)

Instructor, Department of Communication, University of California, San Diego, 2008-2011

“Law, Communication, and Freedom of Expression (Part I)”

“Law, Communication, and Freedom of Expression (Part II)”

“Introduction to Communication as a Social Force”

Program Manager, Institute on Global Conflict and Cooperation, University of California, 2010-Current

Research Associate, John F. Kennedy School of Government, Harvard University, 2008-2009

Research Fellow, Institute on Global Conflict and Cooperation, University of California, 2006-2008
ABSTRACT OF THE DISSERTATION

Networks, Deregulation, and Risk:
The Politics of Critical Infrastructure Protection

by

Ryan Nelson Ellis

Doctor of Philosophy in Communication

University of California, San Diego, 2011

Professor Robert Horwitz, Chair

The dissertation, “Networks, Deregulation, and Risk: The Politics of Critical Infrastructure Protection,” engages post-9/11 debates over the role of public policy and novel technologies in crafting and maintaining resilient infrastructure networks against the threat of terrorism. In the wake of the terrorist attacks of 9/11, networks of communication, electric power, and transportation (as well as others) became sites of anxiety, debate, and, ultimately, intervention. The dissertation examines three separate infrastructures—the postal system, freight rail transportation, and the electric power grid—and considers the ways in which new regulations and technologies of control are introduced in an effort to counter the presumed disruptive impacts of terrorism. The study follows two related lines of inquiry: First, it examines the relationship between deregulation—the selective and incomplete restructuring of infrastructure regulation in deference to the market—and infrastructure vulnerability; Second, it considers how after
9/11 a cross-section of actors from inside and outside the sphere of traditional national security policymaking intervene in defining the terms on which critical infrastructure protection unfolds. The dissertation argues that deregulation, despite its other merits, creates new forms of infrastructure vulnerability. Critically, the partial deregulation of price-and-entry controls over infrastructure operations transformed the architecture of these networks in ways that make infrastructures both more efficient and more vulnerable to large-scale failure. By foregrounding the role of policy and law in shaping infrastructure networks, the dissertation demonstrates how previous regulatory regimes provided a hedge against large-scale failure and how deregulation, unintentionally, conspired to create networks beset by vulnerabilities. Additionally, the dissertation highlights the surprising conclusion that in the aftermath of 9/11 the perceived risk of terrorism ultimately contributes to the democratization of infrastructure governance. Drawing from the work of Ulrich Beck and notions of securitization, the dissertation examines how risk serves as a resource for otherwise marginalized interest groups to open infrastructure governance to a range of voices. In the reviewed cases, risk does not serve to support the suspension or diminution of democratic practices, but on the contrary enlivens particular aspects of practices of democracy.
Foreword: The Place of Communication within the Dissertation

The dissertation both theoretically and analytically is indebted to the traditions of communication research and provides a reworking of some key themes and approaches. In a general sense, the dissertation ricochets between two competing notions of communication—what James Carey defines as the transmission and ritual perspectives—by focusing on the intersection of information and communication technologies (ICTs) and larger technical-systems (such as transportation and energy), as well as the way in which these technologies become objects of public deliberation and reflection. Juggling these aims is a difficult but needed undertaking. To understand infrastructure vulnerability and new forms of control, it is critical to develop both an appreciation for the organization of ICTs—how and why they are ordered as they are—and an understanding for how such systems enter into the processes of public debate. A brief explanation of how the current study is situated within the confines of the discipline of communication, and why the discipline of communication is uniquely well-positioned to address the questions taken-up in this study, follows. The ideas and themes discussed below will be, of course, dealt with greater detail and nuance in the proceeding pages, but a short orienting section can provide some useful ballast.

James Carey outlines two distinct approaches to understanding communication. Carey argues that public perception and scholarly inquiry have primarily conceived of communication as a form of “transmission”: a means of sending and receiving information to enact control. Within this perspective, communication is inseparable

---

from the notion of control. Here, communication is seen as a way of acting at a distance, of exercising control through the sending of messages encoded within a particular medium. This view is evidenced within a long-line of communication research, including effects research and cybernetics, and collapses the distinction between information and communication. In Carey’s reading, this view of communication became increasingly central in popular and scholarly thinking after the advent of the telegraph. Here, for the first time, the means of communicating appeared to fulfill the dream of control at a distance. The telegraph did not so much separate the formerly conjoined processes of transportation and communication, but rather inverted the relationship between the two. As the telegraph and electronic communication emerged during the middle and late-19th century, wired forms of communication became central to controlling expanding, fast-moving, coordinated systems, of which the railroads are the most obvious example.³

In many ways, this study is well-situated within this trajectory of communication research. The study is primarily concerned with forms of control, broadly conceived, and their use in conjunction with geographically expansive, integrated, infrastructure networks. It engages in explicating how technologies of control are enmeshed within larger infrastructure systems. In this regard the dissertation most closely resembles and draws from James Beniger’s brilliant and flawed *The Control Revolution: Technological and Economic Origins of the Information Society.*⁴ Beniger provides a useful insight into the scope of ICTs that enable interconnected and geographically expansive systems to

---

operate. Beniger’s work develops a broad notion of ICTs that includes seemingly disparate inventions as formal bureaucracies and the multi-unit firm, office technologies such as bills of lading and the typewriter, and intellectual technologies such as operations and marketing research, and provides a useful model for the current study. Beniger’s reconceptualization of the so-called information revolution—redubbed as the “Control Revolution” as a longitudinal process that began in response to the exigencies of the industrial revolution during the early and middle of the 19th century and gradually transformed the realms of production, distribution, and consumption—redefines the scope of phenomena that can be considered to be information and communication technologies while drawing attention to the long-standing importance of such technologies to the making of the modern systems of transportation, manufacture, and consumption. In doing so, Beniger examines the interaction between macro and micro level changes and argues that greater forms of control both higher and lower levels of operation—at the level of firms and national economies, and the individual worker and consumer—are needed to domesticate the effects of industrial processes.

In a similar vein, the dissertation, as will be plain, considers technologies of control at multiple levels of operation and is particularly keen on tracing the intersection of how changes interact between macro and micro levels. The dissertation also adopts Beniger’s broad reconceptualization of ICTs by examining not only the material (and in some respects obvious) forms of information and communication technologies such as electronic control systems, computer networks, and software platforms, but also considers regulatory regimes, the structure of firms, and the intellectual or decision
technologies such as cost-benefit analysis. It will be argues that an expansive view is needed to sort out the multiple transformations that have beset infrastructure operation since the 1970s. These changes, alternately seen as technological and political, transform the structure of infrastructure networks. It is a central argument of the study that the process of regulatory reform begun during the 1970s that introduced market mechanisms more fully into the operation of infrastructures of differing type and the diffusion of new forms of ICTs throughout these networks are entwined. An understanding of deregulation, the occasionally misleading but nonetheless useful catch-all under which the political restructuring of infrastructures can be grouped, helps explain the particular and selective way in which new ICTs are deployed within infrastructures. The dissertation argues that the linked processes of deregulation and technological change deliver greater efficiency at the cost of creating new forms of vulnerability. New post-9/11 efforts to create secure infrastructures are, in some ways, attempts to cope with the unintended consequences of these transformations. Often, the search for new forms of security turns toward ICTs not only as sites of vulnerability, but also as technological “fixes” that hold out the promise, illusory though it may be, of seamless error-free networks.

Another relevant strain of research within the transmission model can be found in studies by Herbert Schiller, Daniel Headrick, Thomas Streeter, and Janet Abbate that emphasize the role of national security concerns in molding ICTs. The dissertation also

---

finds commonality with these studies that have attempted to sketch the interconnection, increasing during the late 19th and 20th century, between military concerns and ICTs. Notably, inquiries into the way in which the military aided or shaped the development of telegraph networks, radio, satellite, and other forms of electronic communication provide a partial illustration of what this study hopes to accomplish. Generally, these studies argue that the aims of national security became at key moments salient variables in determining how ICTs would develop. This study, likewise, examines how national security concerns, the threat of terrorism in particular, weigh on the current organization and operation of ICTs. Unlike these studies, however, the present work is interested in ICTs not as stand-alone enterprises, but as key components of other infrastructures.

The dissertation also identifies limitations within the transmission model in general and Beniger’s approach specifically. Beniger fails to acknowledge a clear sense of power, culture, and contingency as a part of the control revolution. Beniger’s reconceptualization of ICTs and focus on the processing of information as the fundamental activity of all purposive entities, elides important questions of directionality, power, and contingency. In his reading, how goals are selected (in effect, what is being controlled) who is controlling whom (or the social and political dimension of control), and why certain forms of control emerge over other seemingly viable options are largely unexplored. Despite praising and drawing insight from Max Weber, Beniger avoids questions of substantive rationality. Beniger is relentless in parsing new forms of control, but is less adept in explaining how the larger goals of these enterprises were devised,

challenged, and in some instances transformed. Debates during the 19th and 20th centuries continually questioned the purposes of the new forms of business. The telegraph, to use an obvious example, was the subject of a half-century debate about whether or not its aim was public edification or private profit. Determining the larger purpose of this system had great significance in defining how it operated as means of control. Additionally, Beniger explicitly draws theoretical insight from cybernetics, yet he largely ignores one of the central preoccupations of Norbert Wiener: the larger societal implications of new forms of control. Beniger is baffled by critics of bureaucratization and sees such attacks as misguided, while briefly passing over the differential benefits and costs of new forms of control with the notation that: “the Control Revolution …represented the beginning of a restoration – although with increased centralization – of economic and political control.” In part such a lack of curiosity stems from the essentially functionalist perspective adopted throughout; if new forms of control represent natural inevitabilities, there is little sense in agonizing over who does and does not benefit as a result of such innovations. Finally, Beniger’s account reads like a victor’s history, omitting any sense of alternate possible paths and forms of development. The notion that, perhaps, other forms of control could have developed is absent; the rich history of failure finds no home in his account. As a whole, these limitations signal potential pitfalls that lie along the path through which this study proceeds; care is taken not to repeat what I perceive to be important oversights.

---

7 Beniger. The Control Revolution.7, 59-60.
Beniger’s failings stem ultimately from a failure to recognize communication as anything other than a technique of control. Yet, for Carey another notion of communication can be recovered from older, archaic traditions that view communication as a form of ritual. This ritual perspective considers communication not only as discrete forms of information exchanged to enact control, but rather as a form of shared beliefs. In highlighting the limitations of the transmission view, Carey underscores the importance of considering communication as a form of culture. Drawing from Weber, John Dewey, and Clifford Geertz, Carey argues that communication is not only, or perhaps even centrally, instrumentation, but rather is a shared symbolic process within which reality is produced, maintained, and transformed.

To this end, the dissertation draws from treatments of risk that note its social and symbolic dimensions. Central in this regard is the work of Ulrich Beck and his thesis concerning what he designates as “risk society.” For Beck difficult to define forms of risk offer a type of political potential that can lead to new forms of engagement and challenges. Beck’s work is situated within the ritual perspective of communication insofar as it considers risks to be sites of open-ended definitional struggles. The dissertation considers how infrastructures are not only beset by and rely on ICTs, but are also objects of societal reflection and deliberation. That is, they are objects ensnared within a larger cultural framework that designates them certain values and meanings. The dissertation considers how highly visible events such as the terrorist attacks of 9/11

---

have helped redefine how infrastructures are viewed by policymakers, operators, and the public. The redefinition of infrastructures as loci of uncertain danger opens a host of new, difficult to resolve, questions concerning the role of the state and private companies in achieving security. An understanding of how such forms of uncertainty are resolved is one of the central aims of the current study.

Carey has, I believe, correctly identified a central tension found within the discipline of communication. The tension between divergent approaches to thematizing what communication is can be seen in the broad scope of work grouped within the confines of the discipline. The dissertation moves between these two views in a productive manner. Untangling the roots of infrastructure vulnerability and way in which security is uneasily grafted into the operation of infrastructure networks requires an understanding of the breadth of communication scholarship. Communication studies, in its multiple and often disjointed iterations, provides a useful point of departure.
Chapter 1: Introduction: Post-9/11 Anxiety and the Problem of Critical Infrastructure Protection

Infrastructures are, for the most part, boring; by their very nature, they support other forms of activity and do not call attention to their organization. Well-functioning systems of communication, transportation, and power are taken for granted and hidden in plain sight as mundane features of daily life. While their operation is of vital interest to the workers and regulators engaged in keeping the lights on, moving the mail, or operating waste-water systems, it is rare that the minutia inherent to their operation—questions of standards, technologies, labor practices, and regulatory authority tucked behind drab and esoteric jargon—becomes a visible point of reflection and engagement for those not “inside” their operation. Despite their apparent seamlessness, however, infrastructures are, in nearly all instances, sites where the play of power operates. Important contestations and battles, long since waged, are encoded and built within infrastructures and, over time, become obscured as normal.

Occasionally, however, this typical situation is inverted: the dull and boring workings of infrastructure appear interesting, pressing, and of a more general concern. Failures, outages or feared outages, illuminate the once hidden and politicize what were previously considered to be “technical,” neutral and expert, domains. The threat of terrorism, in the aftermath of the terrorist attacks of September 11th, 2001, operated in just this manner. After the attacks, infrastructures were recast as sites of vulnerability,

---

danger, and fear. Seemingly insignificant questions concerning baggage handling, the tiny pores of paper envelopes, the operation and ownership of ports, and a host of other examples, were placed on the political agenda. No longer taken for granted, infrastructures became visible, debated, and, in some instances, remade.

The dissertation examines what is often referred to as “critical infrastructure protection”—the post-9/11 revisiting of infrastructure security. In the wake of the terrorist attacks of 9/11, networks of communication, electric power, and transportation (as well as others) became sites of anxiety, debate, and, ultimately, intervention. The dissertation puts these security debates into a broader historical frame of reference. I argue that infrastructures in the U.S. underwent two transformations beginning in the 1970s, deregulation and the information revolution, and that these transformations created new forms of systemic vulnerability that reduced resiliency. While delivering many benefits, including increased efficiency, new services, and reduced costs, deregulation and the subsequent integration of new forms of information and communication technologies into infrastructure systems unintentionally created conditions that make these systems more prone to large-scale failures. These new vulnerabilities lie outside of normal operating conditions; they are not challenges to the regular operation of infrastructures, but rather represent new low-probability, high-consequence dangers and risks. In the aftermath of 9/11 these vulnerabilities became objects of explicit concern and negotiation. The possibility that malicious actors could attack these networks and

---

2 A note on terminology: throughout the dissertation the term “infrastructure” is used, as is colloquial, to refer to both the material networks and the key services operating these networks. In instances where infrastructures and infrastructure services depart in important respects, the distinction will be noted.
exploit their vulnerabilities became a public issue. How to reconcile the uncertainty of terrorism with the varied costs of security is a difficult balancing act. The stakes, for workers, users, operators, regulators, local, state, and federal authorities both were and remain quite high: the difference between doing too much and too little is difficult to know.

The dissertation examines post-September 11th debates about infrastructure security through a consideration of three separate networks: the postal system, the electric power grid, and freight rail transportation. I follow two related lines of inquiry. First, I examine the relationship between deregulation and vulnerability. Post, electric power, and rail, like most infrastructures, underwent a process of partial deregulation beginning during the 1970s and 1980s that relaxed price-and-entry regulation in deference to market-based measures. Selective deregulation transformed both the architecture of these networks and the sets of relationships structured around and within these different infrastructures. Deregulation delivered a number of tangible benefits. In the aftermath of regulatory reform, infrastructures operated with greater efficiency; redundancies were eliminated, and, for some key services, prices dropped. The process of deregulation, importantly, also shaped how new forms of information and communication technologies (ICTs) would be adopted and integrated into the operation of the postal system, the electric power grid, and, to a lesser degree, the freight rail network. Deregulation and the information revolution were not only contemporaneous; rather, the unfolding process of deregulation in important ways shaped and drove precisely how new communications technologies would be integrated. The dissertation argues that these processes of
restructuring—related and entwined as they were—created new forms of vulnerability and imperiled resiliency. They remade the operation of the postal system, electric power grid, and freight rail network in ways that make large-scale, irregular, failures more likely. Exploring this relationship—between deregulation and infrastructure vulnerability—occupies the first half of the dissertation.

Next, I consider how in the wake of the terrorist attacks of 9/11 these vulnerabilities became objects of political debate and, ultimately, intervention. I turn to investigate what we might term the bureaucratization of risk.\(^3\) I use the term bureaucratization to capture the way in which the abstract threat of terrorism—the idea of a possible terrorist attack at some point in the future—is written into stable and routine practices of security expressed in new regulations, codes, practices, and technologies.

Post-9/11 new worst-case scenarios concerning attacks targeting domestic infrastructure networks circulated widely in the press, academic circles, industry, and different levels of the government. The attacks of 9/11 and the subsequent distribution of anthrax through the U.S. mail vividly illustrated how infrastructure networks could serve as a means of causing great harm, panic, and economic damage.\(^4\) The mere possibility of catastrophic loss became an acute source of concern for the state, operators, and a host of groups affiliated with or interested in the operation of these networks. I examine the debates and

\(^3\) A brief clarification: Frank Knight offers the central distinction between risk, which is calculable, and uncertainty which resists calculation. Here, I use risk is the colloquial sense of perceived danger or harm rather than the more specific sense indicated by Knight. In Knight’s terms, the process under examination here and in Chapters 5, 6, and 7 would be more properly termed the bureaucratization of uncertainty. Frank Knight. *Risk, Uncertainty, and Profit.* 1921. New York: Kelley & Millman, 1957.

\(^4\) The anthrax attacks were, at the time, seen as a terrorist attack. Sheryl Gay Stolber and Judith Miller. “A Nation Challenged: The Response; After a Week of Reassurances, Ridge’s Anthrax Message is Grim.” *New York Times.* Oct. 26, 2001.
new practices of security in three realms: biological terrorism and the postal system; cyber-security and the electric power grid; and finally the transportation of toxic materials via railroads. The threat of terrorism casts a certain amount public attention back on the operation of infrastructures and, in the process, serves a democratizing function by elevating the “sunk” politics of infrastructures above ground. Now, the constitutive political choices embedded within the architecture of infrastructures become clear. In each case, how to mesh the exigencies of expansive, open, networks with the perceived needs of security was an open question subject to the competing interpretations of multiple constituencies. Drawing from the work on risk by Ulrich Beck and others, I examine the way in which new practices of security—technologies, policies, and regulations—are objects of contestation and debate. The idea of terrorism provided a range of groups, including labor, environmentalists, local governments, and others, with a resource to legitimate and advance their claims about the ordering of infrastructure systems. Here, risk serves as a type of political capital to be appropriated by diverse actors, including those often outside of typical discussions of national security. How this capital is used, how the salient but difficult to define risk is melded or translated into novel forms of security, is ultimately, an empirical question to be considered on a case-by-case basis. Like Beck, I elide the easy assumption that new forms of risk, in every instance, support the reification of power, but instead offer the possibility, however contingent, of inverting, undermining, or remaking politics.

The remaining portion of the introductory Chapter provides a more detailed overview of the topic of the dissertation, the relevant theoretical materials that are central
to the undertaking, and the structure of the work. The dissertation should be of interest to a broad range of scholars working in various fields. Ultimately, the dissertation is rooted in the traditions and theoretical insights of the field of communication. However, it is my hope that the work will also prove useful to those working within sociology, political science, economics, science and technology studies, and the field of security studies.

Overview: Structure of the Study

The dissertation is divided into two broad sections: the first section examines the process of deregulation and how it ultimately remade the postal system, freight rail network, and electric power system in ways that increased vulnerability; the second section considers post-9/11 debates and interventions oriented toward infrastructure security within each of these selected infrastructures. A short outline of the sections and chapters follows.

Section I: The Political Origins of Infrastructure Vulnerability

During the 1970s and 1980s, price-and-entry regulations were dramatically, and quite unexpectedly, largely dismantled.\(^5\) Infrastructures were partially liberalized—opened up to increased competition—and rate regulation was relaxed. This process of deregulation transformed the rules governing infrastructures that had been in place for decades (or longer) and remade the relationship between the state, network operators, and private capital. Deregulation attempted to remove or relax the role of the state in overseeing key aspects of infrastructure services and introduce, however imperfectly, 

---

market-based competition. At its most basic level, the revision and in some cases rescission of regulations governing price and entry attempted to replace administrative controls with market rationality.⁶

Deregulation impacted a broad swath of services, including post, telecommunication, transportation (rail, trucking, and air), electric power, and was in each instance significant and partial. It was significant insofar as it represented a historic repudiation of the basic compact that attempted to strike a balance between the broad aims of social policy, heterodox interests, and particular economic and technical characteristics of infrastructures in governing the operation of these services; it was partial in that it did not completely undo or eliminate state supervision or intervention in the operation of infrastructures—numerous regulations were left intact and still exist. While deregulation mainly focused on industry-specific price-and-entry regulations, it left in place so-called social regulations, such as those governing environmental pollution and occupational safety and health.⁷ In all instances, deregulation was a political process of restructuring that sought, with different degrees of success, to remove the regulations governing price and entry in infrastructure services.⁸ Price-and-entry regulations or “economic regulation,” governed who could provide infrastructure services and set the rates such firms could charge for service. In aim and practice, price-and-entry regulations created monopolies or quasi-monopolies in the realm of communications

---


⁸ On deregulation as a political project, see Horwitz. *The Irony of Regulatory Reform*. 
(including postal service), power, and transportation. Legally enforced limits on entry discouraged competition. In turn, rate regulation attempted to ensure that providers did not abuse monopoly power, propped up cross-subsidies to support the expansion and performance of the network in ways that were deemed to be socially desirable and in the "public interest," and provided firms with operating revenue to fund expansive systems. Price-and-entry regulation offered a quid pro quo; providers were given protections from competition, which was understood to be destructive within these particular industries, while in exchange rate regulation checked predatory activity and sought to support certain socially desirable forms of service.

In this section of the dissertation, I revisit the process of deregulation within the postal industry, electric power, and freight rail transportation. In each case, new legislation during the 1970s and 1980s greatly transformed the institutional framework within which these services operated. The Postal Reorganization Act of 1970 began the process of transforming postal service by paving the way for the partial liberalization of the postal market, phasing out direct subsidies from the federal government for general operation and undercutting the continuation of cross-subsidies between different categories of service in a movement towards rates that more directly reflected market conditions. In electric power, the Public Utility Regulatory Policies Act of 1978 (PURPA) began the tentative creation of competitive electricity markets and the splitting of vertical monopolies controlling generation, transmission, and distribution. Finally, in rail service, the Railroad Revitalization and Regulatory Reform Act of 1976 and the
Staggers Rail Act of 1980 transformed the industry and allowed shippers and carriers to set rates privately and for greater exit and entry in the industry.

In all cases, these changes undid aspects of price-and-entry controls in an effort to place these services on something closer to a market-based model. Deregulation led to the reorganization of the architecture of these networks and, at the same time, transformed the relationships among the various parties invested in the service. These changes, despite their other merits, created or extended new forms of vulnerability and imperiled resiliency by eliminating redundancies and creating tightly-coupled or complex material networks prone to single-point, large-scale, failures. Here, the contradictions of deregulation emerge: the pursuit of new efficiencies, at the same time, creates or enlarges forms of forms of vulnerability. Importantly, deregulation shaped how new information communications technologies would be adopted. In some cases, the information revolution simply provided a new means of achieving old tasks; networked computers replaced informational work that was previously done with different tools. In other instances these new tools allowed for new possibilities of control and coordination that were previously unavailable. These technologies, however, created new forms of interdependencies and opened these networks to new possibilities of failure. Additionally, the transformation of the relationships between different key parties involved in the provision of service imperiled resilience. The creation of new forms of competition and private rate-setting within domains that were previously subject to enforceable, legal, limits on the number of providers tipped the balance of power between providers and certain large-volume users of infrastructural services. In doing so, the
added costs of security are exceedingly difficult to pass on to customers that, now, have the ability to choose from a cross-section of providers. Seen in this light, the successes of deregulation are qualified: the efficiencies and lower tariffs they achieved came at the price of greater vulnerability and brittleness. Likewise, from this angle the inefficacies of the old regulatory model, much decried during their day, appear to have offered some overlooked benefits. Stable, monopoly or quasi-monopoly markets, characterized by redundancy incidentally were resilient and provided a check against systemic vulnerabilities.

Underscoring the obscured or overlooked failings of deregulation by drawing a connection between the process of regulatory reform and infrastructure security challenges certain basic assumptions about infrastructure vulnerability. Rather than seeing infrastructure vulnerability as alternately the product of years of not-so benign neglect, the autonomous logic of technological or economic progress, or an inevitability of private ownership, focusing on the transformations associated with deregulation highlights that vulnerabilities are politically and historically marked; they spring in part from critical and specific political decisions that are linked to the upheavals in infrastructure governance that occurred during the 1970s and 1980s.\(^9\) Such challenges or problems reflect, then, not inevitabilities but instead the particular consequences of the

---

political project of deregulation. Further, placing infrastructure vulnerability within the
countext of deregulation notes a connection between forms of regulatory intervention that
are usually seen as separate and distinct. Here, there is a connection between the
loosening or removal of economic regulation and the later need for further social or
behavioral regulatory interventions—undertaken in the name of security after the terrorist
attacks of 9/11—to make the operation of infrastructures palatable. Thus, contemporary
efforts to create policies and practices of security, which are dealt within the second half
of the dissertation, appear to be an attempt to reconcile the contradictions of deregulation.
Ironically, as others have pointed out in relation to a different set of questions,
deregulation requires additional and ongoing regulation in a newly uncertain
environment.\textsuperscript{10}

In sketching this argument, the dissertation provides an overview of the history,
operation, and consequences of regulation across different infrastructures and examines
the dynamics and consequences of deregulation. I will then examine in detail the process
of restructuring as it unfolded within specific areas of post, freight rail, and electric
power. Theoretically, the dissertation considers infrastructures to be examples of what
Thomas Hughes describes as socio-technical systems; they are more than just
compilations of hardware and technical components.\textsuperscript{11} Rather, they are made of the
intersection of technologies, politics, economics, and culture. In examining the durability
and later transformation of regulations, the dissertation argues that regulation and

\textsuperscript{10} For example, see Patricia Aufderheide. \textit{Communications Policy and the Public Interest: The

\textsuperscript{11} Thomas. P. Hughes. 1983. \textit{Networks of Power: Electrification in Western Society, 1880-1930}. Baltimore:
deregulation can be best understood within the context of path dependent explanations and neo-institutional accounts of institutional creation, durability, and flux. In assessing how regulation and deregulation shaped infrastructure networks, I turn toward the work of Charles Perrow. Perrow’s work on normal accidents provides a framework for considering the relationship between network structure and vulnerability. Perrow introduces the categories of “tight coupling” and “complexity” and provides a general theoretical framework through which to examine and assess the structure of infrastructures as organized under regulation and deregulation.\footnote{Charles Perrow. *Normal Accidents: Living with High-Risk Technologies.* 1984. Princeton, NJ: Princeton UP, 1999.} Perrow offers a general vocabulary and set of concepts that can be used to examine post, rail, and electric power.

The first section of dissertation comprises two chapters.

*Chapter 2: Stumbling Toward Resilience: Revisiting and Rethinking Infrastructure Regulation*

This Chapter introduces the work of Charles Perrow and provides a general account of regulation as it operated across post, rail, and electric power. This Chapter provides an overview of how regulation shaped the structure of the postal system, freight rail network, and electric power in ways that provided a hedge against large-scale failures. Politics, in the guise of regulatory oversight, played a key role in shaping infrastructures. In post and electric power, regulation was central to creating systems where control was distributed throughout the network, a form of loose coupling. In freight rail, regulation supported a linear system by buttressing a redundant rail network.
and discouraging the transportation of large quantities of dangerous chemicals through the rail network.

Chapter 3: The Political Origins of Infrastructure Vulnerability: Regulatory Reform, Tight Coupling, and Complex Systems

In this Chapter, I examine the process of deregulation and how it transformed the structure of infrastructure systems in ways that made them vulnerable to large-scale failure. The Chapter introduces theoretical treatments of path dependence in order to provide an account of why seemingly inviolable regulations began during the 1970s to be redrawn in favor of a closer embrace of markets. Deregulation reordered the incentives and rules governing infrastructure operation and, consequently, initiated the transformation of the postal network, freight rail transportation, and electric power from systems characterized by loose coupling and limited complexity into complex and tightly-coupled networks. Deregulation reordered the postal network and encouraged the adoption of centralized, automated, processing hubs and elevated the power of large industrial mailers in setting postal policy. The reorganization of freight rail, in a similar fashion, eliminated redundant track, increased shipments of dangerous chemical hazards, and empowered shippers in setting rates. Likewise, the creation of competitive markets for electricity eliminated redundancies and led to the adoption of new common forms of ICTs and centralized control. Through this Chapter, I argue that increased vulnerability is a hidden cost of deregulation. Regulatory reform pursued new efficiencies, but at the same time undermined the historic support that regulation offered for the creation and maintenance of resilient systems.
Section II: Infrastructure Inversion and the Bureaucratization of Risk

In the aftermath of the attacks of September 11th, 2001, to some, the sinister appeared to be lurking in the familiar. Water purification plants, power systems, ports, subways, and bridges—previously taken-for-granted systems tucked into the dull background of daily life—were now reprioritized as sites of critical vulnerability and possible terrorist attack. The seemingly endless string of so-called “soft targets” provided would-be terrorists with a host of potential opportunities and public officials with a daunting, near impossible, challenge. As the newly created Executive Branch Office of Homeland Security argued in its 2002 National Strategy for Homeland Security: “Our society presents an almost infinite array of potential targets that can be attacked through a variety of methods” and that “[t]oday’s terrorists can strike any place, at any time.”13 Policymakers quickly noted that the nation’s infrastructure systems—interconnected systems of communication, transportation, and power (among others)—offered malicious actors a prime target for causing destruction. These systems provide the foundation for economic, social, and political life. Additionally, the military relies on the availability and reliability of these networks for its activities. Disruption of these networks could result in grave damage to life and property, impair economic activity, and disrupt national security operations. Homeland Security Presidential Directive-7 (2003) succinctly states the case:

Terrorists seek to destroy, incapacitate, or exploit critical infrastructure and key resources across the United States to threaten national security, cause mass casualties, weaken our economy, and damage public morale and confidence…Critical infrastructure and key resources provide the essential services that underpin American society. The Nation possesses

---

numerous key resources, whose exploitation or destruction by terrorists could cause catastrophic health effects or mass casualties comparable to those from the use of a weapon of mass destruction, or could profoundly affect our national prestige and morale. In addition, there is critical infrastructure so vital that its incapacitation, exploitation, or destruction, through terrorist attack, could have a debilitating effect on security and economic well-being.\textsuperscript{14}

Government at the federal, state, and local level initiated new policies and procedures designed to protect key infrastructure systems against the perceived threat of ongoing terrorist action. In tandem, the industries operating or using infrastructures considered new practices in an effort to prevent additional attacks. In addition to a set of Executive Orders issued in the immediate aftermath of 9/11, both the PATRIOT Act of 2001 and the Homeland Security Act of 2002 focused on the protection of critical infrastructures as a key challenge in the domestic front of the war on terror.\textsuperscript{15} Together these acts elevated the protection of infrastructure networks to a national priority, called for increased cooperation and information sharing within and between the public and private sectors, and positioned the newly created Department of Homeland Security as the lead agency in devising a strategy for creating secure, resilient, and robust infrastructures.

The new legislation and executive orders, however, provided, at best, broad guidelines. The President’s \textit{National Strategy for Homeland Security} and \textit{National Strategy for the Physical Protection of Critical Infrastructures and Key Assets} both clearly articulated a preference for private, market-based, solutions to the challenges of


domestic security and voiced caution for direct government intervention in most cases.\textsuperscript{16} The \textit{National Strategy for Homeland Security} noted that it was U.S. policy to “[allow] the market to solve homeland security shortfalls whenever possible.”\textsuperscript{17} The \textit{National Strategy for the Protection of Critical Infrastructures and Key Assets} stated that in addressing the protection of infrastructures it was important to “[e]ncourage market solutions wherever possible and compensate for market failure with focused government intervention” and “regulatory directives or mandates should only be necessary in instances where market forces are insufficient to prompt the investments necessary to assure critical infrastructure … protection.”\textsuperscript{18} Yet, determining when such conditions actually existed and how, then, to blend the needs of security with the exigencies of ongoing operation remained a challenge fraught with difficulties. Identifying the cases where direct government intervention was needed, for the most part occurred outside of the direct ambit of DHS and within other regulatory bodies, Congressional deliberations, and through negotiations between vested stakeholders. Creating new security regulations and practices became a contentious site of debate; new initiatives undertaken in the name of security could, potentially, impose new administrative burdens and costs to the providers of services, as well as the users that had come to rely on the smooth, reliable, and cost-effective networks. Terrorism presented operators and policymakers with a difficult to define variable: the uncertain possibility of a future attack targeting networks that, in some instances, had never actually been directly targeted by previous terrorist


\textsuperscript{17} OHS. \textit{National Strategy for Homeland Security}. 4.

\textsuperscript{18} OHS. \textit{National Strategy for the Physical Protection of Critical Infrastructures and Key Assets}. 12.
aggression was not easily assessed. For some, the odious hassle of increased security was difficult to balance against a threat that might never materialize.

In this section of the dissertation, I consider the post-9/11 creation of new forms of control within the postal system, electric power grid, and freight rail transportation. The perceived threat of terrorism created a crisis of control for infrastructures generally and the above trio specifically. Here, the consequences of deregulation emerge as public problems. During this anxious moment, infrastructures ceased, if only briefly, being taken for granted and became visible, debated, and open to remaking. This crisis was two-fold: it concerned both how to combine the needs of security with the operational requirements of modern systems of communication, power, and transportation; and it concerned who would control the operation of these networks.

More clearly, post-9/11 debates concerning infrastructure security unfolded across two related axes. In the main, the perceived threat of terrorism introduced an operational challenge: the perceived threat of malicious actors targeting these vital systems called for new policies, technologies, and regulations that could reduce vulnerability and support resiliency, while still maintaining a high-level of service. Yet, at the same time these debates were about the proper relationship between the state and more-or-less private enterprises. The melding of national security, or “homeland security” as its domestic iteration was christened, with the workings of the private sector engenders conflict among infrastructure operators, particular industries that are large users of these services, and the government. Additionally, other groups that were on the periphery and/or

---

marginalized in relation to questions of infrastructure operation—including local
governments, environmental groups, and labor—saw an opportunity to advance or voice
their own perspectives and agendas. Drafting and implementing new regulations and
practices devoted to security challenged the autonomy of private actors and added new
costs and burdens. Seen from this perspective, post-9/11 debates concerning
infrastructure protection were not merely technical, but were also directly linked with the
larger question of the proper role of the state in infrastructure policy.20

In this section, I examine the bureaucratization of risk. I examine how the
uncertain and abstract threat of terrorism is written into stable and routine practices of
security expressed in new regulations, practices, and technologies. In examining the
creation of new forms of control, I examine how, by whom, and to what end new forms
of security are constructed. I consider three separate instances where national security
and infrastructure operation intersect. I focus on the creation of bioterrorism security
practices for the postal system in the wake of the anthrax attacks of 2001, the crafting of
erouting standards for the transportation of toxic materials across the freight rail
network, and the creation and implementation of the first mandatory cybersecurity
standards for the bulk electric power grid. After the anthrax attacks of October, 2001

20 To a degree, the challenges presented by contemporary infrastructure security are a recurrent theme in
U.S. history. In the U.S., infrastructures have long been viewed as a key component of national security.
However, the aims of national security have mainly been marshaled to support developmental projects and
the creation of new systems. Indeed, in communications, radio technology, satellites, computers, and what
would eventually morph into the internet all were given critical early and direct support from the
government under the guise of furthering the aims of national security. Famously, the national interstate
highway system was created in part as a means of creating a quick channel for the circulation of troops and
supplies in the event of a war. Hence its formal name: the Dwight D. Eisenhower National System of
Interstate and Defense Highways. During World War II the federal government used its powers to create
new power plants to support production for the war effort. In the wake of 9/11, however, the challenge was
not focused on creating new or parallel systems but on revisiting systems already in place.
Postal planners attempted to create new security practices to combat future instances of biological terrorism. Eventually, two new information and communication systems were interlaced within the operation of the postal network: the biohazard detection system (BDS) and intelligent mail. Both systems promised greater control and increased surveillance over certain portions of the postal network. The freight rail network was also seen as a possible target of terrorism. New rerouting regulations for the transportation of toxic materials via rail were debated in response to the possible threat of terrorism and eventually enacted by the Department of Transportation, Pipeline and Hazardous Materials Safety Administration in 2008. Finally, the Energy Policy Act of 2005 gave the Federal Energy Regulatory Commission new powers to create and enforce mandatory cybersecurity regulations for the bulk electric power grid. These new regulations were drafted over a period of several years and finally implemented in 2008. In each instance, new practices and regulations were contested by numerous parties and attempted to address structural vulnerabilities linked with deregulation.

The theoretical work of Ulrich Beck anchors this section. Beck argues that risk can, in some instances, draw attention and politicize domains typically considered to be outside of legitimate political discussion. Additionally for Beck, certain categories of risk—forms of uncertainty that are both publicly salient or available but that resist easy quantification—are also sites of new forms of political engagement. For Beck risks are neutral; in and of themselves they do not compel anything. Rather, organizations,

---


22 Beck’s notion of risks is more closely aligned with what Knight terms uncertainty. See, Knight. *Risk, Uncertainty, and Profit.*
individuals, or coalitions of the concerned can take an entrepreneurial role and seize the possibilities offered by public attention and anxiety to engage in political action.

Crucially, such engagements hold the possibility of challenging the status quo and upend familiar loci of power, including prevailing political, economic, or scientific elites.

Risks, then, can serve a political role; they can politicize domains that are often seen as being outside of the realm of politics, while also providing a catalyst for new political collectives and interventions. In this reading, widespread risks do not represent the foreclosure of democratic practice, but offer its possibility.

Terrorism, post-9/11, surprisingly, operated in this fashion. Infrastructures became central to ongoing political debate and new interventions. Risk served to enact a type of inversion: it foregrounded the importance of systems that were typically held in the background. In post, electric power, and freight transportation, the threat of terrorism either drew new parties into the deliberation about infrastructure operation or potentially enabled groups already involved in their operation to a greater degree. In each case, new practices of security—regulations, technologies, and polices—were created and implemented. This section of the dissertation examines in detail these new interventions.

This section of the dissertation is divided into four chapters:

Chapter 4: Risk, Politics, and Culture: Risk Society and Reflexive Modernization

This Chapter provides introduction to the work of Beck and the treatment of risk as a political resource. Beck argues that we are entering a period of “reflexive modernization” where the successes of modernization are now creating novel forms of risk that, in certain circumstances, can foster new types of political engagement. As risks
become political problems, they can potentially lead to reflection: a turning back to consider the origins and production of risk. It is within this moment that Beck locates the democratic possibilities of risk. For Beck, the political volatility of certain categories of risks—what makes them potentially transformative political resources—is inseparable to their positioning as cultural artifacts.

In this Chapter the notion that risks can be catalytic and serve as a resource for democratization will be detailed and explained. Beck works within and against a number of different traditions. Most significantly, he presents risks as both socially constructed and as having a material reality. For Beck, the social and political implications of risks are always an empirical question. To understand if risk supports democratization or, alternately, the reification of power, it is necessary to examine how, by whom, and to what end risk is marshaled.

Chapter 5: Delivering the Mail: Terrorism, Efficiency, and the Politics of Post-Anthrax Bioterrorism Security

In this Chapter, I examine the 2001 anthrax attacks and the creation and implementation of new forms of postal security designed to combat biological terrorism. The Chapter considers first how the anthrax attacks created a crisis of control that destabilized the postal network, and next turns to examine the creation of two new information and communication systems, the Biohazard Detection System (BDS) and Intelligent Mail, employed in an effort to create a “secure” network. The attacks revealed the limitations and drawbacks of deregulation and, for a moment, served to democratize postal governance. The attacks highlight the ways in which the pursuit of new
efficiencies enabled by regulatory reform created not only new efficiencies within postal operations but also new forms of vulnerability. At the same time, the anthrax attacks briefly served to democratize the governance of postal service. The threat of biological terrorism, perhaps somewhat counter-intuitively, fostered the democratization of postal governance. Yet, while diverse postal stakeholders were given a voice in drafting new security protocols, the prevailing hierarchies of power that dominate postal politics—the centrality of large volume mailers and the ongoing subordination of labor—are ultimately encoded within the organization and operation of new control technologies designed to counter bioterrorism.

*Chapter 6: Green Security? The Environmental Movement, the Transportation of Hazardous Materials, and the War on Terror*

In this Chapter, I examine the creation of new regulations governing the transportation of hazardous materials via rail. In the wake of 9/11, a coalition of environmental activists and local public officials appropriated the threat of terrorism to advocate new stringent hazmat regulations over the objection of the rail industry, the chemical lobby, and the Bush Administration. Activists were able to appropriate the vernacular of security—seize on the salient though difficult to define threat of terrorism—in order to substantively change transportation policy. Here, I examine the fight for new rerouting legislation through regulatory hearings, the courts, and Congress.

*Chapter 7: The Creation of Mandatory Cybersecurity Reliability Standards: Autonomy, Interdependency, and Electric Power*
In Chapter 7, I examine the creation of the first set of mandatory cybersecurity reliability standard for the electric power grid. In this Chapter, I examine the post-9/11 redefinition of “security” within electric power and the battle between industry and federal regulators over the authority to define new mandatory cybersecurity regulations. Though the reliability development and approval process is structured to support the autonomy of industry to devise standards as they see fit, federal regulators work to carve out a significant role for federal oversight of cybersecurity. In the process, regulators, quite surprisingly reject industry’s preference for lax administrative regulations and, instead, work to secure comprehensive cybersecurity standards.

Conclusion: The Politics of Critical Infrastructure Protection

The concluding Chapter provides a brief review of the central findings and claims of the dissertation. In concluding, the dissertation considers the relationship between deregulation and infrastructure vulnerability, as well as post-9/11 efforts to create new forms of security. Here, I consider the overlooked benefits of regulation in providing resilience and the tension between economic deregulation and vulnerability. I also consider the way in which, perhaps quite surprisingly, the perceived risk of terrorism operates as a resource of democratization within infrastructure governance. Following the earlier discussion of Beck, the dissertation concludes with a consideration of how, by whom, and to what end new forms of infrastructure security have been created.
Chapter 2: Stumbling Toward Resilience: Revisiting and Rethinking Infrastructure

Regulation

Before the 1970s, price-and-entry regulation governed the operation of infrastructures in the United States. Though regulations functioned in slightly different ways across different sectors, they generally created stable monopolies or cartels and engendered support for a host of socially valued policy goals, such as the extension of services and non-discriminatory rates, through imposed positive obligations and control over rates. Economic regulation, however, also provided an overlooked benefit: resilience. Regulation militated not only against the deleterious effects of concentrations of economic power, but also against dangerous concentrations within the architecture of infrastructure services and systems. Price-and-entry regulation, both directly and indirectly, supported linear organization and loose coupling by promoting distributed control, redundancy, and diversity. These qualities provided a hedge against large-scale and systemic failures. Regulation pushed development in directions that were at odds with a narrow conceptualization of economic efficiency but, at the same time, supported resilient networks: in the postal network processing and distribution of the mail—command and control functions—were dispersed throughout the network, rather than centralized in a few large facilities; the regulation of the railroads discouraged the transportation of large, bulk, concentrations of hazardous materials through the rail network in favor of smaller shipments carried by truck and, at the same time, supported a redundant rail system that allowed for the transportation of hazardous materials on paths that skirted densely populated areas; the regulation of electric power supported
distributed control within territorial monopolies and cultivated command and control information and communication technologies. These qualities were, for the most part, unintended byproducts of regulation. Later, during the political upheavals of deregulation, it was these very qualities that came under attack as being elements of wasteful inefficiency and, to some, which signaled the need for reform. Ironically, these perceived failures of regulation provided, to differing degrees, a margin of security and safety that post-9/11 infrastructure interventions now seek to recapture.

*The Importance of Politics: Infrastructures as Socio-Technical Systems*

Politics play a key role in shaping the architecture and operation of infrastructure networks. Infrastructures are what Thomas Hughes describes as “socio-technical systems”; they are more than just compilations of hardware and technical components.¹ Rather they are made of the intersection of technologies, politics, economics, and culture. They do not spring forth according to the immutable laws of economic efficiency or technological determinism, but instead are shaped and molded by the interplay of competing sets of historical forces that ultimately give them form. Politics and regulation play central roles in defining and shaping infrastructures; they compel and encourage certain configurations, while forbidding or discouraging others. It is only through a historically-informed understanding of the particularities of regulation that we can make sense of why infrastructures appear as they do. Only once we have foregrounded regulation can we understand why, for example, the Post Office continued to rely on

---

manual sorting long into the 1960s after more efficient automated equipment became available; or, why for decades the cost of shipping a ton of grain via rail was far less expensive than shipping a ton of industrial chemicals; or why electric power systems were organized into a set of local and regional systems, rather than larger configurations. Without examining the regulatory process, these and many other “quirks” of infrastructure design and operation remain difficult to parse. To understand how and why infrastructures adopted resilient forms, and later turned toward increasingly vulnerable configurations, it is crucial to start by examining the role regulation played in directly and indirectly shaping these systems.

*Normal Accidents: Complex Networks, Tight Coupling, and Vulnerability*

Before moving to discuss how regulation supported the construction and maintenance of infrastructure systems that were, in key respects, impervious to large-scale vulnerabilities, it is first necessary to provide a conceptual framework for assessing the characteristics that define resiliency and vulnerability of large technical systems. The work of Charles Perrow on normal accidents offers a useful point of departure for considering the properties of infrastructures that make them open to systemic or large-scale failures. Perrow’s framework is elastic: it is a general “theory of systems” that can be used to examine disparate systems; and its central categories can accommodate the interesting theoretical insights from the recent application of graph theory from the fields of mathematics and computer science to the analysis of infrastructure vulnerability.²

---
Importantly, Perrow’s work makes explicit the connection between institutions, technological structure, and power that is either absent or underplayed in other analyses of the vulnerability of technical systems but which is indispensable to understanding the larger causes of vulnerability.

Perrow’s work on normal accidents initially derived from an examination of industrial accidents before being expanded to assess the impacts of natural disasters and terrorism. As Perrow notes, the same qualities that render systems prone to accidents also increase the possibility of failures from natural disasters and malicious, intentional, disruption from acts of sabotage or terrorism.\(^3\) Perrow offers a novel interpretation of what causes accidents. Rather than seeing large-scale failures as simply the product of operator error or caprice, he redirects attention towards the different ways that system architecture can create the conditions for large-scale failure.\(^4\) Perrow labels these types of failures “normal accidents” because the roots of failure are embedded within the structure of the system itself.

Perrow’s framework revolves around two central concepts: complexity and coupling.\(^5\) For Perrow, differing degrees of complexity or coupling within a system

---


\(^4\) In this fashion, Perrow’s understanding is in line with treatments that consider technologies as containing a set of affordances that make different types of activities more or less likely. See, Yochai Benkler. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven, CT: Yale UP, 2006.

determines the difference between resilient systems that can tolerate disruptions and systems prone to large-scale failures. Complex, as opposed to linear, systems are defined by reliance on common mode components linking multiple different elements within a system; or by the proximity between different, unrelated, systems. Complexity creates unexpected interactions that can lead to systemic failures. The failure of common mode equipment, for example, does not lead to a simple, localized, serial failure in the next step of a sequential process, but can have cascading failures across the system. In linear systems, components are organized in sequence such that a failure of any single component simply breaks the chain (think here of a simple assembly line system) and causes a minor disruption until the failing component can be replaced. In complex systems, however, such a failure is much more consequential and goes beyond a minor disruption. The failure of a common mode connection linking different portions of a process, such as interconnected computer systems controlling billing, production, and distribution for a manufacturer, not only causes a delay in one small isolated or containable area, but rather spin outward in unexpected ways crippling various interconnected areas of operation and leading to a systemic-failure. Likewise, proximity can also create complexity. The proximity of different systems or sub-systems can lead to unexpected interactions across different systems. Proximity fosters interconnection between unrelated systems whereby a failure within one system or sub-system jeopardizes the formally distinct but spatially near system; failure in system X causes a

---

6 Perrow also discusses the role that indirect information can play. Perrow. *Normal Accidents*. 73.
7 Ibid. 72-84.
8 Ibid.
disruption in system Y, simply by virtue of co-location.\textsuperscript{9} The close proximity between massive oil drilling systems and fragile and volatile bodies of water, for example, provides a vivid and all-too-readily available example of the pitfalls of proximity. The failure of oil extraction equipment can destabilize and damage the environment within or near which it resides; at the same time, the volatility of the body of water can lead to the failure of the equipment. Proximity, as this simple case highlights, can lead to unexpected interactions across systems with large consequences. Proximity between different systems increases complexity in that it injects another variable, the action of the adjacent system, for which the neighboring system is not designed to cope.

Perrow’s second key concept, coupling, refers to the degree of slack in a system. Perrow offers a distinction between loose and tightly-coupled systems.\textsuperscript{10} Tightly-coupled systems are characterized by time-dependent processes that cannot be slowed down, invariant operation, and little independence of components.\textsuperscript{11} The strict conditions of time-dependence eliminates any moment for the system to regroup, while invariant operations do not admit substitutions or “ad hoc” changes. But most significantly, tightly-coupled systems favor centralized and rigid control aggregated at higher levels as opposed to distributed control spread-out through the system. This centralization of control limits the adaptability of systems by removing the ability of components to act independently in the face of changing conditions. Under this model, disruptions that undermine centralized controls create system-wide failures specifically because the

\textsuperscript{9} Ibid.
\textsuperscript{10} Ibid. 81-100.
\textsuperscript{11} Ibid. 93-94.
components of the system are not able to exercise flexible control; once the system is disrupted or centralized controls are undermined, the system fails.\textsuperscript{12}

Underlying Perrow’s discussion of linear/complex and loose/tightly-coupled systems is a skepticism and antipathy toward dependences. Echoed in his historical work on the rise of organizations, Perrow sees increasing dependence both within and between systems, inter- and intra-dependences, as offering new possibilities for failure.\textsuperscript{13} The benefits of increasing complexity and tightly-coupled systems are, in some cases, plain: they require less space, resources, and appear, if the costs and consequences of failure are ignored or overlooked, to be more economically efficient.\textsuperscript{14} Yet, dependences imperil the ability of systems to absorb the shocks of small failures and, instead, create conditions conducive to systemic failure.

Recent work approaching the analysis of infrastructure vulnerability through a focus on network structures fits well within Perrow’s general framework.\textsuperscript{15} Such work highlights the vulnerability associated with network structures containing dense hubs or nodes. Scale-free and hub-and-spoke networks characterized by a limited number of dense nodes connected outward with scattered points also create the conditions for

\begin{itemize}
\item \textsuperscript{12} Ibid. 81-100.
\item \textsuperscript{14} Perrow. \textit{Normal Accidents}. 88.
\end{itemize}
systemic failure. Under normal operating conditions, these networks work well, failures or disruption at the margins impacting any of the many scattered and loosely connected nodes are of little consequence and create a localized failure. But if a dense hub fails the larger system is destabilized. Within these networks, dense nodes shoulder an extra share of the freight, making them critical to overall system performance and reliability.\textsuperscript{16} The dense nodes of network analysis are, ultimately, little different from Perrow’s reading of common mode equipment and complexity. Perrow’s framework, however, aids an informed understanding that is largely absent within network analysis. His discussion of proximity underscores the importance of understanding the interaction between and across system; likewise, his interest in the relationship between politics, power, and technology supports analysis that moves beyond a discussion of networks as abstract geographies characterized by links and hubs, to a more complete understanding of the roots of failure.

This brief overview of Perrow’s use of complexity and coupling will be fleshed out below as the key concepts are applied to the organization of the postal system, freight rail network, and electric power system. Perrow’s work is inherently qualitative; it is not a rigid formula for diagnosing resilience and vulnerability, but rather provides a useful working vocabulary to examine different systems and their changes over time. It is important to bear in mind that Perrow’s categories—the space between linear and complex, loose and tight coupling—are distinctions of degree, not kind.\textsuperscript{17} A close look at

\textsuperscript{16} Albert, Jeong, and Barabási. “Error and Attack Tolerance of Complex Networks.”
\textsuperscript{17} Perrow. Normal Accidents. 78.
different systems can locate them on a plot between these different points, but it does not lead to absolute, either/or distinctions. Yet, the consequences of where on this sliding scale a system falls are crucial; the distinction can mean the difference between an accident and a catastrophe, a local emergency and a national crisis.

*The Postal System: Regulation and Distributed Control*

Historically, the regulation of the postal system supported a network structure characterized by dispersed command and control—a loosely-coupled network with decentralized, labor intensive, information processing—and worked against the widespread adoption of internal improvements afforded by new information processing technologies, such as mechanization and automation, that would centralize control. Congressional oversight of postal rates and budget favored short-term planning and fixes at the expense of long-term projects and investments. As a consequence, the regulatory structure restrained modernization efforts seeking to remake the postal network into a tightly-coupled system. The structure of regulation worked to obstruct the Post Office from investing in research and development and the acquisition of new sorting and processing technologies that would eliminate distributed processing and cut the cost of processing the mail, increase throughput, and lessen the need to continually add new workers to process growing mail volume. In this fashion, the architecture of the network embodied the structure and particularity of postal regulations: the outlines of the network, its vagaries of form and durable oddities, reflected the structure of regulation.
In many ways, the support for distributed processing and the barriers presented to the adoption of mechanization and automation were side-effects of a regulatory regime that supported other sets of benefits. Historically, postal regulation undergirded a thriving national arena of communication and commerce and designed into the system of rates and classifications broad support for public information—newspapers, magazines, and the correspondence of public officials, most prominently—while extending service across the nation.

**National Operation as Regulation: The Structure of Postal Regulation**

In the United States, postal service, unlike other network infrastructures, developed as a state-run enterprise. Postal service, unlike private systems of rail and electric power, was not subject to oversight by independent regulatory commission, but directly operated by the federal government through the Post Office Department. In keeping with the tradition of state-run postal systems, the Articles of Confederation provided the Continental Congress with the exclusive right of establishing and regulating post offices. The adoption of the Constitution, likewise, did not introduce any significant discontinuity. In language similar to the Articles of Confederation, Article 1, Section 8 of the Constitution reserved for Congress “the Power…to establish Post Offices and post Roads.” It was not until the landmark Postal Act of 1792, however, that U.S. postal

---

18 From the 1790s-1820s the postal system was operated by the General Post Office. In the 1820s, the General Post Office was renamed the Post Office Department. For simplicity, the Post Office Department will be used throughout. James. W. Cortada. *The Digital Hand: How Computers Changed the Work of American Public Sector Industries*. Vol. III. New York: Oxford UP, 2008. 164.

policy would find clear expression and definition: the Act clarified the power of the federal government to operate the postal service by making it a civil offense to conduct private postal services in direct competition with the Post Office; codified the Congressional role in establishing new routes; established a new structure of rates benefiting public information; and included strict protections against opening and tampering with the mails. With the Postal Act of 1792, U.S. postal policy sparked what historian Richard John identifies as a “communications revolution” within antebellum America that led to the wide-distribution of public information and the extension of service across the nation.20 Both Federalists and Jeffersonian Republicans supported the new emphasis: Federalists seeing it as a useful tool for sowing the bonds of national community across a geographically expansive area; Republicans seeing it as a means of keeping tabs on the workings of the remote central government.21

Though postal service operated under a different institutional structure than other regulated infrastructures, state operation and regulation ultimately functioned in a similar fashion. Indeed, public operation functioned as a type of regulation: it limited

---


competition within a defined sector of the economy; it controlled rates not through the market but through administrative processes; legal monopoly protections supported cross-subsidies in support of the public good, here defined as broad access and discounts for public information; and the Post Office Department retained some power to define, enforce, and adjudicate complaints regarding rules of service.22

Power to oversee the operation of the Post Office was shared between the Executive Branch and Congress. The President played a key role in the operation of the Post Office through the appointment of the top staff of the Post Office Department—the Postmaster General, Deputy Postmaster General, and Assistant Postmasters General—and through the appointment of local postmasters.23 The scope of the postal monopoly and the terms on which private firms could engage in postal services were fixed by federal law under the Private Express Statutes.24 The postal monopoly, generally, only extended to prevent the for-profit routine carriage of letters over postal routes and functioned as a revenue-protection measure to ensure the financing of the postal system.

---

22 For example, the Post Office exercised limited discretion concerning classification decisions and determinations concerning whether or not particular articles could be entered into the mails. See, Dorothy G. Fowler. Unmailable: Congress and the Post Office. Athens, GA: University of Georgia Press, 1977; Kielbowicz. “A History of Mail Classification.”

23 The Postmaster General was a key political appointee and a served as a member of the cabinet between 1829 until reorganization in the 1970s. The President controlled the appointment of local postmasters, key sources of political patronage, both through direct appointment and indirectly through the Postmaster General. John. Spreading the News. 206-256; Wayne E. Fuller. The American Mail: Enlarger of the Common Life. Chicago: Chicago UP, 1972. 285-330.

as a whole and the maintenance of cross-subsidies for valued forms of content.\textsuperscript{25}

Crucially, Congress held the power to define the establishment and extension of routes—determine the scope of the postal network—establish different classes of mail, set wages for employees, and fix rates for the various classes. In these ways, Congress set the basic terms of service for the Post Office and controlled both key elements of cost (including the scope of service and compensation) and revenue (through the fixing of rates).

Congress also exerted direct control over postal service through the appropriations process. Once annual appropriations to offset postal deficits became common and central to postal revenue after the 1820s (see below), the role of Congress in postal affairs expanded. Contingent upon Treasury financing to offset annual losses, Congress scrutinized and examined the operations and accounts of the Post Office, including major capital outlays, and the Post Office was compelled to seek approval for large projects and undertakings.\textsuperscript{26}

\textit{Regulation and Distributed Processing: Congressional Control, Information and Communication Technologies, and Network Organization}

Postal regulation before the beginning of regulatory reform in 1970 subsidized a network that, seen through a narrow lens, contained structural inefficiencies; it supported a loosely-coupled network with distributed control rather than a “small-world” network

\textsuperscript{25} The postal monopoly changed over time, admitting a number of exceptions at various points. Exemptions included: non-commercial carriage; the carriage of pre-paid stamped mail by third parties; special messenger services; letters accompanying cargo. The postal monopoly also did not cover non-letter mail such as newspapers, or postal service in areas where the Post Office did not operate. USPS. “Report on Universal Service.”

with control concentrated in a few dense hubs. The Post Office, between World War II and 1970, lagged behind other large-scale, labor-intensive, operations in terms of adoption of new technologies and, instead, relied on increasing labor to offset growing postal volume. While comparable industries and firms adopted new mechanized and, later, automated, technologies, the Post Office continued operating as it had for nearly a century, relying on hand processing and an army of trained clerks to continually route and sort the mail as it passed through the network. The continuation and durability of antiquated methods of processing and distributing the mail were linked to the broader structure within which the postal system operated. The Post Office did not have the resources needed to pursue the widespread adoption of new ICTs that would consolidate and streamline the postal network. As in other infrastructures, the architecture of the network embodied the particular regulatory regime under which it operated.

In the century before regulatory reform the work of sorting and processing the mail changed remarkably little. Clerks and carriers situated throughout the postal network routed and sorted the mail repeatedly by hand based on complicated memorized “schemes” delineating postal routes and transit points. The Post Office initially tested mechanized equipment for handling and sorting the mail as early as the 1920s and it was available by the 1950s; automated equipment followed during the mid-1960s. Yet,

---

27 Small-world and scale-free networks exhibit clustering within which a few central nodes or hubs densely connect many other points in the network. For a discussion of the differences between different network variations, see Gorman. *Networks, Security and Complexity.* 34-40.

28 Mechanization and automation targeted each of the different stages of mail processing, including: culling, separating letters into various sizes and classes; facing, placing the letters in the same direction so that they can be read; canceling the postage; sorting, separating the mail into clusters for the next step of distribution and transportation. USPS. *United States Postal Service: An American History, 1775-2006.* May
postal workers remained largely unaided by mechanical or automated equipment before reorganization in 1970; rather they processed letters by hand and sorted mail by adopting the “peek and poke” method, “peeking” at the address then “poking” letters into designated bins or cases for transfer based on memorized schemes. The President’s Commission on Postal Organization charged with reviewing postal operation, observed in 1968:

In most offices men and women lift, haul and push mail sacks and boxes with little more mechanical assistance than the handcart available centuries ago. In the electronic era, the basic sorting device remains the pigeonhole case into which letters are placed, by hand, one by one. The basic parcel post container is the canvas sack, filled, lifted and “dumped” by human labor… the Post Office relies almost totally upon human labor for bulk materials handling and distribution—the transfer of mail between work stations, and the loading and unloading of vehicles.29

In 1969, a year before the passage of regulatory reform, a stunning 90% of all mail was still processed by hand and less than 5% of the largest 300 postal facilities were mechanized.30 The average letter traveling through the postal network was sorted and resorted between 10 and 18 times, in 5 to 6 different post offices, before reaching its destination.31 Five hundred and fifty-two sectional sorting centers, large facilities linked by air routes, served as the backbone of this system, connecting regional, city, and branch post offices.32 In manual sorting, clerks and carriers perform informational work based

---

31 POD. *Postal Progress.* 5.
32 The Post Office introduced the Sectional Center System in 1963 in conjunction with ZIP codes as a means of reducing handling costs. Yet, the postal system still relied on distributed processing throughout
on memorized data sets. Under this model, processing could not be usefully condensed or centralized within a limited number of dense postal hubs. The reliance on workers’ manual ability and memorized information for routing and sorting limited the flow of throughput and the number of distribution points that any single facility could serve. Manual sorting placed a limit on the centralization of control. As a result of the reliance on manual sorting, the postal network pre-reorganization was characterized by distributed processing and sorting; letters would have to be continually sorted and resorted into more fine-grained levels of specificity as they moved through the network.

The Post Office confronted increasing mail volume, particularly the sharp spike post-World War II, by adding clerks to the rolls in increasing numbers; as postal volume grew sharply through the 20th century, so too did the number of clerks employed to sort the mail (see Table 2-1). Between and 1940 and the 1968, both total mail volume and clerks employed roughly tripled: total volume jumped from 27.7 billion to 79.5 billion pieces, while clerks employed rose from 97,000 to 308,000. The increase in clerks far outpaced total growth in postal labor: in 1930, clerks accounted for 27% of total workers; by, 1968, clerks, the central figures in routing and sorting the mail, accounted for 42% of

---

33 Clerks during the 1960s could reliably sort letters for roughly 55 different postal facilities at a rate of less than 1,000 letters per hour. During the same period a team of clerks using advanced mechanized multiple position letter sorting equipment could sort 28,000 letters per hour into 277 different separations. Baxter. *Labor and Politics*. 70-71.

the total postal workforce; between 1930 and 1968 total non-clerk postal labor increased by 71.6%, while clerks over the same period increased by 230.5% (see Table 2-2).\textsuperscript{35}

New forms of equipment that came on line during the 1950s and 1960s held out the prospect of a postal system where fewer workers were involved in sorting more mail, for a greater range of destinations, more quickly and cheaply. Mechanization and automation offered a more economic and long-term solution to the problem of processing the mail by supporting labor reductions and the concentration of sorting and processing in large centralized hubs (in order to capture the benefits afforded by economies of scale). New technology offered the possibility of largely deskilling the process of sorting and processing the mail, while expanding throughput and decreasing unit costs.\textsuperscript{36}

Mechanized and automated equipment simplified the work of clerks by reducing the cognitive burdens associated with sorting the mail. By embedding such knowledge in the equipment, a rotating cast of comparatively deskilled (and cheaper) workers could ably complete the same tasks. By the 1960s, mechanized and automated equipment offered significant savings: mechanized equipment used in culling, stacking, and cancelling mail cost $1.20 per thousand letters, compared against $3.50 per thousand for manual processes; automated processing and sorting operated at $2.00 per thousand letters, mechanized equipment operated at $3.42 per thousand letters, and manual sorting and processing at $4.20 per thousand letters.\textsuperscript{37} The difference in throughput reveals similar


\textsuperscript{36}See, Baxter. *Labor and Politics*.

disparities: the average clerk processed 600 letters per hour; workers aided by mechanization averaged 2,800 letters per hour; while automated equipment processed 30,000 letters per hour. Additionally, as noted above, mechanized and automated equipment allow mail to be sorted and divided into a greater number of divisions—before mechanization clerks sorted mail into 55 distinct separations, while mechanized equipment allowed for 277 divisions—and eliminates the need for multiple sorts throughout the network. Indeed, the savings made possible by mechanized and automated equipment are only possible if economies of scale are captured by lowering unit costs by aggregating large volumes of mail in central locations.

The regulatory structure of the Post Office, however, actively frustrated the widespread adoption of new technologies and, in its place, supported the continuation of the status quo. At the most general level, the postal model of monopoly protection and deficit financing through annual appropriation buffered the Post Office from competitive pressures. The private express statutes reserved for the Post Office a guaranteed market which, combined with the regularity of annual appropriations, reduced the incentives faced by management to pursue the most cost-effective means of processing the mail. Appropriations became common after the 1820s and offset roughly 15% of total postal

---

40 Ibid.
41 The contention that the political structure of postal service slowed the pace and terms of adoption of new technologies, detailed below, is central to the observations of the President’s Commission on Postal Organization and is echoed in Baxter. See, President’s Commission on Postal Organization. *Towards Postal Excellence*; Baxter. *Labor and Politics*. 
costs in the following decades. In the late 1960s, a Presidential Commission formed to study the Post Office drew an explicit connection between the appropriations and inefficiencies, remarking that: “the practice of making up losses from the Federal Treasury removes much of the incentive for efficient operation. There is no need to control costs if a supplemental appropriation may be expected as a matter of course.”

While public pressure historically pushed the Post Office to innovate in providing new services for consumers, like rural free delivery, a comparable movement to push the Post Office to pursue internal improvements did not develop; mechanization and automation were “back of the house” issues that did not, directly, touch the experience of the mailing public. The incentives to innovate to reduce operating costs were somewhat muted.

