SHIFTING CONTEXTS AND INCREASING COMPLEXITY IN THE AMERICAN ELECTRIC POWER INDUSTRY

by

Benjamin A. Crabb

A Plan B project submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

Bioregional Planning

Approved:

__________________________  ________________________
Richard E. Toth, M.L.A.        Donald L. Snyder, Ph.D.
Major Professor                Committee Member

__________________________
Joseph A. Tainter, Ph.D.
Committee Member

Department of Environment and Society
College of Natural Resources
Utah State University

2010
# Table of Contents

1. Introduction .................................................................................................................. 1
   Organization of this paper .......................................................................................... 4

2. A primer on electric power systems ............................................................................. 7
   Four functions: generation, transmission, distribution, and consumption .................. 7
   Objective: Meet peak load instantaneously .................................................................. 8
   The characteristics of loads ....................................................................................... 9
   Reliability ................................................................................................................... 13
   Characteristics of T&D costs ...................................................................................... 15
   Summary ..................................................................................................................... 16

3. The electric power industry and Supply-Side Sustainability theory ......................... 18
   Complexity and problem-solving ............................................................................. 18
   Managing for outputs versus managing for system integrity ...................................... 21

4. Industry development, 1880s – 1960s ...................................................................... 24
   Technical context ....................................................................................................... 25
   The earliest systems .................................................................................................. 25
   Critical early improvements in generation and transmission .................................... 25
   Critical early improvements in consumption ........................................................... 29
   Social context ........................................................................................................... 32
   The question of regulatory oversight: monopoly status and rate of return regulation .......................................................... 32
   The question of ownership: public vs. private ......................................................... 34
   Cultural aspects: utility culture ............................................................................... 35
   Cultural aspects: larger society .............................................................................. 38
   Increases in complexity ............................................................................................ 40
   Holding companies and PUHCA ............................................................................ 40
   Social equity and rural electrification ...................................................................... 42
   Chapter four summary .............................................................................................. 43

5. Shifting contexts and increasing complexity, mid 1960s to early 1990s ..................... 46
   Shifting technological context .................................................................................. 47
An end to generator improvements through economies of scale .......... 47
Experiences with nuclear power ............................................................. 49
Reliability problems ............................................................................. 50
Despite increased complexity, reliability problems persist............... 54

Shifting social and macroeconomic contexts ...................................... 57

Environmentalism ............................................................................... 57
OPEC oil embargo, inflation, and declining growth rates .................. 59
The role of utility culture ...................................................................... 63
PURPA and the undermining of the traditional industry structure ......... 66

Chapter five summary ........................................................................ 68

6. Attempts to sustain the industry through restructuring, plans for a “super grid”, and an emerging distributed utility paradigm .......................................................... 71

Restructuring: background and context .............................................. 71
Implementation of restructuring .......................................................... 73
Proposals for super grids ....................................................................... 73
Summary of restructuring efforts to sustain the industry ...................... 86
Model of industry function, 1960s – present ........................................ 87
An emerging industry paradigm: the distributed utility approach .......... 89

Challenges: social acceptability and regulatory support ..................... 93

Chapter six summary ........................................................................ 95

7. Conclusion ...................................................................................... 98

Works cited ........................................................................................ 102
## List of figures

1. The four functions of an electric power system .............................................................. 7
2. Annual load duration curve .......................................................................................... 9
3. Load curves smooth with aggregation .......................................................................... 11
4. Peak period increases with number of customers, with diminishing returns ............. 12
5. Peak load per customer drops with more customers, with diminishing returns .......... 12
6. NERC Interconnections Map ..................................................................................... 13
7. US Electrical Transmission Lines .................................................................................. 14
8. Diminishing marginal returns to complexity ............................................................... 20
9. Thermal efficiency of fossil-fueled steam turbine power plants or units, 1882-1994 .......... 26
10. Incremental cost of generating plants, 1950-1968 ....................................................... 27
11. US residential average real price of electricity, 1890-2000 .......................................... 28
12. Incremental cost of transmission capacity, 1950-1983 .................................................. 30
13. One argument in favor of natural monopoly status .................................................... 33
14. Electricity was viewed as an emancipator and life-enhancer ...................................... 36
15. Model of industry function, 1880s-1960s ................................................................... 45
17. Protesters demonstrating outside the gates of Seabrook Nuclear Power Plant ............ 51
18. NERC regions and balancing authorities ..................................................................... 52
19. Big blackouts come with "indirect" costs ..................................................................... 53
21. A hierarchy inverter: low-level actors can impact high-level processes .................... 60
22. Historically high growth rates have settled around 2% annually .............................. 61
23. The industry's predictions were often simply the extrapolation of past trends .......... 64
24. Maximum and average size of new generating units, 5yr rolling average, 1900-2000 ....... 65
25. ISOs and RTOs .................................................................................................................. 75
26. Increasing complexity during the era of industry restructuring .................................. 76
27. Status of restructuring by state, January 2010 .............................................................. 79
28. Residential customers may not prefer choice .............................................................. 80
29. Model of industry function, 1960s-present .................................................................. 88
30. Traditional utility response to demand increases is to build new facilities ............... 91
31. The value of distributed generation to the utility system .......................................... 92
32. Slow, lumpy capacity upgrades overshoot demand in three ways................................ 92
33. Model of industry function: current opportunities ..................................................... 97
Chapter 1: Introduction

In its description of the electric service industry’s importance to the United States, the Encyclopedia of American Industries declares the industry “runs America” (Gale, 2008). Indeed, the extent and pervasiveness of electric power in modern life can hardly be overstated. The economies of modern industrial societies and the lifestyles of their citizens have, over the course of little more than a century, become highly dependent upon access to abundant and low cost electricity. Electrical power offers its users an odorless, shapeless form of energy that has been applied felicitously to tasks ranging from large industrial processes to small personal whims. Meanwhile, withdrawal of access to electric service severely curtails most forms of social and economic relations.

In the United States, electric power is produced by mostly privately-owned generation plants and delivered to consumers over an extensive transmission and distribution network popularly known as “the grid.” While historically the industry has been highly regulated, based on the conception of utility firms as natural monopolies, in recent years the electric power industry has undergone fundamental changes. Restructuring legislation has opened the generation sector of the industry to competition nationwide. With utility firms no longer vertically integrated, responsibility for the maintenance and upkeep of the grid has been thrown into question.

Meanwhile, new, smaller generators have developed in recent decades that, when located at advantageous locations in the power system and used in concert with modern information and control technologies, can deliver electric service that is economically competitive with that delivered by the very large generation plants and extensive transmission and distribution network traditionally utilized by the industry (Ilic, Black, & Prica, 2007; Sovacool & Hirsh, 2007; National Commission on Energy Policy, 2004, p. 95; Willis & Scott, 2000). This
opens the possibility for an electric power industry of the future that may little resemble that of
the present. These changes are occurring amid growing national and worldwide interest in
curbing the emission of greenhouse gases, of which the electric power industry is a primary
producer. Prospects for the future of the electric power industry, therefore, are characterized
by high levels of uncertainty related to organizational structure, choice of technologies, and the
potential future costs of compliance with environmental regulations.

In this context, the United States is exploring ways to increase its use of renewable
energies for electricity production. Since the areas of highest wind and solar potential are
remote and sparsely populated, government and industry actors have put forward proposals for
the construction of vast expansions to the national grid to bring clean energy from the plains
and the deserts to cities and densely populated coasts. Challenges to the implementation of
such a plan are plentiful. Beyond the question of who will pay and how much such plans would
cost, successfully implementing a vastly expanded transmission network would also require the
navigation of a complex web of regulations, land use restrictions, and jurisdictional variability
In return for the successful navigation of these obstacles, proponents of a so-called Super Grid
claim that expected benefits include a much greener electric power system, improved service
reliability, and expanded opportunities for energy entrepreneurs (Kaplan, 2009, p. 12).

While these goals seem laudable, to attain them will nonetheless require a massive and
coordinated mobilization of resources that will reaffirm our society’s commitment to a
dependence on the nationwide grid for the provision of our electricity needs. While this may
seem an innocuous observation, the structure and function of our energy infrastructure hold
important implications for the structure and function of our social and economic systems. In
particular, we should be well informed about the long-term management implications of
administering a continental-scale electric power grid and about how such a grid would align with the country’s social and technological contexts. Essentially, we should look before we leap.

As a contribution to this effort, this paper takes a historical view of the development of the electric power industry from its beginnings in the 1880s up to the present. Using the theoretical perspective of Allen, Tainter, and Hoekstra described in Supply-Side Sustainability (2003), this paper analyzes how the big, interconnected grid emerged from particular social and technological contexts and how those contexts have changed over time in ways that have challenged the long-term viability of what I call the Big Infrastructure model of the industry. By the “Big Infrastructure” model of the industry, I refer to its traditional physical organization, characterized by large, usually remotely sited central generation plants, and an extensive, interconnected transmission network used to deliver power to consumers.

Let us consider what constitutes big electrical infrastructure. A typical “big” power plant produces power on the order of magnitude of $10^9$ watts (a watt is a unit of power), whereas the average US household’s electricity use is about $10^3$ watts, meaning that a typical thermal power plant produces around a million times more power than the average household uses (Lovins, 2002, p. 35). I would therefore call such a power plant “big,” but big and small are relative terms, of course. Relative to the needs of a factory that requires $10^7$–$10^8$ watts, our power plant would be much less large. For the purposes of this paper, however, when I refer to the “Big Infrastructure” model of the industry, I am referring to the use of very large generation plants, (such as the plant in the example above), and to the continent-wide power grid that delivers electricity from such plants to end users who can be hundreds of miles away.

This paper will argue that the Big Infrastructure model of the industry might be better conceptualized as an output of obsolete socioeconomic and technological contexts, rather than as the technological context from which to drive modern strategies of industry management.
Accordingly, the major emphasis of this paper will be on the broad sweep of the historical development of the industry’s socioeconomic and technological contexts and on the strategies used by the industry to maintain its effectiveness in those shifting contexts. Where necessary, admittedly nontechnical attention is paid to the costs and characteristics of particular technologies.

**Organization of this paper**

Following this introduction, chapter two provides the reader with a basic description of the physical structure and function of electric power systems. Concepts covered include the four basic functions of such systems (generation, transmission, distribution, and consumption), the characteristics of electricity supply and demand, and some costs associated with the maintenance of a transmission and distribution system.

Chapter three gives an overview of the theoretical perspective of *Supply-Side Sustainability*. The chapter presents a discussion of how societies tend to approach problem-solving by investing in organizational complexity, and of how the benefits of increasing complexity, are subject to diminishing returns. This means that a society’s or institution’s pursuit of a management strategy characterized by investments in increasing levels of complexity eventually leads it to a severely constrained capacity to continue solving problems. In order to avoid this scenario, Supply-Side Sustainability theory recommends that management efforts aim to maintain the systemic contexts that produce desired conditions rather than attempting to manage for the desired conditions themselves.

Chapter four covers the historical development of the industry from the 1880s through the 1960s. It describes how great economies of scale in generation plants and natural monopoly status for utility firms provided the basic technological and social contexts of the industry during
this period. From these contexts came the development of the grid and of a utility culture that viewed itself as uniquely, even divinely, qualified to deliver the benefits of electricity to a grateful and growing nation.

Chapter five describes how the industry’s technological and social contexts shifted in the 1960s and 1970s and how the industry responded to these changes. The primary indicators of this change were the industry’s inability to continue driving down costs through the use of ever-larger generation plants, the rise of environmentalism and an activist public, and a reduction in electricity consumption growth rates. Faced with an abundance of new and unforeseen problems arising from these contexts, industry and government invested in increased and new forms of complexity to maintain the functioning of the now-essential Big Infrastructure model of the industry. These included more activist state regulatory commissions, new R&D groups, enhanced reliability oversight, and the Department of Energy. These forms of complexity helped the industry survive the troublesome 1970s and 1980s with its general form intact.

Chapter six describes continuing efforts to sustain the Big Infrastructure model of the industry from the early 1990s to the present. During this period, restructuring legislation facilitated by technological advancements in smaller generation technologies and a cultural shift towards the right opened the generation sector of the industry to much more competition. These changes broke utilities’ vertically integrated control over the industry and raised a host of new problems related to the management of the big, interconnected grid. Aware that investment in the grid has been lacking in recent years, government and industry groups have begun to devise proposals calling for massive grid upgrades to “green” the power industry and to shore up the reliability of electricity service (US Department of Energy, 2008; Kempton et al., 2010). The chapter concludes with an overview of an emerging paradigm of industry organization, called the distributed utility approach, which holds promise as a way to realign the
structure and function of the electric power industry with the social and technological contexts of modern America.

Chapter seven is a concluding chapter which presents a summary and analysis of the historical development of the power industry. The emphasis of the chapter is on the sustainability implications of the trend of increasing levels of complexity needed to sustain the Big Infrastructure model of the industry.
Four functions: generation, transmission, distribution, and consumption

Today’s electric power systems can be described by four basic functions: generation, transmission, distribution, and consumption (see Figure 1).

Figure 1: The four functions of an electric power system: generation, transmission, distribution (the distributed line is labeled as “Feeder” here), and consumption. (Willis, 2004)
The amount of power that the generation function (in colloquial terms, a system’s power plants) can produce is called the system’s generating capacity. Generators have historically benefitted from economies of scale: electricity is generally cheaper to produce in large quantities. Transmission and distribution (in colloquial terms, a system’s power lines) are the means by which electricity is delivered from generator to consumer. A higher capacity power line can transmit more electricity than a lower capacity line. Transmission lines have high capacity and transmit power over relatively long distances. Distribution lines have relatively lower capacities, transmit electricity over shorter distances, and deliver electricity to the end user. The devices that connect generators to transmission lines, transmission lines to distribution lines, and distribution lines to end-users are called transformers.

**Objective: Meet peak load instantaneously**

The purpose of an electric power system is to deliver to its users a reliable supply of on-demand electric power. Since electricity cannot be economically stored like other commodities, making interruptions in the supply of electricity “instantaneously disruptive” (Lovins & Lovins, 1982, p. 38), electric power systems must be scaled to accommodate the relatively rare occurrence of the highest level of demand on the system, called the system’s peak load (see Figure 2). Meeting the load, whether peak or off-peak, requires balancing the supply of electricity delivered to the system with the demand for electricity being drawn from the system by end users.

Supply and demand imbalances can cause serious problems ranging from performance deficiencies in electrical devices to blackouts and damaged transmission and distribution (T&D) infrastructure. In order to reliably and economically plan for and deliver satisfactory system
performance, an understanding of the characteristics of loads is therefore of great importance to utilities.

The characteristics of loads

A system’s load is the aggregated electrical demand of all of the customers on a system. The shape of a load curve over time depends on those customers’ electricity usage levels and patterns. Let us imagine a hypothetical electric power system. For simplicity’s sake, let us imagine this system’s customers use electricity only to run their refrigerators. With only one customer, the system’s load is equal to the amount of electricity that the single refrigerator requires. Whenever the customer’s refrigerator turns on, the load jumps upward. When the refrigerator turns off, the load drops back to zero. Add another customer, and the load reflects the demands of two refrigerators which do not turn on and off in perfect synchrony. The load on the system is at its peak during the relatively rare occasions when both refrigerators are running and, likewise, less frequently drops down to zero total load when both units sit idle. In
essence, this load is more diverse since the multiple demands on it are not perfectly coincident. Add more customers, and the load becomes more diverse still. With 10 customers, the chance that all refrigerators will be drawing electricity at the same time (creating a maximal peak load) is probabilistically small. With 100 customers, the chances are practically nil. Applying this concept to the fluctuations of numerous electrical devices operating in rough proportion to the diurnal activity patterns of households, load curves such as those seen in Figure 3 emerge. In these curves, we see that with more customers comes a smoother, more predictable load, while each new customer adds proportionately less and less to the peak load level.