By the 1950s and early 1960s, however, postal management actively sought to innovate and transform the postal network through mechanization and automation as strategy to reduce costs. Repeatedly, in Annual Reports and Congressional testimony, the Postmaster General underlined the importance of investing in new technology and called for a strong commitment to support postal efforts to implement “postal modernization,” a phrase synonymous with mechanization and automation. The lack of

---

42 Between 1838 and 1968, postal revenues were in the red for all but 17 years. The amount of the deficit fluctuated from year-to-year but the President’s Commission calculated that as of 1968, the cumulative postal deficit accounted for 15% of total postal costs. During the 1960s, shortfalls were, however, growing; in 1967 the deficit was $1.17 billion, or 19.1% of the total postal budget. President’s Commission on Postal Organization. Towards Postal Excellence. 22.
43 Ibid. 37.
innovation, then, was not entirely or perhaps even centrally a question of missing
incentives or will on the part of postal management, or a side-effect of a lazy monopolist
or complacent public agency, but rather was tied to the interplay between Congressional
oversight of postal finances and the resources required for the adoption of new capital-
intensive technologies. Recall that Congress controlled postal finances both by setting
rates and, more important, by controlling the postal budget through the appropriations
process. The Post Office had little discretion or control over how collected revenue was
used: annually, the revenue collected by the Post Office through rates was deposited in
the Federal Treasury; at the end of the year Congress appropriated the collected revenue
with an additional sum to close the accumulated yearly deficit back to the Post Office
with binding limits on how the funds were to be used. Control over the different aspects
of postal finances, rates and appropriations, were fractured among different
Congressional committees with little coordination; the House and Senate Committees on
the Post Office and Civil Service set postal rates, while the Appropriations Committees
allocated funds for the Post Office.46 Congressional control worked to maintain the status
quo and keep the Post Office afloat, but it also starved it of the capital needed to engage
in postal modernization by keeping rates low and by refusing to provide appropriations to
cover expensive long-term investments.47

46 On this point, see Arthur D. Little. “The U.S. Post Office and Organizational Options for its
Improvement.” President’s Commission on Postal Service. Towards Postal Excellence. Report of the
47 President’s Commission on Postal Service. Towards Postal Excellence. 35-36; Fuller. The American
<www.prc.gov/tsp/55/enterprise.pdf>. 63-68.
Historically, postal rates under Congressional control stayed flat and did not rise to meet expanding costs. Once established, rates developed an inertia that was difficult to dislodge: Congress was loath to undertake the unpopular and highly visible step of passing legislation increasing postal rates and the industries reliant on cheap postage to distribute books, magazines, catalogues, and other goods actively lobbied for their continuation at the expense of growing revenue shortfalls.\(^{48}\) For their part, the Post Office, too, was sensitive to the political machinations required to raise rates and would frequently refrain from requesting a general rate increase during an election year or in close succession with a previous rate hike, regardless of actual need.\(^{49}\) As a result, postal rates did not rise to keep pace with the rising postal deficit. For example, first class postage was only adjusted seven times in the one-hundred years before 1968, while the majority of all second class mailings traveled through the post at rates of less than 2 cents in 1968, roughly the same rate they would have been charged over a century previous.\(^{50}\)

Collected cost data played a minor role in establishing postal rates and were routinely

---


ignored—trumped by politics—during the legislative process (see Table 2-3 and Table 2-4).

Most significantly, Congress wielded the appropriation process to exert tight control over how the Post Office allocated its resources. The Post Office, unlike private industry, did not have discretionary authority over its collected revenue. Rather, collected postal revenue was funneled through the Treasury and appropriated by Congress back to the Post Office, with added sums to bridge the accumulated deficit. As part of this process, Congress reviewed major proposed projects and appropriated the postal budget into eight different categories (see Table 2-5) between which the Post Office could not, by penalty of law, transfer funds. Though appropriations rose above total revenue to offset the deficit, the drain imposed on postal finances through low rates did have a real impact on total funds available and, consequently, efforts to modernize. The Appropriations Committees used collected revenue as a baseline when determining the total amount of funds to appropriate and attempted to limit, to the greatest extent possible, the level of appropriated funds above collected revenue. In defining the postal budget, Congress routinely pursued a strategy of incremental increases in funds—minor year-to-year increases—as opposed to major new allocations. The Post Office held a legal monopoly that granted it nearly complete market control, yet, by virtue of

---

51 Before the inauguration of the Cost Ascertainment System (CAS) in 1926, the Post Office did not collect comprehensive cost data on its services. Even after the introduction of CAS, however, data on cost was disputed and ignored as Congress fixed rates. See, Foster Associates. “Rates and Rate-Making.” 2-8 – 2-13.

52 Before 1951, funds were made in 58 separate appropriations. Legislation in 1951 simplified the appropriations structure into 8 accounts. The Post Office, before 1968, did have limited authority to transfer funds between accounts in amounts that could not exceed more than 2.5% of either donor or acceptor appropriations. Arthur D. Little. “A Description of the Postal Service Today.” 6-33 – 6-34.

53 Ibid. 6-87 – 6-88.
Congressional control over rates and budget, it had a paucity of resources to invest in internal improvements.

In reviewing postal finances, Congress concerned itself most centrally with figures associated with inputs, cost data, at the expense of more thorough studies of output and efficiency.\textsuperscript{54} As a result, appropriations were overwhelmingly assigned to operations, labor and other day-to-day costs, while all but ignoring investments in long-term capital projects, such as building new postal facilities or acquiring new equipment, or investment in research and development that offered upfront outlays and delayed benefits. The refusal of Congress to provide such funds reflects both the general challenges associated with elected public officials and projects with delayed benefits, as well as circumstances unique to the Post Office. Congressional reluctance to look beyond the narrow exigencies of the moment and engage in long-term planning stems in part from the familiar disjunction of time horizons confronting members of Congress and the benefits of long-term investment. Members of Congress faced the immediate pressure to maintain postal service \textit{and} to limit federal postal spending relative to other pressing national needs. Investing in R&D and new postal technology, however, only yields gains in later years—as investments eventually reduce unit costs and take a bite out of postal expenses—when the members of Congress that initially supported the investment would perhaps be out of office.\textsuperscript{55} At the same time, postal labor, a significant

\textsuperscript{54} Ibid. 6-30.
and well-organized interest group, actively lobbied against investment in new technology.\textsuperscript{56} Though labor publicly supported the aims of mechanization and automation, they moved to obstruct the process in order to prevent job reductions.\textsuperscript{57}

The figures concerning Post Office investment in developing new technologies and funding large capital projects under Congressional control are telling. After WWII, the Post Office created a Research and Development (R&D) Bureau to help spur innovation in mail operations. Yet, even after the beginning of R&D spending in 1950, funding remained nearly microscopic relative to the size of postal operations. Between 1950 and 1970 (before reorganization), the Post Office, at the instruction of Congress, invested an average of .17\% of total obligations on R&D. During this term the Post Office never invested more than .47\% of total obligations for R&D in any single year (see Table 2-6 and Table 2-7). Postal investment in large capital projects followed a similar, if slightly less dramatic, outline. Between 1955, when efforts to modernize got off the ground, and 1968, Congress set aside only $520 million for the Post Office to apply to large capital projects (including, mechanization, automation, and new facility construction), an average of 1.16\% of total obligations to large capital projects per year (see Table 2-8 and Table 2-9). In nine of those years, investment in capital projects fell below 1\% of total obligations. Nearly 2/3 of this spending was devoted to creating new

\textsuperscript{56} By the late 1960s, 87.5\% of postal workers were represented by 1 of 12 craft unions. The largest two unions, the National Association of Letter Carriers and the United Federation of Postal Clerks, represented well over half of all postal employees. President’s Commission on Postal Organization. \textit{Towards Postal Excellence}, 18-19.
\textsuperscript{57} They did this not only by lobbying Congress to limit the amount of money appropriated for investment in new technology, but also by refusing to fully cooperate with some equipment field tests. See, Baxter. \textit{Labor and Politics}, 73-74.
postal facilities, while the remaining 1/3 was spent on mechanization projects.\(^{58}\) As a result, the Post Office at the end of the 1960s fell well behind private industry in terms of assets per employee. By 1968, the Post Office had roughly an asset-per-employee ratio of $700 ($700 in assets for every one employee), while Fortune’s 500 largest industrials had a minimum asset-per-employee ratio of $5,760.\(^{59}\) Firms with a similar degree of labor intensity, such as merchandizing firms and the transportation firm Consolidated Freightways had an asset-per-employee ratio between $4,000 and $5,000. According to any measure, postal investment in capital projects lagged seriously behind comparable enterprises.\(^{60}\) In 1968, the Post Office and the President’s Commission on Postal Organization estimated that a new investment of $5 billion—10 times the amount invested in the past thirteen years—was needed to make-up for chronic underinvestment.\(^{61}\)

*Regulation and Network Structure: The Post Office*

Congressional control of postal finances dampened and shaped the pace of technological change within the postal system and, by extension, molded the larger structure of the processing network. The broader regulatory structure governing the postal system—valuable in providing a national market for communication, subsidizing the diffusion of news and information, and undergirding the integration of rural

---

\(^{58}\) Figure calculated from data in, Arthur D. Little. “A Description of the Postal Service Today.” Table 4.1.11. 4-23.

\(^{59}\) Calculated from data in POD. Annual Report of the Postmaster General, 1969. Table 804. 246; Arthur D. Little. “A Description of the Postal Service Today.” Table 4.1.11. 4-20 – 4-23.

\(^{60}\) Arthur D. Little. “A Description of the Postal Service Today.” 4-20.

\(^{61}\) President’s Commission on Postal Organization. Towards Postal Excellence. 27.
communities into the life of the nation—served as a barrier to the widespread adoption of new technologies that could reduce costs and transform control within the network. The lack of innovation in new technologies of automation and mechanization, postal modernization as it was described during the 1950s and 1960s, was not a failing of postal management per se, but, rather, a side-effect of Congressional control and the structure of postal finances.

Postal regulation supported a loosely-coupled network structure. By favoring labor over new technologies, regulation buttressed a network where control was distributed and flexible. While this model was not the most cheap or economically efficient way to transport the mail available, it did provide a network structure that was resilient. The distribution of processing throughout the network meant that disruptions—poor weather in a certain city, the closing of an office due to flood—were local and did not have system-wide consequences. By supporting the processing of mail across the network and preventing the pooling of resources into a limited set of points, regulation provided system stability.62

Regulating the Railroads: Redundancy and the Unintended Benefits of Rate-Policy

While postal regulation supported loose coupling, regulation of freight railroads provided a hedge against complexity. Regulation inadvertently helped discourage the transportation of large quantities of dangerous chemicals in close proximity to population centers, limiting interactions of co-location and complexity. Unlike hazardous materials

62 Perrow highlights in passing the postal system as an example of a loosely-coupled system, though he does not provide an analysis of the network. Perrow. Normal Accidents. 97-99.
regulations specifically tailored to ensure the safe transportation of dangerous materials, economic regulation grew out of the late-19th century efforts of key shippers and the public to check the new, then virtually unrivaled, power of the railroads over the terms of the economy and the agitation of the railroads themselves to limit ruinous competition within the industry. The intersection of divergent interest groups led to the creation of the Interstate Commerce Commission (ICC), the first modern regulatory agency, in 1887 and mechanisms of regulation that stood in place for nearly 90 years. Regulation prevented the most egregious forms of monopoly abuse and stabilized the industry by preventing destructive competition in the form of rate wars, rebates, drawbacks. Yet, as a side-effect of regulation, control over rail rates helped shift the transportation of large quantities of hazardous materials during the middle decades of the 20th century from carriage by rail to smaller quantities carried by truck. Regulatory control over rates, abandonments, and mergers also underwrote an expansive and redundant rail network above what was strictly economically justified. The redundant network allowed for multiple shipping pathways connecting various points situated along the network and, as a result, made it possible to route hazardous shipments across routes that avoided traveling near dense areas. Together, regulation helped to create a rail network that discouraged the transportation of hazardous materials in large accumulations and allowed for those remaining, residual, shipments of hazardous materials still traveling by rail to skirt densely populated areas.

*State Support and Subsidy: Early Railroad Growth and Development*
Government involvement in U.S. railroads predates federal regulation. Indeed, it is difficult to find a moment when the state was not engaged with the railroads. Before the Civil War, the states and federal government played a key role in fostering the development and growth of railroads in an effort to link local communities with expanding markets of trade and spur commerce. Merchants and shippers called for direct public subsidies in support of the new emerging system of transportation. Direct public subsidies to private railroads, particularly before the 1860s, were a key source of capital. In addition to direct subsidies, growing rail networks benefited from public support in the form of the use of eminent domain law to provide rights of ways and the proffering of lucrative mail contracts. After modest development in the 1830s, the railroad boom that began in the middle of the 1840s brought rapid expansion first to New

---


64 Railroads provided many obvious benefits over other modes of transportation. Unlike other forms of overland transportation, such as canals, the railroads were quicker, cheaper, and far more reliable—particularly in difficult weather conditions. Alfred D. Chandler Jr. *The Visible Hand: The Managerial Revolution in American Business*. 1977. Cambridge, MA: Harvard UP, 1993. 86.

65 Despite support for subsidy, these same groups rarely called for public ownership of the railroads during the early period of development. Only a small number of railroads in the United States were ever operated as public enterprises and by 1850, with few exceptions, these were transferred to private hands. The unit costs of rail service were far lower than other competing modes of transportation; in part because railroads could make multiple trips in the period of time that it took other modes, such as canals, to complete one trip, and railroads provided great predictability. Chandler. *The Visible Hand*. 82. Frederick C. Clark. *State Railroad Commissions, and How they May be Made Effective*. Baltimore: Guggenheimer and Weil Co., 1891. 12-14.


England, and then the south, and west. By 1854, the main intersectional trunk lines—the Erie, the Baltimore and Ohio, the Pennsylvania, and the New York Central—provided links between east and west.\textsuperscript{68} During the middle decades of the 19\textsuperscript{th} century, railroads grew steadily (see Table 2-10) and included nearly 100,000 miles of track by 1880.

**The Need for Regulation: Monopoly Power and Ruinous Competition**

It is difficult to overstate the transformative impact of railroads on 19\textsuperscript{th} century America. The railroads reordered established, and mainly local, patterns of trade and fed the growth of a national market dominated by forms of mass distribution and production.\textsuperscript{69} The railroads provided a new unparalleled locus of power in American life; they could, through the setting of rates and terms of access, determine the entire fortunes of a city or region. On some lines, railroads operated as a monopoly and, as a result, wielded incredible power to control markets for goods.\textsuperscript{70} Even in markets with modest competition, skewed rates, benefitting certain producers or shippers over others, could elevate and impoverish in equal measures.\textsuperscript{71} The rush to use public funds and grant rights of ways to support the railroads was initially an attempt by local communities to secure the benefits afforded by rail transportation and was, in some cases, seen as a matter of

\textsuperscript{68} Chandler. *The Visible Hand*. 83.
\textsuperscript{70} Often, railroads set local rates above through rates in an effort to capture long-haul traffic. This long/short divide disadvantaged farm producers and merchants situated on interior lines who were effectively bypassed in favor those situated at the ends of the lines. Hoogenboom and Hoogenboom. *A History of the ICC*. 4-5.
\textsuperscript{71} Undisclosed rebates and drawbacks (rebates on competitor’s freight) were used by the railroads to capture certain large industrial shippers and drew immense disapproval from competitors that were denied such benefits. Famously, Standard Oil was a beneficiary of such practices. Ibid. 5.
survival. Yet, once the railroads arrived, discriminatory practices offered bitter frustration. After raising funds to attract service, the promise of cheap access to markets often proved to be an illusion. Shippers, including farmers, merchants, and independent oil producers viewed these practices as violations of basic notions of fair play and agitated for rail regulation as a means of redress.

Shipping interests were not alone in seeking regulation as a palliative to the new problems introduced by the growth of the rail networks: the railroads, too, sought regulation. Here, the problem was not too little competition, but too much: competition in areas with multiple lines proved to be ruinous. In a seeming paradox, the railroads faced both too little and too much competition. As Ari and Olive Hoogenboom note, “[r]ailroads were the first large American business and were absolute monopolies in large areas, yet fiercely competitive at many points.” Through lines and segments connecting urban centers were often served by multiple railroads. Railroads are characterized by high fixed costs and the pressure to secure high volume through low rates to defray those

---

72 For example, thirteen hundred farmers spread across eight different Wisconsin counties mortgaged their farmers to raise $1.5 million to attract rail service. Stone. *The Interstate Commerce Commission*. 2-3; Hoogenboom and Hoogenboom. *A History of the ICC*. 6-7.


costs was often difficult to ignore. On competitive lines, railroads engaged in periodic rate wars, sending rates downward, and offered large users generous benefits in the form of rebates and drawbacks in an effort to secure market share. As a result, maintaining profitability on competitive lines proved difficult and major railroads frequently faced bankruptcy. To combat the destructive effects of competition, railroads first turned toward voluntary associations—formal and informal pools and rate agreements to fix rates at certain levels (and prevent drawbacks and secret rebates), divide traffic among pool members, and share revenue—and later, during the 1880s, undertook massive campaigns to build larger end-to-end networks in an attempt to secure steady streams of traffic and eliminate competition (see Table 2-10). Neither strategy, however, succeeded: voluntary agreements were not legally binding and members broke agreements when it suited their interests, dissolving the pools; while enlarged, end-to-end networks, also failed to eliminate competition.

75 Under conditions of high-fixed costs increased volume does not dramatically alter cost. Hence, where competition exists, the inclination to slash rates to attract greater volume is acute. One estimate stated that during the 1880s as much as two-thirds or rail costs did not vary with volume carried. Chandler. The Visible Hand. 134; Hoogenboom and Hoogenboom. A History of the ICC. 2-3.

76 These practices were self-defeating and led to cycles of boom and bust. Freight revenue per ton mile declined between 1870 and 1900 (just before ICC regulation became effective in maintaining stability) from 1.88 cents in 1870, to 1.22 cents in 1880, .94 cents in 1890, and finally .73 cents in 1900. Large customers on competitive lines extracted rebates of up to 50% of the published rate as railroads scrambled to maintain traffic levels. During the rate war of 1876 and 1877, eastbound freight rates fell from one dollar to 15 cents per hundred pounds, while freight rates heading west fell from 75 to 25 cents. Keeler. Railroads, Freight, and Public Policy. 21-22; Hoogenboom and Hoogenboom. A History of the ICC. 2-3; Kolko. Railroads and Regulation. 7-12; Chandler. The Visible Hand. 134-143; Horwitz. The Irony of Regulatory Reform. 60-61; George W. Hilton. “The Consistency of the Interstate Commerce Act.” Journal of Law and Economics 9 (1966): 90.

77 The first pools were formed as early as the 1850s. The major pools were created during the 1870s first as informal agreements between regional competitors, and then as formal administrative federations containing executive, legislative, and judicial branches. Hoogenboom and Hoogenboom. A History of the ICC. 4-5; Kolko. Railroads and Regulation. 8-9; Chandler. The Visible Hand. 135.

78 Of the eight pools in operation in 1879, all but one rapidly failed. Chandler argues that consolidation and expansion helped buffer the ruinous effects of competition and, as a result, helped stabilize the industry to a
The Limits of State Regulation and the Advent of Federal Regulation

Before federal regulation, the states’ attempts to impose control over the railroad operation through charters and state commissions were ineffective. The Supreme Court considered the legality of state controls over the railroads in the landmark case of *Munn v. Illinois* (1876) and affirmed the contention, long established as a part of English Common Law, that certain types of private enterprises, “common callings” or “common carriers,” are imbued with a public interest due to their importance to commerce or scarcity and, as such, cease to be solely private. In particular, the Court reasoned that infrastructure services, such as railroads, are imbued with a larger significance and responsibility because they “stand…in the very ‘gateway of commerce’”; that is, they are the platforms upon which commerce operates. The railroads however, easily

greater degree than federal regulation. Yet, as Chandler later notes, these efforts did not eliminate rebates or drawbacks, two key elements of ruinous competition. Indeed, ruinous competition continued after the railroads pursued the creation of integrated systems. Chandler. *The Visible Hand*. 142-187; Kolko. *Railroads and Regulation*. 7-19; 64-101.

79 The first state commission was established in 1832 in Connecticut, though commissions remained irregular until the 1860s. Between 1860 and 1887, 20 state Commissions were in operation; in New England all states, with the exception of New Hampshire, had state commissions, while the Grange successfully pushed for strong state commissions in the Middle West. Generally, state commission attempted to ensure that railroads operated as common carriers and did not refuse service, charged reasonable and nondiscriminatory rates, and did not discontinue or abandon service without the consent of those served. These provisions were, however, often self-enforcing and as result translated into little practical control. The majority of state commissions were weak; of 20 state commissions, only 8 had the power to set mandatory rates. The remaining commissions acted as advisory boards with little power of enforcement. Benson. *Merchants, Farmers, and Railroads*. 1-2; Hoogenboom and Hoogenboom. *A History of the ICC*. 7; Kolko. *Railroads and Regulation*. 16.

80 The case considered the status of state controls over the economy and, specifically, the Granger laws. The case focused on grain elevators, but was directly linked to the railroad issue. As the Court opinion clarifies: “Property does become clothed with a public interest when used in a manner to make it of public consequence, and affect the community at large. When, therefore, one devotes his property to a use in which the public has an interest, he, in effect, grants to the public an interest in that use, and must submit to be controlled by the public for the common good, to the extent of the interest he has thus created.” The common law principle of common callings, or as it became known within the realm of transportation, common carriers, was applied to a wide swath of activities, including canals, turnpikes, millers, inns, and railroads. Munn v. Illinois, 94 U.S. 113 (1876); Horwitz. *The Irony of Regulatory Reform*. 49.

81 Ibid.
circumvented state control over rates by adjusting rates on interstate traffic. The final nail in the coffin of state regulation came in *Wabash v. Illinois* (1886) when the Supreme Court struck down interstate rate regulation by the states as inconsistent with the Interstate Commerce Clause of the Constitution.83

**The Interstate Commerce Commission: Regulating the Railroads**

In the wake of *Wabash*, Congress passed the Interstate Commerce Act of 1887 and created the Interstate Commerce Commission to regulate the national railroad system. The Act provided a general mandate for the ICC to enforce basic common carrier principles including nondiscrimination and fair and reasonable rates, while leaving the ICC room to define the specifics of regulations.84 Federal regulation offered a compromise between shipping interests and the railroads.85 The ICC eliminated the worst abuses of monopoly power and provided shippers with access to rail service on predictable, reasonable, and equitable terms. At the same time, the ICC also served the interests of the railroads: the Commission stabilized rates, provided an avenue for collective ratemaking, ensured a stable rate of return for the railroads, and eliminated the unpredictable nature of regulation by individual states.86


The Act created the ICC as an independent regulatory body headed by five commissioners serving staggered terms and provided the Commission with a general mandate to enforce common carrier obligations in the public interest. The Act charged that railroads must provide nondiscriminatory service, charge just and reasonable rates, adhere to published rates, and limited the use of pools. The ICC operated on a case-by-case basis and had latitude to enact the provisions of the Act. What precisely constituted fair and reasonable rates, for example, was an open question. The ICC represented a new mode of federal public authority that melded the various different functions of government—executive, legislative, and judicial. Initially, the courts moved to limit the power of the ICC through a series of decisions that struck a near fatal blow to the ICC and did not benefit either railroads or shippers. In response, Congress passed a series of laws restoring the powers the Commission was intended to have. By the end

---

87 Ibid. 22.
90 Kolko. Railroads and Regulation. 45-63.
91 After some initial success in enacting rail regulations, a series of court decisions in the 1890s circumscribed the power of the Commission. In 1889, during the first appeal of an ICC decision, the presiding Federal Circuit Court ruling insisted on rehearing and examining evidence previously submitted. As a result, the courts routinely considered new evidence and reevaluated evidence already presented to the ICC during the appeals process. This practice undercut the authority of the Commission and gave the courts tight control over the regulatory process. During the 1890s, the Supreme Court further reduced the power and effectiveness of the ICC. In Maximum Freight Rate (1897) the Court stripped the ICC of the power to set rates, while in Trans-Missouri (1897) the Court argued that cartel agreements and joint rates, which the ICC had moved to adopt and support, violated the Sherman Antitrust Act. Horwitz. The Irony of Regulatory Reform. 11; Stone. Interstate Commerce Commission. 7-8; Hoogenboom and Hoogenboom. A History of the ICC. 30-31; Kolko. Railroads and Regulation. 71-83.
92 These acts were: The Elkins Act (1903); The Hepburn Act (1906); and the Mann-Elkins Act of 1910. The Elkins Act made the published rate the legal rate and made deviations, such as rebates, punishable by law. The Act also allowed for joint or collective ratemaking by the railroads. The Hepburn Act gave the ICC the power to replace existing rates and made ICC decisions binding, shifting power away from the courts and back to the ICC. It also provided the ICC with the power to appoint examiners and agents to inspect railroad accounts. Finally, the Mann-Elkins Act gave the ICC the power to suspend rates and determine if circumstances between routes were substantially different enough to warrant pricing
of the first decade of the 20th century, the ICC now could set and suspend rates, enact binding decisions, engage in direct scrutiny of rail operations, and allow collective or joint rates filed by the railroads.

The ICC provided the railroads with many of the benefits that voluntary associations failed to provide. The ICC stabilized the industry by limiting competition and maintaining rate levels that could support existing operations. Legally binding rates eliminated competition based on price and supplemented the direct control of entry through licensing procedures with indirect control over entry (as new entrants could not employ below-cost pricing as a way of gaining market share). Additionally, by enforcing published rates, the ICC limited the drain, estimated at 10% of revenue before effective ICC regulation, associated with rebates and drawbacks. In determining reasonable rates, the ICC set rates at levels that provided a stable rate of return for the

93 Performance under regulation increased: between 1888 and 1910, the percentage of railroad stocks paying dividends rose from 39% to 67%; during the same period the average rate of dividends paid increased from 2.1% to 5%. Kolko. *Railroads and Regulation*. 232.

94 Although pooling was formally prohibited by the Interstate Commerce Act, the ICC, from its very beginning, achieved the desired end: the elimination of rate competition. In drawing regulations, the ICC allowed, informally at first and then legally, collective rate making through rate associations, and eventually set maximum and minimum rate-thresholds. Eventually, the Transportation Act of 1920 legalized pools, subject to ICC approval. Hoogenboom and Hoogenboom. *A History of the ICC*. 18-29, 60-61, 109; Kolko. *Railroads and Regulation*. 58-63, 100.

railroads; in the words of one Commissioner, railroads ought to earn a “living wage” that attracted investment and operate comfortably.\(^\text{97}\)

The ICC was not, however, simply a channel for railroad interests. On the contrary, regulation also addressed the concerns voiced for decades by shipping interests. The elimination of generous deals to valued shippers, such as drawbacks and rebates, in deference to published rates created a level playing field for shippers.\(^\text{98}\) Most notably, the call for just and reasonable rates and the prohibition of discriminatory service assured that shippers would not face the destructive caprice of the railroads or the wielding of monopoly power on non-competitive lines. To this end, the ICC capped the upper limit of rates and set comparable rates for the shipment of comparable goods.\(^\text{99}\)

*Regulating Redundancy: Cross-Subsidies, Mergers, Abandonments, and Excess Track*

Regulation was ultimately conservative: it sought to maintain the existing scale of the national rail network through limits on abandonments (discontinuation of service), rates, and the encouragement of strategic mergers. Rather than fostering consolidation and the elimination of excess plant as some hoped and suggested, ICC regulations buttressed the expansive network.\(^\text{100}\) A combination of factors—including the generous

---

\(^\text{97}\) This principle was consistently put into practice by the ICC and appeared as early as 1890. The 1914 ICC ruling on a proposed general rate increase in, in what is known as the “5 Percent Case,” formalized this policy. The 1920 Transportation Act fixed the rate-of-return at 6%. Stone. *The Interstate Commerce Commission*. 21; Kolko. *Railroads and Regulation*. 54, 210-217.


\(^\text{100}\) One of the key aims of the Transportation Act of 1920 was to introduce a plan to consolidate the railroad system into a more efficient network. Such a plan was commissioned but ultimately ignored. Keeler. *Railroads, Freight, and Public Policy*. 26.
solicitation of local communities, outright speculation, and the push by railroads to create integrated systems during the period before effective regulation—led to the creation of a rail network characterized by over-building (see Table 2-10). The sharpest growth of new track occurred during the 1870s and 1880s, when roughly 125,000 miles or 50% of eventual peak rail mileage in the United States was constructed.\textsuperscript{101} Contemporary observers noted that only a third of the added track was justified by increased traffic, with the rest being redundant.\textsuperscript{102} The network peaked in 1916 with 250,000 road-miles, and declined gradually through the 1970s as the overall share of inter-city freight traveling by rail plummeted (see Table 2-11).\textsuperscript{103} Reviews conducted during the 1970s estimated that roughly 75% to 80% of the nation’s rail system as it existed before deregulation was redundant and unsupported by traffic.\textsuperscript{104}

The Transportation Act of 1920 transferred control over abandonments from the states to the ICC. The ICC, as the states did previously, favored the continuation of service over abandonment and obligated railroads to maintain unprofitable services as part of their common carrier duty to operate in the public convenience and necessity.\textsuperscript{105}

Typically, if any party, such as mainly local shippers or rail customers, objected to a

\textsuperscript{102} Redundant track included both small branch lines unsupported by traffic and duplicate or parallel lines connecting two points. Hilton. “The Consistency of the Interstate Commerce Act.” 92-93.  
\textsuperscript{103} Rodrigue, Comtois, and Slack. The Geography of Transport Systems.  
\textsuperscript{104} The redundancy and excess capacity of the railroad system was not, then, simply a byproduct of the growth of other modes of transportation (such as air and truck). Rather, it was established during the 19th century growth of the network. The development of competing modes merely reinforced these qualities. Richard C. Levin. “Regulation, Barriers to Exit, and the Investment Behavior of Railroads.” Studies of Economic Regulation. Cambridge, MA: MIT Press, 1981. 184.  
service change the application would be rejected. The abandonment process itself was costly, slow moving, and complex. As a result, railroads rarely attempted to secure abandonments unless success was inevitable and, instead, chose to carry unprofitable services.

Rate regulation worked in concert with control over abandonments to support the maintenance of the existing outlines of the network. Rail rates were not set according to price. Indeed, the ICC interpreted the non-discriminatory language of the Interstate Commerce Act to mandate that shippers of comparable goods on high and low-density lines should not be charged different rates. In this manner, the ICC used its control over price to enact subsidies between rates on profitable densely trafficked lines and otherwise unprofitable low-density lines.

The ICC also used its control over mergers and traffic patterns to support the continuation of the expansive and redundant network. The Transportation Act of 1920 provided the ICC with power to approve mergers between lines. The ICC used this power to join weaker and stronger lines, rather than to consolidate excess capacity, and required that merged lines maintained pre-merger traffic patterns. The policy of maintaining pre-merger traffic flows was implemented as early as 1922 and formalized in

---

107 In the decade before deregulation, only 1% of total track was put up for abandonment. Of proposed abandonments, nearly all, 97.5%, were approved by the ICC. Levin. “Regulation, Barriers to Exit, and the Investment Behavior of Railroads.” 216-217; Dougan. “Railway Abandonments, Cross-Subsidies, and the Theory of Regulation.” 302.
the ICC’s ruling in the case *Detroit, Toledo & Ironton R. Co. Control* (1950) as a regulatory standard known as “DT&I conditions.” Regulatory oversight of mergers and imposition of DT&I conditions further supported redundancy by working to prevent the outright consolidation of track through mergers and by preventing the diversion of traffic flows to the most efficient routes. Here again, regulations worked to maintain the historic outlines of the network, rather than support efficient reorganization of the network into a dense web of heavily trafficked lines running between population centers.

*Rate Regulation and Cross-Modal Substitution*

While control over abandonments, rates, and mergers supported an expansive network across which freight moved, the ICC’s use of value-of-service rather than cost-of-service pricing in setting rates discouraged the transportation of large-quantities of hazardous chemicals via rail. Value-of-service pricing set rates in accordance with demand elasticities. For goods where demand fluctuated with price changes, rates were kept low; for those goods where demand did not vary with price, higher rates were imposed. Value-of-service pricing reflected the early agitation of farm interests in securing regulatory relief and had the practical effect of setting very low rates for cheap bulk agricultural products and high rates for more expensive manufactured goods. Since transportation costs accounted for a greater percentage of low cost commodities, demand demonstrated far greater elasticities for these products compared against higher costing commodities.

---

goods. In this manner, value-of-service pricing amounted to value-of-commodity pricing in transportation regulation.\textsuperscript{111}

Value-of-service regulation, along with more general cross-subsidies between high and low-density lines outlined above, contributed to the steep decline in the share of inter-city freight carried by rail through the 20th century. During the 1920s, railroads carried roughly 75\% of all intercity freight; by 1980, their share fell to 27\% (see Table 2-11) while inter-city motor carriage increased. For example, between 1939 and 1967, the share carried by railroad dropped from roughly 62.4\% to 41.6\%, while motor carrier share more than doubled from 9.7\% to 22.1\%.\textsuperscript{112} The structure of rail rates hastened modal shifts. Value-of-service pricing and cross-subsidies work well when users have few alternative options. But, throughout the middle-decades of the 20th century motor carriage emerged as a nimble competitor and as an avenue for shippers to opt out of the regulatory burdens associated with rail regulation. Though the ICC regulated trucking under the Motor Carriers Act of 1935 and adopted similar value-of-service principles in setting rates, the imposed regulatory burden for trucks was substantially less onerous when compared against rail. Value-of-service pricing in the abstract is designed to accommodate the availability of other modes of transportation. Yet, in practice, the ICC

\textsuperscript{111} Value-of-service pricing encouraged, by design, the widest possible utilization of the rail network and specifically encouraged rural farmers to ship via rail. In this way it also supported the continuation of an expansive network; whereas cost-of-service pricing would have discouraged the use of costly rural services, the adoption of value-of-service pricing encouraged wide use of the network. Stone, The Interstate Commerce Commission, 13; Keeler, Railroads, Freight, and Public Policy, 14; Ann F. Friedlaender and Richard H. Spady, Freight Transport Regulation: Equity, Efficiency, and Competition in the Rail and Trucking Industries. Cambridge, MA: MIT Press, 1981. 1-3.

refused to lower rail rates to accommodate the development of new modes of transportation and kept rail rates high even in the face of the steady erosion of market share. Additionally, the federal government provided support to trucking as it invested billions of dollars into creating the interstate highway system and improving roads during the 1950s and 1960s. These efforts helped nurture the growth of trucking services through indirect subsidy and establish trucking as a viable shipping alternative. Additionally, a substantial portion trucking developed outside of ICC regulation. Trucks owned and operated by the shipper of the commodity, private carriage, were not subject to ICC regulation and carriers of particular agricultural products were also exempt. It is unsurprising, then, to see the steady decline of freight traveling by rail and the growth of the share of inter-city freight traveling by truck. Indeed, trucking captured portions of inter-city freight that, absent regulation, would have been more cheaply handled by rail, such as traffic moving at distances above 100 miles. In this manner, regulation begat not only modal substitution but the misallocation of traffic from rail to truck.

The general picture of falling shares of inter-city freight allocated to rail and increasing carriage by truck was mirrored in the transportation of chemicals. Value-of-service regulation, by design, charged different rates to different commodities; the relative burdens of regulation were distributed differently. Rail regulation inflated rates by roughly 20% for the shipment of chemicals, near the average for all commodities, and

---

the mark-up above marginal cost assigned to rail shipments of chemicals were double that of corresponding truck rates.¹¹⁵ Though reliable figures are hard to come by, the Bureau of Census’ Commodity Transportation Survey, conducted as a part of the larger Economic Census during 1963, 1967, 1972, and 1977 provides a snapshot of the changing modal distribution for the transportation of industrial chemicals before deregulation. Between, 1963 and 1977, the share of industrial chemicals transported by rail declined from 55.4% to 33.4%, while the portion carried by truck increased from 28.8% to 39%; declines in line with more general modal shifts (see Table 2-11 and Table 2-12).¹¹⁶ There were, and are, stark differences in terms of how chemicals are moved via rail and truck. Typically, individual rail shipments move at roughly four times the size of truck shipments and travel greater distances between point of origin and destination.¹¹⁷ Yet, the structure of rail rates made it more attractive to producers and consumers to turn toward smaller shipments moving by truck.

*Rail Regulation: Network Structure, Proximity, and Hazardous Materials*

It is somewhat ironic to reinterpret the much decried inefficiencies of rail regulation through the lens of security. The particularities of rail regulation supported the decline of the use of rail shipments for inter-city freight. This substitution was not,

however, without its unexpected benefits. The drag placed on rates for different commodities by regulation helped remove a portion of large shipments of chemicals from the rails. Transportation by truck historically has a higher accident rate than transport by rail. Yet, despite this fact, transportation of hazardous materials via rail is more dangerous. The transportation of massive shipments of chemicals via rail cars creates the possibility of a catastrophic accident, a rare but exceedingly damaging event, while the transportation in small batches by truck create more frequent, though less consequential accidents. The risk associated with hazardous materials is tied to the properties of the carried material, the amount of material present, exposed population, and frequency of event.\(^\text{118}\) The structure of rail rates pushed the transportation of chemicals toward smaller batches of materials and shifted the risk away from a catastrophic event, toward a series of localized failures.

Ultimately, railroad regulation supported what Charles Perrow would describe as a linear system. Here, regulation worked to increase the distance (enlarge proximity) between large quantities of dangerous materials and dense populations. In doing so, regulation supported independence between two separate systems: the transportation of large quantities of dangerous materials and population centers. Regulation created incentives for the adoption of smaller shipments of hazardous materials, checked the larger aggregation of such materials, and also supported a rail network that wound through the nation making circuitous pathways possible. Those remaining chemicals that

did travel by rail moved across a redundant network containing track and routes that not only hewed close to cities and centers of trade, but extend through the nation’s less populated corners. In this respect, regulation created conditions that allowed hazardous materials to move at a remove from population centers.

A Network of Networks: Electric Power, Regulation, and Control

State and federal regulation of electric power promoted a loosely-coupled network of vertically organized local monopoly providers. Regulatory controls over entry organized the electric power system as a “network of networks,” to steal the oft used description of the internet. Power was produced and consumed within circumscribed regions and control was distributed throughout the network of investor owned utilities (IOUs)—nested within each vertical system—while interconnections with adjacent systems provided an added backstop in case of localized outages and failures. IOUs took charge of the three phases of electric power—generation, transmission, and distribution—for a limited geographic area and coordinated these operations through internal, rather than inter-firm, systems of control. The specific structure of utility regulation encouraged IOUs to invest heavily in new ICTs and unique systems of coordination and control; rate-of-return regulation provided the foundation on which proprietary, idiosyncratic, and isolated control systems would be developed and deployed.

State Regulation of Electric Power

The regulation of electric power, like the regulation of the railroads, did not emerge over the opposition of private industry, but rather found vocal and important
champions from within industry. Pioneering entrepreneurs in the business of electricity promoted the notion of natural monopoly—that electric power was best operated by a single provider—and solicited state regulation as a way of increasing productivity and market power, and as a means of deflecting the surging challenge from municipally owned and operated power providers.¹¹⁹ Progressive civic groups and reform governors joined with industry advocates in arguing for the necessity and virtue of state regulation. For reformers, state regulation offered a way to combat the endemic local corruption attached to municipal control. The ultimate success of state regulation of private monopolies was largely due to the intersection of the efforts of both industry and reformers; political maneuvering, less than technological inevitability, accounted for the institutionalization of the form.¹²⁰ Between 1907 and 1922, 32 states created state regulatory commissions to oversee the operation of private monopolies and the basic structure of the industry was set firmly in place.¹²¹

In the decades following Thomas Edison’s launch of the Pearl Street station in New York City in 1882, different approaches to organizing the provision of electricity jostled for supremacy. In addition to the well-documented battle of the currents, between

¹¹⁹ A “natural monopoly” is an industry with economies of scale where average cost decreases as the size of producing firm increases. Under these conditions, it is most beneficial to have one firm operating. See, Alfred Kahn. “The Role and Definition of Competition: Natural Monopoly.” The Economics of Regulation: Principles and Institutions. Vol. 2. Institutional Issues. New York: Wiley and Sons, 1971. 113-171.
Edison’s direct current (DC) and Westinghouse’s alternating current (AC) systems, different schemes for providing power—including local competition between multiple private firms, municipally owned and operated systems, on-site “isolated” systems, and cogeneration—flourished.\textsuperscript{122} The eventual institutionalization of state regulation of private vertical monopolies owed much of its success to the efforts of Samuel Insull. Insull served as Thomas Edison’s secretary from 1881-1892 before taking over as the head of the then small Chicago Edison Company and, eventually, steered industry advocacy in favor of state regulation.\textsuperscript{123} Insull’s great innovation lay in recognizing and then successfully realizing the opportunities afforded by monopoly service and new forms of steam generation.\textsuperscript{124} Insull was a proponent of “centralized systems,” large generation units linked to a set of residential customers through a transmission and distribution network.\textsuperscript{125} In retrospect, his approach seems deceptively simple. Electric power cannot be readily stored, generation and consumption must occur simultaneously, and demand varies greatly. Residential consumers during the early 20\textsuperscript{th} century, for example, used electric power mainly for lighting and, as a result, demand for power

\textsuperscript{123} Hirsh. \textit{Power Loss}. 13.  
\textsuperscript{124} Insull was an early booster of the steam turbine as a replacement for the reciprocating steam engine. Steam turbines utilized roughly one-tenth the space of reciprocating steam engines and, eventually, were able to generate far greater amounts of power. Ibid.  
\textsuperscript{125} This model of electricity provision, which Edison in effect inaugurated with the Pearl Street Station, is often described as centralized, as it pairs generation and a network of users. However, it is important to note that degree of centralization was qualified and limited. As will be discussed below, control was organized at the level of individual power systems and the electric power grid, in aggregate, comprised a web of functionally discrete and autonomous linked systems. Note: transmission lines are high voltage lines running from generating plants to substations; distribution lines are lower voltage lines that connect directly to end users.
would spike in the evening hours after work and decrease during the day. To meet demand, utilities would have to invest in generating capacity to meet peak load (demand) that would sit dormant during non-peak hours. Insull recognized that capturing demand diversity was essential to realizing economies of scale and maximizing the return on investment in large steam turbines. Rather than catering to a homogenous market where demand varied in lock-step, Insull sought out a mix of different customers with complementary demand curves (patterns of use) such that demand would be spread somewhat evenly. To take a basic example, by catering to both industrial consumers, whose demand spiked during the day, and residential users, whose demand peaked during at night, the electric system could be more effectively utilized and greater returns earned. 

Competition from multiple outlets and sources of electrical power, however, frustrated attempts to build a diverse portfolio of customers. Industrial customers cheaply produced their own power on site, a form of cogeneration, as a by-product of industrial processes, while isolated plants, on-site home and business generation systems, were widely used. In other cases, cities offered franchises to multiple competitive

---

126 Insull divided customers into 11 different classes of consumers in order to analyze the diversity of demand. He particularly prized ice manufacturers because their demand was at its lowest on cold days when demand from other customers was at its peak. Hughes. *Networks of Power.* 217-218.


private providers or operated their own municipal power systems.\textsuperscript{129} Municipal systems, however, presented the greatest challenge to the viability of private power systems. Between 1895 and 1906, publicly owned systems grew at twice the rate of private systems and accounted for 30 percent of the nation’s electrical suppliers, as cities turned toward public ownership as an alternative to private operation.\textsuperscript{130} Insull subverted these forms of competition by offering industrial customers and consumers drawing power at non-peak hours deep discounts, effectively siphoning off the most prized customers, and pursued the acquisition of rivals to limit local competition.\textsuperscript{131} But Insull saw these strategies as limited and partial; he pushed for state regulation as a way of checking the growing swell in favor of municipal power and limiting the availability of alternate suppliers of power.

Insull and his associates promoted state regulation through the two most powerful industry trade associations, the Association of Edison Illuminating Companies (AEIC) and the National Electric Light Association (NELA), and the National Civic Federation (NCF), a national progressive organization where he served on the executive committee.\textsuperscript{132} Insull advanced the idea of regulation as a useful alternative to municipal power in a speech before the NELA as early as 1898, where it was received with horror from members of industry who rejected the notion of regulation almost as an article of faith. Gradually over the next several years, however, as Insull’s associates would rise to

\textsuperscript{129} Chicago, for example, issued 20 non-exclusive, competitive, franchises between 1882 and 1905. Hirsh. \textit{Power Loss}. 14


\textsuperscript{131} Hirsh. \textit{Power Loss}. 13-14.

dominate industry trade groups, the industry came to embrace his views. Utilities found powerful allies in public reformers and the progressive movement. Governors Robert La Follette, Charles Evan Hughes, Hiram Johnson, and Woodrow Wilson, despite their many differences, viewed municipal power as bastions of local corruption and saw state commissions as an important remedy that could act as fair dealers and provide franchises in an uncompromised manner. Progressives, too, accepted the basic public position of the industry—that electric power was a natural monopoly—and saw state regulation as a means of harnessing the benefits of electricity for the public good. State regulation emerged from this confluence of interests, joining businessman and reformers. In 1907, the NCF promoted Insull’s position and drafted model legislation chartering state commissions; Wisconsin, New York, and Massachusetts adopted regulatory laws creating state commissions based on the NCF’ model legislation during the year (the Wisconsin law was directly taken from the drafted text), while other states quickly followed their lead and state commissions became standard.

The Outline of State Regulation

---

State commissions provided the basic regulatory mechanism that institutionalized the private operation of vertical monopolies. State commissions, in broad strokes, controlled entry through the granting of exclusive franchises for power companies to operate within a finite geographic territory.\(^{137}\) Regulation foreclosed alternate possibilities and pathways: the birth of state regulatory commissions effectively sucked the air out of the municipal power movement, and explicitly supported vertically organized systems of power that integrated generation, transmission, and distribution at the expense of other, different, models such as cogeneration, on-site power, and competitive provision.\(^{138}\) State regulation deflected the challenge of municipal ownership by legitimating private ownership and addressing, to a degree, complaints about private ownership and operation. The total number of municipal power systems peaked in 1922, with 2,500 utilities providing 5% of the nation’s generating capacity.\(^{139}\) The model that became common—vertically organized private monopolies—was not, to be clear, necessarily the most efficient: cogeneration, on-site systems, and municipal ownerships all had their merits and a clear, unequivocal, technical case against them cannot be made. State regulation created new barriers to alternate forms of power production and generation, while providing private vertical power systems with generous benefits.\(^{140}\)


\(^{139}\) Hirsh. “Emergence of Electrical Utilities.”

As in other industries, regulation operated as a quid quo pro that provided both private capital and the public with benefits. Licensed companies, most importantly, were protected from competition and state commissions set rates that provided a stable, guaranteed, return on investment (see below). Power companies were now easily able to aggregate diverse sets of users and attract outside investment, a necessity due to the high fixed costs associated with electric power systems and a notoriously difficult achievement under the previous conditions of hit-and-run competition and possible municipal takeover. Regulation also enabled private companies to use eminent domain in siting power plants, lines, and other facilities. The public also reaped substantial rewards from the regulatory compact. In concert with exclusive franchise, power companies were essentially treated as common carriers, obligated to build out systems to serve all customers at fair, non-discriminatory, rates and provide reliable service.

Regulation and Control: Vertical Monopoly and ICTs

State regulation shaped how electric power systems deployed ICTs as vital elements of control in two key respects. First, state regulation supported a model of loose coupling and distributed control. Regulation created a series of geographically-bound autonomous vertical monopolies through control over entry. Each of these units coordinated the flow of power through their systems through intra-firm control networks. Second, state regulation supported the adoption of new ICTs. Rate-of-return regulation encouraged and rewarded investment, even overinvestment, in new capital. As a result,

---

electric power utilities invested heavily in the development and acquisition of ICTs customized to the particular needs of the utility.

State regulatory commissions supported the creation of vertical monopolies that combined all three components of the provision of electricity—generation, transmission, and distribution. Under this model, each individual utility system controlled the flow of power through its network via central dispatch and internal control networks. ICTs were critical to this process: distributed power networks, networks combining off-site generation and the distribution of load to multiple endpoints, require ongoing measurement of voltage, frequency, ampere, as well as information about the state of the physical system, including circuit breakers, switches, turbines, and boilers.\(^\text{143}\) Electric power, as noted above, is a unique commodity in that it cannot be stored; it epitomizes the notion of “just-in-time” operation: generation and consumption must occur simultaneously. To achieve this balance, the different components of the system must operate in concert through coordination. Disjunction of supply and demand—generation above or below demand due to unplanned component failures or unforeseen swings in supply or demand—creates variations in voltage and current that undermine the security of the system and can lead to failure.\(^\text{144}\) As electric power systems increased in size to

---

\(^{143}\) Hughes. *Networks of Power*. 374.

incorporate multiple generators and larger distribution and transmission networks, coordination challenges increased apace.\textsuperscript{145}

During the 1920s and 1930s, the problem of network stability dominated the preoccupations of engineers and researchers. As vertical monopolies, the load dispatcher and dispatching center of each utility performed close monitoring and control over their individual networks through a range of control technologies designed both to forecast expected demand and measure current system performance. A central task for electric power operators was estimating fluctuations in demand. Demand fluctuates both predictably, in accordance with time of day, seasonal variation, and other stable changes that influence power consumption, and due to difficult to predict short-term variables, such as atypical weather. Utilities created detailed tables and statistical measurements, informational work, of historic trends of demand and capacity, in order to plan ahead for estimated required load. At the same time, increasingly sophisticated control devices provided system operators with readings about the current status of the network—readings about output and load, the status of generating units and switches—and allowed them to make system adjustments as needed from a centralized control room linked to generators and switches. These systems of control allowed the electric power systems to be centrally managed. Dispatchers linked to remote portions of the system through dedicated telegraph and telephone lines, carrier communications over power delivery facilities, and later microwave channels, received detailed readings of the network and,

\textsuperscript{145} Nagel. “Operating a Major Electric Utility Today.” 986.
when needed, manipulated circuit breakers, generators, and switches located at a great
remove in order to achieve ongoing balance between supply and demand.\textsuperscript{146}

Sate regulation supported distributed control in that each distinct vertical
monopoly controlled its own network. Stepping back to survey the state of electric
power, what appears is not a single network, but rather a collection of autonomous,
centrally managed, territorial monopolies.\textsuperscript{147} At the most basic level, the provision of
vertical monopolies supported internal control networks—the complex of ICTs wound
through an electric power system to provide measurement and manipulation. Alternate
schemes, to separate generation, transmission, and control on a competitive basis, as were
tried during the early days of electric power (see above) and would later make a
comeback in the wake of deregulation, require inter-firm control networks open to each
of the different players.

More subtly, rate-of-return regulation supported the development and deployment
of specialized ICTs for power system operations. Under rate-of-return regulation,
research and development and capital costs were included in calculating the rate base
upon which utilities were guaranteed a certain percentage return, while labor and fuel

\textsuperscript{146} As Hughes notes, the central dispatcher had to be something of a historian, understanding patterns of use
for the area under their control. As early as 1910, central dispatch regularly had access to compiled records
displaying hourly system output going back several years that could be matched against demand. Hughes. 
\textit{Networks of Power}. 215-215. For an overview of the use of real time technologies in electrical power, see
Nathan Cohn. “Recollections of the Evolution of Realtime Control Applications to Power Systems.”
discussion of the history of control technology as it relates to electrical power and beyond, see Stuart 
Bennett. \textit{A History of Control Engineering, 1800-1930}. New York: Institution of Electrical Engineers,
Engineers, 1993.

costs were not included in the rate base and were directly passed onto the consumer. Regulation, as a blunt instrument, created a system of incentives that encouraged utilities to seek out and invest heavily in new, costly, capital projects and research (though they often bypassed research in favor of capital) regardless if they provided certain economic benefits. Under rate-of-return regulation, firms could invest in capital intensive projects and technologies as a way to expand the rate base and, consequently, increase profits. Historically, state regulatory commissions rarely challenged or disallowed utilities from including particular capital costs in the calculation of the rate base, nor were they quick to lower rates if new cost savings materialized. Under this model, the electric utility industry became the largest industry in the U.S. in terms of gross capital assets. Unlike the Post Office, where regulation supported labor over technology, the regulation of electric power encouraged investment in technologies above all else. Key beneficiaries and supporters of rate-of-return regulation were the equipment manufactures that

---


150 State regulatory commissions did not have the staff to provide such detailed oversight. Nor was there a real call to do so until the upheavals of the 1970s and the advent of regulatory activism, which will be discussed in the next chapter. This gap, between shifting costs and unadjusted, stable, rates, is known as regulatory lag. Hirsh. *Technology and Transformation*. 112-113, 150-152; Anderson. *Regulatory Politics and Electric Utilities*. 71.


152 The replacement of manual controls with automated, computerized, controls reduced but did not eliminate the hand of the operator. In addition to the programming and maintenance, even closed loop automated power systems required a degree of operator oversight and contained interfaces to this end. Nagel. “Operating a Major Electrical Utility Today.” 988.
supplied industry, General Electric and Westinghouse most obviously. In this light, it is unsurprising to find that utilities provided early support and were early adopters of analog and digital computers. Initially, computers offered utilities a way to study the stability of large power systems through models and in assisting in solving complex differential equations. Later, beginning in the 1950s, utilities integrated open and closed loop computer systems into their operations to aid system stability, economic dispatch, demand forecasting, contingency planning, and other functions.

During the early 1920s and early 1930s, General Electric was a key source of support for the work of Vannevar Bush and his colleagues at MIT’s Electrical Engineering Department. Their work on the impact of transients—short-term, passing, phenomena such as lightning strikes, increases or decreases in load or generation, and short circuits—on system stability led to the development of two early analog computers: the network analyzer and the differential analyzer. The network analyzer served as a programmable scale model for electrical systems that could be used to simulate power

---

153 Hirsh, Technology and Transformation, 84.
154 This is not to say that such technologies did not improve utility performance or that utility managers did not see their adoption as a good faith effort to improve operation. Rather, the structure of incentives as designed under rate-of-return regulation simply made adoption more attractive regardless.
system behavior, while the differential analyzer simplified the work of computing complex differential equations associated with the analysis of power system operation.\textsuperscript{157} MIT offered both devices for use to commercial clients, such as General Electric, American Gas and Electrical Services, and other utilities, and were later reproduced by General Electric and other universities.\textsuperscript{158} General Electric’s main competitor in the electric equipment market, Westinghouse, developed its own computational equipment as well. After WWI, Westinghouse built DC and AC calculating boards that could approximate network behavior and during the 1940s, in partnership with Caltech, built the analog computer Anacom to study system stability.\textsuperscript{159} The differential analyzer, in particular, was used by the University of Pennsylvania in developing ENIAC, the first digital computer.\textsuperscript{160}

During the 1950s, electric power companies began to adopt digital computers directly into their day-to-day operations for system management, rather than as tools to provide modeling and off-line analysis of system stability. Leeds & Northrup, IBM, GE,

\textsuperscript{157}The study of transients was central to the work Bush, and the Electrical Engineering Department of MIT conducted during the 1920s and early 1930s. The network analyzer used small transformers as generators and motors, artificial lines, and a plugboard from a telephone exchange to allow reconfiguration. The network analyzer reproduced a 200-mile system within 50 square feet and differed from other earlier models and AC and DC calculating boards in that it was programmable: it could be reconfigured to represent and model different systems. The differential analyzer was the latest in a line of mechanical calculating machines that included the earlier intergraph. For an overview the creation of the network analyzer and the differential analyzer, see Mindell. \textit{Between Human and Machine}. 143-158; Owens. “Vannevar Bush and the Differential Analyzer.” 63-95; Hughes. \textit{Networks of Power}. 376-377; Aristotle Tympas. “From Digital to Analog and Back: The Ideology of Intelligent Machines in the History of the Electric Analyzer, 1870s-1960s.” \textit{IEEE Annals of the History of Computing} 18.4 (1996): 42-48.

\textsuperscript{158}Mindell. \textit{Between Human and Machine}. 152-153, 157-158; Tympas. “From Digital to Analog and Back.” 45.


\textsuperscript{160}Hughes. \textit{Networks of Power}. 377.
Westinghouse, as well as other suppliers provided equipment and systems utilizing digital computers for utilities during the 1950s and 1960s to help manage their systems.\footnote{One of the early computer systems designed for power systems, Leeds and Northrup’s “Early Bird,” was available in 1954 and subsequently adopted by a host of utilities. In 1961, Louisiana Power and Light opened what was considered to be the first fully automated power plant, its Little Gypsy station outside of New Orleans, with equipment designed by Daystrom. Through the 1960s, power plants adopted digital computer systems to control operations with equipment from a variety of suppliers. For contemporary news accounts, see Gene Smith. “Computer to Get New Power Role.” \textit{New York Times}. July 6, 1958. F1; “Daystrom to Build System for Utility.” \textit{New York Times}. Mar. 12, 1959. 40; Gene Smith. “Automation Due in Power Plants.” \textit{New York Times}. Feb. 5, 1961. F1; “Computer Industry Chalks Up a Big Day.” \textit{New York Times}. Apr. 4, 1963. 67; Gene Smith. “Power Industry Adds Computers.” \textit{New York Times}. Oct. 8, 1964. 63.} Computer systems replaced the direct wire and basic telemetry systems then in use. Distributed computer networks called Supervisory Control and Data Acquisition (SCADA) systems linked remote terminal units (RTUs) with a master computer through a communications channel, such as telephone line or microwave. A single RTU collected data from multiple instruments and transmitted the information back to the master unit for logging and analysis. The master unit contained a graphic interface and returned commands back to the RTUs based on programmed commands or operator discretion (the systems contained both open and closed loop elements). SCADA systems provided control over complex, geographically distributed, interdependent systems.\footnote{Cohn, et al. “On-Line Computer Applications in the Electric Power Industry.” 81;Robert H. Miller. \textit{Power System Operation}. New York: McGraw-Hill, 1983. 90-97; Jerry Russell. “A Brief History of SCADA/EMS.” <http://scadahistory.com/index.html>.} They automated some of the work of system operations and provided operators with better control and information about system operations. Specifically, modified SCADA systems were used to preserve system stability and economic dispatch. One of the central functions of newly installed computer systems was to match supply (load) and demand (generation). These systems monitored system performance and made corresponding
adjustments to meet fluctuations, through what is known as Automatic Generation
Control (AGC), and also archived historical data on system performance to create
detailed forecasts that helped guide the adjustment of generators and performed
contingency analysis to alert operators as to possible critical system failures. These
systems also were used to execute economic dispatch, a process for determining the most
optimal, cheapest, combination of the multiple available generators connected to the
power system. Computer programs and systems compared the total required load against
the measurement of incremental cost for the generating assets based on information on
fuel costs, unit costs, transmission loss, and performed dispatch to meet the total power
need at the lowest total cost.

Taken together, state regulation undergirded the creation of vertical monopolies
containing intra-firm control networks, while also making available capital for
investment in the development and deployment of increasingly sophisticated ICTs to
operate these networks. These new ICTs, however, were integrated as intra-system
controls. Such support amounted to a hedge against large-scale failure, as control was
linked to system operation.

163 Initially, different computer systems were dedicated to these different functions. However, by the 1960s
SCADA systems integrated AGC and other applications into single, multipurpose, systems. Miller. Power
Controlled Power Systems: Part II—Area Controls and Load Dispatch.” IEEE Spectrum (May 1965): 72-
177-184.
164 Cohn. “Recollections of the Evolution of Realtime Control.” 157-159; Friedlander. “Computer-
Controlled Power Systems.” 72-89. On the development of economic dispatch techniques and computers,
Economic Scheduling and Control of Power Systems.” Transactions of the American Institute of Electrical
Engineers 74.3 (1955): 1136-1142.
Towards Super Power? Federal Regulation and the Preservation of Loose Coupling

Federal regulation intervened at a critical moment in the development of electrical power and checked the consolidation of scattered local systems into centralized regional systems or even a national power system. During the interwar years competing movements sought to transform the organization of electric power and exploit the new possibilities of interconnection that the war effort revealed. These movements sought to replace the loosely interconnected network of networks—interconnected autonomous vertical monopolies, what Thomas Hughes describes as something akin to a “confederation of nation-states”—with a single system under central control covering vast regions or the entire nation. Ultimately, the Public Utility Holding Company Act of 1935 (PUHCA) provided a break on these competing moves toward consolidation and maintained the independent and loosely-coupled architecture of electric power as it previously existed.

WWI: Creating an Interconnected Network

Before World War I, most electric power systems lacked interconnection; power systems were not linked to one another, but rather were configured as isolated networks operated by a single utility. During WWI, however, adjacent power companies began to enter into power pools in order to share power between independent systems to meet the growing demand driven by the war effort. Mobilization created spikes in demand

---

165 Hughes. *Networks of Power*. 325.
166 Hughes. *Networks of Power*. 292.
167 The war drove demand for electricity and stressed the capacity of providers, particularly in areas undertaking the production of munitions and shipbuilding. In 1917, Army engineers working under the newly formed War Industries Board conducted a survey of U.S. generating capacity in order to determine if
and created a new urgency for access to power. To this end, the War Industries Board along with state and federal authorities encouraged, and in some cases, directly ordered interconnection. The interconnection of adjoining systems through new high voltage transmission lines offered a way of maximizing capacity with little investment or the long lead-times of creating new generating capacity. Improvements in high voltage transmission lines during the first decades of the 1900s steadily increased the carrying capacity and distance of lines and made such interconnections possible. Through interconnection, surplus power could be used between systems; as one system approached peak capacity, the excess power of an interconnected underutilized system could be brought on line as a supplement. Just as interconnecting different customers within a city proved a boon for electric power companies, interconnections between systems served to increase load factor by matching diverse demand portfolios across broader regions. Interconnection provided not only a backup during periods of peak demand, but also offered a reserve in case of an unexpected outage; if a generator failed or a component of the system stopped working, power could be solicited and drawn from the electric power facilities were sufficient to support mobilization. Even though new, fast-tracked, construction projects would increase generating capacity by 10% in two years, the study forecasted dire shortfalls that could cripple the war effort. Leonard DeGraaf. “Corporate Liberalism and Electric Power System Planning in the 1920s.” The Business History Review 64.1 (1990): 15; Thomas P. Hughes. “Technology and Public Policy: The Failure of Giant Power.” Proceedings of the IEEE 64.9 (1976): 1363; Hughes. Networks of Power. 289-292.


170 The voltage of transmission lines doubled every 10 years during the early 1900s. In 1900 50,000 V lines were commercially available; in 1910, 110,000 V lines were available; and in 1920, 220,000 V were in use. Hughes. “Technology and Public Policy.” 1362.
an interconnected system to satisfy demand.  

Importantly, systems participating in power pools and interconnections, however, remained independent and loosely coupled: they controlled the flow of power through their individual network and remained autonomous.

_Super Power, Giant Power, and the Holding Companies_

War-time interconnection proved the viability, and in some cases, the desirability, of connecting previously separate systems. After the war, throughout the 1920s and early 1930s, numerous public and private schemes attempted to follow this trend to its logical conclusion and push toward consolidating loosely-conjoined local networks into a set of vast regional systems or even a single national power system. From within government, the Department of the Interior and the Governor of Pennsylvania championed creating the foundations for a vast regional or national power system under projects that would, in effect, replace local power systems with a centralized public power system. At the same time, the consolidation of ownership within electric power through the use of highly-leveraged holding companies sought to effect the same basic transformation, though, importantly, with the caveat of substituting the dominance of private industry in place of possible public control and oversight. These interwar movements and counter movements were a contest over who would control the newly reorganized power network—whether the state or private capital would dominate the (seeming) future of electric power.

---

During the 1920s, the Department of the Interior and, later, Commerce Secretary Herbert Hoover, along with the Governor of Pennsylvania, Gifford Pinchot, advanced public plans to reorganize and consolidate electric power under public direction as a way of rationalizing electric power and extending service to rural communities, while avoiding the pernicious influence of private capital through the growing “power trusts.” These plans, known as “Super Power” and “Giant Power” respectively, differed in their specifics but both advanced the erasure of the scattered sets of individual power systems into larger regional aggregations or even a single national system under central direction by a new governmental entity.173 The plans received broad coverage in the press but faced bitter opposition from the utility industry.174 Industry attacked both projects, even the more modest Super Power plan championed by Hoover, and charged that the public plans touting state and federal supervision of new interconnected systems were “communistic.” They worked to block the advancement of the plans.175

172 On the differences of the plans, see DeGraaf. “Corporate Liberalism and Electric Power.”
173 Super Power sought to integrate electric power systems stretching from Boston to Washington, D.C. into a single system boosted by additional large plants and transmission lines under the control of a new advisory board, the North America Super Power Board, which would assume the role of coordinating the distribution of power across the newly interconnected system of utilities. Giant Power, as envisioned, was a more radical public intervention. It called for a new Giant Power Board and the creation new plants situated at mine mouths in Pennsylvania that could export power to load centers throughout the state, and, eventually New York and New Jersey as well, across high voltage transmission lines. Under this system, the state would reserve the right to take over private facilities in 50 years and a more strict policy for setting rates, one that was not as hospitable to utilities, would be implemented. DeGraaf. “Corporate Liberalism and Electric Power.” 1-12; Hughes. Networks of Power. 296-313.
174 Initially, private utilities voiced some support for Super Power when it was first discussed in 1919. Quickly, however, as capital markets loosened and holding companies began having success at integrating operations outside of government control industry turned against plans for Super Power. DeGraaf. “Corporate Liberalism and Electric Power.” 11.
175 Hughes. Networks of Power. 308.
At the same time that Super Power and Giant Power were under evaluation, private industry undertook its own program of interconnection and consolidation outside of direct state stewardship. Beginning in earnest during the 1920s, highly-leveraged holding companies consolidated ownership of multiple power systems. Though holding companies had been involved with utility companies since the early days of electricity, it was not until after WWI that they began to dominate the industry. In addition to fostering leveraged structures, holding companies also notoriously overvalued assets and padded expenses as a way of circumventing state regulations on profits. This gambit provided handsome returns for those that controlled stock at the upper levels of the pyramid. The holding company movement reached its peak during the early 1930s when 16 holding companies controlled 90% of all the electricity in the U.S. This


177 Between 1922 and 1927 the number of holding companies increased from 102 to 180, while the number of operating companies actually declined from roughly 6,000 to 4,000. In the following years, the trend toward consolidation continued as the ownership of more than half of all private utility companies changed hands. From 1927 to 1930, there were an astounding 2,500 changes of ownership, the majority involving holding companies gaining control over independent operators. Hirsh. “Emergence of Electrical Utilities”; EIA. Public Utility Holding Company Act of 1935. 5-6.