Since peak load determines the size (and therefore much of the cost) of an electric power system, a smooth, predictable load curve that avoids an unnecessarily high peak can render major savings in costs. Management efforts to this end are called “load smoothing” (Willis & Scott, 2000, p. 59).

Historically, two concepts have been employed in load smoothing efforts: load factor and diversity factor. Load and diversity factors are important to our discussion because they are an approximation of an electric power system’s “natural” characteristics, and because the pursuit of benefits associated with improved load and diversity factors was a major incentive for the expansion and interconnection of early electric power systems in the United States.

The first concept, load factor, refers to the percentage of the time that a given load is present on a system. A load factor close to one means that a load is nearly always present, such as would result from an industrial process that runs around the clock. A load factor close to zero means that a load is rarely present, such as might arise from the usage patterns of some residential customers. Utilities prefer load factors closer to one, because this reflects a predictable, dependable load. Utilities can be confident that generation capacity built to serve
this load will be highly utilized. In contrast, generation and T&D capacity built to serve an infrequently-present load will often sit idle and underutilized. The second concept is diversity factor: the degree to which a system’s customers demand electricity at different times. Or, more formally: the ratio of the sum of the individual maximum demands to total maximum demand. Higher diversity factors indicate more predictable loads, allowing more effective and economical system planning.

Figures 4 and 5 represent the dynamics of load and diversity factors, which improve with additional customers on a system. In Figure 4, the duration of the system’s peak load is a proxy for load factor. In Figure 5, coincidence factor (coincidence factor = 1/diversity factor) is a proxy for diversity factor. Of particular note in these representative graphs is the severely diminished physical marginal return to additional customers beyond approximately 100. This suggests that power systems can capture most of the benefits of improved load and diversity.
factors by serving a population of customers *numbering merely in the hundreds*. To put this in perspective with the current scale of the electric power industry in the United States, the three major interconnections in the country (see their spatial scope, Figure 6) each serve tens to hundreds of millions of customers.

![Figure 4: Peak period increases with number of customers, with diminishing returns. (Willis, 1997, p. 66)](image1)

![Figure 5: Peak load per customer drops with more customers, with diminishing returns. (Willis, 1997, p. 66)](image2)
Two more concepts are important to establish with regard to the structure and function of electric power systems: (1) the requirements for system reliability, and (2) the costs of transmission and distribution infrastructure.

Reliability

Modern users of electric power systems demand highly reliable service, which is ensured on a tactical level by an assortment of high-tech safety and contingency equipment, such as circuit breakers, switches, and voltage regulators. These components are designed to maintain stability in operations and arrest isolated failures before they become widespread.
Of greater interest for our purposes is an understanding of a more strategic approach to system reliability. Historically, a basic rule of thumb, known as the “7-10% rule,” has been employed in this regard. Under this rule, no single generator should provide more than 7-10% of the total generation capacity of a system. By following the 7-10% rule, utilities can maintain system functioning despite the loss of any one generator. Through the first eight decades of industry development, when major economies of scale were realized through continual scaling-up of new generating units, the biggest and most efficient generators would have violated the 7-10% rule if they were operated on small and isolated systems. This provided a major incentive to utility managers to pursue system expansion and interconnection. By having access to a shared transmission network, even the biggest generators would not violate the 7-10% rule. Eventually, isolated systems merged into regional grids, then into major continental-scale interconnections, forming the current structure of the transmission grid in the United States (see Figure 7).

Figure 7: US Electrical Transmission Lines (Argonne National Lab, 2009)
Characteristics of T&D costs

A brief overview of transmission and distribution costs will provide us with a basic understanding of the nature of transmission and distribution costs in the electric power industry. A basic understanding of the costs of T&D is important to our discussion because a major distinction in the alternative futures open to the electric power industry is between a large-scale expansion of the T&D network or the pursuit of a more localized and decentralized form of the electric power industry. The account of T&D costs that follows therefore focuses on the costs that would be most salient to decision-making concerning these alternatives.

The first two “laws of T&D” from the Power Distribution Planning Reference Book are relevant to this sort of overview:

1. It is more economical to move power at high voltage.
2. The higher the voltage, the greater the capacity, and the higher the aggregate costs.

“Thus, high voltage lines, while potentially economical, cost a great deal more than low voltage lines, but have a much greater capacity. They are only economical in practice if they can be used to move a lot of power in one block – they are the giant economy size, but while always giant, they are only economical if one truly needs the giant size” (Willis, 1997, p. 4)(italics added).

These “laws” are important to our understanding of the electric power industry because the formative decades of industry development were characterized by major economies of scale in generating units and high growth rates in electricity consumption, two characteristics that legitimized the use of high capacity T&D infrastructure.

A few additional aspects of T&D costs warrant our attention. Firstly, T&D infrastructure is capital-intensive, and capacity upgrades frequently cost more than the original equipment (Willis, 1997, p. 29). This means that utility planners generally install a capacity that will accommodate forecasted load growth for a service area over a long time horizon. Additionally, a capacity margin of about 20% of the peak load is allotted to T&D components as a safety
margin against unexpectedly high loads. As a result of these conditions, T&D systems are ideally overbuilt for current system needs (often by margins of around 50% above existing load). (Willis, 1997, p. 29). This planning practice works well enough, provided growth in electricity consumption is robust enough to fully utilize the new line capacity, and utility finances can absorb the costs associated with the high initial costs of the line. However, if load growth does not materialize as planned, line capacity goes unused and utilities are burdened with what, in retrospect, are unnecessarily high investments in system infrastructure. T&D planning is not the only utility planning process characterized by this type of risk; a similar analysis can be applied to generation planning, which makes clear some of the potential benefits of small-scale investments over large, lumpy investments in all parts of the utility planning process. Utility planning, in other words, is characterized by uncertainty and high stakes.

Secondly, T&D lines also incur line losses, which can be thought of as a line’s operating costs. Line losses are a function of voltage and distance and are seldom more than 8% of total power transmitted, but the present worth of the lifetimes losses on a major T&D component can “often [be] more than the original capital cost of the unit.... Similar loss-capital relations exist for all other levels of the T&D system, with the ratio of losses’ costs/capital cost increasing as one nears the customer level” (Willis, 1997, pp. 24, 33). These losses can be mitigated by upgrading affected components to offer less impedance to electric current, though these upgrades, as noted above, are expensive.

Summary

This chapter has described the basic components of an electric power system, including generation, transmission, distribution, and consumption. Important concepts from this chapter are: (1) the consumption function as the *raison d’etre* for an electric power system, (2) the diminishing returns to additional users on a system beyond about 100 with regards to improved
load and diversity factors, (3) the 7-10% rule concerning the relation of the size of generators to system reliability, and (4) the capital intensity of T&D infrastructure, T&D upgrades, and line losses.

These concepts will help us interpret the historical development of the US electric power industry and better assess the alternative future development scenarios that could be pursued. Before delving into an exploration of the industry’s historical development, the next chapter describes the theoretical perspective that will inform the description and analysis of the industry’s development. It is to that task that we now turn our attention.
Chapter 3: The electric power industry and supply-side sustainability theory

This paper’s treatment of the electric power industry can be considered an assessment of the long-term viability of management efforts aimed at maintaining the industry’s ability to deliver adequate, reliable, affordable electric service to the nation’s consumers. Viewed from this perspective, this paper analyzes historical problem-solving efforts. To guide this analysis, I use the theoretical perspective of Supply-Side Sustainability, as articulated in the 2003 book of the same name (Allen, Tainter, & Hoekstra, 2003). In the book, the authors develop two major principles which are useful for comprehending the long-term viability of problem solving efforts (Allen, Tainter, & Hoekstra, 2003, p. 59). The first is “to understand the difference between managing for the outputs of a productive system and managing for the integrity of the system” (Allen, Tainter, & Hoekstra, 2003, p. 59). The second concerns the historical development of problem-solving institutions and the effect of complexity on their efficacy. To understand how these principles relate to our analysis of the electric power industry, let us first consider the concepts of problem-solving and complexity and then turn to the difference between managing for outputs versus managing for system integrity.

Complexity and problem-solving

Societies and their institutions are problem-solving systems whose goal is the maintenance of desired conditions (Allen, Tainter, & Hoekstra, 2003, p. 59). To achieve this goal requires solving problems. A primary problem-solving strategy employed to this end has been for a society or institution to invest in complexity. Complexity in this sense refers to “differentiation and organization or to increasing organization” (Allen, Tainter, & Hoekstra, 2003, p. 62) and often looks like new or increased forms of bureaucracy, specialization, and status. For example, after the 9/11 terrorist attacks, the federal government created the
Department of Homeland Security to safeguard the nation’s air transport capabilities and to combat terrorist networks. On a smaller scale, a business faced with the need to grow or adapt to an evolving marketplace may find the need to hire or train new IT, legal, or logistics specialists, or to create new divisions within the company. In order for these new and specialized roles to be useful, they must function in an integrated fashion. This requires an enhanced flow of organized information. Complexity therefore also involves an increased level of production, synthesis, and direction of information.

As a problem solving tool, complexity often works. By increasing the level of structural differentiation in an institution or society, complexity simplifies and makes more efficient the behavior of individuals, who can perform specialized roles in a more integrated organizational structure (Allen, Tainter, & Hoekstra, 2003, p. 64). Complexity can therefore allow a society or institution to address new problems in a more efficient, organized way. A major problem with complexity as a problem-solving tool is that it incurs costs. As they complexify, societies must devote resources to the creation and maintenance of new forms of organization and specialization and to the regulation of behavior and the production and control of information (Allen, Tainter, & Hoekstra, 2003, pp. 62-63). In return, complexity delivers an enhanced capacity for solving problems, such as improved means of producing resources, processing information, or defense. Complexity can therefore be conceptualized as an investment.

Following an economically rational approach, societies tend to initially invest in relatively simple, inexpensive forms of complexity that allow a more effective exploitation of readily available resources (Allen, Tainter, & Hoekstra, 2003, p. 65). The initial high returns on these investments legitimize the complexity used and contribute to a positive feedback effect that enables further exploitation of available resources. Eventually, however, new problems tend to arise that the initial forms of complexity are unable to effectively solve, such as the exhaustion of the most
readily available resources. To address these problems, new forms of complexity are devised, but incur higher costs and lower returns relative to earlier forms of complexity. As a society or institution invests in additional and new forms of complexity, more and more of its resources must be devoted to paying for the complexity itself, while society gains proportionately less and less problem-solving capacity in return. Allen, Tainter and Hoekstra characterize this process as the problem of diminishing returns to complexity (2003, p. 66) (see Figure 8).

The problem of diminishing returns to complexity holds significance for our understanding of the long-term viability of the electric power industry because problem-solving systems characterized by diminishing returns will eventually reach a point – barring the development of new energy subsidies or a more efficient use of existing resources – at which they are no longer capable of effectively meeting new challenges. This corresponds to the point C2, B2 in Figure 8. As the problem-solving system continues to invest in increasingly costly and

Figure 8: Diminishing marginal returns to complexity. (Allen, Tainter, & Hoekstra, 2003, p. 151)
diminishingly effective complexity, it becomes vulnerable to collapse: a drastic and rapid simplification of structure and the inability to continue producing desired conditions, goods, and services. This corresponds to point C3, B1 in Figure 8. Since access to electric power is too valuable of a service for modern societies to do without, an electric power industry that becomes severely constrained in its problem-solving capabilities will likely be the beneficiary of major governmental subsidies, preventing a collapse, but incurring high costs to society. An electric power industry that can continue to deliver desired services to society without needing to depend on direct social subsidies is therefore desirable, and to plan for a future in which the industry can do so requires us to analyze the industry with an appreciation for the role of complexity.

Managing for outputs versus managing for system integrity

From our discussion above, the problem of diminishing returns to complexity would appear to be a nearly insurmountable roadblock to the attainment of management efforts that are viable over the long term. However, *Supply-Side Sustainability* provides perspectives on how managers can align their efforts to, as much as possible, take advantage of services that are freely available from the systemic contexts in which management efforts take place. In this way, management efforts are subsidized by the natural system. This relieves managers from investing in burdensome complexity aimed at ensuring the production of ever-fluctuating levels of outputs. To focus on maintaining a given level of outputs is to constantly mitigate myriad small problems, leading inexorably to increased levels of complexity. In contrast, management efforts that align with the constraints and processes of the larger natural system will be less able to strictly manage the precise levels and types of outputs but will benefit mightily from the decreased amount of overall complexity required.
As an example of this approach, Allen, Tainter, and Hoekstra describe cottonwood forest restoration efforts on the floodplains of the Rio Grande river in central New Mexico (2003, pp. 19-20). Instead of attempting to sustain the forest through planting of individual trees over thousands of acres (which would require a highly organized and costly campaign), a supply-side sustainability approach would utilize natural overbank flooding that occurs each spring, which decays detritus and allows a new generation of cottonwood seedlings to take root.

Another example comes from ancient history (Allen, Tainter, & Hoekstra, 2003, pp. 132-136). Faced declining revenue for administering its empire, the Byzantines of the 7th century devised a way for its expensive army to support itself. The Byzantine imperial family gave away tracts of land to soldiers in exchange for hereditary military service, while halving military pay and eliminating many forms of administration and social services. As a result, the population was relieved of extremely burdensome taxes, and a new class of farmer-soldiers was formed which had obligations not to wealthy landowners but to the state. With sufficient surpluses from their farmlands, farmer-soldiers could afford to have others work their fields while they took part in military campaigns. Fighting for their own lands, they also had greater incentive to repel invaders. By aligning the interests of the peasantry with the interests of the empire, the Byzantines avoided many of the costs of burdensome complexity and gained sustainability until the neglect of the peasant-soldiers contributed to a slow decline from the 11th to the 15th centuries.

This concept is relevant to our analysis of the electric power industry in several ways, which I have alluded to in the above discussion of complexity and in the treatment in Chapter 2 of the characteristics of electric power systems. During the industry’s formative decades from the 1880s to the 1960s, great economies of scale in generation plants and the existence of monopoly-based regulation of utility firms served as the most important aspects of the
industry’s technical and social contexts. From these contexts arose naturally the big, interconnected grid and a sustainable business model called the grow-and-build strategy.

During the 1960s and 1970s, however, the industry’s technological and social contexts began to fundamentally change, while the structure of the big grid and the ingrained utility culture of the industry’s formative years did not. As a result, management strategies aimed at sustaining the Big Infrastructure model of the industry since the 1960s have been characterized by increasing levels of complexity. While this research cannot definitively establish a trend of diminishing returns to this complexity, it does qualitatively describe intensifying management efforts that have been able only to maintain the status quo of electric reliability and price.

In the chapters that follow, I present a historical analysis that attempts to illuminate the social and technical contexts in which the electric power industry developed and the increasingly complex efforts aimed at sustaining the industry. I separate this analysis into three parts: the establishment of the Big Infrastructure model of the industry from the 1880s to the 1960s, a period of technological stasis and social and economic changes from the 1960s to the early 1990s, and the years from the early 1990s to the present, characterized by attempts at industry restructuring and the development of new technologies facilitating new conceptualizations of industry structure and function, such as the concept of the “distributed utility.”
Chapter 4: Industry development, 1880s-1960s

As outlined in chapter two, any electric power system must perform a set of basic functions: generation, transmission, distribution, and consumption. In the first section of this chapter I describe the technical context from which the Big Infrastructure model of the industry emerged during the period from the 1880s to the 1960s by outlining the essential trends in these functions. Emphasis is placed on the increasing scale and efficiency of generators, and the importance of knowledge advancements concerning the industry’s cost structure.

I then examine the social context in which the industry emerged by describing some aspects of the business and cultural environment of the times. I focus on the concepts of high rates of growth, rate of return regulation, and natural monopoly status as the essential components of the business context. This context also helped give rise to a unique utility culture. Early forms of complexity such as the Rural Electrification Administration and the Public Utilities Holding Company Act of 1935 were characterized by high returns, bringing the benefits of electrification to rural areas and helping ensure socially acceptable business practices by the industry.

The scenario that emerges is one of generally positive feedback loops: early advancements in generation, transmission, and accurate metering allowed utilities to manage their technical risks and profitably expand their service areas, while lowering costs for customers. A satisfied public and rate of return regulation allowed utilities to manage their organizational or business risks and incentivized system growth, which was in turn fueled by continual engineering improvements in generators.
Technical context

The earliest systems

The first electric power system in the United States was developed by Thomas Edison and installed in New York City in 1882. Though imperfect, it served as the model for other Edison systems that appeared in several cities in the 1880s. For our purposes, only the most basic characteristics of these systems are of interest. The generators used were loud, inefficient machines, capable of providing only a relatively small amount of power to a relatively small, isolated area (Hirsh, 1989, p. 16). Even if generators had been more powerful, transmission capabilities were very limited. Since early systems used direct current in their transmission lines, long-distance delivery was impractical due to the inability of economically raising and lowering the voltage at which electricity was transmitted.

At the sales end of the business, customers were charged on a per-bulb basis (the earliest systems were used almost exclusively for lighting), without regard to how much electricity the customer used or when they used it (Hirsh, 1989, p. 17; Platt, 1991, p. 79). This produced the then-perplexing result of decreasing earnings with increasing electricity usage, a riddle that would later be solved by an improved understanding of how peak demand and load and diversity factors function as major determinants of overall system costs.

Critical early improvements in generation and transmission

Aware of the enormous commercial and technical potential of electricity, engineers set about improving the performance of the earliest “dynamos.” Generators were soon manufactured that were less noisy, more efficient, and larger, and which could produce electricity at a lower unit cost. Following conservative engineering practice, manufacturers produced machines that enjoyed economies of scale in such prodigious fashion that it was
thought the trend would continue indefinitely. In the words of a GE engineer in 1930, “The art is still progressing very rapidly, and the indications are that it will continue to do so indefinitely” (Hirsh, 1989, p. 38). The engineer’s optimism is hard to criticize: generating units enjoyed continual efficiency improvements and declining marginal costs in power production from the 1880s through the mid-1960s (see Figures 9 and 10). In order to realize the benefits of these improved generators, however, the transmission function had to evolve to more efficiently deliver power over longer distances.

Figure 9: Thermal efficiency of fossil-fueled steam turbine power plants or units, 1882-1994. (Hirsh R., 1999, p. 57)
In the 1890s, the transmission function was revolutionized by the development of transformers and the use of alternating current. By enabling system operators to raise and lower the voltage of electricity, the use of these technologies meant that power could be economically delivered much farther from generation plants than was previously possible. This allowed utilities to access a broader base of customers, improving load and diversity factors and making feasible the use of very large generation plants that produced power at diminishing marginal costs (Hirsh, 1989, p. 20). In short, improved “transmission facilities turned scale economies on the plant level into scale economies on the system level” (Hirsh, 1989, pp. 20, 44).

The use of long distance transmission lines also helped consolidate power into the hands of a few big utility companies, whose cheaply produced power could drive out local power companies who could not afford the biggest, newest, and most efficient generators. As transmission lines snaked farther from generation plants, previously isolated systems were
interconnected in order to boost overall system reliability and to access the cheap power available on the developing grid. By about 1940, most of the country received its power from one of three interconnected regional transmission grids (Hirsh, 1989, p. 142; Nye, 1990, p. 388).

The grid can therefore be conceptualized as a form of complexity that arose in response to accessible, high quality power from continually improving electric generation plants. The grid simplified and constrained the activities of localities since purchasing power from the grid was now easier and less expensive than producing it locally. Coordination and maintenance of the grid also required the development of social complexity. Organizations called power pools formed that oversaw the coordination of system planning and operation (Casazza & Delea, 2003, p. 98). These early forms of complexity were of high return, bringing electric power at decreasing costs to nearly every corner of the country (see Figure 11).

Figure 11: US residential average real price of electricity, 1890-2000. (Lovins, 2002, p. 21)
The success story of the integration of power systems into a large, interconnected grid was driven largely by the decreasing costs of generators and by the improvements to load and diversity factors that came with an expanded geographical scope. However, the very success of economies in generating units, and the attention that was paid to this function of the industry, helped obscure the costs of transmission and distribution. As a result, data on full costs of T&D infrastructure, especially with regard to maintenance and operation, are scarce (Lovins, 2002; Willis, 2004), but a few sources help us outline basic trends. After improvements in transmission technologies “throughout the power industry’s first several decades” (Hirsh, 1989, p. 43), at least one source reports data that suggest that the marginal costs of transmission infrastructure were non-diminishing from 1950 through the early 1980s (see Figure 12) (Hyman, 1983). This suggests that while early advances in transmission technology helped enable the development of the grid by turning economies at the generation plant level to economies at the system level, the costs of additional grid construction from 1950 to 1980 were mitigated by neither decreasing marginal transmission costs nor, beginning in the late 1960s, from decreasing marginal generation costs (Hyman, 1983), undercutting the technical context from which the grid emerged.

**Critical early improvements in the consumption function**

Recall that early utility managers were perplexed by the problem of diminishing earnings with increasing electricity usage. A more complete understanding of the concepts of peak demand and load and diversity factors, and the development of accurate metering technology, helped the industry resolve this problem:
“The demand meter showed [utility magnate Samuel] Insull that electric companies did not work “just as the gas companies do.” The storage tanks of the gas companies allowed them to even out their production schedules and make maximum use of the equipment on a twenty-four hour cycle. In this way, gas companies could keep their capital investments in central station machinery to a minimum, because there was no need for expensive but little-used equipment to meet brief periods of peak demand on the system. In contrast, electric companies had to keep an instant balance between demand and supply or suffer service blackout and equipment damage….

Helping Insull break free of his mentor’s teachings, the demand meter gave him the insight to recalculate the equation between the electric utility’s costs and the customers’ bills” (Platt, 1991, pp. 83-84).

By employing a two-tiered rate structure that allowed utilities to charge customers based on their contribution to the system’s peak demand (which determined the requisite scale and overhead costs of the whole system) as well as, at a much lower rate, the total amount of electricity they consumed at all other times, the Wright demand meter allowed a more proper accounting of utility costs while strongly encouraging “every type of customer to use more
energy” (Platt, 1991, pp. 84-85). In essence, Wright and Insull had come to understand an
important part of the technical context of electric power systems:

“Insull... displayed a sophisticated understanding of the primary importance of building the size and diversity of demand on the generating equipment. Only then could the central station achieve true economies of scale and undersell the competition of gaslight and self-contained systems” (Platt, 1991, p. 86).

Armed with an understanding of the industry’s cost structure, utilities went about soliciting industrial, municipal, and other large customers who would boost their load and diversity factors while using electricity in significant amounts (Nye, 1990, p. 236). To utility owners and managers, this strategy resulted in more customers and revenue, fewer competitors, and lower production costs per unit of electricity. To the targeted big daytime users, electricity costs dropped and they could simplify their operations by abandoning their own electricity production projects (Hirsh, 1989; Nye, 1990).

Utility companies were not only interested in capturing big industrial and municipal customers. For economies of scale to be realized and big technologies to be used, systems had to experience “absolute growth in capacity requirements” (Hirsh, 1989, p. 80). This meant encouraging consumption across the board, which was embarked upon through advertising and glorifying the labor-saving potential of electricity use (Nye, 1990). Manufacturers such as GE and Westinghouse, meanwhile, began developing a host of consumer goods such as lights, appliances, and motors that could be integrated into the daily lives of American homes and businesses. People found these devices convenient, symbols of status, and increasingly affordable. Electricity use soared, growing at an annual average rate of 12% from 1900 to 1920 and at 7% from 1920 to 1973 (Hirsh, 1989, p. 82).
Social context

Given the appropriate technical context (improving generators and the capacity for long distance transmission) and an understanding of its structure and function (load and diversity factors and the need for growth in electricity consumption), electric utilities could by the early 1900s generate, transmit, and charge for electricity in such a way that they could profitably expand operations to take advantage of improvements in generating unit economies of scale. Because electric power infrastructure is highly capital intensive, however, for utility companies to be profitable required a social context in which they could control their risks and be reasonably sure of an acceptable return. This required the resolution of questions of regulatory oversight, and of public versus private ownership. Additionally, the electric power industry developed in such a way that utility managers and the general public developed strongly held ideas about electricity and its management. These cultural aspects formed an important part of the social context in which decisions about industry management were made.

The question of regulatory oversight:
Monopoly status and rate of return regulation

Early entrepreneurs and executives in the electric power industry realized that competition between systems resulted in poor overall economic performance due to the industry’s very high capital requirements. Reasoning that the business of delivering electricity was a natural monopoly, they proposed a bargain to state governments. In return for legal monopoly status over a service area, private utilities would submit to regulatory oversight to ensure they were serving the public interest (Hirsh, 1989). Appreciative of the economic arguments, state governments accepted the bargain and began in 1907 to form public service commissions to regulate the utility companies (Rudolph & Ridley, 1986, pp. 38-40; Hirsh, 1989, p. 22).
Figure 13: One argument in favor of natural monopoly status: In New York City, “there was a vast and intricate network of wires over all of the city” (Nye, 1990, p. 47). Overhead Telephone and Telegraph Wires on Broadway, 1890 (CUNY, 2010).

Regulation was based on state commissions’ assessments of a fair rate of return on the utilities’ capital expenditures. However, public service commissions had small staffs and few resources. Since utility companies were successfully and continually bringing down costs, easy approval of utility rate cases was de rigueur, essentially guaranteeing utilities of their ability to recoup costs and make a fair profit. Their risk exposure essentially eliminated, utilities set about pursuing a grow-and-build strategy to expand their service areas and bottom lines.
The question of ownership: public vs. private

Whereas the electric power industry developed in the Soviet Union through the forced centralized planning of the state (Lenin famously declared that “Communism equals Soviet power plus the electrification of the whole country” (Read, 2005, p. 230)) and in Europe in nationalized forms under the umbrella of the welfare state (Nye, 1990, p. 299), in the United States the industry took on a decidedly more decentralized and privatized approach in congruence with America’s more decentralized governance structures and strong capitalist ideology (Nye, 1990, pp. 139-140). In the US, it was local governments who held the power to permit private businesses within their boundaries, and planning across jurisdictions was difficult (Nye, 1990, p. 139). This made the development of large, interconnected, and publicly-owned electric systems difficult (Rudolph & Ridley, 1986, p. 47). Moreover, city governments struggling to provide services for their growing populations often did not have the means to fund a capital intensive electric utility system, while private firms had a profit incentive to do so (Nye, 1990, p. 302).

Nonetheless, in the early decades of industry development, publicly owned municipal systems were competitive with private utility firms. There are at least three reasons for this. Firstly, generators in the industry’s early decades were not huge, making them affordable to many municipalities. Secondly, municipal systems at least ostensibly aligned the values of the electric service provider with those of its customers. In light of electric utilities’ treatment as natural monopolies, the protection afforded to the public by a utility whose mission was to serve the public, not maximize profit, meant that the municipal system could avoid the complexity associated with external regulation. Thirdly, municipal utilities were not obliged to pay dividends to their shareholders, allowing them to raise debt capital at a lower cost than privately owned utilities. This advantage was undermined by the financial panic of 1907, in
which investor confidence in municipal bonds was destroyed just as state commissions were being formed and electric generators were growing bigger, more efficient and, from the standpoint of overall capital costs, more expensive (Hirsh, 1989, p. 22). Despite spirited efforts around the country to protect, install, or reclaim service territories for municipally owned systems (Rudolph & Ridley, 1986), the general trend of the industry from its earliest decades was toward private ownership (Nye, 1990, p. 140).

Privately owned utilities, while not benefiting from the reduced cost of capital or inherent alignment of values, were advanced by their ability to raise more capital than municipalities. Privately owned utilities could also expand their operations beyond the boundaries of any given city, provided proper permitting. Since profits were regulated at an allowable rate of overall expenditures, bigger utility systems could produce more profits for their owners than could smaller private or public firms, giving private utilities an incentive to expand. Using the financial innovation of the holding company, or by simply taking over subsidiary systems, large utilities could raise the needed capital to acquire the newest, biggest, most efficient, and most expensive generators (Hirsh, 1989). In the face of more efficiently produced electricity delivered via an interconnected transmission system, municipal systems were simply priced out (Rudolph & Ridley, 1986, p. 44).

**Cultural Aspects: utility culture**

As electricity emerged as an important presence in American life, a culture developed around the engineers who designed, built, and employed the new electrical infrastructure. This culture viewed itself as composed of a set of technically skilled elites, who worked for the benefit of society and the liberation of humanity from toil and the primitive past (Hirsh, 1989, p. 32). This can be seen in the advertising campaigns of utilities from the period (see Figure 14), as
Figure 14: Electricity was viewed as an emancipator and life-enhancer. Billboard posters for the Public Service Company of Northern Illinois, 1924, and a newspaper advertisement for the Commonwealth Edison Company, 1925 (Platt, 1991)
well as in the proclamations of those engaged in the electric power industry. For example, a booklet of the Jovians, a fraternal society of electrical industry managers, engineers, and technicians, declared in 1899 that “Electricity occupies the twilight between spirit and matter. Electricians are all proud of their business. They should be. God is the Great Electrician” (Nye, 1990, p. 161).

This lofty view of the electric power industry was encouraged by Industrial Age attitudes towards the ability of machines and of the men who designed and operated them to solve the problems faced by society. A 1926 Saturday Evening Post article entitled “Nine Slaves for Every Citizen” concluded that the creators of the machine age, more than any other group, had liberated the common man: “When mechanical slaves replace human drudgery the general welfare of mankind must inevitably be advanced” (Barbour et al., 1982, p. 27). Similarly, at the dedication of a new steam turbine generating plant in Cincinnati in 1925, a utility man, New York Edison’s Owen D. Young, explained its larger meaning:

“This is the way America must solve her problem of maintaining higher wages than any other country in the world and at the same time keeping her goods competitive in foreign markets. We must put more energy back of the worker…. we shall banish from the farmers’ homes drudgery which in the earlier days killed their wives. We have come here to dedicate a power plant – an instrument of utility. Is it only that? Perhaps it is a temple.” (Barbour et al., 1982, pp. 26-27)

The growing presence of electricity in the homes and businesses of America created the need to produce skilled workers to oversee the industry, formalizing and legitimizing engineering as a profession in the United States. Beginning in the 1880s, new schools of engineering were established by many state universities, in which engineers could gain the technical skills and enculturation needed for success in the industry (Hirsh, 1989, p. 32). While “as late as 1885 a self-trained man could hold a responsible position as an electrical engineer, in less than a generation universities such as MIT, Purdue, Cornell, and Michigan produced
thousands with engineering degrees” to meet the “sudden emergence of power engineering” as a profession (Nye, 1990, p. 160).