178 Holding companies not only served as tools for speculation and the overvaluation of operating companies, they also provided some legitimate benefits. They attracted investment to the capital intensive industry, and allowed the pooling of technical and managerial resources between previously failing smaller utilities and larger units. Hirsh. Technology and Transformation. 23-24; William Emmons III. “Franklin D. Roosevelt, Electric Utilities, and the Power of Competition.” The Journal of Economic History. 53.4 (1993): 883.

change in political economy produced an echo in the structure of the physical networks, as operational integration followed economic consolidation. Transformations of the ownership and control over IOUs fed the growth toward increasingly centralized electric power system. Holding companies sought to effect the same changes as outlined in the proposed Super Power and Giant Power programs—consolidation and integration of scattered geographical networks into a single operating unit, knitted together with high voltage transmission lines fed by large generation facilities—though, of course, under the auspices of private authority and not public supervision. Once holding companies swallowed up previously independent operators, the independent systems were transformed into larger units under centralized control operating as one network.\footnote{\textit{“Section 11 (b) of the Holding Company Act.”} 1111; DeGraaf. \textit{“Corporate Liberalism and Electric Power.”} For a larger discussion of the rise and breadth of holding companies, see Hughes. \textit{Networks of Power.} 363-403.} For example, the United Corporation, chartered in 1929 by J.P Morgan & Company in partnership with two other investment houses, controlled five major holding companies and, by extension, numerous utilities across the northeast and southeast. Through such financial arrangements, J.P Morgan in effect controlled more than one third of the nation’s electric supply.\footnote{Beder. \textit{Power Loss.} 62; Hughes. \textit{Networks of Power.} 400-401.} After incorporation, holding companies began the process of interconnecting their utilities across the east coast into a system that looked much like the proposed Super Power system and rumors of a possible national monopoly under private control flourished.\footnote{Hughes. \textit{Networks of Power.} 400-401; Hughes. \textit{“Technology and Public Policy.”} 1368.}
During the 1920s and 1930s, reformers and the industry struggled to define who would control the reorganized system of electric power. Both movements took for granted the inevitability and desirability of merging formerly distinct networks into larger systems; but differed as to the terms on which such a change should unfold. Pinchot saw his model of public intervention as an explicit reaction and rejection to the growth of holding companies and private power. For him, the growing empires of electricity presided over by holding companies mimicked “an enchanted evil spider … hasting to spread his web over the whole of the United States” and presented a clear choice: “either we must control electric power…or its masters and owners will control us.” These movements attempted to nudge electric power systems toward a tightly-coupled architecture; geographically bound independent systems began to give way to increasingly larger regional operations with control—and as a result, risk and vulnerability—aggregated at correspondingly higher levels.

Federal Regulation: Public Utility Holding Company Act of 1935

Federal legislation blocked these competing movements and preserved the loosely-coupled structure of electric power systems. The Public Utilities Holding Company Act (PUHCA) of 1935 rolled back the clock, as it were, and pushed back against the drive toward consolidating systems into larger aggregations. The Act was a conservative resolution to the competing public and private movements toward consolidation. While private utilities parried and defeated Super Power and Giant Power, holding companies were not left untouched. The PUHCA, like other pieces of New Deal

legislation, struck a balance between preserving private capital and limiting the more deleterious effects of private power.

Public dissatisfaction against the power holding companies had been growing since the mid-1920s. At the behest of Congress, the Federal Trade Commission (FTC) began a high-profile public investigation into the holding company issue in 1928. The FTC’s 6-year study uncovered and publicized a myriad of financial irregularities and abuses at the hands of holding companies. The Great Depression hit holding companies, and their leveraged structure, particularly hard, and forced many into bankruptcy and ruined scores of investors in the process. Franklin D. Roosevelt capitalized on the rising outrage over holding companies, then seen as symbols of financial avarice that had helped capsize the economy, and targeted the control over electricity by the power trusts as a central part of his campaign platform in 1932. Roosevelt made his ire known, stating that holding companies were “a corporate invention which gave a few corporate insiders unwarranted and intolerable powers” and that he stood against the “Ishmaels and the Insulls, whose hand is against everyman’s.”

185 Though none quite as spectacularly as Insull’s Middle West Utilities Company. In 1929 shares of Insull’s company traded at $57, but were worth only 25 cents in 1932 before it declared what was at the time the largest bankruptcy in history. The damage was stark: 100,000 stockholders lost roughly $4 billion, including company employees who had been exhorted to invest in the company. The Middle West Utilities Company was unique only in the size of the failure, as the Great Depression drove holding companies, and the larger economy, to ruin. Beder. Power Play. 64; Emmons. “Franklin D. Roosevelt, Electric Utilities, and the Power of Competition.” 883-885; EIA. “Public Utilities Holding Company Act of 1935.” 1-8; Hirsh. Power Loss. 40; “Section 11 (b) of the Holding Company Act: Fifteen Years in Retrospect.” 1089.
As part of the New Deal, FDR moved to implement new federal legislation to break up holding companies.\footnote{Roosevelt attacked the utility problem through three directions: 1. PUHCA sought to break up holding companies and return them to effective state control; 2. Tennessee Valley Authority attempted to spur regional economic development and lower electricity prices through the introduction of regional competition via the new public project; 3. The Rural Electric Administration made available loans to assist rural cooperatives in setting up their own power systems. Though each of these initiatives was important, I focus on PUHCA as it most directly impacted the architecture of electric power networks. Emmons. “Franklin D. Roosevelt, Electric Utilities, and the Power of Competition.” 885-886. On the run up to the PUHCA see also, Hughes. Networks of Power. 392-393.} State commissions were unable to exert real regulatory oversight over the holding company empires, as their ownership of multiple franchises in different states amounted to interstate commerce and, as a result, placed them beyond the effective ambit of state authorities.\footnote{EIA. “Public Utilities Holding Company Act of 1935.” 1; Emmons. “Franklin D. Roosevelt, Electric Utilities, and the Power of Competition.” 883-885.} The PUHCA of 1935 attempted to reform the utility industry and return regulatory power to the states. The Act enabled the Securities and Exchange Commission (SEC) to oversee dividends, loans, security issuances, property transfers, approve new acquisitions, and impose uniform accounting provisions on utility companies.\footnote{“Section 11(b) of Holding Company Act.” 1094.} The central provision of the Act, section 11(b), the so-called death sentence, attempted a more thorough reform of utility operation. Section 11(b) called for the break-up of interstate holding companies into various, single, geographically contiguous operating companies. The vast empires under holding company control, made up of scattered utilities operating across states, were no longer tolerated. Now, companies could only operate a single “integrated system” that served a limited area.\footnote{“Section 11(b) of Holding Company Act.” 1094-1104; Hirsh. “Emergence of Electrical Utilities.”} Section 11(b) sought to limit the on-going accumulation and unification of scattered regional systems by holding companies, and, rather, bracket the operation of
utilities to a single system that could, now, be effectively regulated under the jurisdiction of a state regulatory commission. PUHCA, in effect, returned to the states the power to control entry.

The Act did not mandate the abolition all holding companies and made allowance for the retention of small, otherwise, uneconomical systems. Yet, the intent and eventual implementation of the Act attacked the growing “bigness” of utility systems and the consolidation of power behind holding companies, and to support local—that is state—control of limited systems. The SEC narrowly interpreted the Act and held that the breadth of a system alone, the size of a system, could provide adequate grounds for ordering divestment. The Act reversed the trajectory of the holding companies toward greater integration of scattered systems and led to the separation of hundreds of systems back into independent entities. Between 1935 and 1950, the number of holding companies registered with the SEC dropped to less than 20, while over 750 systems with assets valued at over $7 billion were divested from various holding companies. The PUHCA, along with state regulation, preserved and supported a patchwork network of

---

191 Multiple systems could be retained if a system (a) could not otherwise be operated without a substantial economic burden; (b) resided in the same or adjoining state; (c) the advantage of local management would not be impaired. In practice, the SEC developed a strict interpretation of the new law and required that systems be fully integrated and serves a single area (known as the “single area” interpretation). “Section 11(b) of Holding Company Act.” 1100-1104.

192 The number of holding companies registered with the SEC dropped to less than 20, while over 750 systems with assets valued at over $7 billion were divested from various holding companies. For a thorough discussion of section 11(b), its scope, purpose, and eventual interpretation, see “Section 11(b) of the Holding Company Act.”

193 Ibid. 1104.


vertical monopolies serving limited areas. The drive toward tight coupling, for a time, receded in the face of new regulatory provisions.

*Loose Coupling: Regulation and Control*

State and federal regulation maintained a loosely-coupled network of interconnected electric power systems. Under the legal framework created during the early decades of the 20th century, state regulatory commissions provided franchises creating autonomous vertical monopolies; while federal regulation, through the PUHCA and the auspices of the SEC, checked the growth and consolidation of these systems into the large configurations. In this manner, regulation supported distributed control: each system maintained hierarchical control over its segment of the larger network and oversaw the coordination of generation, transmission, and distribution. Rate-of-return regulation also supported the development and adoption of increasingly sophisticated ICTs to operate these systems. The structure of state and federal regulation maintained a patchwork confederation of systems, a network of networks, interconnected for reliability and support, but functionally distinct. In doing so, regulation provided a hedge against large-scale failure and systemic vulnerability. Distributed control prevented the aggregation or accumulation of risk; failures of any single utility system are limited to the scale of its operating territory. Though vertical monopolies were linked into larger regional grids, these linkages were designed such that they provided access to additional resources without the cession of control.

*Stumbling Toward Resilience: Revisiting and Rethinking Infrastructure Regulation*
Reexamining the regulation of infrastructures through the lens of normal accident theory reveals a largely unintended benefit: resilience. The historical experience of regulation helped push these systems into particular configurations. Infrastructures did not develop outside of or in opposition to the political arena. Rather, “politics” were instrumental in guiding and shaping how these networks evolved. Indeed, the particularities of regulation became embedded within the very structure and operation of these technical systems. Regulation operated differently in each sector and, consequently, left different imprints on each system. The structure of regulation in post, for example, starved the postal system of capital for technical improvements and supported distributed processing and a reliance on labor. In electric power, however, regulation operated quite differently and offered generous support for investment in new technologies, while preventing the growth of large centralized systems. This makes it difficult to offer sweeping statements about the relationship between regulatory oversight and technological development; in some cases, regulation supported heavy investment in new technologies, in others it made no such allowances. Yet, in each case examined, regulation did buttress resilient configurations.

To be sure, at times these configurations were not the most efficient model possible. Yet, entwined with these inefficiencies were benefits that were easy to overlook and ignore. Regulation built into infrastructures a degree of resilience by supporting loose coupling and linearity and, at times, buffering the push toward tight coupling and increasing complexity. By supporting distributed control in the postal and
electric power systems, and independence within the transportation of chemicals, regulation helped create and maintain systems that could cope with localized failures. These systems were not, to be certain, free from their share of disruptions or failures. However, the architecture of the systems helped to ensure that such failures were limited in scope and scale. Yet, while politics could undergird resilient systems it could in a quick breath turn to undermine them. Beginning in the 1970s, deregulation remade how infrastructures operated and, in the drive to capture new efficiencies, subverted the historic support for resilience. While the restructuring of infrastructures delivered new efficiencies, it did so at a price.
Tables

Table 2-1: Post Office Mail Volume and Clerks Employed, 1910-1968

<table>
<thead>
<tr>
<th>Year</th>
<th>Total Mail Volume (Billions)</th>
<th>Clerks (Thousands)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1910</td>
<td>14.8</td>
<td>49.8</td>
</tr>
<tr>
<td>1920</td>
<td>23</td>
<td>68</td>
</tr>
<tr>
<td>1930</td>
<td>27.8</td>
<td>93.2</td>
</tr>
<tr>
<td>1940</td>
<td>27.7</td>
<td>97.6</td>
</tr>
<tr>
<td>1950</td>
<td>45</td>
<td>168.2</td>
</tr>
<tr>
<td>1960</td>
<td>63.6</td>
<td>171.9</td>
</tr>
<tr>
<td>1968</td>
<td>79.5</td>
<td>308.1</td>
</tr>
</tbody>
</table>

Table 2-2: Postal Employment, 1930-1968

Postal Employment, 1930-1968 (in Thousands)

<table>
<thead>
<tr>
<th>Year</th>
<th>Total Employed Postal Clerks</th>
<th>Total Employed Postal Workers (All Categories)</th>
<th>Total Employed Postal Workers (Excluding Clerks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1930</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1940</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1950</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1960</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1968</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

196 Note: Mail volume not available for 1920, datum from 1923 is substituted in chart. POD. Annual Report of the Postmaster General, 1969. Table 801. 236-241; President’s Commission on Postal Organization. Towards Postal Excellence. 16, 170.

Table 2-3: Postal Operating Deficit, 1926-1969\textsuperscript{198}

<table>
<thead>
<tr>
<th>Year</th>
<th>Operating Deficit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1926</td>
<td>$(400,000,000)</td>
</tr>
<tr>
<td>1929</td>
<td>$(200,000,000)</td>
</tr>
<tr>
<td>1932</td>
<td>$0</td>
</tr>
<tr>
<td>1935</td>
<td>$200,000,000</td>
</tr>
<tr>
<td>1938</td>
<td>$400,000,000</td>
</tr>
<tr>
<td>1941</td>
<td>$600,000,000</td>
</tr>
<tr>
<td>1944</td>
<td>$800,000,000</td>
</tr>
<tr>
<td>1947</td>
<td>$1,000,000,000</td>
</tr>
<tr>
<td>1950</td>
<td>$1,200,000,000</td>
</tr>
<tr>
<td>1953</td>
<td>$1,400,000,000</td>
</tr>
<tr>
<td>1956</td>
<td>$1,600,000,000</td>
</tr>
<tr>
<td>1959</td>
<td>$1,800,000,000</td>
</tr>
<tr>
<td>1962</td>
<td>$2,000,000,000</td>
</tr>
<tr>
<td>1965</td>
<td>$2,200,000,000</td>
</tr>
<tr>
<td>1968</td>
<td>$2,400,000,000</td>
</tr>
</tbody>
</table>

\textsuperscript{198} The operating deficit represents expense over income. For an explanation of how the postal deficit is calculated, see, POD. Annual Report of the Postmaster General, 1969. 166-168. Table data collected from POD. Annual Report of the Postmaster General, 1969. Table 801. 236-241.
Table 2-4: Postal Deficit as Percentage of Total Annual Postal Expenditures, 1926-1969

<table>
<thead>
<tr>
<th>Year</th>
<th>Deficit as Percentage of Total Annual Postal Expenditures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1926</td>
<td>20%</td>
</tr>
<tr>
<td>1929</td>
<td>15%</td>
</tr>
<tr>
<td>1932</td>
<td>10%</td>
</tr>
<tr>
<td>1935</td>
<td>5%</td>
</tr>
<tr>
<td>1938</td>
<td>0%</td>
</tr>
<tr>
<td>1941</td>
<td>5%</td>
</tr>
<tr>
<td>1944</td>
<td>10%</td>
</tr>
<tr>
<td>1947</td>
<td>15%</td>
</tr>
<tr>
<td>1950</td>
<td>20%</td>
</tr>
<tr>
<td>1953</td>
<td>25%</td>
</tr>
<tr>
<td>1956</td>
<td>30%</td>
</tr>
<tr>
<td>1959</td>
<td>35%</td>
</tr>
<tr>
<td>1962</td>
<td>20%</td>
</tr>
<tr>
<td>1965</td>
<td>15%</td>
</tr>
<tr>
<td>1968</td>
<td>10%</td>
</tr>
</tbody>
</table>

All table data collected from POD. *Annual Report of the Postmaster General, 1969*. Table 801. 236-241.
Table 2-5: Post Office Department Appropriation Categories

<table>
<thead>
<tr>
<th>Appropriation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Administration and Regional Operations</td>
<td>• Headquarters administration;</td>
</tr>
<tr>
<td></td>
<td>• Field inspection service;</td>
</tr>
<tr>
<td></td>
<td>• Regional operation;</td>
</tr>
<tr>
<td></td>
<td>• Postal data centers.</td>
</tr>
<tr>
<td>2. Research, Development, and Engineering</td>
<td>• Direct operation;</td>
</tr>
<tr>
<td></td>
<td>• Engineering support services;</td>
</tr>
<tr>
<td></td>
<td>• Contract research and development.</td>
</tr>
<tr>
<td>3. Operations</td>
<td>• Administration of postal installations;</td>
</tr>
<tr>
<td></td>
<td>• Labor (postmasters, supervisors, clerks, carriers, mailhandlers);</td>
</tr>
<tr>
<td></td>
<td>• Maintenance service;</td>
</tr>
<tr>
<td></td>
<td>• Vehicle service.</td>
</tr>
<tr>
<td>4. Transportation</td>
<td>• Surface and air transportation.</td>
</tr>
<tr>
<td>5. Building Occupancy</td>
<td>• Rent;</td>
</tr>
<tr>
<td></td>
<td>• Fuel;</td>
</tr>
<tr>
<td></td>
<td>• Site acquisition.</td>
</tr>
<tr>
<td>6. Postal Supplies</td>
<td>• Supplies and Services.</td>
</tr>
<tr>
<td>7. Plant and Equipment</td>
<td>• Federal building improvements;</td>
</tr>
<tr>
<td></td>
<td>• Vehicles;</td>
</tr>
<tr>
<td></td>
<td>• Mail processing equipment;</td>
</tr>
<tr>
<td></td>
<td>• Postal support equipment;</td>
</tr>
<tr>
<td>8. Postal Public Buildings</td>
<td>• Sites, design, and expenses;</td>
</tr>
<tr>
<td></td>
<td>• Construction.</td>
</tr>
</tbody>
</table>

Condensed from, Arthur D. Little. “A Description of the Postal Service Today.” Table 6.3.4. 6-35.
Table 2-6: Post Office Research and Development Spending, 1950-1969

Table 2-7: Percentage of Total Postal Obligations Devoted to Research and Development, 1950-1970

\[ \text{R&D Spending} \]

\[ \text{R&D as Percentage of Total Obligations} \]

\[ \text{POD. Annual Report of the Postmaster General, 1969. Table 804. 246.} \]

### Table 2-8: Annual Postal Investment in Large Capital Projects, 1955-1967

<table>
<thead>
<tr>
<th>Year</th>
<th>Investment in Large Capital Projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1955</td>
<td>$-</td>
</tr>
<tr>
<td>1956</td>
<td>$-</td>
</tr>
<tr>
<td>1957</td>
<td>$20,000,000</td>
</tr>
<tr>
<td>1958</td>
<td>$40,000,000</td>
</tr>
<tr>
<td>1959</td>
<td>$60,000,000</td>
</tr>
<tr>
<td>1960</td>
<td>$80,000,000</td>
</tr>
<tr>
<td>1961</td>
<td>$60,000,000</td>
</tr>
<tr>
<td>1962</td>
<td>$40,000,000</td>
</tr>
<tr>
<td>1963</td>
<td>$20,000,000</td>
</tr>
<tr>
<td>1964</td>
<td>$40,000,000</td>
</tr>
<tr>
<td>1965</td>
<td>$60,000,000</td>
</tr>
<tr>
<td>1966</td>
<td>$80,000,000</td>
</tr>
<tr>
<td>1967</td>
<td>$100,000,000</td>
</tr>
</tbody>
</table>

---

In accounting, the Post Office divided large capital expenditures into three categories: land and buildings; fixed mechanization; and non-fixed mechanization. Fixed mechanization includes equipment that is facility-specific and involved in materials handling, such as conveyor belt systems and sack sorting systems. Non-fixed equipment is defined as equipment used in processing the mail and that can be used across facilities, included here are letter sorting machines, facer-cancellers, stacker feeders, as well as other pieces. Arthur D. Little. “A Description of the Postal Service Today.” 4-19 – 4-20.
Table 2-9: Percentage of Postal Obligations Devoted to Large Capital Projects, 1955-1967

Percentage of Postal Obligations Devoted to Large Capital Projects, 1955-1967

0.00% 0.50% 1.00% 1.50% 2.00% 2.50%

---

204 Arthur D. Little. “A Description of the Postal Service Today.” Table 4.1.10. 4-23; POD. Annual Report of the Postmaster General, 1969. Table 804. 246.
Table 2-10: U.S. Rail Mileage, 1840-1920\textsuperscript{205}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{U.S. Rail Mileage, 1840-1980}
\end{figure}

*Note: Peak mileage in 1916.

Table 2-11: Percentage of Inter-City Freight Carried by Rail, 1930-1980\textsuperscript{206}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Percentage of Inter-City Freight Carried by Rail, 1930-1980}
\end{figure}

\textsuperscript{205} Mileage information drawn from Chandler. \textit{The Visible Hand}. 82-83, 89; Rodrigue, Comtois, and Slack. \textit{The Geography of Transport Systems}.

Table 2-12: Percentage of Industrial Chemicals by Mode of Transportation, 1963-1977

Chapter 3: The Political Origins of Infrastructure Vulnerability: Regulatory Reform, Tight Coupling, and Complex Systems

Beginning in the 1970s, price-and-entry regulations governing infrastructure services were dramatically, and quite unexpectedly, largely dismantled. Liberalization and reform opened infrastructures to competition and relaxed state control and supervision over rates. The process of deregulation transformed the rules governing infrastructures that had been in place for decades (or longer) and remade the relationship between the state, network operators, and private capital. Deregulation attempted to reduce the role of the state in overseeing key aspects of infrastructure services and introduce, however imperfectly, market-based competition. At its most basic level, the revision and in some cases recision of regulations governing price and entry attempted to replace administrative controls with market rationality.

Deregulation impacted a broad swath of services, including post, telecommunication, transportation (rail, trucking, and air), electric power, and was in each instance significant and partial. It was significant insofar as it represented a historic repudiation of the basic compact that attempted to strike a balance between the broad aims of social policy, heterodox interests, and particular economic and technical characteristics of infrastructures. Yet, it was also partial in that it did not completely undo or eliminate state supervision or intervention in the operation of infrastructures—numerous regulations were left intact and indeed still exist. Though the particular dynamics of the forces driving the push for reform and the specifics of how deregulation unfolded differed across infrastructural sectors, in all instances deregulation was a
political process of restructuring that sought to relax economic regulations governing price and entry in infrastructure services in favor of a closer embrace of the market.

Ultimately, deregulation transformed the architecture of infrastructure networks and services. These changes, despite their other merits, created or extended new forms of vulnerability and imperiled resiliency. Previous regulatory regimes undergirded infrastructures that were resilient to large-scale failures (see Chapter 2). Deregulation created new forms of vulnerability through the elimination of redundancies and through creation of tightly-coupled and complex material networks prone to single-point, large-scale, failures. Deregulation, as a project of infrastructure reform, sought to eliminate waste and slack from infrastructures and deliver newly efficient systems. Yet, interwoven with these new-found efficiencies are significant vulnerabilities. From this angle the inefficacies of the old regulatory model, much decried during their day, appear to have offered some overlooked benefits.

At the same time that this political process of reorganization unfolded, new information and communication technologies became central to the operation of other infrastructures. As geographically expansive networks, ICTs have always been a key cog in the operation of the postal system, rail network, and the electric power grid.¹ However, starting in the 1960s new forms of computer-based, networked, technologies began to be integrated into the operation of infrastructures. The diffusion of these

---

technologies and deregulation were not merely contemporaneous. Rather, the exigencies of deregulation shaped the terms on which these new ICTs would be adopted, particularly within electric power and postal service. These entwined transformations, deregulation and what is often dubbed the “information revolution,” created new conditions of systemic vulnerability.

Outline of the Chapter

In this section of the dissertation, I revisit the process of deregulation within the postal industry, electric power, and freight rail transportation. In each case, new legislation during the 1970s and 1980s greatly transformed the institutional framework within which these services operated. The Postal Reorganization Act of 1970 began the process of transforming postal service by paving the way for the partial liberalization of the postal market, phasing out direct subsidies from the federal government for general operation, and undercutting the continuation of cross-subsidies between different categories of service in a movement towards rates that more directly reflect market conditions. In rail service, the Railroad Revitalization and Regulatory Reform Act of 1976 and the Staggers Rail Act of 1980 transformed the industry by allowing shippers and carriers to set rates privately and for railroads to define the outlines of the networks upon which they operate. Finally, in electric power, the Public Utility Regulatory Policies Act of 1978 (PURPA) began the tentative creation of competitive electricity

---

2 The term “deregulation” captures the basic ethos animating the move toward a model that afforded greater reliance on the market over state-intervention, if not the empirical reality. To be clear, deregulation was qualified and partial; some regulations remained in place. Ironically, in some instances “deregulation” required the proliferation of new regulations to maintain the theoretically free market.
markets and the splitting of local vertical monopolies controlling generation, transmission, and distribution.

Before delving into in-depth examinations of how deregulation unfolded within each sector, a more general theoretical framework drawn from the literature of path dependence will be introduced. Theories of path dependence offer a heuristic for understanding both the lasting durability of price-and-entry regulations and the eventual upheavals of deregulation.

Path Dependence: Regulatory Stability and Change

Deregulation was a seismic and surprising turn that stunned many observers. For decades, price-and-entry regulations were seen as stable and seemingly inviolable constructs that, once in place, were extremely difficult to dislodge or alter. Broad consensus and support from a variety of constituencies appeared to make any substantive adjustment to the basic terms of the regulatory compact extremely difficult, if not impossible. Yet over the course of a fairly short period of time the basic regulatory provisions overseeing transportation, communications, and power were radically altered in ways that would have been hard to predict even a few years previous.3

Theories of path dependence provide an explanatory framework that is useful in helping to make sense of these changes. Initially developed within economics to describe quixotic phenomena that challenged conventional theory, path dependence provides a useful tool for explaining the characteristics that made price-and-entry regulations durable and for providing an informed account of why, during the 1970s and early 1980s,

---

they became redrawn in critical respects. Regulations, once in place, carried a type of momentum or inertia that was difficult to alter. The outlines of regulation were encoded (or ossified) in law and the particulars of regulation that sat at the center of a web of complementary institutions. In turn, complementary institutions provided support for the continuation of regulation. Examining the interconnections between these complementary institutions underscores the durability of regulation, but can also help explain eventual change and revision. By considering how the interconnections between regulations and other complementary institutions began to fray and slacken, the “misfiring” of mechanisms of reproduction, we begin to see the openings of new spaces for change.

Path Dependence: Outlines of a Theory

Path dependence appeared during the 1980s within economics before quickly migrating and finding application within sociology, political science, and other related fields and sub-fields engaged in historically informed institutional analysis. The theory underwent important modifications concerning definitions of contingency and normative

---

assumptions relating to inefficiency as the general insights contained within path-dependent accounts gained currency across the social science.\(^5\) As a general approach, path dependence traces the interconnection between historically remote moments and focuses on mechanisms of “lock-in” or positive feedback that work to sustain particular outcomes. Path dependence, as a heuristic tool, embodies not only the bland assertion that “history matters,” but rather suggests a systematic approach for understanding how and why certain contingencies become durable and difficult to dislodge; it rests on the insight that certain significant, contingent, moments shape and determine later stages of a sequence through self-reinforcing mechanisms.\(^6\) By drawing attention to the importance of critical moments and mechanisms of “lock in,” studies of path dependence call for more than simple descriptions of what happened, or what Jeffery Haydu refers to as “plain historical common sense,” and instead highlight why certain formations persist.\(^7\)

Path-dependent accounts are defined by three common features: (1) initial contingency; (2) the presence of mechanisms of lock-in; (3) later durability or fixity of a chosen path.\(^8\) All treatments of path-dependent processes are defined by moments of initial contingency. At some early point, there are competing options or paths available to be selected; within such contingent moments, multiple outcomes are possible and different factors can influence which is chosen. After initial selection, however, the


process becomes increasingly deterministic and choice is constrained; once a particular
path is chosen, switching tracks, as it were, becomes more difficult (although not
impossible). At its core, path dependence rests on this disjunction between initial
openness and later rigidity.\(^9\) In its application, path dependence provides an account of
how history winnows down choice and constrains actors at later moments from diverting
toward a different set of possibilities.

The deterministic quality of the sequence results from the presence of self-
reinforcing mechanisms of positive feedback that work to support the selected track and
discourage other options or directions.\(^10\) In elaborating and defining how certain forces
create durability, path dependence becomes a useful theoretical tool with broader
application. It is through the identification of common self-reinforcing processes or
mechanisms of lock-in that path dependence provides an explanation of why certain
accidents of history are reproduced.\(^11\) Arthur identifies generic forms of self-reinforcing
mechanisms that function to support “lock in” in the economy: large set-up or fixed costs;
learning effects; coordination effects; and self-reinforcing expectations.\(^12\) Focusing on
these mechanisms offers a way of detailing historically informed explanations of
sequential processes and the development and maintenance of institutions.

In broadening the application of path dependence beyond economics, the
rigidities of the theory associated with its initial development and application were

\(^10\) Brian Arthur, one of the central figures in the development of path dependence within economic,
identifies four different generic forms of self-reinforcing mechanisms: large set-up or fixed costs; learning
effects; coordination effects; and self-reinforcing expectations. Arthur. *Increasing Returns and Path
Dependence in the Economy*. 112.
\(^12\) Arthur. *Increasing Resturns and Path Dependence in the Economy*. 112.
softened; the focus on normative assumptions of efficient and inefficient outcomes was
shed, as was the equation of contingency with randomness. This loosening opens the
applicability of the theory to a greater range of possible problems or cases and,
additionally, resituates it as a usefully explanatory framework exclusive of assumptions
concerning the relative utility or value of a given outcome. In this reworked form, path
dependence can be used to explain persistence, explaining why society moved in a certain
direction and the implications of such a move, without engaging in explicit judgments of
the desirability of such outcomes. At the same time, the move away from an excessive
focus on initial randomness transformed the understanding of contingency as it relates to
path dependence. In the early economics literature, contingency referred to moments
within which multiple outcomes were possible and the selection of one such outcome was
based on random or near-random factors. However, subsequent uses of the theory
reconceptualized contingency in a vein similar to “openness.” During these significant
contingent moments or critical junctures multiple paths are indeed available, but the
reasons behind the selection of a particular path need not reside within a conceptual black
box, cut-off from explanation or investigation. As Paul Pierson notes, path dependent
processes are rarely defined by completely random moments of initial contingency where
“anything goes.” The fact that there may be explanations of why a path was chosen
rooted in historical particularity and detail does not disqualify a certain sequence as being
path dependent. Indeed, many attempts to understand path-dependent processes focus

---

16 Pierson. Politics in Time. 51
attention on understanding the forces at play that shaped the selection of one path over another.\textsuperscript{17} Rather, the important factor remains the important role that identifiable mechanisms of lock-in play in reproducing and extending this initial condition forward in time in a manner that is increasingly difficult to change.\textsuperscript{18}

*Infrastructure Regulation as Path Dependent Processes*

Douglas North’s work on institutions is perhaps the most influential reworking of the economics literature on path dependence for social scientists and is central, along with the work of Paul Pierson’s work on political institutions, to the contention developed here that price-and-entry regulations are characterized by path dependence.\textsuperscript{19} North’s *Institutions, Institutional Change and Economic Performance* (1990) draws explicit connection between the work done by Brian Arthur and Paul David on technological development and the analysis of institutions. For North, institutions are defined as “any form of constraint that human beings devise to shape human interaction” and function as both formal and informal “rules of the game” that govern spheres of action.\textsuperscript{20} The same mechanisms of lock-in or positive feedback that Arthur identifies with path-dependent process are, in North’s reading, found within institutions. Institutions endure, in large part, because they are difficult to initially create and because a host of organizations or interests adapt to the specific system of incentives offered and consequently serve to militate against change: institutions create powerful continuances. Institutions, too, face


\textsuperscript{19} Ibid. 26.

\textsuperscript{20} North. *Institutions, Institutional Change and Economic Performance*. 4.
large setup costs, generate learning effects or coordination effects, and adaptive expectations.\textsuperscript{21} Developing and implementing a structure of rules that applies to diverse sets of actors entails difficult problems of collective action.\textsuperscript{22} Nowhere is this more visible than in the crafting of new forms of legislation and political institutions within a democracy. As Paul Pierson notes, getting a needed majority to agree on the importance of an issue and also agree to a particular remedy is an obvious collective action problem. Democratic majoritarian legislative practices, in fact are structured to discourage both hasty action and easy revision. That is, in general terms they favor both inaction and the status quo.\textsuperscript{23}

Critically for North, institutions create interdependencies between institutions and organizations that result in positive feedback. Institutions are, always, particular: they offer a set of incentives and opportunities that can be pursued. In North’s reading organizations, firms, government agencies, trade unions, churches, voluntary associations, or any other type of “interest” adapt to and, eventually, become invested in the vagaries of institutional structure.\textsuperscript{24} This symbiosis between institution and related organizations creates interdependences that are resistant to change.\textsuperscript{25} This interdependent web of institution and organizations creates, as North puts it, “massive increasing returns.” The institution affords advantages to certain organizational actors; these actors, in turn, offer both active and tacit support for the ongoing provision of the institution.

\textsuperscript{21} Ibid. 94.
\textsuperscript{22} The example North gives, the drafting and ratification of the U.S. Constitution, underscores the challenges of creating formal institutions. Ibid. 95.
\textsuperscript{24} Ibid. 1-10.
\textsuperscript{25} Ibid. 1-10, 83-110.
Political institutions, like institutions in general, are beset by interdependencies that generate positive feedback. Policies and law provide opportunities to which social actors adapt and seek to pursue. This complement then, operates in a self-reinforcing manner: entrenched interests defend their position and support the status quo. This dynamic, of high setup costs and interdependences, is largely self-sustaining and offers an explanation of how institutions, absent any claim of efficiency or general societal benefit, endure through time.

Regulations, like all political institutions, face high setup costs. During formative periods before the imposition of regulation, there are many different possibilities for organizing and arranging systems of communication, transportation, and power. That is, the shape, form, and direction of the systems are contingent. The eventual systems that emerged were not inevitabilities, but rather reflect the outcome of strategic actions undertaken by different parties with a vested stake in the particulars of the systems. Regulations are central to this process; they provide fixity to the systems and are instrumental in selecting and promoting a certain type of postal network, railroad, and

---

26 This complementarity, in some instances, is not benign; they can work to reproduce and extend asymmetries of power. Political institutions can be used to support the power of certain groups or interests while also diminishing the political capabilities of others. In such cases, power accumulates in a self-sustaining or circular way. What sets North’s and Pierson’s readings apart from general interest group theory, is that they do not take the allotment of interest groups as a given, but rather point toward how both institutions and interest groups develop in a interdependent manner. Similar to how users become invested in a technological format and standard, so too do organizations become invested in an institutional structure. These observations say nothing of the actual desirability or efficiency of a given institution. Instead, it highlights how persistence is achieved.

27 Alternate pathways and possibilities existed for each. The British model provided the foundation and clear forerunner for the U.S. postal system, though the U.S. Post Office sharply departed with tradition. In railroading, state regulation and competitive provision both vied with federal regulation before the advent of the ICC as possible ways of structuring governance. Finally, during the early days of electric power different ways of providing power—from on-site sources as opposed to central generation stations, or from municipally owned and operated stations or competitive markets—angled for supremacy. See Chapter 2.
electric power system. The outlines of infrastructure regulation are written into federal and, in some cases, state law after long, contentious, debates, and the eventual compromise of divergent interests.\textsuperscript{28} Once encoded in law, substantive tinkering is difficult. The structure of legislative politics makes it far easier to accept regulation as it is rather than undertake the difficult work of revisiting old debates, building a new consensus or coalition of support, and renegotiating new terms. In this respect, the path-dependent quality of price-and-entry regulations do not substantially differ from any other legally enshrined institutional form. Following North and Pierson, the high setup costs inherent in implementing new legislation provides a bias, once initially in place, in maintaining the status quo. Encased in legislative mandate, regulations adopt a degree of “stickiness” that makes continuation, rather than change, the norm.

More interestingly, however, once in place price-and-entry regulations generate positive feedback or reinforcement through a set of constituencies that advocated its continuation and deflected efforts at change. Infrastructure service providers, consumers, including key industrial users, and regulatory bodies all derived tangible benefit from regulation and, in turn, worked to maintain the status quo against periodic challenges. Regulations protected incumbent infrastructure providers from competition and, in some cases, the unpredictability of volatile popular unrest and agitation. Unlike forms of social regulations that industry chafed at as onerous, price-and-entry regulations were not punitive toward the industries that they governed, but rather, in the main, provided protection and stability. While cross-subsidies burdened different users to different

\textsuperscript{28} See Chapter 2.
degrees, the presumed added premium—the tax of regulation—was relatively minor, if not trivial, and offset by the provision of stable service. Regulatory bodies also received obvious benefits from the proffering of regulation: ongoing relevance. In the language of path dependence, regulations generated positive feedback that acted as mechanisms of reproduction. The dense web of complementary institutions and interests make substantive revision quite difficult. This complementarity, combined with general “stickiness” affiliated with law, gave infrastructure regulation inertia and momentum.

*Path Dependence and Deregulation*

Path dependence is not only, however, a theory of stability and continuity. One frequent challenge to the utility of path dependent explanations is that they take later moments of a sequence as over-determined and impervious to change. On the contrary, as North highlights path dependent institutions, as social products, are always open to change. While change is difficult and alternative options less available at later points in time than during earlier conjectural moments, actors can work to overcome institutional durability: inertia is not inevitability. Mechanisms of positive feedback can, over time, become less durable and start to fray. During these moments new opportunities and spaces for actors to intervene and possibly create new paths and sequences emerge. These moments are, in their own right, critical junctures. Additionally, external shocks, in the form of revolutionary new technologies, the emergence of new political movements, or catastrophic events (to name but a few) can also subvert the rigidities of path dependent processes. These external ruptures can provide the impetus for a redrafting of the seemingly taken-for-granted state of things. In either case, whether
aided by the breakdown of once stable mechanisms of positive reinforcement or through the presence of external shocks, as North underlines, change is always a social process that requires actors to seek out and advocate for a new vision.

By the 1970s, the sources of positive feedback that sustained regulation began to weaken; the interdependencies between complementary institutions began to fray and snap. The general consensus supporting regulation was challenged both by liberal-left activists that saw current regulation and regulatory bodies as “captured” by industry and conservative-right groups that viewed economic regulation as distortions of the free market. Regulation was conceptually recast less as a palliative than a problem. Regulated industry and key consumers, too, grew disenchanted with regulation. Price-and-entry regulations work reasonably well during periods of stable cost and limited competition. During the 1970s, however, costs raced upwards while new forms of competition outside of the regulated system emerged. Spikes in costs created great stress as previously minor cross-subsidies became increasingly weighty and engendered push-back from consumers. The introduction of new forms of competition further undermined the benefits of regulation. Once certain users are able to opt out of the regulated system, or at the very least regulated markets become contestable, it becomes very difficult to pass along the added costs of regulation. The migration of high-value users of services leaves regulated providers with the most expensive and least desirable users. In many cases, regulations no longer offered providers the stable operating margins of monopoly

---

30 Ibid.
service, but rather provided a straitjacket limiting cost-cutting and internal improvements.

It was against this backdrop that once seemingly inviolable regulations began to be renegotiated. To understand the specifics of deregulation as it unfolded within the postal sector, freight rail, and electric power, we now turn to examine each sector in detail.

_Deregulating Postal Service: ICTs, Tight-Coupling, and Vulnerability_

The U.S. postal system underwent a process of selective deregulation that began with the passage of the Postal Reorganization Act of 1970 and subsequently unfolded in a series of administrative proceedings stretching into the early 1980s. Regulatory reform created new forms of vulnerability by remaking the postal system into a tightly-coupled processing network. While the previous structure of regulation supported distributed, labor-intensive, processing, deregulation undid these supports and subsequently led to the adoption of automation and centralized processing within a collection of dense hubs (see Chapter 2). Framed as a way of removing the deleterious effects of “politics” on postal affairs, reorganization transferred the operation of the postal system from the Cabinet-level Post Office Department to the independent government establishment, the United States Postal Service (USPS). Importantly, reorganization eliminated Congressional-control over rates and expenditures and rejected what had been the core of postal regulation, cross-subsidies, in favor of market-based rates. Reform shaped the terms on which new forms of information and communication technologies (ICTs) would (or would not) be integrated into postal operations and at the same time circumscribed the role of the state in intervening in postal markets in deference to the presumed exigencies of the market. Ultimately, deregulation liberalized portions of the postal network and,
most importantly, increased the efficiency of the system by finally freeing management to adopt new technologies to aid in the processing and distribution of the mail. Yet, this process was not an unqualified success; these changes also injected new forms of vulnerability into postal operations and created the conditions for large-scale failure. As the 2001 Anthrax attacks would vividly highlight, the successes of reorganization were not without its drawbacks.

**Postal Service and the Push for Comprehensive Reform**

The push to reform postal service came from postal management and leadership. By the mid-1960s the Post Office faced an economic and operational crisis characterized by rising costs, an expanding deficit, and declines in the quality of service. Post-World War II economic prosperity fed a spike in mail volume that overtaxed the postal infrastructure: between 1945 and 1970, annual mail volume leapt from 37.9 billion to 84.8 billion pieces and total postal costs ballooned from roughly $1.1 billion to $7.3 billion.31 Revenue, however, did not keep pace, rising from $1.1 billion to only $6.3 billion over the same period, leading the Post Office to accumulate a deficit in excess of a billion dollars by the late 1960s (see Table 3-1).32

Postal equipment had not been updated in decades and the existing network of post offices was structured around increasingly obsolete rail lines, which made offices difficult to access via the now

---


prominent modes of truck and rail of transportation.\textsuperscript{33} The seemingly dire state of affairs was given visible expression with the breakdown of the Chicago Post Office under the weight of rising volume and outmoded plant and operations in 1966. As 10 million pieces of backlogged mail sat immobile in the office and in trailers lining the surrounding streets, management considered the drastic step of burning the accumulated mail in a desperate effort to “reset” the system.\textsuperscript{34}

To then Postmaster General Lawrence O’Brien, the breakdown was symptomatic of a larger looming crisis: increasing volume, rising costs and deficit, and outmoded infrastructure and technology would make large-scale breakdowns like Chicago the norm. In his estimation, the Post Office was in a “race with catastrophe” and regulatory reform provided the only answer.\textsuperscript{35} O’Brien pressed the issue of comprehensive postal reform during his tenure in office, over the reluctance of President Johnson, and secured the creation of a Presidential Commission to examine and offer recommendations on

\textsuperscript{33} Congress consistently refused to increase investment in postal infrastructure (in the form of creating new facilities or investing in new technologies to modernize sorting and transportation) through rate increases (which were never popular) or appropriations. As a result, postal clerks sorted the vast majority of mail by hand as they had for decades. For a complete discussion of the relationship between the structure of regulation and mail processing, see Chapter 2. The decline in rail service was sharp and challenged the ability of the Post Office to move the mail quickly and efficiently. Post offices had been strategically located adjacent to rail lines in order to assist quick processing and distribution and railway postal cars provided in-transit sorting. With the decline of rail and the subsequent shift to truck and air, the postal network became increasingly difficult to access and the railway postal cars became mostly a curiosity. In 1929 there were 224,324 miles of passenger line rail service, while in 1967 this number had shrunk to 69,108 miles. The Post Office’s fleet of railway postal cars which sorted mail in transit also became obsolete, declining from 5,087 in 1910 to 1,476 in 1967. United States. President’s Commission on Postal Organization. \textit{Towards Postal Excellence: The Report of the President’s Commission on Postal Organization}. Washington: GPO, 1968, 170.

\textsuperscript{34} Eventually, extra workers were hired, others worked overtime, and the situation was corrected. Kathleen Conkey. \textit{The Postal Precipice: Can the U.S. Postal Service be Saved?} Washington, D.C.: Center for the Study of Responsive Law, 1983. 37; President’s Commission on Postal Organization. \textit{Towards Postal Excellence}. 11-14.

\textsuperscript{35} President’s Commission on Postal Organization. \textit{Towards Postal Excellence}. 1968, 12.
postal reform in 1967. The Commission, known as the Kappel Commission after its chairperson, former head of AT&T Frederick Kappel, released its final report in 1968 and provided the basic outlines for reform legislation. O’Brien was able to secure a comment in support of postal reform in President Johnson’s farewell address and worked in tandem with President Nixon’s Postmaster General Winton Blount to advance the Kappel Commission’s recommendations and secure support from Congress, mailers, and the postal unions as the co-chair of the bipartisan Citizens Committee for Postal Reform.

The idea of reforming the basic structure of postal governance was not completely novel. Periodically, different groups agitated for dramatic reform. Nor was a postal deficit unprecedented; though the deficit spiked during the 1960s, revenue shortfalls were the norm for over a century. Historically, however, complementary institutions supported the continuation of the status quo and blunted calls for comprehensive change.

The unique form of postal service created in fits and starts during the late 18th and 19th

---


 centuries became invested in other, mutually supporting, institutions. The publishing industry, political parties, and national voluntary associations relied on cheap second-class postage for their operation, while political parties also relied on the dispensing of postmasterships as a key source of political patronage.  

In each instance, these interests provide a key source of “feedback” supporting the continuation of postal status quo. This mutual support, as much as ideological commitment to an ideal of service, helped ward-off challenges to the continuation of postal service in its unique iteration.

This matrix of support was, however, fragile and contingent. By the 1960s, generous postal rates were not quite as important to publishers as they had been in previous decades; alternate communication and distribution channels eroded the

---

41 Richard John and Richard Kielbowicz note the important role that postal rates played in the development of a national market for information and the publishing industry. See, Richard R. John. *Spreading the News: The American Postal System from Franklin to Morse*. Cambridge, MA: Harvard UP, 1998; Richard B. Kielbowicz. *News in the Mail: The Press, Post Office, and Public Information, 1700-1860s*. New York: Greenwood, 1989. Theda Skocpol notes that the spread of voluntary associations on a national scale mimicked the structure of the federal government and was supported by cheap second-class postage. Theda Skocpol. *Diminished Democracy: From Membership to Management in American Civic Life*. Norman, OK: University of Oklahoma Press, 2003. After Andrew Jackson took office in 1829, the Post Office functioned as an adjunct to the growing national political parties and as a key means of allocating spoils. Postmaster positions, representing each district in the nation, were lucrative posts (particularly when combined with operating a store or other business) and an easy way of keeping the political machinery well oiled. John. *Spreading the News*. 206-256; Fuller. *The American Mail*. 285-330. Civil service reform and the passage of the Pendleton Act in 1883 did little to separate the connection between the postal bureaucracy and the spoils system. Postmaster positions were exempt from the reform and were still doled out according to political favoritism. For example, Postmaster General Wanamaker replaced over 48,000 of the total 56,315 postmasters during his tenure (1889-1993) in office. Wayne E. Fuller. *RFD: The Changing Face of Rural America*. Bloomington, IN: Indiana UP, 1964. 86. Clerks and city mail carriers, the bulk of the urban work force, however, did come under civil service regulation during the early 1900s. Fuller. *The American Mail*. 310-330.

42 Special interests played a key role in postal rates and classifications. Importantly, however, these interests were not only external to postal policy but were creations of postal policy. See, Jane Kennedy. “United States Postal Rates, 1845-1951.” *Ph. D. Dissertation*. Columbia University, 1955.

43 Though challenges were frequent, none was quite as serious as that presented by the private express companies during the middle decades of the 19th century. On how the linkages between the Post Office, political parties, and the publishing industry helped to maintain the unique iteration of postal service, see Priest. “The History of the Postal Monopoly.”; Richard R. John. “Private Mail Delivery in the United States During the Nineteenth Century: A Sketch.” *Business and Economic History* 2.15 (1986): 135-147.
significance of the postal system particularly as a channel for newsgathering. Most importantly, during the 20th century the tight connection between the postal bureaucracy and partisan politics frayed. Carriers and clerks were placed under the civil service system (as opposed to postmasters who were still political appointees) and became an independent unionized political force that agitated for better working conditions and wage increases. The growing activism of labor transformed the relationship between partisan politics, and hence Congress, and postal labor: the postal workforce morphed from a political resource into a liability. By the late 1960s, postal reform appeared to be an attractive way for Congress and the Executive Branch to divest themselves from the headaches and diminishing benefits that flowed from involvement in postal politics.

Through occasionally tense negotiations with mailers and labor, the Nixon administration secured the passage of the Postal Reorganization Act (PRA). Critically, postal unions, initially a key roadblock to passing reform legislation, agreed to throw their support behind reorganization in exchange for wage increases and binding arbitration. Commercial mailers also came onboard once Postmaster Generals O’Brien and Blount and the Nixon Administration put their weight behind reform and regulatory restructuring appeared to be a real possibility. The promise of reducing postal costs

---

44 Letter carrier unions and voluntary associations were founded in the wake of the creation of free city delivery in 1863. By the early 1960s, the Post Office was one of the most heavily unionized areas of government with 87.5% of all workers belonging to one of the craft unions. Although, these workers could not legally strike, Executive Order 10988 in 1962 recognized unions within federal agencies. President’s Commission on Postal Organization. Towards Postal Excellence. 18-22, 114. A detailed history of the unionization of the postal workforce can be found in Sterling D. Spero, The Labor Movement in a Government Industry: A Study of Employee Organization in the Postal Service. New York: Arno, 1971.

45 In 1970 postal workers engaged in a wildcat strike over wages and working conditions that severely impacted service. As part of negotiations to end the strike, the unions and administration worked out the deal to secure passage of the PRA.
served as a carrot for mailers leery of losing generous rates in the face of escalating costs and declining service.

*The Postal Reorganization Act of 1970*

The Commission’s report provided the basic blueprint for the Postal Reorganization Act of 1970 (PRA). The Commission identified “politics” as the problem hampering the Post Office. The fracturing of authority between Congress and the Executive Branch meant that postal managers had no control over costs, revenue, or investment (see Chapter 2). In the Commission’s view, Congressional authority to set rates was an unquestioned disaster that favored political expedience over sound business discretion and economic data. Their recommendations sought to address these problems by transforming the Post Office into an independent self-supporting corporation owned by the federal government. The postal corporation, as they called it, would be free to open and close offices, set rates in accordance with sound data and study, borrow and invest in new technologies, and operate with the efficiency and flexibility of a business. The Kappel Commission argued that the reorganization of postal service into a business-like model would lead to the elimination of the postal deficit, lower rates, and improved service. While the Commission did not explicitly favor the repeal of the postal monopoly, it did advocate that rates and service should more closely follow the dictates of the market.

---

46 Postmaster General O’Brien played a key in shaping the ultimate recommendations of the commission. See, O’Brien. Lawrence F. O'Brien Oral History Interview XXI.

47 The Kappel Commission argued that reorganization would reduce postal costs by 20% without impacting levels of service. President’s Commission on Postal Organization. *Towards Postal Excellence*. 5-6.

48 The Commission noted that “only a Post Office quick to identify and meet market needs can successfully serve a changing economy.” President’s Commission on Postal Organization. *Towards Postal Excellence*. 3.
The Commission’s recommendations were written into law as the PRA.\(^49\) The PRA abolished the cabinet-level Post Office Department and established the United States Postal Service (USPS) as an independent government agency, and the Postal Rate Commission (PRC), a body to review postal rates and offer recommendations, to govern postal policy. The PRA also began the process of phasing out federal appropriations and required the USPS to break even.\(^50\) A new Board of Governors would head the USPS and be subject to staggered appointment (to ensure that no single President could stack the board).\(^51\) The USPS would operate with little direct Congressional oversight; it could initiate rate increases, borrow up to $15 billion, invest as in new technologies and improvements as they saw fit, and generally operate with a more or less free hand. Some general language, however, was written into the law noting the importance of rural service, equitable rates, and the continued support for cultural, literary, and informational content. Finally, the Act created a new body, the PRC, to serve as a board of review for rate proceedings to ensure that rates were fair, based on evidence, and in accordance with the aims of the PRA.

\(^{49}\) The major stumbling block to the passage of the PRA came from the postal labor unions. Roughly 90% of the over 700,000 postal employees belonged to one of the various postal labor unions. By the late 1960s, the postal work force was overwhelmingly concentrated in urban areas, 10 large city post offices employed nearly a quarter of all workers and 15% of the total offices employed 88% of all workers. The objections of the unions were dropped after the strike in early 1970 that crippled the post office. In negotiating the end of the strike, union officials traded their support for reform legislation for wage increases. Concerned mailers eventually supported reform rightly surmising that removing the process of rate setting from Congress would benefit them. John T. Tierney. *The U.S. Postal Service: Status and Prospects of a Public Enterprise.* Boston: Auburn, 1988. 15-31; Conkey. *The Postal Precipice.* 52-58; President’s Commission on Postal Organization. *Towards Postal Excellence.* 102.

\(^{50}\) Appropriations were phased out gradually. Currently, appropriations covering “revenue foregone” are only provided for discounted mail for the blind and overseas absentee ballots. See, United States. Congressional Research Service (CRS). “The Postal Revenue Forgone Appropriation: Overview and Current Issues.” RS21025. Dec. 28, 2005.

\(^{51}\) The Board of Governors is comprised of nine appointed board members and two further members – the Postmaster and Deputy Postmaster General – appointed by the other members of the Board.
The Politics of Postal Rates: Cost, Cross-Subsidies, and Value

The PRA transformed the basic outlines of postal governance by shifting authority from Congress to the newly created USPS. Congressional control over postal finances historically supported distributed processing through the network. The transfer of control to USPS, however, undid these supports. The intention of this transfer was to remake postal service in the guise of a typical business enterprise that could invest in new technologies, set rates, and define service at the discretion of the Board of Governors in line with the demands of the market. Yet, the PRA included vague language acknowledging that the Postal Service did, indeed, have a public service mission and retained the Private Express Statutes. Sorting out how to reconcile these different aspects of postal service, how the USPS would function as a business and a public service, was left muddled by the PRA and subsequently worked out through a series of contentious landmark rate cases and administrative hearings during the 1970s and early 1980s. It was here that the process of deregulation proceeded as large volume mailers attacked and helped discredit what had been the cornerstone of postal regulation for over a century, cross-subsidies between different forms of content, in favor of market prices. In the drab language of economic theory and legal minutia, substantial and dramatic changes to the constitution of postal service occurred. Importantly, the rejection of cross-subsidies

\[52\] The PRA did not, to be clear, eliminate politics from postal affairs, but rather changed the venue within which different interest groups, including mailers and labor, fought over the manifestly political issues relating to cost, classification, and terms of service. Now, these debates occurred through formal rate cases and administrative hearings that featured arcane economic studies and precise legal justifications. See, Tierney. The U.S. Postal Service; Conkey. The Postal Precipice.
introduced new notions of discriminatory pricing into the setting of postal rates that in
turn impacted the Postal Service’s adoption of new ICTs.

The PRA established an ambiguous set of guidelines for determining rates and
classifications. It phased out direct appropriations and stated in part that “each class of
mail or type of mail service bear the direct and indirect postal costs attributable to that
class or type plus that portion of all other costs of the Postal Service reasonably
assignable to such class or type.” It further states that “[t]he value of the mail service
actually provided each class or type of mail service to both the sender and recipient”
should be a factor in setting rates. In other words, both cost and “value” are to be used
in setting rates. However, it was unclear how costs would be determined (how
attributable costs would be ascertained and the remaining assignable costs proportioned)
and what would define value. Resolving these issues was central to the first several rate
cases inaugurated under the PRA.

Immediately, during the first case (R71-1) instituted in 1971, different coalitions
of mailers began challenging the Board of Governors’ allocation of costs and
corresponding proposed rates. The debate in this and in subsequent cases centers
around two questions: (1) the method used to determine attributable costs (those costs
directly related, and thus charged, to a class of mail); (2) the logic or method of fixing
assignable costs (those shared, institutional, costs to be divided across the classes).

54 An administrative law judge was used by the PRC to preside over the first two rate cases held after
reorganization. Thereafter the PRC held the hearings directly.
55 Dockets R71-1, R74-1, and R76-1 (PRC) provide exhaustive discussions of the different proposed
costing methodologies and their implication for different categories of mailers. United States. Postal Rate
Board of Governors (with the support of publishers of newspapers, magazines, and other publishers that comprised second-class mail) in effect proposed to maintain cross-subsidies based on using a short-run costing methodology to determine cost attribution, and “inverse elasticity” and value of service justifications for assigning costs. The USPS model of short-run costs defined as attributable costs that directly fluctuated with changes in volume. This left a significant amount of postal costs (51% during the first rate case, 55% in the next) as institutional or common costs and, consequently, preserved a significant amount discretion for the Board of Governors in setting rates. The institutional costs of providing service (those costs not directly attributable to any one class) were then assigned to classes of mail according to the elasticity (the degree to which volume fluctuates with changes in rates) and the presumed public value of content. By adopting inverse elasticity and value of service (i.e., content regulation) as pricing guidelines the basic system of cross-subsidies could be maintained.

First-class mailers (the Association of Public Utility Mailers and the National Association of Greeting Card Producers, most prominently), third-class mailers (such as...
the Association of Third Class Mailers and Mail Advertisers Service Association) countered that the use of inverse elasticity and assessments of value linked to content in rate setting was unfairly discriminatory and challenged the Postal Board of Governors.\textsuperscript{58}

They pointed out that the PRA called for rates that supported all of the costs that could be attributed to a particular class of mail and only a reasonable share of the remaining (assignable) costs. First and third-class interveners advocated a method of long-run costing that attributed nearly all costs. According to this model, first and third-class rates would not be subject to such a significant share of institutional costs and would decline, while second-class rates would dramatically increase. First-class mailers argued that as a “captured” class of mail subject to monopoly restrictions, postal rates under the terms of the PRA were discriminatory and unfair. Inverse elastic pricing, in their estimation, maintained cross-subsides: the monopoly class of mail, letters, would always be less sensitive to cost fluctuations and, as a result, shoulder the burden of institutional costs. First-class mailers, accordingly, argued that such a pricing scheme violated the terms of the Act. Third-class mailers also argued that by adopting a short-run costing methodology third-class rates were greatly inflated (due to paying for other forms of service). Both argued that value of service considerations should be content neutral.\textsuperscript{59}

That is, demand and not assumptions of the broader importance of the content of certain types of mailings should serve to define value under the PRA; the market, not substantive

\textsuperscript{58} For the various interveners’ complaints and claims of discriminatory pricing, PRC. Postal Rate and Fee Increases, 1971. 63-112.

\textsuperscript{59} Administrative Law Judge Seymour Wenner presided over the first two rate cases in concert with the PRC. He was very sympathetic to the claims of first-class mailers and challenged the continuation of cross-subsidies, noting disapprovingly: “If someone gets something for nothing someone else pays for it.” Wenner. “Chief Administrative Law Judge’s Initial Decision, R71-1.” 44.
judgments, should evaluate worth. Taken as a whole, these interveners argued that cross-subsidies and redistribution had no place in postal service and that discretion should be limited to the greatest extent possible. They advocated strict cost-of-service considerations as the single factor in setting rates and fought against interpreting the value-of-service provision of the PRA as supporting preferred content.\textsuperscript{60}

Disputes over cost attribution and assignability are about the linked concepts of the Board of Governors’ autonomy, the validity of cross-subsidies, and the meaning of value in determining rates; they are about substantive regulation and its role in the organization of the postal market. In each of the first three cases, the same problems appeared. The presiding administrative law judge and the PRC accepted the basic outlines of the arguments of first and third-class mailers and pushed the Board of Governors to develop models accounting for greater portions of costs (and, thus, limit the discretionary power of management). The Board of Governors did so reluctantly, but refused to assign all costs; the subsidies were reduced, but not completely abolished.\textsuperscript{61}

Second-class rates rose significantly, but the Board of Governors maintained some limited discretion over rates.\textsuperscript{62} Although the more radical proposals to have full-cost


\textsuperscript{61} The USPS assigned the following % of total costs as attributable: R71-1, 49%; R74-1, 45.1%; R76-1, 55.1%. The chief examiner countered in R74-1 by attributing 70.6%, which the PRC modified to 53.5%. In R76-1, the PRC modified the Board’s figure to 60.4% (see Commission 1977a, 59-67 and the dockets of each case). During the fourth rate case which began in 1977, the PRC attributed 74.5% of costs in agreement with the Board. Tierney. Postal Reorganization. 130; Tierney. The U.S. Postal Service. 162.

\textsuperscript{62} The National Association of Greeting Card Publishers filed suit in 1975 challenging the discretionary authority of the Board to determine costing methodology and assign rates (apportion institutional costs). The case eventually was joined with a similar challenge from UPS and progressed to the Supreme Court. In
accounting shape rates and only use demand factors (i.e., fluctuations of volume) as the only guides to value were averted, cost become the major factor is shaping rates. More importantly, first and third-class mailers were able to formerly introduce anti-trust principles and apply them to the setting of rates for all classes of mail. The foundational premise of reorganization, to give the postal bureaucracy increased autonomy and control, was quickly limited; the redistributive system of cross-subsidies was successfully challenged and the ability of the state to assign different rates based on distinctions of the value of content was greatly reduced.

National Association of Greeting Card Publishers, et al. v. United States Postal Service (1983) the Court upheld the authority of the Board to use discretion in setting rates. In other words, full-cost accounting was not necessary. But by this time first and third class mailers (along with the private delivery firms) had already succeeded in pushing for a greater amount of cost attribution, successfully challenged the system of cross-subsidies, and attacked content-based assessments of value. Despite the victory for the Board of Governors, the basic terms of service were already radically transformed. National Association of Greeting Card Publishers, et al. v. United States Postal Service, 462 U.S. 810 (1983).

Implicitly and explicitly the challenge to the authority of the Board of Governors to use discretion to assign costs was linked with a narrow view of value. Wenner, echoing the sentiments of those seeking to reduce the freedom of the Board of Governors, stated: “value of service considerations should in the main deal with demand factors. The demand for a class reflects the evaluation placed by both the sender and recipient.” Wenner. “Chief Administrative Law Judge’s Initial Decision, R71-1.” This opinion is a radical rebuke to way in which postal policy had been viewed for well over a century. In viewing value so narrowly, Wenner was arguing that postal service and indeed communication have little public value and instead are only significant as a matter between two contracting parties. A general disdain for evaluations of content is repeated throughout his opinions in R71-1 and R74-1. The PRC similarly adopted this dim view of content-related considerations of value in R76-1, but did not go so far as to saw that the use of discretion was, in and of itself, discriminatory. PRC. Postal Rate and Fee Increases, 1975. 75-156. Many decried this development, noting that it would help hasten the decline of small, marginal, publications. Arthur Schlesinger Jr. saw that rising second-class rates transformed the USPS into the “judge, jury, and executioner of the nation’s intellectual life”. Qtd. in Tierney. Postal Reorganization. 129. A June 16, 1975 Time magazine editorial also argued that the shifting of costs from first to second class mail would irrevocably damage small, marginal publications. It noted that second-class rates had increased over 100% since 1971 and that greater increases were forthcoming. “Postal Nightmare.” Time. June 16, 1975.

The new hostility toward the reallocation of coasts is expressed most clearly by Wenner. In his opinion in R74-1 he writes: “[t]he Postal service has become a tax-collacting agency, collecting money from first-class mailers to distribute to other favored classes” and that this situation needs to be undone as quickly as possible. Seymour Wenner. “Chief Administrative Law Judge’s Initial Decision, R74-1.” PRC. Postal Rate and Fee Increases, 1973. Docket No. R74-1. 13. Anti-trust principles were applied to Post Office’s parcel delivery service since the inception of parcel post.

The recently passed Postal Accountability and Enhancement Act provides a further elaboration of these principles by dividing postal operations between competitive and market dominant lines and calling for greater cost attribution and limitation on cross-subsidies. Postal Accountability and Enhancement Act. Dec. 20, 2006. P.L. 109-435.
New Technologies and New Opportunities: The Postal System and the Information Revolution

The introduction of new notions of non-discriminatory pricing into the setting of postal rates and the discrediting of cross-subsidies were important victories for large volume mailers. In addition to lessening the share of costs shouldered by first and third-class mailers, the outcome of the early rate cases provided mailers with new power to dictate the direction of postal policy. Postal reorganization was designed in large part to aid management’s adoption of new technologies and increase their control over key terms of service. Under the PRA, USPS was promised the autonomy to raise capital, through self-supporting rates and barrowing, and invest as it saw fit. Yet, the successful introduction of principles of non-discriminatory pricing and the attack on cross-subsidies circumscribed this autonomy. Now, large volume mailers are able to essentially wield a veto over postal policy by raising the objection that monopoly revenue generated through first-class rates subsidized other forms of service which resulted in both an unfair tax on the captured class of mailers and unfairly subsidized Postal Service activity in ostensibly competitive markets (such as electronic mail, as discussed below).

Quickly, during the 1970s mailers seized on the discrediting of cross-subsidies and new centrality of cost in the setting of rates to successfully argue for the partial liberalization of the upstream postal network. Mailers pushed through new workshare discounts on the premise of cost avoidance that, in effect, reduced the reach of the postal monopoly to the last mile and allowed bulk mailers and third party pre-sort aggregators to
deliver prepared mailings to select postal facilities close to their final destination at reduced rates.\textsuperscript{66}

Importantly, mailers further used these new principles to dictate the terms on which the Postal Service would adopt new ICTs. Through the 1970s and 1980s, the Postal Service pursued two strategies concerning new ICTs: market growth and expansion through the introduction of an electronic mail system, a proto e-mail service dubbed E-COM ("electronic originated mail"); and cost reduction and improved efficiency through increased reliance on mechanization and automation. Mailers, in direct conflict with the wishes of Postal Management, pushed the Postal Service to abandon efforts to use ICTs for a hybrid "electronic mail service" and supported the use of new technologies as a means of controlling costs and improving the efficiency of the basic postal network. By steering Postal Management away from innovative new forms of service and toward a program of internal improvements structured around mechanization and automation, large volume mailers helped to create a tightly-coupled, centralized, processing network.

\textit{An Electronic Postal System? The Failure of E-COM}

The USPS purposed E-COM during in 1978 as a means of countering the long-term forecasted electronic diversion of hardcopy mail by new electronic messaging

\textsuperscript{66} Pre-sorting discounts were first suggested by Readers Digest Corporation and the Association of Public Utility Mailers (APUM) during the 1971 rate case and were given a separate classification hearing in 1973. PRC. "Opinion and Recommended Decision, MC73-1." PRC. Mail Classification Schedule, 1973. Docket MC73-1. Washington, D.C.: PRC, 1976. 11. For the initial call, see PRC. Postal Rate and Fee Increases, 1971. 112-120. The postal unions, such as the American Postal Workers Union (APWU), repeatedly challenged that these rates offer discounts above cost-avoidance and transfer productivity gains to particular users at the expense of the general system. For an overview of the history of these categories and critiques of their effectiveness, see United States. General Accounting Office (GAO). "U.S. Postal Service: A Primer on Postal Worksharing." GAO-03-927. July 2003.
systems and electronic fund transfers. The Postal Board of Governors request to alter the mail classification schedule outlined E-COM as an electronic/print hybrid that would integrate electronic transmission and data processing with the delivery of print letters.

Under the proposed plan, the USPS would contract with a telecommunications carrier to provide service between 25 “serving post offices” (SPOs) and a designated point of reception. Once an electronic message was sent to the Service, it would be forwarded to one of the SPOs (in electronic form), printed, enveloped, and delivered. The USPS would be acting as a telecom reseller, leasing telecom services from a provider (such as Western Union), and bundle electronic transmission service with data processing and letter delivery as a single product sold to a consumer. E-COM was intended to be a first...