Considered by government, business, and the populace as the only legitimate planners of the high-tech, important, and burgeoning electric power industry, power engineers filled the ranks of utility management positions from its earliest days (Hirsh, 1989). However, after the establishment of state utility commissions and monopoly status in the first decades of the 20th century, utility companies became less competitive and managerial culture became more entrenched and conservative (Hirsh, 1989; Gordon, 1991, p. 399). After decades of reliable performance based on its dominant design, the electric power industry became viewed as a conservative, unexciting industry whose major problems had been solved long ago (Hirsh, 1989). This resulted in an industry whose managers were ill suited to respond to the challenges that began to change the industry in the 1960s, and also helps us understand the appeal of modern proposals for new incarnations of the grid – super grids – that hearken back to, and reinforce, the industry’s traditional Big Infrastructure model.

**Cultural aspects: larger society**

By the 1920s and 1930s, private utilities’ grow-and-build strategy had made electricity widely accessible to urban and suburban Americans, and the federal government’s rural electrification programs of the 1930s brought power to the nation’s rural population. In short, the Big Infrastructure model of the industry worked well to bring electricity to all Americans because it was the best available way of delivering power at an economical price. Accordingly, while debates raged over the merits and demerits of public versus private ownership of the industry, progressives and industrialists alike agreed that the best way to build an electric power system was “to generate on a huge scale in “superpower” stations and to link large territories
into continuous networks” (Christie, 1972, p. 481). This paper makes the assumption that the very success of the Big Infrastructure model of the electric power industry has caused both dependency upon it (Charles Ross declared of the historical success of electricity by the late 1960s: “Almost imperceptibly, electricity had become the master, not the slave.” (Ross, 1973, p. 48)) and a lack of awareness of its peculiarity and recent historical pedigree. Writing in 2001 about how Americans experience energy systems, James Williams explains:

“ Pipelines, transmission lines, aqueducts, and highways have become so necessary and expected that daily life and ordinary nature is unimaginable without them. Energy systems are quite simply "second nature." And therein lies a large part of the present problem. Economists, businesspeople, and some politicians think of energy as a commodity to be generated in abundance in a low-cost deregulated market, but the average American consumer sees electricity and natural gas, as the novelist D. J. Waldie puts it, "as something real, like air and sunlight." People know intellectually that technology and technological systems are the tools with which they interact with the non human world and that such interactions can be quite rational, but once technological landscapes are in place people fold them so completely into their psyches and day-to-day lives that their interactions with those very technological landscapes confound theoretically rational market behavior” (Williams, 2001, p. 627).

While these uncritical views of energy systems may not characterize the perspectives Americans had of electricity in the first decades of the 19th century, when electricity was treated more as magic and marvel than established and boring (Nye, 1990; Bary, 2009), they nonetheless describe how those perspectives have evolved into present day understandings. These understandings are important to contemporary decision-making because the expectation of continuing to enjoy lifestyles increasingly dependent on electricity in a country protective of its cherished values of political decentralization and individualism, and increasingly forced to confront environmental and climate change issues, will likely require a more critical analysis of energy systems than simply viewing them as “second nature.” This is especially true when considered from the perspective of supply-side sustainability, which suggests the necessity of critical understandings of a system’s context in order to be able to avoid the problem of
diminishing returns to complexity. It is to the increases in complexity that occurred during the early decades of industry development, which produced high returns, that we now turn our attention.

Increases in complexity

By the 1930s, the industry had developed a dominant design that worked well for all stakeholders, based upon the grow-and-build strategy (Hirsh, 1989, p. 25). The success of this design was due largely to its alignment with the social and technical contexts of the time. However, problems soon arose related to the concentration of economic power and need for a level of social equity. In addition to state regulatory agencies, the responses to these early problems can be conceptualized as early forms of complexity arising in the management of the industry. By ameliorating the over-concentration of power in holding companies and by creating new organizations to deliver power to the rural areas of the country, these new forms of complexity had high returns. In later chapters we will explore new forms of complexity that arose in the Big Infrastructure model of the industry that delivered perhaps lower returns. First, a short history of holding companies, the Public Utilities Holding Company Act of 1935 (PUHCA), and the push for rural electrification.

Holding companies and PUHCA

In order to raise the enormous amounts of capital needed for electric infrastructure, and to route more wealth to utility owners, private utility companies developed an innovative financial tool - the holding company - that allowed them to leverage small amounts of assets through elegant pyramidal ownership schemes (Hyman, 1983, p. 79; Hirsh, 1989, p. 24). Developed by larger utilities and financial services firms as a means to help small utilities raise
capital, holding companies would purchase the relatively risky securities issued by a diversity of smaller firms and subsequently issue their own more attractive securities, whose lower risk ratings were derived from the portfolio (or bundle) of the small firm securities (Hirsh, 1989, pp. 23-24). Holding companies also provided valuable engineering and management services to the small utilities, including facilitating the interconnection of contiguous systems for reliability benefits (Hirsh, 1989). In short, this financial innovation allowed small utilities to raise needed capital, provided investors with attractive risks and returns, and facilitated the interconnection of systems, while concentrating economic power into relatively few hands.

While the holding company concept worked well in a number of respects, many holding companies were swept up in the speculative fever of the 1920s. Holding companies were soon engaged in businesses unrelated to the electric power industry, and the electric systems that they did own were frequently geographically scattered, undermining some of the ostensible benefits to member systems of being part of a holding company (Rudolph & Ridley, 1986, p. 53). The economic concentration also continued, and by 1932 sixteen holding companies controlled fully 75% of the power produced in the United States (Moorhouse, 1986, p. 66). Competition between these companies was guided more by short-term financial concerns than by engineering or long-term economic interest:

“As one contemporary Wall Street analyst described it, the holding company directors were “wizards of financial chicanery,” skilled at “double-shuffling, honey-fudging, horn-swoogling and skullduggery.” Needless to say, with Wall Street analysts unable to understand or follow the labyrinthine dealings of the holding companies, state regulatory agencies were hopelessly lost.” (Rudolph & Ridley, 1986, p. 53)

Concerned by the growing influence of the holding companies, Gifford Pinchot, Teddy Roosevelt’s former chief forester and governor of Pennsylvania, observed in the 1920s that

“Nothing like this gigantic monopoly has ever appeared in the history of the world. Nothing has been imagined before that remotely approaches it in the
thorough-going, intimate, unceasing control it may exercise over the daily life of every human being within the web of its wires.” (Rudolph & Ridley, 1986, p. 47)

With holding companies straying farther and farther from providing beneficial services to the industry, the federal government, over the howling protests of industry executives, passed the Public Utilities Holding Companies Act (PUHCA) of 1935 (Rudolph & Ridley, 1986, p. 77). PUHCA forced holding companies to divest themselves of their holdings in companies outside of their primary industry while limiting the scope and risk of their permissible financial dealings. Holding companies were not wholly outlawed, however. Since regulators recognized the benefits to be had in economies of scale, PUHCA allowed holding companies to exist as long as the systems of their subsidiaries were geographically contiguous (Hirsh, 1989, p. 24).

**Social equity and rural electrification**

A shortcoming of the privatized electric power industry was its inability or unwillingness to deliver power to sparsely populated rural areas due to the high costs of T&D infrastructure and low load density. For example, in 1935, 53 years after Edison’s first New York City system went into service, only 11 percent of the nation’s farms had electricity (Rudolph & Ridley, 1986, p. 80; Nye, 1990, p. 299). By this time, electricity had begun to be viewed as, in Franklin Roosevelt’s words, “no longer a luxury, it is a definite necessity” (Nye, 1990, p. 304). The low rate of rural electrification therefore gave rise to concerns for social equity and the preservation of the economic viability of the nation’s breadbasket. Stirred by these concerns and brandishing a New Deal appetite for big public works projects, the federal government in the 1930s enacted legislation and created new organizations to bring the benefits of electricity to the countryside (Hirsh, 1989; Nye, 1990).

These organizations included the Rural Electrification Administration (1935), the Tennessee Valley Authority (1933), and federal power projects such as the Colorado River
Storage Project and the Bonneville Power Project. Rural electrification proved economically feasible due to low cost power available from major federal generation projects and by the creation of hundreds of rural electric cooperatives through the use of innovative low interest, longer-term financing tools subsidized by state and federal governments (Rudolph & Ridley, 1986, p. 83). As a result of these new forms of complexity, by 1945, 75% of the nation’s farms had electricity (Hirsh, 1989, p. 32).

Chapter four summary

During the first eight decades of industry development, a basic structure and function of the industry emerged, containing social and technical components. The interrelationships between these components are seen in model form in Figure 15. By managing their sources of technical and business risk in supportive social technical contexts, utility firms pursued a strategy of growth and expansion that gave rise to the interconnected national electric grid.

During this time period, power engineering emerged as a discipline, and advancements in the scale and efficiency of generators were great. These successes were complemented by important advancements in transmission technologies, which allowed utilities to transmit their ever-more-cheaply produced power ever farther from generation plants, to an electricity-hungry public. With costs declining and growth rates high, the technological context of the industry made construction of the Big Grid a felicitous and rational way of electrifying the country.

While engineering advances made the electrification of the country possible, the utility firms who performed the task did so in a broader social context which determined the particular form of the American electric power industry. Viewing electricity as a nearly magical and perhaps divinely-inspired phenomenon rich with the potential to liberate the common man and
woman from lives of toil, the America of the early decades of industry development was eager and appreciative for access to electric power. Those who worked in the power business were likewise proud to be involved in an exciting and important new industry.

Cognizant of the immense cultural and commercial potential of widespread electrification, industry and government produced several early forms of complexity that would allow the industry to develop in an economical and socially acceptable manner in capitalist America. First, rate of return regulation based on the conceptualization of utilities as natural monopolies created relatively small public service commissions, which allowed private utility firms to deliver the benefits of cheaply produced electricity to consumers without having to engage in costly competition. Rate of return regulation and the original public service commissions can be therefore be considered high-return forms of complexity.

In the 1930s, additional forms of complexity were created to rectify some of the problems that had arisen in industry development: socially inequitable access to electricity, and the extreme concentration of wealth and power through holding companies. These problems were ameliorated by the creation of rural electrification agencies and programs, such as the REA and the TVA, and by the passage and enforcement of PUHCA. The REA, the TVA, and PUHCA demanded substantial administrative coordination and oversight, and did not facilitate the delivery of electricity to as many people as did the creation of Public Service Commissions and the early successes of the discipline of power engineering. In these respects, these efforts can be considered high-return forms of complexity, but perhaps not quite as high-return as the earlier forms.
Figure 15: Model of industry function, 1880s-1960s
Chapter 5: Shifting contexts and increasing complexity, mid 1960s to early 1990s

The fortunes of the electric power industry looked rosy in the halcyon days of the 1950s and early 1960s. The industry’s successful grow-and-build strategy was based on historically dependable high rates of growth, continual engineering improvements, and social approval (Hirsh, 1989). However, technological stasis and societal changes beginning in the 1960s and 1970s fundamentally changed these drivers, causing significant economic and operational problems for the industry. Meanwhile, an ingrained and conservative utility culture which had developed around the traditional grow-and-build business strategy rendered industry leaders ill-prepared to adjust to changing social and technological contexts. New forms of complexity arose to address these challenges. Agencies and regulations were developed to ensure the reliability of the grid, perform R&D, oversee the development of nuclear power, protect the environment, and stimulate improvements in efficiency (Hirsh, 1999; Casazza & Delea, 2003).

While these efforts allowed the Big Infrastructure model of the industry to persist, the new technological and social contexts gave rise to restructuring programs such as Public Utilities Regulatory Policy Act of 1978 (PURPA) and the Energy Policy Act of 1992, which undermined the long-term validity of the traditional organizational structure of the industry and helped set the stage for the deregulation programs of the 1990s and 2000s.

In this chapter, I describe the basic outlines of the changing technical and social contexts in which the electric power industry operated from the 1960s to the early 1990s, present a review of the new forms of complexity that arose to address these problems, and describe how one of these efforts, PURPA, opened the generation sector of the industry to competition. By spurring the development of small-scale generators and undercutting utilities’ monopolistic control over the industry, PURPA made feasible a reconceptualization of the industry based on
economically viable smaller-scale technological components, in contradistinction with the traditional Big Infrastructure model of the industry.

**Shifting technological context**

**An end to generator improvements through economies of scale**

Through the industry’s formative decades, utilities and manufacturers had enjoyed a mutually beneficial relationship based upon the competitive manufacturing sector’s ability to deliver ever-larger and more efficient generating units, which were in turn utilized in the industry’s powerful grow-and-build strategy (Hirsh, 1989, p. 37). Employing these generators became a cornerstone of utilities’ success in delivering ever-cheaper power to customers, and utilities continually pushed manufacturers to produce bigger and more economical units. By the 1950s and 1960s, the pressure to scale up generator size led manufacturing firms to abandon the conservative design-by-experience engineering strategy that had previously undergirded their success (Hirsh, 1989, p. 58).

The design-by-experience strategy had emphasized thoroughly testing incremental improvements to existing designs before moving products to commercialization. In its stead, manufacturers began employing a design-by-extrapolation strategy which bypassed the accumulation of experience in favor of designing new generators based on the extrapolation of past trends (Hirsh, 1989, p. 64). This strategy proved problematic, creating generators with high failure rates stemming from their “sheer size and complexity” (Hirsh, 1989, p. 96) and from the natural limits of the metals with which they were made (Hirsh, 1989, p. 109). As a result, for the first time in history, the marginal cost of producing electricity ceased to decline, indicating that the industry had “forfeited its primary means of improving productivity and mitigating difficult economic problems” (Hirsh, 1989, p. 141) (see Figure 16). Since declining marginal costs of
electricity production – enabled by the use of ever-larger generating units – had been both the cause of, and subsidy for, the construction and maintenance of the continental-scale grid, the forfeiture of this powerful economic force represented a fundamental and problematic shift in the industry’s technical context.

In attempts to ameliorate the problem of the shift in the industry’s technological context, new organizations were formed by industry and government. Industry groups included R&D organizations such as the Electric Research Council (1965) and the Electric Power Research Institute (EPRI) (1972), which were organized to augment the insufficient research of utilities and manufacturers (Hirsh, 1989, pp. 135-136). While providing important services, these groups also left the industry’s traditional assumptions essentially unquestioned: “EPRI managers still
reaffirmed the basic tenets of load growth and the need for large-scale technology to supply the increasing demand for electricity” (Hirsh, 1989, pp. 136, 138). New governmental organizations were created, as well. After World War II, nuclear energy was viewed as a possible new source of cheap electricity and became the beneficiary of much federal support. To oversee its development, the Atomic (1946-1974) and Nuclear Regulatory Commissions (1974) were created, and the Price-Anderson Act (1957) was passed to provide liability insurance to nuclear power developers. In the 1970s, the Energy Research and Development Administration (1974-1977) and the Department of Energy (1977) were created to perform energy-related R&D, with special concern for national security (Casazza & Delea, 2003, p. 147). Despite these new forms of complexity, engineers could not improve the thermal efficiencies or economies of scale of big power plants. Perhaps the most spectacular of these shortcomings was the industry’s experiments with nuclear power.

**Experiences with nuclear power**

Besieged by stasis in improvements to large-scale generators and increasingly compelled to deal with a public concern for environmental issues, utilities turned to nuclear power as a potential solution (Hirsh, 1989, pp. 124, 132). Since the end of World War II, research and development in nuclear power had been the beneficiary of generous federal support, and a famous 1954 prediction claimed that nuclear power would produce electricity in the near future at rates “too cheap to meter” (Cohn, 1997).