---


69 PRC. “Opinion and Recommended Decision, MC78-3.” 24-30.
cautious step into providing electronic mail services, with the possibility of more expanded, fully electronic transmission and delivery hosted by the USPS to follow.\textsuperscript{70}

Large volume mailers and telecom companies, however, undermined E-COM and insured its failure. During PRC review of the proposal, associations of first-class mailers (including utility companies and greeting card manufactures) and private telecom companies (such as Graphnet Systems, MCI, Xerox, AT&T, and a host of others) argued that the risky new service would invariably and unfairly be subsidized with revenue from first-class rates, making first-class rates discriminatory by taxing regular users of first-class mail to support the risky and costly new venture and give the Postal Service an unfair leg up in competing with private telecom providers of similar electronic communications services.\textsuperscript{71} That is, interveners charged that the capital intensive E-COM would both unfairly burden first-class mailers and place private telecom services at a competitive disadvantage.\textsuperscript{72} In response to these arguments, the PRC radically

\textsuperscript{70} E-COM was supported by President Carter and the Postal Board of Governors as a way of extending and expanding postal service. See Administration’s Policy Statement July 19, 1979 included as Appendix B in Postal Rate Commissioner James Duffy’s “Dissenting Opinion and Recommended Decision, MC78-3.” The statement reads in part: “… as long as physical delivery through the mails exists as a primary means of communications to a large segment of the population, the USPS should take advantage of electronic communications to improve its service.” The Board of Governors, the PRC, and the Congressionally created Commission on Postal Service (created to review postal service) all concluded that integrating “electronic mail” and traditional letter mail was advisable. James Duffy. “Dissenting Opinion and Recommended Decision, MC78-3.” PRC. Mail Classification Schedule, Proposal 1978. Docket MC78-3. 1979; PRC. Mail Classification Schedule, Proposal 1978; Commission on Postal Service. \textit{Report of the Commission on Postal Service}. Vol. 1.


reconfigured the Postal Service’s proposal and, in its place, outlined a reworked E-COM service doomed to fail.\textsuperscript{73} Under the PRC’s E-COM plan which was put in place, the USPS had no control over the electronic transmission of messages, the costs of service would be borne fully by users of the service (cross subsidies would not be allowed), and total operational and start-up costs had to be recovered quite early, during the third and fourth years of operation.\textsuperscript{74} The reworked E-COM went live on January 2, 1982 and, despite steady growth in volume, operated for slightly over two years before being discontinued.\textsuperscript{75} The move to separate electronic transmission resulted in high transmission costs and the requirement to recover all costs during the third and fourth year led the PRC to outline 200% to 300% increases in rates for 1984.\textsuperscript{76} The Postal

\textsuperscript{73} The stinging dissenting opinions of Chairmen Duffy and Kieran O’Doharty make this point plain. Duffy notes that despite adopting the Board of Governors terminology “a careful reading of the majority opinion will reveal its true intent – to preclude Postal Service entry into electronic mail and to permanently trap the service in a subordinate role from which it cannot escape… it has recommended a data-processing, envelope-stuffing service, not an electronic mail service; [it] has rejected rather than accepted the concept of electronic mail service; [it] has contrived to postpone rather than encourage electronic mail service; [it] has done so by concealing rather than revealing the true important distinction.” O’Doharty’s dissent was no less forceful, noting that: “[t]he opinion issued by the majority today is a brazen attempt to arrogate the basic powers and responsibilities expressly reserved by statute for the management of the postal Service…It is a bold rejection of the Postal Service’s primary role in the development of new mail services for the American Public.” Duffy, “Dissenting Opinion and Recommended Decision, MC78-3.” 21; Kieran O’Doharty, “Dissenting Opinion and Recommended Decision, MC78-3.” PRC. Mail Classification Schedule, Proposal 1978. Docket MC78-3. 1979. 1.

\textsuperscript{74} PRC. “Opinion and Recommended Decision, MC78-3.” 1-59, 269-183.

\textsuperscript{75} In the first 6 months of operation 660,000 messages were sent. But by July E-COM was handling 172,000 messages weekly. OTA. Implications of Electronic Mail and Message Systems. 12.

Service surmised that the rate hikes would be fatal to the service and, rather than raising rates, abandoned the service. The future of the Postal Service would not involve e-mail.  

Mechanization and Automation: Centralized Processing and ICTs

While large-volume mailers and their associated allies seized on notions of unfair competition and discriminatory pricing to prevent the USPS from capitalizing on the emerging electronic communication market, the Postal Service did adapt new forms of ICTs for the purpose of processing mail. Mechanization and automation, the use of computerized systems in sorting the mail, became a regular feature of postal service during the 1970s and 1980s. The Postal Service invested over $6 billion in new equipment after reorganization and transformed the way in which mail moved through the postal network.

The Postal Service first invested heavily in mechanization before turning in the early 1980s toward automation. During the 1970s the Postal Service invested over $1 billion in new mechanized equipment, mainly multi-position letter-sorting machines and single-position letter-sorting machines (MPLSMs and SPLSMs), to help control labor costs. During the year before regulatory reform a mere 10% of total letters were processed on mechanized equipment and the process of sorting mail remained labor intensive. By the 1980, however, the Postal Service’s commitment showed real progress as 80% of total mail was now processed by mechanized equipment (see Table 3-2).  

Automation followed in earnest during the early 1980s. The old Post Office Department provided initial research and development funding to Philco and IBM during the 1960s for what would become the central technology in automating postal sorting and handling: Optical Character Readers (OCRs). OCRs perform a number of operations: they scan letters (or other material), “read” the address written on the piece, translate this information into a barcode, and spray the barcode on the piece for later sorting. Once bar-coded, bar code sorter machines (BCS) scans the code and sorts the mail for distribution. The Kappel commission discussed automation in its report as a possible key source of cost reduction. But, it was not until the Deputy Postmaster General’s Task Force on Mail Processing Systems offered its recommendations in 1976 that a clear automation strategy was developed. The USPS invested $5 billion into new automation equipment between 1980 and 1997, eventually acquiring over 13,000 pieces of hardware. The Postal Service rolled out first-generation OCRs in 1982 and updated versions, known as Multi-line OCRs or MLOCRs, in 1988. The key difference between these two systems relates to the amount of information that can be read and synthesized. While the earlier OCRs only read the lowest address line (including city, state, and zip code), MLOCRs read the entire address, match the address against an electronic database

83 Ibid.
containing a compilation of address information (names and addresses), and then apply a barcode containing the relevant information. Letters that cannot be read by the MLOCRs are supported by remote encoding centers. Video cameras take snapshots of letters that cannot be processed by MLOCRs, (due to illegible handwriting or other forms of presentation that cannot be identified) and transmit the image to remote encoding centers where operators read the address, manually input a barcode, and transmit the data back to the sorting facility where it is printed on the letter.\textsuperscript{84} MLOCRs and remote encoding centers allow for a greater degree of depth in sorting and enable the organization of mail into delivery sequence (known as delivery point sequence or DPS) without repeated sorting and resorting throughout the network.\textsuperscript{85}

The Postal Service’s commitment to and investment in mechanization and automation reflects its autonomy from the strictures of Congressional oversight. Importantly, unlike with E-COM, mailers did not voice strong opposition to internal improvements designed to lower processing costs. The $6 billion investment undertaken between 1970 and 1997 dwarfed any comparable investment undertaken by the Post Office Department under Congressional supervision. Between 1955 when modernization efforts first got off the ground and 1968, the Post Office Department invested $520 million total in large capital programs, including mechanization, automation, and new facilities and other equipment. Of this amount only a paltry $173 million was devoted to mechanization and automation. By comparison, under USPS stewardship, between 1970


\textsuperscript{85} DPS was first used in 1991. GAO. “Postal Service: Automation is Taking Longer.” 10.
and 1997, 10 times as much as was spent on total capital projects during the previous 13 year term was spent on automation and mechanization alone. When comparing investment in automation and mechanization only, USPS invested a staggering 34 times the amount of money spent between 1955 and 1968 in the two decades plus after regulatory reform.86 Though revenue and expenses also rose sharply after reorganization—the Postal Service by the late 1970s was dealing with significantly more mail, revenue, and expenses than the previous Post Office Department—the commitment toward automation and mechanization as a percentage of total annual obligations showed a dramatic and telling increase. Between 1955 and 1968, the Post Office spent on average .38% of total obligations on mechanization and automation annually, while between 1970 and 1997 the Postal Service devoted an average of .77% of annual total obligations.87

Mechanization and automation helped improve productivity and efficiency as postal volume exploded. Automation and mechanization speed up the throughput of mail processing and reduce the amount of workers needed to process the mail. Manual sorting processes 600 pieces per hour, mechanized sorting increases the rate of throughput to 2,800 per hour, and OCRs sort letters at a rate of roughly 30,000 per hour.88 In 1970, the Post Office Department processed 85 billion letters with a workforce of slightly over 501,000 employees, at a rate of 167,000 pieces per employee. By 1995 USPS carried

86 Figure calculated from data in Arthur D. Little. “A Description of the Postal Service Today: Appendix A to the General Contract.” President’s Commission on Postal Service. Towards Postal Excellence. Report of the General Contractor. Annex Contractor’s Reports Vol. 4. Table 4.1.11. 4-23; GAO. Postal Service: Automation is Restraining but Not Reducing Costs. 10.
87 Ibid.
88 GAO. Postal Service: Automation is Restraining but Not Reducing Costs. 10-12.
180 billion pieces of mail with roughly 875,000 workers, at a rate of 206,552 per employee. Automation and mechanization helped reduce the number of clerks employed, dropping from 308,000 and 42% of total workers in 1968, to 274,000 and 31% of total workers in 1995. By the middle of the 1990s, the Postal Service was processing over 659,000 letters annually per clerk employed, compared against roughly 275,000 per clerk during the late 1960s. Automation and mechanization both reduced the total number of workers needed and, importantly, reduced the number of full-time, career workers. Previously, sorting the mail relied on knowledgeable clerks (occasionally aided by mechanized “keying” equipment) and carriers to recall addresses, zip codes, and other date for disambiguation. Automated sorting deskills the work of clerks (in sorting the mail for distribution between offices) and carriers (in having to correctly sort letters for delivery) from an expanded knowledge of particular information concerning routes, addresses, and frequently changing addresses and allows temporary workers to approximate the performance of career employees. Finally, in terms of cost, new technologies helped reduce unit costs when adjusted for inflation. The comparative costs for sorting letters are $3 per thousand using automation, $19 using mechanization, and

81 The adoption of mechanized equipment before reorganization was limited. President’s Commission on Postal Organization. Towards Postal Excellence. 24.
82 As of 2003, 13% of postal employees are non-career workers. From 1980-2003, temporary labor increased by 292% (from 25,000 to 98,000) while during the same period career employees increased by only 13% (from 643,000 to 729,000). USPS repeatedly affirms its commitment to reduce costs through a smaller labor force and the deployment of temporary staff. Census Bureau. Statistical Abstract of the United States, 2004-2005. Table No. 1112.
$42 when manually sorted. In 1970, it cost the Post Office 35 cents in adjusted 1995 figures per piece of mail, while in 1995 unit costs were only 28 cents.

Functionally, automation transformed the sorting of mail and the organization of the postal network. The mechanization and automation program facilitated the centralization of sorting operations and the streamlining of the postal logistics network. While political pressure forces USPS to maintain an expansive (and expensive) network of retail post offices, sorting and processing have, since reorganization, become the province of a dwindling number of critical hubs. Before reorganization, the sectional center system divided the country into 552 sub-state districts, each containing a sectional center that served as a processing and distribution hub that linked to surrounding post offices (from which mail was collected and delivered) as well as other sectional centers. This network of hubs was reorganized in accordance with the new efficiencies automation afforded. As of 2001, the sectional centers had been roughly halved into 237 Processing and Distribution Centers (P&DCs), which serve as the key points in-transit for mail as it moves through the postal system and as entry points where bulk mail is

---


96 The number of post offices faced a sharp decline in the early part of the 20th century with the introduction of the automobile. Since reorganization, the number of post offices has hovered around 30,000 with little change. The PRA specified that no post office could be closed due to an operating deficit. The USPS instituted a moratorium on closing small rural post offices between 1998 and 2002. Robert H. Cohen, Matthew Robinson, Renee Sheehy, Tom Sharkey, John Waller, and Spyros Xenakis. "The Conflict about Preserving Small Rural Post Offices: Differences in the Distribution of Pharmacies and Post Offices." Presented at Winton M. Blount Symposium on Postal History. Washington D.C., Nov. 2006.

commingled with the general mailstream. The pursuit of efficiencies provided by centralizing processing and handling in PD&Cs reconfigured the postal network into a more tightly integrated spoke-and-hub network. These centralized facilities now are the critical hubs that link roughly 40,000 post offices, 350,000 collection boxes, 136 million unique delivery addresses into a national network that process 680 million pieces of mail daily. Individually P&DC sit as the gateways into large cities and regions and connect local post offices with the broader postal network.

The Triumph of Large-Volume Mailers: The Deregulation of Postal Service

The remade postal service is an example of selective and incomplete deregulation. While the letter monopoly still exists and the USPS is a public entity, substantive regulation is largely absent. The equation of value with market determinations is at odds with the foundational ethos of postal service that made explicit the aim to provide services in excess of what the market would provide. By 1984, with abandonment of E-COM, the idea that postal service could embody such notions appeared anachronistic. Currently, the service is in a precarious position: automation and the explosion of ad mail that fed the phenomenal growth in volume during the 1980s and 1990s provided some

---


100 Figures are approximate and current to the fall of 2001. USPS. Emergency Preparedness Plan. 1-3.

cover, but long-term trends suggest a dire position. Costs continue to rise, while volume, finally, appears to be stalling and actually falling.\footnote{102}

Despite the failure of E-COM, the postal service was not completely shut out of the information revolution. Regulatory reform did enable postal management to adopt new ICTs into their operation. But the ethos of deregulation guided and limited precisely how these new technologies would be used, ultimately pushing the Postal Service to use ICTs as a tool to streamline the processing network while blocking expansion into growing electronic communications markets. Mechanization and automation provided some clear productivity and efficiency gains. The deployment of new ICTs advanced in tandem with the recalibration of the postal network into a more tightly integrated hub-and-spoke network structured around processing and distribution centers in order to capture newfound economies associated with increased throughput and lower unit costs.

*The Limitations of Centralized Processing: Tight Coupling and Modes of Failure*

The reorganization of the postal network into a system of interlocking dense hubs that process and route the mail has, without question, improved the efficiency of the postal network. Over the course of the past several decades, postal volume has sharply
risen while costs and levels of service have remained fairly consistent or improved. As a result of automation more mail is now processed in fewer locations, quicker, and at a comparatively lower price. Yet, while the benefits of automation—and by extension of regulatory reform—are plain, improved efficiency is not without its costs. By moving from a distributed to a centralized or tightly-coupled processing model, the postal network, now, is more vulnerable to infrequent large-scale failure. The hub-and-spoke architecture of small world networks work well under normal operating conditions. While random failures at single points within the network, for the most part of little concern, a failure that occurs within a dense hub can have an inversely large impact on the larger network. To provide a germane example, the disruption of mail service along any single postal route, home, or even branch office would not in all likelihood impact the corresponding service of those connected to the network. The failure would be minor and relatively trivial. Yet, a failure at one of the 200 plus processing hubs that now make up the backbone of postal service would result in a non-trivial failure. The density of the connections associated with centralized processing plant creates choke points where failures have the possibility of impacting a large swath of connected users (i.e., those linked to and served by the incoming and outgoing mail flows of the plant). The aggregation of processing activities in fewer locations creates new interdependencies within the network that also increase the possibility of large-scale failure. While reform created a more efficient postal service it also created the new possibility of systemic

failure that had been largely absent. The possibility of such a failure, however, ceased to be purely theoretical during the 2001 anthrax attacks (see Chapter 5).

*The Deregulation of Railroad: Complexity and Proximity*

As outlined in detail in Chapter 2, ICC regulation of freight transportation provided a measure of resilience. Though largely as an unintended side-effect, the structure of freight regulation discouraged the shipment of large quantities of dangerous chemicals in close proximity to densely populated areas. Freight regulation supported resilience through two different means: (1) it encouraged the substitution of smaller shipments of dangerous materials carried via truck over the transport of bulk rail shipments; (2) it provided a circuitous rail network through which residual rail shipments could move outside of close proximity to densely populated areas. With regards to the shipment of dangerous materials, freight regulation supported the maintenance of what Charles Perrow would describe as (mostly) discrete linear systems.

During the 1970s and early 1980s, however, selective deregulation took aim and eventually undid these historic supports. Similar to post, deregulation freed management to pursue the elimination of redundancies and the creation of more efficient network. In the ensuing years, the scope of the rail network was radically reduced and patterns of inter-modal commodity distribution shifted, leading to a spike in the volume of large quantities of dangerous materials moving by rail. Deregulation, unquestionably, saved the failing railroad industry by allowing railroads to shed costly track and recapture traffic that had been diverted to other modes of transportation. But increased efficiency carries a price: by eliminating seemingly inefficient track and encouraging bulk
shipments of chemicals via rail, deregulation undermined the overlooked and incidental security benefits intermixed with regulation.

*The Push for Rail Deregulation*

The process of deregulation unfolded through a series of new laws and administrative decisions during the 1970s that culminated in the Staggers Rail Act of 1980. Deregulation radically altered the status quo of freight regulation by granting railroads wide latitude to set competitive rates and determine the scale of the networks they operated. The push for deregulation came largely from the railroads themselves as widespread economic challenges eroded the positive benefits secured through regulation, while well-publicized criticisms from a cross-section of right and liberal-left observers increasingly painted regulation as both anti-business and anti-consumer.

The push for deregulation was startling; for decades the railroads were the fiercest supporters of ICC authority and the continuation of regulations. Likewise, consensus between the ICC, shippers, and Congress historically stood in support of the status quo. Indeed, for a time regulation worked well in securing tangible benefits for the railroad industry and the public. Yet, by the 1970s, regulation no longer appeared to work quite so well. The steady introduction of intermodal competition offered the public cheaper and more attractive alternative shipping options which led to the erosion of rail profits. At the same time, economists attacked regulation as an inefficient relic, while liberal-left consumer advocates challenged regulation as a form of corporate welfare. It was against this complicated backdrop that deregulation unfolded.

*The Durability of Freight Regulation*
Historically, ICC regulation provided clear and obvious benefits for railroads. Price-and-entry controls protected railroads from destructive competition and the vagaries of regulation on a state-by-state basis, and guaranteed a generous rate of return. At the same time, regulation was not simply a proxy for railroad interests, but also provided shippers, and by extension the public, with clear benefits. Regulation protected shippers from discriminatory rates and unfettered rent-seeking on monopoly lines, and maintained an expansive network that provided the public with access to a national system of freight transportation. Under value-of-service pricing and the general system of cross-subsidies, to be sure, certain shippers and segments of the population benefited more than others—farmers and shippers of cheap bulk products received rates that were quite low when compared against expensive manufactured goods and the maintenance of an extensive rail system was a boon to rural interests—but in general ICC regulation of freight transportation provided broad benefits to both railroads and the public.\footnote{104}{For a detailed discussion, see Chapter 2.}

Railroads, key shippers, and the ICC in turn, supported the continuation and at times extension of ICC regulation. In the face of periodic challenges, the railroads were staunch advocates and protectors of ICC authority.\footnote{105}{For a clear discussion of the railroads proffering of key support for the ICC, See, Samuel P. Huntington. “The Marasmus of the ICC: The Commission, the Railroads, and the Public Interest.” \textit{Yale Law Journal} 61.4 (1952) 467-509.} Shipping interests, particularly farmers and the U.S. Department of Agriculture (USDA), also supported the continuation and extension of ICC regulation.\footnote{106}{James C. Nelson. “The Effects of Entry Control in Surface Transport.” \textit{Transportation Economics}. 1965. NBER. 381-422; Ari Hoogenboom and Olive Hoogenboom. \textit{A History of the ICC: From Panacea to Palliative}. New York: Norton, 1976.} Finally, the ICC itself became an interest group...
wedded to and protective of the status quo. In a familiar pattern, regulation encased in law became subject to self-reinforcing mechanisms that helped ensure its durability. In a crude calculus, regulation and the ICC supported railroads and shippers, while railroads and shippers supported regulation and the ICC.

Intermodal Competition and Regulation: Cracks in the Consensus

Widespread economic problems, coupled with growing anti-regulatory sentiment from industry, observers, and government, pushed deregulation as a means to save the railroads. The introduction of intermodal competition destabilized the system of freight regulation. Cross-subsidies and value-of-service pricing function well in monopoly markets with little opportunity for consumer exit. By securing a large market and limiting alternative options, the added costs of regulation are spread across a large number of users and customers are compelled to pay only a relatively small markup. In the case of freight rail regulation, for example, the costs of maintaining an expansive rail network and supporting low rates for bulk agricultural commodities were distributed across customers and shippers of other commodities. The introduction of competition, however, undermines this system. Cross-subsidies and value-of-service pricing create incentives for users to opt out of the regulated system in favor of services not saddled with the added cost burdens due to an absence of regulation or a comparatively smaller markup. As customers flee to other services, the system of regulation becomes increasingly tenuous: operators are left with low value customers and are forced to now

---

107 Hoogenboom and Olive Hoogenboom. A History of the ICC. 118.
spread costs across a smaller fraction of customers, in turn creating even greater incentives for customers to opt out.\textsuperscript{108}

The introduction of new modes of transportation, motor carriage most clearly, functioned precisely in this manner. The growth of trucking service led to the substitution of traffic from rail to motor carriage, particularly for high value manufactured goods that shouldered the greatest burden of rail costs under value-of-service pricing.\textsuperscript{109} The railroads, in concert with large trucking firms and unionized labor leery of hit-and-run competition from smaller outfits, successfully urged for the extension of ICC regulation over motor carriage.\textsuperscript{110} For the railroads, placing motor carriage within the ICC’s ambit attempted to secure their market position and the preservation of the rate structure. The Motor Carriers Act of 1935 extended ICC regulation of entry and rates to motor carriage, but contained a number of exemptions that in time proved significant. As a concession to farming interests, the Act exempted the motor carriage of unprocessed farm products—fruits, vegetables, and livestock—from ICC regulation. Additionally, the Act only covered common carriers, and exempted private carriers and contract carriers.\textsuperscript{111}

Under the Motor Carriers Act, the ICC sought to maintain the viability of the railroads

\textsuperscript{108} On the difficulty of maintaining cross-subsidies in the face of opportunities for bypass, see, Horwitz. \emph{The Irony of Regulatory Reform}. 7-8.


and motor carriage through rate parity. That is, the ICC attempted to maintain the basic structure of rates that included cross-subsidies and value-of-service pricing by preventing either mode from drastically slashing prices in intermodal competition. Yet, the exemptions proved all important. A thriving motor carriage sector operating outside of regulatory constraints siphoned freight traffic from rail service and steadily limited the share of inter-city freight traveling by rail. The ICC’s attempt to protect railroads via rate parity, ironically, hastened the shift from rail to motor. Rail had a real advantage in terms of cost over motor for long-haul shipments. However, the ICC’s devotion to parity prevented railroads from competing on price and forced them into service competition. In service competition, motor, with its flexibility, had a clear advantage that rail could not easily best.

The impact of intermodal competition on the health of railroads was stark. In 1935, railroads carried 75% of all intercity freight, but by the late 1970s, that number slipped by more than half to a low of 27%. Motor carriage increased sharply during this period and a sizable share—above 40% by the 1970s—operated exempt from regulation. As traffic declined, railroad profits plummeted. During the 1940s, railroads return on investment averaged a decent 4.1% before steadily declining to 2%

---

114 Ibid.
116 GAO. *Railroad Regulation*. 12.
during the late 1970s, showing profits lower than any other major U.S. industry (see Table 3-3). Bankruptcy threatened to sink the national rail network. By the late 1960s and early 1970s, nearly a quarter of total track was in bankruptcy and key major lines in the Northeast and Middle West, including Penn Central, Milwaukee Railroad, and Rock Island, plunged into bankruptcy. Railroads, long advocates of regulation, now became vocal champions for deregulation as a means of salvaging the industry from ruin. Through the 1970s, railroads lobbied and agitated for deregulation and increased freedoms to set rates and abandon track as a way of salvaging the industry through the newly (1970) formed Committee for Efficient Transportation (COMET), renamed the Committee Urging Regulatory Reform for Efficient National Transportation (CURRENT) in 1975.

At the same moment that intermodal competition wrecked havoc on the rail system, observers from across the political spectrum began to revisit the role of the ICC and the desirability of regulation. A series of high-profile works challenged the effectiveness and necessity of regulation under the ICC. Beginning in 1959, with John R. Meyer and colleagues’ *The Economics of Competition in the Transportation Industries*, a steady stream of economic studies argued that ICC regulation caused the railroad industry

---


118 The failure of specific railroads presented a more general problem. The railroad network is interdependent, freight moves across and between different carriers through exchanges to reach its end point. Thus a failure of a sizable carrier has effects that impair the larger network. AAR. “A Short History of U.S. Freight Railroads.”; Darius W. Gaskins. “Regulation of Freight Railroads in the Modern Era: 1970 – 2010.” Review of Network Economics 7.4 (2008); GAO. Railroad Regulation. 11.

119 These committees were formed under the guidance of the Council of Economic Advisors and released a “Bicentennial Bill of Rights for railroads” on the nation’s bicentennial Stone. The Interstate Commerce Commission. 54-55, 85.
to lag behind other industries, inflated prices and cut into profits, and led to the inefficient use of transportation facilities. Studies in the 1960s and 1970s estimated that yoke of unprofitable track associated with ICC regulation of freight railroads led to $1 billion in waste annually. The rates charged by exempt motor carriers provided ample evidence of the limitations of regulation, typically running at between 10% to 36% less than comparable regulated shipments. In 1970, Robert C. Fellmeth’s high-profile and sensational The Interstate Commerce Omission, published under the imprint of the Ralph Nader Study Group, attacked the ICC as an example of a regulatory agency “captured” by industry that served the interests of the railroads at the expense of the public. Paired with Gabriel Kolko’s recent revisionist history, Railroads and Regulation, 1877-1916 (1965), these works charged that regulation did not ensure the public interest as much as thwart it. Criticisms from both the left and right argued that regulation, once seen as a basic solution to the unique economic problems posed by characteristics of the industry, was at best an outmoded relic that now promoted inefficiency and at worst the handmaiden of the industry that it was ostensibly designed to regulate. These criticisms found sympathy in Congress and the Executive branch beginning with the election of

President Kennedy. By the late 1960s and early 1970s, regulation appeared both anti-industry and anti-consumer. During the Presidential campaign of 1976, Jimmy Carter seized on the issue as a key part of his platform, promising to undertake the deregulation of transportation. “Deregulation,” as a policy option, was firmly ingrained in mainstream policy discussions.

*The Deregulation of Freight Railroad Transportation*

The deregulation of freight rail transportation proceeded through new legislation and, to an extent, administrative decisions. In the immediate wake of the failure of rail service in the northeast, Congress passed the Regional Rail Revitalization Act of 1973 (known as 3R) to reorganize the seven bankrupt lines into what would be known as Consolidated Rail Corporation (Conrail). Sweeping reform and deregulation followed with the Railroad Revitalization and Regulatory Reform Act of 1976 (4R) and the Staggers Rail Act of 1980. 4(R) enacted Conrail and provided limited relaxation over rates, while Staggers further relaxed control over rates and reordered the procedures for securing abandonments and divesting under-used track. In addition to enacting Conrail, 4(R) specified just and reasonable rates would now only have to be equal to or greater than variable cost of providing service. Additionally, the Act specified that rates would not be considered excessive unless a finding of conditions of market dominance, i.e., lack

---

125 The Landis report prepared for President-elect Kennedy in 1960 offered a strong critique of the regulatory agencies and Congress issued the similarly stinging Doyle Report of the Senate Transportation Study Group in 1961. In 1962, President Kennedy issued a special Transportation Message to Congress, arguing that current regulations were “a chaotic patchwork of inconsistent and often obsolete legislation” in need of reform. Stone. *The Interstate Commerce Commission.* 48-54.


of competition. Finally, the Act also granted railroads the ability to adjust rates by 7% (the “yo-yo” provision) during the year without onerous hearings and provided the ICC with the power to exempt portions of service if, in their estimation, regulations would be an undue burden and serve little useful public purpose.\textsuperscript{128}

Between 4(R) and Staggers, the ICC used its new found autonomy to further deregulate portions of freight transportation. In 1978, in response to a request from Southern Pacific Railroad, the ICC unanimously voted to remove regulation over rates for fruits and vegetables, commodities exempt from regulation when carried by motor carriage. The ICC also began to allow contract rates, private rates negotiated between shipper and carrier, in certain limited situations.\textsuperscript{129}

Staggers extended the limited freedoms granted under 4(R) and amplified by the ICC in the intervening years by providing railroads increased freedom to set competitive rates and eliminate under-performing track. Staggers exempted 2/3 of all rates from max rate regulation and in cases where railroads had market dominance rates could be set at 160% of variable cost (with provisions to eventually escalate to 180%). The railroads, however, remained common carriers and could not refuse service. The Act provided a greatly narrowed definition of market dominance which placed the burden on shippers to prove captivity and benefited the railroads. Contract rates were explicitly made legal, while joint rates, a form of antitrust exemption, were no longer allowed. Importantly, the policy on abandonments was also changed in favor of the railroads. Under Staggers, the ICC was directed to rule on abandonments within 75 days if uncontested and within 225

\textsuperscript{128} Stone. The Interstate Commerce Commission. 77-78.
\textsuperscript{129} Ibid. 88-90.
days if requiring an investigation. Previously, there were no limits on the length of
time governing rulings and hearings averaged a staggering 400 days. Staggers, as the
culmination of nearly a decade’s work to relax regulation, reordered the landscape within
which railroads operated. Now, for the first time in nearly a century, railroads had the
ability to set rates in competitive markets and define the scope of their networks with a
relatively freehand.

Post-Staggers: Railroads Revitalized

The sweeping regulatory changes which crested with Staggers revitalized and
salvaged the rail industry. The relaxation of control over abandonments and rates
allowed the rail industry to eliminate unprofitable track, recapture significant traffic
volume, and pull back from the brink of ruin. In the decades following Staggers, total
U.S. rail mileage sharply declined from 200,000 miles at the close of the 1970s to
140,000 miles by 2006, levels comparable to the rail network of the 1880s. The most
significant declines resulted from Class I railroads (the largest railroads) abandoning 1/3
of total track, reducing the mileage they operated from roughly 150,000 to 94,000 miles
between 1980 and 2006. The reduction of the expansive network helped reduce

130 Stone. The Interstate Commerce Commission. 113-141; GAO. Railroad Regulation. 15; Gaskins.
131 4(R) also attempted to speed up the abandonment process. Stone. The Interstate Commerce Commission.
97.
<http://www.aar.org/PubCommon/Documents/AboutTheIndustry/Overview.pdf>; Jean-Paul Rodrigue,
133 Dramatic consolidation of Class I railroads also followed Staggers. In 1980 there were 40 Class I
operating costs by nearly 60% by eliminating maintenance costs of non-remunerative track.\textsuperscript{134} No longer saddled with unprofitable lines or obligated to build cross-subsidies into rates, railroads successfully reversed decades-long traffic declines by slashing rates.\textsuperscript{135} Post-Staggers, rail rates for all commodities fell by 55% and, correspondingly, the volume of intercity freight carried by rail, measured in ton-miles, doubled.\textsuperscript{136} Railroads not only added volume but also recaptured a significant market share from competing modes. Recall, as noted in Chapter 2, by the end of the 1970s, rail’s share of total intercity freight had fallen to a historic low 27%. By 2010, railroads had regained 43% of total intercity freight and were a viable competitor with other modes of transport.\textsuperscript{137} Profitability followed reduced costs and increased traffic. Industry rate-of-return rebounded from the dismal low of an average level of 2% in the 1970s to 8% during the first decade of the 21st century (see Table 3-3).\textsuperscript{138}

Regulatory reform helped to overcome the dire state within which the industry was mired during the 1970s. The new power to shed unprofitable track and set rates was a boon to both industry and consumers: rates fell, volume and market share improved, and profits rose. In the aftermath of Staggers, U.S. railroads were, without a doubt, more efficient. The density of the U.S. rail system rapidly increased, as more product moved

\textsuperscript{134} In concert with the relaxation of abandonments, after 1980 the ICC moved to eliminate Detroit, Toledo, and Ironton conditions (DT&I) which compelled railroads to preserve patterns of freight movement in the wake of mergers. Gaskins. “Regulation of Freight Railroads.”
\textsuperscript{135} Significant savings came from reductions in labor as well as the elimination of unprofitable track. Carlton and Picker. “Antitrust and Regulation.”
\textsuperscript{136} The use of contract rates helped, in part, drive the decline in rates. By 1988, 60% of all Class I freight moved under contract rates. AAR. “Overview of America’s Freight Railroads.”; Stone. The Interstate Commerce Commission. 161.
\textsuperscript{138} Data from AAR. “A Short History of U.S. Freight Railroads.”
across a smaller network and costs of service decreased.\textsuperscript{139} Recent evaluations of the impact of deregulation have tabulated the annual savings directly attributable to regulatory reform to be between $7 and $10 billion, far in excess of even the most optimistic predictions during the run-up to reform.\textsuperscript{140}

\textit{Deregulation and Complexity: Revisiting an Efficient Rail Network}

The triumphs of deregulation, however, came at a price. Deregulation, as in other infrastructure sectors, improved the efficiency of the rail network at the expense of creating new (or enlarging existing) vulnerabilities and undermining resilience. Just a postal reorganization enabled management to create a more efficient network, so too did the deregulation of rail embolden railroads to optimize rail operations and network. As part of the general improvement of market share and volume gains, deregulation supported increasing rail shipments of hazardous materials and fostered intermodal substitution from small hauls traveling via motor to large bulk rail shipments. Additionally, regulatory reforms that hastened the reduction of the scale of underused, unprofitable track, made it more difficult to ship dangerous materials along pathways that skirt densely populated areas. The pruning of the rail network helped reduce costs and make the rail network a more efficient system of pathways and interconnections. However, by eliminating thousands of miles of track, now, rail shipments of dangerous materials find it increasingly difficult to travel across paths that are not in close proximity to dense population centers. The process of deregulation bound together two largely

\textsuperscript{139} For a detailed discussion of the impact of deregulation on density, see Bitzan and Keeler. “Economies of Density.”

\textsuperscript{140} Ibid. 173-175.
discrete systems—the shipment of large quantities of dangerous materials through the rail network and population centers—into a complex formation. This shift eroded the observed security benefits associated with regulation.

Gains made by rail in capturing intercity freight are reproduced within the market for the transportation of dangerous materials. The post-Staggers recovery by railroads of general freight volume and market share is matched by increasing volume and share of hazardous materials shipments. Post-Staggers, the total volume of hazardous materials transported by rail doubled just as total freight transported by rail also doubled during the same period. By 2007, rail and truck divided a nearly equal share of transportation of intercity hazardous materials, rail carrying 28.5% and truck carrying 32.2% of ton-miles. Significantly, rail rebounded to carry a significant share of the most dangerous commodity-type, and those which would be the focus of the most intensive security interventions, poisonous liquids and compressed gases labeled as toxic-inhalation-hazards (TIH; see Chapter 6). Looking closer at specific commodity type, in 2007, rail carried 51.5% of basic chemicals, 57.9% of toxic poisons (Class 6 Hazardous Materials), and 40.7% of gases (Class 2 Hazardous Materials); while motor carriage transported

---

141 Post-Staggers, during the 1990s, the Department of Transportation reconfigured how they tabulated commodity flow, after surveys during the 1980s were considered incomplete and abandoned due to quality concerns. As part of this reconfiguration, now the economic census provides fairly detailed and specific information about commodity flow.


29.3% of basic chemicals, 31.3% of toxic poisons, and 37.5% of gases.\textsuperscript{144} The figures were all record highs for rail and record lows for trucking since the tabulation of hazardous material data began in the 1990s, with the exception of Class 2 explosives where trucking appeared to gain some ground. Though cost is not the only factor determining shipping decisions, deregulation plainly helped encourage the adoption of rail as mode for transporting dangerous materials, as it did for commodities more broadly.

The substitution from truck to rail changes the hazard associated with the transportation of dangerous goods. A number of key factors define the risk of transporting hazardous materials: frequency of disruption (accident or intentional release), size of shipment, toxicity or volatility of substance, and population density in area of release.\textsuperscript{145} Railroads are prone to less frequent accidents than motor carriage, but the consequences of accidents involving rail shipments of dangerous materials far outstrip truck accidents. The characteristics of hazardous materials shipments differ in important respects between modes. Rail shipments of hazardous materials are coupled

\textsuperscript{144} Commodity flow data from the economic census are not available for the 1980s. Additionally, the categories listed changed in the mid-1990s, making direct comparisons between eras difficult (for example, in the 1990s flow data about hazardous materials was collected, while in the 1960s and 1970s no such data was collected). However, the basic picture indicated—of increasing rail volume and increasing share of transported commodities both generally and for dangerous materials—appears to be valid. In 1993, when data and categories comparable to those used in previous surveys were last employed, rail carried 59.9% of total ton-miles for industrial chemicals compared with just 20.7% carried by truck. This shows a sharp uptick of the trend observed during the 1960s and 1970s when rail’s share declined from a high of 67.1% to 57.4% and truck increased its share from a low of 15.6% to 24.9%. DOT. “Hazardous Materials.” 2007 Economic Census. Tables 1a and 6a; DOT. “Hazardous Materials.” 2002 Economic Census. December 2004. Tables 1a and 6a. DOT. 1993 Economic Census. Oct. 1996. United States. Census Bureau. Commodity Transportation Survey. Economic Census. Washington, D.C.: GPO, 1963, 1967, 1972, 1977.

into much larger bulk shipments and travel greater distances, while trucks transport small quantities between comparatively short distances. The average rail shipment of hazardous materials is 87.81 tons and travels 578 miles, compared against 4.82 tons and an average of 59 miles for shipments via truck.\textsuperscript{146} Accident rates involving rail shipments of hazardous materials are significantly lower than for truck. Despite comparable ton-mileage, incidents involving truck are 16 times more likely.\textsuperscript{147} However, a single accident of an average rail shipment of TIH is an order of magnitude more severe in terms of loss of life and injury than a single average truck shipment due to the differences of shipment size.\textsuperscript{148} That is, per average incident, 10 times as many individuals are exposed, injured, and fatally exposed during a rail accident as compared with a motor accident.\textsuperscript{149} Outliers, extreme cases involving rail shipments, further skew the scale of the risk involved with rail transportation. Rail shipments of large quantities of TIH, and hazardous materials more generally, are a low-probability, high-consequence event. Though rare in occurrence, a single accident involving rail shipment can result in catastrophic consequences.\textsuperscript{150} For example, as a recent much publicized report from Dr. Jay Boris of the Naval Research Laboratory noted, a release from a rail shipment of TIH


\textsuperscript{148} Railroad incidents involving TIH account for 78\% of total injuries and fatalities during a 10-year period. Brown, Dunn, and Policastro. \textit{A National Risk Assessment}. 4, 179.

\textsuperscript{149} Ibid. 148-150.

\textsuperscript{150} For an extensive discussion of the cross-modal comparisons of the risks of transporting hazardous materials and TIH, see: Brown, Dunn, and Policastro. \textit{A National Risk Assessment.}
in a densely populated area under ideal weather conditions could result in 100,000 fatalities.\textsuperscript{151} In a sense, the danger of transporting TIH and hazardous materials via rail is understated, dramatically so, by the historical record due to the relative infrequency of a serious disruption under ideal dispersion and exposure conditions. While accidents involving truck shipments of hazardous materials occur with greater frequency, they are, on average, localized and small-scale. Shipments via rail exhibit opposite characteristics—they are infrequent and potentially catastrophic—and present a different set of challenges.

The derailment of railcars transporting chlorine in Graniteville, South Carolina provides a glimpse into the possible scale of a major accident involving TIH. In January, 2005 an accident involving railcars carrying chlorine led to 9 deaths, the hospitalization of over 500 residents, and the displacement of over 5,000 people living near the accident site. The total costs from the accident were estimated at $126 million.\textsuperscript{152} The accident, though, could have been far worse. The derailment occurred during the middle of the night near a small town. A similar accident near a population center could have been catastrophic.

Here, the security costs of deregulation can be seen. The changes ushered in through deregulation have, for the most part, delivered a host of benefits for shippers, the


railroads, and the public. Yet, the improved efficiency of the rail network is not without its costs. The revitalization of rail service has led to the increasing use of rail as a mode of transportation for the shipment of dangerous materials. The move toward larger shipments via rail reduces the likelihood of frequent small-scale incidents but increases the long-term possibility of a catastrophic event. Additionally, the slashing of excess track across the rail network undercuts the added security afforded by an expansive rail network. A lean network that increasingly runs between heavily trafficked points and dense population centers, limits the routing options afforded to shipments of dangerous materials. Under these conditions, skirting densely packed population centers—that is utilizing circuitous routes that eschew the most efficient paths in favor of routes with comparatively low population densities—is made more difficult and less likely due to the elimination of seemingly excess track. In Perrow’s terms, deregulation encouraged complexity over linearity. Regulatory reform inadvertently worked to fuse in close proximity two (mostly) independent systems in ways that made new types of large-scale failure possible.

*The Deregulation of Electric Power: Tight Coupling and Complexity*

The partial deregulation of electric power remade the resilient network-of-networks into an increasingly fragile system open to large-scale failures. Under regulation, electric power developed as a system of (mostly) privately owned utilities operating vertical monopolies (combining all three phases of electric power, generation, transmission, and distribution) serving a limited territory. Regulation supported

---

153 For a discussion of population density and the risk of hazardous materials transportation, see Brown, Dunn, and Policastro. *A National Risk Assessment.*
resilience by encouraging distributed control—each local utility maintained control for only one small corner of the larger interconnected power grid—and underwriting the development and use of sophisticated ICTs. Restructuring, however, undid these supports and encouraged the centralization of control and transformed the use of ICTs in favor of standardized and accessible systems. Both shifts made the interconnected electric power system more susceptible to large-scale failures. Just as in post and rail, economic restructuring carries an often overlooked price: vulnerability.

The Grow-and-Build Consensus: Post War Stability

Between World War II and the late 1960s, the regulated system of electricity production worked remarkably well. The Public Utility Holding Company Act of 1935 (PUHCA) and state regulation provided privately owned utilities, investor-owned utilities (IOUs), with monopoly protection and a guaranteed rate of return, while imposing basic limits to protect the public from the volatility endemic to the holding company scandals of the 1920s and 1930s, deliver reasonable rates, and expand reliable service. A broad consensus between the public, utilities, and regulators supported the status quo—what historian Richard Hirsh refers to as the “utility consensus”—for decades. Under this consensus, utilities flourished, power consumption increased, and rates fell as new technological advances led to decreasing production costs. As with other infrastructures, regulations, once in place, obtained an inertia and momentum that made them difficult to dislodge or substantially redraw.

---

Under regulation, utilities pursued what is described as the “grow-and-build strategy” of promoting increased consumption and lower costs through technological improvements. Utilities invested in new large generators and increased thermal efficiencies to lower unit costs. In simple terms, the same amount of fuel could produce 50% more energy in the 1960s than it could during the 1940s. At the same time, utilities lowered unit costs by investing in larger fossil fuel generators. Thanks to these advances, as well as improvements in transmission and distribution, productivity increased by 5.5% per year, between 1889 and 1963, outpacing all other sectors of the economy. New efficiencies translated into precipitously declining consumer prices: residential prices fell from (adjusted) 453 cents per KWh in 1927 to 62 cents in 1927, 47 cents in 1937, 22 cents in 1947, to 9 cents in 1969. Utilities further promoted consumption and demand (and increasing capacity) through declining block rates that fell with increasing consumption. Unsurprisingly, demand jumped as prices fell.

---


156 Recall, as noted in Chapter 2, the structure of rate-of-return regulation rewarded investment in new technologies.

157 During the late 1940s plants converted roughly 22% of a fuel’s content to energy, but by the 1960s advances in steam temperatures and turbines allowed plants to operate with a thermal efficiency of 33%. Hirsh. Power Loss. 56.

158 Capacity of the largest units jumped from 190 MW to 575 MW between 1950 and 1960, while average cost of new plant construction dropped from $173 per KW to $149 per KW in adjusted dollars. Hirsh. Power Loss. 56.

159 Ibid. 47.


increasing by 12% annually between 1900 and 1920 and 7% annually between 1920 and 1973.\textsuperscript{162}

Under the grow-and-build model, both consumers and utilities were satisfied and saw little reason to revisit the basic terms of service. Regulators, for their part, saw little need to be more than passive observers of utility behavior as long as industry and consumers found few points of contention.\textsuperscript{163} Indeed, a mass of support developed around the utility consensus that made revision both difficult and largely unthinkable. The utilities and the public both received tangible benefits under the devised system, while ancillary interests, including manufacturers of electrical appliances, hardware manufactures that supplied utilities with equipment, investors, and regulators also supported the maintenance of the status quo.\textsuperscript{164} In a familiar pattern, regulation appeared insulated from change. After the tumult and upheavals of the early decades of electrical power which had pitted consumers, industry, and government into increasingly confrontational configurations, the utility consensus provided a respite.

\textit{Cracks in the Consensus: Technological Stasis, the Oil Crisis, and the Emergence of Environmentalism}

Support for the status quo, however, came unwound during the late 1960s and early 1970s as technological improvements appeared to reach their limits, the 1973 oil crisis rocked energy markets, and the newly sprouted environmental movement


\textsuperscript{163} Hirsh. \textit{Technology and Transformation}. 85-86.

\textsuperscript{164} Hirsh. \textit{Power Loss}. 1; Hirsh. \textit{Technology and Transformation}. 82-86.
challenged escalating power consumption. These developments undermined the foundations upon which support for the grow-and-build strategy rested and led to the cautious introduction of new competition into the field of electric power generation with the passage of PURPA.

By the late 1960s and early 1970s, steady improvements in thermal efficiencies and generating unit size appeared to hit their limits and provide diminishing returns. Utility managers could no longer rely on improving equipment to boost productivity in support of falling prices and increased supply. Technical limitations could not have arrived at a less opportune time for utilities. The 1973 energy crisis upturned energy markets globally and hit the U.S. electricity market hard. U.S. utilities had increasingly adopted petroleum as a fuel source after 1960, increasing the total share of generation from petro from 6.9% in 1962 to 20% in 1973. During the crisis, price of oil leapt from $1.77 per barrel to $12 and led to widespread inflation and stagnation across the economy (stagflation). The shortage and rising price of oil, helped reverse the trend toward falling electricity rates, as residential rates (in unadjusted terms) increased from 2.1 cents per KWh in 1969 to 4.4 cents a decade later (industrial prices, likewise,

---


167 The Organization of Petroleum Exporting Countries (OPEC), in retaliation for support of Israel in the Yum Kipper War, cut oil production, increased prices, and subsequently embargoed oil shipments to the U.S.


169 Ibid. 60.
increased from .9 to 2.6 cents per KWh over the same period). The energy crisis, in an unprecedented development, led to an increase of consumption of a then unthinkable paltry .1% from 1973 to 1974 and a very modest 1.9% the following year.\textsuperscript{170}

The emergence of the environmental movement in the 1960s and increased activism in the 1970s revisited the promotion of continued growth of power consumption. The Clean Air Act (1963) and the National Environmental Policy Act (1970), which created the Environmental Protection Agency, reflected the growing awareness of environmental issues and increasing influence of organized activists (as did the celebration of the first Earth Day on April 22, 1970).\textsuperscript{171} The National Environmental Policy Act, subsequently amended in 1974 and 1977, provided public agencies with powers to review and scrutinize environmental impacts of new projects and offered forums for public comment. In the realm of electric power, now, regulatory review of utility projects became contentious as opponents both from the environmental movement and the consumer movement intervened.\textsuperscript{172} Through the 1970s, activist successfully challenged the grow-and-build model and introduced principles of conservation and demand management (limiting consumption) into utility planning.\textsuperscript{173} These gains would be extended in subsequent years and aided by PURPA.\textsuperscript{174}

\textsuperscript{170} Ibid. 61.
\textsuperscript{171} Ibid. 64-65.
\textsuperscript{172} Ibid. 170-175.
\textsuperscript{173} Regulatory commissions, long deferential to utilities, began to in some instances deny cost recovery and proposed projects. The Environmental Defense Foundation (EDF) won a major victory during the 1976 California Public Utilities Commission’s review of a proposed Pacific Gas & Electric (PG&E) nuclear plant by introducing alternate demand forecasts based on conservation. Ultimately, PG&E incorporated some of EDF’s suggestions and scrapped the planned project. The EDF’s victory was not an isolated event. Hirsh. \textit{Technology and Transformation}. 151-152.
\textsuperscript{174} For a complete discussion of the relationship between the emergence of environmental activism and electric power from which the above is drawn, see Hirsh. \textit{Power Loss}. 64-68.
Together—technical stasis, the oil crisis, and environmental activism—challenged the calcified structure of electric power provision and created new openings for incremental change. The grow-and-build strategy found support and generated positive feedback as long as technological improvement supported increasing returns, cheap raw materials remained available, and environmental impacts were overlooked. However, once these basic conditions shifted, change, even if only seemingly marginal, became possible. No longer was it possible, or perhaps even desirable, to continue the status quo. President Carter attempted the ambitious task of redrafting energy policy through an expansive and detailed National Energy Plan, introduced in 1977, in an effort to reduce reliance on foreign oil and mitigate harmful environmental effects.\(^\text{175}\) Carter’s sweeping plan would be divided into different bills and picked apart in contentious and fractious debate. Eventually, a watered-down version would pass into law.\(^\text{176}\) One little remarked on aspect of the larger plan which was lost amid the larger debates, Section 210 of the Public Utilities Regulatory Policies Act, however, would begin to introduce new competition into electric power and set the stage for more thoroughgoing deregulation.\(^\text{177}\)

**New Entrants: PURPA and the Rise of Independent Power Generation**

Section 210 of PURPA intended to stimulate non-traditional, nonutility, forms of power to reduce dependence on foreign oil imports and lessen environmental damage.

\(^\text{175}\) Imported petroleum accounted for an increasing share of domestic consumption. In 1973 imported petroleum accounted for 36% of domestic consumption, but by 1976, imports accounted for 42% and showed signs of increasing growth. The plan included taxes and incentives in favor of conservation, nuclear power, and coal, the granting of new federal power to force utility interconnection, and the end of the promotional block rate structure. Ibid. 73-75.

\(^\text{176}\) Ibid. 73-80.

\(^\text{177}\) Section 210, what turned out to be the key provisions of PURPA, was overlooked during the debate. Ibid. 81-88.
Section 210 created a guaranteed market for power from cogeneration and renewable sources by forcing utilities to purchase power from qualifying facilities (QFs) at the avoided cost at which the purchasing utility would have spent to generate energy. Before PURPA, regulated utilities were not only the exclusive suppliers of electricity to the public through transmission and distribution facilities, but they also were the sole market for power produced by cogeneration facilities. Taking advantage of their unique position, utilities had little reason to offer cogeneration facilities decent rates or open their network to outside suppliers. Section 210 of PURPA, sought to remove these barriers and open utilities to outside supply. At the time, incumbent producers saw the provision as trivial. Section 210 was working at margins of industry—non-utility sources amounted to less than 4% of total capacity in 1978 – but would in time prove revolutionary.

Section 210 created a new category of power producers, QFs, with a guaranteed market and little regulatory restriction. Critically, Section 210 specified that in tabulating avoided costs, incremental costs rather than the more slight marginal costs of production were to be used. In providing direction as to how to calculate avoided costs, the Federal Energy Regulatory Commission (FERC) interpreted Section 210 in the most generous manner possible for QFs. FERC specified that avoided cost would be calculated by using

---


179 They controlled a monopsony market within which they were the only buyer.

180 Hirsh. Power Loss. 81-88.

181 Ibid. 81-82.
fuel and capital costs of plants built after 1978, costs that were in excess of pre-1978 costs, in setting rates to be paid to QFs. The decision to offer a more expansive definition of avoided cost that included fuel and capital costs in tabulating the prices utilities had to pay led to generous terms for QFs.\(^{182}\) PURPA also exempted QFs from PUHCA and state laws regulating rates and did not require them to register as utilities.\(^ {183}\)

PURPA offered the first opening up of the vertical monopolies that dominated electricity provision for the better part of a century; by legislative fiat it removed barriers to entry and created a robust market for nonutility power and aided growth of independent power producers. As states implemented the law with FERC guidance, the total nonutility electricity delivered to utilities increased dramatically from 6.034 billion KWh in 1979 to 28.03 billion in 1985 and 116.538 billion in 1990, moving from 3.5% of installed capacity to 5.5% in 1990.\(^ {184}\) By 1990, 7% of all electricity generated in the U.S. came from nonutilities, up from 2.9% in 1979.\(^ {185}\) PURPA, however, was only the first step in a broader push toward the deregulation and transformation of electric power.


\(^{182}\) QFs could buy power at residential rates from utilities, which were based on average costs that were significantly less than avoided costs, while turning around and selling power back to utilities at higher rates based on avoided costs. Ibid. 86-91.

\(^{183}\) Ibid. 87


PURPA discredited the notion that electricity was in all aspects a natural monopoly, the traditional justification for regulation. Importantly, the growth of independent power producers created a new interest group—QFs—that would advocate for the relaxation of regulations and the move toward wholesale and retail electricity markets by the early 1990s. Further, the disparity between prices offered by regulated utilities and QFs tempted large customers with the promise—or at least the possibility—that deregulation would lead to lower prices and induced them to join independent producers in lobbying for deregulation. It was this ad-hoc coalition of independent power producers and large industrial consumers that would be instrumental in pushing for deregulation.

Under PURPA, independent power producers took the lead in developing and implementing new small-scale efficient gas-turbines and other innovative technologies. Cogeneration plants by the 1990s could beat utility prices by 5 to 15%. Free market economists for decades attacked economic regulation in general and the regulation of electric power in particular as wasteful. The success of PURPA served a tangible example to which they could now highlight. Unregulated power, so it seemed, could work just as well if not better than regulated utilities. The success of the QFs undermined the widely-accepted notion that electric power generation was a natural monopoly, defined by economies of scale and difficulties of attracting capital investment, and supported competitive reform as a plausible alternative.

---

While PURPA created a real-world laboratory to test the viability of competitive power generation, it also created a new motivated constituency: independent power producers. The growth of QFs during the 1980s created a politically motivated interest group that, once established, pushed to solidify and extend their position. PURPA, by introducing independent power producers into the formerly staid world of utilities, in the words of Richard Hirsh, “altered the demographics of power within the utility system.”\(^{188}\) Independent power producers lobbied hard for deregulation over the objections of entrenched members of the utility industry, which feared challenges to their position, in order to expand market access and reduce limits on operations.\(^{189}\) Under PURPA, independent power producers sold power directly to utilities through agreements undertaken on terms defined by Section 210 (and state laws implementing the section).\(^{190}\) As QFs grew in stature, they wanted broader access to the markets and, by extension, the electric power grid. They also wanted relaxation of limits on size and fuel use to which non-cogeneration QFs were subject (keeping with the spirit of PURPA to support alternative forms of energy and small-scale production).\(^{191}\) Utility control over transmission and distribution networks limited QFs’ access to larger markets. QFs could not easily sell power to end consumers or to other remote utilities so long as utilities maintained exclusive control over their transmission networks, which linked to other utilities through the electric power grid and distribution networks which connected end

\(^{188}\) Hirsh. *Power Loss.* 119.
\(^{189}\) The Edison Electric Institute (EEI), Alliance of Energy Suppliers would push for deregulation, while entrenched utilities (IOUs) formed Electric Reliability Coalition to fight opening transmission and distribution networks to competition. Beder. *Power Play.* 88; Hirsh. *Power Loss.* 245.
\(^{191}\) Ibid.
users. Size limits, more obviously, limited the scale of commercial QFs and presented a barrier to growth.\textsuperscript{192}

Large industrial consumers joined independent power producers in pursuing the introduction of wholesale and retail competition. As occurred in other industries, such as rail, the introduction of partial competition at the margins created new inducements for players to advocate further deregulation. The success of QFs under PURPA tempted large consumers of power to push for competition as a way of gaining access to cheaper supplies of power. Under the partial liberalization afforded by PURPA, QFs in some states began producing power at levels lower than the retail prices offered by utilities (which were based on averaged costs and reflected historic investments). In certain markets, a wide disparity opened between wholesale and retail prices. In areas with high retail rates, such as California and the Northeast, a gap of 3-4 cents emerged between retail and wholesale prices.\textsuperscript{193} Large consumers, represented through the Electricity Consumers Resource Council (ELCON), lobbied for wholesale and retail competition in order to capture lower prices and, in effect, opt out of the regulated electric power market.\textsuperscript{194}

The advocacy of independent power producers and ELCON found a hearing within George H. Bush’s administration. The Energy Policy Act of 1992 embodied most,


\textsuperscript{194} Hirsh. \textit{Power Loss.} 244-245; Joskow. “Restructuring, Competition and Regulatory Reform in the U.S. Electricity Sector.”; Beder. \textit{Power Play.} 87-88.
though not all, of their wishes. The Act opened the power grid, the loosely interconnected network of networks that stitched together different utilities, by mandating wholesale competition among power generators and created a new category of power producer, exempt wholesale generators (EWG).\textsuperscript{195} Though the precise rules governing such competition would be formalized by FERC, the law transformed the power grid into a marketplace. Now, utilities had to open their transmission network to independent producers, allowing generators to sell power across the network to different utilities and service companies. The Energy Policy Act in effect broke up the old utility system of vertical monopolies by opening generation fully to competition: no longer would a single utility control generation, transmission, and distribution. Through competitive wholesale wheeling, power could be sold across the grid and supplied to different end providers. The law, however, did not mandate retail competition but allowed each state to decide the retail question on a state-by-state basis.\textsuperscript{196} The creation of EWGs represented an enlargement of the QFs established under PURPA. Unlike QFs under PURPA, EWGs were not bound by efficiency requirements, nor were EWGs subject to PUHCA regulations. EWG, for example, did not have to abide PUHCA restrictions on the use of debt and could be owned by utilities outside of their franchise area.\textsuperscript{197}

In the years following the Energy Policy Act of 1992, independent power generation (QFs, EWG, and other forms of nonutility power) grew rapidly as competitive wholesale markets replaced the staid utility model. Nonutility power accounted for 7.5%  

\textsuperscript{196} Hirsh. Power Play. 241-244.  
\textsuperscript{197} Hirsh. Power Loss. 243.
of total power generated (229,676.33 GWh) in 1990, 10.70% (358,958.77 GWh) in 1995, 20.96% (786,721.67 GWh) in 2000, 38.97% (1,580,577.19 GWh) in 2005, and 39.93% (1,576,918.20 GWh) in 2009 (see Table 3-4). Additionally, a thriving wholesale market for power developed through which utilities, independent generators, and third party power marketers sold electricity. In 2001, for the first time, the resale power market—electricity resold to a distributing entity or power marketing company—outstripped total power generated: in 2001, 3,736,643.65 GWh were generated, while 7,345,319 GWh were sold on the resale market (see Table 3-5). The more controversial introduction of retail competition proceeded with caution. As of 2008, only 14 states adopted retail competition, while well-publicized scandals and failures led 8 different states to suspend retail competition.

*Electricity Deregulation: Tight Coupling, Complexity, and New Vulnerabilities*

Deregulation transformed the interconnected electric power system in ways that introduce or enlarge large-scale vulnerabilities. The embrace of markets transforms, both directly and indirectly, the use of ICTs within day-to-day operations. The push toward market liberalization is matched with higher levels of centralized control. In order to create open market access and coordination, new institutions, independent service

---


operators (ISOs) and regional transmission organization (RTOs), assume control over what were formerly multiple independent systems. The rise of centrally managed ISOs and RTOs supplants the balkanized network-of-network architecture previously supported by regulation (see Chapter 2). More subtly, the dismantling of the utility system reorders the context within which firms operate and eliminates the incentives built into rate-of-return regulation to encourage long-term investment (even overinvestment) in new, expensive, equipment, in favor of bottom-line considerations. In tandem with restructuring, electric power companies increasingly turn away from isolated and customized control systems in favor of internet-enabled commercial-off-the-shelf (COTS) control systems. Both movements create new vulnerabilities and possibilities for large-scale failure that were discouraged under regulation.

Consolidated Control: ISOs, RTOs, and Organized Regional Markets

FERC Orders 888, 889, and 2000 set out the basic terms on which competitive wholesale markets would replace the utility model and encouraged the development of centrally managed or organized markets. In these Orders, FERC attempts to eliminate the barriers that incumbent control of transmission networks pose to wholesale trading. FERC requires utilities to functionally “unbundle” their networks by filing open-access transmission tariffs and offering non-discriminatory access to transmission networks—

---


202 Under the Energy Policy Act, FERC now has the power to order utilities to open transmission facilities to independent producers (interconnection) even if such an extension requires expansion of transmission facilities. Joskow. “Deregulation and Regulatory Reform.” 129-133; Hirsh. Power Loss. 244.
that is allow interested parties access to transmission services on terms comparable to that of the operating utility—and participate in an “open access same-time information system” (OASIS) linked through the internet offering all market participants up-to-date capacity and price data.\textsuperscript{203} To satisfy these requirements, FERC Orders 888, 889, and 2000 strongly support, though do not mandate, the adoption of a regional organized market-structure overseen by ISOs and RTOs (henceforth ISO/RTOs).\textsuperscript{204} In organized markets, ISO/RTOs perform key functions: they host day-ahead and real-time spot wholesale markets; provide all participants open, non-discriminatory, access to transmission service and real-time information about grid conditions through the posting of system information and prices via OASIS; and, importantly, assume operational control of transmission networks under their jurisdiction.\textsuperscript{205} ISO/RTOs are subject to FERC regulation and operate independently from any market player.\textsuperscript{206} The ISO/RTO model has been widely adopted over the past decade and a half such that ISO/RTOs now control the bulk of the U.S. electric power system (see Figure 3-1). As of 2010, 7 ISO/RTOs currently coordinate and control electric power service for the majority of the country, covering 35 states, including all of New England, the Mid-Atlantic, the

\textsuperscript{204} “ISO” and “RTO” are used interchangeable by FERC, though RTOs historically have been multi-state in nature. ISO/RTOs oversee what are known as “organized markets” as opposed to markets lacking a central coordinator/operator which are organized through bilateral contracts. GAO. \textit{Electricity Restructuring: FERC Could Take Additional Steps to Analyze Regional Transmission Organizations’ Benefits and Performance}. GAO-08-987. Sept. 2008. Fn. 3; ISO/RTO Council. \textit{State of the Market Report: 2009}. 2009. \url{<http://www.isorto.org>}.  
\textsuperscript{206} With one exception: ERCOT only operates within the state of Texas and is overseen by the state utility commission.
Midwest, and large swaths of the Southwest and California. These 7 ISO/RTOs now control power for 194 million customers, more than 2/3 of all electricity customers in the U.S., and oversee 57% of total U.S. generation capacity.

ISO/RTOs provide the centralized control and coordination needed to ensure that the interconnected system remains stable. Under the ISO/RTO model, multiple independent local utilities voluntarily cede control of their transmission networks to the governing ISO/RTO to create a larger single control area—in effect a larger interdependent system (see Figure 3-1). ISO/RTOs assume operational authority over the accumulated transmission network and adopt many of the control functions that were previously the domain of individual utilities, most importantly the critical functions of maintaining real-time balance between supply and demand, forecasting, and dispatch. Under regulation, control was located at the level of the individual utility. Utilities relied on internal, intra-firm, control and communication networks to maintain real-time

\[\text{References}\]


210 As noted in Chapter 2, electric power, due to its unique characteristics, requires real-time control and coordination to balance demand and supply. Electricity does not move through interconnected networks in a point-to-point, switched fashion but rather moves according to Kirchhoff’s Laws and requires that total generation and consumption must be balanced. ISO/RTOs now are entrusted to provide reliability through balancing, dispatch, and real-time system monitoring. Joskow. “Deregulation and Regulatory Reform.” 116; 2010 ISO/RTO Metrics Report; ISO/RTO Council. State of the Market Report: 2009.
monitoring, balance, and dispatch over integrated networks (combining generation, transmission, and distribution). Under deregulation, however, ISO/RTOs now take over responsibility for these functions by replacing intra-firm controls with centralized, inter-firm, control located at the regional, ISO/RTO-level. Whereas under the utility model control was distributed widely and nested within in each separate utility, ISO/RTOs aggregate control over many formerly distinct and relatively independent systems. For example, the creation of ERCOT consolidated 10 separate control areas under its authority, while the Midwest ISO assumed control over 34 previously independent and separate transmission owners across the Middle West.

The consolidation and centralization of control within electric power finds close parallel in the postal network. Recall, in order to improve processing efficiency, USPS adopted new mechanization and, later, automated technologies and restructured the internal processing network around a limited number of large processing centers. The adoption of the managed ISO/RTO model evidences the same logic. In both post and electric power, regulation undergirded distributed processing throughout the network. In order to capture new efficiencies, however, loosely-coupled network structures gave way to tight coupling.

The transformation of electric power into a tightly-coupled system dominated by a few points of control aggregates risk in new ways. ISO/RTOs champion the apparent

---

212 Ibid.
improved reliability which they deliver. ISO/RTOs note that consolidating and centralizing control over what where formally multiple independent systems reduces the number of transactions and controllers involved in the provision of service. The ISO/RTO council, a collaborative comprised of all operating ISO/RTOs with the exception of ERCOT, claim that “ISOs/RTOs enhance reliability through their large geographic scope – by dispatching generation over a broad region they reduce the number of decision makers managing the grid, which simplifies coordination and improves reliability.” Yet, while consolidating control may arguably increase day-to-day reliability by simplifying control, it also increases the possibility of large-scale failure. Within the restructured network topology, ISO/RTOs stand out as dense hubs central to the operation of vast regional interdependent power systems. Breakdowns in ISO/RTO control—due to operator error, intentional attack or sabotage, technological failure, or other cause—jeopardize the large interdependent electric power networks which they manage. The balkanized structure of electric power under regulation militated against these possibilities by underwriting diverse loci of control. The enlargement of systems under single authority—that is moving from distributed to central control—erases these benefits. The failure of NYISO, CAISO, or ERCOT, unlikely though it may be, undermines power not for a single community but for virtually an entire state; while the failure of larger ISO/RTOs threatens to destabilize service across multiple states. The new interdependences created by ISO/RTOs correspondingly

increase the stakes of failure dramatically. In seeking the smooth functioning of competitive markets, the process of deregulation creates conditions sympathetic to large-scale failure.

*Deregulation, Control Systems, and Complexity*

Restructuring also more subtly helps transform how firms involved in the provision of electric power use ICTs. The injection of competition into the electric power sector supports the adoption of standardized and accessible control systems by reordering the basic set of incentives under which firms operate. Rate-of-return regulation, as noted in Chapter 2, encouraged firms to invest heavily in capital and adopt sophisticated ICTs (as well as other forms of equipment more generally) for operations management. Under competition, however, markets create incentives to maximize short-term efficiency and reduce costs. Firms facing competitive pressures are replacing customized control systems operating on isolated communication networks in favor of standardized commercial-off-the-shelf (COTS) systems operating on common operating systems, such as Windows, linked to the internet or other multi-use networks. Now, control systems are more standardized and accessible.

---

215 Functional unbundling replaced rate-of-return regulation with competitive markets for generation (while leaving rate-of-return regulation in place for transmission and distribution in those areas where retail competition was not introduced).


Supervisory Control and Data Acquisition (SCADA) systems are computer-based centralized control systems used to oversee remote processes in the provision of electric power. SCADA systems are the “digital brains” used in all phases of electric power—generation, transmission, and distribution—to monitor and control real-time distributed processes from a control center. SCADA systems collect data from remote locations and send response commands from a control center to field devices to enact physical operations for valves, breakers, capacitors, and transformers based on programmed or operator controls. Typical SCADA systems are made up of hardware and software components including a master terminal unit (MTU) and human-machine-interface (HMI) located at the control center, remote terminal units (RTUs) and/or Programmable Logic Controllers (PLCs) located in the field, and communications links connecting the control center to field devices (see Figure 3-2). Historically, utilities invested in customized SCADA equipment (hardware and software) designed for their specific operations. These systems, referred to as “legacy systems,” operated across isolated


217 SCADA systems are used in a number of different industries to provide central control and monitoring of remote processes. In electric power, SCADA systems are commonly a part of energy management systems (EMS), which additionally contain automatic generation control (AGC), dispatch, and forecasting systems. Barnes, Johnson, and Nickelson. Review of Supervisory Control and Data Acquisition Systems (SCADA). 13.

218 SCADA systems are used to monitor distributed processes, while distributed control systems (DCS) are used to monitor and control site-specific operations. DCS are also used in electric power, specifically at generation sites, but are often linked to SCADA systems to provide coordinated control. OE. National SCADA Test Bed Program. 1; GAO. Critical Infrastructure Protection: Challenges and Efforts to Secure Control Systems. 8-9; NIST. Guide to Industrial Control Systems (ICS) Security. 2-1.

communication channels and relied on specific, and obscure, protocols. Beginning in the late 1990s, however, SCADA systems began to change. New standardized COTS equipment manufactured by a handful of vendors is replacing legacy systems. New COTS equipment is increasingly designed to operate on standard multipurpose commercial Operating Systems (OS), such as Windows, and standardized protocols, while also offering connectivity to the public internet and corporate intranets.220

Restructuring indirectly aided the transition to the new control system architecture. Deregulation transformed the basic model of utility operation and, just as importantly, investment. Rate-of-return regulation rewarded long-term, capital intensive, projects. Public utility commissions for decades were deferential to utility requests to undertake new projects. Indeed, increasing capital was one way to expand the rate base and hence returns.221 The move to competitive markets, however, undercuts these supports. Under competitive markets, firms—including utilities operating competitive generation, independent generators, and third party firms—no longer are rewarded by investing in new capital intensive projects, but rather are rewarded for low costs and lean operations.222 Without promised returns and in the face of competitive pressure,
investment priorities shift toward low cost opportunities and short-run profitability. In this context, firms began to adopt cheaper COTS systems.\textsuperscript{223}

The new control system architecture offers new forms of large-scale vulnerability. Increased standardization and accessibility create conditions for cascading failure. Esoteric and isolated control systems offer defense through obscurity: targeted attacks would have a difficult time uncovering faults in the little-used systems to exploit or gaining access to systems operating on isolated channels. Legacy systems were, of course, not perfect or immune to failure.\textsuperscript{224} Yet, the lack of interconnection and standardization ensured that potential failures were small-scale and localized—in effect, the specificity of control systems limited the scale of failures. Increased standardization and accessibility offers a trade-off: it enlarges the community of users working to identify and address flaws, while increasing the possibility of large-scale failures. The flip side of making information about control systems available, and systems more accessible through internet and intranet connections, is that they now become more prone to malware—worms and virus targeting exploits in the multipurpose OS or specific application—and intentional disruption. Standardization makes information, and hence information about system weaknesses, more widely distributed and allows interested parties to create malware taking advantage of these exploits. At the same time, the adoption of common OS, leads control systems open to the raft of malware designed to


\textsuperscript{224} The limited circulation of information about customized systems limits the community of skilled technicians working to identify flaws and improve reliability.
target OS-level vulnerabilities. Within this ecology, failures are quite different. They no longer stay within a specific system, but, rather, travel across open communication networks and find shared vulnerabilities endemic to common systems.  

The risks associated with control system vulnerabilities exist not only in the abstract. In 2003, the so-called “Slammer Worm” targeting Microsoft SQL server found its way into the FirstEnergy’s Davis-Besse nuclear power plant. The worm impacted computer systems globally and spread to the Davis-Besse plant through unsecured connections between FirstEnergy’s corporate network and the internet. The worm disabled the plant’s safety monitoring system for nearly 5 hours, as workers scrambled to restore system operation. Luckily, at the time of the incident the plant was idle. Nonetheless, the case underscores the challenges of securing control systems. Here, the combination of standardized equipment and accessibility led to a potentially very serious system failure.

Conclusion: The Political Origins of Infrastructure Vulnerability

Deregulation creates new contradictions within infrastructure services. Political restructuring of price-and-entry control made infrastructure services both more efficient


and more vulnerable. Seen in this light, the presumed successes of deregulation are qualified: efficiencies and lower tariffs are achieved at the price of greater vulnerability and brittleness. Likewise, from this angle the inefficacies of the old regulatory model, much decried during their day, appear to have offered some overlooked benefits.

Regulation, for its myriad charged failings, built into infrastructures a degree of resilience. Stable monopoly or quasi-monopoly markets incidentally provided a check against systemic vulnerabilities and supported resilient configurations. Economic regulation, enacted for a host of other reasons, supported the creation and maintenance of systems characterized by loose coupling and linier organization. Despite appearing through the dimmed prism of economic efficiency to be wasteful indulgences, these qualities provided an easily overlooked benefit: resilience. Regulatory reform undermined these benefits in the search for improved operation. In many ways, then, the new or enlarged vulnerabilities sketched above represent not the failing of the project of regulatory restructuring but rather its success: it is with the successful adoption of new efficient processes that new vulnerabilities are created and enlarged.

Foregrounding the connection between political restructuring—the project of deregulation—and infrastructure vulnerability challenges certain basic assumptions about the roots of infrastructure vulnerability. Rather than seeing infrastructure vulnerability as alternately the product of years of not-so benign neglect, the autonomous logic of technological or economic progress, or an inevitability of private ownership, focusing on the transformations associated with deregulation highlights that vulnerabilities are politically and historically marked; they spring in part from critical and specific political
decisions that are linked to the upheavals in infrastructure governance that first began to stir during the 1970s.\footnote{227} Contemporary challenges or problems reflect then not inevitabilities, but instead the particular consequences of the political project of deregulation. In numerous U.S. Government publications, the fact that 85% of the nation’s critical infrastructures are privately owned is repeatedly invoked as a way of explaining vulnerability.\footnote{228} This statement, though accurate, tells only part of story. The assertion elides the important role that regulation and public policy play in shaping and determining how such property is organized.\footnote{229}

The introduced new forms of vulnerability sketched above, however, remain, for a time, largely obscured from view. Hidden in plain sight, the risks associated with centralized control, standardization, and the influx of shipments of hazardous materials across a greatly reduced network go (largely) unnoticed. For the most part, after deregulation the postal system, rail network, and electric power grid operate reliably; minor disruptions occasionally occur and often occasion little more than a raised eyebrow before regaining their footing and returning to normal operating conditions. Yet, while deregulation has not imperiled the reliability of day-to-day operation, it creates the conditions for rare, large-scale, systemic failure. In the aftermath of the terrorist attacks


\footnote{229} It is worth noting that ostensibly free markets are an active political creation, rather than a “natural” state of affairs outside of political and social life. See, Karl Polanyi. The Great Transformation. Boston: Beacon Press, 1944.
of 9/11, the contradictions implicit within infrastructure restructuring emerge from the shadows and are pushed to the fore. No longer obscured, the overlooked costs of deregulation are called to account and the contradictions implicit within restructuring become central objects of political debate and ultimately intervention. It is to this process that the second half of the dissertation now turns.
Table 3-1: Postal Operating Deficit, 1926-1969

Postal Operating Deficit, 1926-1969

Table 3-2: Percentage of USPS Mail Volume Sorted Mechanically, 1968-1980

Percentage of USPS Mail Volume Sorted Mechanically, 1968-1980

<table>
<thead>
<tr>
<th>Year</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1968</td>
<td>10%</td>
</tr>
<tr>
<td>1971</td>
<td>25%</td>
</tr>
<tr>
<td>1974</td>
<td>52%</td>
</tr>
<tr>
<td>1977</td>
<td>64%</td>
</tr>
<tr>
<td>1980</td>
<td>80%</td>
</tr>
</tbody>
</table>

Table 3-3: Average Return on Investment Per Decade, Railroad Industry, 1940s-2000s

<table>
<thead>
<tr>
<th>Decade</th>
<th>Average Return on Investment (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940</td>
<td>2.0%</td>
</tr>
<tr>
<td>1950</td>
<td>2.0%</td>
</tr>
<tr>
<td>1960</td>
<td>2.0%</td>
</tr>
<tr>
<td>1970</td>
<td>2.0%</td>
</tr>
<tr>
<td>1980</td>
<td>6.0%</td>
</tr>
<tr>
<td>1990</td>
<td>8.0%</td>
</tr>
<tr>
<td>2000</td>
<td>10.0%</td>
</tr>
</tbody>
</table>

Table 3-4: Percentage Utility and Non-Utility U.S. Electricity Generation, 1990-2009

<table>
<thead>
<tr>
<th>Year</th>
<th>Utility</th>
<th>Non-Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>90.00%</td>
<td>10.00%</td>
</tr>
<tr>
<td>1991</td>
<td>85.00%</td>
<td>15.00%</td>
</tr>
<tr>
<td>1992</td>
<td>80.00%</td>
<td>20.00%</td>
</tr>
<tr>
<td>1993</td>
<td>75.00%</td>
<td>25.00%</td>
</tr>
<tr>
<td>1994</td>
<td>70.00%</td>
<td>30.00%</td>
</tr>
<tr>
<td>1995</td>
<td>65.00%</td>
<td>35.00%</td>
</tr>
<tr>
<td>1996</td>
<td>60.00%</td>
<td>40.00%</td>
</tr>
<tr>
<td>1997</td>
<td>55.00%</td>
<td>45.00%</td>
</tr>
<tr>
<td>1998</td>
<td>50.00%</td>
<td>50.00%</td>
</tr>
<tr>
<td>1999</td>
<td>45.00%</td>
<td>55.00%</td>
</tr>
<tr>
<td>2000</td>
<td>40.00%</td>
<td>60.00%</td>
</tr>
<tr>
<td>2001</td>
<td>35.00%</td>
<td>65.00%</td>
</tr>
<tr>
<td>2002</td>
<td>30.00%</td>
<td>70.00%</td>
</tr>
<tr>
<td>2003</td>
<td>25.00%</td>
<td>75.00%</td>
</tr>
<tr>
<td>2004</td>
<td>20.00%</td>
<td>80.00%</td>
</tr>
<tr>
<td>2005</td>
<td>15.00%</td>
<td>85.00%</td>
</tr>
<tr>
<td>2006</td>
<td>10.00%</td>
<td>90.00%</td>
</tr>
<tr>
<td>2007</td>
<td>5.00%</td>
<td>95.00%</td>
</tr>
<tr>
<td>2008</td>
<td>0.00%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

Table 3-5: Total U.S. Electricity Generation and Resale, 1998-2009

Total U.S. Electricity Generation and Resale, 1998-2009

Figure 3-1: ISO and RTO North American Operating Map\textsuperscript{235}

The supervisory control and monitoring station typically contains redundant application servers, an engineering workstation, and a human-machine interface (HMI) that collects and logs information obtained from the remote/local stations and sends commands to the remote/local stations in response to events detected by the sensors. The HMI displays status information, including alarm readings or operator attention.

A remote/local station contains a remote terminal unit (RTU), programmable logic controller (PLC), or other controller that receives and interprets signals from the sensors and generates corresponding control signals that it transmits to the control equipment.

Communications network can be the Internet, a public switched telephone network, a wireless network, or a radio network.

Outside vendor (may include vendors, customers, and other business partners)

Figure 3-2: Typical Elements of a Control System

---

236 Figure is taken from, GAO. Critical Infrastructure Protection: Challenges and Efforts to Secure Control Systems. Figure 3: Typical Components of a Control System. 10.
Infrastructures are, for the most part, boring; by their very nature, they support other forms of activity and do not call attention to their organization. Well-functioning systems of communication, transportation, and power are taken for granted and hidden in plain sight as mundane features of daily life. Occasionally, however, this typical situation is inverted: the dull and boring workings of infrastructure appear interesting, pressing, and of a more general concern. Failures, outages or feared outages, illuminate the once hidden and politicize what were previously considered to be “technical,” neutral and expert, domains.

In the aftermath of the terrorist attacks of September 11th, 2001, the threat of terrorism operated in just this manner. The threat of terrorism cast public attention back on the operation of infrastructures and, in the process, elevated the “sunk” politics of infrastructures above ground and created an opening for substantive debate about their operation. Now, after 9/11, the contradictions of deregulation became political problems. Counter-intuitively, the threat of terrorism post-9/11 served as a resource in support of the democratization of domestic infrastructure governance. Rather than axiomatically leading to the diminution or suspension of democratic processes, the perceived risk of terrorism offered the possibility of opening infrastructures to debate and discussion by wide-ranging actors. Groups outside of traditional discussions of national security—including labor, environmentalists, and local community groups—appropriated the risk of terrorism and became central players in determining the conditions upon which

---

infrastructures would operate. The idea of terrorism provided a range of groups with a resource to legitimate and advance their claims about the ordering of infrastructure systems. Here, risk serves as a type of political capital to be appropriated by diverse actors. How this capital is used, how salient but difficult-to-define risks are translated into novel forms of security, is ultimately an empirical question to be considered on a case-by-case basis. In working through the political dimensions of risks, it is imperative to be alert to how, by whom, and to what ends new risks are interpreted and invoked.

In this short chapter I examine the theoretical foundations for considering novel forms of risk as a political resource. An exegesis of the work of Ulrich Beck and his concepts of risk society and reflexive modernization provides the grounding for the empirical examinations of post-9/11 debates about infrastructure security within the postal system, freight rail transportation, and electric power which follow in Chapters 5-7. In surveying Beck’s work I foreground the important role afforded communication and culture, while also recognizing limitations within his framework. For Beck, the political volatility of certain categories of risks—what makes them potentially transformative political resources—is inseparable to their positioning as cultural artifacts. In Beck’s work risks can, in some cases, serve as resources which support democratic engagement. Yet, as will be discussed in detail below, the celebration of the political potential of risks must be qualified and tempered. Beck’s conceptualization of democracy rests on a single-minded embrace of participation as the centerpiece of democratic practice and, as a result leads to a broad attack on all forms of hierarchy, particularly expertise, as anti-democratic. Yet, Beck ignores the vital role that experts
play in a democracy and underplays the real dangers that participation can present to the values of democracy; hierarchies in many instances are central to the preservation of democracy. As noted above, the politics of risk are an empirical question that calls for the close examination of concrete cases.

Risk Society and Reflexive Modernization: New Political Possibilities

Beck’s central thesis posits that society is entering a period of late-modernity characterized by the centrality of new forms of risk. For Beck, the process of modernization increasingly produces new forms of risk that sit at the center of political debate. While traditional political conflicts within modernity centered, to a large degree, on the distribution of “goods,” such as wealth and jobs, Beck now sees the distribution of “bads”—forms of harm and danger—as central. As he writes:

[w]ith the advent of risk society, the distributional conflicts over ‘goods’ (income, jobs, social security), which constituted the basic conflict of classical industrial society and led to attempted solutions in the relevant institutions, are covered over by the distributional conflicts over ‘bads’ … they erupt over how the risks accompanying goods production … can be distributed, controlled, and legitimized.²

Beck argues that as part of this turn, society can now be conceived as “risk society.” The production of new forms of risks, defined by Beck in keeping with the term’s colloquial meaning as forms of harm and danger, are internal to the process of modernization; they are what Anthony Giddens refers to as the “dark side of modernization.”³ The production of risks within risk society is not a failure of modernization, but rather a side-effect of the

processes of industrialization and capitalism. As Beck notes, “[i]n advanced modernity the social production of wealth is systematically accompanied by the social production of risks.” With the ascendance of risk society, society confronts and is preoccupied with the management and control of self-generated risks; society increasingly faces the contradictions of modernization.\(^4\)

Beck’s thesis emerges from his reading of the ecological movement in Germany during the 1970s and 1980s and is generalized into flexible theoretical treatment of contemporary risk.\(^5\) Beck argues that though forms of harm have, of course, always existed, the process of modernization produces novel risks.\(^6\) Now, modernization is producing risks that are expansive—unmoored in time and space—global in scope, potentially catastrophic, and resistant to calculation. These new forms of risk cannot be reduced to a specific location or fixed temporal period and challenge the authority of states operating over a limited territory and narrow time-horizon. The risks associated with nuclear waste and climate change, for example, cannot be reduced to a specific state and extend far into the future. Precisely whose problem such risks are is open to question.\(^8\) Nor can the risks produced by modernization be reduced to easy estimates of frequency or degree of danger. Indeed, Beck suggests that contemporary risks outrun the


\(^{7}\) Ibid. 20-21.

ability to tabulate either their likely occurrence or severity with great confidence. These risks are atypical, infrequently occurring, and quite new, making historical extrapolation difficult. Additionally, certain risks driven by advances in techno-science, such as those linked with new biotechnologies, challenge the ability of scientists to unanimously and precisely describe their effects. In this manner, “risks,” as Beck sketches the term, appear to be what are often described within the literature on risk as “uncertainties.”

These forms of risk have the potential to create institutional turmoil and disruption. Risks that spill over boundaries and cannot be localized within specific jurisdictions undermine the territorial foundations of national structures of authority and call for the development of new institutions. Nor can uncontested statements about the danger they present be offered. Calculability is central to the management of risks by supporting empirical testing and the provision of mitigation measures, including insurance. Absent accepted measures of probability, central institutions involved in managing risks—namely branches of science, political bodies, and insurance—are destabilized. In this manner, risks such as those associated with ecological collapse, nuclear waste, and terrorism, to select three risks Beck regularly invokes, undermine and

---

10 The classic distinction between “risk,” for which probabilistic estimates can be made, and “uncertainty,” for which such estimates cannot, is found in Frank Knight. Risk, Uncertainty, and Profit. 1921. New York: Kelley & Millman, 1957. Beck’s notion of risks is more closely aligned with what Knight terms uncertainty. Indeed, Beck adopts the phrase “manufactured uncertainties” in discussing the types of novel risks with which he is interested. Beck. World at Risk. 50. For a discussion of the disjunction between Beck’s use of the term risk and its usage within economics and the literature on risks analysis, see Lupton. Risk. 7-8.
challenge the ability of existing institutions to control the side-effects of modernization and present a crisis of control.\textsuperscript{12}

Beck argues that the emergence of risk society fosters a process of reflexive modernization. Beck employs “reflexive” in a double sense, aiming to capture the meaning of both “reflex” and “reflect.” The process of reflexive modernization comprises two steps or phases. First, modernization systematically produces risks, but they do not become political or public problems. Here, risks are produced, but remain unrecognized. Second, the risks of modernization emerge as public problems that dominate debates and conflicts.\textsuperscript{13} As Beck writes, during this second moment, “the institutions of industrial society become the producers … of threats they cannot control. What happens here is that certain features of industrial society become socially and politically problematic.”\textsuperscript{14} Within this moment, society wrestles with the contradictions of modernization and in some instances becomes self-critical. It is in this manner that Beck speaks of reflexive modernization as the “modernization of modern society” and a process through which “modernization … is becoming its own theme.”\textsuperscript{15} During the first phase, modernization creates contradictions and self-generated risks. Here, Beck underscores the automatic or reflex-like quality of modernization to generate risks as a


\textsuperscript{14} Ibid. Italics in original.

side effect. During the second phase, risks become political problems and potentially lead to reflection: a turning back to consider the origins and production of novel risks. Within this phase, self-critique is made possible.\textsuperscript{16} The institutions of modernization become subject to debate, scrutiny, and quite possibly remaking. It is in this second move that Beck locates the democratic possibilities of risk.

\textit{Risks and Democratization}

Beck suggests that new forms of risk hold out the possibility for democratization.\textsuperscript{17} The recognition of the risks of modernization draw attention to the political implications of seemingly private and apolitical areas and, in the process, can create openings for new political engagements at odds with the status quo. The democratizing potential of risks operates across two distinct, though related, trajectories. On the one hand, risks once made visible can open previously closed domains considered beyond the reach of political engagement; they can illuminate the political choices and implications buried within the seemingly apolitical fabric of science and technology.\textsuperscript{18} At the same time, risks also function as a resource for otherwise marginalized groups to enter into political discussions.

Beck forwards a view of politics as usual that is dominated by increasing specialization at the expense of substantive political participation. As Beck argues, politics is increasingly colonized by hierarchies—capital, political elites, scientific experts—that crowd-out participation. Here, democracy is impoverished specifically

\textsuperscript{18} Beck. \textit{Risk Society}. 77.
because of the ongoing ascendance of hierarchies. Beck suggests, both implicitly and explicitly, that hierarchies stand in contrast to the ethos of a vibrant democracy.\textsuperscript{19} In doing so, Beck seizes upon particular aspects of democratic life—equality and participation—and locates them as central to the practice of democracy. Risk within this understanding shoulders an important significance. It not only politicizes areas of society that have been placed beyond the scope of political debate and, in the process reveals the politics implicit within such domains; but also provides a resource for those effectively barred from speaking on a certain topic to enter political discussions.

As Beck notes, the political dynamic of risks offers the possibility that for a moment “the unthinkable and unmakeable become possibilities.”\textsuperscript{20} In disrupting the normal course of politics, salient risks operate in an agenda-setting fashion and create what are known within political science as policy windows: moments ripe for new political action.\textsuperscript{21} Areas typically considered private and apolitical, such as business, economics, and science, are now politicized and subject to public scrutiny and discussion; they become legitimate topics of political discussion and intervention. As Beck notes, the recognition of new forms or risk can lead:

  previously depoliticised fields of decision-making [to become]
  politicized… Generally involuntarily and against the resistance of …
  powerful institutions… they are opening up these problems to public
doubt and debates. In global risk society, therefore, subjects and themes
once treated behind closed doors, such as public investment decisions, the
chemical composition of products and medicines, scientific research

\begin{footnotes}
\footnotetext[19]{See, Beck. Risk Society. 19-50; Beck. World at Risk; Beck. “Reinvention of Politics.” 29.}
\footnotetext[20]{Beck. Risk Society. 77.}
\end{footnotes}
programmes and the development of new technologies, are articulated and debated in public.\textsuperscript{22}

The political implications of these seemingly private domains come into focus as new forms of risk are recognized and acknowledged. The recognition of the production of risks within areas considered private and “off-limits” illuminates the inherent politics at play. In doing so, it offers the possibility of opening these areas for participatory politics and some form of collective control.

In tandem with opening apolitical spheres, the recognition of novel risks enables new actors to enter the fray. The uncertainty surrounding the probability and effects of novel risks creates a space where multiple subjects can offer competing claims. New risks allow new types of claims to be offered and taken seriously, as well as new sets of speakers to find a voice. New risks expand the horizon of possibilities and enable “as if” thinking. Once risks are publicly accepted and salient, a range of imagined possible scenarios are now credible; as Beck notes, the subjunctive replaces the indicative mood.\textsuperscript{23}

In other words, novel risks enable a new type of thinking structured around the hypothetical or possible. If the past no longer is taken to offer a clear image of the future—if the future is no longer presumed to mimic the past—the limits of what is considered plausible is stretched. Under conditions of uncertainty, then, varied claims and scenarios of what \textit{could} happen are taken as legitimate. In this moment, new and existing social movements can adopt risks as a clarion call that challenges the accepted wisdom of the status quo and, importantly, challenge power.\textsuperscript{24} In a cumbersome phrase,

\begin{itemize}
  \item \textsuperscript{22} Beck “Global Risk Politics.” 32.
  \item \textsuperscript{23} Beck, Giddens, and Lash. \textit{Reflexive Modernization}. vii; Beck. \textit{World at Risk}. 30, 68-69, 135.
  \item \textsuperscript{24} For a discussion of how social movements can appropriate risk, see Kim Fortun. \textit{Advocacy After Bhopal: Environmentalism, Disaster, New Global Orders}. Chicago: University of Chicago Press, 2001.
\end{itemize}
Beck refers to the “side-effects of side-effects” to underscore how the production of new types of risks that resist easy definition create a crisis of legitimization that underwrites the possible emergence of a new politics. The production of new risks challenges the authority and legitimacy of recognized experts and authorities. Under condition of uncertainty, Beck argues, “the boundaries between scientific and unscientific, between science and politics, and between experts and layman” erode. For Beck, the presence of uncertainty weakens monopolies of knowledge and renders expert claims contestable. As a result, a range of voices, including lay actors and other speakers not typically authorized to speak with authority are enabled to enter into substantive debates. Different groups adopt risks as strategies of legitimation. What can be called “risk entrepreneurs,” concerned individuals or groups, can appropriate the anxiety and fear associated with new risks to advance new political positions and demand access to decision making. Within the vacuum created by these types of risks, then, groups otherwise excluded from the machinations of power can point toward the uncertain and salient risk as a way of legitimating their claims. In this respect, risks serve as a type of capital that is not reducible to existing political, economic, or scientific power.

*Risk, Culture, and Politics*

Beck’s thesis concerning the political possibilities of risks is explicitly cultural. For Beck, risks are not inherently transformative, but rather they offer the potential for

---

29 Ibid. 47-48.
political transformation. Risks serve as a resource open to appropriation; they provide a source of legitimation and external support to which actors of different stripes can highlight and rely in advancing their positions. In Beck’s reading risks are unreliable allies to power; the definition and interpretation of risks does not necessarily follow established currents of power. Rather, risks are open to negotiation and appropriation from diverse sources. What direction transformations linked to the recognition of novel risks will take—in support democratization or the retrenchment of power—is an open, empirical, question. The political potential of risks is simply that, potential. It requires individuals and groups to advocate for change in reference to risks. The potential of risks to serve as a means of democratization lies, for Beck, in the gap between what he, in homage to Marx, terms “relations of definition” and other forms of power. It is in this disjunction, between who defines risks and existing forms of power, that Beck locates the radical potential of risks. In these ways, once recognized novel risks are, in the words of Beck, “politically explosive.”

Beck’s interpretation of risks sits between the poles of strict realism and constructivist accounts. Beck argues that risks have a “both/and” quality that encompasses both material and immaterial (symbolic) dimensions. Beck does not see risks as simply discursive productions, but instead argues that they have a material existence independent of their articulation. As Beck is at pains to point out, modernization is creating new forms of peril, from nuclear hazard, ecological destruction,
new technologies, and even terrorism, which have real consequences and stakes. In this
regard, Beck departs from the work of Mary Douglas and Aaron Wildavsky, which
appears to argue that risks are only products of cultural articulation. In claiming that
risks do have a material reality, even if it is one that is difficult to precisely measure with
certainty, Beck avoids the jarring abdication of responsibility which can be read in
Douglas and Wildavsky’s work.34

Yet, critically, Beck notes that risks are, always, open to interpretation and
mediation. That is, what risks mean, how they are thematized and translated into new
policies and procedures, is an open question. Risks do not in and of themselves compel
action in particular directions. Risks are ambivalent or neutral; they can be used as
symbolic resources for quite different ends.35 As Beck notes: “One can make two
diametrically opposed kinds of assertion… risks inspire paralyzing terror; or…risks
create new room for action. Neither of these statements is true and neither is false.”36

The invocation of novel risks can support a reflexive democratic politics which considers
the sources of new forms of danger and new courses of action, or, alternately, a type of
totalitarianism which impoverishes participation in the name of emergency measures.37

Risks can spur new changes and engagement or paralysis in the face of seemingly

34 In their work, for example, by virtue of claiming risks as exclusively social productions, any sense of
responsibility for ecological destruction is nullified and, in its place, the ecological movement appears to be
engaging in reckless attacks on industry. For a summary of critiques of Douglas and Wildavsky, see
Lupton. Risk. 57. See, Mary Douglas and Aaron Wildavsky. Risk and Culture: An Essay on the Selection of
Technological and Environmental Dangers. Berkeley, CA: University of California Press, 1982; Aaron
What Does This Mean?” Reflexive Modernization: Politics, Tradition and Aesthetics in the Modern Social
World at Risk. 57-59.
36 Beck. World at Risk. 57.
37 Beck. Risk Society. 78-81.
impossible challenges. The key questions for Beck are: how, by whom, and to what end are risks interpreted and marshaled as a resource?³⁸

In eschewing the rigidities of either dogmatic realism or constructivism, Beck suggests that risks exist as both material and immaterial productions. Yet, in considering the political potential of risks to forge new political engagements, Beck highlights the cultural dimensions of risks. In doing so, he notes a clear link between structure and agency. For Beck, to understand the ultimate significance of risks, how they spark and undergird action, requires empirical investigation of how groups mobilize around and deploy the symbolic currency or capital embodied within new risks. Here, Beck allows for risk entrepreneurs who invoke risks as a way of justifying action. Beck emphasizes that once risks become relevant and salient—accepted as legitimate points of reference within the stock of common images, meanings, and arguments that define shared culture—it is at this point that they offer a resource in support of agency.³⁹ Once certain risks are accepted as legitimate within the larger framework of shared and competing meanings that define a common culture, they can now be appropriated by diverse sets of actors to justify action. In this manner, Beck is making a clear connection between culture as a shared set of common, but by no means static or without contradiction, meanings, and culture as a repertoire of available tools. In Beck’s reading culture is shared, interpersonal, and collective; it is a broader structure within which action makes sense and is intelligible. In this fashion, Beck forwards a notion of culture that is

³⁸ Beck. World at Risk. 32.
³⁹ Beck’s discussion of the staging of risks in World at Risk highlights that collective knowledge and recognition of a risk is indispensable to its utility as a resource. Beck. World at Risk. 10-16. See also, Beck. Risk Society. 33-34.
compatible with Clifford Geertz’s interpretive notion of culture as a shared semiotic framework or context within which action can be described.\textsuperscript{40} Yet, in identifying how cultural notions of risk can be instrumentalized as resources of legitimation, Beck shares an affinity with Ann Swidler (and others) in emphasizing how the broader set of available meanings within a culture enable and support autonomy. In her well-worn phrase, culture acts as a “tool kit,” a repertoire of available ideas from which individuals can draw to support actions. As Beck argues, risks empower marginalized groups because they are “new sources of legitimation.”\textsuperscript{41} Risks “provid[e] political movements with the external conditions, strategies and resources for a type of ‘judo’ politics. That is to say…[risks] unleash powerful impulses for change from positions of relative powerlessness.”\textsuperscript{42} Here, culture does not compel action, it is not deterministic or oppressive in a functionalist mode, but on the contrary it enables autonomy. It is through the availability of cultural resources that “strategies of action” can be forged.\textsuperscript{43}

Securitization: Risks as Communicative Resource

The concept of “securitization,” developed most centrally within security studies by Barry Buzan, Ole Waever, and Jaap de Wilde, helps flesh out Beck’s theoretical work and concretize how, in practice, claims concerning risk function as communicative resources.\textsuperscript{44} Securitization is a particular approach within security studies that suggests

\textsuperscript{40} Geertz’s notion of culture is informed most significantly by the work of Max Weber. See, Clifford Geertz. “Thick Description: Toward an Interpretive Theory of Culture.” The Interpretation of Cultures: Selected Essays. New York: Basic, 1973. 3-30.
\textsuperscript{41} Beck. World at Risk. 66.
\textsuperscript{42} Beck. “Global Risk Politics.” 30.
that security is, primarily, a rhetorical device used to justify exceptionality.\textsuperscript{45} This school of thought positions security as, always, an inter-subjective achievement between actors advocating the necessity of exceptional measures and those in a position to accept the legitimacy of such a position. Security, under this rubric, is a speech act of a very particular type; it posits an uncommon threat or problem, the need for the suspension of the normal procedures of politics, and the implementation of some type of emergency measure.\textsuperscript{46} As Buzan and colleagues note: “Security’ is the move that takes politics beyond the established rules of the game and frames the issue either as a special kind of politics or as above politics.”\textsuperscript{47} Securitization, as a framework, identifies the mechanics through which the rhetoric of security operates.

Typically, invocations of security—the process of securitization—are taken to occur under the guise of expanding state power to mobilize against an identified threat.\textsuperscript{48} In this regard, securitization is placed in difference to normal democratic practices.\textsuperscript{49}

Indeed, securitization, as an analytic framework is largely structured around scrutinizing how claims to depart from such processes are enacted. As Buzan and co-authors remark: securitization … means to present an issue as urgent and existential, as so important that it should not be exposed to the normal haggling of politics but should be dealt with decisively by top leadership prior to other issues … it works to silence opposition and has given power holders many opportunities to exploit threats for domestic purposes, to claim the right to handle something with less democratic control and constraint.\textsuperscript{50}

\textsuperscript{45} Ibid.
\textsuperscript{46} Ibid. 26.
\textsuperscript{47} Ibid. 23.
\textsuperscript{48} Ibid. 21.
\textsuperscript{50} Buzan, Weaver, and De Wilde. Security. 29.
This appears to align with Beck’s cautionary note, and perhaps the more conventional view, that large-scale threats are marshaled to reinforce existing blocs of power (e.g., the state) at the detriment of genuine collective forms of decision making (see below). Yet, securitization as a framework defines and describes a general situation—a style of argument and speech—rather than a specific iteration. The general framework can be severed from the standard context from which it appears and still supply explanatory value. In some cases, securitization can work not to silence but amplify voices. In a type of inversion, the exploitation of threats can be put into service of democratic participation.

The joining of Beck with the securitization framework provides an important conceptual apparatus for considering instances within which invocations of security appear to enhance key aspects of democratic participation. Once we take securitization to be a value-free analytic framework, it can serve as a way of helping to concretize some of Beck’s observations. Beck is ultimately describing processes of securitization—the ways in which salient risks are adopted by actors marginalized from centers of power to advance a change in the status quo through the opening of closed domains to political deliberation—with the caveat that he begins from a different starting point. While Buzan and colleagues consider the state of normalcy to be defined by the smooth functioning of democratic processes, Beck starts from a vantage point that suggests that democratic processes are effectively stalled and inoperative. In such cases, the availability of risks offer diverse actors to legitimate security claims—to give them inter-subjective significance—in order to suspend the normal state of affairs and reinvigorate democratic
participation. Securitization helps illustrate the process through which new publics adopt formations of risk to insert themselves into relevant decision-making contexts.

Qualification and Critique

Beck’s basic thesis concerning the political potential of risks is open to challenge and qualification. Typically, risks are understood and presumed to support the suspension of democratic practices; in the face of a new threat, democracy is often said to be luxury that no longer can be indulged. Indeed, the notion that risks can serve as the foundation for democratic revitalization and engagement is critiqued at length by Corey Robin in *Fear: History of a Political Idea*. For Robin, risks are the exclusive province of the powerful; anxiety, fear, and uncertainty only support greater concentrations of power and are put in service to reduce public participation and freedom. In Robin’s reading, risks obscure the political character of situations and justify the suspension of the normal workings of democratic processes in response to the needs of a presumed emergency.

Robin’s argument is not entirely unique—John Mueller in *Overblown* offers much of the same argument in reference to post-9/11 politics—and certainly many examples of how risks have been used to justify exceptionality at the expense of democracy can be offered. Robin’s basic scenario is not precluded by Beck’s reasoning. As Beck admits

---

52 Robin’s discussion and use of “fear” is analogous to Beck’s more colloquial use of risks to mean perception of harm or danger.
53 John Mueller. *Overblown: How Politicians and the Terrorism Industry Inflate National Security Threats, and Why We Believe Them*. New York: Free Press, 2006. The argument that risks are used to justify the extension of power is also evident in the work of Robert Higgs and Naomi Klein. Higgs sees the appropriation of risks as driving the growth of the powerful administrative state, while Klein sees the same process as fostering the evisceration of public controls designed to protect public welfare in favor of powerful corporate interests. See, Robert Higgs. *Crisis and Leviathan: Critical Episodes in the Growth of*
in some cases risks can, and indeed are, deployed in precisely this manner. Yet, what Beck argues, and what Robin does not allow, is that risks are malleable and open. Ultimately, the difference between Robin (and others) and Beck splits on different understandings of culture. Here, culture is not just a scrim through which established power shines. For Beck, risks are cultural resources that can be appropriated and used in different ways. Robin, however, and the broader interpretation that views risks as anti-democratic, admits no such possibility. In his understanding, culture only operates as a means of flattering established power. Beck offers a more nuanced discussion of risks than Robin on this account. By ignoring the ways in which culture can stand in opposition to power, Robin presumes too much.

A more important critique and qualification of Beck’s work, however, can be offered based on his rendering and understanding of democracy. Beck advances two related notions within his work that require further scrutiny: first, that hierarchies stand in opposition to democracy; second, that participation defines democracy. In both cases, Beck’s work is incomplete. As noted, for Beck hierarchies thwart democracy. This is taken as a broad attack on many different iterations of power including, though not limited to, scientific expertise and political elites. Beck yokes together too much and overlooks the important role that hierarchies can, and do, play in sustaining a democracy. His general dismissal of expertise as anti-democratic finds comfort in a long-tradition that looks askance at the place of expertise within democracy. Critiques of expertise note that the presence of experts challenges equality and drowns out the voices of others and, in

some instances, the purported neutrality of experts mask particular interests. Yet, it is sobering to remember the ways in which experts and expertise, when properly executed, are important to democracy. Experts can, and often do, provide a counter balance to the arbitrary wielding of power by offering a record of truth, no matter how uncomfortable, which is occasionally at odds with parochial interests. Expertise is not always a cloak for other powerful interests, but in many instances offers a rejoinder that undermines the improper wielding of power. Experts, too, aid the democratic process by helping to clarify and set the basic grounds upon which sustained discussion can take root. Experts can help prevent cacophony by offering basic shared assumptions upon which discussion can build. In a rush to praise lay participation, it is worth noting that in a representative democracy, experts are more than a necessary evil: they are vital. Beck’s notion of democracy elevates participation as the key democratic value. Yet, in the face of complexity, some manner of expertise and delegation is needed to provide decision makers and the public with information about which to deliberate; when operating well, they provide the neutral set of background information which provides the starting point of discussion. There is a real tension between expertise and democracy, but to charge, a Beck does, that expertise is always anti-democratic presents a too simplistic of view of both expertise and democracy.

Bearing these critiques in mind, Beck’s theoretical evaluation of the relationship between risks and democratic politics is still valuable. Beck highlights the relationship

---

55 Ibid. 499-502.
56 Ibid.
between risks and new forms of political engagements. He illustrates the ways in which risks can serve as a resource to enable participation from a variety of different sources, including those otherwise marginalized. Importantly, Beck demands empirical study to substantiate the importance of risks. Risks are not autonomous, nor are they necessarily the province of either those inside or outside power. In Beck’s estimation, risks represent opportunities, not conclusions. It is the role of the scholar to trace how, by whom, and to what end novel risks are interpreted. It is imperative that examinations remain sensitive the process through which risks are translated into new policies and technologies of control; or what I term the bureaucratization of risks.

*Revisiting Deregulation: Reflexive Modernization and the New Terrain of Post-9/11 Politics*

Beck’s insights and theoretical framework offer a useful point of entry for considering post-9/11 infrastructure politics. After the attacks of September 11, infrastructures were recast as sites of vulnerability, danger, and fear. Post-9/11 new worst-case scenarios concerning attacks targeting domestic infrastructure networks circulated widely in the press, academic circles, industry, and different levels of the government. The attacks of 9/11 and the subsequent distribution of anthrax through the U.S. mail vividly illustrated how infrastructure networks could serve as a means of causing great harm, panic, and economic damage.57 The difficult to calculate risk of terrorism became an acute problem; the possibility, rather than the probability, of

---

57 The anthrax attacks were, at the time, seen as a terrorist attack. Sheryl Gay Stolberb and Judith Miller. “A Nation Challenged: The Response; After a Week of Reassurances, Ridge’s Anthrax Message is Grim.” *New York Times.* Oct. 26, 2001.
catastrophic loss became a source of concern for the state, operators, and a host of groups affiliated with or interested in the operation of these networks.

No longer taken for granted, infrastructures became visible and open to renegotiation. In this moment, the new vulnerabilities introduced or enlarged through deregulation become objects of political debate and, ultimately, intervention. In Beck’s terms, the threat of terrorism engendered a moment of reflexive modernization. Post-9/11 politics grapples with how to reconcile the reigning ethos of neo-liberalism with the perceived imperatives of security. Specifically, the tension between the call for limited state intervention and engagement in the economy and the push for assurances of security and safety came to the fore. As was discussed in the previous two chapters, the decades-long project of political restructuring to reduce key forms of state-supervision and control helped create infrastructures which were both efficient and prone to large-scale, albeit rare, failures. Post-9/11, these contradictions emerge as central political problems. Perhaps deregulation was not such a wonderful bargain after all?

Post-9/11 Security Discourse: Infrastructures as Targets

In the aftermath of the attacks of September 11th, 2001, to some, the sinister appeared to be lurking in the familiar. Water purification plants, power systems, ports, subways, and bridges—previously taken-for-granted systems tucked into the background of daily life—were now reprioritized as sites of critical vulnerability and possible terrorist attack. The seemingly endless string of “soft targets” provide enterprising terrorists with an unyielding supply of potential targets and, correspondingly, public officials with a daunting, near impossible, challenge. Policymakers quickly noted that the nation’s
infrastructure systems—interconnected systems of communication, transportation, and power (among others)—offered malicious actors a prime target for causing destruction and engaging in asymmetric warfare. These systems provide the foundation for economic, social, and political life. Additionally, the military relies on the availability and reliability of these networks for its activities. Disruption of these networks could result in grave damage to life and property, impair economic activity, and disrupt national security operations. In 2003, Homeland Security Presidential Directive-7 (HSPD-7) succinctly stated the case:

Terrorists seek to destroy, incapacitate, or exploit critical infrastructure and key resources across the United States to threaten national security, cause mass casualties, weaken our economy, and damage public morale and confidence…Critical infrastructure and key resources provide the essential services that underpin American society. The Nation possesses numerous key resources, whose exploitation or destruction by terrorists could cause catastrophic health effects or mass casualties comparable to those from the use of a weapon of mass destruction, or could profoundly affect our national prestige and morale. In addition, there is critical infrastructure so vital that its incapacitation, exploitation, or destruction, through terrorist attack, could have a debilitating effect on security and economic well-being.58

Government at the federal, state, and local level initiated new policies and procedures designed to protect key infrastructure systems against the perceived threat of ongoing terrorist action. In addition to a set of Presidential Executive Orders issued in the immediate aftermath of 9/11, both the PATRIOT Act of 2001 and the Homeland Security Act of 2002 focused on the protection of critical infrastructures as a key challenge in the domestic front of the war on terror.59 Together these Acts elevate the protection of

infrastructure networks to a national priority, call for increased cooperation and information sharing within and between the public and private sectors, and position the newly created Department of Homeland Security as the lead agency in devising a strategy for creating secure, resilient, and robust infrastructures. At a more general level, the Acts illustrate one of the turns of post-9/11 national security discourse: infrastructures, now, are viewed as targets in waiting.

Concerns over terrorism and the security of the nation’s infrastructure are, of course, not entirely new. In policy circles, the threat of terrorism was oft considered. Likewise, concern over the security of infrastructure networks was a staple of cold war defense planning. Yet, 9/11 did represent something new. In its wake, the threat of terrorism now became much harder to ignore or treat as a distant priority. An act of terrorism is an exceedingly rare occurrence; as is often pointed out, the average American has a far greater likelihood of perishing in a car accident than in a terrorist attack. But, this comparison, in large part, is beside the point. After 9/11 the threat of terrorism, slight though it may remain, became highly salient. As Beck notes, the subjunctive replaces the indicative mood. It became much easier to imagine future terrorist attacks and to credibly speak of their possibility. The terrorist attacks of 9/11 were important not

---


only for the devastation that they wrought but also because they helped forge new expectations. As Beck writes, the attacks “brought forth something spatially and temporally removed from them, namely, the expectation of terrorism. They created the taken-for-granted belief that, however improbable it may seem, such a thing is really possible…[and] can reoccur anywhere at any time.”

In a similar vein, Cass Sunstein writing in *Worst-Case Scenarios* highlights this very point: the attacks worked in a type of agenda-setting function whereby worst-case scenarios that previously could be viewed as remote fantasies—safely ignored virtual impossibilities—now moved into the horizon of possibility and legitimate concern.

Likewise, Langdon Winner notes, the terrorist attacks punctured the implicit trust that governed the operation of complex socio-technical systems and everyday life. Now, quickly, startlingly, the vulnerability of life became hard to elide. In this manner, the terrorist attacks presented a moment of cultural fissure. Terrorism became a priority and a lens through which to view; terrorism became a taken-for-granted possibility, despite its difficult or perhaps impossible to calculate probability, which can be invoked to legitimate new sets of activities. That is, it became a resource open to appropriation.

*The Bureaucratization of Risk*

In the next three chapters I turn to investigate the bureaucratization of risk. I trace how a range of actors employ the salient and difficult to define threat of terrorism to

---

65 A brief note of clarification: Frank Knight offers the central distinction between risk, which is calculable, and uncertainty which resists calculation. Here, I use risk is the colloquial sense of perceived danger or harm rather than the more specific sense indicated by Knight. In Knight’s terms, the process under
create new forms of security within the postal network, freight rail transportation, and the electric power grid. In each instance, the contradictions of deregulation emerge as problems to be controlled. The vulnerabilities created or enlarged by deregulation, it appears, offer terrorists with attractive targets. I use the term bureaucratization to capture the way in which the abstract threat of terrorism—the idea of a possible terrorist attack at some point in the future—is written into stable and routine practices of security expressed in new regulations, codes, practices, and technologies. The bureaucratization of risk translates uncertainty into a routine case.

In examining the process through which the threat of terrorism justifies new forms of security interventions, I follow Beck’s lead in ferreting out how, by whom, and to what end are new forms of security constructed. It is in answering these questions that we can begin to understand if the perceived threat of terrorism supported the democratization of infrastructure governance or the reification of power. As Beck is clear to note, risk, in and of itself does not compel; understanding the significance and implications of risk are, in each instance an empirical question. It is to this question that the following three chapters are devoted.
Chapter 5: Delivering the Mail: Terrorism, Efficiency, and the Politics of Post-Anthrax Bioterrorism Security

In 2001 an unknown number of letters containing spores of anthrax were sent through the U.S. postal system exposing scores of postal workers and patrons to possible harm.¹ Five individuals died as a result of anthrax exposure (22 were infected), while over 30,000 people were given prophylactic antibiotic treatment. According to estimates, one-third of the U.S. population took extra precautions in handling the mail after the attacks.² Letters containing anthrax were sent to media outlets in New York City and Boca Raton, FL, and government offices in Washington D.C., causing panic in the newsrooms and resulting in the shutdown of the Hart Senate Office Building and the temporary relocation of the Supreme Court. Although the original doctored letters may have only been aimed at a few different addresses and organizations, the effects of attacks were sweeping in scope. Cross-contamination—the spread of anthrax throughout the mailstream, moving from letters intentionally treated with anthrax to other letters, processing equipment, and facilities—greatly expanded the radius of exposure by sending unknown quantities of anthrax across the nation. Eventually, anthrax was found in over


60 different sites, one-third of which were postal facilities. Many of these locations never directly handled or housed a letter intentionally contaminated with anthrax. In the wake of the attacks, anthrax hoaxes involving doctored letters claiming to contain anthrax became common disruptions; the investigative arm of the United States Postal Service (USPS), the United States Inspection Service (USPIS), dealt with over 17,000 such incidents and evacuated over 600 postal facilities in the year following the attacks. The economic costs of the attacks were also staggering: clean-up of U.S. Postal facilities alone ran into the hundreds of millions of dollars, while other related economic costs tied to declines in letter volume, productivity, and new post-attack security measures amounted to nearly $6 billion dollars.

For a period of time, one of the more mundane aspects of daily life, opening the mail, became imbued with danger and unease. As Chief Postal Inspector Lee Heath noted, the attacks were characterized by the use of a “weapon of mass disruption.” While not reaching catastrophic levels of harm, the tragic deaths of five individuals compounded by widespread anxiety and far-reaching economic consequences elevated

the attacks into a crisis. The attacks destabilized the continued flow of mail, imperiled, to a degree, the health of workers and patrons, and introduced potentially ruinous costs for the already-financially-precarious Postal Service. In this sense, the threat of biological terrorism created a crisis of control for the postal system: it threatened to disrupt the processing and distribution of the mail, the health and safety of workers and patrons, and the economic viability of the enterprise.

Chapter Overview and Outline

The 2001 anthrax attacks created a moment of reflexive modernization that revealed the limitations and drawbacks of deregulation and, for a moment, served to democratize postal governance (see Chapter 3 and Chapter 4). The attacks highlight the ways in which regulatory reform created not only new efficiencies within postal operations but also new forms of vulnerability. The transformation of the postal network into a tightly-coupled network unwittingly amplified the anthrax attack in ways that confounded public health and defense experts. During the anthrax attacks, the tightly-coupled structure of the network fostered cross-contamination of postal facilities and helped undermine existing security plans. Here, the success of deregulation reappears as a public problem; or, as Beck might offer, the contradictions of deregulation are laid bare. At the same time, the anthrax attacks briefly served to democratize the governance of postal service. The threat of biological terrorism, perhaps somewhat counter-intuitively,

---

fostered the democratization of postal governance. The threat of terrorism disrupts the usual structure of interest group politics and opens the workings of the postal system to a greater range of participants, including postal labor unions, than had been previously admitted. Contrary to other domains, the threat of terrorism led to a more pluralistic, inclusive, deliberative process, rather than greater secrecy, consolidation of power, or a quieting of voices. Yet, while diverse postal stakeholders were given a voice in drafting new security protocols, the prevailing hierarchies of power that dominate postal politics—the centrality of large volume mailers and the ongoing subordination of labor—are ultimately encoded within the organization and operation of new control technologies designed to counter bioterrorism. While the anthrax attacks did create a brief moment of reflexivity and discontinuity, the new practices of security eventually implemented within the network embody the familiar structure of power characteristic of postal politics.

In this chapter, I examine the 2001 anthrax attacks and the creation and implementation of new forms of postal security designed to combat biological terrorism. The chapter considers first how the anthrax attacks created a crisis of control that destabilized the postal network, and next what we might call the bureaucratization of risk: how the uncertainty of biological terrorism is translated into new routine practices of security. To this end, the chapter is divided into three main sections. The first section provides a brief narrative overview of the anthrax attacks of 2001 and examines how the

---

8 For a discussion of the transformation of postal service, see Chapters 2 and 3.
9 A brief clarification: Frank Knight offers the central distinction between risk, which is calculable, and uncertainty which resists calculation. Here, I use risk in the colloquial sense of perceived danger or harm rather than the more specific sense indicated by Knight. In Knight’s terms, the process under examination here and in Chapters 6 and 7 would be more properly termed the bureaucratization of uncertainty. Frank Knight. *Risk, Uncertainty, and Profit.* 1921. New York: Kelley & Millman, 1957.
tightly-coupled network structure amplifies the attacks in ways that challenges the efficacy of pre-attack planning. The second section examines the work of the innovative Mail Security Task Force in devising strategies to combat the threat of biological terrorism. Here, the typical structure of postal governance gave way to a more inclusive forum. Finally, the third and central section examines politics encoded within the Biohazard Detection System (BDS) and Intelligent Mail (IM), new systems of control and surveillance implemented after the attacks to address the threat of biological terrorism.

*The Anthrax Attacks of 2001: Systemic Risk and the Problem of Networks*

In the fall of 2001, spores of anthrax flit across the nation causing illness, panic, and disruption. Though letters packed with anthrax were only sent to a few specific addresses, spores of the bacteria spread across the postal system and wound up in over 60 different locations in six different states.\(^1^0\) Anthrax appeared in disparate places—the Federal Reserve Board, Howard University, Walter Reed Hospital, rural Connecticut, to name a few—linked by the common thread of postal service.\(^1^1\) Defense planners and public health officials worried about the possibility of an anthrax attack since at least World War II; for decades, extensive efforts attempted to model and prepare for an attack by relying on a blend of epidemiological information and inference. Yet, despite over 50 years of concern, the anthrax attacks of 2001 confounded expectations and presented new, difficult, challenges. Planners long envisioned complex scenarios involving mass

---

exposure in a limited geographic area. In this instance, however, anthrax could not be reduced to a finite geographic location. Rather, the attacks spread small quantities of anthrax across the expansive postal network; it was a system-wide or network problem. The architecture of the network transformed the attacks from a site-specific challenge reducible to the handful of targeted addresses into a case of systemic failure; the topology of the postal system unwittingly amplified the attacks. Deregulation transformed the postal network and made it possible to process large volumes of mail quickly and cheaply in large processing and distribution centers. Yet, at the same time these efficiencies created new vulnerabilities and possibilities of large-scale failure that the anthrax attacks exploited.

Uncertain Origins and Limited Exposure: Initial Cases, September 18-October 17

Two separate batches of letters containing anthrax, totaling between 5 and 7 mail pieces, were postmarked on September 18th and October 9th in Trenton, NJ. The first set of letters was sent to media establishments, including the network news departments of ABC, CBS, and NBC, the New York Post, and the tabloid publisher American Media Inc. (AMI), while the October letters were sent to the offices of Senators Tom Daschle and Patrick Leahy. The presence of anthrax in the mail, however, was not immediately apparent. The scope and scale of the attack gradually became known as multiple cases of anthrax appeared and as letters containing spores of the disease were found. During the

---

12 The letters sent to NBC, the New York Post, and both Senators were recovered. Investigators established the presence of additional letters based on environmental testing and additional evidence, though not all presumed letters were recovered. Dellafera, Affidavit in Support of Search Warrant; GAO, “U.S. Postal Service: Better Guidance Needed to Ensure Appropriate Response to Anthrax Contamination.” GAO-04-239, Sept. 2004. 12.
first few weeks, public officials thought that the attacks were narrowly targeted and placed at risk only those few individuals that had been sent a letter intentionally filled with anthrax. The discovery of cross-contamination during mid-October, however, transformed the attacks from a site-specific and relatively limited concern into a systemic and network-wide issue.

The first indication of a possible attack using anthrax occurred on October 4th when Bob Stevens, a photo editor at AMI in Boca Raton, FL was confirmed by the Florida Department of Health to have contracted inhalation anthrax. Initially, Tommy Thompson, Secretary of the Department of Health and Human Services, speculated that Stevens may have contracted the disease naturally through drinking contaminated water on a recent camping trip to North Carolina. Anthrax does occur naturally and is most commonly associated with the processing of animal hair. However, the Stevens case raised suspicions: cases of inhalation anthrax have been exceedingly rare in the U.S; in the 20th century there were only 18 documented cases of inhalation anthrax and none since 1978. Thompson, and other public officials, cautioned against jumping to the conclusion that the case was the result of a deliberate act or terrorism before more information could be collected. In response to the confirmation of the Stevens case,
Thompson remarked that “it is an isolated case…There is no terrorism” and Dr. Jeffery Koplan, director of the Center for Disease Control (CDC), noted that there were no indications of a larger outbreak and that terrorism was only one of many possible explanations.\textsuperscript{18}

The possibility that the Stevens case could have been the result of natural causes was dispelled on October 7\textsuperscript{th} as spores of anthrax were found in the AMI building and one of his co-workers was confirmed to also have contracted the disease.\textsuperscript{19} The spread of anthrax was directly linked to the postal system on October 12\textsuperscript{th}, as an NBC staffer was confirmed to have contracted anthrax from opening a letter containing the substance addressed to Tom Brokaw.\textsuperscript{20} In the ensuing days, an additional letter containing anthrax was found at the New York Daily News, and cases of infection were confirmed at ABC and CBS that presumably resulted from contact with infected mail.\textsuperscript{21} On October 15\textsuperscript{th}, anthrax spread to the nation’s capital as a staffer at Sen. Daschle’s office opened a letter containing anthrax (an additional letter treated with anthrax sent to Sen. Leahy was found unopened in the following days).\textsuperscript{22} Once it was established that anthrax had been sent through the mail, government officials quickly labeled the sending of the letters as an act of terrorism, but the actual source of the attacks remained uncertain and open to much debate inside and outside of the government.\textsuperscript{23}

\textsuperscript{18} Ibid. 20-21.
\textsuperscript{19} Ibid. 23.
\textsuperscript{20} Ibid. 27.
\textsuperscript{21} Ibid. 28-40.
\textsuperscript{22} Ibid. 32, 61.
\textsuperscript{23} While there was near unanimity in calling the mailings an act of terrorism, government officials alternately speculated that the mailings may be domestic in origin, linked with the terrorist attacks of 9/11, or affiliated with Iraqi weapons programs. See, ibid.
At this point, public health officials still saw the danger posed by the anthrax letters to be limited to persons who received and opened a letter intentionally treated with anthrax. In keeping with this assumption, during these early days of the attacks USPS issued guidelines for handling mail and identifying a suspicious letters (those with misspellings, no return address, stains, and other seemingly anomalous characteristics), and made available gloves and masks to workers upon request.\textsuperscript{24} In a note to customers on Oct. 17\textsuperscript{th}, the Postal Service urged: “[t]he mail is safe! People shouldn't stop using the mail because of these isolated incidents… Everyone, in the mailing community, as well as the American public, should exercise common sense.”\textsuperscript{25} Likewise, during Congressional testimony Postmaster General Potter advised postal patrons to be vigilant and recognize mail that was out of the ordinary. The familiar letters, bills, greeting cards, and catalogues, however, in Potter’s estimation, were safe.\textsuperscript{26} A \textit{New York Times} editorial counseled restraint: “A logical assessment of the odds of harm helps. Should a person inadvertently open a letter containing anthrax spores, he or she can prevent the onset of disease by promptly seeking medical treatment.”\textsuperscript{27} The editorial assumed, like public health and government officials, that only those opening a letter intentionally treated with anthrax were at risk. Despite the assurances that the attacks were isolated and no cause for hysteria, false alarms led to evacuations of office buildings, newspapers, doctors’


\textsuperscript{25} USPS. “USPS Message to Customers.”

\textsuperscript{26} John E. Potter. Testimony. \textit{Oversight of the U.S. Postal Service: Ensuring the Safety of Postal Employees and the U.S. Mail}. 107\textsuperscript{th} Cong., 1\textsuperscript{st} Sess. 86-134, 2001. 111.

offices, and business across the U.S and abroad during the first weeks of October.

Unidentified white powders that, in most instances, were harmless substances like household detergent now triggered emergency response measures.28

_Cross-Contamination, Network Interdependencies and Systemic Risk: October 18th and Beyond_

The understanding of the risks presented by the anthrax mailings changed shortly after the discovery of the Daschle letter. Over the next several weeks it became clear to those tracking the cases that anthrax spores could, and indeed did, travel from well-sealed envelopes during the transportation, sorting, and distribution of letters. Cross-contamination—the spread of anthrax spores from an initially treated letter to others pieces of mail, processing equipment, and facilities—transformed the anthrax attacks from a site-specific, limited, though frightening, set of incidents into a system-wide network problem. Spores of anthrax appeared in postal facilities, homes, and offices that never had contact with one of the initial letters. Nine postal workers contracted the disease as a result of their proximity to the mailstream, while two individuals in New York and Connecticut died as a result of cross-contamination. Cross-contamination challenged earlier assertions that only letters that were purposefully treated with anthrax were dangerous. U.S. Surgeon General, Dr. David Satcher, reconsidered the assumption by public health officials that cross-contamination did not present a risk: “Until a week ago, I think all of the experts would have said [that cross-contamination was not a

danger]. The fact of the matter is, we were wrong.”29 In the wake of verified instances of cross-contamination, Postmaster General Potter also offered a dour qualification of earlier statements downplaying or discounting the risks of cross-contamination, stating that the safety of the mail could not be guaranteed.30

Strong indications of cross-contamination first came on October 18th as two postal facilities that handled the mail for AMI in Florida and the Senate mailroom, where the Daschle letter was handled, tested positive for anthrax.31 Deborah Willhite, VP with USPS, stated that the finding of spores in the Senate mailroom was “the first time our concerns that anthrax might be able to escape an unopened envelope were sort of confirmed.”32 After postal workers in the Trenton, NJ and the Brentwood, MD processing and distribution centers became ill with anthrax, both facilities were closed. Tests established that both processing and distribution facilities contained potentially dangerous levels of anthrax spores and required decontamination. Two cases of fatal anthrax infection further pointed toward the problem of cross-contamination. A hospital worker at the Manhattan Eye, Ear, and Throat Hospital died on October 31 and on November 21 an elderly Connecticut resident died due to inhalation anthrax presumed to have been caused by cross-contamination. Testing eventually found anthrax in over 23 different postal facilities, including the massive New York City Morgan and the Wallingford, CT processing and distribution centers (Table 5-1). A host of government buildings—including the Supreme Court building, State Department headquarters, the

30 Ibid.
31 Ibid. 38-40.
32 Ibid. 40.
Pentagon, CIA offices in Langley, VA, the Hart Senate building, the Russell Senate Building, the offices of Senators Ted Kennedy and Christopher Dodd—also tested positive for anthrax.\(^{33}\)

Cross-contamination occurred while the anthrax letters were handled by high-speed automated sorting and processing equipment within originating and destination postal processing and distribution centers (P&DCs). All of the letters appear to have first been mailed from a Trenton, NJ drop box, collected, and then sent to the regional Trenton Processing and Distribution Center (P&DC) for routing. Next, the letters were sent to various downstream destination P&DC for distribution to local post offices and delivery: the AMI letter was routed through the West Palm Beach P&DC; the NYC letters were sent through the Morgan P&DC; and the D.C.-area letters were processed at the Brentwood P&DC (see Chart 5-1 and Chart 5-2). Within the Trenton P&DC, three automated machines processed the letters: advanced facer canceller system machines (AFCS) canceled their postage; optical character readers (OCR) encoded barcodes corresponding to their intended destination; and delivery barcode sorters (DBCS) sorted them into different groupings for distribution. At the destination P&DCs, the letters were again run through automated machines—DBCS—that sorted them for final distribution into delivery sequence. Automated equipment processes letters at rates of up to 30,000 letters per hour and, in the process, compresses letters in a manner that can lead spores to escape from a well-sealed envelope though minute tears or pores in the paper. During

\(^{33}\) Though 60 different locations tested positive for anthrax, it is likely that testing did not identify all of the sites that contained anthrax and accurately map the scope of cross-contamination. For a discussion of the limitations of the testing process, see GAO. “Anthrax Detection: Agencies Need to Validate Sampling Activities in Order to Increase Confidence in Negative Results.” GAO-05-251. March 2005.
these various points, spores of anthrax spilled out of the envelopes and traveled through the air contaminating other pieces of mail, machines, and work spaces, and exposing workers to harmful levels of anthrax.

Cross-contamination rendered hollow many of the invocations of public health and postal officials during the early days of the attack. Pleas for caution and vigilance were not entirely useful. Cross-contaminated mail showed none of the unusual or suspicious markings, such as misspellings or the lack of a return address, to which the public was told to be alert. On the contrary, cross-contaminated mail bore no indicators that it might present a danger; anthrax spores were not visible to the naked eye and attached themselves to the familiar mix of letters, junk mail, and magazines that filled mailboxes across the country. Cross-contamination transformed the attacks into a case of network-failure; the postal system itself was contaminated, potentially exposing all those associated with its operation—including postal workers and anyone receiving mail, essentially the entire population—to harm. CDC Director Dr. Koplan offered a bleak assessment of the challenges of cross-contamination, stating that “[t]here seems to be the potential for not just hundreds and just thousands, but tens of thousands and maybe more letters to be potentially at risk for some level of cross-contamination.”

Network Interdependencies and Systemic Risk: Topology and Failure

Cross-contamination transformed the anthrax attacks into a system-wide, operational, problem. Processing and distribution centers (P&DCs) are the dense nodes of the postal network. The 237 P&DCs within the network serve as hubs that link

---

34 Ibid. 73.
roughly 40,000 post offices, 350,000 collection boxes, and 136 million unique delivery addresses into a national network that process 680 million pieces of mail daily.\textsuperscript{35} P&DCs rely on automated equipment to sort and process high throughput of incoming and outgoing regional mail and are key sites where pre-sorted, “bulk” mail, is comingled with single-piece mail. The size and importance of these facilities to the postal system is hard to overestimate (see Table 5-2). The Morgan facility, for example, is the central hub for nearly all incoming mail for New York City, covers an entire city block—measuring at 2.2 million square feet—and processes 12 million pieces of mail per day.\textsuperscript{36} The Brentwood facility provides all incoming and outgoing mail processing for Washington, D.C. (zip codes 20000-20099), connects with 56 local USPS facilities (including branch post offices and stations), and serves 177 different federal facilities.\textsuperscript{37} The Trenton P&DC has a similar profile; it collects and processes all incoming and outgoing mail for the region, covering 50 local post offices, and sends mail to 12 other P&DCs downstream.\textsuperscript{38}

The centralization of processing and distribution activities and the introduction of automation technologies, as discussed previously, greatly increased (and continues to increase) the efficiency of the postal system; these network reorganizations allow the

postal system to process high throughput of mail volume more quickly, in fewer locations, at a lower cost. The replacement of distributed processing with centralized processing was one of the vital goals and achievements of regulatory reform. However, centralization into dense hubs also inadvertently increases the scope, and risks, of cross-contamination. Once spores of anthrax circulated within a P&DC, they could—and did—jump to other P&DCs and post offices and transform a local incident into regional and even national problem (see Chart 5-2). In this case, the tight-coupling of the postal network amplified the anthrax attacks and transformed the attacks into a much larger operational problem: it expanded the radius of both actual and potential exposure. USPS assessed mail-flow data to create a list of at-risk facilities that received significant quantities of mail from or were otherwise in close connection through supply exchanges with the Brentwood and Trenton P&DCs (data from the other impacted P&DCs, were not compiled) and found 179 facilities to be at-risk.39

Cross-contamination potentially exposed millions of pieces of mail to spores of anthrax and created a backlog of mail that could not be immediately delivered. The Brentwood and Trenton P&DCs, for example, processed approximately 85 million pieces of mail after being heavily contaminated by spores of anthrax before being closed.40 The discovery of cross-contaminated mail also created a back-log of millions of pieces of mail caught within contaminated P&DCs.41 USPS sent the mail housed within closed


41 Postal records did not count the exact number of letters from contaminated P&DCs. GAO. “United States Postal Service: Information on the Irradiation of Federal Mail in the Washington, D.C., Area.” Briefing to
and contaminated P&DCs to private contractors in Lima, OH and Bridgeport, NJ for sanitation by irradiation and then hand-sorted and examined the mail before final delivery. In some cases, delivery of irradiated mail to federal agencies and locations in Washington, D.C. was delayed for three months.