Instead of being too cheap to meter, nuclear power plants “became an overwhelmingly complex form of technology as a result of imposed safety regulations and management problems. Exemplifying the lack of simplicity in equipment and management are the 90,000 drawings needed to describe a typical unit, or the 8 million sheets of paper required to support
its quality-assurance program…” (Hirsh, 1989, p. 97). Nuclear plants were also plagued by regulatory burdens at federal, state, and local levels, as well as strident public opposition (see Figure 17). However, since utilities could only recoup their costs on projects by finishing them, at which point the costs would be passed on to customers in the form of a modified rate base (Chase, 1988, p. 50), they would persist in pursuing drastically over-budget nuclear projects. The results were generally poor: in 1984, the Department of Energy reported that 77% of operating nuclear plants cost at least double their originally estimated prices (Hirsh, 1999, p. 172). While nuclear power now provides about 20% of America’s electricity, no new power plants have been ordered since the 1970s (EIA, 2007), resulting in the culmination of a process The Economist has called going from too cheap to meter to “too costly to matter” (Economist, 2001). As TVA director S. David Freeman remarked in 1984:

“[the] whole [utility] industry had blind faith that nuclear was another form of power. They [utility managers] had grossly underestimated the difficulties of the nuclear technology and the safety problems. It was a massive blunder of the technological elitists. There was a worship of nuclear scientists. It’s a situation where I think contemporary history shows that the common sense of the ordinary citizen who was objecting to nuclear power… has turned out to be closer to the truth than the collective judgment of [manufacturers] GE and Westinghouse and the Joint Committee on Atomic Energy…. Participatory democracy came up with a better result than the feudalistic centralized decision making process.” (Hirsh, 1989, p. 165)

**Reliability problems**

Unfortunately for the industry, losing its primary productivity driver was exacerbated by reliability problems stemming from the functioning of the grid. Beginning with the Northeast Blackout of 1965, when a cascading power failure caused 30 million people to lose power over 80,000 square miles of the northeast US and Canada, the massive grid, which had been developed piecemeal over the last eight decades, began to display serious reliability problems. While the industry attempted to downplay the significance of the blackout,
additional large blackouts affected other areas of the country in subsequent years (Hyman, 1983, p. 109; Hirsh, 1989, p. 133).

The industry reacted by creating new organizations aimed at improving grid reliability while attempting to avoid additional government oversight. While relatively small reliability organizations called balancing authorities had evolved as local systems expanded in the early decades of industry development, the large interconnected grid of the 1960s required reliability oversight at a higher level, since the behavior of any actor on the grid would affect all other participants (Casazza & Delea, 2003, p. 98). To serve this need, numerous regional and one national council, today called the North American Electric Reliability Corporation (NERC) (see Figure 18), were created in the late 1960s (Casazza & Delea, 2003, p. 187). The costs associated
with these efforts had not been considered in the original process of grid interconnection and marked the “beginning of an expenditure program that had to be independent of the direction of demand, for what utility people like to call “non-revenue producing” plant” (Hyman, 1983, p. 109).

Despite these efforts, blackouts continued to occur. When a major blackout resulted in looting and riots in New York City in 1977 (see Figure 19), the first provisions for electric reliability oversight were enacted in federal legislation, though these provisions were never implemented (NERC, 2007a). NERC continued to be an industry organization without the ability to enforce its reliability standards and recommendations.
After the September 11, 2001 terrorist attacks and the August 14, 2003 blackout that affected over 50 million people in the eastern US and Canada, the federal government took steps in 2005 to officially safeguard critical infrastructure, and NERC was contracted by the Federal Energy Regulatory Commission (FERC) to perform this duty for the electric power grid (NERC, 2007a). It was not until 2007, when FERC approved NERC reliability standards, that the US electric power industry operated under mandatory, enforceable reliability standards (NERC,
2007a). In the future, additional layers of complexity may be necessary to continue to effectively administer grid functions. For example, FERC is currently (2010) considering consolidating a recent proliferation of balancing authorities to facilitate the incorporation of more electricity from renewable energies into the grid (Stoel Rives LLP, 2010).

A few words about FERC are in order. FERC is an independent agency within the Department of Energy and is charged with regulating interstate electricity transfers, among other duties. Created in 1977, FERC overtook the duties of the Federal Power Commission (FPC), which had been formed in 1935 with the duty to promote, and to sometimes mandate, the interconnection and coordination of power systems (Casazza & Delea, 2003, p. 140). From a larger historical perspective, then, the FPC was created to encourage interconnection at a time when this resulted in big economic and reliability benefits, while the FERC was created to oversee fair economic transactions across the resulting nationwide, interconnected system. In the 1990s and 2000s, in pursuit of more fair and economic transactions, many of FERC’s major management efforts were aimed at creating open and non-discriminatory access to the interstate transmission network so that the electric power industry could benefit from deregulation and competition. As we will see in Chapter six, these efforts have been only marginally successful, and have also raised a host of problems concerning the planning, maintenance, and expansion of the transmission network, among other issues.

**Despite increased complexity, reliability problems persist**

Despite the emergence of numerous reliability oversight organizations, NERC data from 1984-2006 show that this increasing complexity has not reduced the frequency of large blackouts in the US (Hines, Apt, & Talukdar, 2009, p. 5249). To understand why this is the case, let us briefly explore the behavior of the grid with respect to the frequency of large blackouts.
Recent research (Carreras et al., 2003; Carreras, Newman, & Dobson, 2004; Dobson et al., 2007) suggests that the grid displays properties of a complex “self-organizing critical system.” The interactions between the components of a complex system are too numerous to model individually and too few to model using statistical estimates. The interactions are therefore “not effectively computable” (Robinson, 2002). A self-organizing critical system “perpetually steer[s] itself toward a dynamic equilibrium, where small perturbations have long term effects” (Robinson, 2002). In the power grid, this equilibrium is determined (from a technological perspective) by the opposing forces of component loading (due to load growth) versus component upgrades (Carreras, Newman, & Dobson, 2004, p. 1738). These technical processes are, of course, driven by social trends of increased demand for electricity versus the willingness and ability to invest in grid upgrades. When the weakening effects of load growth are not sufficiently matched by the strengthening effects of component upgrades, blackout risk increases.

While most blackouts are small, the “non-computable” behavior of the big grid means that small, localized perturbations produce large blackouts at rates that “are much more frequent than might be expected” (Carreras, Newman, & Dobson, 2004, p. 1739) by traditional risk analysis (Robinson, 2002). These very large failures bring with them the costs not only of lost power and damaged equipment but also very high indirect costs, such as those arising from social disorder (Carreras et al., 2003). Moreover, piecemeal efforts to prevent small blackouts may actually increase the risk of big blackouts (Dobson et al., 2007, p. 10). All of this suggests that effective Big Grid management should address the global dynamics of the system by engaging in “considerable and sustained investments in both generation and transmission” (Carreras et al., 2003).
While this approach may be sound from a technological perspective, even the authors who recommend it recognize that “it is not clear to what extent the industry, regulators, or the public are prepared to spend money to avoid rare events, even if the risk and consequent economic impact of these rare events are high” (Carreras et al., 2003). From the perspective of Supply-Side Sustainability, the approach of making endless investments in grid infrastructure without sufficient socioeconomic support amounts to managing for outputs, rather than context. For example, in the socioeconomic context of an American grid that “is balkanized, with about 200,000 miles of power lines divided among 500 owners” (Wald, 2008), the prospect of coordinated system-wide transmission upgrades is exceedingly problematic. In contrast, an approach more suited to the socioeconomic and technological contexts of the power grid could employ localized generation to alleviate the stress of load growth on transmission components (Dobson et al., 2007), and/or could strive to build a system “comprising a collection of small disconnected regions” which physically could not exhibit large-scale failures (Roy et al., 2001).

While a more detailed discussion of the complex dynamics of the grid are beyond the scope of this paper, the key idea here is that the grid has displayed non-diminishing frequencies of blackouts despite an increase in the level of complexity intended to ensure grid reliability. Additionally, in this section we have begun to introduce the concept of utilizing localized and smaller-scale electric resources as a way to manage electric power systems in congruence with technological and socioeconomic contexts. These are key ideas of the emerging power industry paradigm of the “distributed utility” (Feinstein, Orans, & Chapel, 1997; Lovins, 2002), which we will briefly explore in Chapter six.
Shifting social and macroeconomic contexts

While shifting technological contexts hampered the successful operations of big generators and the big transmission network, social changes likewise undercut the continued viability of the industry’s old grow-and-build strategy. In particular, the rise of environmentalism, the OPEC oil embargo, and a reduction in the growth rate of electricity consumption uniquely challenged the continued viability of the traditional industry paradigm and culture. As a result, while the industry managed to survive the period from the 1960s to the early 1990s with the Big Infrastructure model of the industry intact, it did so while having to accommodate increasing costs associated with environmental protection and social acceptability. One result of the social context of the times was the passage in 1978 of the Public Utilities Regulatory Policies Act (PURPA), which allowed non-utility power producers to enter the electric power business. By breaking the complete monopoly control that utility firms had over the industry, PURPA directly undermined the industry’s organizational structure and encouraged the development of new, smaller generating technologies.

Environmentalism

In the 1960s and 1970s, a growing environmental movement began to challenge the Industrial Age model of the world as a passive “mechanical vehicle” (Deese, 2009, p. 72) to be comprehensively managed by human overseers. In the context of increasing industrial pollution, the arrival of the atomic bomb, and the existential discomfort of the Cold War, various lines of environmental thought argued that many social and environmental problems stemmed from the manipulative, materialist, short term, and inappropriately objective Industrial Age approach to relationships with self, others, and the natural world. By irresponsibly consuming resources for short-term profit or progress, environmentalists warned that humanity was
dooming itself to a future of scarcity, despoliation, and hardship, as well as to an unfulfilling, exploitative present. Accordingly, environmentalists counseled humanity, or at least western nations, to adopt strategies that would conserve, preserve, reserve, or otherwise more prudently manage humanity’s interactions with the natural world.

While environmentalist perspectives ranged from spiritualist to materialist to overtly sociopolitical, the electric power industry experienced the environmental ethos mostly through challenges on three fronts (Ringleb, 1986, p. 212), each of which increased the level of complexity associated with industry management. Firstly, environmentalists became involved in the legislative process at state and federal levels, finding success in the passage of numerous environmental laws and regulations. One author counted the passage of 47 new “Major Federal Environmental Protection Statutes” from 1969-1980, with attendant increases in costs related to environmental compliance, administrative and agency hearings, and litigation (see Figure 20) (Moorhouse, 1986, pp. 214-215).

Secondly, environmentalists challenged siting and operating licenses of generation and transmission facilities. Since siting laws and ordinances exist at local, state, and national levels, agencies at any of these levels had the jurisdiction to “fully impede a new project” (Ringleb, 1986, p. 191). Since the electric power industry relied so heavily on the utilization of huge infrastructure, the ability of individuals or small groups at relatively local levels to block major projects meant that, from a hierarchical perspective, higher-level projects (e.g., a power plant capable of powering tens of thousands of homes) could be managed by lower-level actors (e.g., local government or property owners), an inversion of effective management (see Figure 21) which also undermines the value of organizational complexity (Smart, 2010; Whitehurst, 2010).

An Edison Electric Institute (an industry lobbying group) employee complained in 1980, “A basic unfairness lies in the fact that the energy developer must win every battle, whereas his opponent need win only one” (Ringleb, 1986, p. 192).

Thirdly, environmentalists challenged the rate structure of utilities, contending that the traditional practice of encouraging consumption through declining rate structures was socially undesirable due to the attendant increases in pollution and resource use. Instead, environmentalists sought penalties for higher consumption levels, in direct conflict with the old grow-and-build strategy, which depended heavily on high growth rates in electricity consumption.

**OPEC oil embargo, inflation, and declining growth rates**

In addition to the activism of environmental groups, the industry was also faced with the serious challenges of inflation and the 1973 OPEC oil embargo, which drove up industry costs in all categories. Since the enormous generating units ordered by utilities required long construction times and heavy financing, inflation and frequent construction delays meant large
Utilities had to request rate hikes, which regulators, aware of an increasingly powerful constituency of conservation and consumer advocates, were reluctant to rubberstamp (Hirsh, 1999, p. 175). Rate hikes that were granted raised the ire of all customers, especially those less concerned with the health of the environment than with the health of their pocketbooks (Hirsh, 1989, p. 147).
Meanwhile, increasing industry costs and decreasing purchasing power contributed to a decline in the growth rate of electricity consumption for the first time since the Great Depression. Utility managers, convinced of the virtues of continual growth, mistakenly identified this shift as a temporary aberration (Hirsh, 1989, p. 127). After growing at an average annual rate of 12% from 1900 to 1920, and 7% from 1920 to 1973 (Hirsh, 1989, p. 82), electricity usage grew at an average rate of about 2.4% from 1973 to 2008 (see Figure 22) (EIA, 2009b). Since the combination of high growth rates and improving economies of scale in generators was integral to the early success of the industry, the decline in growth rates signaled a fundamental change in the business context for the electric power industry.

Figure 22: Historically high growth rates have settled around 2% annually. Source: EIA, 2009b.
These changes highlighted a previously underappreciated vulnerability of the Big Infrastructure model of the industry. Since big power plants and long distance, high voltage transmission lines require major outlays of capital, disruptions to the successful deployment of these projects incur great costs for their owners. In addition to the evolution of bigger and more complex generation and transmission systems, and the rise of environmentalism and an activist public, the uncertainties posed by geopolitical and macroeconomic processes represent another hurdle over which big projects must propel themselves.

One could argue that smaller projects would be equally burdened by such challenges, but this is not altogether accurate. Since smaller projects have shorter lead-up times to full deployment, require less financing, fewer resources, and less space than bigger projects, they inherently face less uncertainty and may encounter fewer active opponents. Additionally, the predicted future for which smaller projects are designed is more likely to come to fruition than is the necessarily more distant future predicted by big project designers. For example, enormous power plants are built in expectation of meeting the future need to service an enormous demand for electricity. If such demand does not materialize, then some proportion of a big power plant’s capacity sits idle, ready to serve a need that does not exist.

On the other hand, the construction of a smaller plant represents a more short-term prediction about the future. This short-term prediction is likely to be more accurate than the long-term forecast and, even if it turns out to be inaccurate, the amount of the smaller plant’s capacity that sits idle is small comparatively and therefore incurs fewer economic penalties. While a large plant is sometimes clearly a more appropriate choice than a small plant (such as in a scenario of major generator economies of scale, high growth rates, and a supportive public), a big project nonetheless cannot escape the risks associated with attempting to predict the future
over a relatively long time-horizon. In the context of an uncertain future, big projects are therefore at a disadvantage relative to small projects.

**The role of utility culture**

Caught in a matrix of unanticipated technological and socioeconomic problems, utility companies were largely unprepared to respond effectively. Trained as power engineers and enculturated in the traditional, grow-and-build model of the electric power industry, utility executives felt that the public was unprepared to guide the development of a complex technological system and that utility leaders should be trusted to manage the industry, as they had for decades (Hirsh, 1989, p. 146). Convinced of the continued validity of their traditional grow-and-build strategy, utilities promoted energy usage through declining rate structures and prepared for the incorporation of more big infrastructure by increasing orders with manufacturers (Hyman, 1983, p. 107; Hirsh, 1989, p. 105). These actions were reflective of the utility culture’s assumption that growth was good and big machines were its necessary means. To utility managers, the proof of the validity of this philosophy was in the pudding: nearly a century of declining prices for electricity. As the director of the TVA remarked in 1977 on that organization’s culture, “low priced electricity was almost a religious belief down here... believed as a matter of faith” (Hirsh, 1989, p. 129). During a temporary jump in growth rates to 9% in the late 1960s, utilities panicked, swamping manufacturers with new orders in the late 1960s and early 1970s (Hirsh, 1989, pp. 101-102). When manufacturers couldn’t keep up with orders due to labor disputes and the shortcomings of the design-by-extrapolation strategy, utilities had to accept economic losses and petition for rate increases from state commissions.

One reason for these unhappy circumstances is that, after decades of relative success under the guaranteeing watch of government regulators, industry leaders had neglected to
develop effective monitoring or predicting mechanisms. Instead, utilities’ strategic planning approach was to simply extrapolate past trends and assume that the future would conform to expectations (which, until the 1960s, it largely did) (see Figure 23) (Hirsh, 1989, p. 179). As one utility manager put it, “if you had a straight edge, you were a load forecaster” (Hirsh, 1989, p. 128). In a period characterized by societal support, high growth, continual economies of scale in generating units, and vertically integrated utility firms, planning was non-contentious, unified, large grain, and generally successful. In contrast, the suite of challenges that the electric power industry encountered in the 1960s and 70s rendered planning a contentious and risky process, subject to the conflicting and active participation of many stakeholders and unfolding in an uncertain macroeconomic and regulatory context (Hirsh, 1999, p. 69).

As a result, by the early 1980s, the electric power industry had begun to modify its operations. Some utilities, hoping to avoid the high costs and long time-horizons of big infrastructure projects, initiated a variety of demand-side management and energy efficiency efforts designed to avoid, as much as possible, the necessity of building at all, signaling a stark change from the grow-and-build strategy (Hirsh, 1989, p. 157). When new generation capacity was needed, utilities began building smaller facilities and, after the implementation of PURPA, were sometimes mandated to purchase power from small, non-utility power producers who could undersell the utilities (Hyman, 1983, p. 118; Hirsh, 1989, p. 168). As the biggest generators fell out of favor, the average size of new generators began a trend of decreasing size, which continues to the present (see Figure 24) (Lovins, 2002, p. 25). Hoping to facilitate additional reform in the industry, in 1978 the federal government passed the Public Utilities Regulatory Policies Act (PURPA), which, by allowing non-utility firms to enter the electric power business, undermined the traditional organizational structure of the industry and set the stage for the much more far-reaching restructuring programs of the 1990s.

![Figure 24: Maximum and average size of new generating units, 5yr rolling average, 1900-2000. Big power plants have been fading since about 1970. (Lovins, 2002, p. 25)](image-url)
**PURPA and the undermining of the traditional industry structure**

In the wake of the OPEC oil embargo of 1973, and in the context of increasing American dependence on foreign sources of energy, President Carter vowed to make conservation and increased reliance on domestic energy cornerstones of his energy policy (Hirsh, 1999, p. 73). Accordingly, he presented his National Energy Plan to Congress in 1977. Congressional committees eventually broke the Plan into five separate bills, of which PURPA was one. The most contentious sections of PURPA, which attracted the majority of industry, lobbyist, and Congressional attention, concerned the conversion of power plants to burn domestically abundant coal instead of foreign and increasingly expensive (but cleaner burning) oil and gas (to which, interestingly, utilities had previously been encouraged to convert for environmental reasons), and to amend utilities’ rate structures to encourage conservation (Hirsh, 1999, pp. 74-75, 81). Consequently, a much overlooked section entitled “Cogeneration and Small Power Production” was included in the final bill without an appreciation for what would be its far-reaching effects on the industry’s traditional business model (Hirsh, 1999, p. 88).

Upon PURPA’s passage, the section – Section 210 – created a new class of participants in the electric power industry, called “qualifying small power production facilities,” or more commonly, “qualifying facilities” (QFs). QFs were defined by their relatively small size and by their efficiency, fuel type, reliability, and other characteristics. Cogeneration facilities were a type of QF that was particularly encouraged. In contrast to other kinds of power plants whose waste heat is expelled into the atmosphere, cogeneration facilities use the heat produced in the process of generating electricity for industrial or other purposes: they cogenerate both power and heat. To use the heat effectively, cogeneration facilities have to be located relatively near to where the heat is needed; otherwise the heat dissipates. Cogeneration plants also tend to be small, as this allows the quantity of heat generated to be fully used and allows the plants to be
sited close to end-users. In this way, overall energy efficiency of cogeneration plants can be dramatically increased from that of large scale, electricity-only power plants: the best large power plants have efficiencies of about 35-40%, while cogenerators had efficiencies of better than 50% in the 1970s (Hirsh, 1999, p. 124), and up to 70-90% by the 1990s (Hirsh, 1999, p. 247).

A most important provision of Section 210 compelled utilities to purchase electricity from QFs at their “avoided cost,” or the cost the utility would have incurred in producing the power itself. This guaranteed a market for electricity produced by QFs and exempted them from having to petition state commissioners for sufficient rates of return, as regulated utilities had to do, meaning that QFs could make profits commensurate with their efficiency relative to utility firms (Hirsh, 1999, p. 87).

Drawn to a lucrative new market, independent power producers proliferated in the decades that followed the passage of PURPA: in 1979, non-utility generators controlled only 2.9% of the nation’s power resources; by 1995 they controlled 8.1% (Hirsh, 1999, p. 114). Under the supportive provisions of PURPA and a variety of state-level incentives, research and development in smaller-scale generating technologies increased significantly. As a result, costs for solar, geothermal, fuel cell, and other generating technologies dropped significantly, rendering many of these technologies competitive with large, centralized plants (Hirsh, 1999, p. 115). In 1985, analysts at the Congressional Office of Technology Assessment reported that PURPA-stimulated projects had begun to serve “as the principal test bed for first generation commercial applications of many new generating technologies” (Hirsh, 1999, pp. 115-116). The impact of the success of small-scale technologies is described by industry historian Richard Hirsh:

“PURPA did more than motivate development of novel generating technologies: it spurred creation of radical technologies that reduced the control over electricity production held by utility managers. This impact of the new technologies stemmed from their small scale and cost-effectiveness. The
compact size of quickly manufactured gas turbines, for example, meant that independent power companies could locate prepackaged generation units near load centers soon after the need for power became apparent. Customers did not need to wait for regulated utilities to complete construction of large, centralized plants – plants that often came on line later than expected and at high cost. At the same time, the use of small, localized generators minimized transmission and distribution costs and sometimes allowed customers to employ the by-products of electricity generation for industrial purposes. In this so-called distributed utility network – somewhat of a throwback to the days of industrial self-generation in the early 1900s – customers reduced their dependence on regulated utilities and derived valuable benefits” (Hirsh, 1999, p. 117).

The success of small-scale, PURPA-inspired generation facilities therefore eroded the concept of the necessity of natural monopoly status for utility firms by proving to be cost-effective in competition with large-scale plants that had previously benefitted from overwhelming economies of scale. Since they were owned by non-utility actors, the success of QFs also fundamentally challenged, and began to change, the traditional organizational structure of the industry. These trends would be further borne out in the deregulation programs of the 1990s and 2000s, although these programs also created new problems with respect to the responsibility and capacity for the administration, maintenance, and expansion of the grid, and for the continuing viability of the Big Infrastructure model of the industry, which we turn to in the next chapter.

Chapter five summary

Beginning with the massive 1965 blackout, the period from the mid-1960s to the early 1990’s was characterized by an impressive suite of problems for the electric power industry. In addition to reliability problems on the electric grid, efficiency improvements to generators dried up, manufacturers had difficulties filling orders, growth rates in electricity consumption
declined, and an increasingly activist populace challenged utility plans and operations on several fronts.

These problems reflected significant changes in the social and technological contexts of the electric power industry. No longer could utilities take for granted an acquiescent public that could be treated as a passive player in the industry. No longer could they assume 7% rates of growth in electricity consumption, or the rubberstamping of rate requests by state service commissions. Moreover, a volatile geopolitical and macroeconomic landscape made long-term cost projections uncertain, confusing utility plans for infrastructure with long service lives. In short, beginning in the 1960s, the industry found itself increasingly operating in a social context characterized by an engaged, environmentally concerned populace and an uncertain economic and regulatory environment.

The technological context in which the industry operated also began to fundamentally change during this time. Most importantly, power engineers could no longer find ways to improve the efficiency of power plants by increasing their size, signaling an end to the primary method by which utilities had historically mitigated difficult economic problems. Additionally, the power grid had become so big and complex that large-scale blackouts began occurring with surprising frequency. Finally, a host of efficient, small generators were developed during this time, undercutting some of the historic justification for utilities’ monopoly status and the validity of the Big Infrastructure model of the industry.

To address the problems presented by these changed social and technological contexts, many new forms of complexity were created which had not been necessary just decades earlier. Utility and manufacturing R&D organizations, reliability oversight bodies, the Department of Energy, environmental regulatory and interest groups, and new classes of power producers were all created during the 1960s, 70s, and 80s. Despite financial hardship and some necessary
retooling of basic operating procedures and assumptions, the new programs, laws, and organizations helped the Big Infrastructure model of the industry survive relatively intact. The problems encountered during this period also helped set the stage for the restructuring programs of the 1990s and 2000s, which would introduce important changes to the organizational structure of the industry.
Chapter 6: Attempts to sustain the industry through restructuring, plans for a “super grid,” and an emerging distributed utility paradigm

Restructuring: background and context

In the decade that followed the implementation of PURPA, a flood of small, efficient, and successful PURPA-inspired technologies and actors entered the electric power industry. The success of these entities eroded the historical justification for utilities’ natural monopoly status, which had been based largely on great economies of scale in generating units. It also caused industry experts to begin considering the costs and benefits of encouraging greater competition in the electric power industry (Hirsh, 1999, p. 262). In addition to the technological advances that made these new ideas feasible, broader cultural shifts were underway, as well, which influenced the eventual implementation of broad industry restructuring.

In the 1980s and 1990s, a political environment emerged in the United States that was to some degree a backlash against the cultural trends of the previous two decades. This political environment championed free market ideology and the virtues of individual responsibility. It found intellectual support in the work of influential economists who argued that policy should be based on sound economic principles and free market mechanisms wherever possible, and viewed government oversight of industries as likely to produce inefficient and inequitable results (Hirsh, 1999, p. 230). From the perspective of these thinkers, increasingly activist regulators involved in administering social or environmental policy were attempting to, in the words of Carter Administration economist Alfred Kahn, “play the role of a paternalistic central government planner.” (Hirsh, 1999, p. 233) This role, of course, is in broad disagreement with the “exalted principles of decentralization and personal responsibility” (Hirsh, 1999, p. 225) that characterize the American political and social system.
These free market views were part of a larger cultural shift which, under the Carter and Reagan Administrations, helped deliver various degrees of deregulation – with generally positive results – to the airline, natural gas, petroleum, financial services, railroad, and telecommunications industries (Winston, 1998; Hirsh, 1999, p. 226). Moreover, electric utility systems in England, Norway, and Chile began deregulation programs in the late 1980s and early 1990s, suggesting that the American electric power industry could be eligible for deregulation, as well (Hirsh, 1999, p. 227).

In this cultural, political, and technological context, the federal government passed the Energy Policy Act of 1992 (EPACT92). EPACT92 sought to bring the benefits of a market-based approach to electric power industry management by mandating competition among power generators at the wholesale level (Hirsh, 1999, p. 239) (the wholesale level refers to sales to utility companies. In contrast, the at retail level power is sold directly to consumers). This meant that a new category of non-utility generators, called exempt wholesale generators (EWGs), could enter the electric power business, using the transmission network to sell their electricity to utility companies. Utility companies would then distribute the power to their customers. Moreover, EPACT92 allowed states to, at their discretion, enact “retail wheeling” programs, more commonly known as customer or retail choice programs. Under these programs, customers could choose their power producers from among all of the firms sending power to the grid.

Inspired by economic philosophy that touted the virtues of decentralized competition, EPACT92 brought about major, even fundamental, changes in the organizational structure of the industry and has delivered in some cases significant gains in generating efficiency (Goto & Tsutsui, 2008). These changes, however, were somewhat misaligned with the technological and social contexts of the big, interconnected grid. In particular, by dis-integrating the utility firms
who had previously overseen the planning, expansion, and maintenance of the grid, EPACT92 introduced the need for new organizations to be created to oversee grid planning and operations. Moreover, while virtues of decentralized economic competition inspired the passage of EPACT92, these principles conflict to some degree with the enormous scale of the Big Infrastructure model of the power industry, whose complex configuration of components stretch across multiple jurisdictions and beg for high-level, centralized administration for effective performance. Additionally, the expansion of competition to the retail level has encountered problems associated with the ability of powerful market players to manipulate price of electricity and disturb system performance, and with people’s desire for simplicity, not choice, in their interactions with electricity providers. EPACT92, therefore, while attempting to sustain the industry through the introduction of free-market based management, has to some degree overlooked the larger social and technological contexts of the electric power industry. In the sections below I explore some of the ways in which restructuring has affected the industry, the problems these efforts have created, and the new forms of complexity that have arisen as a result.

**Implementation of restructuring**

EPACT92 was implemented by FERC in 1996 through Orders 888 and 889, which compelled utilities to allow open and non-discriminatory access to the interstate transmission network (Casazza & Delea, 2003, p. 148). This functionally meant that utility firms were no longer vertically integrated through their four functions (generation, transmission, distribution, and retail sales). While utilities could still own interstate transmission lines, they had to (create and) pay new access tariffs in order to transmit power over the lines. These tariffs were to be applied to all electricity producers in the interest of allowing non-discriminatory access to the
transmission network (Casazza & Delea, 2003, p. 149). The system would be administered through an online service called OASIS: Open Access Same-time Information System.

To functionally unbundle their generation, distribution, and retail functions from their transmission function, utilities could pursue one of two routes: they could retain the same corporate structure but strictly separate the costs of the generation, distribution, and retail functions from the transmission function of their firms, or they could go a step further and form separate entities called independent system operators (ISOs) (Casazza & Delea, 2003, p. 167). ISOs would operate the lines for utility firms through the use of an OASIS system and were meant to ensure open and non-discriminatory access. ISOs were created in several parts of the country, but since FERC Orders 888 and 889 did not specify how wholesale electricity markets were to be structured, regional ISOs took different approaches to grid management. This hampered the smooth functioning of the intended national electricity wholesale market.

Moreover, in the years following Orders 888 and 889, it became clear that discrimination still existed in accessing the transmission network. As a result, in 1999 FERC issued a corrective mandate, Order 2000, which sought to remedy this situation.

Order 2000 specified in greater detail the requirements for independent transmission operations (Casazza & Delea, 2003, p. 169) and requested that utilities place their transmission facilities under the control of new bodies called regional transmission organizations (RTOs). RTOs were meant to supersede the fractured policies of the ISOs by encompassing broader geographic areas and having enhanced transmission planning powers. In practice, it seems that ISOs and RTOs are very much analogous institutions, and they have formed an oversight council, the ISO/RTO Council (IRC), to attempt to coordinate their efforts. In any case, implementation of FERC Order 2000 was again highly variable across the country, reflecting the difficulty of trying to manage the interconnected grid in the decentralized political landscape of the United
States. Two main objections emerged from the states with regard to FERC’s attempt to create large, multi-state RTOs (Casazza & Delea, 2003, p. 170). First, state regulators voiced concerns about a loss of autonomous control over the utilities serving customers in their states (Casazza & Delea, 2003, p. 170). Secondly, states that already enjoyed lower-cost electricity were concerned that a multi-state RTO and a nationwide electricity marketplace would allow their low-cost power to be diverted to other regions (Casazza & Delea, 2003, p. 170). As a result, utilities in some states did not join ISOS/RTOs (see Figure 25).