The anthrax attacks underscore the relationship between network structure and forms of failure. Hub-and-spoke network architecture offers a number of benefits. When economies of scale and/or increased speed of processing are achieved through the centralization of processing activities within a limited number of dense or critical “hubs,” such a model makes the network more efficient by eliminating redundancies. Yet, the same model is, to a degree, vulnerable to single-point, or cascading, failures. Failures within these critical nodes can branch out and have system-wide impacts. Distributed processing and redundancies may not provide obvious benefits under normal operating conditions, but such a network—characterized by relative decentralization and an expansive number of less-densely interconnected hubs or nodes—is fault tolerant: failures within nodes with limited connectivity are isolated and localized. In this sense, the increased density of interconnections within any set of central hubs or nodes correspondingly increases system-wide vulnerability.

---

The reorganization of the postal network into a tightly-coupled system, as already discussed, was an outcome of the process of partial deregulation that unfolded during the 1970s and early 1980s. Previously, before regulatory reform, the structure of postal regulation supported labor-intensive distributed processing and discouraged the adoption of centralized plants. The transformation of the material postal network, under normal operating conditions, improved postal performance according to most metrics. Yet, the transformation, bound the network flows together in ways that, inadvertently, create systemic risk. Unwittingly, the anthrax attacks of 2001, exploited the particular network topology of the postal system to greatly expand the radius of both real and potential anthrax exposure from a handful of targets to, in effect, anyone coming into contact with the mail.\footnote{Unwittingly in the sense that it does not appear that the mailer of the anthrax letters intended for cross-contamination. The letters were well-sealed with tape.}

*Old Fears and New Challenges: Bio-Preparedness before the Fall of 2001*

The anthrax attacks of 2001 were, in many ways, unsurprising. Since at least World War II, defense strategists and public health officials worried about the possibility of a biological attack. In fretting about such a possibility, anthrax, due to its durability and relative ease of production, often topped the list of likely agents that could be used.\footnote{As two analysts for U.S. Army Medical Research Institute of Infectious Diseases (USAMRIID) summarized: “Anthrax, in the minds of most military and counterterrorism planners, represents the single greatest biological warfare threat,” Cieslak and Eitzen. “Clinical and Epidemiological Principles of Anthrax.” 552. See also, Donald Henderson. “The Looming Threat of Bioterrorism.” *Science* 283.5406 (1999): 1279-1282; Inglesby, et al. “Anthrax as a Biological Weapon: Medical and Public Health Management.” 1735-1745; Tom Mangold and Jeff Goldberg. *Plague Wars: A True Story of Biological Warfare*. New York: St. Martin’s, 2000.} Planners assumed that an attack using anthrax would fall into one of two categories: a mass casualty scenario whereby a large, concentrated, population was exposed to a large
number of spores; or, limited exposure through the mail where a small number of individuals in a single room or building would be exposed. These assertions of “what-if” were based in part on analogy, inference, intelligence gathering, and scientific research. The hypothetical cases blended facts, recent events, and fears into two coherent formulations depicting how an attack might unfold. These two different scenarios or normative cases guided preparedness efforts and led to the development of protocols, knowledge, and resources designed to confront the challenges they seemed to entail.

The 2001 attacks, however, fell uneasily between these two models and challenged the basic assumptions that undergirded planning and preparedness efforts before 2001. Cross-contamination spread anthrax in small, yet potentially dangerous, quantities across the geographically expansive postal network. Rather than subjecting a bounded population to exposure, cross-contamination potentially exposed everyone that came into contact with the postal system to some degree of harm. Cross-contaminated mail also rendered the spread of anthrax invisible to the naked eye. As a result, the attacks created new challenges for postal labor, management, mailers, and public health officials.


Thomas V. Inglesby, writing in *Emerging Infectious Diseases* in 1999, offers a long-form narrative of what a potential large-scale attack might look like. Inglesby sets the scene:

On the evening of November 1, a professional football stadium before an audience of 74,000. The evening sky is overcast; the temperature mild, a breeze blows from west to east. During the first quarter of the game, an unmarked truck drives along an elevated highway a mile upwind of the stadium. As it passes the stadium, the truck releases an aerosol of
powdered anthrax over 30 seconds, creating an invisible, odorless anthrax cloud more than a third of a mile in breadth. The wind blows the cloud across the stadium, and onward for miles over the neighboring business and residential districts… the anthrax is detected by no one.\footnote{Thomas V. Inglesby, “Anthrax: A Possible Case History.” \textit{Emerging Infectious Diseases} 5.4 (1999): 556.}

Inglesby’s narrative is representative of how an attack was most commonly conceptualized before the fall of 2001. Similar accounts and speculative formations circulated within U.S. defense, public health, and scientific communities for a half-century. The “mass-casualty scenario” appeared in public information campaigns during the 1950s produced by the Federal Civil Defense Administration, Department of Defense (DOD) and U.S. Army and was subject to a succession of tests and studies undertaken by the U.S and World Health Organization (WHO) during the 1960s and 1970s.\footnote{The U.S. tests attempted to determine the possible scope of an anthrax attack. For example, one series of tests simulated the dispersal of anthrax in a crowded New York subway during rush hour. Another military trial over the Pacific Ocean in the 1960s simulated the release of anthrax from a plane. Similarly, an influential World Health Organization (WHO) study in 1970 took for granted that an anthrax attack would target a populous area via a bomb or plane mounted aerosol. See, James C. Pile, et al. “Anthrax as a Potential Biowarfare Agent.” \textit{Archive of Internal Medicine} 158.5 (1998): 429-434; Inglesby, et al. “Anthrax as a Biological Weapon, 2002: Updated Recommendations for Management.” 2236-2252; World Health Organization. \textit{Health Aspects of Chemical and Biological Weapons}. World Health Organization: Geneva, 1970. 84-100; Elizabeth Fee and Theodore M. Brown. “Preemptive Biopreparedness: Can We Learn Anything from History?” \textit{American Journal of Public Health} 91.5 (2001): 721-726.}

In the post-Cold War period, the model was cited again to caution against the insecure stocks of bio-weapons in the former Soviet republics, the threat posed by so-called “rogue nations” (such as Iraq), and terrorist organizations.\footnote{United States. Office of Technology Assessment (OTA). \textit{Proliferation of Weapons of Mass Destruction: Assessing the Risk}. OTA-ISC-559. Washington, D.C.: GPO, 1993; Henderson. “The Looming Threat of Bioterrorism.” 1279-1282.}

Advocates of bio-preparedness and defense were a vocal minority during the middle part of the 20\textsuperscript{th} century, continually warning and drawing attention to what they perceived to be an overlooked threat.\footnote{Elizabeth Ethridge. \textit{Sentinel for Health: A History of the Centers for Disease Control}. Berkeley, CA: University of California Press, 1992; Fee and Brown. “Preemptive Biopreparedness.” 721-726; Russ Zajtchuk and David Franz. “Biological Terrorism.” Committee on Counterterrorism Challenges for Russia} To this end, the CDC created the Epidemic
Intelligence Service (EIS) in 1951 with the explicit goal of preparing for the possibility of a biological attack. The EIS placed trained epidemiologists in the field at a number of locations throughout the country to discern the early-warning signs of a bio-attack (or an otherwise atypical outbreak of disease). By 2000, multiple sectors of government embraced the mass-casualty scenario and interpreted and translated it into corresponding preparedness efforts. Planners assumed that an anthrax attack would unleash a flood of sick and worried people that would overwhelm the capacities of local health care institutions. During this period, preparedness efforts focused on creating surge capacity in the public health infrastructure through the creation of the National Pharmaceutical Stockpile and rapid-deployment teams that could be sent to a particular site of a biological attack to aid local officials. Additionally, public officials and emergency
responder games “mass casualty scenarios” of bio-warfare and bioterrorism in training exercises intended to simulate the challenges of such an event.\textsuperscript{54}

\textit{Hoax or Harbinger?: Limited Exposure through the Mail, 1998-2001}

Planning for an anthrax attack did not only look toward the worst case scenario described above. Although such thinking dominated discussions of bio-weapons and the risk they presented, a series of hoaxes during the late 1990s pointed toward the possibility that the postal system could provide the means to carry out an anthrax attack. Letters purporting to contain anthrax appeared in government offices, health clinics, and private businesses throughout the U.S and Canada during this period.\textsuperscript{55} Although none of these letters actually contained anthrax spores—they were doctored in most instances with a harmless powder—the hoaxes did point toward a possible method of attack that was far different from the continually recycled “mass-casualty” scenario.

CDC and USPS officials treated these letters as a serious but minor concern: a localized problem that could be dealt with through the isolation of the recipient of the letter and those in the immediate vicinity of its opening. In 1999, the CDC responded to the hoaxes by releasing interim guidelines defining how public health officials should respond to alleged incidents of anthrax in the mail. The guidelines recommended that

\textsuperscript{54} These games took the form of table-top exercises within which agency heads and other public officials walked through the steps of responding to a novel event or field events using a wide number of participants to simulate the chaos that confronts responders across the different levels of government. Between 1995 and 1998, 201 such exercises were undertaken, with two-thirds being dedicated to chemical and/or biological attacks. One notable simulation held in 1998 considered the impact of a disgruntled lab worker releasing anthrax in New York City, killing over 140,000. Importantly, these simulations focused on the use of bio-weapons according to the familiar “mass-casualty scenario.” Joby Warrick and Joe Stephens. “Before Attack, U.S. Expected Different Hit.” \textit{Washington Post}. Oct. 2, 2001.

only the environment in direct contact with the letter and its contents should be decontaminated and outlined how, in the event of a credible incident, decontamination of skin and clothing should proceed. Several months later, USPS adopted these principles and issued the new management instruction, “Emergency Response to Mail Allegedly Containing Anthrax,” outlining a new policy concerning how to handle possible anthrax contamination. USPS policy specified that in the event of the discovery of a suspect piece of mail, workers were to notify their supervisors, isolate the item, and wash their hands. Canadian defense researchers conducted the only known laboratory tests of the distribution of anthrax through the mail prior to the fall of 2001 and echoed the conclusions reached by the CDC and USPS. Through lab tests, the researchers confirmed that opening a letter with anthrax could result in the exposure of lethal quantities of anthrax to both the individual that opened the letter and those in the immediate area. The test did not, however, assess the risk that the letter posed as it traveled throughout the mailstream, instead only focusing on the moment when the letter was opened.

Normative Cases and their Limits: Between Two Scenarios

58 Ibid.
59 As Canadian Defense researchers assessing the risk posed by letters treated with anthrax concluded with prescience weeks before the actual attacks: “It is only a matter of time until a real anthrax letter arrives in some mail room.” Kournikakis, et al. “Risk Assessment of Anthrax Letters.” 13.
60 Ibid. 12.
61 Ibid. 13.
Establishing guidelines and anticipating the challenges of an event that is without a direct historic analogue is an interpretive and imaginative challenge. There were no documented cases of using anthrax as a bio-weapon in the U.S. before the 2001 attacks. Decades-long concern with this possibility, however, posited two normative cases, mass casualty and limited exposure, of what a possible attack might look. The various efforts to prepare for an anthrax attack in many cases had benefits that were applicable to areas beyond bio-war (for example the creation of the EIS). Yet, the partiality of the strategies cannot be ignored. Indeed, the actual events of the fall of 2001 threw into relief the limits and shortcomings of such preparatory efforts. The actual attacks looked much different from the hypothetical models. Rather than dealing with large numbers of sick or a limited, bounded, number of “at risk” individuals, cross-contamination of the postal network subjected everyone that came into contact with the mail to some possibility, however small, of exposure. Cross-contamination dislocated and unmoored the attacks from a set of finite, at-risk, sites into a network problem. Postal workers and others impacted by cross-contamination did not know they had been exposed to possibly deadly levels of anthrax until after they became ill. The flood of hoaxes and false alarms that followed the discovery of anthrax in the mail, now presented serious operational challenges as well. Once innocuous spots of dust and powder lingering on pieces of mail became, in the weeks and months after the attacks, suspect; unidentified particles offered reason enough to close and evacuate postal facilities across the country.62

---

62 The United States Postal Inspection Service dealt with over 17,000 hoaxes and suspicious cases, and evacuated over 600 postal facilities, in the year following the attacks. USPIS. *Annual Report of Investigations of the United States Postal Inspection Service*, 12.
Creating a Safe Postal System: Focusing Events and Risk Entrepreneurs

The anthrax attacks served to politicize the postal system and introduced a new set of concerns and problems through which postal management and other stakeholders navigated. The attacks disrupted the usual structure of interest group liberalism and institutional claims and opened postal governance. The attacks served as a “focusing event” that elevated the issue of postal security to a place of prominence for both postal management and, to a degree, the broader public. Suddenly, the threat of biological terrorism to postal security became a pressing public problem that could not be adequately contained by the practices of security already in place. Postal security became an object of debate, reflection, and ultimately action. The anthrax attacks placed security at the top of the agenda for groups such as management, labor, and large volume mailers that were already routinely involved in the day-to-day operations of the postal system. But, the attacks also pulled groups not typically or frequently involved in the minutia of the postal system into the ambit of postal politics: Congress, the Presidential Office of Science and Technology Policy, U.S. District Courts, defense contractors, and local governments. These newcomers also played a role in defining the contours of new forms of postal security. The inner-workings of the postal system are for the most part obscure to all but the handful of groups wed to the system. The anthrax attacks disrupted this normal state of affairs and, briefly, pushed postal matters to the fore. In these ways, the attacks fostered a moment of reflexive modernization.

---

Still, although Congress, the courts, and others not typically concerned with the day-to-day operation of the postal system played important roles in defining the shape of postal security (see below), by and large postal labor and large volume mailers (that is, the mailing industry) were the central figures in driving new policy interventions. Both groups acted, toward at times differing ends, as “risk entrepreneurs” that attached themselves to the issue and seized on the relative salience of the risk to argue for new extra forms of security. These groups had the most invested in the ongoing viability of the postal system and were central in crafting new solutions that sought to reconcile the safety of postal workers and patrons with the ongoing operation of a low-cost, high-throughput, ubiquitous postal network.

Postal workers, both through their union representatives and individually, argued that the cross-contamination created a workplace-safety issue and aggressively pushed postal management to take extra precautions. To some postal workers, the anthrax attacks placed them on the “front lines” of a new and uncertain conflict against terrorism. The American Postal Workers Union (APWU), the largest union representing postal workers, was the most outspoken of the various unions. The APWU filed three lawsuits in U.S. District Court to mandate further testing and the closure of a variety of postal facilities during the fall of 2001. The union publicly charged the postal service with ignoring the health and safety of their workers and led a worker walkout at the Morgan P&DC. The APWU insisted that none of the workers it represented would

---

65 The APWU represents clerks, maintenance, motor vehicle, and support divisions.
work in a facility found to contain anthrax and filed suit seeking the closure of processing and distribution facilities in New Jersey, New York City and Fort Lauderdale, FL that were kept open despite being contaminated. The lawsuits sought an expanded program of testing and the closure of any facility found to contain anthrax. USPS was initially reluctant to shutdown facilities over what it, in consultation with public health officials from the CDC, considered to be minor amounts of anthrax. The Postal Service pushed for continued operation, while certain contaminated workspaces within the facilities were closed for cleaning. APWU President William Burrus publicly charged postal management with offering different levels of safety for government officials and postal workers. Burrus, and other postal workers, pointed out that the Supreme Court and Congressional offices were closed at the first hint of possible anthrax exposure, while postal facilities, in some cases, remained open even after tests established the presence of spores. As Burrus framed the issue: “Are the lives of postal workers less valuable than the lives of U.S. senators? Are the lives of postal workers less valuable than the lives of Supreme Court justices? What value do our lives have?” William Smith, President of New York Area Metro Postal Union, drew the same conclusion, offering: “I realize that Morgan employees are not Supreme Court justices or senators or Congress, but they are God's children; they have the same right to life as the aristocrats” and that “[n]o one piece

of mail is worth a human life.” Brentwood postal worker Keith Beckett was equally blunt, charging that “postal management only cares about moving the mail, not its employees.” Burrus additionally called on postal employees to refuse to work and request administrative leave in the event that they viewed their workplace as unsafe. In New York, during the height of tensions over the Morgan P&DC, one-third of workers stayed out of work. While only one lawsuit was successful and one facility, in New Jersey, was closed by court order and subject to additional tests, the actions of the APWU pressured USPS to take new steps to counter the threat of biological terrorism.

Postal labor was not alone in calling for new security measures. The mailing industry, through trade associations such as the Association for Postal Commerce, Mailers Council, the Direct Marketing Association, and others, pushed the Postal Service to recognize the potentially ruinous effect that the anthrax attacks could have on the industry. Where labor saw the threat of biological terrorism as an issue of workplace safety, the mailing industry saw possible economic ruin. Large volume mailers—heavy, bulk users of the mails for commerce including direct mailers, catalogue merchants, and other companies that rely on mailings as a central component of their businesses—worried that a loss of confidence in the safety of the mail, either by recipients who no longer opened so-called junk mail, catalogues, or other solicitations, or by advertisers that

---

69 Finn. “Union Chief’s Battle Is With the Postal Service.”
contracted with direct mailers to reach the mailboxes of possible customers, could undermine their business models. Direct mailers are mainly third party vendors that create and coordinate advertisement or marketing mailings on behalf of clients in a variety of industries and promote their services based on the response rates that such mailings generate and the timeliness of the delivery of the message. After the anthrax attacks, direct mailers worried that response rates would decline as recipients became reluctant to sort through non-personal mailings. They also worried that as offices instituted new restrictive mail policies concerning incoming mail, possible delays of timed-advertisements, promoting a special event such as an upcoming weekend sale, would render them anachronistic. Mailers worried that their customers would turn to other outlets—such as the telemarketing and the internet—in order to reach their target audience. Catalogue companies, the greeting card industry, credit card companies, as well as many others, echoed the same concerns and worried that the anthrax attacks would, rightly or wrongly, scare-off potential customers.

Yet, while both labor and industry sought new forms of postal security, mailers offered an important caveat. They argued that any new form of security should be cost neutral, that is not increase mailing costs or postage rates, and avoid adding to processing and delivery times.

---

76 Schultz. “Direct Hit: Mail War.”; “Reader’s Digest Hurt by 9/11 and Anthrax Scare.”
77 Paul M. Alberta. “USPS Expected to Seek as Much as $10 Billion in Aid From Congress.” Direct. Nov. 7, 2001; GAO. “Highlights of GAO’s Conference on Options to Enhance Mail Security and Postal
new added costs of security and supported federal subsidies to offset any added security costs. Direct Marketing Association (DMA) President H. Robert Wientzen articulated the industry position, arguing that the costs of new forms of postal security amounted to “a national security and a national commerce issue, which should be paid for by the national government.” Likewise, DMA Vice President Jerry Cerasale argued that absent federal funds, new forms of security would lead to a “rate increase so great that it would force many of [mailers] out of business.” Executive Director of the Alliance of Nonprofit Mailers, an industry trade group, Neal Denton echoed this view, noting that without federal appropriations “we may be seeing the end of the national mail network as we know it.”

In discussing security, the limitations and drawbacks of a concentrated market for postal services are evident. As discussed previously, since the 1970s, the market for postal services become increasingly contestable and concentrated (see Chapter 3). The rise of electronic messaging systems and electronic banking services, combined with the disastrous and aborted effort by the Postal Service to enter the electronic messaging market with its proto-email, E-COM, made large volume mailers more important than ever. As other users flocked to new outlets, the importance of large volume mailers only increased, effectively providing bulk mailer with a de facto veto over postal changes. The threat by mailers to resist any possible move to recover the costs of security through


79 Alberta. “USPS Expected to Seek as Much as $10 Billion in Aid From Congress.”
80 Ibid.
rate hikes is not simply, or only, overheated rhetoric. Indeed, there are other avenues open to traditional businesses that rely on the mail. A contestable and concentrated market makes additional costs difficult to pass to consumers. With other options available, large customers aligned together can offer credible threats of exit that management must take seriously. The combination of the availability of substitution services (such as electronic communication) and a well-organized group of customers that account for a large share of business creates a situation of near-capture where key customers gain some control over the setting and direction of postal policy. Focusing back on the issue of security, this combination challenged or constrained the ability of management to recover new costs of security through postal rates.

Postal service management was receptive to these concerns and suggested that the very survival of the postal service rested on the ability to restore confidence in the mail, ensure the health of postal workers, and limit the impact of new security costs on postal rates. In November, 2001 Postmaster General John Potter responded to the concerns laid out by both labor and mailers, and asked Congress for $5-6 billion to cover the direct and indirect costs of the anthrax attacks. Potter defined the direct costs as costs linked to clean-up and new security efforts and indirect costs ascribed to declining mail use. The Postmaster General articulated the position expressed by the mailing industry and warned that the postal rate-base could not and would not shoulder the burden of increased postal rates designed to cover the costs of security and would, possibly, turn toward alternative forms of communication (such as e-mail) unless federal appropriations off-set these new costs.

81 CRS. “The U.S. Postal Service Response to the Threat of Bioterrorism through the Mail.” 1.
Postal service, Potter asserted, represents not only a vital national infrastructure, but also had become “the front lines of a new war”; without an influx of funding, the Postmaster General further testified, service levels would be cut and the continuation of universal service would be sacrificed in the name of safety.\(^2\) Potter was partially successful: Congress would not provide funding to makeup losses of revenue, but USPS did receive $1.265 billion to off-set some of the new costs of enhancing the safety of the mail system and reduce the risks to patrons and workers through a series of appropriations distributed between 2001 and 2006.\(^3\)

*Seeking Security: Reflexivity and the Mail Security Task Force*

The attacks disrupted, for a moment, the typical structure of power within the postal system. The usual structure of interest group liberalism became unmoored as security took on a new urgency. In addition to lobbying Congress for funding, the Postal Service convened a new task force uniting senior postal management and representatives from the Postal Inspection Service, U.S. Postal Service Inspector General, labor unions, large volume mailers, the Presidential Office of Science and Technology Policy, and the CDC to draft new policies and procedures to combat the spread of biological terrorism. The formation of a new task force to oversee planning for postal security represented a


\(^3\) The USPS does not disclose the total costs of security but notes that the appropriations do not cover the total costs. The appropriations were distributed as follows: An Emergency Presidential Appropriation of $175 million, November 5, 2001; $500 million from the Department of Defense Emergency Supplemental Appropriations Act, 2001; $87 million from the Supplemental Act for Further Recovery From and Response To Terrorist Attacks on the United States, 2002; $503 million as a part of the Consolidated Appropriations Act, 2005. USPS. *2006 Comprehensive Statement on Postal Operations.* Washington, D.C.: USPS, 2006. 61.
novel departure from the familiar mold of postal politics. Postmaster General John Potter created the Mail Security Task Force (MSTF) in the midst of the attacks on Oct. 15th, 2001 and placed Chief Postal Inspector Kenneth Weaver (head of the Postal Inspection Service) in charge.84 The MSTF’s mission was to “[e]nhance the safety and security of all postal employees and contractors and improve customer confidence in the mail.”85 The steering committee of the MSTF comprised of the presidents of the postal labor unions and management associations, senior postal management, and the USPS Inspection Service. The MSTF met daily from October until the end of 2001 and then continued weekly meetings.86 The MSTF divided into seven working groups: Communicating and Messaging; Safety and Security in the Workplace; Mail Center Security; Contingency Planning; Mail Screening; Mail Preparation; Surface Transportation Security.87 The work of the MSTF culminated in the creation of the Postal Service’s Emergency Preparedness Plan in March of 2002, the blueprint for post-anthrax postal security that charted how new control technologies would be integrated into the postal network.

The operation of the MSTF offered a break from the Postal Service’s typical decision-making framework and embodied an understanding that the threat of biological terrorism presented an exceptional case that required novel forms of engagement. The MSTF is unique in that it was a forum that brought together stakeholders including

86 Ibid.
87 Ibid.
management, labor, and large users outside of the periodic adversarial rate cases. For
decades, large mailers held the ear of postal management in setting postal policy through
their involvement in the Mailers’ Technical Advisory Committee, a public-private
partnership run jointly by representatives of the mailing industries and USPS.88 The
Mailers’ Technical Advisory Committee serves as a steering committee for postal policy
that is closed to union participation.89 The MSTF, in its inclusion of both industry and
labor, however, was a novel disruption in postal policy that represented, at least
nominally, the democratization of decision making within postal politics: it opened
participation more fully to groups, the postal labor unions, previously shut out of direct
decision making and previously were, instead, solely reliant on lobbying and
participation in rate-cases. It also brought experts in the specialized technical fields, such
as weapons detection and epidemiology, into deliberations with key postal stakeholders.
In this manner, the attacks not only drew attention to the material workings of the postal
system, but also constituted a reflexive moment that, in a limited manner, democratized
postal governance.

The convening and collection of participants in the MSTF offered recognition of
the importance and limitations of technical expertise. The issues considered by the
MSTF called for expert judgments across a number of diverse fields—including biology,
epidemiology, and engineering—but substantive judgments about postal security
necessarily required weighing different values, between cost and safety, operational

88 Mailers’ Technical Advisory Committee. “Mailers’ Technical Advisory Committee Charter.”
efficiency and security, which could not be written off as merely technical. At a certain point, discussion of security move beyond the question of what does or does not “work” and began to involve more broadly drawn questions about the relative importance of workplace safety and economic viability, the equity of the assignment of costs, and the importance of privacy. By including not only experts in the different relevant fields, but also labor unions and certain customers in these deliberations, USPS recognized that substantive decisions about postal security encompass more than the application of technical accounting: it requires judgments about values. In this respect, the MSTF represented a model of decision making that was at once participatory and expert.

In charting a post-anthrax security strategy and drafting the *Emergency Preparedness Plan*, the Task Force considered a host of available and in-development technologies, policies, and process alterations that could improve mail security and the confidence of customers in the safety of the mail. The MSTF looked to balance worker and customer safety with the continued operation of a high-throughput, ubiquitous, low-cost postal system. To this end, the Task Force considered the effectiveness of possible new technologies and process changes in combating the spread of harmful biological agents through the mailstream, as well as the economic costs and operational impacts stemming from any such changes. Initially, in late 2001 and early 2002, the MSTF considered creating different sets of rates corresponding to different “risk” levels, eliminating street collection boxes and anonymous mail, and instituting a system of video recording and identification verification within post offices that would link each piece of
mail to a specific individual. However, these far-reaching revisions of postal operations were discarded as being too disruptive, impractical, and costly. Instead, the MSTF sought technologies near maturation that could be deployed without greatly adding new costs to mailers or delaying mail processing and distribution. By the spring of 2002, the MSTF concluded that no single technology could reduce the threat of additional biological attacks. Rather, the MSTF proposed the adoption of different complementary sanitation, prevention, detection, and identification technologies.

The work of the MSTF culminated in the completion in March 2002 of the Emergency Preparedness Plan. In the document, the MSTF outlined a detailed plain charting how USPS would confront the challenges of biological terrorism. The plan called for the limited adoption of sanitation technologies into the postal network on a trial basis, and pushed for the system-wide adoption of new preventative technologies that could help reduce the threat of cross-contamination, new detection technologies that could identify the presence of harmful biological agents, and identification technologies that could create better transparency in postal operations and allow greater track-and-trace capabilities.

*From Sanitation to ICTs: Evolving Approaches to Security*

The plan’s call for a limited role for sanitation technologies marked a reversal from what was initially anticipated to be the centerpiece of post-anthrax postal security. At first, the MSTF and postal planners considered sanitation technologies, irradiation in

---

90 GAO. “Highlights of GAO’s Conference on Options to Enhance Mail Security and Postal Operations.”;
USPS. *Emergency Preparedness Plan*.
particular, to be the most promising option for combating the spread of biological agents through the mail and hoped to implement system-wide irradiation.\textsuperscript{92} Irradiation is a sanitation process that uses targeted ionizing radiation to kill or neutralize harmful agents that is routinely used in industrial settings, such as food and medical equipment sterilization.\textsuperscript{93} After the discovery of anthrax in the Washington, D.C. area, and the targeting of members of Congress, USPS entered into agreements with two independent contractors, Titan Corporation and Ion Beam Applications, to irradiate mail stored within cross-contaminated postal facilities and to treat all mail destined for Washington, D.C. Zip Code 202 through 205 (covering most federal addresses and agencies).\textsuperscript{94} This limited use of irradiation was conceived as only the first step before moving to full network sanitation. Limited irradiation began in November, 2001 and the first batches of irradiated mail were delivered to Congressional Offices in January, 2001.\textsuperscript{95} USPS shipped mail to contractor facilities in Lima, OH and Bridgeport, NJ for irradiation, and then distributed the mail to its final destination. Rerouting mail for irradiation, however, added new delays to mail processing and distribution. During the first months of the program, letters traveling to the above Zip Codes were delayed by up to three months. Later, mail subject to irradiation on average took over twice as long to reach its final


\textsuperscript{95} General Counsel of the Office of Compliance. “Investigation of the Health Effects of Irradiated Mail.”
destination (8 days, as opposed to 2 or 3 days). Due to these delays (and contractor costs) USPS and MSTF did not consider rerouting and off-site irradiation as a viable option for a more expanded program of sanitation. Rather, USPS sought to test the feasibility of brining sanitation in house and situating irradiation equipment within P&DCs across the nation. USPS bought eight electron-beam (e-beam) sanitation systems from Titan Corp. under a $40 million contract for use in irradiating Washington, D.C. mail and as a way to test the feasibility of system-wide deployment.

The use of irradiation, however, quickly ran into operational problems, health concerns, and public opposition that prevented system-wide adoption. Irradiation neutralizes harmful biological agents by disrupting cell structures. Yet, as became clear as the first batches of mail were sanitized, the same process damages a variety of different items routinely sent through the postal system; seeds, medicines, plastics, (such as credit cards), electronics, and other materials experience varying degrees of damage when irradiated. Even simple letters and paper mailings can be damaged by the process; in some instances irradiation embrittles and discolors paper. Irradiation sanitized the mail, but at the expense of damaging, and in some cases rendering useless, the contents of the mail. Shortly after the introduction of limited irradiation, mailers within the MSTF Mail Screening Working Group argued against rolling out widespread

---

98 USPS. Emergency Preparedness Plan. 2-3, 4-25.
sanitation and suggested that, at the very least, commercial mail should be re-routed to avoid both the damages and added processing times linked with irradiation.101

Federal employees receiving irradiated mail also voiced complaints in early 2002 that their mail was causing nausea, headaches, nosebleeds, and other forms of sickness. The Congressional Office of Compliance opened an investigation into the possible harmful effects of irradiated mail in February, 2002 at the request of Senator Charles Grassley and House of Representatives employees.102 That same month, the National Institution for Occupational Safety and Health (NIOSH) also opened an investigation into the possible side-effects of handling irradiated mail at the request of the Senate Sergeant at Arms Office and the Chief Administrative Officer of the House of Representatives.103 Both the Office of Compliance and NIOSH conducted their studies during the same period during which the MSTF was considering how to approach post-anthrax security. The Office of Compliance concluded that “handling irradiated mail for substantial periods of time may be the cause, or a contributing cause, of adverse health symptoms reported by a significant number of Legislative Branch employees.”104 NIOSH reported that irradiated mail may be the cause of some of the reported symptoms, but that there was no evidence of long term health effects.105 Although both groups issued their reports in July, 2002, after the completion of the Emergency Preparedness Plan, the complaints

102 General Counsel of the Office of Compliance. “Investigation of the Health Effects of Irradiated Mail.”
of federal employees receiving irradiated mail created controversy over the health effects of the program and caught the attention of the MSTF and postal planners.

Local governments, however, offered perhaps the largest barrier to system-wide irradiation. After purchasing eight e-beam irradiation systems in October, 2001, USPS entered into two contracts to construct facilities to process sanitized mail and house the irradiation equipment near Washington, D.C. In early November, 2001 USPS leased a building in Temple Hills, MD to process sanitized mail. One week after starting to remodel and outfit the building, local officials shut off the building’s utilities in response to concerns over the safety of bringing sanitized mail into the county. As a result, postal officials terminated the contract and took a loss of $600,000 on incurred remodeling expenses. Local officials similarly derailed USPS plans to operate an irradiation facility in Montgomery County, MD. USPS leased a facility within the county to house the eight e-beam systems and irradiate the mail. However, that plan was also abandoned due to local pressure, resulting in a loss of roughly an additional $400,000.106

The combination of operational problems, health concerns, and local opposition derailed efforts to implement system-wide irradiation. By the spring of 2002, USPS and MSTF moved away from system-wide sanitation and embraced a strategy that included a mix of prevention, detection, and identification technologies. The *Emergency Preparedness Plan* outlined a minor role for sanitation technologies and called for the limited deployment of the already purchased e-beam systems to replace off-site...

---

irradiation and treat federal mail. The plan also set into motion the installation of new ventilation and filtration systems within P&DCs as a preventative technology to limit the circulation and recirculation of spores of anthrax and other airborne contaminants.

Importantly, the plan outlined a new role for information and communication technologies (ICTs) within the postal system as technologies of control and surveillance. ICTs held out the promise of security without the myriad drawbacks associated with sanitation. They seemed to hold out the promise of providing security without upsetting day-to-day operations. Through the Emergency Preparedness Plan, the MSTF outlined how new ICTs would be threaded within the postal network as distinct systems of detection and identification that, respectively, could identify the presence of biohazards and provide postal officials with greater forensic information through new track-and-trace capabilities. Over the course of the next several months, USPS deployed these new technologies, known as the Biohazard Detection System (BDS) and Intelligent Mail (IM), system-wide and, in the process, transformed the operation of the postal network. Though these systems avoided the problems associated with sanitation, a certain formation of politics is inscribed within these technologies of control. Both systems, in different ways, embody the triumph of large volume mailers at the expense of the general public

107 The plan held out hope that a suitable facility could be constructed in the D.C. area with local support. Even this modified and limited plan, however, never came to fruition. In 2003, the USPS transferred 6 of the 8 purchased e-beam systems to other public agencies. Between 2005 and 2008, the USPS spent $1.5 million conducting site surveys and environmental assessments in an effort to lay the ground work for an irradiation facility in Washington, D.C., before abandoning the project in favor of the continuation of outsourcing sanitization for government mail. Currently, irradiation remains outsourced and is used on a limited basis only for incoming federal mail. GAO. “United States Postal Service: Information on the Irradiation of Federal Mail in the Washington, D.C., Area.” 39-40; USPS. 2003 Comprehensive Statement on Postal Operations. Washington, D.C.: USPS, 2003. 103.

public and postal labor. While the anthrax attacks served to politicize in a novel manner the postal system and led to the tangible, if brief, democratization of postal governance, the ultimate new security measures that resulted were shot through with typical hierarchies of postal politics.

*Biohazard Detection System: Toward a Bifurcated Postal System*

The Biohazard Detection System (BDS) splits the mailstream into two separate and roughly parallel postal networks, integrated for the last mile and delivery; one for individuals or small organizations sending mail from their homes, street collection boxes, or post offices; another for large volume, commercial or bulk, mailers sending mail in large collated batches at discounted rates. The BDS, and its associated costs and extra handling, apply only to the first, individual, network, while large volume mailers effectively opt out of the costs and extra handling tied to new security procedures and, in its place, have the option of subscribing to a voluntary security program. The BDS, in its situation within the complex postal network, treats as “safe” the post of large, institutional, mailers, while scrutinizing the mail from individuals, homes, and other small entities. The BDS also carries a more or less explicit promise to institutional mailers: in the event of another anthrax incident, their mailings will be expedited and not subject to additional screening, delays, or costs. The BDS, in these ways, embodies the caprice of large volume mailers; it provides a type of security designed to reassure the public, but does so without adding to the postage rates or delivery times of bulk mailings. Sunk within the technical minutia of the system and its operation lies explicitly political
judgments concerning which categories of mail qualify as “safe” and “unsafe,” and who is responsible for paying the freight of postal security.

**BDS: Operation and Development**

The BDS is an automated detection system that tests air particles above particular mail processing equipment for the presence of anthrax and, in the event of a positive result, triggers both on-site and off-site alarms. The system is placed in P&DCs and collects and tests air samples released from individual letters as they are run through advanced facer canceller system machines (AFCS). The system is made up of numerous components, including a hood that fits over the AFCS and collects samples, a transmission line connecting the hood to a cabinet containing the equipment to test the sample, and a site controller—a networked computer terminal linked to the cabinet and other on-site and off-site information systems—that oversees system operation and relays positive tests results. As letters are run through AFCS, pinch belts squeeze individual letters and, in the process, a small plume of air and particulate matter is released into the hood covering the AFCS. In the words of one postal official, as letters travel through AFCS they nearly “explode” with dust and other matter as they are squished through the belts of the system. The BDS samples this plume by immersing the collected particles into a sterile liquid and injecting the liquid into a self-contained cartridge that is then

---

111 USPS. “Biohazard Detection System: National Communications Plan.”
analyzed for the presence of anthrax. The process uses polymerase chain reaction (PCR) to amplify the sample by creating multiple copies of the collected DNA in order to allow for greater testing accuracy and comparison of the sample against the DNA structure of anthrax (a match signals the presence of anthrax). The BDS continuously collects samples as the AFCS operates and then tests for the presence of anthrax. The assessment of the collected samples takes from 1 hour to 90 minutes; during this initial period the samples are tested and any positive results are automatically re-tested for confirmation. In the result of a positive test, the site controller triggers an audible and visual alarm within the postal facility, the processing of mail stops, and the facility is evacuated. Simultaneously, the site controller sends a notification off-site to USPS headquarters and USPIS and, subsequently, DHS and local first responders are notified. Finally, the test cartridge is removed by USPIS and is re-tested by the CDC laboratory response network for confirmation.

The BDS is a novel information and communication system that automates biohazard detection and, at the same time, does not decrease the throughput of mail processing. The system allows AFCS to operate at regular rates and, absent a positive result, does not delay the mail. The use of the BDS in P&DCs is intended to maintain the efficiencies achieved through the concentration and automation of processing and sorting, while eliminating the possibility of a cascading, systemic, failure. Although

116 USPS OIG. “Biohazard Detection System.”
117 USPS. “Biohazard Detection System: National Communications Plan.”
118 USPS. “Biohazard Detection System: National Communications Plan.”
USPS initially sought a real-time system of detection, the BDS is what it calls a “near-real time” system that can detect contaminated mail before it is sent downstream to other postal facilities and points of delivery.\(^{119}\) The system combines and automates the amplification of DNA (through PCR) with detection. PCR and detection are, of course, not new processes. But, the automation and combination of these processes is relatively new and the BDS is the first large-scale application.\(^{120}\)

USPS and vendors developed and tested the BDS equipment in 2002 and 2003 before completing national implementation in December, 2005. Initially, Northrop Grumman and Lockheed Martin conducted preliminary tests at the Dulles P&DC in March and April of 2002 to examine the airflow rates associated with mail processing equipment in order to develop a working detection system.\(^{121}\) After testing, USPS approved Northrop Grumman to proceed to the prototype phase.\(^{122}\) Northrop Grumman placed the first BDS in Baltimore, MD P&DC for testing in June, 2002.\(^{123}\) After the successful field test in Baltimore, the USPS Board of Governors approved $279 million for Phase I of the BDS in October, 2002. Under Phase I, USPS bought over 700 detection systems from Northrop Grumman for limited trial use in 15 P&DCs across the

---

\(^{119}\) USPS OIG. “Biohazard Detection System.”

\(^{120}\) BioWatch, instituted in 2003, is another large-scale biohazard detection program initiated after 2001 in order to provide early warning of an act of biological terrorism. BioWatch collects air samples from major urban areas and CDC labs test the samples daily for the presence of biological pathogens. This system is not, however, automated. See, Curtis Chan and Erica Pan. “Anthrax in the Air? Medical & Public Health Response to Environmental Biological Agent Detectors.” San Francisco Department of Health. <http://sfcdc.org>.

\(^{121}\) USPS OIG. “Postal Service’s Efforts to Implement Prevention and Detection Technology.” “Appendix A: Comments of Thomas Day, Vice President, Engineering United States Postal Service.”


\(^{123}\) USPS. “Biohazard Detection System: National Communications Plan.”

The system was (and is) costly and procured under unique contracting conditions. USPS will not release total costs of the development, implementation, and operation of the system. However, according to audits conducted by the USPS Office of Inspector General and published figures from the Washington Post, total system costs (including consumable materials) to date range between $1.05 and $1.4 billion dollars.\footnote{Spencer S. Hsu. “Anthrax Alarm Uncovers Response Flaws.” Washington Post. March 17, 2005; USPS OIG. “Biohazard Detection System Consumables.”} A significant portion of the incurred expenses are annual operating costs of above $100 million.\footnote{Elizabeth Lester, Gregory Bearman, and Adrian Ponce. “A Second Generation Anthrax Smoke Detector: An Inexpensive Front-End Monitor that Detects Bacterial Spores.” IEEE Engineering in Medicine and Biology 23.1 (2004): 130-135; USPS OIG. “Biohazard Detection System Consumables.”} A portion of these costs have been paid through appropriations, but the rest of the costs, including annual operating costs, are treated as operating expenses and recovered through postal rates.\footnote{Since 2001, the USPS has received four appropriations totaling $1.265 billion in federal funds for emergency preparedness. These funds have been used to pay for a portion of a number of different projects, including the BDS, ventilation and filtration systems, irradiation, clean-up, and testing. It appears based on available USPS cost data and OIG audits that less than half of these funds have been used to pay for the BDS system. USPS. Emergency Preparedness Plan. 6-5; USPS OIG. “Postal Service’s Efforts to Implement Prevention and Detection Technology.”; USPS. 2006 Comprehensive Statement on Postal Operations. 61; USPS OIG. “Biohazard Detection System Consumables.”; USPS. “Decision Analysis Report: Biohazard Detection Systems.” 5; USPS. “Response to the House Subcommittee on Transportation, Treasury, Housing and Urban Development, the Judiciary, District of Columbia and the Senate
certified under the Support Anti-terrorism by Fostering Effective Technologies Act of 2002 (SAFETY Act). The SAFETY Act empowers DHS to designate certain homeland security technologies for liability protection. Northrop Grumman initially held up the development of the BDS by demanding indemnification. In response, postal management considered either budgeting extra funds to provide liability protection for Northrop Grumman or legislative remedies. The SAFETY Act addressed this problem and provided the manufacturer with liability protection.

**BDS and a Bifurcated Mail System: Safe and “At-Risk” Mail**

The BDS only “scans” collection mail, letters sent from collection boxes on the street, homes, small businesses, and individuals, and does not apply to the mailings of large volume mailers. Large institutional mailers account for the majority of mail sent through the postal network and send their mailings in presorted bundles at discounted “presort” rates calculated based on the cost avoidance achieved by performing some extra preparation and sorting work. This mail bypasses the BDS and is not subject to scrutiny. The BDS operates by collecting samples from the AFCS. This canceling equipment only handles collection mail; bulk mailings are integrated by a different set of automation technologies at a later step in the supply chain for sequencing (though in the

---


132 Roughly 75% of all mail is sent with some worksharing discount. GAO. “U.S. Postal Service: A Primer on Worksharing.” GAO-03-927. July 2003. The *Emergency Preparedness Plan* provided a lower estimate and suggested that 60% of total mail volume is comprised of bulk business mailings and could bypass detection. USPS. *Emergency Preparedness Plan.* 1-3.
same facilities.\textsuperscript{133} By linking the BDS and AFCS, the postal service creates two distinct mailstreams divided by differing levels of security and allows large volume mailers to opt out of new security practices.

The decision of where to place the BDS is not reducible to a set of neutral expert judgments, but rather rests on certain assumptions about how anthrax would, in the future, enter the mailstream and which categories of mail are “at risk.”\textsuperscript{134} The decision to place the BDS on AFCSs linked to the assumption that collection mail, and not other categories of mail, is suspect.\textsuperscript{135} AFCS are the first automated equipment used to process collection mail, but they are not the only machines capable of spreading spores of anthrax throughout mail facilities and leading to cross-contamination. Indeed, during the fall 2001 spores of anthrax appear to have been aerosolized not only on the AFCS, but also by Delivery Bar Code Sorter (DBCS) that collate both collection and bulk mailings.\textsuperscript{136} The decision to only test these particular automation technologies, then, follows from the assignment of collection mail as “at risk” and bulk mailings as safe.

\textsuperscript{133} USPS OIG. “Postal Service’s Efforts to Implement Prevention and Detection Technology.” “Appendix A: Comments of Thomas Day, Vice President, Engineering United Sates Postal Service.”

\textsuperscript{134} The USPS did collect intelligence to assess the threat of another bioterrorist attack using the mail. But the threat assessment was not completed by the independent contractor until May, 2003 and presented to the USPS until June, 2003, 9 months after the outlines of the BDS were approved by the USPS management. USPS. “Decision Analysis Report: Biohazard Detection Systems.”; USPS. “Update on Biohazard Detection System in Response to OIG Report# DA-MA-02-001.” July 7, 2004.


Large volume mailers argued since the initial consideration of irradiation that their mail was safe (see above). In their view, only single piece, individual, mailings present security risks. USPS accepts this position and organized the BDS accordingly. Yet, while the majority of the initial 2001 anthrax letters were sent via street collection boxes, the Florida mailing was, it appears, sent by parcel (a category that is not currently subject to the BDS). More importantly, while collection or individual mailings are perhaps the most likely path through which anthrax could enter the postal network, it is by no means certain that if another attack was to be plotted that it would directly follow the outlines of the 2001 attack. Along these lines, the USPS OIG criticized the exclusive application of the BDS on collection mail and called for an assessment of all categories of mail. The APWU also critiqued the partiality of the BDS. APWU President William Burrus during Congressional Testimony, noted in forceful terms:

Let me say a word about the present effort to provide detection equipment. [The BDS] will go on specified postal equipment, not all of the equipment… [Most mail] is commingled in postal facilities throughout this country. Over 50 percent of that mail bypasses the Postal Service system … there are over 200 private consolidation plants in existence in this country processing American’s mail. They hire low-wage workers without background checks. It’s very possible for a terrorist to be hired by one of these companies. That mail would never come through … biodetection equipment. It will go directly to the letter carrier, to the bag, to the American customer, to the American citizens.

The APWU echoes the assessment made by the USPS OIG by pointing out, in fact, anthrax could enter the mailstream from bulk mailers. The APWU, rather than seeing collection mail as the only category of mail that represents a potential source of harm, 

138 USPS OIG. “Biohazard Detection System.”
argues that the partial liberalization of postal markets and the attendant growth of consolidation services introduces new avenues for subverting the security of the postal network. In critiquing the partiality of the BDS, both the USPS OIG and APWU point out the limitation of assuming that the particulars of the 2001 dictate how a future attack may unfold. For both groups the notion put forth by both USPS and large volume mailers that the commercial bulk mail is not “at risk” is suspect. Rather, in calling for a more expanded program of detection that considers all categories of mail these groups highlight the political character inscribed within the BDS.

As touched on above in discussing irradiation, as postal stakeholders considered what form or forms new postal security measures would take, large volume mailers sought to avoid any added security burdens. While they supported a new security posture as a means of quelling possible market disruptions, they actively pushed USPS to create an uneven system of security that would not add new costs or handling to their mailings.140 The organization of BDS within the postal network embodies this point of view. By bypassing the BDS, large volume mailers avoided not only the extra scrutiny offered by the system, but also the added costs of security. Presort rates, as noted above, are calculated largely by tabulating cost avoidance: the presumed amount of savings USPS receives on account of the preparation work undertaken by the mailer or, as is often the case, third-party mailing houses that collate the mailings of multiple mailers to achieve the maximum presorting discount. Since presorted mail bypasses the BDS, the extra costs of security are another avoided cost that is deducted from (or not added to)

140 See, Alberta. “USPS Expected to Seek as Much as $10 Billion in Aid From Congress.”; GAO. “Highlights of GAO’s Conference on Options to Enhance Mail Security and Postal Operations.”
their rates. Rather, the costs of the system are borne by small mailers, who do not have the option of opting out of the added costs of security, and the public at large, through homeland security appropriations. While large volume mailers benefit from added levels of security in terms of increased public confidence in the mail, and indeed solicited such new measures based on this logic, they are largely immune from their costs.

In place of involvement with the BDS, large volume mailers have the option of enrolling in a voluntary mail security program. In early 2002, the MSTF Mail Preparation Working Group initially floated the idea of a voluntary security program for business mailers as a way for business mailer to circumvent added security measures.\(^\text{141}\) The *Emergency Preparedness Plan* affirmed the commitment to a dual approach to security, where collection mail was subject to detection and added treatment and large mailers had the option of engaging in a voluntary security program.\(^\text{142}\) Subsequently, the USPIS put into practice first a pilot program, the Trusted Mail Program (TRUMP), and then in October 2003 established as the B.2.2 Security Initiative (named after the subheading under which the principles of the program are described in the Appendix to *Emergency Preparedness Plan*).\(^\text{143}\) The B.2.2 program is completely voluntary and is run by the USPIS. As part of the program, mailers that volunteer to participate are randomly selected by the Inspection Service for audit. The Inspection Service examines different aspects of the mailers operation, including hiring practices, on-site emergency plans,


\(^{142}\) See, for example, USPS. *Emergency Preparedness Plan.* “Appendix.” H-14 -H-17.

access controls, and provides the mailer with a written summary report. The USPIS does not require a follow-up after providing the report and does not compel any change as a result of the review. The scores of the individual mailers are confidential and are not disclosed.  

The decision of where to place the BDS, then, was in these respects imbued with a larger significance. Deciding where to locate the equipment, rested on assumptions concerning which mail is “at risk” and determining who is responsible for the provision of security. The organization of the BDS reflects the wishes and ultimately power of large volume mailers to opt out of new security procedures. While institutional mailers joined labor in supporting new security efforts as a means of assuaging customer concern over fears of anthrax, they limited the scope of these efforts and worked to displace the costs of security onto the public at large (through appropriations) and small and individual mailers. The BDS, in its selective deployment and distribution of costs, creates a tiered system of security within which large volume mailers are free from added burdens. The pairing of the voluntary B.2.2 mailer security initiative with the BDS highlights the cleavages present within postal security. B.2.2, in its voluntary nature and lack of enforcement measures, is a well-meaning but essentially toothless program that provides a symbolic check on the security practices of bulk mailers. Finally, and importantly, the BDS creates a precedent concerning how biological terrorism through the mail would be dealt. In creating the program, mailers were assured that, in the event of a future biohazard incident involving the mail their mailings would not receive

\[144\] Ibid.
additional, extra, handling, but would be considered “safe.” Thus, institutional mailers would avoid any future costs, delays, or scrutiny.

Intelligent Mail: The Belated Information Revolution

The Biohazard Detection System represents only one prong of the post-anthrax postal transformation. In addition to the BDS, postal management adopted a suite of services known as Intelligent Mail (or IM). Plans for Intelligent Mail predate the anthrax crisis and were initially developed as a way for the postal system to adapt to the changing communications landscape characterized by low-cost, networked computing. After the anthrax attacks of 2001, however, Intelligent Mail was repositioned, in part, as a novel security measure to combat the threat of biological terrorism. By the time Intelligent Mail moved beyond development and started to be integrated into postal operations during the mid-2000s, the multiple goals of creating new streams of revenue, cutting costs, and increasing security were firmly in place. Intelligent Mail, at its most basic, is a suite of ICTs that extract, store, and manipulate information about the postal network. The program attempts, at once, to provide new streams of revenue for the postal service through the promotion of value-added services marketed to large volume mailers; help cut costs and streamline postal operations by providing a raft of real-time data that can ferret out inefficient operations and enable better supervision of labor; and provide investigators with forensic data of how mail circulates through the network.

147 IM conjoins three typically discrete domains of surveillance: the workplace, consumptive behavior, and security. Lyon sees the convergence of formerly separate types of surveillance as characteristic of post-9-11
While the BDS is characterized by the absence of participation by large volume mailers and the creation of a bifurcated postal system, Intelligent Mail represents an attempt to actively solicit, court, and cater to large volume mailers. While large volume mailers are able to opt out of the BDS, Intelligent Mail is an attempt to provide them with new services and products. At the same time, Intelligent Mail represents the ongoing attempt by the Postal Service to more tightly survey postal workers and move towards a flexible labor force that limits the role of unionized workers in favor of temporary employees. Intelligent Mail in these ways embodies the familiar hierarchies of contemporary postal politics; through its aims and operation, it reproduces the conjoined triumph of large volume mailers and the marginalization of organized or unionized labor.

*The Invention of Intelligent Mail: Remaking the Postal Service for the Information Age*

Intelligent Mail was first conceived and developed by two different collaborative working groups that brought together postal management and different sectors of the mailing industry to confront the challenges and opportunities that new technologies present to postal service. The “Intelligent Document Task Force” (ID Task Force) first introduced and promoted the idea of intelligent mail during the mid and late-1990s. The ID Task Force was comprised of postal management (from technical, strategic, and legal divisions within USPS) and representatives of postal equipment manufacturers. The ID Task Force was headed by Maynard H. Benjamin, President and CEO of the Envelope Manufacturing Association (an industry trade group), and Patrick McGee, USPS security engagements generally. See, David Lyon. *Surveillance after September 11*. Malden, MA: Polity, 2003.
Manager, Retail Marketing. Pitney Bowes (a postal equipment manufacturer), Mead/Westvaco (a packaging company), International Paper Company, Neopost (a postal equipment manufacturer), DuPont, and E-Stamp (a company then involved in PC postage) rounded out the Task Force. The ID Task force worked from 1996 until 2004 exploring new technologies that could be integrated into the postal network to adapt and thrive within a changing communications environment. The Task Force set out, as its initial report made clear, to examine how paper communication could avoid becoming obsolete in the age of the internet. The ID Task Force examined products and technologies in development at Sandia National Laboratory, MIT Media Lab, IBM, Hewlett-Packard, Motorola, Kodak, Xerox, and a number of other sites working on new communication technologies that could have application to the postal environment. The ID Task Force argued that the future of the postal system rested on the creation of new, “intelligent documents” that move beyond traditional mail. The ID Task Force remarked that:

An intelligent document, or intelligent mailpiece, is merely a document that does more than convey information to a human reader. It may be a document or mailpiece that initiates an electronic transaction, confirms delivery, tracks itself within a delivery system or provides information on its own disposition. Intelligent documents or mailpieces are therefore both “human readable” and machine readable.

---

149 Ibid. 35.
151 Intelligent Document Task Force. The Internet of Things.
At its most basic, the ID Task Force sought to devise and promote ways of encoding greater quantities of information on mailpieces that can be read by different forms of electronic equipment. The ID Task Force sketched a future postal service that could find new forms of information-rich barcodes, radio-frequency identification (RFID), organic semiconductors, or image imbedding as a part of a mailpiece. Such technologies could, in the estimation of the ID Task Force, provide new “on-document intelligence” that is machine readable and could be of value to mailers or postal management through improved logistics management and narrowly tailored services. The ID Task Force advocated that USPS confront possible declines in mailing brought about through electronic diversion by repositioning how it uses information. The ID Task Force argued that USPS should consider information as a “strategic asset” and move away from a volume dependent business model to one centered on new value-added services structured around the commodification of information.

The ideas and novel approach introduced by the ID Task Force were taken up and filtered into a more clearly defined program by the Mailing Industry Task Force. In 2001 Chief Executives of 11 different companies drawn from different corners of the postal industry and the USPS Deputy Postmaster General created a Mailing Industry Task Force (MITF) to find ways to integrate mail into the changing communications landscape. In their words, they sought to find out if “mail remains a viable communications channel in

\[154\] Ibid. 34.
an era of profound technological change.”¹⁵⁶ The MITF was chaired by Pitney Bowes Chairman and CEO, Michael J. Critelli and USPS Deputy Postmaster General. The Steering Committee drew members from across the postal industry, including: large volume mailers such as American Express and Capital One; marketing companies, including ADVO, Acxiom, Young & Rubicom; catalogue and greeting card manufactures; and communication companies, such as Siebel Systems, Symbol Technologies, and R.R. Donnelley.¹⁵⁷ The MITF took the basic concept underling the ID Task Force’s push for an “intelligent document”—the encoding of new forms of machine-readable information within mailpieces to create new products, better logistics management by both mailers and the postal service—and strongly advocated the pursuit by USPS of an “Intelligent Mail” platform. The MITF adopts the ID Task Force’s definition of an “intelligent document” and rechristens it as “Intelligent Mail.” The MITF describes Intelligent Mail as “the use of data-rich, machine-readable bar codes to make each mailing piece unique” that would “allow the mailing industry to compete by including data that ‘lives’ with the mailpiece or package.”¹⁵⁸ The MITF scaled back some of the more outré ideas pursued by the ID Task Force, such as placing organic semiconductors on pieces of mail, and focused on the use of new information-rich barcodes as a means to collect, store, and use new forms of information drawn from the mailstream. The MITF positions IM as a tool primarily for commercial mailers. In their estimation, IM would provide commercial mailers with greater control over their

¹⁵⁷ Ibid. 2.
¹⁵⁸ Ibid. 19.
mailings, allowing them to time and target marketing campaigns with better precision, and provide continuous feedback on the effectiveness of their mailings through a closed-loop system (where information about the mailings status would be returned to the sender). The MITF also envisions the new data collected from IM as a tool to enable better use and oversight of USPS resources. By collecting better data about how the mail moves from point-to-point through the postal network and offering better forecasting models of the required resources, the MITF saw that IM provides USPS with the opportunity to better analyze and manage its logistic operations.

*Intelligent Mail as Security Technology*

The MITF released its first set of recommendations on October 15, 2001, just as the anthrax crisis was unfolding. IM quickly was repositioned as not only a way of creating new forms of revenue and cutting costs, but also as a new security tool. The ID Task Force scrambled to consider the possible security applications of their work. In their first report after the anthrax attack, the ID Task Force concluded that the technologies under review could provide an added layer of security. New barcodes could identify the sender and receiver of pieces of mail, as well as provide postal and law enforcement officials with the ability to track and trace postal flows through the network. Such innovations would limit the anonymity of the mail and provide forensic information concerning when and how an infected mailpiece moved through the particular nodes of

---

159 Ibid. 19-20.
160 Ibid. 19.
161 Ibid.
the postal network and which additional letters were, as a result, at risk of cross-contamination. These security applications were, to the ID Task Force, a “logical extension” of these new information and communication technologies.\(^{163}\) The MITF likewise noted the same security features of IM in its first update on its activities relating to Intelligent Mail.\(^{164}\) The Mail Security Task Force follows the lead IM advocates and ultimately suggests the adoption of IM as a security tool. The March 2002 *Emergency Preparedness Plan* notes that although IM predates the anthrax crisis, its adoption could help deter and prevent further attacks.\(^{165}\) The *Emergency Preparedness Plan* followed the findings of the ID Task Force and argues that by encoding information concerning sender and receiver of the mail, and creating data charting how individual mailpieces move through the network, IM creates greater accountability of mailers and assists post-attack response (such as quick re-routing of mail) and investigation (through data mining).\(^{166}\)

One month after the submitting the *Emergency Preparedness Plan* to Congress, USPS released its transformation plan, a major strategy document covering all aspects of the postal business that mapped out the path the Postal Service would pursue during the decades ahead. The *Transformation Plan* focused on Intelligent Mail as one of the centerpieces of its new strategy.\(^{167}\) Shortly thereafter, USPS started to adopt and implement Intelligent Mail. In 2003, USPS Postmaster General John Potter created a new position

---


\(^{165}\) USPS. *Emergency Preparedness Plan.* ES-5, 3-2, 5-1, 6-1.

\(^{166}\) Ibid. 3-2, 5-1, 6-1.

within the organization, Senior Vice President Intelligent Mail and Address Quality, and created an Intelligent Mail corporate plan defining IM strategy. The Postal Service’s Intelligent Mail plan adopted many of the findings of the ID Task Force and MITF and called for the eventual unique identification of all pieces of mail and end-to-end (from mailer to recipient) track-and-trace capabilities for all mailings. This strategy, which the Postal Service, christened “total mail visibility,” would support the introduction of new informational products, help cut costs, and increase the security of the postal network. Operationally, the Postal Service’s IM plan called for the use of new, information-rich, barcodes that could store greater quantities of information on mailpieces, new scanning equipment to record the circulation of mail through the postal network, and computer systems that could store and manipulate large quantities of collected data. IM first started to be used on a limited basis in 2006 and expanded into full service in 2009.

Intelligent Mail: Operational Characteristics and Aims

Operationally, IM rests on the adoption of new barcodes, scanning equipment, and computer networks to manage the raft of new collected data. The USPS corporate plan defines IM as “the placement of information-rich code on all mail, aggregates of mail [such as palates and trays], and business forms, enabling end-to-end visibility into the mailstream.” The system rests on the continual active and passive scanning of new IM barcodes placed on both individual pieces of mail and aggregates, such as sacks and

---

170 GAO. “U.S. Postal Service: Intelligent Mail Benefits May Not be Achieved if Key Risks are Not Addressed.” GAO-09-599. May 2009. 28.
containers containing collated mailings, as mail moves through the postal network to generate real-time information. Barcodes have been used for decades in postal processing. However, the IM barcode differs from previous postal barcodes in that it has the capacity to include far greater quantities of information and replaces the need for multiple barcodes—before IM over 30 different USPS barcodes were in use for different services and a single piece of mail would often have multiple different barcodes printed on its exterior.\textsuperscript{172} IM barcodes are what are known as 4-state barcodes that contain 4 differently situated bars: a full bar, ascender, descender, and tracker (see Figure 5-1 and Figure 5-2).\textsuperscript{173} The barcode is 65 bars long and contains 31-digits of mailpiece data (a string of 31 numbers) within the code. The IM barcode includes multiple forms of data in a single code, including: a mailer ID, a unique number identifying the specific mailer; barcode ID, denoting which type of presort discount the mail is eligible; service type ID, defining what value-added service the piece is requesting; a serial number, identifying the unique piece of mail; and a routing code, providing information about where the piece is to be delivered.\textsuperscript{174} The matching of a unique mailer identification number and a 6-digit serial number allows each mailer to track up to 1 billion unique mailings at any single moment.\textsuperscript{175} The IM barcode is created on the mailpiece by the mailer and condenses the multiple barcodes that previously specified type of discount, change of address information, specialty service, and other forms of data into a single code. As mailpieces

\textsuperscript{172} GAO. “U.S. Postal Service: Intelligent Mail Benefits May Not be Achieved if Key Risks are Not Addressed. 2.
\textsuperscript{174} Ibid. 4.
\textsuperscript{175} Previous postal barcodes, for example, could only track 1 million unique pieces from a single mailer at one time. See, Melissa Campanelli. “USPS Touts Four-State Barcode, OneCode Products.” \textit{DM News}. April 7, 2006.
and aggregates are introduced into and then moved throughout the network, a series of scans read the barcodes and collect data in real-time. Automated processing equipment performs passive scans continually during operation, while postal workers (including carriers and others) use hand-held scanners, known as Intelligent Mail Devices (or IM Devices), to perform active scans throughout the system. All data from both passive and active scans are then collected and stored through the Postal Services information-backbone service, Intelligent Mail Visibility Service. The collected information is then made available to customers and management through networked computers.

**IM and New Revenue Streams**

As a revenue generating tool, IM seeks to expand the range of products that USPS markets to corporate mailers and counter the loss of revenue linked to recent declines in volume growth. The Postal Service described some of the different value-added services that it intends to build from IM in its corporate plan. The Postal Service envisions IM as providing mailers with a host of different services and opportunities, including:

- confirmation and proof of final delivery;
- start the clock information, which alerts mailers when their mail is first accepted by USPS;
- track-and-trace services that enable mailers to time delivery of advertisements and marketing programs;
- return services, that alert mailers to incoming mail as soon as it is introduced into the network (a form of advanced notification);
- and address correction and change of address information. The Postal Service tailored and marketed these offerings—some now in effect, others only

---

177 Ibid. 29.
planned—to commercial mailers as a way of capturing a larger share of the ad market.\textsuperscript{178}

In explaining the new benefits of the services enabled by IM, USPS describes how better time of delivery information will allow marketers to engage in targeted messaging from multiple forms of media (in other words, to better coordinate media buys and ad spending) and link mailings with sales and promotions.\textsuperscript{179} The Postal Service also envisions IM as enabling mailers to better manage their logistics; by providing incoming data (advanced notification of when returning bills or notices were sent), companies can better forecast staffing needs, limit unnecessary mailings of reminders and second notices (while a payment or response is already in process). Finally, by providing a more detailed picture of how customers respond to particular ads and marketing messages, USPS argues that IM will help mailers more effectively customize ad campaigns and project cash flow and sales responses.\textsuperscript{180}

With IM, the Postal Service is responding to the unique economics of postal service and reality of stagnant and declining mail volume (see Chapters 2 and 3). Postal service is characterized by rising fixed costs. These costs currently account for 44\% of total postal costs and the labor associated with the delivery of the mail, last-mile costs, makes up nearly half of all fixed costs.\textsuperscript{181} Each year the postal service adds roughly 1.7 million new delivery points (addresses) and, as a result, postal costs rise annually

\textsuperscript{178} This point is taken for granted within the USPS’s “Intelligent Mail Vision” and reinforced through the description product offerings that are, without exception, described in terms of their value to large, commercial, mailers. See, USPS. Ibid. 13.

\textsuperscript{179} Ibid. 34.

\textsuperscript{180} Ibid. 13.

independent of volume fluctuations. For decades, the explosion in advertising and
commercial mail—added by discounted postage for bulk mailers—helped defray these
escalating costs and buffer the effects of electronic diversion. However, by the late
1990s and early 2000s, postal volume started to level off and actually decline due, at least
in part, to the growth of electronic channels of communication. Postal management, as
a result, faced the grim prospect of perhaps permanent rising deficits and revenue
shortfalls. It was in this context that IM as a value-added service was conceived. As
the ID Task Force phrased it, IM was designed to help transform the Postal Service “from
an older volume driven model of business activity to a more contemporary value driven
business model.” While value-added services currently contribute a fraction of total
postal revenue, Postal management sees this segment, unlike the traditional mail market,
as an area of growth that can counteract volume declines.