In short, EPACT92 and FERC’s attempts to implement it did not result in a more smoothly run electricity system. Instead, the decentralized political landscape of the country caused different grid management programs to be adopted in different parts of the country, while truly non-discriminatory access to transmission lines proved elusive (Lambert, 2006, p. 98). As a result, the major goal of restructuring – “to bring more efficient, lower cost power to

Figure 25: ISOS and RTOs
the Nation’s electricity consumers” (Casazza & Delea, 2003, p. 148) – remains somewhat unfulfilled. This can be seen in the distinctive trends of the national average price for electricity and the number of “major orders and regulations” issued by FERC in recent years (see Figure 26). While the national average price of electricity has remained essentially unchanged since Orders 888 and 889 took effect in 1996 (EIA, 2009a), FERC has issued a preponderance of new “major orders” aimed at more effectively managing the Big Grid (FERC, 2010), indicating increasing complexity just to maintain system performance.

Figure 26: Increasing complexity during the era of industry restructuring (EIA, 2009a; FERC, 2010)
While a comprehensive treatment of industry restructuring is beyond the scope of this paper, a few more aspects of the process are worth noting. First, the opening of the transmission network to many new, non-utility power producers raised serious questions about the responsibility and capacity for long-term system planning. Previous to industry restructuring, utility planning was a relatively straightforward process performed by one entity, either a utility firm or an association of utilities. This entity would use load forecasts to plan for generation needs, and transmission upgrades could then be planned to accommodate the increased generation capacity. In this way, system planning could be performed cohesively, if not highly efficiently (big upgrades based on aggregated, long term load forecasts always incur efficiency penalties as load grows to fully utilize new capacity) (Casazza & Delea, 2003, p. 107).

In contrast, under restructuring, market forces determine when, where, and how much privately-owned generation capacity is to be constructed. However, responsibility for upgrading the transmission network is unclear, since this infrastructure is treated as a kind of common property. As a result, new transmission capacity under restructuring has significantly lagged behind new generation capacity (Hirst 2000, 2004).

Transmission planning problems are exacerbated by the uneven policies of ISOs, RTOs, and multiple state and local jurisdictions. As industry analysts Cassaza and Delea put it:

“...while the goal of designing a system to facilitate a wholesale power market is commendable, the means for doing so have become extremely complex due, in large part to structural changes being implemented to establish the market.

Transmission planning used to involve dealing with a more or less consistent pattern of power flow from known sources to known load pockets. Even then, in many areas the time it took to plan and get approval for new transmission lines could take many years, even when there were relatively few regulatory jurisdictions involved. Going forward the level of uncertainty has increased dramatically...” (Casazza & Delea, 2003, p. 115)

All of this means that answers to questions of who, where, when, and how grid upgrades will occur are all uncertain. These uncertainties are compounded by the fact that the
industry has not historically emphasized the importance of transmission and distribution costs, focused as early industry leaders were on improving generation (Lovins, 2002, p. 77). As a result, accounting practices related to T&D costs varied widely among utilities, and today no comprehensive dataset exists on the full costs of the grid, especially with regard to maintenance and operation (Lovins, 2002, p. 77; Willis, 2004, p. 24). Additionally, access to current transmission upgrade plans for analysis and review by academics or other analysts is often curtailed due to national security or commercial sensitivity concerns (Hirst, 2004).

Second, electricity markets may prove “uniquely vulnerable to poor operation and even failure” because of the extremely short length of the market interval over which supply and demand must be balanced (Kwoka & Madjarov, 2007, p. 35). The most well known example of this is the spectacular market failures of California’s power system in 2001 due to the exploitation by energy traders of arbitrage opportunities in the fluctuations of energy prices (Blumsack, Apt, & Lave, 2006). Since electric power systems must perfectly balance supply and demand on a nearly instantaneous time scale, the ability of some major market players to significantly affect this balance led to electricity shortages, rolling blackouts, and opportunities for price gauging (Blumsack, Apt, & Lave, 2006). Partly because the Big Grid spans such enormous areas, the failures of the California system allowed electricity prices to soar across the whole western US in the summer of 2001 (Joskow, 2001, p. 378). In the wake of the California experience, numerous states suspended their customer choice programs, since it was this state-level deregulation program that allowed problematic arbitrage opportunities to arise (see Figure 27).
Interestingly, while customer choice has been touted by proponents of deregulation, there is little evidence that customers prefer choice in energy providers (Brennan, 2007). Instead, people seem to prefer simplicity when it comes to energy choices (see Figure 28). This can be seen in the fact that very few customers in states that enacted customer choice programs actively changed providers (Casazza & Delea, 2003, p. 154; Brennan, 2007). Complicated electricity systems can also cause information overload and poor outcomes (Brennan, 2007). This is especially true for small residential or commercial customers who do not want to invest their scarce resources in analysis of potential electricity providers (Brennan, 2007, p. 1617).

Figure 27: Status of restructuring by state, January 2010. Source: EIA
People also generally do not want to be impacted by the presence of energy infrastructure in their local areas. The “Not In My Back Yard” (NIMBY) concept summarizes this attitude to neighborhood, regional, or eminent-domain instances of unwanted land use changes. NIMBYism can be considered a rational, risk averse strategy employed by individuals to deter the possibility of harm to the value of their property (Fischel, 2001) and is particularly comprehensible when viewed with an appreciation of an American society that highly values individual property rights. Effective NIMBY-based opposition by organized local residents can derail or delay infrastructure projects, especially plans for long-distance transmission lines (Benjamin, 2007, pp. 40-41). Should such a project be pushed through over the objections of
state, local, or citizen groups to enable construction of locally unwanted but regionally or
nationally needed transmission lines, a likely result could be strained relations between the
groups involved and an increased likelihood of future problems, legal challenges, or even
sabotage (Lovins & Lovins, 1982). For example, in a protracted protest from 1974 to 1978
against a 430 mile transmission line planned and built from North Dakota to Minnesota,
hundreds of farmers confronted surveyors and police in their fields, then toppled towers and
damaged the lines with rifle fire after legal channels of opposition were exhausted. Despite
their lengthy and costly protests, however, the line was built and became operational (Casper &

Using the Supply-Side Sustainability definition of sustainability, this form of industry
management would be better classified as demonstrating system resilience, not sustainability.
Resilience refers to a system’s ability to adjust its configuration and function under disturbance,
without regard to peoples’ preferences (Allen, Tainter, & Hoekstra, 2003, p. 26). Sustainability,
on the other hand, arises from maintaining desired conditions. Since the desire for electric
service must be weighed against the desire for political protections of individuals and private
property, mandated major electricity infrastructure would, for at least some people, represent
unsustainability. However, it is precisely plans for major new infrastructure that seem to be
dominating proposals for the future of the grid.

Proposals for super grids

As we have noted, while deregulation has enabled a proliferation of new generators to
become involved in the electric power industry, transmission upkeep has not kept pace with the
new additions to generation capacity, especially for large-scale transmission lines that would
connect major regions (Hirst & Kirby, 2002, p. 54; Hirst, 2004; Renner, 2005). While different
policy and economic fixes have been proposed to remedy this situation (Hirst, 2000; Hirst & Kirby, 2002; Renner, 2005; Kwoka & Madjarov, 2007; Contreras et al., 2009; Sauma & Oren, 2009), there seems to be widespread agreement that beefing up the Big Grid is necessary.

This viewpoint is reinforced by industry groups such as the Electric Power Research Institute and the Edison Electric Institute (Lerner, 2003, p. 13), as well as by current and former high level government actors such as former Vice President Al Gore (Gore, 2009), the current Obama Administration (Department of Energy, 2009), and former Secretary of Energy Bill Richardson who, in the wake of the August 2003 blackout, declared that the United States was a “major superpower with a third-world electrical grid” (Suellentrop, 2003). The Big Grid upgrade argument is also present in popular media (Time, 2003; Cook, 2009; Garber, 2009; Johnson, 2009; NPR, 2009). One such article, from Scientific American magazine, touted the benefits of a “Continental SuperGrid” composed of “Super-Cables” (Grant, Starr, & Overbye, 2006). In echoes of the overwhelming confidence of utility managers and nuclear developers who foresaw nothing but benefits and success stemming from the pursuit of ever-larger generation plants, the authors of the Scientific American article betrayed boundless optimism in the wisdom of building a Super Grid. They noted that while “mustering the social and national resolve to create it [the Super Grid] may be a challenge” and that “It is difficult to estimate the cost of a multidecade, multigenerational Super Grid effort,” these costs would obviously be outweighed by the final success of the Super Grid: “one can judge the ultimate benefits: a carbonless, ecologically gentle domestic energy infrastructure yielding economic and physical security” (Grant, Starr, & Overbye, 2006, pp. 78, 83).

Estimating the cost of major upgrades to the national grid over a long time-horizon is fraught with uncertainty, but a preliminary estimate by the American Society of Engineers places the estimated cost of needed grid upgrades at $1.5 trillion dollars by 2030 (Buhrman, Reed, &
Albers, 2009). In at least partial anticipation of such costs, the Department of Energy announced in July of 2009 $750 million to subsidize large transmission projects in the US (Department of Energy, 2009). To facilitate the planning and construction of these lines, NERC has weighed in with a few recommendations. After noting that the entities to be sustained by transmission upgrades are local systems, the reliability oversight body recommends investing in new forms of complexity: “The majority of the proposed transmission projects are for local system support... New regional planning entities, adequate pricing incentives, and improved, streamlined approval processes... must be developed to deal with the need for new transmission lines for an open market” (Hirst, 2000, pp. 78-79).

Judging from the overwhelming support for the notion of constructing a nationwide Super Grid, this author cannot argue with the necessity of such an effort to maintain the Big Infrastructure model of the industry. However, judging from the history of increasing complexity associated with maintaining the grid in changed social and technological contexts, there is the possibility that the new forms of complexity and increasing subsidies necessary for the maintenance of the Big Infrastructure model may contravene the recommendations of Supply-Side Sustainability for successful long-term management viability. Recall that the benefits associated with improved load and diversity factors improve only incrementally beyond about 100 customers. Moreover, given the current availability of efficient small-scale generators and advanced information technology (Watanabe, Kishioka, & Carvajal, 2005), some authors are advocating a consideration of decentralized and localized electric power systems as a means to sustainably provide electric services (Markvart, 2006; Ilic, Black, & Prica, 2007; Morris, 2009; Sebitosi & Okou, 2010). At the same time, to the extent that the costs and beneficiaries of grid upgrade efforts are socially acceptable, successful Super Grid construction does present a potentially sound management strategy; but if other options present ways of
reducing complexity by aligning management efforts with present social and technological contexts, while also avoiding the inflexibilities associated with very large projects, plans to build a Super Grid may be poorly informed.

Two more concepts are worth exploring briefly with regards to the construction of a continental Super Grid. First is the old engineering mindset that assumes that society is a passive actor in the technological/social system that is the electric power industry and that use of ever-bigger tools is a preferred way to solve problems. This mindset can be seen in several accounts of how the industry has historically and contemporarily approached problem-solving. Writing about the culture that took root during the industry’s formative decades, historian Richard Hirsh notes:

“Because of their technical training, high-level executives cherished values that differed from managers who understood only “dollars and cents” concerns.... Big, “neat,” and exciting technologies often caught their imagination, distracting them from purely economic considerations... trained as engineers, these people [utility managers] developed a sense of community traditions that they were unlikely to discard after becoming managers... This background provided a standard approach toward solving problems – the “engineering method” – that attempted to isolate a system and control its variables” (Hirsh, 1989, pp. 71-73).

Accounts written in 2006 and 2007 describe how the engineering method remains the industry’s standard approach to addressing problems by focusing on isolated technological fixes without considering larger, system-wide concerns or societal aspects of the industry:

“Response to the August 2003 blackout followed narrow engineering lines. One company did not follow rules. One RTO did not know what was going on. The solution: obey rules, rewrite rules when required, and make sure that everyone follows them. That does not address the question of whether the grid should be designed and operated in such a way that an operating error in Ohio (if that was the cause) could bring down the entire Northeast. Nor does it ask whether non-grid solutions (such as distributed generation located near consumers or arrangements to cut off designated consumers, with compensation for inconvenience) could provide equal or better reliability at a lower cost” (Hyman, 2006, p. 49).

“Utility engineers make prodigious efforts to avoid blackouts and especially to avoid repeated blackouts with similar causes. These engineering responses to a
blackout... include repair of damaged equipment, more frequent maintenance, changes in operating policy away from the specific conditions causing the blackout, installing new equipment to increase system capacity, and adjusting or adding system alarms or controls” (Dobson et al., 2007, p. 9).

The second concept worth considering in our discussion of plans for a Super Grid is the inflexibility of big plans and big projects. As the electric power industry discovered in the time period from the 1960s to the 1980s, ever bigger power plants could not continue to produce the benefits that their proponents presumed they would deliver. Nuclear power plants proved to be far more complex than expected and social resistance far greater. Cost overruns resulted from the inability of utilities to bring big, inflexible projects on-line, while the decentralized political landscape of the United States thwarted plans for easy expansion and planning. If the country chooses to pursue a major expansion of the grid along the lines of a “Continental Super Grid,” it will embark on a very large project whose outcome and costs cannot be effectively predicted and which may be derailed by any number of unforeseen or unpredicted social or technological problems. For example, one such problem could be “loop flows” of electricity in transmission lines, which can fill up the line’s effective capacity without delivering power to customers (Lerner, 2003, p. 9). Writing about how this phenomenon can create problems for grid expansion plans, science writer Eric Lerner observed in 2003 that

“experts outside the utility industry point to serious drawbacks in the build-more solution... For one, it is almost impossible to say what level of capacity will accommodate the long-distance wholesale trading. The data needed to judge that is now proprietary and unavailable in detail. Even if made available to planners, this data refers only to the present. Transmission lines take years to build, but energy flows can expand rapidly to fill new capacity...” (Lerner, 2003, p. 13).

A perhaps more easily overlooked, and more easily pooh-poohed, risk facing the big grid are the strong electromagnetic disturbances that arise from solar storms which occur approximately every 50 years (Hyman & Hyman, 2006, p. 29; Odenwald & Green, 2008). These disturbances could overheat and destroy big transformers, causing continental scale blackouts
which would require weeks or months of power restoration efforts (Odenwald & Green, 2008). While smaller grids would not be exempt from the effects of such storms, they may enjoy an advantage in time required for recovery, since the whole national grid would not have to be serviced.

**Summary of restructuring efforts to sustain the industry**

Industry restructuring efforts designed to encourage more efficient production of power, more competition, and the avoidance of regulatory activism have created a new crop of problems related to the management of the Big Infrastructure model of the industry. Foremost among these problems include the challenges of equitably administering a market-based approach to electricity provision and delivery, and the major uncertainties related to the responsibility for grid upgrades and maintenance. Both of these problems have social aspects that exacerbate challenges related to their effective management. For example, with regard to administering market-based approaches to electricity provision, states are reluctant to participate in a system which may adversely affect their autonomous control over utility firms and their access to lower than average priced electricity. Likewise, plans to upgrade the grid are burdened by the number and variety of jurisdictions and potentially activist property owners impacted by long distance lines.