IM and the Control of Labor

---

182 USPS. United States Postal Service Transformation Plan. 5, “Appendix F.”
183 USPS. United States Postal Service Transformation Plan. L-6; GAO. “U.S. Postal Service: A Primer on
Postal Worksharing.” 25.
184 Postal volume grew at an average annual rate of 5.6% during the 1980s and declined to 2.5% in the
1990s. In 2001 total mail volume declined. USPS. United States Postal Service Transformation Plan. 1.
185 For a detailed description of the precarious financial position of the Postal Service during this period,
see CRS. “Postal Service Financial Problems and Stakeholder Proposals.” GAO. “Major Management
Challenges and Program Risks: U.S. Postal Service.”
Impacting the Future of Mail. ii.
187 From 2000 to 2007, what the USPS defines as “Special Services” including insured, registered, and
certified mail accounted from between 2.9 and 3.6% of total Postal Service Revenue (the lowest figure in
terms of revenue and percentage contribution occurred in 2000 and the highest revenue and percentage
contribution occurred in 2007). If Priority and Express Mail receipts are included as value-added services,
the percentage of total revenue contribution for the same dates expands to between 11.1 and 12.1% of total
revenue. No matter how calculated, these specialty services contribute a fraction of total postal revenue.
USPS. United States Postal Service Transformation Plan. “Appendix K”; USPS. Strategic Transformation
As a control technology, USPS aims to use IM as a means of managing labor outside of the office and support the increased use of temporary, non-union, workers.\textsuperscript{188} In these ways, IM seeks to not only support the creation of new streams of revenue, but also reduce costs. “Total Mail Visibility” not only enables mailers to peek inside the postal network and track the circulation of their mailings; it also allows postal management to extend surveillance of postal workers and finely tailor how labor is deployed. Beginning in the early 1990s, postal automation transformed the responsibilities of letter carriers and, in the process, created new management challenges. Starting in 1992, USPS implemented new automated letter sorting systems that could process letters into what is called “delivery point sequence” (DPS). DPS mail eliminates some of the in-office work of letter carriers by sorting mailings into the exact order in which it will be delivered.\textsuperscript{189} Previously, letter carriers spent a significant portion of their workday sorting mail destined for their particular route into the delivery order (this is called “casing” the mail). During the mid-1980s, the typical 8-hour work day of a letter carrier was divided equally between 4-hours in office casing the mail and 4-hour “street time” delivering the mail to addresses.\textsuperscript{190} The introduction of DPS in 1992 and subsequent program expansion reduces the amount of time carriers spend in office, while proportionally increasing their out-of-office street time. By 2005, the division of the

\textsuperscript{188} IM is a type of “location technology” that is designed to create a digital record of movements that are typical beyond direct supervision. See, Lyon. \textit{Surveillance Studies: An Overview}. 43.


\textsuperscript{190} GAO. “U.S. Postal Service: Progress Made in Implementing Automated Letter Sequencing, but Some Issues Remain.” 5.
average carrier’s day shifted to 2-3 hours in office, and 5-6 hours street time. The ongoing fine-tuning and extension of postal automation is expected to lead to the even greater shift in the balance between office and street time. Eventually, USPS hopes to reduce in office time to .5 to 1 hour with a street time of either 7 or 7.5 hours. The elimination of manual casing supports greater street time and, as a result, longer routes covering a greater number of addresses. Longer routes, in turn, support cost reductions by decreasing the number of carriers needed to shoulder mail delivery.

Yet, increased street time carries with it managerial challenges. The benefits of automated DPS can only be realized if carriers maintain a high pace of work on the street and do not shirk their duties. The independence afforded out-of-office workers creates just such opportunities. To prevent the loss of efficiency gains, USPS has turned to ICTs such as IM as a way of creating new forms of data about carrier performance and enhancing control. IM offers managers a tool to survey behavior on the street and measure the performance of carriers. As a part of the IM program, carriers are equipped with an IM Device, a handheld scanner used to record information about the mail (such as the delivery of a registered letter) and about the carrier. As carriers move through their

---

194 USPS forecasts that the reduction of in-office time from the high-mark of 4 hours during the mid-1980s, to between one and a half hour would allow a single carrier to deliver mail to nearly twice as many addresses. In 1988, the average route contained 470 delivery points. The USPS forecasts that a carrier with a 7 to 7.5 hour street time could manage a route containing between 750-800 delivery points. GAO. “U.S. Postal Service: Progress Made in Implementing Automated Letter Sequencing, but Some Issues Remain.” 5; USPS. “USPS Corporate Flats Strategy.”
195 See, GAO. “U.S. Postal Service: Progress Made in Implementing Automated Letter Sequencing, but Some Issues Remain.”
day, they use IM Device to scan what are known as “managed service points,” a series of barcodes arranged in sequential order throughout their routes on specific locations (such as a mailbox), and barcodes affixed on value-added, deliverables (insured packages, registered letters, pieces receiving delivery confirmation). These active scans create a digital record of the carrier’s behavior on the street that are used to measure performance, adjust workload, and provide general supervision. As a control technology, IM provides operational data that extends the view of management outward, beyond the office, in a routinized and detailed form. IM attempts to ensure that carriers are not accruing “under time,” non-productive hours on the street that are hidden behind the cloak of distance. IM, then, attempts to ensure cost savings by providing a check against the unintended consequences of automation—increased worker autonomy—and ensures that the gains of automation are realized.

The use of IM as a control technology is also tied to the use of temporary, casual, labor. Since the mid-1980s, the Postal Service has increasingly relied on temporary workers as a means of controlling costs. Non-union, non-career, temporary workers provide the Postal Service with a cheap alternative to career workers. Non-career employees work for lower wages and, importantly, in most cases do not qualify for the benefits afforded career employees. Labor costs account for the overwhelming majority of postal costs—78.4% in 2001—and a significant (and growing) portion of total costs are allocated to benefits (see Table 5-3).

---

197 Ibid.
198 GAO. “U.S. Postal Service: Mail Delivery Efficiency Has Improved, but Additional Actions Needed to Achieve Further Gains.” 33-35.
In an effort to check the growth of these costs, USPS has attempted to replace career positions with temporary workers. Since the mid-1980s, the growth of casual workers has outpaced the growth of career employees. From 1987 to 2007, total career employees declined by 8.8%, while non-career employees increased by 106.5%. As a share of total employees, non-career employees grew from 6.1% of total employees in 1987 to 12.9% in 2001.199 IM supports the use of non-career workers in two related ways. First, as is the case with career carriers, IM provides a record of performance outside the office. This is even more important for temporary workers that are employed for short periods of time and do not have established records of performance.200 The adoption of casual employees carries a certain amount of uncertainty concerning their performance (there is not a record from which to base assessments, and they are not wed to the institution). IM provides management with the ability to quickly assess individual behavior and performance, without the aid of long-term, annual, measures. The added supervisory data procured through IM provides the means to ensure reliability despite the churn associated with temporary employees. Second, IM is designed to allow better forecasting of volume and workload fluctuations.201 Some of the benefits of temporary labor, such as short-term appointments without guarantees, match well with fluctuating workload. By providing better data concerning the ebb and flow of postal operations, IM allows management to match temporary workers with temporary spikes in workload.

200 For example, casuals and rural substitute carriers comprise the largest portion of USPS non-career workers and can only be employed for limited periods of time. CRS. “U.S. Postal Service Workforce Size and Employment Categories, 1987-2007.”
201 USPS. “Intelligent Mail Vision.” 9-12, 32-33.
This type of fine tailoring allows for the maximization of the value of non-career, temporary, employees.

**IM and Security**

As a security tool, the value in IM also rests in its creation of a data record concerning how the mail moves through the postal network. The Postal Service’s *Emergency Preparedness Plan* describes IM as a preventative technology that also provides law enforcement with forensic information about the postal system. The security applications of IM, as outlined by USPS, are two-fold: as a preventive measure, real-time data about the processing and distribution of mailings can be used to isolate mail that may have been cross-contaminated and re-route other mailings around troubled network nodes; additionally, the track-and-trace data collected through IM can help investigators attempt to recreate how a certain suspect letter moved through the network. The collection of real-time data about the postal network supports the ability to engage in adaptive re-routing during a crisis to limit system-wide disruptions. By identifying the mail that is possibly at risk for cross-contamination, IM helps prevent a localized security incident from becoming a cascading failure that spreads across the postal system. IM, in this regard, complements the BDS and aims to maintain the efficiencies of centralized processing while checking the possibility of cascading failure. While the BDS detects the presence of anthrax within P&DCs, IM supports subsequent adaptive re-routing and isolation of cross-contaminated mail. Additionally, IM reduces the anonymity of the mail somewhat by identifying mailers with individual pieces of

---

202 USPS. *Emergency Preparedness Plan.* ES-5, 3-2, 5-1, 6-1.
mail, and provides better forensic data about how certain pieces of the mail were entered into the mailstream.\textsuperscript{203} In the estimation of USPS, such data provides investigators with useful information that can help them piece together how an attack unfolded. In both instances, the security components of IM are inseparable from routine operation; the security features are not additive, but rather are part of common operating profile of IM. The only real distinction between IM’s security features and its role as a source of new revenue or operational efficiencies rests on how collected data are used. In all cases, the collection of data occurs in the same fashion; what differentiates the different aspects of IM is the later purpose to which the information is then applied.

The positioning of IM as a security technology, to be clear, was something of an afterthought. Members of the postal industry and USPS developed IM for quite different purposes in response to issues unrelated to the specific threat of biological terrorism or even postal security. Yet, as IM was being reviewed and considered for adoption, the threat of biological terrorism was used by both industry and the Postal Service as a justification for pursing the project. In the context of the debates about postal security, IM emerged as a possible new “fix.”\textsuperscript{204} Equally significant, however, is that the logic of IM supports multiple divergent aims. The same qualities that make the collection of real-

\textsuperscript{203} Ibid.
time postal data attractive as a commodity and as an operational tool—increased customization, visibility, and control—also make it a useful security instrument.

**IM and the Politics of Control**

Intelligent Mail, like the Biohazard Detection System, is inscribed with a certain politics: it is not simply a value-free tool. IM supports and is central to the reinvention of the postal system as a flexible, informational, enterprise. With IM, the Postal Service attempts to commodify information about the postal system in an effort to cater to the presumed needs of large volume mailers. IM, since its inception, is pitched toward capturing a greater share of ad dollars and providing institutional mailers with value-added services. In this way, the innovations associated with IM represent the prioritization of large mailers over small-scale and individual mailers. As mail delivery becomes a contestable market with exit opportunities, the system is consigned to reverse cross-subsidies and move to cater to large volume mailers to keep them in the system.

The reinvention of the postal service as an informational enterprise also includes a revision of the role of labor. IM provides a new means of managing and controlling the autonomy of workers out of the office and supports the ongoing adoption of casual, temporary, workers at the expense of higher-salaried workers with a portfolio of benefits. As a control technology, IM is central to the Postal Service’s strategy of remaking the role of labor within the organization.

**Conclusion: Reflexive Modernization and the Limits of Risk Politics**

The 2001 anthrax attacks illustrate the power of network failures to destabilize the gridlock of politics as usual and the limitations of these moments of fissure. The attacks
elevated the sunk politics of postal service, briefly, above ground. The reorganization of the postal network into a tightly-coupled, dense network reliant on centralized automated sorting and processing deliver operational efficiencies that, under normal operating conditions, help move massive quantities of mail in a timely and reasonably cost-efficient manner. Yet, this network realignment also, unwittingly, serves to make the network more vulnerable and amplify the effects of the anthrax attacks. The dense nodes of the automated postal network expand the scope of both actual and possible cross-contamination outward and transformed what might have otherwise been a localized set of disruptions into a system-wide, cascading, failure. The successes of reorganization were turned on their head by the anthrax attacks as the obscured systemic vulnerabilities built within the architecture of the system now became difficult to ignore public problems. As we will see in Chapter 6 and Chapter 7, post-9/11, the successes of deregulation reemerge as problems.

The attacks presented a reflexive moment that led to the partial democratization of postal decision making. Following Beck, the attacks helped not only cast attention onto the overlooked political dimensions at play within structure of postal service, but also opened postal decision making to actors typically marginalized. The newly created Mail Security Task Force departed with traditional modes of postal governance and brought together in a deliberative setting labor, large users, management, and relevant experts to discuss how to confront the cross-cutting challenges of postal security. In this sense, the anthrax attacks were a catalyst for a new mode of inclusive postal politics. The attacks, however, only created an opportune moment and were not singularly deterministic; the
attacks, absent outside intervention, compelled very little. Postal labor and large volume mailers seized on the high-profile nature of the attacks and acted as risk entrepreneurs pushing postal management to more seriously confront the challenges of cross-contamination. The anthrax attacks created, if only for a fleeting moment, political capital that these groups adopted and marshaled to different ends.

Yet, while the MSTF represents a break from continuity, the new technological systems of control situated within the postal system to confront the challenges of biological terrorism embody the familiar hierarchies of postal politics. The MSTF turned to ICTs as a means of grafting security into the workings of the postal network. New technologies of control seek to maintain the efficiencies of the network, while limiting the possible systemic disruptions of another anthrax attack. New control technologies are now a regular feature of postal network. These ICTs, however, are not apolitical or neutral, but instead are imbued with a system of values. The BDS creates a bifurcated postal system that allows large volume mailers to opt out of new security measures, while the general public and small mailers shoulder the new imposed burdens of security. The BDS collects new forms of data about the mailstream in an effort to sift the harmful letters from the otherwise regular flow of mail. The system relies on the presumption of a distinction between safe and unsafe mail that privileges large volume mailers. In a different fashion, IM creates new forms of information about the postal network, mailers, and crucially, labor. IM ties together the new emphasis and prioritization of security with a focus on creating new value-added services for large volume mailers and the expanded control of workers. The politics inscribed within these control technologies are
consistent with the characteristic constellations of power found in the contemporary postal system. Recall, reorganization not only transformed the material processing network, but also elevated large volume mailers into the central position of postal power. This favored position finds expression within the assemblage of technologies designed to counter the threat of biological terrorism.

Ultimately, the catalytic potential of the anthrax attacks is qualified. The salient risk of biological terrorism led to the democratization of formerly closed processes of governance, but the eventual bureaucratization of risk into routine security practices embody the familiar hierarchies of postal politics. Access does not always equate to a leveling of power, nor does it promise a particular set of outcomes. The melding of new practices of security within the postal service illustrates the disjunction between inclusive processes of deliberation and their outcome.

Turning toward examinations of post-9/11 security interventions in rail and electric power, we see different outcomes. In these cases, the perceived threat of terrorism not only draws attention to the overlooked political dimensions encoded within the structure of technological systems, but is successfully leveraged to create durable opportunities for public comment and engagement. In these cases, as we will see, the bureaucratization of risks unfolds in a different fashion; rather than reinscribing familiar hierarchies of power, a significant reordering takes place.
<table>
<thead>
<tr>
<th>Facilities that Processed Initial Anthrax Mail</th>
<th>Letter(s) Processed</th>
<th>Cross-Contaminated Downstream U.S. Postal Facilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trenton/Hamilton P&amp;DC</td>
<td>2 NYC Media Letters, Daschle and Leahy Letters, AMI Letter (Unrecovered, Suspected) N=5 (Uncertain)</td>
<td>• Indianapolis Critical Parts Center &amp; Repair Facility&lt;sup&gt;207&lt;/sup&gt;  • Jackson Main Post Office  • Princeton Main Post Office  • Princeton-Palmer Square Station  • Rocky Hill Post Office  • South Jersey P&amp;DC  • Southern Connecticut P&amp;DC (Wallingford)  • Trenton Station E</td>
</tr>
<tr>
<td>Brentwood P&amp;DC</td>
<td>Daschle and Leahy Letters N=2</td>
<td>• Dulles Post Office  • Friendship Station  • Indianapolis Critical Parts Center &amp; Repair Facility  • Kansas City Stamp Fulfillment Services  • Pentagon Station  • Raleigh P&amp;DC  • Southwest Station N= 8</td>
</tr>
<tr>
<td>Morgan P&amp;DC</td>
<td>NYC Media Letters N=2</td>
<td>None</td>
</tr>
<tr>
<td>West Palm Beach P&amp;DC</td>
<td>AMI Letter N=1 (Uncertain)</td>
<td>• Blue Lake Post Office  • Boca Raton Main Post Office  • Greenacres Post Office  • Lake Worth Post Office  • Lucerne Post Office N= 5</td>
</tr>
<tr>
<td><strong>Total Facilities that Processed Anthrax Mailing:</strong> N=4</td>
<td><strong>N=5</strong></td>
<td><strong>Total Cross-Contaminated Facilities:</strong> N=19</td>
</tr>
</tbody>
</table>

---


<sup>206</sup> Not all of the initial letters were recovered. Based on epidemiological and environmental evidence, it appears that at least 5 letters treated with anthrax were sent through the mail.

<sup>207</sup> The Indianapolis Critical Parts Center & Repair Facility received cross-contaminated parts from both the Trenton/Hamilton P&DC and the Brentwood P&DC.
Chart 5-1: Simplified Model of Mail Processing and Distribution Network\textsuperscript{208}

### Table 5-2: Characteristics of Five Processing and Distribution Centers Involved in 2001 Anthrax Attacks

<table>
<thead>
<tr>
<th>Facility</th>
<th>Facility Size (in square feet)</th>
<th>Mail volume/capacity per day</th>
<th>Number of employees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trenton</td>
<td>282,000</td>
<td>4 million</td>
<td>960</td>
</tr>
<tr>
<td>Brentwood</td>
<td>684,000</td>
<td>5 million</td>
<td>2,490</td>
</tr>
<tr>
<td>Morgan</td>
<td>2.1 million</td>
<td>12.5 million</td>
<td>5,000</td>
</tr>
<tr>
<td>West Palm Beach</td>
<td>185,000</td>
<td>6 million</td>
<td>930</td>
</tr>
<tr>
<td>Wallingford</td>
<td>350,000</td>
<td>3 million</td>
<td>1,120</td>
</tr>
</tbody>
</table>

Chart 5-2: Cross-Contamination of U.S. Postal Facilities

---

Figure 5-1: Features of Intelligent Mail Barcode

---

Figure 5-2: Sample Intelligent Mail Barcode

---

### Table 5-3: Labor Costs as Percent of Total USPS Costs

<table>
<thead>
<tr>
<th>Year</th>
<th>Benefits</th>
<th>Wages</th>
<th>Total Labor Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>8.8%</td>
<td>76.1%</td>
<td>84.9%</td>
</tr>
<tr>
<td>1987</td>
<td>15.8%</td>
<td>67.4%</td>
<td>83.2%</td>
</tr>
<tr>
<td>2001</td>
<td>22.3%</td>
<td>56.1%</td>
<td>78.4%</td>
</tr>
</tbody>
</table>

---

Chapter 6: Green Security? The Environmental Movement, the Transportation of Hazardous Materials, and the War on Terror

Introduction: A Crisis of Control

On September 13, 2005, 50 protesters lay silently in front of the U.S. Capitol as a man dressed in Hazmat suit surveyed the field of mock fatalities. In the background, a 40-by-12 replica railroad tank car emitted smoke and a persistent hiss. Painted on the outside of the car, and reproduced on shirts worn by the protesters, was the slogan “Reroute and Phase-Out Chemicals of Mass Destruction” (see Image 6-1). Greenpeace organized the protest at the nation’s Capital to press the federal government to enact stringent security regulations governing the transportation of large quantities of hazardous materials. The tableau was intended to represent the aftermath of a terrorist attack targeting a 90-ton rail car filled with dangerous chemicals. Through this act of theater, Greenpeace sought to highlight the failure of the federal government to enact new legislation or regulations protecting the transportation of hazardous materials from terrorist attack. As Greenpeace Legislative Director Rich Hind noted, the simulation at the Capital was designed to force Congress to imagine the effects of a real chemical disaster.\(^1\) In an odd twist of post-9/11 politics, environmentalists argued that the Bush Administration was too lax in its engagement in the “war on terror”; in their estimation, the government was largely ignoring the threat of terrorism.

---

Post-9/11, the rail industry faced a crisis of control. Rail shipments of large quantities of hazardous materials passing near or through urban areas offer terrorists an inviting target. Each year, 1.7 million shipments of dangerous materials travel across a domestic rail network that is largely open, unsupervised, and often operating in close proximity to cities. An attack targeting a large shipment could have dramatic results. In widely quoted testimony, Dr. Jay Boris, Naval Research Lab, detailed that an attack on an 90-ton rail car of chlorine could, under the right circumstances, kill 100,000 people in a densely packed urban area in as little as 30 minutes. Rail cargoes of hazardous materials are, in the words of then-Senator Joe Biden, “rolling weapons of mass destruction.”

Federal law, regulations, and industry voluntary practices have historically been preoccupied with rail “safety” and are pitched to prevent and mitigate unintentional accidents. After 9/11, however, security became a new and pressing concern. Calls for the introduction of tighter security measures to confront the threat of terrorism came from a variety of corners; the environmental movement, local and state officials, members of Congress, and the rail industry each advocated new enhanced rail security for hazardous materials. What form such new measures would take, however, was an open question. Though near consensus existed that the pre-9/11 status quo on rail safety needed

---


rethinking and remaking, what direction such reforms should take was deeply, and at
times bitterly, contested.

In the following chapter I examine the debate over the rerouting of rail shipments
of hazardous materials. After the terrorist attacks of 9/11, environmental groups, such as
Greenpeace, the Sierra Club, and Friends of the Earth, advocated strict new security
regulations concerning the routing of hazardous materials shipments. Pointing toward the
threat of terrorism, these groups backed local city councils and state governments, as well
as some prominent federal legislators, in seeking to enact strict security regulations
directing railroads to reroute hazardous cargoes away from urban areas. Environmental
activists seek rerouting as a way of reducing the risks associated with transporting
hazardous cargoes in close proximity to cities, and also as an indirect means of achieving
their long-standing goal of limiting the use of toxic materials, particularly chlorine, in
industrial production. By limiting flows of toxic materials via routing restrictions,
environmental activists seek to make it more difficult to access industrial chemicals and
encourage the adoption of safer substitutes. The rail and chemical industries chafed and
pushed back against post-9/11 calls for routing controls. The chemical industry plainly

---

7 On the issue of rerouting, the rail and chemical industries aligned to reject government intervention. On other key issues, however, it the rail and chemical industries sharply disagreed. Railroads support legislation supporting the adoption of inherently safer technologies and the relaxation of common carrier obligations which would have allowed them to refuse to carry hazmat cargo. On both these issues, the chemical industry actively challenged the position of the railroads. See, Association of American Railroads (AAR). Statement for the Record. United States. House of Representatives, Committee on Homeland Security, Subcommittee on Transportation Security and Infrastructure Protection. Dec. 12, 2007; AAR.
rejects rerouting on the grounds that it adds costs and limits the availability of its products. The railroads’ objections are somewhat more nuanced. Railroads seek to maintain autonomy to determine the material flows of their networks and utilize the most profitable shipping routes. For railroads, hazardous materials are an unwelcome nuisance. Common carrier obligations require railroads to haul hazardous materials, whether they prefer to or not. Though the most dangerous hazardous materials which sit at the center of rerouting debates, toxic-inhalation-hazard (TIH) materials, account for a tiny fraction of total rail cargo, mandatory rerouting threatens rail profitability. TIH materials, like other hazardous cargoes, travel as part of manifest freight trains carrying all sorts of other cargo (that is, they do not travel independently on dedicated trains). Rerouting shipments of hazardous cargo, then, require rerouting of all other commodities included in the assembled train. Barring revision of common carrier obligations,
railroads reject calls for rerouting. In countering calls for mandatory rerouting, the chemical and rail industry promote loosely structured voluntary guidelines.

Contemporary debates about rail hazmat security consider a range of possible remedies and mitigation measures, including improved tank car design, the introduction of positive train control (PTC), the adoption of inherently safer technologies (ISTs), and stricter personnel controls for rail employees. Here, I focus on rerouting because it is the most divisive issue and because the antagonisms at play, between local and federal government, activists and industry, are most sharply defined. Additionally, the contradictions of deregulation documented in Chapters 2 and 3 are most acute within the rerouting debate. During the late 1970s and 1980s, deregulation revitalized freight railroads and pulled the industry from the brink of ruin. In the process, however, deregulation increased the flow of large shipments of hazardous cargo across a shorn and lean rail network. Reorganization supported the development of an increasing complex network which yoked together large shipments of hazardous materials and cities. As in the other surveyed infrastructures, deregulation promotes the development of efficient networks at the expense of increasing large-scale vulnerability. The debate on rerouting directly confronts and struggles with this legacy.

*From Safety to Security: Rail Security and Hazardous Material Regulation in Historical Perspective*

Immediately after 9/11, public officials looked toward the transportation of hazardous materials as a likely target of terrorist attack. Richard Falkenrath, former-

---

Deputy Homeland Security Advisor to President Bush, observed in testimony before Congress that “[o]f all the various remaining civilian vulnerabilities in America today, one stands alone as uniquely deadly, pervasive and susceptible to terrorist attack: toxic-inhalation-hazard industrial chemicals.”

New laws, regulations, and voluntary practices expanded the scope of traditional rail safety measures to include a new emphasis on rail security. The shift from nearly exclusive concern with accidents to a perspective that takes seriously the possibility of intentional disruptions, resulted in the creation of new agency, the Transportation Security Administration (TSA) and the enlargement of the regulatory responsibilities and purview of the Department of Transportation (DOT).

Also, as part of the initial recalibration towards rail security, the Association of American Railroads (AAR), the largest rail industry trade association, introduced new voluntary security measures covering the handling of hazardous materials. Though the new powers assigned to DOT and TSA enabled them to introduce new far-reaching security plans, the first new rail security regulations were minor and provided the railroads with wide-latitude in compliance.

Historically, a mix of federal legislation, regulations, and industry voluntary practices governed rail safety and the transportation of hazardous materials. Safety rules and practices were structured to counter accidents and largely ignored the possibility of intentional disruptions. Voluntary regulation of hazardous materials began in 1907 with the creation of the Bureau of Explosives by the American Railway Association (ARA),

---

an industry trade association. In the following year, at the behest of the ARA, Congress passed legislation providing the Interstate Commerce Commission (ICC) with authority to regulate shipments of hazardous materials. The ICC promptly granted the Bureau of Explosives power to develop and enforce hazardous materials regulations, covering explosives and a wide-range of other hazardous cargoes. The Bureau of Explosives created regulations concerning labeling, handling, container specifications, and performed periodic inspections. In 1966, the ICC’s authority to regulate hazmat transportation was shifted to DOT. Subsequently, the Federal Railroad Safety Act, the Hazardous Materials Transportation Control Act of 1970, and the Hazardous Materials Transportation Act of 1975 consolidated the authority of DOT and provided the modern legislative authority for the regulation of rail safety and hazardous materials transportation. Contained in Title 49 of the Code of Federal Regulations (49 C.F.R), DOT developed complex regulations governing the packing, handling, labeling, and

---


17 OTA. Transportation of Hazardous Materials. 146-147.

 designation of hazardous materials. Further, DOT developed engineering standards for tank cars and hazmat containers, and possessed the authority to levy penalties for non-compliance. Within DOT, the Research and Special Programs Administration (RSPA) took responsibility for developing hazmat regulations, while the Federal Railroad Administration enforced regulations. The Association of American Railroads (AAR), the successor to the ARA, developed voluntary standards and conducted inspections complementing federal standards.

To a large degree, this assemblage of regulations and voluntary practices worked quite well; as AAR highlights, 99.99% of all rail hazmat shipments reach their destination without incident. Since 1980, rail accidents decreased by 75% and accidents involving releases of hazardous materials declined by 91%. Though the adoption of a more narrow definition of what constitutes a reportable hazmat accident in the early 1980s slightly inflates the degree of improvement, routine rail shipments of hazardous materials are generally safe and accidental releases are rare.

In the wake of 9/11, however, the possibility of a terrorist attack on rail shipments of hazardous materials became an area of public concern. Previously, when security was considered at all, concerns were limited to unauthorized access to rail facilities by

---

19 Over 30,000 different materials are identified within federal hazardous materials regulations. OTA. *Transportation of Hazardous Materials.* 147.
21 OTA. *Transportation of Hazardous Materials.* 162.
trespassers and vandals. Authorities fretted over accidents—and the resulting liability—caused by the use of rail facilities for sport by recreational motor-sports enthusiasts (what is known as the “ATV Problem”) and trespassing. Post-9/11, however, concerns arose over the possibility that terrorists would target large shipments of toxic materials, in particular TIH materials such as chlorine, traveling in close proximity to densely populated urban areas. The perceived threat of terrorism introduces “salient uncertainty” into rail operations. Safety measures designed to confront rail accidents, or relatively minor incidents of unauthorized access, are ill-suited to mitigate acts of terrorism. Indeed, in some instances the exigencies of safety and security are antithetical. For example, clear, recognizable, placards identifying hazardous cargoes have been a cornerstone of hazmat transportation safety practices and regulations for decades. By clearly identifying shipments through widely-accepted markings, first responders, rail labors, and others are able to ensure proper handling and quick response to spills and accidental releases. Yet, the same transparency which aids safety precautions can provide malicious actors with a bull’s-eye.

As concern turned away from accidents and toward intentional acts of sabotage, the accumulated safety record of railroads in transporting hazardous materials became, if not moot, somewhat incidental to the debate. If the terrorist threat represents something new—an evolving hazard just now emerging—then historical data is of little guide. Fairly rigorous data on rail accidents and hazmat incidents have been collected for

27 Ibid.
decades. From these data, accident probabilities can be derived with a degree of confidence. However, if a new hazard—terrorism—which previously did not target rail hazmat cargoes is added to the equation, the historical record no longer provides a clear glimpse of the future: there is no guarantee that the future will mimic the past. In this gulf, imagined possibilities replace calculated probabilities: imagined worst-case scenarios become legitimate.\textsuperscript{28} In discussing security, recent accidents involving hazmat cargoes become portentous. Indeed, advocates for increased rail security invoke recent accidents—including, accidents in Minot, ND, Graniteville, SC, and Baltimore, MD—as anecdotes emphasizing the general danger associated with rail shipments of hazmat materials, but also to highlight that the destruction from recent accidents would pale in comparison to a coordinated targeted attack on a shipment in an urban area.\textsuperscript{29}

\textit{Rethinking Rail Security Post-9/11: Voluntary Measures and the Transportation Security Administration}

Initial industry and federal action to address security concerns commenced shortly after 9/11. At the start of the war in Afghanistan on October 7, 2001 all rail shipments of hazardous materials within the U.S. were suspended for 72 hours out of fear of sabotage.\textsuperscript{30} The rail industry, through AAR, quickly formed a security task force to

\textsuperscript{28} On this point, see the discussion of Ulrich Beck in Chapter 4.
review and offer recommendations on hazmat transportation. At the federal level, Senator Jon Corzine introduced the Chemical Security Act of 2001 on October 31, 2001. The Bill sought to address the threat of terrorism to chemical facilities and carriers of large quantities of chemicals. The proposed legislation would have given the EPA authority to mandate tighter security regulation of facilities producing or storing certain chemicals and promote the elimination of transportation of such chemicals. Paul Orum, Director of Working Group on Community-Right-to-Know, singled out the chemical and rail industries for criticism in comments in a November 12, 2001 Washington Post story highlighting gaps in chemical facility and hazmat transportation security. Though the Corzine Bill stalled, the Aviation and Transportation Security Act, enacted November 19, 2001 created TSA within DOT and tasked TSA with overseeing new security procedures across all modes of transportation, including freight rail. TSA initially, however, focused on aviation security. Voluntary efforts by the rail industry did however seek to address the issue. In December, AAR implemented new security controls as part of its newly created “Terrorism Risk Analysis and Security Management Plan.” As part of the new security management plan, AAR created the Railway Alert Network (RAN) to share information between railroads and intelligence agencies relating to credible threats and security issues. DOT issued an advisory notice in February 2002 echoing AAR’s

31 In addition to hazmat security, AAR’s security taskforce examined protection of physical infrastructure, operational security, interaction with military, and communication and information technology security. Ibid. 298.
voluntary measures and encouraged railroads to review their security plans—including background checks, en route and facility security. Though there was a brief spike in interest by industry and the federal government in rail security in the months immediately following 9/11, without specific threats or an attack directed toward rail, attention and action tapered off.

*The Homeland Security Act of 2002 and HM-232*

Rail security was again thrust to fore in 2002 as the FBI issued a public warning and environmental activists pressed the issue. New legislation expanded the authority of DOT to oversee the security of hazmat shipments, augmenting their existing authority to oversee the safety of such shipments, while new regulations imposed minimum new burdens on railroads. In October 2002, the FBI provided the first specific warning to law enforcement, rail companies, and the public indicating that Al Qaeda was specifically targeting U.S. rail shipments of hazardous materials. The FBI alert sourced the warning from information gleamed from captured detainees. Previously, concerns over rail security and terrorism were substantiated only through inference. Now, for the first time, the FBI stated in public that railroads and shipments of hazardous materials in the U.S. were a target of terrorist action. Greenpeace, Friends of the Earth, and the Sierra Club pushed the issue of hazmat transportation security. The use and transportation of hazardous chemicals had been a longstanding issue of concern and activism for

---


environmentalists; environmental groups now invoked the threat of terrorism in calling for greater public controls over toxic cargoes.\textsuperscript{39} Activists appropriated the language of national security in order to legitimate their claims; they acted as “risk entrepreneurs” by seizing the availability and salience of fears of terrorism to forward their argument. Environmental groups supported Sen. Corzine’s introduction of chemical security legislation again in 2002 and pushed for more extensive controls than the voluntary actions taken by AAR. Environmental groups’ legislative agenda rested on two main goals: support for rerouting of toxic shipments away from cities and support for ISTs that would reduce or eliminate the reliance on toxic chemicals in manufacture (and of course in the process eliminate the need for the transportation of such materials).\textsuperscript{40}

Though the Corzine legislation did not advance in either 2001 or 2002, new legislation and modest regulations touching on hazmat transportation security did move forward. The Homeland Security Act of 2002 created the Department of Homeland Security (DHS) and moved TSA under DHS authority. Importantly, the Homeland Security Act also amended portions of Title 49 of the U.S. Code relating to the scope of DOT’s power. The Homeland Security Act enlarged the responsibility of DOT to include rail and hazmat security in addition to safety.\textsuperscript{41} Despite shifting TSA to the newly created DHS, the Act explicitly delineated rail security and hazmat transportation security


\textsuperscript{40} “Hell On Wheels”; Greenpeace. “Re-routing Will Eliminate Catastrophic Risks.”

as areas within DOT’s authority. The Act provided DOT with a mandate to address security and expanded the scope of DOT’s power.

Despite the broad new mandate, the first set of regulations introduced by DOT under new powers granted by the Homeland Security Act continues industry’s preference for lax, mostly-voluntary, measures. The new regulations, known as HM-232, acknowledge the danger represented by rail shipments of hazardous cargo near urban areas, but exhibit great deference to rail industry prerogatives. In the publication of the Final Rule, DOT notes that “the September 11th terrorist atrocities to [sic] indicate the heightened risk of terrorism with which we all now live and the need to reassess and address security vulnerabilities in all areas of our public and private lives.”

Further, DOT recognizes that the threat of terrorism requires adapting hazmat regulations to the unique challenges of intentional disruption. Nonetheless, HM-232 only places loose requirements on rail carriers of select hazardous cargoes to draft and have on file a security plan and increase training for workers. Security plans must be in writing and, at minimum, address en route security, personnel practices (i.e., hiring practices and background checks), and unauthorized access. Crucially, the regulations do not identify any specific measures to be taken, nor require railroads to address the routing of hazardous materials through or near large population centers. While the regulations require the development of some type of security plan, what form such a plan takes, or how it is put into action, was left largely to the discretion of each individual railroad.

---

44 Ibid.
Ultimately, environmental groups were dissatisfied with HM-232. They hoped that under the sign of security, DOT might introduce new strict rules governing the transportation of dangerous materials.\footnote{Greenpeace. Submitted Comments. RSPA. Hazardous Materials: Enhancing Rail Transportation Security for Toxic Inhalation Hazard Materials. RSPA. Docket RSPA-2004-18730. Oct. 18, 2004.} Yet, the first set of regulations largely left substantive decisions concerning rail security in the hands of industry.\footnote{Greenpeace charged that the Bush Administration’s close ties with industry interfered with more substantive regulations. At the time, former CSX Chairman John W. Snow served as treasury secretary. Spencer S. Hsu. “D.C. May Ban Hazardous Shipments; Anti-Terror Proposal Pits Environmentalists Against Rail Officials,” \textit{Washington Post}. Jan. 24, 2004. On the relationship between the rail and chemical industries and regulation, see also Public Citizen. \textit{Homeland Unsecured: The Bush Administration’s Hostility to Regulation and Ties to Industry Leave America Vulnerable}. Oct. 2004.} The enlargement of the scope of DOT’s powers did provide the nascent possibility of revisiting and ultimately revising how toxic materials flow through the rail network. However, at first pass the rail industry and the chemical industry won victories. New regulations allowed the status quo to continue and left the private sector with the ability to decide through which routes toxic materials moved. Yet, taken as a whole, early action on hazmat security is not entirely inconsequential. By adopting legislation giving DOT clear power to act on issues of rail hazmat security, the possibility remains that new more stringent regulations can be introduced at a later date; the power of DOT to act on hazmat security remains, despite the limitations of initial regulations. After the disappointment of the new DOT regulations, environmental activists turned their attention and efforts to a new avenue: the pursuit of city and state rerouting legislation.

\textit{Local Action: Mandatory Rerouting and Preemption}

After the passage of the first set of new DOT regulations, activists turned their attention to city and state governments. If the federal government would not take hazmat
transportation security seriously, perhaps cities and states would? While rerouting legislation stalled in Congress and the first set of new regulations proved a disappointment, environmental groups pursued local and state-wide measures to mandate the rerouting of toxic materials away from urban areas. The aim of this strategy was twofold: the piecemeal pursuit of local controls could, in aggregate, limit the shipment of hazmat cargoes near dense population centers; and the passage of legislation could goad the federal government to enact rerouting legislation and regulations. With the support of Greenpeace, Friends of the Earth, and the Sierra Club, at least 14 different cities and states—including Washington D.C., Boston, Cleveland, Chicago, Buffalo, St. Louis, Memphis, Las Vegas, and the state of California—considered rerouting legislation. 47 Washington, D.C. passed the first local ban in 2005. The subsequent legal battle waged over the D.C. ban served as a test case for the legality of local rerouting laws. Though a temporary injunction from the U.S. Court of Appeals prevented the D.C. law from taking effect, the local strategy did pay dividends. The D.C. ordinance ultimately spurred new federal legislation and regulations that offered significant concessions to the demands of environmental activists.

*The Terrorism Prevention in Hazardous Materials Transportation Emergency Act of 2005*

Washington, D.C. offered an attractive test case for local rerouting restrictions. The city’s role as the nation’s Capital and its symbolic significance made it difficult to argue that it was not an attractive terrorist target; of course, the attacks of 9/11 targeted

---

47 Paolino. “All Aboard.” 158-159.
the city as did the anthrax attacks. Additionally, CSX, one the largest U.S. railroad companies, operates two rail lines which transport hazardous materials through the city: the “I-95 line,” the north-south main line stretching from New England to Florida; and the “B&O” east-west main line, connecting with Chicago and St. Louis. CSX shipments of 90-ton chlorine tank cars run at their closest a mere 4 blocks from the U.S. Capitol, well within the range of exposure from an intentional or accidental release.

The D.C. city council first considered enacting a local ban on rail shipments of certain hazardous materials in 2003. After introducing legislation in 2003 and again 2004, the D.C. City Council agreed to table any action while DOT and TSA formulated comprehensive new security regulations. Federal regulators suggested that HM-232 represented only the first step in crafting a new regulatory framework for hazmat transportation and that additional more strict regulations would be forthcoming. After two years of delusory waiting for comprehensive regulations, the D.C. City Council, led by Councilmember Kathy Patterson, viewed the lack of a thorough-going federal action to be an abdication of responsibility. Councilmember Patterson noted that despite

---

48 Regarding the exceptional status of Washington, D.C. as a terrorist target, City Councilmember Kathy Patterson noted that travelers in and out of Regan National Airport alone face unique restrictions. Only when arriving or departing form Regan, are travelers prevented from leaving there seats for 30 minutes before take-off and landing. Patterson. “Councilmember Patterson’s Statement on Introduction.”
50 Hsu. “D.C. May Ban Hazardous Shipments.”
52 In August, 2004, DOT issued a request seeking comments on how to improve the security of hazmat transportation. Though comments were due by October, DOT did not take any further steps to introduce new regulations. RSPA. “Request for Comments.” Docket RSPA-2004-18730. 50988-50994; Hsu. and Horwitz. “Hazmat Rerouting Decision Delayed.”
“massive danger” of a terrorist attack targeting large shipments of dangerous materials, since 9/11 “[n]o enforceable or regulatory action has been taken by the Federal Government to eliminate this threat.” Councilmember Patterson introduced the Terrorism Prevention in Hazardous Materials Transportation Emergency Act of 2005 on January 28, 2005. On Feb. 1, the D.C. City Council voted 10-1 in support; two weeks later, on Feb. 15, Mayor Anthony Williams signed the Act into law making D.C. the first city to enact a local ban on rail shipment of hazardous materials.

The D.C. Council succinctly states the case for mandatory rerouting in legislative findings accompanying the Act. They note that the unique position of D.C. as a center of American politics and history make it an inviting terrorist target and that “[a] terrorist attack on a large-quantity hazardous material shipment near the United States Capitol … would be expected to cause tens of thousands of deaths and a catastrophic economic impact of $5 billion or more.” Further, the Act underscores the inaction of the federal government to address the hazard: “The federal government has not acted to prevent the terrorist threat resulting from the transportation of dangerous quantities of ultra-hazardous materials near the Capitol.” The deferential regulations contained in HM-232, coupled with the promise by federal regulators to act soon, rang hollow for the D.C. City Council.

53 Patterson. “Councilmember Patterson’s Statement on Introduction.”
56 Ibid.
57 Patterson. “Councilmember Patterson’s Statement on Introduction.”
The Act prohibits shipments of large quantities of explosives, flammable gases, or TIH materials within a designated 2.2 mile radius of the Capital, known as the “Capital Exclusion Zone.” The Act does, however, contain possible exceptions: the Act allows shipments during emergencies; and authorizes the D.C. Department of Transportation to issue permits allowing transport if no alternative route exists or if alternate routes are determined to be cost prohibitive. The Act is plainly directed at CSX: both CSX’s main north-south and east-west lines fall within the Capital Exclusion Zone. The intent behind the Act is to force CSX’s hand to enter into interchange agreements with other railroads to reroute specified cargoes across other non-urban lines and limit the shipment of hazardous materials through the city.

*Legal Challenge: CSX v. Williams*

On Feb. 16, 2005, the day after Mayor Williams signed the Act into law, CSX filed suit in U.S. District Court for the District of Columbia seeking permanent or temporary injunction of enforcement of the Act before its scheduled start date of April 20th. In broad strokes, the case pitted the federal government and the railroads against environmental activists and local governments interested in restricting hazmat transportation. The Department of Justice argued in support of CSX, while the Sierra Club joined Washington, D.C. in arguing in support of the Act before the District Court.

---

60 Specifically, the D.C. Council wanted CSX to enter into an interchange agreement which would have allowed Norfolk Southern Railroad to carry identified hazardous cargoes on its line 50 miles west of the District. Councilmember Statement. Patterson. “Councilmember Patterson’s Statement on Introduction.”
and on appeal in the U.S. Court of Appeals. CSX enjoyed support from the rail and chemical industries; AAR, Norfolk Southern, and the National Industrial Transportation League (an industry trade group), the American Chemistry Council, and others joined CSX as amicus curiae. The case hinged on whether or not local communities could assert their prerogative to enact rerouting legislation. In the words of presiding District Court Judge Emmet G. Sullivan the case involves “the longstanding principles of federalism applied to one of the government’s newest and most important roles—protecting our Nation from threat of terrorism.” The case asks, what, if any, role should local communities have in crafting ordinances to combat terrorism? Specifically, could the District of Columbia “pursuant to its traditional police power… prohibit the rail carriage of certain hazardous materials”? Or, does federal law take precedent and prevent enforcement of the Act? The precise legal question in the case turned on the limits of federal preemption of local law.

In the suit, CSX charges that if allowed to stand the D.C. Act would open the door for copy-cat local legislation, crippling the nation’s rail system from operating and halting interstate shipments of vital materials. CSX argues that the Act: (1) violates the Commerce Clause of the Constitution; (2) is preempted by the Federal Rail Safety Act.

---

64 Ibid. 4-5.
the Hazardous Materials Transportation Act, and the Interstate Commerce Commission Termination Act; (3) and is at odds with the legislative authority given to Washington, D.C. under the Home Rule Act.\textsuperscript{67} The question of preemption is central in both the District Court ruling, and the eventual ruling by U.S. Court of Appeals for the District of Columbia. The Federal Rail Safety Act and the Hazardous Materials Transportation Act each contain language indicating that federal law preempts—that is takes precedent and quiets—local and state legislation under certain conditions. The purpose of preemption, as spelled out under the amended Federal Rail Safety Act, is to ensure that “[l]aws, regulations, and orders related to railroad safety and laws, regulations, and orders related to railroad security shall be nationally uniform to the extent practicable.”\textsuperscript{68} The Homeland Security Act of 2002 explicitly retains the preemption provisions of both the Federal Rail Safety Act and the Hazardous Materials Transportation Act.\textsuperscript{69} The preemption provisions, however, identify circumstances under which local jurisdictions can enact more strict rules and regulations. The Federal Rail Safety Act allows states to regulate rail safety and security “until” the federal government “prescribes a regulation or issues an order covering the subject matter of the State requirement.”\textsuperscript{70} If the federal government has not yet acted in a specific area of rail security or safety, states can act. The preemption provision of the Federal Rail Safety Act further creates a safe harbor for states to act \textit{even if} the federal government has acted, if the state law, regulation, or order:

\begin{footnotesize}
\begin{enumerate}
\item \textsuperscript{67} Moore. “Federalism vs. Terrorism.” 772.
\item \textsuperscript{68} 49 U.S.C. § 20106.
\item \textsuperscript{70} 49 U.S.C. § 20106; See also, CSX Transportation, Inc. v Anthony A. Williams, 2005 U.S. Dist. LEXIS 6569 (D.D.C. 2005). 22.
\end{enumerate}
\end{footnotesize}
“(1) is necessary to eliminate or reduce an essentially local safety or security hazard; (2) is not incompatible with a law, regulation, or order of the United States government; and (3) does not unreasonably burden interstate commerce.” The preemption provisions contained in the Hazardous Materials Transportation Act echo the Federal Rail Safety Act and preempt state law when it is not possible to comply with both state and federal regulation, or if state regulation creates a barrier to compliance with federal regulation.

Thus, in examining the D.C. Act, the Courts look at two basic questions: (1) Has the federal government already taken action on the topic covered within the D.C. Act?; and if so, (2) does the D.C. act fall within allowable safe harbor provisions? If the federal government has not yet acted, then the D.C. rerouting ban is (potentially) on firm legal ground. Additionally, if the federal government has already acted, but the D.C. ban addresses a local hazard in a manner that not at odds with federal regulation nor burdens interstate commerce, the D.C. Act can be saved. The District Court and Court of Appeals answer these questions differently and, as a result, offer two different readings of the case.

**CSX v. Williams: U.S. District Court, District of Columbia**

The District Court denied CSX’s motion seeking a preliminary injunction preventing the enforcement of the D.C. Act and ordered full proceedings considering a permanent injunction to commence after the law took effect. In Judge Emmet G. Sullivan’s opinion, the status of federal activity is paramount. In examining new federal

---

regulatory efforts, the District Court identifies a current gap in federal regulations that leaves open the door for state rerouting. The Court notes that rail hazmat security is an essentially new task that stands outside traditional safety regulation.\textsuperscript{74} For the District Court, the threat of terrorism presents a new hazard and challenge to rail operations. In this regard, regulations directed toward routine operational risks of transporting hazardous materials via rail are distinct from regulations focused on the deliberate targeting of hazardous cargoes by terrorists: safety and security are different domains. Looking at federal efforts to regulate security since 9/11, the Court finds little evidence that such efforts have yet come to fruition.\textsuperscript{75} In the Court’s view HM-232, discussed above, does not qualify as a regulation or order “covering the subject matter of the State requirement.”\textsuperscript{76} Despite CSX and the Department of Justice’s argument to the contrary, the Court does not accept that the requirement to develop what are basically voluntary security plans preempts the District Act.\textsuperscript{77} The Court notes that the text of HM-232 identifies the regulation as merely a partial first step and that comprehensive and authoritative federal hazmat transportation security regulations will be forthcoming.\textsuperscript{78} Here, the Court identifies a gap in federal action that allows states to act; until such further comprehensive and authoritative security regulations covering the security of rail shipments of hazardous cargoes are introduced, states are free to introduce legislation.\textsuperscript{79}

\textsuperscript{74} Ibid. 31-32.
\textsuperscript{75} Ibid. 37.
\textsuperscript{76} The Court notes that DOT estimates that compliance with HM-232 will occupy 50 hours of work, roughly one week of work by a single employee. In the Court’s reading, such a minor effort can hardly be deemed to “cover” the complex issue of hazardous material security. Ibid.
\textsuperscript{77} Ibid.
\textsuperscript{78} Ibid.
\textsuperscript{79} Ibid.
The District Court also examines if the D.C. Act prevents compliance with current federal regulations. Here again the Court considers HM-232. The Court finds no reason why CSX cannot at the same time comply with the general requirement to craft a security plan and abide the D.C. Act’s requirement to reroute hazardous cargo away from the Capital Exclusion Zone. In the Court’s view it is not impossible to do so. CSX and the Department of Justice argue that HM-232 is explicitly designed to provide railroads with flexibility to craft security plans tailored to their unique circumstances. In their view, by requiring rerouting, the D.C. Act hinders this flexibility and, as a result, stands in conflict with existing federal regulations. However, the Court notes that the D.C. Act supports the general safety and security aims that underlie federal regulation. The Court finds it unpersuasive that the general regulation requiring the development of a security plan is at odds with local rerouting by D.C. Rather, the D.C. measure is a gap-filling measure allowable within the federal statutory framework.80

CSX v. Williams: U.S. Court of Appeals

CSX and the Department of Justice appealed the District Court decision and were granted a temporary injunction by the U.S. Court of Appeals for the District of Columbia.81 In reviewing the District Court’s ruling, the Court of Appeals again examines the question of preemption and, specifically, HM-232. The Court considers if HM-232 qualifies as action covering the same subject matter as the D.C. Act. The Court of Appeals notes that “[t]o ‘cover the subject matter,’ HM-232 must ‘substantially

80 Ibid. 39-41.
subsume’ the subject matter of the state law, not merely ‘touch upon’ or ‘relate to’ it.”82

On this point, the Court of Appeals departs from the District Court. The Court of Appeals asserts that the provisions in HM-232 requiring railroad companies to devise security plans which address en route security qualifies as federal action covering the same subject matter. The preemption provision of the Federal Rail Safety Act allows the court only to determine if the regulation covers the subject matter, not gauge their relative effectiveness.83 The Court of Appeals notes that in crafting HM-232 DOT considered more stringent rerouting measures, but ultimately settled on deferring to the informed judgment of the railroads as to how best secure hazardous materials shipments. Regardless if the remedy is imperfect, in the Court’s view HM-232 is federal action on rail hazmat security. On this point the Court of Appeals departs from the District Court. While the District Court places important weight on the notion that HM-232 was presented as the first in a succession of comprehensive rules, and that promised subsequent regulation had not yet materialized, the Court of Appeals more narrowly examines if there has been any form of action. The Court reasons that D.C. City Council’s complaint is not that the federal government has failed to act, but that they have not acted in the manner or to the degree to which they would like.84 Nonetheless, the D.C. Act does not qualify as state action designed to fill an unattended gap in federal policy.

82 Ibid. 7.
83 Ibid. 9.
84 Ibid. 8.
Turning toward the safe harbor provisions which allow more stringent state action even if federal regulations or laws have already been enacted, the Court of Appeals examines if the D.C. Act (1) is necessary to reduce a local security hazard; (2) is not incompatible with a law, regulation, or order of the United States government; and (3) does not unreasonably burden interstate commerce. The Court determines that it does not appear that the D.C. Act satisfies these three conditions.85 Terrorism, for the Court, is a national problem motivated by a hatred of the United States as a nation.86 The exception for local hazards is only designed for situations which uniform national standards cannot address and local communities are uniquely positioned to act.87 For the Court, federal policy can—and indeed has in HM-232—address the security of hazardous rail cargoes. Next, the Court asserts that HM-232 is incompatible with the D.C. Act. The preference articulated in HM-232 for voluntary measures is at odds with the requirement to reroute. The Capital Exclusion Zone limits the discretion expressly authorized under HM-232 insofar as it requires that railroads incorporate rerouting around D.C. as part of their security plans. The District Court and the Court of Appeals strike different balances between autonomy and potential compliance: for the District Court it is not impossible that a railroad’s security plan could accommodate rerouting requirements imposed by the D.C. Act; for the Court of Appeals it is not impossible that a security plan developed by a railroad would determine the safest route of transport moves through D.C. In this manner, for the Court of Appeals the D.C. Act and HM-232 are in conflict.88 Finally, the

85 Ibid. 9
86 Ibid. 10
87 Ibid. 9
88 Ibid. 10.
Court links the Act with substantial burdens on interstate commerce. The Act is model legislation that is, if allowed to stand, likely to be replicated by cities and states across the county. The cumulative effective of a number of such local statues would be to severely disrupt the interstate trade of hazardous materials.  

Thus, the Court determines that the D.C. Act does not fulfill the safe harbor conditions outlined by the Federal Rail Safety Act and is preempted.

The Court of Appeals granted a temporary injunction preventing the D.C. Act for taking effect and remanded the case back to District Court for full hearing. The Court of Appeals ruling was a serious blow to D.C. and the environmental movement, but it was not the final word on the Act. The case proceeded in District Court to consider the merits of a permanent injunction. Though the Court of Appeals indicated a strong likelihood that they would rule in favor of an eventual permanent injunction, the Act remained in ongoing litigation. Despite the fact that the Act was prevented from taking effect and did not serve as model legislation to be enacted on a city-by-city basis, it did, however, push the federal government to act and address rerouting. The ongoing legal uncertainty surrounding the Act, pushed DOT to propose new rail security regulations that offered some concessions to the environmental movement. In this fashion, the pursuit of local rerouting laws was not fruitless; the D.C. Act was a qualified success for its backers.

**Federal Rerouting: Initial Proposed Regulations**

---

89 Ibid. 10-11.
90 The concurring opinion of Judge Karen LeCraft Henderson notes that the Act would also be preempted by the Hazardous Materials Transportation Act on similar ground. CSX Transportation, Inc. v. Anthony Williams, (D.C.C. 2005) (Karen LeCraft Henderson, Concurring) No. 05-5131.
The ongoing litigation over the D.C. Act spurred the introduction of new federal rerouting regulations which offered important concession to those seeking rerouting, but stopped short of implementing mandatory rerouting of hazardous cargoes around urban centers. The implemented regulations fall between the loose voluntary regulations of HM-232 and the more prescriptive regulations put forward in the D.C. Act and supported by environmental groups. Importantly, new federal legislation enacted during the review process of the new security legislation includes key provisions sought by environmentalists over the objections of the rail industry.

After the Court of Appeals granted a temporary injunction against the D.C. Act, the District Court began proceedings on a permanent injunction.91 Though CSX and the Department of Justice won an important victory in securing the temporary injunction, the case was not yet full adjudicated. Further, as made clear in the District Court’s initial ruling, presiding Judge Emmet G. Sullivan appeared to be sympathetic to D.C. and the Sierra Club’s arguments in support of the Act. Against this backdrop of legal limbo and uncertainty, DOT introduces new federal regulations governing how railroads select routes transporting hazardous cargoes.92 The DOT regulations seek to remove any ambiguity concerning whether or not the federal government has of yet implemented regulations fully covering the subject matter of the D.C. Act. As the District Court

---

initially found, DOT promised on introduction of HM-232 that additional regulations would be forthcoming. The relative lack of subsequent action, in the District Court’s estimation, creates a gap in federal regulation which allows states and cities to introduce mandatory rerouting regulations.93 With the new set of regulations, DOT sought to make it perfectly clear that federal regulations preempt the D.C. Act as well as similar legislative efforts by cities and states elsewhere.94

DOT’s Pipeline and Hazardous Materials Safety Administration (PHMSA), the successor to the RSPA, proposed new regulations directing railroads to conduct periodic reviews of routes used to ship hazardous materials and examine the viability of additional single alternate commercially practicable route.95 Under the December 21, 2006 Notice of Proposed Rulemaking, the regulations direct railroads to evaluate the safety and security of the surveyed operating and surveyed lines and make routing decisions based on the review. The expressed purpose of the regulation is to ensure that railroads select reroutes that have the fewest safety and security risks.96 In the summary included with the Notice of Proposed Rulemaking, DOT briefly describes that the regulations: require rail carriers to compile annual data on specified shipments of hazardous materials, use the data to analyze safety and security risks along rail transportation routes where those materials are transported, assess alternative routing options, and make routing decisions based on those assessments.97

The outlined regulations call for railroads to annually collect commodity data and review all routes on which large quantities of explosives, radioactive materials, and TIH cargoes...
travel; and additionally review a single alternate commercially practicable route.98 Railroads then must select the route that presents the fewest safety and security risks for hazmat transportation.99 The proposed regulations provide specific guidance as to how railroads should conduct route reviews. The language contained in proposed Appendix D to 49 CFR 172 lists 27 different “rail risk analysis factors” to be considered during annual evaluations of current and alternate routes. The list of factors is wide-ranging, including quality of track, traffic density, volume of hazmat material shipped, as well as known threats and mitigation measures in place.100 Railroads must include in writing analysis of these different 27 elements in their annual reviews.101 Additionally, the proposed regulations provide FRA with enforcement powers; as part of their routine review and evaluation of rail operations during site visits, FRA is enabled to review the routing analysis conducted by the railroad and, upon findings of non-compliance impose alternate routes, operational modifications, or levy civil penalties.102

As initially introduced, the regulations represent a modest tightening of the voluntary measures imposed under HM-232. While HM-232 required the development of security plans that took into consideration en route security, the new proposed regulations instruct that railroads engage in annual review of rail routing decisions and evaluate a single additional route through which hazmat materials could potentially

98 Ibid. 76839-76840.  
99 Ibid. 76842.  
100 Ibid. 76849.  
101 Ibid.  
102 Ibid. 76844.
alternately run. Importantly, the proposal also empowers FRA to levy fines or impose alternate routes in instances of non-compliance.

**Federal Rerouting: Modifications and Review**

The proposed regulations underwent important modifications between initial proposal in April 2006 and enactment of the Final Rule in November 2008. During the review and comment period, environmental groups successfully marshaled support for the introduction of new clauses expanding the scope of route review beyond a pro forma single alternate route analysis, providing local communities with a mandated role in routing decisions, and made explicit the importance of proximate urban areas in safety and security analysis. Through submitted comments and support for new federal legislation, environmental groups were able to ensure that the enacted regulations were not simply an insurance policy designed to head off the D.C. Act in the event of an unfavorable outcome in the judicial process.

Environmental groups, mainly Greenpeace and Friends of the Earth, joined cities and states in pushing for modifications to the proposed regulations. In particular, they pushed for regulations to expand the scope of the required analysis of alternative routes. As initially proposed, as part of their annual review railroads only need to examine a

---

single alternative route—identified as the “next most commercially practicable route over which it has authority to operate”—for each route carrying specified quantities of hazardous cargoes. Greenpeace and Friends of the Earth pushed for analysis which was not limited to the single most economically alternative, but rather included all alternative routes and also considered the possibility of interchange agreements with other railroads. These modifications, if enacted, would greatly expand the scope of rail review. Here, they would have to consider all possible alternatives and also look beyond the lines over which they have direct authority to consider interchange agreements.

Activists also sought to increase the involvement of local communities in routing decisions and make explicit the importance of high-consequence targets adjacent to rail lines in conducting route analysis. Though the proposed regulations explicitly affirm preemption, activists sought inclusion of a mandate requiring railroads to consult with local communities in identifying en route risks. Activists wished for a clause in the regulations which would carve out a space for local participation. Additionally, though the regulations would not mandate rerouting around cities, activists, as a consolation, sought to add language to the regulations identifying the significance of urban areas in making safety and security determinations.

The environmental movement and a host of cities and states offered these modifications during the open comment period following the Notice of Proposed

Rulemaking. More importantly, not taking any chance that DOT would ignore their comments in deference to greater flexibility for the rail industry, they backed new federal legislation compelling the modifications. While under DOT review, the newly Democratic-controlled Congress passed the Implementing Recommendations of the 9/11 Commission Act of 2007 (The 9/11 Commission Act). Title 15, Subtitle D of the Act, Surface Transportation Security: Hazardous Material and Pipeline Security, directs DOT to include many of the modifications activists were seeking. The Act sought to speed up the slow-moving regulatory process by directing DOT to publish its final rule within the next 9 months. Substantively, the Act imposed burdens on how the new regulations would operate. The Act required that the eventual final regulations compel railroads to consider multiple routes, rather than a single alternative route, and possible interchange agreements in assessing alternative routes. In another nod to activists, the Act specified that route analysis take into account and consult with state, local, and tribal governments about significant “high-consequence” targets located near rail lines on which hazardous cargoes travel. The 9/11 Commission Act ensured that DOT would have to accept the concerns articulated by environmentalists and local communities.

Though rerouting legislation had been introduced numerous times in Congress previously with little success, the 9/11 Commission Act represented a triumph for Greenpeace and

108 See, Fn 103.
111 Ibid. § 1551(a).
112 Ibid. § 1551(d).
113 Ibid. § 1551(c)(h).
Friends of the Earth. The Act stopped short of mandating rerouting around urban areas, but it made explicit the importance of considering urban areas in routing decisions, expands the scope of alternate routes railroads examine and from which railroads must select the safest and most secure route, and provides a place for local communities in routing decisions.

Federal Rerouting: Qualified Victory and Compromise

In response to the 9/11 Commission Act, DOT introduced substantial revision to the final regulations. The new regulations located in 49 CFR 172 and 174 move away from the essentially toothless and lax regulations introduced as HM-232 and the modest regulations introduced in the Notice of Proposed Rulemaking. The final regulations require railroads transporting specified hazardous cargoes to examine alternate practicable routes and interchange agreements during annual security and safety reviews.

Rather than only calling on railroads to examine a single alternative line, and look to routes fully under their control, the rule expands the scope of review greatly and, in the process, increases the likelihood that alternative routes will be identified and employed. The revised final regulations also make explicit that local cities and states must be consulted in making routing decisions. Railroads must provide a single point of contact to integrate the information provided by local communities and consult their estimations of risks accompanying different rail routes.

Finally, 49 CFR 172, Appendix D

---


specified that high consequence targets (read: cities) are to be considered in determining safe and secure routes.\textsuperscript{116}

Though railroads fought these changes, the final regulations were not entirely inhospitable to their prerogatives.\textsuperscript{117} Under the new terms, mandatory rerouting was not put in place and preemption was made explicit and unambiguous.\textsuperscript{118} In this fashion, new regulations spared railroads from hostile local rerouting efforts. The regulations also preserved a certain amount of autonomy for the railroads in examining alternate routes. The 27 factors identified in 49 CFR 172, Appendix D covered are not organized in any schematic way.\textsuperscript{119} Though each factor must be addressed in writing and is subject to FRA review, the regulations do not provide a clear framework for weighing the relative importance of listed factors. Indeed, the list is perhaps expansive enough to allow railroads wide-latitude.

\textit{Conclusion: The Politics of Security}

As in the other examined case studies, the central question or set of questions to discuss when examining the bureaucratisation of risk returns to politics: Who defines the new techniques of control employed to domesticate the uncertainty of terrorism? What values are embedded within these new policies and procedures? Glancing back to the previous discussion of Beck, in global terms, are the new security interventions instances of a type of democratization or on the contrary a reactionary retrenchment of power?

Ultimately, the introduction of new rail hazmat security regulations represent a qualified,

\begin{flushright}
\begin{footnotesize}
\textsuperscript{117} PHMSA. “Final Rule.” Docket PHMSA-RSPA-2004-18730. 72183.
\textsuperscript{118} Ibid. 20761-20762, 20766-20768.
\textsuperscript{119} Ibid. 20766.
\end{footnotesize}
\end{flushright}
though important, triumph for environmental activists and their allies in local and federal government.

The final new regulations governing the security of rail cargoes of hazardous materials do not include mandatory rerouting. Yet, in important respects, activists are key authors of new security regulations, often winning concessions over the objections and reluctance of industry, the Executive Branch, and the Department of Transportation. The final DOT security regulations move significantly away from the loose voluntary approach initially championed by industry, the Bush Administration, and the Department of Transportation. By appropriating the language of national security and building coalitions with local politicians at the city and state level, as well as key Democratic senators, environmentalists were able to write into new security regulations mandatory consultation with local governments, annual review of expanded alternate routes and interchange agreements, and consideration of urban areas in route assessments. These are not trivial changes. Taken together, these provisions offer significant transformation of the governance of hazardous materials transportation. New regulations impose burdens requiring railroads to conduct annual review of routing options and justify the selection of routes based on safety and security, while FRA enforcement creates liability for noncompliance by the railroads. Additionally, the introduction of a new mandate for DOT to regulate rail security keeps alive the possibility for further action in the future. In these ways, activists are modestly successful as risk entrepreneurs; the threat of terrorism

provided them with an opportunity to intervene in rail policy and allowed them to shift the tenor of regulations in significant ways.

Post-9/11, the threat of terrorism destabilized rail operations and called into question the existing regime of safety regulations and practices which had evolved for nearly a century. The fight over rerouting is about more than the specific minutia of hazmat policy; rather, it is also about the role of the public in shaping policy. Here, activists were able to appropriate the vernacular of security—seize on the salient though difficult to define threat of terrorism—in order to substantively change transportation policy. Though industry retained key elements of their agenda in the final regulations, the threat of terrorism did open rail policy to scrutiny by the broader public. In early 2000, Greenpeace activists scaled a fence of a Dow Chemical plant outside Baton Rouge, LA to draw attention to the security risks associated with toxic chemicals. Their “raid” drew little attention or substantive change in how chemicals are produced, transported, or used.\textsuperscript{121} After 9/11, however, there concerns appeared far more legitimate and reasonable.\textsuperscript{122} Security provided activists with a say in policy which they might not have otherwise had.