It appears that efforts to upgrade the Big Grid will be characterized by the necessity for major government subsidies and new organizations dedicated to handling the problems of transmission planning and regional grid coordination. This unfolding of events lends credence to the proposition that the social and technological contexts from which the grid emerged in the early decades of the 20th century are now significantly changed and that attempting to sustain
the Big Infrastructure model of the industry can be usefully understood, from a *Supply-Side Sustainability* perspective, as managing for outputs instead of managing from contexts.

As has been alluded to throughout this paper, an emerging industry paradigm based upon smaller technologies, greater problem-solving flexibility, and in potentially greater congruence with the decentralized political landscape of the United States, presents a model for industry management that may avoid many of the burgeoning forms of complexity currently burdening Big Grid management efforts. After revisiting our model of industry function, the final section of this chapter will briefly explore some of the basic concepts and planning approaches of this “distributed utility” paradigm.

**Model of industry function, 1960s – present**

Updating our model of industry function, we can conceptualize utility firms as still being driven by the need to control technical and business risks and to earn a reasonable return on their investments (see Figure 29). However, new problems now burden the relationship between the components of the model. We can see that the interruption of the process of improvements to generators through economies of scale and the increase in complexity arising from social requirements for more reliable and environmentally friendly power have contributed to challenges related to organizational and technical risk mitigation, and to the ability of utility firms to be profitable.
Figure 29: Model of industry function, 1960s-present
An Emerging Industry Paradigm:  
the Distributed Utility approach

This paper has described the historical development of the electric power industry with respect to the changing social and technological contexts in which it functions and the increasing levels of complexity undertaken in management efforts to administer the Big Infrastructure model of the industry. While we have touched on the topic of system planning briefly earlier in this chapter, an additional treatment of planning approaches will be instructive in this section to highlight the differences between a distributed utility approach to system planning and a traditional one. The summary of Feinstein, Orans, and Chapel provides a succinct account of the essence of these differences:

“The fundamental idea within the distributed utility concept is that particular local load forecasts can be satisfied at least cost by avoiding or delaying the more traditional investments in central generation capacity, bulk transmission expansion, and local transmission and distribution upgrades. Instead of these investments, the distributed utility concept suggest that investments in local generation, local storage, and local demand-side management technologies can be designed to satisfy increasing local demand at lower total cost” (1997, p. 155).

The key concept here is the scale differences in the level of interest at which the distributed utility and traditional planning approaches focus their attention. While a traditional approach uses aggregated load forecasts and the installation of large central generation plants, a distributed approach is interested in local conditions and costs. In order to pursue this more fine-grained approach to system planning, a distributed utility approach must be able to access and process a greater amount of information related to the area and time specific costs of the power system (Feinstein, Orans, & Chapel, 1997, p. 155). Owing to advancements in information and control technologies, this capability is technologically and economically feasible, although it depends to some degree on the cooperation of the users of electricity (Wald, 2009). By integrating a thorough understanding of a system’s area and time specific
costs with rigorous ways of approaching risk management that are informed by financial and economic theory, such as portfolio theory and real options analysis, the added value that comes from a smaller investment’s greater flexibility can be incorporated into a utility’s analysis of how and when to invest in capacity upgrades (Hoff, Wenger, & Farmer, 1996).

By striving to make electrical resources the right size and located at the right place in an electric power system, the distributed utility approach aims to capture a host of potential benefits, including a reduced reliance on the grid, reduced environmental consequences when compared with traditional power generators, lower costs, and improved reliability (Feinstein, Orans, & Chapel, 1997, pp. 145-146; Carley, 2009). A central concept of the distributed utility approach is the necessity of careful accounting in order to be able to account for all of the benefits that relatively smaller, modular resources can deliver when placed at locations strategically close to end users. A 2002 book by the Rocky Mountain Institute was hailed as the “Book of The Year” by The Economist for outlining 207 supposed benefits of distributed (decentralized) resources, claiming that a proper accounting of their benefits “raises their value by a large factor, often approximately tenfold, by improving system planning, utility construction and operation (especially of the grid), and service quality, and by avoiding societal costs” (Lovins, 2002).

A full treatment of the costs and benefits of a distributed utility planning approach is well beyond the scope of this paper, but a few figures can help us intuitively understand the value that smaller, well placed technologies can deliver to electric power systems. These are Figures 30 and 31, which show how a small distributed generator placed near a local increase in load can help defer costly investments all the way up the chain of the electric power system (Hoff, 1996; Hoff, Wenger, & Farmer, 1996; Kahn, 2008). The distributed generator can also provide “voltage support” to the existing grid, helping shore up system reliability (Hoff, Wenger,
& Farmer, 1996). By utilizing relatively small upgrades in generation capacity, a distributed utility approach can more closely match capacity upgrades with demand growth, thus avoiding the possibility for capacity overshoot of actual system needs (see Figure 32).

![Figure 30: Traditional utility response to demand increases is to build new facilities (bolded graphics). (Hoff, Wenger, & Farmer, 1996, p. 138)](image-url)
Figure 31: The value of distributed generation to the utility system. (Hoff, Wenger, & Farmer, 1996, p. 138)

Figure 32: Slow, lumpy capacity upgrades overshoot demand in three ways (the blue, yellow, and red portions of this graph). (Lovins, 2002, p. 122)
Challenges: social acceptability and regulatory support

The value of distributed resources is being increasingly recognized by some players in the electric power industry, but the widespread adoption of distributed resources faces resistance from an ingrained utility culture, which tends to view a distributed utility approach as either (1) not economically feasible or (2) not technically feasible to incorporate in the larger electric power system. Sovacool and Hirsh call these objections part of an “energy myth” (2007, p. 145), which obscures the reality of what are primarily social and regulatory barriers to the incorporation of distributed resources. In fact, the historically ingrained culture of utility managers, which produced in its adherents a preference for big, centralized generators as the best way to build an electric power system, comprises an important part of these social barriers, since these preferences discourage interest in employing new or novel options for the provision of electric services (Sovacool & Hirsh, 2007, p. 161).

Moreover, the general ignorance of most Americans about the systems involved in electric service provision (a 1978 California survey question “where does electricity come from” was most commonly answered with, “from the socket in the wall” (Sovacool & Hirsh, 2007, p. 161)) means that most people oppose plans for new electric infrastructure because they “do not realize that new plants of any type appear necessary to provide additional electricity” (Sovacool & Hirsh, 2007, p. 161). Interaction with and perception of power infrastructure is considered undesirable by most people, while the benefits of distributed resources to power systems do not factor in to their general positions of opposition. Because of the small scale and dispersed configuration of its components, a distributed utility approach to system planning may encounter more opposition from the public than would plans to build a few remotely sited plants. However, remembering that over time people tend to fold technological landscapes into their psyches and begin to view them as “second nature” (Williams, 2001), widespread social
ignorance about power systems and opposition to power infrastructure could fade as distributed resources bring such infrastructure from relative invisibility to the foreground of daily life.

Two more aspects of the social challenges to the implementation of a distributed utility approach to industry planning will conclude our treatment of the topic. These aspects are needed forms of new complexity. Much as early forms of complexity, such as monopoly status for utility firms and the interconnection of contiguous systems, allowed the Big Infrastructure model of the industry to develop effectively in the early decades of the 20th century, a few forms of high-return complexity could allow a distributed utility paradigm to effectively develop in the early decades of the 21st century. First, modifying rate of return regulation would provide an incentive for utility firms to pursue the most cost-effective investment strategies, which may likely entail a greater interest in smaller, distributed resources. Under rate of return regulation, utilities have little incentive to invest carefully, because they make profits commensurate with their level of capital expenditures. Under a performance-based regulatory framework, utilities would have an incentive to minimize their costs because their revenues would be fixed at a given amount for each customer served. The more efficiently the utilities could deliver service to their customers, the higher their level of profits (Lovins, 2002).

Second, the development of nationwide interconnection standards for distributed resources would enable manufacturers to drive down costs by scaling up production of standardized units, benefitting from economies of scale in the production process. This would also allow utilities to more effectively plan for the incorporation of distributed resources into their power systems without compromising system reliability or imposing high administrative costs.
Chapter six summary

The restructuring process begun by the passage of PURPA in 1978 was extended to a much broader field of independent power producers by the Energy Policy Act of 1992. By opening the nation’s interstate transmission network to non-discriminatory access by utility and non-utility power producers, EPACT92 fundamentally changed the organizational structure of the US electric power industry. Nearly two decades after the passage of EPACT92, most utilities (depending on the state) are still regulated by state service commissions and are granted monopoly control over a geographic service area. However, they are no longer vertically integrated companies with complete control over a set of generation, transmission, distribution, and consumption infrastructures.

To administer the operation of, and planning for, the Big Grid in this new organizational context has required the creation of an array of new planning organizations and the issuance of a multitude of new federal mandates aimed at more equitably administering the now open-access grid. Additional organizations will likely have to be created to oversee grid upgrades, as responsibility for that crucial task has been thrown into question by restructuring legislation. Meanwhile, the real price of electricity has remained largely unchanged under restructuring, confounding one of its central objectives.

In addition to the challenges and new complexities brought about by restructuring legislation, the period since the enactment of PURPA and EPACT92 has witnessed new developments in efficient small-scale generators and information and control technologies. This technological context makes increasingly feasible an approach to industry planning and operations based on locally sited, distributed resources (see Figure 33). Such a “distributed utility” paradigm would not eliminate the need for existing forms of complexity that have
developed to administer the Big Infrastructure model of the industry. It may, however, present a way to align the scale of electrical infrastructure with the decentralized political landscape of the United States and thereby avoid some of the complexity associated with planning, building, and maintaining current or future forms of the big, continental-scale grid.

The potential for the avoidance of such complexity may be especially beneficial when the development of the industry is considered from a historical perspective. Since the social and technological contexts from which the Big Infrastructure model of the industry emerged have changed significantly in the last fifty years, the amount of complexity necessary to maintain industry function has commensurately increased. This trend suggests that the electric power industry may be following a trajectory of diminishing returns to complexity in its management efforts. Intensifying our commitment to the Big Infrastructure model of the industry may therefore prove problematic in the long-term compared to a distributed utility approach.
Figure 33: Model of Industry Function: Current Opportunities
Chapter 7: Conclusion

This paper has analyzed the historical development of the electric power industry in America with particular emphasis on the role of complexity and the importance of changing social and technological contexts. Using the theoretical perspective of Supply-Side Sustainability to guide this analysis, several trends have been elucidated which may hold significant implications for current investment decision-making and the long-term viability of industry management efforts.

The American electric power industry emerged in the late 19th and early 20th centuries in particular social and technological contexts that facilitated the successful establishment and growth of the industry and the widespread electrification of the country. In particular, the technological context was characterized by rapidly improving generation and transmission technologies that incentivized the use of ever bigger infrastructure and rapid system expansion, both geographically and in terms of system capacity. The social context likewise supported the rapid growth of the industry by granting the protection of monopoly-based regulation to utility firms and due to the enthusiastic desire for electricity by the general public.

Because both social and technological contexts facilitated the establishment and growth of what I have called the Big Infrastructure model, the industry developed through the mid 20th century in a desirable manner for all of its stakeholders. Utility managers and engineers felt they were benefitting society (and perhaps glorifying God) through their status as a group of technically skilled elites. Investors and utility owners made healthy and dependable profits. Governments protects their citizens and the ideals of capitalism, while the general public enjoyed access to much appreciated and increasingly affordable electric power. During this time a few, high-return forms of complexity – state service commissions, rural electrification
programs, and PUHCA – ensured the social acceptability of the industry and the equitable
distribution of access to electricity.

Beginning in the 1960s and 1970s, however, the social and technological contexts of the
industry began to shift, leading to new and increased forms of complexity aimed at sustaining
the viability of the Big Infrastructure model of the industry. Instead of remaining passive
players, consumers became highly involved in the industry’s planning and decision-making
process, challenging plans and permits for undesirable projects and helping to enact a host of
environmental regulations. State service commissions, reflecting the sentiments of their
constituents, likewise became more active and stringent in their treatment of the industry.
While its social context had suddenly become far less supportive and dependable, the industry’s
technological context likewise shifted in problematic ways. Advancements in economies of
scale in generators dried up, indicating that utilities had lost their primary means of mitigating
difficult economic problems. The huge and highly interconnected grid began to display serious
reliability problems. Efforts to develop nuclear power were also plagued by the complexity of
the engineering process and the social resistance to that source of energy.

While both its social and technological contexts had shifted, the industry could not have
shifted the form of the grid into some more flexible form better suited to a more uncertain
future. Additionally, the conservative industry management culture that had become ingrained
during earlier decades of dependable and supportive contexts was resistant to the prospect of
change. The need to address the problems presented by shifting contexts was therefore filled
by a host of new forms of complexity that had not been necessary just decades earlier. These
included new R&D organizations, reliability oversight bodies, environmental and nuclear
regulatory bodies, and the Department of Energy. By increasing the complexity of the industry,
these efforts allowed the industry to survive into the 1990s with its general organizational form intact.

The problems of the 1960s through the 1980s brought to light some of the deficiencies of utilities’ monopoly status in changed technological and social contexts. Cognizant of a technological context in which small generators could be competitive with big power plants, and embedded in a social context characterized by a preference for competition and deregulation, the federal government opened the transmission network to non-discriminatory access in 1992. While enabling some efficiency gains in generators, restructuring of the industry also had the effect of creating a proliferation of new players and organizations in the industry. It has also necessitated a series of additional mandates and laws aimed at ensuring that access to the grid is indeed administered fairly. While some analysts see restructuring as having been beneficial and other see it as detrimental, on a national scale the average price of electricity and the rate of large blackouts have remained relatively stable. This suggests that while the electric power industry has been successful in maintaining the status quo of reliability and price, it has done so under conditions of significantly increased complexity.

While the early forms of complexity created to solve the problems of the electric power industry were characterized by high returns, it appears that the extensive new forms of complexity that have developed since the 1960s and 1970s may be characterized by diminishing returns. Although this paper cannot thoroughly and quantitatively assess the return to complexity throughout the history of the industry, nor does it provide an assessment of the multitudinous concerns over the costs and benefits of distributed resources, a broad interpretation of historical trends of context and complexity suggests that a distributed utility approach to industry planning and development may present one way of aligning industry management efforts with current technological and social contexts.
While a distributed utility approach to industry planning would require the development of new forms of complexity in its own right, it is feasible that these investments may be of high enough return to warrant the undertaking. Moreover, should the costs of complexity related to the maintenance of the Big Infrastructure model of the industry continue to rise, and if distributed systems are indeed more well suited to modern technological and social contexts, then the costs of transitioning to a distributed utility paradigm would be comparatively low and prospects for the long-term sustainability of the services delivered by the America’s electric power industry commensurately high.

At the very least, we should learn more about the costs and benefits of distributed utility approaches to industry planning. A more thorough understanding of the area- and time-specific costs of electric power systems, especially with regard to the costs of the grid, will help guide future management decisions. Additionally, investing some portion of the enormous costs that would have to be dedicated to the construction of a “Super Grid” into experimental distributed utility projects would help furnish real-world experience with the potential development of a new structural paradigm for the industry. Making a few such small investments that are rich with the potential to help avoid major and irreversible new courses of action and investment would be societal resources well spent.
Works Cited


Lerner, E. J. (2003, October/November). What’s wrong with the electric grid? *The Industrial Physicist, 8-13.*


http://www.nytimes.com/2009/12/14/us/14meters.html?_r=1&emc=eta1