\textsuperscript{121} Eric Pianin. “Toxic Chemicals’ Security Worries Officials.”
\textsuperscript{122} Ibid.
Images

Image 6-1: Chemicals of Mass Destruction

Chapter 7: The Creation of Mandatory Cybersecurity Reliability Standards: Autonomy, Interdependency, and Electric Power

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. Thousands abandoned their typical commute and set out on foot, streaming across the Brooklyn Bridge on their way home; while others decided not to brave the hectic conditions and bunk wherever they could. That day, the largest blackout in the nation’s history stretched across eight states and parts of Canada, leaving over 50 million people without power. Shortly after 4:00 p.m., 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, Midwest, and Ontario. Speaking in San Diego, President Bush promised a thorough review of the nation’s electrical system.

Though not caused by malicious action, the specter of terrorism informed the eventual response to the blackout. The blackout spurred the expansion of federal regulatory powers governing the nation’s electrical system and the first set of mandatory

---


cybersecurity regulations for the electric power grid. Federal regulators from the Federal Energy Regulatory Commission (FERC) and the electric power industry haggled over the terms of new regulations and the limits of newly granted federal authority. Somewhat surprisingly, FERC leveraged its limited powers to ultimately implement cybersecurity regulations significantly tighter than standards proposed by industry. FERC bucked industry preference for limited administrative regulations containing ample opportunities for industry to opt out and, instead, worked to secure comprehensive cybersecurity controls.

In this chapter, I sketch the development and implementation of new cybersecurity regulations for the electric power grid. The new regulations, known as Critical Infrastructure Protection Reliability Standards 002 through 009 (CIP-002 through CIP-009), confront the entwined challenges unleashed by deregulation and the dynamics of technological change. Deregulation remade the provision of electric power from a loosely-coupled system of vertical utility providers into a competitive market of interdependent players offering generation, transmission, and distribution.⁶ Deregulation, both directly and indirectly, transformed the use of information and communication technologies (ICTs) within electric power in ways that opened the power system to new forms of large-scale failure (see Chapters 2 and 3).⁷ After the 2003 blackout, the limits of competitive markets and the regime of voluntary standards became apparent. The joint U.S.-Canada Task Force examining the causes of the blackout identified the changing institutional arrangements, i.e., the move to competitive markets, as a contributing factor.

---

⁷ Ibid. 132.
to the outage and a source of continued vulnerability. More specifically, the Task Force identified the new adoption and deployment of ICTs within electric power as key area of concern. The blackout occasioned a revisiting of the consequences of deregulation and a search for polices to confront the new vulnerabilities tethered to reorganization. The blackout opened the system of electric power to a reevaluation; in this fashion, it engendered a moment of political reflexivity in which deregulation became a topic of debate, active concern, and ultimately new interventions.

*The August 14, 2003 Blackout: Revisiting Electric Power*

The August 14, 2003 blackout spurred an investigation into the stability of the interconnected electric power system and, ultimately, led to the adoption of a new structure of governance for the industry. The Task Force charged with investigating the blackout offered a series of recommendations that provided the blueprint for new legislation formalizing the authority of the North American Reliability Council (NERC), an industry trade association, and outlining a limited role for federal regulators in setting mandatory standards for the electric power grid.\(^8\) Though the Task Force concluded that the blackout was not caused by malicious activity or an act of terrorism, they took the threat of terrorism seriously and made the need for cybersecurity regulations central to its recommendations.

The August 14, 2003 blackout was the worst power outage in the nation’s history: over 50 million people in the U.S. and Canada lost power, as 508 generators at 265 plants

---

\(^8\) NERC was renamed the North American Reliability Corporation in 2006.
fell off-line impacting 61,800 megawatts (MW) of electric load. The blackout’s economic impact was staggering; estimates place losses from the outage at between $4 and $10 billion dollars. On August 15th, President Bush and Canadian Prime Minister Jean Chrétien announced the formation of the U.S.-Canada Power System Outage Task Force to investigate the causes of the blackout and offer recommendations to improve the reliability of the electric power system. The Task Force investigated both the blackout’s contributing causes and conducted a broad review of the electric power system by interviewing energy providers and experts, and reviewing collected system data. U.S. Secretary of Energy Spencer Abraham and Herb Dhaliwal, 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. 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.

---

9 In terms of MW and people impacted, the 2003 blackout far outstripped pervious blackouts, including the 1965 Northeast Blackout and the New York City blackout of 1977. The 1965 blackout, for example, impacted an estimated 30 million people and involved 20,000 MW. For historical comparisons with other blackouts, see, U.S.-Canada Power System Outage Task Force. *Final Report.* 1-2, 73-74, 103-107.

10 Ibid. 1.

11 Ibid. 1-2.

12 Phase I investigated the causes of the blackout; phase II examined requirements needed to prevent future blackouts. Ibid. 1

13 Ibid. 1.

The Task Force released its interim report in November, 2003 and final report in April, 2004. The final report identifies a mix of technical failures, human errors, and institutional factors as contributing causes. The blackout was initially set in motion by transmission-line failures and faulty system monitoring in Ohio. These failures cascaded through the system, throwing transmission lines and generators off line in quick succession. Importantly, the Task Force identifies a decidedly low-tech contributing cause to the blackout: faulty vegetation practices. Improper maintenance of tree growth near power lines helped spread the outage. The failure was not, however, merely technical; at critical moments, key human errors failed to stop the cascading failure before it was too late. Finally, the Task Force identifies lax and ambiguous voluntary regulations underling the blackout. In examining the practices of FirstEnergy, the Midwest Independent System Operator (MISO) PJM Interconnection, key players in the outage, the Task Force found multiple violations of NERC’s voluntary reliability standards, and other standards so ambiguous to make compliance difficult. In the Task Force’s reading, multiple overlapping technical, human, and institutional causes contributed to the widespread outage.

_Terrorism, Cybersecurity, and the U.S.-Canada Power System Outage Task Force: An Accident or a Cyber Attack?_

---

17 Ibid. 45-49.
18 Ibid. 18-19, 45.
19 Ibid. 18-19, 48.
20 Ibid. 17-20, 22.
The Task Force examined another possible source behind the massive outage: terrorism. The possibility that the blackout could have been caused by a terrorist act informed the Task Force’s work and, ultimately its recommendations. Consistent with the make-up of the Task Force and broad mandate to examine possible sources of future failures, the Task Force moved beyond the causes of the 2003 blackout to offer a more general vulnerability assessment. In the process, the Task Force helped transform or shade the meaning of “security” within discussions of electric power from a preoccupation with system stability to a meaning more closely aligned with national security, malicious activity, and contemporary fears of terrorism.

The power outage initially sparked fears of a terrorist attack. Post-9/11, major infrastructure failures were greeted with foreboding: Could this be another terrorist attack? On the day of the blackout, the New York City Policy Department activated its terrorism-response plans and deployed armed teams to protect high-profile targets, such as the New York Stock Exchange. A few weeks earlier in July, U.S. and Canadian government officials issued warnings to the energy sector based on collected intelligence indicating that Al Qaeda might attack power plants. 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

---

22 Rashbaum. “Responding to a Blackout, but on Alert for Terrorism.”
is that this was not a terrorist act.” Yet, concerns lingered. On August 18th, Al-Hayat, an Egyptian news outlet, reprinted a communiqué attributed to al-Qaeda claiming credit for the power outage. Later in the month and in September, similar claims appeared in various news outlets.

The Task Force was oriented toward and examined seriously charges that the outage could be the product of 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 then-Secretary of the Department of Homeland, Security Tom Ridge. The selection of Secretary Ridge highlighted the prioritization of security and terrorism in the Task Force. Similarly, the creation and designation of the Security Working Group as one of three working groups within the Task Force underscores the prominence of security concerns within the inquiry. The Security Working Group examined collected intelligence and 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.

Redefining Security: The U.S.-Canada Taskforce Considers Cybersecurity

26 Ibid.
27 Ibid. 1.
28 Ibid.
29 Ibid. 134-135.
Yet, though the Task Force found no concrete evidence of foul play, they did identify significant cyber vulnerabilities within electric power. These vulnerabilities, in the Task Force’s estimation create an opening for enterprising terrorists to cause great harm. In keeping with its mandate to offer a broad review of reliability, the Security Working Group seized on cyber vulnerabilities as a significant area in need of improvement.30 The final report notes 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.”31 Recall, the reorganization of electric power under competitive terms fostered the consolidation of control over system operation and, in the process, creates new points of vulnerability. Centralized ICTs are critical to the move from vertical monopolies to competitive markets. At the same time, regulatory reform reversed incentives that historically supported diverse and isolated control systems, here too creating new possibilities of failure (see Chapter 2 and Chapter 3). The Task Force zeros in on these challenges, noting that SCADA and control systems are particularly susceptible to corruption. For the Task Force, the migration of ICTs within electric power from stand-alone proprietary systems to common platforms operating across corporate networks and the internet creates new cybersecurity challenges for electric power.32 The combination of standardized equipment and increased connectivity through the internet creates interdependences within electric power that can be exploited to undermine the functionality of the grid. The report readily accepts that electric power

30 Ibid. 131-138.
31 Ibid. 132.
32 Ibid.
is a target, noting: “[t]he generation and delivery of electricity has been, and continues to be, a target of malicious groups and individuals intent on disrupting this system.”33 In addition to targeted attacks, the spread of general worms and viruses designed to wreck havoc indiscriminately, rather than target electricity control systems specifically, can pose a threat to electric power systems.34 The Task Force, both implicitly and explicitly, wrestles with the challenges presented by the reorganization of electric power. The increased reliance on centralized controls and common or standard ICTs operating on accessible networks creates new vulnerabilities which, in the Task Forces estimation, require immediate action.

The Task Force’s discussion of cybersecurity reworks traditional notions of what “security” means within electric power.35 Historically, within electric power “security” is defined as: “[t]he ability of the electric system to withstand sudden disturbances such as electric short circuits or unanticipated loss of system elements.”36 Security, in this sense, relates to system stability and reliability. To ensure a secure system within this framework meant planning for and having on hand adequate resources, particularly generating capacity, to meet fluctuating demand and keep the system in a state of balance matching generation and consumption. Under this rubric, a secure system is one that has adequate capacity to meet demand. The Task Force, however, also adopts the more colloquial notion of security that means something closer to “defense against intentional

33 Ibid. 132.
34 The report notes that FirstEnergy, a key player during the 2003 blackout, had significant problems with the “slammer” worm in January 2003. The worm attacked control systems at FirstEnergy’s idle Davis-Besse nuclear plant. Ibid. 132-133.
35 There is confusion and overlapping usage within the report itself. See, for example Ibid. 23.
36 Ibid.
attack.” Here, the idea of a “secure” system is quite different; rather than emphasizing adequate resources, security is linked with authorized access, defensive postures designed to repel or discourage sabotage, and the broader interests of national security. By emphasizing the threat of terrorism and malicious intentional disruption, the Task Force expands the meaning of security. Now, the Task Force argues, planning for security must confront the possibility of enterprising intentional disruption.

The move to consider targeted attacks within the catchall of electric power security and reliability is a significant complication. Historically, operators relied on what is known as the “N-1 Criterion” to guide secure and reliable operation. Long used by operators and written into NERC Operating Policy 2.A—Transmission Operations, N-1 directs operators to tailor planning and administration to withstand the loss of the single most critical transmission or generation asset.37 N-1, or normal operating conditions minus one, is designed to support systems that operate at a state that can at any moment withstand the single worst contingency without interruption. N-1 is intended to ensure that systems operate with a margin of fault tolerance that can absorb the single worst outage without lapsing into a large-scale cascading failure.38 During day-to-day operation controllers continually reevaluate and adjust system performance to confront changing conditions. For example, if the worst single contingency occurs, operators are instructed to adjust to the new N-1 state—that is, adapting to survive what is now the

37 Ibid. 8-9, 13.
38 Ibid.
single largest contingency within the reconfigured system conditions—within 30 minutes.\textsuperscript{39}

The expansion of the definition of security to include prevention of intentional disruptions and protection of cyber assets challenges the utility of N-1. N-1 works well to confront random failures. It is assumed that when these sorts of infrequent failures occur they will only trip a single asset at a time.\textsuperscript{40} Intentional malicious attacks, however, do not follow a random distribution; they can occur at multiple locations and target multiple assets simultaneously to maximize the scale of harm. In confronting terrorism, multiple simultaneous contingences, rather than only the single worst contingency, appear as a legitimate possibility.\textsuperscript{41} The perceived threat of terrorism, as elsewhere, expands the range of credible scenarios. Likewise, the deployment of standardized cyber tool and the increasing reliance on the internet as a backbone undergirding the electric power system further undermines the utility of N-1. Standardized ICTs and the reliance on the internet create new interdependencies that open electric power to the possibility of simultaneous multiple failures across and within systems. As NERC’s Chief Security Officer notes, the deployment of new technologies within electric power create possibilities of horizontal disruption, i.e., disruptions spreading across and within systems that impact multiple assets, that are not well accounted for under N-1.\textsuperscript{42} For example, a

\textsuperscript{39} Ibid.
\textsuperscript{42} Assante. “RE: Critical Cyber Asset Identification.”
worm or virus that impacts operating tools relying on a common operating system or widely-used SCADA architecture would most likely spread through interconnected networks and disable many different components situated throughout the power grid, rather than a single isolated system. N-1 is designed to confront random outages. However, terrorism and the new information and communication-landscape present different challenges. The possibility of targeted simultaneous attack or horizontal failures of standard computer equipment go beyond the scope of N-1. In these ways, planning for intentional disruption generally and cybersecurity specifically challenges the standard model of ensuring system reliability. In the Task Force’s estimation an overhaul of the governance of electric power is overdue.


The U.S.-Canada Power System Outage Task Force’s final report presents both general and specific recommendations for overhauling the governance of electric power and confronting the challenges of cybersecurity. The recommendations undergird new U.S. legislation, the Energy Policy Act of 2005, which implements the Task Force recommendations. The recommendations and subsequent legislation sought to preserve the role of industry as a steward of self-regulation in the development of new reliability standards, including those designed to cover cybersecurity. In this fashion, the Task Force and the Energy Policy Act were conservative responses to the blackout and identified cybersecurity vulnerabilities. Rather than undertaking a fundamental transformation, both the recommendations and the Act attempt to buttress the power of
industry to self-regulate by supporting mechanisms through which industry-developed reliability standards become legally enforceable, while limiting the role of federal regulators in developing reliability standards.

The Task Force’s final report contains 46 recommendations, covering four substantive areas: Institutional Issues Related to Reliability; Support and Strengthen NERC’s Actions; Physical and Cyber Security of North American Bulk Electric Power Systems; Canadian Nuclear Power Sector. At a global level, the recommendations call for an overhaul of the process of developing and enforcing reliability standards. For decades, voluntary measures and reliability standards governed electric power. The Task Force, however, notes that enforcement of these NERC standards is lacking and that they are ignored with little punishment. As an industry trade group, NERC standards do not carry the force of law and provide few tools to address non-compliance. The Task Force seeks to improve the ability of NERC to enforce reliability standards, while preserving the role of industry as an essentially self-regulating entity. The Task Force recommends a tiered regulatory structure that would allow NERC to internally create and enforce mandatory reliability standards, and provide federal regulators, FERC, with limited power to accept or reject the NERC-developed standards. Under the Task Force’s model, federal regulators would certify an “electric reliability organization” (ERO) representing industry to develop and enforce mandatory reliability regulations.

---

Through the ERO, industry would develop reliability standards and enforce compliance. The Task Force made plain in its final report that the ERO-concept was designed and proposed with the understanding that NERC would move into the role. The role of federal regulators under the proposed plan is narrowly drawn. Under the Task Force’s recommendations, FERC would have the authority to accept or deny proposed reliability standards, but not draft or develop regulations; they would defer to industry in crafting reliability standards. Within this scheme, FERC’s approval gives industry-developed regulations legal standing. The proposed governance structure enhances the power of industry to set reliability standards and closes gaps in enforcement endemic under the voluntary model. Yet, industry still largely controls the development of reliability standards. Under the proposed guidelines, the ERO, that is to say NERC, controls substantive debate concerning reliability standards and is imbued with the authority to draft guidelines based on internal debate within industry. FERC, on the contrary, has limited authority to approve or remand proposals, but not develop substantive regulations.

In parallel with its comprehensive recommendations, the Task Force takes special note of cybersecurity. 13 of 46 recommendations specifically address cybersecurity. The Task Force recommendations concerning cybersecurity on the whole largely refrain from suggesting specific measures. Rather, the Task Force calls for industry to begin to address the issue through corporate information technology (IT) plans and IT

---

47 Ibid. 141-142.
48 Ibid. 140-142.
49 Ibid.
management strategies, as well as periodic reviews of cyber vulnerabilities. A few recommendations, however, are more narrowly tailored: the Task Force calls for the introduction of tools to monitor wireless and remote intrusion and the use of controls to limit access to sensitive equipment. The Task Force also recommends the adoption of NERC’s cybersecurity standards as part of the drive toward mandatory reliability standards. In a coincidence, on August 13, 2003, the day before the blackout, NERC implemented its first voluntary cybersecurity standards, Urgent Action Standard 1200 (UA 1200). The Task Force did not provide a review of UA 1200, but, nonetheless, advocated its adoption as a mandatory reliability standard as a first step in addressing cybersecurity.

The recommendations, overall, show great deference to industry prerogatives and propose a change in the structure of governance, which on review represent a modest alteration of the status quo. The Task Force explicitly supports the continuation of self-regulation by industry; its recommendations emphasize the importance of continuing to develop reliability standards within the confines of NERC. The Task Force’s suggestion that NERC’s voluntary standards adopt the imprimatur of law responds to the apparent incompleteness of enforcement under voluntary regulations. Likewise, though the Task Force calls for an expansion of the concept of security to include cybersecurity, it leaves

50 Ibid.
51 Ibid.
52 Ibid.
ample room for industry to define the tenor and form that such regulations would ultimately adopt.

The recommendations were written into law with little adjustment in the Energy Policy Act of 2005.\textsuperscript{55} Title XII of the Act, known as the Electricity Modernization Act of 2005, for the first time grants authority to create mandatory reliability standards.\textsuperscript{56} The Act amends section 215 of the Federal Power Act and adopts the Task Force’s endorsement of a bifurcated system with industry, through the ERO, in the central position of defining and enforcing mandatory reliability standards. The Act calls for the development of mandatory reliability standards for the bulk-power system and, in particular, calls for new cybersecurity standards.\textsuperscript{57} The Act specifies that reliability standards cover all 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 Act, generation and transmission fall within the “bulk-power system,” while distribution to local entities is not covered. The Act specifies that reliability standards are designed to support the operation of the interconnected power system such that sudden disturbances will not lead to cascading failure.\textsuperscript{58}

The Act grants FERC limited powers and carves out a significant space for industry to control drafting and enforcement of reliability standards. The Act calls for FERC to certify the ERO. Under the term of the Act, any entity can apply to FERC to become the ERO. In order to be certified as an ERO, the applicant must (1) have the

\textsuperscript{56} Ibid. § 1211.
\textsuperscript{57} Ibid.
\textsuperscript{58} Ibid.
ability to develop and enforce reliability standards for the bulk-power system; (2) operate independently from specific users and providers; (3) assure stakeholder representation within the standards development process; (4) provide impartial enforcement of reliability standards.\(^5^9\) The Act’s call for independence from users and providers is designed with the aim of preventing a single group or entity from dominating the ERO, not limiting industry involvement. On the contrary, indeed as the Act makes clear the ERO is to be an industry forum within which standards will be developed and enforced. 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; they cannot draft reliability standards, but rather only approve or remand. Additionally, FERC’s review criteria are narrow. The terms of Act require FERC to give “due weight to the technical expertise” of the ERO when examining a proposed standard.\(^6^0\) While FERC is empowered to review and approve standards, they are 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.\(^6^1\) But, again, FERC cannot dictate the content or terms of the standard and must defer to the technical competence of the ERO. The Act also specifies that the ERO is responsible for enforcement and can levy penalties on any covered entity found to be in non-compliance, subject to FERC review.\(^6^2\)

\(^{59}\) Ibid.  
\(^{60}\) Ibid. \(^{61}\) Ibid. \(^{62}\) Ibid.
The Act also explicitly embraces an expansive notion of reliability that incorporates cybersecurity within its purview. The Act requires the ERO to develop reliability standards that apply to intentional cyber attacks or what it terms a “cybersecurity incident.”63 In this fashion, it adopts the extended notion of security discussed within the Task Force’s final report. The Act defines 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.”64 Though the Act generally left the content of reliability standards to be determined, it specifically singles out cybersecurity as an area where the ERO must develop standards. The prioritization of cybersecurity is in keeping with the Task Force’s recommendations.

Similar to the discussion of hazardous material transportation in Chapter 6, the threat of terrorism is used to justify an expansion of regulatory powers into a new domain: security. Unlike the expansion of the Department of Transportation’s power to regulate the security of hazardous materials, however, the Energy Policy Act of 2005 took pains to limit the role of federal regulators in devising new security regulations. The Act created a tiered regulatory structure that supports the centrality of industry in the regulatory process. Though on its face the creation of new legislation enabling mandatory reliability standards appear to be a drastic expansion of government power over electric power operations, a more close examination reveals that the role of federal

---

63 Ibid.
64 Ibid.
regulators, FERC, is intended to be marginal. FERC is instructed to defer to the technical expertise of industry and cannot draft reliability standards. Rather, they are reduced to approving standards developed internally through industry consensus or remanding standards. Both the Task Force’s recommendations and the subsequent legislation intend to provide industry with the tools to better self-regulate, not greatly expand federal involvement. In this manner, the push for mandatory reliability standards generally, and standards covering cybersecurity specifically, were industry-friendly initiatives that sought to increase, not diminish, the power of industry over the operating terms of the bulk-power system. Yet, surprisingly, federal regulators did not sit idly by and rubber stamp industry-approved reliability standards. On the contrary, they leveraged their limited power to substantially redefine the terms of cybersecurity regulations.

*Mandatory Cybersecurity Regulations: Regulatory Activism and Transformation*

Unsurprisingly, under the new terms of the Energy Policy Act, NERC sought recognition as an ERO and moved to convert its voluntary guidelines into mandatory reliability standards. Though the process set-up under the Energy Policy Act is structured to provide NERC, that is, industry, with control over reliability standards, enterprising federal regulators from FERC leveraged the limited power available to extract significant changes to cybersecurity regulations. Through the approval process, NERC and FERC clashed over the degree of industry autonomy, the specificity of new

---

regulations, and the penalties associated with non-compliance. The freshly minted regulatory process at this point had not yet been defined or tested. In seeking control, both industry and federal regulators attempt to carve out roles and delimit authority in order to set the tenor and contours of the new relationship between federal regulators and industry. FERC engagement in regulatory activism helped transform cybersecurity regulations over the objections of industry from little-more than loose guidelines into prescriptive regulations.

*Implementing the Energy Policy Act and Defining FERC’s Role*

FERC Order 672, issued Feb. 3, 2006, implements 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. In the Order, FERC asserts a prominent role in defining reliability standards and rejects calls to defer to expertise and judgment of the ERO. FERC interprets the Energy Policy Act broadly as providing a key role for FERC that goes well beyond simply affirming industry reliability standards. FERC’s interpretation cuts against the intent behind the initial Task Force recommendations and, to a large degree, the apparent purpose of the Energy Policy Act. The interpretation is certainly plausible within the boundaries set forth by the Act, but it is a generous interpretation of role of FERC under the Act. Most importantly, in defining the meaning of “due weight,” the importance of flexibility within mandatory standards, and the scope of remand authority, Order 672 consistently

---

66 FERC. Order 672.
offers interpretations of the Energy Policy Act in favor of an enlarged federal role at the expense of the power of the ERO.

FERC reiterates the basic criteria for the ERO introduced in the Energy Policy Act, but goes further in defining the relationship between FERC and the ERO in the development of reliability standards. Order 672 specifies that for an applicant to be certified as the ERO it must be independent, impartial, and open to all industry stakeholders.67 FERC does not dictate a specific governance structure but rather leaves it open for applicants to determine.68 The ERO is to be the only entity empowered to develop and enforce reliability standards.69 Under the outlined terms, the ERO will develop reliability standards through open deliberative process involving industry stakeholders and submit them for approval by FERC.70

FERC makes clear that it views the process of review as a serious matter that, potentially, can shape the direction and content of reliability standards. FERC cautions that it will closely review proposed reliability standards to ensure that the standard does not amount to bland “lowest common denominator” standards that appease all corners of industry.71 The Energy Policy Act charges FERC to review proposed reliability standards to determine if they are just, reasonable, not unduly discriminatory or preferential, and in the public interest, while also giving due weight to the technical expertise of the ERO. There is a tension between the call for FERC to defer to the

67 Ibid. 48. 54-60.
68 Ibid. 54-60.
69 Ibid. 8.
70 Ibid. 105-140.
71 Ibid. 10.
expertise and prerogatives of industry on the one hand, and FERC’s duty to perform oversight. In Order 672, FERC makes clear how they intend to resolve these competing values. It asserts that it does not view the call for “due weight” to be a carte blanch grant of authority to the ERO. On the contrary, FERC rejects calls from industry to presume approval for reliability standards developed by the ERO and remand only in instances of observed procedural violations by the ERO (i.e., instances when the ERO deviates from its own guidelines). FERC refuses to accept the proffered claim that “due weight” establishes a rebuttable presumption—that a finding is to be taken as true unless a specific objection against it is offered—in favor of all proposal emerging from the internal deliberations of the ERO. FERC rejects the contention that a standard that passes through the internal ERO development and vetting process automatically passes muster as just, reasonable, and non-discriminatory.72 Such a claim, advanced by members of industry, if accepted would have effectively shut FERC out of substantive decisions.73 FERC refuses to relinquish any claim to examine offered standards. Instead, it stakes a bold claim to authority that relies on a broad interpretation of the Energy Policy Act. Rather than taking the Energy Policy Act to offer FERC an administrative role in making sure the ERO does not commit procedural violations, FERC holds out a substantive role for itself in examining the content of reliability standards.

Likewise, elsewhere in Order 672, FERC rejects the notion that approved reliability standards should in each instance offer general guidelines as opposed to

72 For a discussion of industry claims in favor of a strong ERO unencumbered by substantive review and FERC’s rejection of this position, see Ibid. 50-54, 136-138.
73 Ibid. 136-138.
detailed requirements. FERC rejects that approved reliability standard must avoid
prescription or that FERC may only review the ultimate aims of proposed standards at the
to the expense of an analysis of the precise mechanisms through which such aims are supported.
Counter to the claims of some members of industry, FERC notes that in some cases
regulations will have to move beyond specifying general goals and leaving the
mechanisms employed open to the discretion of each individual covered entity, but also
concern the specific procedures and mechanisms to be used to reach a goal.\textsuperscript{74} FERC
asserts that to ensure that reliability standards are technically sound, uniform, and
unambiguous an examination of implementation will in some cases be indispensable.\textsuperscript{75}
While FERC accepts the general contention that reliability standards will, in some
instances, admit flexibility, it will not accept that this need renders scrutiny or review of
specific operations as beyond the scope of its duties.\textsuperscript{76} FERC, here, is signaling that
regulations will in some cases touch on both the “what” and the “how” of operations,
and, consequently, FERC review can and will take both under examination.\textsuperscript{77} FERC
rejects NERC’s contention that regulations should avoid prescription in favor of
specification of desired outcomes, while leaving it up to each regulated entity to devise
means to achieve outlined goals. In affirming the principle that regulations and
Commission review of proposed regulations will on occasion necessarily concern
implementation, FERC enlarges the scope of its authority to conduct review and signals

\textsuperscript{74} See for example, comment of PacifiCorp. Ibid. 136.
\textsuperscript{75} Ibid. 107-109.
\textsuperscript{76} Ibid.
\textsuperscript{77} Ibid.
that it can and will involve itself in substantive issues involving the implementation of reliability standards.

Finally, in Order 672 FERC asserts its prerogative that the terms of the Energy Policy Act allow it to remand proposed standards with the inclusion of a deadline for ERO resubmission of revised reliability standards. FERC clarifies the scope of remand authority in accordance with the broad interpretation of its powers found throughout Order 672. Though, as industry argues through NERC, FERC cannot draft reliability standards, by asserting the right to remand with a deadline FERC carves a proactive role for itself in the drafting process.\footnote{Ibid. 148-150.} By remanding a proposal with a detailed explanation, FERC’s hand can weigh heavily on the ERO such that they are de facto authors.

Taken together, Order 672 offers a broad interpretation of the powers afford to FERC under the Energy Policy Act and moves to qualify the scope of authority assigned to the ERO. The Energy Policy Act and the recommendations of the Task Force intended to ensure that industry would have decisive say over the material content of enforceable 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 engages in regulatory activism expanding its role. Rather than being content to sit on the side-lines and offer rote approval of standards developed by industry, FERC interjects itself into the substantive work of defining new standards. Through its interpretation of the meaning of “due weight,” understanding of importance of flexibility and definition of the scope of remand authority, FERC, at each turn consistently offers
readings that support a greater federal role at the expense of the authority of industry and the ERO. Order 672 signals to industry that the redrafting of existing voluntary reliability standards into new mandatory regulations will not be a pro forma exercise. Indeed, as would soon become clear, FERC intends to leave its stamp on reliability standards.

**Critical Infrastructure Protection Reliability Standards: Security, Autonomy, and Interdependencies**

Shortly after filing for certification as an ERO, NERC submitted its first batch of proposed mandatory cybersecurity standards, known as Critical Infrastructure Protection (CIP) Reliability Standards 002 through 009-1 (CIP-002-1 through CIP-009-1), for FERC review. FERC notes that industry’s preference for general administrative regulations that left substantive decisions concerning the implementation of cybersecurity measures up to each player would not be acceptable. FERC’s review highlights four issues and concerns that dominate the approval process: (1) the applicability of the concept of “reasonable business judgment”; (2) the ability of covered entities to accept

---


risk in lieu of compliance; (3) the overall specificity of CIP standards; (4) and the assigned penalties for non-compliance.\textsuperscript{82} Through the approval process, FERC uses its asserted powers to push mandatory reliability standards in the direction of tighter controls and limited discretion for covered parties.\textsuperscript{83} On each issue, FERC pushes against industry preference for porous and vague administrative standards carrying light penalties and, instead, advocates more rigorous cybersecurity reliability standards. Underlying FERC’s arguments on each of the above four issues is a concern and recognition of the unique challenges of interdependent security. For FERC, implicit and explicit grants of discretion threaten to subvert the utility of mandatory cybersecurity regulations and put the security of the interconnected bulk-power system at risk.

NERC’s proposed mandatory CIP reliability standards create a framework for the identification and protection of critical cyber assets to support the reliable operation of the bulk electric system.\textsuperscript{84} In initial filing, NERC outlines the intended scope of CIP reliability standards by defining the key terms “cyber assets,” “critical assets,” and “critical cyber assets.” It defines “cyber assets” as “programmable electronic devices and


\textsuperscript{83} The approval process for reliability standard involves four general steps: (1) the ERO develops consensus standards through working groups and balloting; (2) FERC conducts a preliminary technical review of the proposed standard (or standards) and issues a report detailing findings; (3) FERC issues a formal Notice of Proposed Rulemaking and opens the standard to public review and discussion; (4) FERC order to either accept or remand the standard.

communication networks including hardware, software, and data.” Critical assets are defined as “[f]acilities, systems, and equipment which, if destroyed, degraded, or otherwise rendered unavailable, would affect the reliability or operability of the Bulk Electric System.” Finally, “critical cyber assets” are described as “[c]yber Assets essential to the reliable operation of Critical Assets.” Each standard applies to the entities playing a role in the generation and transmission of electric power. As drafted, CIP reliability standards are designed to prevent and respond to malicious disruption or attempted disruption of critical cyber assets. CIP-002 through CIP-009 offers 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.” CIP standards require responsible entities to produce security plans, access control policies, incident response and recovery plans. More specifically, regulations also call for the use of anti-virus and anti-malware utilities for all critical cyber assets (what it refers to as “Malicious Software Prevention”), the creation of a secure electronic perimeter around all critical cyber assets, 

---

86 Ibid.
87 Ibid.
88 Ibid.
89 The standards do not cover local distribution. Each CIP standard identifies 11 responsible entities subject to the rules: reliability coordinator; balancing authority; interchange authority; transmission service provider; transmission owner; transmission operator; generator operator; load serving entity; NERC; regional reliability organizations. NERC. “Cyber Security: Standard CIP-002-1—CIP-009-1.”
and electronic and physical access controls (for an overview of each reliability standard, see Table 7-1).  

FERC notes that the standards are much improved from voluntary standards UA-1200, which were previously in effect. UA-1200 covered much of the same ground as CIP, 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 non-compliance. More troubling, UA-1200 explicitly exempted control systems, perhaps the most important and significant cyber asset, from cybersecurity standards. The proposed mandatory reliability standards move away somewhat from these limitations. CIP standards include control systems within its scope, provide for enforcement through the ERO, rather than self-certification, and assign monetary penalties for non-compliance. Yet, despite these improvements, FERC raises important objections to the CIP reliability standards that cut to the heart of their effectiveness.

Rewriting CIP Reliability Standards: Reasonable Business Judgment and Acceptance of Risk

FERC’s objections take issue with language inserted into the text providing industry with ample opportunity to in effect opt out of regulation at its choice. Each CIP standard contains text indicating that entities subject to the requirements should use “reasonable business judgment” in applying the standard. Each CIP standard notes:

---

91 NERC. “Cybersecurity: Standard CIP-002-1—CIP-009-1.”
“[e]ntities should interpret and apply Standards CIP-002 through CIP-009 using reasonable business judgment.”93 NERC clarifies the meaning of “reasonable business judgment” in a Frequently Asked Questions (FAQ) document submitted as part of its filing seeking approval. In it, NERC notes that the phrase has a 200-year history in business and corporate common law and here is 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.”94 Historically, the principle has been applied to protect corporate officers and board members from judicial review. Under the accepted legal meaning of the phrase, corporate board members and directors are granted wide-discretion in their official capacity to act on behalf of the business, and are obligated to act in what they perceive to be the best financial interests of the company.95 As such, they are protected from judicial review of their decisions so long as they do not act with gross negligence. As NERC describes:

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.96

93 NERC. “Cybersecurity: Standard CIP-002-1—CIP-009-1.”
96 NERC. “Frequently Asked Questions (FAQs):”
The principle of “reasonable business judgment” protects the autonomy of corporate board members and officers and limits liability for decisions undertaken with the understanding of furthering the best interests of the company.  

NERC and others argue for the importance of pairing reasonable business judgment with CIP reliability standards. For NERC, Pacific Gas and Electric (PG&E), and San Diego Gas and Electric (SDG&E), and others, the inclusion of language asserting the importance of reasonable business judgment reserves the right of responsibility entities to weight the presumed costs and benefits associated with any new security measures. In their estimation, security should be carefully weighed within the broader scope of the costs and benefits of each individual company. Without such a qualification, as they note, the possibility of prohibitively expensive new forms of security exist. Likewise, Tampa Energy and Progress Energy and others, assert that the reasonable business judgment qualification is necessary to provide responsible entities the needed flexibility to implement CIP standards in the ways that are best suited for their operating profile.

In the context of CIP standards, reasonable business judgment provides regulated parties with the opportunity to disregard (so-called) mandatory reliability standards if, in

---

their estimation doing so conflicts or is unsupported by the financial best interests of the company. The inclusion of the phrase protects all regulated parties from effective oversight by limiting the ability of both FERC and the courts to review action of regulated parties. The inclusion of the text in the CIP standards allows firm to opt out of compliance when it does not, in their view, suite the bottom line. In FERC’s view, such a standard is inappropriate within the context of CIP. The language provides generous opportunity to opt out of reliability standards and offers far too much discretion for regulated parties. For FERC, there is an obvious tension between the pursuit of individual financial wellbeing and the over-all security of the bulk-power system. The aim of CIP reliability standards is to promote health and resilience of the interconnected electric power grid, not the best interests of any single company’s shareholders. A FERC remarks, “[b]usiness convenience cannot excuse compliance with mandatory Reliability Standards.” Indeed, mandatory reliability standards are called for precisely because these two interests—parochial financial interests and security—diverge.

The interdependent nature of the power grid, and more specifically the interconnection of information and communication networks and technologies connecting

---


different parties, renders the reasonable business judgment standard as inappropriate in this context.\textsuperscript{104} The challenge of interdependent security is central to FERC’s objections. Weak cybersecurity standards by one party involved in the bulk-power system threaten to undermine the security of other interconnected parties and the larger grid. As FERC notes, “[c]yber 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 system can affect the entire grid.”\textsuperscript{105} For example, a worm or virus infecting the control systems of a single player could spread across common communication channels infecting others. Here, the “weakest link” problem which follows from the interdependent nature of the grid and inadequate security jeopardizes the larger community involved in the provision of generation and transmission services.

On similar terms, FERC rejects language included in a number of CIP standards allowing regulated entities to issue a “statement accepting risk” in lieu of complying with outlined requirements.\textsuperscript{106} Text allowing the substitution of acceptance of risk as opposed to meaningful compliance is included in CIP-003 and CIP-007. In CIP-003, responsible entities are allowed to issue a statement of “acceptance of risk” in instances where they fail to comply with required comprehensive security plans.\textsuperscript{107} In the same vein, CIP-007 allows parties to opt out of the requirements to close all unused ports (connections), adopt

\textsuperscript{104} FERC. “Final Rule.” FERC. Docket RM 06-22. 7380.
\textsuperscript{105} FERC. “Notice of Proposed Rulemaking.” FERC. Docket RM 06-22. 43977.
\textsuperscript{107} NERC. “Cybersecurity: Standard CIP-002-1—CIP-009-1.”
regular patches for known vulnerabilities involving critical cyber assets, or use anti-
malicious software protection. 108 CIP-007, one of the few regulations with a prescriptive
requirement (see below), is undercut by the inclusion of “acceptance of risk” language.
In each case, parties covered by the reliability standard have the option of issuing a
statement accepting risk rather than complying with the specific requirements of the
standard. 109

NERC’s support for an acceptance of risk provision follows its general advocacy
of autonomy for regulated entities. In its view, the allowance of acceptance of risk does
not allow an entity to do nothing, but rather allows responsible parties to acknowledge
the inevitable limitations of security strategies. 110 Tampa Electric and Idaho Power join
NERC on similar grounds supporting the provision. In their view, the elimination of all
risk is neither possible nor, when costs are taken into account, necessarily desirable. To
that end, the inclusion of language authorizing acceptance of risks allows regulated
parties to determine an acceptable level of assumed risk. 111

FERC rejects these claims. The inclusion of clauses allowing for the acceptance
of risk undermines the effectiveness of mandatory reliability standards by allowing
entities to opt out of key requirements. 112 The proposed standards allow responsible
entities to, should they so choose, decide to ignore cybersecurity vulnerabilities. As
FERC remarks, there are “no controls or limits on a responsible entities use” of the

---

108 Ibid.
109 Ibid.
Regulatory Commission Staff Preliminary Assessment of the North American Reliability Corporation’s
“acceptance of risk” exception.\textsuperscript{113} The acceptance of risks does not carry with it a pledge to work to mitigate relevant vulnerabilities; indeed it allows entities to ignore such vulnerabilities. Nor is the acceptance of risk tempered by even a modest requirement calling on entities to justify the decision to accept risk or defining narrow circumstances when such a decision would be acceptable. The proposed standards place no obligation to explain why the risk is accepted and, it appears, can be invoked for essentially any reason. As FERC discerns:

\begin{quote}
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.\textsuperscript{114}
\end{quote}

FERC notes that the inclusion of an option to accept risks is particularly troubling in the context of the interdependent power system. Echoing its concerns regarding reasonable business judgment, FERC remarks 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.”\textsuperscript{115} The acceptance of risk by one party, then, impacts others regardless if they too would consent to accept such risks.\textsuperscript{116}

Concerning both the notion of reasonable business judgment and acceptance of risk, FERC offers clear objections. In each case, the inclusion of the provisions is designed to allow regulated parties wide-latitude in deciding if they will comply. The

\textsuperscript{113} FERC. “Notice of Proposed Rulemaking.” FERC. Docket RM 06-22. 43980.

\textsuperscript{114} Ibid.


\textsuperscript{116} Ibid.
offering of nearly unbound discretion runs against the basic notion of mandatory reliability standards and is particularly troubling in light of the interdependence found within electric power. For FERC, implementing meaningful cybersecurity standards requires striking both provisions.

Rewriting CIP Reliability Standards: Specificity and Penalties for Non-Compliance

FERC’s additional major areas of concern relate to the specificity of the proposed standards and the penalties in place for non-compliance. For FERC, CIP standards focus almost exclusively on general, vague, outcomes at the expense of offering specific guidance concerning how to meet goals. In its estimation, the regulations too often amount to exercises in bookkeeping devoid of substantive content. In response, FERC advocates more specificity and, in some instances, prescriptive standards. Additionally, FERC takes aim at the proposed penalties for non-compliance and works to ramp-up the severity of sanctions. As proposed, the penalties are modest and are suitable for breach of administrative standards; for FERC, the proposed penalties amount to little more than a slap on the wrist and do little to encourage compliance. In both cases, FERC seeks to extract significant changes to the proposed standards. In its reading, the over-all generality of CIP standards combined with excessively modest penalties undercuts the effectiveness of mandatory reliability standards. In keeping with its critique of reasonable business judgment and acceptance of risk, FERC argues that, as presented, the CIP standards fail to ensure the security of the bulk-power system.

---

FERC raises significant objections to the generality of the CIP standards. In most cases, the standards refrain from specifying processes or mechanism for achieving identified goals. FERC and NERC split on the proper role of regulations. For NERC, ideal regulations specify outcomes and goals, while deferring to regulated parties to devise the best strategies to meet these goals. In concert with this view, the proposed CIP standards promote general goals and leave the “how” up to each individual entity. For example, CIP-002, calls for the use of “risk-based” methodology in developing cybersecurity plans. Yet, the reliability standard offers no interpretation of what such a methodology might entail. For FERC, the lack of substantive guidance throughout the proposed regulations makes the standards little more than requirements to keep documents on file. Each CIP standard includes various requirements and corresponding measures to be used in assessing compliance. Of the 41 total requirements included within CIP-002 through CIP-009, 38 of the corresponding measures relate to the possession of documents. In these instances, assessing compliance will have nothing to do with the substance of an entities action under the standard, but rather will check to make sure that proper documentation was created and saved. The proposed CIP standards, in this regard, afford regulated parties wide-latitude. Rather than compelling specific action or activities, the standards indicate areas of concern and leave it open to each entity to determine how best to satisfy general reliability goals. FERC, again, views

120 Ibid. 43981.
121 Ibid. 43974.
122 NERC. “Cybersecurity: Standard CIP-002-1—CIP-009-1.”
the grant of exceptional discretion as a fatal weakness of the proposed standards. Without more specific guidance, each entity can interpret the requirements as they wish and, in the process, undermine the vulnerability of the interconnected bulk-power system.124 While FERC recognizes the occasional utility of providing flexibility, it sees that the proposed standards admit too much.

Finally, FERC also takes issues with the penalties assigned for non-compliance. NERC proposes modest penalties for violations of CIP standards. With all reliability standards, penalties for non-compliance are set by NERC according to scales set by what are known as “Violation Risk Factor.” Each requirement and subrequirement of a mandatory reliability standard is assigned a corresponding violation risk factor of either “lower,” “medium,” or “high.”125 These categories affix upper and lower limits of monetary penalties that can be assessed for infractions. After a violation of a requirement, NERC determines on a case-by-case basis a penalty from within the predetermined range outlined by the assigned corresponding violation risk factor. The definition of a requirement’s violation risk factor, then, is quite important: it sets the base and limit on penalties for non-compliance.126 Under NERC’s guidelines, a high violation risk factor indicates that the violation of the requirement could directly contribute or cause system instability and a cascading failure. A medium violation risk factor is

reserved for requirements which, if violated, would impact the interconnected power system abut are unlikely to lead to cascading failure or instability. Lower violation risk factors are assigned to requirements that are only administrative in nature. The violation of a requirement with a low risk factor is not expected to impact the capability of the bulk-power system.\textsuperscript{127} The difference among the fixed range of penalties is significant: infractions of low violation risk factor requirements carry fines between $1,000 and $25,000; violations of medium level requirements range between $2,000 and $335,000; and violations of high level requirements run from $4,000 to $1 million.\textsuperscript{128}

For CIP standards, NERC purposes to assign the overwhelming majority of requirements—85%—lower violation risk factors.\textsuperscript{129} NERC classifies mandatory cybersecurity standards as little more than administrative requirements. For FERC, this interpretation fails to recognize the seriousness of the challenges of cybersecurity or the importance of the requirements.\textsuperscript{130} The requirements in CIP-002 identify and catalogue all critical assets and critical cyber assets, for example, are designated with a low and medium violation risk factor respectively.\textsuperscript{131} Yet, as FERC notes, these requirements are not merely administrative: they control how, or even if, the other reliability standards will be applied. Improper identification of assets jeopardizes the effectiveness of reliability standards in a very real and meaningful way. To designate otherwise, undermines the

\textsuperscript{127} For a description of the different categories of violation risk factors, see NERC. “Request of the North American Electric Reliability Corporation for Approval of Violation Risk Factors for Version 1 Reliability Standards.” 8-9.


\textsuperscript{129} FERC. “Final Rule.” FERC. Docket RM 06-22. 7440.

\textsuperscript{130} Ibid.

utility of cybersecurity regulations.\textsuperscript{132} FERC calls for changes to 43 of the 162 proposed violation risk factors proposed for CIP standards, increasing 6 to “high” and the rest to “medium.”\textsuperscript{133}

FERC’s call for more specific CIP standards, and push for tougher penalties are designed to ensure the effectiveness of the proposed standards. By seeking to make the regulations more precise and creating real consequences for non-compliance, FERC attempts to rework what were lax standards that provided industry with ample discretion to effectively interpret reliability standards as they wish. Underlying FERC’s critique of the proposed standards, both here and in the previous discussion of reasonable business judgment and accepted risk, is recognition of the unique challenges of interdependent security. While in each case NERC and its allies argue for greater flexibility and deference to individual regulated entities, FERC seeks to move regulations closer to uniform standards. In an interdependent environment, non-compliance or uneven, idiosyncratic, interpretation of requirements exposes interconnected players to risk: the weak link undermines all. FERC’s criticisms attempt to move the proposed standards towards more prescriptive standards that can be applied uniformly. Additionally, by ratcheting up the baseline penalties, from a max of $25,000 for most infractions to caps of $335,000 and $1 million, FERC attempts to make sure that reliable standards carry penalties large enough to encourage compliance.


\textsuperscript{133} Note: each CIP reliability standard contains multiple requirements and sub-requirements, each receiving a violation risk factor. For a list of proposed violation risk factors and FERC’s recommend changes, see NERC. “Request of the North American Electric Reliability Corporation for Approval of Violation Risk Factors for Version 1 Reliability Standards.”; FERC. “Notice of Proposed Rulemaking.” FERC. Docket RM 06-22. 44010-44013.
Mandatory Cybersecurity Regulations: The Success of FERC

FERC uses its limited powers to enact substantive changes to the proposed cybersecurity regulations. FERC approves the proposed standards but, at the same time, uses its authority to direct NERC to make significant changes and resubmit. FERC rejects the position advanced by NERC and industry that federal regulators were limited to a narrow review of the proposed standards. Rather than accept that the standards developed through the internal vetting process necessarily are just, reasonable, and in the public interest, FERC takes a hard look at the substance of the standards. The final standards reflect the changes that FERC sought through the approval process. Though it could not draft new regulations, it could approve standards and direct NERC to develop specific modifications. Along these lines, FERC ensured that all references to “reasonable business judgment” and “acceptance of risk” were scrubbed from the eventual mandatory standards and directed NERC to add specificity to vague provisions. Further, FERC cautions NERC that enforcement of standards would not be limited to only the proper maintenance of documents, but that compliance will have to take into account the substantive requirements of the standards as well. Finally, FERC increased the cap on monetary penalties for 43 specific requirements in an effort to ensure that non-compliance could result in meaningful consequences. Taken together, FERC transformed the proposed mandatory reliability standards from vague

134 For example, FERC directed NERC to provide further guidance as to what a risk-based methodology might entail and called for the reworking of CIP-005 to specifically include a direction to regulated parties to adopt defense in depth posture (layering of multiple security procedures). FERC. “Final Rule.” FERC. Docket RM 06-22. 7383-7384, 7394, 7417.
135 Ibid. 7376.
136 Ibid. 7440-7442.
administrative standards containing multiple avenues for non-compliance into something far closer to comprehensive regulations.

Conclusion: Regulatory Activism and Security

The development of the first mandatory cybersecurity regulations for the electric power system is an interesting case. As elsewhere, new interventions struggle to confront the contradictions of deregulation. The move toward unbundled competitive markets for electric power inadvertently introduces new forms of vulnerability into the bulk-power system. After the blackout of 2003, experts reviewing the bulk-power system underscore the need to create cybersecurity standards to confront the threat of terrorism and intentional disruption. Here, there is a moment of reflexivity within which the tension between security and the market becomes an active object of debate and intervention.

Yet, what form such interventions would take remained an open question. While the Task Force and the Energy Policy Act outlined what were basically conservative interventions, FERC undertook regulatory activism to carve out a significant role expanding federal oversight of electric power. FERC seized on the ambiguity within the Energy Policy Act between calls for review in the public interest and deference to industry and offered a broad interpretation of the scope of federal power. FERC flexes its power during the approval process of the CIP reliability standards and successfully transforms what were lax regulations into comparatively tight controls.

The case of cybersecurity regulation provides an interesting juxtaposition with earlier discussions of regulation. Throughout the dissertation, we have consistently seen industry turn toward regulation as a way of reconciling internal problems. In electric
power, for example, utilities seek state regulation as a means of consolidating diverse users and preventing destructive competition. In rail, as well, we see industry actively solicit regulation by the federal government as a means of banning damaging wars over rates. Yet again, during the early 20th century railroads further joined together to seek legally enforceable regulations covering hazmat transportation. The case of cybersecurity regulation appears, at first, to fit this general outline. Industry seeks regulatory authority to enforce mandatory reliability standards and overcome compliance issues associated with voluntary standards. Yet, the development of cybersecurity regulations takes an unexpected turn. Rather than simply enforcing industry-approved provision, federal regulators take their role to review proposed standards seriously. New legislation designed to give industry power, is interpreted by federal regulators in a different manner to support a robust federal role. Once on the books, agencies and the public can turn to new regulatory authority to press industry.

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 regulator against each other. FERC ultimately succeeds and ensures the adoption of standards closer to its view. This success is a type of democratization. Though actors outside the government-industry nexus do not pay a significant role in the drafting of cybersecurity regulations—the debate is mostly inside baseball between industry and government—the establishment of a precedent that stakes out a substantial role for FERC in overseeing
regulatory standards is a potentially significant victory for the public. Robust FERC review ensures that electric power will be open to some outside scrutiny. Likewise, by overlaying FERC’s process of regulatory review, which affords transparent decision-making and opportunities for public comment, on the development of mandatory reliability standard, space for public discussion and review of the operation of electric power is reserved.
## Tables

### Table 7-1: Proposed Cybersecurity Standards: CIP-002–CIP-009

<table>
<thead>
<tr>
<th>Reliability Standard</th>
<th>Title</th>
<th>Purpose</th>
<th>Key Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIP-003-1</td>
<td>Security Management Control</td>
<td>Development and implementation of minimum security management controls to protect Critical Cyber Assets.</td>
<td>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.</td>
</tr>
<tr>
<td>CIP-004-1</td>
<td>Personnel and Training</td>
<td>Provide training and risk assessment for all personnel having access to Critical Cyber Assets.</td>
<td>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.</td>
</tr>
<tr>
<td>CIP-005-1</td>
<td>Electronic Security Perimeter(s)</td>
<td>Identification and protection of an Electronic Security Perimeter(s), within which all Critical Cyber Asset reside.</td>
<td>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.</td>
</tr>
</tbody>
</table>

### Table 7-1: Proposed Cybersecurity Standards: CIP-002–CIP-009 (CONT.)

<table>
<thead>
<tr>
<th>Reliability Standard</th>
<th>Title</th>
<th>Purpose</th>
<th>Key Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIP-006-1</td>
<td>Physical Security of Critical Cyber Assets</td>
<td>Creation of physical security program for Critical Cyber Assets.</td>
<td>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.</td>
</tr>
<tr>
<td>CIP-007-1</td>
<td>Systems Security Management</td>
<td>Definition of methods, processes, and procedures for securing systems determined to be Critical Cyber Assets.</td>
<td>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.</td>
</tr>
<tr>
<td>CIP-008-1</td>
<td>Incident Reporting and Response Planning</td>
<td>Identification, classification, response, and reporting of Cyber Security Incidents.</td>
<td>1.) Develop and implement incident response plan; 2.) Document Cyber Security Incidents.</td>
</tr>
<tr>
<td>CIP-009-1</td>
<td>Recovery Plans for Critical Cyber Assets</td>
<td>Development and implementation of recovery plans for Critical Cyber Assets.</td>
<td>1.) Create and annually review recovery plans; 2.) Exercise plan annually; 3.) Maintain backup information for successful restore of Critical Cyber Assets.</td>
</tr>
</tbody>
</table>

---

Conclusion: The Politics of Critical Infrastructure Protection

The dissertation engages post-9/11 debates over the role of public policy and novel technologies in crafting and maintaining resilient infrastructure networks against the larger historical backdrop of infrastructure deregulation. In proceeding pages, I have followed two related lines of inquiry: First, I examined the relationship between deregulation—the selective and incomplete restructuring of infrastructure regulation in deference to the market—and infrastructure vulnerability; Second, I considered how after 9/11 a cross-section of actors from inside and outside the sphere of traditional national security policymaking intervene in defining the terms on which critical infrastructure protection unfold.

Through an examination of the postal system, freight railroads, and the electric power system, I foreground the important role that politics plays in shaping and defining the contours of infrastructure networks and services. Indeed, as we have seen, the outlines of infrastructures are shaped in critical ways by the structure of regulation. More specifically, regulation, for all of its ascribed failings, built within infrastructures a degree of resilience and provided a hedge against large-scale failure. In the pursuit of new efficiencies, however, deregulation transformed infrastructures and remade them into increasingly vulnerable configurations.

Additionally, the examination of post-9/11 security interventions underscores what is perhaps a surprising or counter-intuitive finding. Here, the perceived threat of terrorism supports the partial democratization of infrastructure governance. Risk, in these instances, serves as a resource to open formerly closed domains and provide
otherwise marginalized groups with access to substantive decisions about the operation of infrastructures. In the examined cases, risk does not serve to support the suspension or diminution of democratic practices, but, on the contrary enlivens particular aspects of practices of democracy.

*The Political Origins of Infrastructure Vulnerability: Revisiting Regulation and Deregulation*

In providing a historical account of the development of the above noted trio of infrastructures, it is clear that the particularities of regulation are encoded and built within the architecture of infrastructure networks. In examining the connections between the regulation and network structure, I note that regulation built into the postal system, the freight rail network, and electric power a degree of resilience that provided a hedge against large-scale failure. In turning to examine the consequences of deregulation within each infrastructure, we see that the pursuit of new efficiencies undercut these supports for resilience and enlarged and created new forms of vulnerability. Rather than seeing current vulnerabilities as the inevitable byproduct of technological development or economic laws, the close examination of the dynamics of deregulation illustrates that contemporary vulnerabilities are linked to a set of contingent political decisions undertaken beginning in the early 1970s to reorder the governance of infrastructure systems in favor of competition. Critically, the partial deregulation of price-and-entry controls over infrastructure operations transformed the architecture of these networks in ways which, ultimately, made infrastructures both more efficient and more vulnerable to large-scale failure. By foregrounding the role of policy and law in shaping infrastructure
networks, I have demonstrated not only how previous regulatory regimes provided a hedge against large-scale failure, but also how deregulation, unintentionally, conspired to create networks beset by vulnerabilities. In placing contemporary debates about infrastructure vulnerability within this larger historical context, I argue for a revisionist view of regulation and note that the vices of regulation—redundancy, distributed processing, excess capacity—now, seen through the prism of security, appear to be virtues.

Regulation, despite its charged failings, supported the creation and maintenance of loosely-coupled and linear systems. In the postal system, Congressional control over key aspects of finances supported a network characterized by distributed command and control. The structure of regulation worked to prevent the adoption of centralized processing and the integration of new forms of automated and mechanical processing. Here, regulation starved postal management of the resources needed to undertake comprehensive reorganization of the processing and distribution of the mail. Distributed processing, despite, or better because of, its inefficiencies prevented large-scale failure. By vesting control throughout the network, rather than at a few central transit points, the network could tolerate isolated failures without jeopardizing the larger system.

Likewise, the regulation of electric power, too, provided checks against vulnerability. A mix of federal and state laws checked the growth of utilities and limited the size and scope of interconnected power systems. As a result, electric power developed into a series of loosely-joined territorial monopolies. Like the postal system, regulation supported distributed, rather than centralized control. Each vertical monopoly
retained control over its operating area and, as a result, limited the potential scale of failure. Interestingly, rate-of-return regulation operated in a fashion quite at odds with the regulation of post. While postal regulation promoted reliance on labor and afforded almost no funds for the development of acquisition of new technologies, rate-of-return regulation pushed utilities to heavily and even over-invest in cutting-edge technologies. Throughout the early and middle decades of the 20\textsuperscript{th} century, regulation supported the development of sophisticated and idiosyncratic technologies of control within electric power. As we have seen, regulation provided a check against cascading failure by both promoting the functional autonomy of vertically integrated system, and by encouraging the development of unique and diverse instruments of control. As in the postal system, regulation supported a loosely-coupled network structure resistant to large-scale failures.

Finally, as we see in the case of freight rail transportation, here again regulation offers overlooked benefits. Regulation over price and entry supports a redundant and expansive network while also discouraging the transportation of large shipments of hazardous materials. The structure of ICC regulation incidentally supported a rail network that discouraged hazardous flows in close proximity to urban centers.

Regulatory support for resilient networks, however, is undermined by deregulation. Post, electric power, and rail, like most infrastructures both domestically and globally, underwent a process of partial deregulation during the 1970s and early 1980s that relaxed price-and-entry regulation in favor of market-based measures. Restructuring, despite its other merits, created or extended new forms of vulnerability that imperiled resiliency. Deregulation created new forms of vulnerability through the
elimination of redundancies and through creation of tightly-coupled material networks prone to single-point, large-scale, failures. Seen in this light, the successes of deregulation are qualified: efficiencies and lower tariffs are achieved at the price of greater vulnerability and brittleness. Likewise, from this angle the inefficacies of the old regulatory model, much decried during their day, appear to offer overlooked benefits. Stable monopoly or quasi-monopoly markets incidentally provided a check against systemic vulnerabilities and supported resilient configurations.

Underscoring the overlooked failings of deregulation challenges certain basic assumptions about infrastructure vulnerability. Rather than seeing infrastructure vulnerability as alternately the product of years of not-so benign neglect, the autonomous logic of technological or economic progress, or a side-effect of private ownership, focusing on the transformations associated with deregulation highlights that vulnerabilities are politically and historically marked: they spring in part from specific political decisions linked to the upheavals in infrastructure governance that occurred during the 1970s and 1980s. Contemporary challenges or problems reflect, then, not inevitabilities but instead the particular consequences of the political project of deregulation.

Further, placing infrastructure vulnerability within the context of deregulation notes a connection between forms of regulatory intervention that are usually seen as separate and distinct. Here, there is a connection between the loosening or removal of economic regulation and the later need for further social or behavioral regulatory interventions—undertaken in the name of security after the terrorist attacks of 9/11—to
make the operation of infrastructures palatable. Thus, contemporary efforts to create policies and practices of security appear as an attempt to reconcile the contradictions of deregulation. Ironically, deregulation requires additional and ongoing regulation in a newly uncertain environment.

**Infrastructure Inversion and the Bureaucratization of Risk**

In turning to the dissertation’s second line of inquiry, I considered how in the wake of the terrorist attacks of 9/11 the vulnerabilities associated with deregulation emerge as objects of political debate and intervention. After 9/11, the tension between economic deregulation and security becomes stark. I have traced what I describe as the bureaucratization of risk: the way the uncertain risk of terrorism is translated and built into regulations and technologies that remake the operation of larger infrastructures. Surprisingly, we see that the perceived risk of terrorism supports the partial democratization of infrastructure governance. While post-9/11 politics in a myriad of other domains appear to be openly hostile toward what we take to be basic democratic norms, a different picture emerges within the narrow confines of infrastructure security that I have examined. Rather than finding the abandonment of democratic practices, we see in some instances its revitalization. The threat of terrorism sparks a moment of reflexive modernization within which the contradictions of deregulation are now taken to be political problem. Critically the salient uncertainty of the threat of terrorism, offers a diverse range of otherwise marginalized groups with access to political decision making.

Post-9/11, the overlooked political dimensions buried within infrastructures, becomes evident and hard to overlook. Risk in this way breeds infrastructure inversion:
infrastructure networks and the political choices inscribed within their structure are pushed to the fore. In post, rail, and electric power, contemporary security debates each wrestle with the contradictions of deregulation. In the proceeding pages, I have considered how, by whom, and to what end new form of control have been created and introduced to infrastructure operation. In doing so, we note some interesting parallels and divergences. In each case, marginalized groups work to appropriate risk as a means of enlarging their role in infrastructure governance. In post, labor is able to find a position in the Mail Security Task Force and engage with management, large volume mailers, and subject matter experts in creating post-anthrax postal security. After having been shut-out of substantive decision making for decades, this is a major achievement. Elsewhere, environmental activists and local and state governments use the threat of terrorism to push for new rerouting requirements for the transportation of hazardous materials via rail. In electric power, federal regulators reject their position on the regulatory sidelines and maneuver into a position to have a significant say in the content of new security regulations. In both freight rail and electric power, the enlargement of regulatory authority makes additional action in the future possible and creates a durable space for public participation. In each case, the threat of terrorism opens infrastructure governance to new parties and allows new voices to legitimate their claims.

Yet, despite these parallels, important differences remain. Significantly, in the postal system, the lasting importance of the democratization of governance is limited. Here, new control technologies ultimately reproduce the traditional hierarchies of power found in postal politics. New systems of control continue the subordination of labor in
favor of large volume mailers. Increased access for labor does not translate into the structure of control technologies. In rail and power, however, risk entrepreneurs—environmental activists, local and state public officials, and federal regulators—are successful in not only gaining access to decision making but also in creating new forms of control more closely aligned with their interests. That is, they are able to have a clear hand in authoring new forms of control.

How then do we account for this difference? In post, the broadening of governance occurred through the informal and ad hoc the Mail Security Task Force, while in rail and power, debates about security occurred within regulatory hearings and involved new legislation expanding the regulatory power of the federal government. These are important differences: The expansion of regulatory authority presents not only regulatory agencies but also the public with new opportunities. In both rail and power, the granting of new regulatory powers covering “security” provides a broad and vague mandate. Now the public can point to new enabling legislation to buttress their claims and demand regulatory action. Even if their initial calls for regulatory interventions fall short of their aim, the possibility of reconsideration at a later date remains. In the postal case, legislation played almost no role. The Task Force was indeed open to participants, but ultimately labor had little recourse to dispute the eventual technologies selected.

There is a real gap here between access, or at least meaningful access, and power. Unlike the examples examined in rail and electric power, labor is not able to leave its mark on the new sets of technologies adopted.

Conclusion: The Politics of Critical Infrastructure Protection
The approaching ten-year anniversary of 9/11 provides an appropriate moment for reflection. In the dissertation, I have examined the way in which the threat of terrorism created a crisis of control for domestic infrastructures. In a modest way, I have examined a small corner of post-9/11 politics. But, as 9/11 continues to recede in memory, how will the changing perception and salience of threat of terrorism translate to infrastructure security? To be sure the transformations linked to deregulation create new possibilities of failure from a range of causes beyond terrorism. But, without question the specter of terrorism has informed and given credence to new security interventions addressing identified vulnerabilities. Will the so-called post-9/11 “new normal” give way and render anachronistic the push for infrastructure security?

As always, how novel risks are translated into new forms of control is an empirical question. Examining the degree to which concerns over terrorism will continue to provide actors with a resource to legitimate claims about infrastructure vulnerability is an interesting question. Moving forward, will the legitimacy of fears of terrorism endure? Or will, the concerns of the past decade over infrastructure security appear as a passing fad? Tracing how new forms of security are built into infrastructures will remain an ongoing project. Yet, equally interesting is the fate of current post-9/11 security interventions. Will the USPS continue to pump resources into sophisticated equipment designed to confront the threat of biological terrorism? Will the new powers assigned to federal regulators to regulate security endure? The dissertation has been in part a sustained examination of how forms of salient uncertainty provide a resource for new
interventions. What happens to these interventions once the perceived threat recedes is an open and interesting question.
References


-- -- --. “USPS Expected to Seek as Much as $10 Billion in Aid From Congress.” *Direct*. Nov. 7, 2001.


Auerswald, Philip, Lewis Branscomb, Todd M. La Porte, Erwann Michel-Kerjan, Eds.  


Clark, Frederick C. State Railroad Commissions, and How they May be Made Effective. Baltimore: Guggenhiemer and Weil Co., 1891.


Douglas, Mary and Aaron Wildavsky. Risk and Culture: An Essay on the Selection of


Friedlaender, Ann F. *The Dilemma of Freight Rail Transportation.* Washington, D.C.:
Brookings Institute, 1969.


Munn v. Illinois, 94 U.S. 113 (1876).


-- -- --. “The Effects of Entry Control in Surface Transport.” Transportation Economics. 1965. NBER. 381-422.


GPO. 2005.


United States. Congress. General Counsel of the Office of Compliance. “Investigation of  


-- -- --. “The Postal Revenue Forgone Appropriation: Overview and Current Issues.”  

-- -- --. “The U.S. Postal Service Response to the Threat of Bioterrorism through the  

RS22864. April 22, 2008.

United States. Department of Commerce (DOC). *The Postal Crisis: The Postal Function  

<http://www.eia.doe.gov/totalenergy/data/annual/>.


by Type, 2007. Figure 3.  


-- -- --. “Open Access Same-Time Information System (formerly Real-Time Information Networks) and Standards of Conduct.” Order 889. 75 FERC 61,078. Docket RM95-9-000. April 24, 1996.

-- -- --. “Promoting Wholesale Competition Through Open Access Non-discriminatory Transmission Services by Public Utilities and Recovery of Stranded Costs by
Public Utilities and Transmitting Utilities.” Order 888. 75 FERC 61,080. Docket RM95-8-000 & RM94-7-001. April 24, 1996.


“What is a Qualifying Facility?”


Changes in Electricity-Related R&D. GAO/RCED-96-203. 1996.


“U.S. Postal Service: Intelligent Mail Benefits May Not be Achieved if Key Risks are Not Addressed.” GAO-09-599. May 2009.


