PUBLICATION

Networked Industries: Patterns in Development, Operation, and Regulation

P. H. Longstaff

March 2000

Program on Information Resources Policy

Center for Information Policy Research

Harvard University

P. H. Longstaff is an Associate Professor at the Newhouse School of Public Communications, Syracuse University. Professor Longstaff is also a lawyer and consultant who has advised clients in the print, broadcast, and tele-communications industries and holds a Master's of Public Administration from the Kennedy School of Government, Harvard University (1994).

ISBN 1-879716-60-7 **P-00-2**

March 2000

PROGRAM ON INFORMATION RESOURCES POLICY

Harvard University **Center for Information Policy Research**

Affiliates

Ameritech
AT&T Corp.
Australian Telecommunications Users Group
Bell Atlantic
BellSouth Corp.
The Boeing Company
Booz-Allen & Hamilton, Inc.
Carvajal S.A. (Colombia)
Center for Excellence in Education
CIRCIT at RMIT (Australia)
Commission of the European Communities
CyberMedia Convergence Consulting
CyraCom International
DACOM (Korea)
ETRI (Korea)
eYak, Inc.
Fujitsu Research Institute (Japan)
GNB Technologies
Grupo Clarin (Argentina)
GTE Corp.
Hanaro Telecom Corp. (Korea)
Hearst Newspapers
Hitachi Research Institute (Japan)
IBM Corp.
Intel Corp.
Kavner & Associates
Korea Telecom
Lee Enterprises, Inc.
Lexis-Nexis
Eli Lilly and Co.
Lucent Technologies
John and Mary R. Markle Foundation
McCann North America
MediaOne Group
Merck & Co., Inc.
Microsoft Corp.
MITRE Corp.
Motorola, Inc.
National Security Research, Inc.
National Telephone Cooperative Assoc.
NEC Corp. (Japan)
NEST–Boston
Nippon Telegraph & Telephone Corp. (Japan)
NMC/Northwestern University
Qwest Communications International, Inc.
Research Institute of Telecommunications and Economics (Japan)
Revista Nacional de Telematica (Brazil)
Samara Associates
SK Telecom Co. Ltd. (Korea)
Strategy Assistance Services
TRW, Inc.
United States Government:
- Department of Commerce
  - National Telecommunications and Information Administration
- Department of Defense
  - Defense Intelligence Agency
  - National Defense University
- Department of Health and Human Services
  - National Library of Medicine
- Department of the Treasury
  - Office of the Comptroller of the Currency
- Federal Communications Commission
- National Security Agency
- United States Postal Service

Acknowledgements

The author gratefully acknowledges the following people who reviewed and commented critically on the draft version of this report. Without their consideration, input, and encouragement, this study could not have been completed:

John W. Berresford
Richard J. Byrne
Michael Cartney
Robert Carson Godbey
Richard John
Vincent Mosco
Kurt Yeager

These reviewers and the Program's Affiliates, however, are not responsible for or necessarily in agreement with the views expressed here, nor should they be blamed for any errors of fact or interpretation.

Research assistance was provided by Melani Keni and Robert Rubi, graduate students at the Newhouse School of Public Communications, Syracuse University.

This report was made possible with the assistance of the Development and Leaves Committee of the Newhouse School of Public Communications, Syracuse University.

Contents

Acknowledgements ... ii

Executive Summary ... ix

Chapter One **Introduction to Networked Industries** ... 1

Chapter Two **Networks and Networking: An Overview** ... 3

Chapter Three **Characteristics of Networks** ... 11

3.1 Senders ... 11

3.2 Receivers ... 13

3.3 Channels ... 14

3.4 Transport ... 14

3.5 Traffic and Payload ... 15

3.6 Security ... 15

3.7 Signalling ... 16

3.8 Scheduling ... 17

3.9 Terminals ... 19

3.10 Ancillaries ... 19

3.11 Common Problems ... 20

3.11.1 Bottlenecks ... 20

3.11.2 Access ... 21

3.11.3 Small Loads vs. Large Loads ... 22

3.11.4 Short Haul vs. Long Haul ... 22

Chapter Four **Regulating Networks** ... 25

4.1 General Considerations ... 25

4.2 Effects of Competition (It's Not "Deregulation") ... 28

4.3 Effects of Competition on Rates and Service ... 30

Chapter Five **Transportation** ... 33

5.1 Roads, Bridges, and Ferries ... 33

5.2 Waterways ... 38

5.3 Railroads ... 41

5.4 Airlines ... 47

Chapter Six **Communications** ... 51

6.1 Postal System ... 51

6.2 Telephone System 55
6.3 The Internet and Computer Networks 61
6.4 Broadcast, Cable, and Satellite 64

Chapter Seven Energy 69
7.1 Natural Gas 70
7.2 Electricity 74

Chapter Eight Conclusions 83

Acronyms 85

Illustrations

Figures

2-1 Point-to-Point Network ... 4
2-2 Point-to-Multipoint Network ... 5
6-1 The Telecommunications Network ... 60
6-2 Television Networks ... 64
6-3 Cable Networks ... 65
6-4 Direct Broadcast Satellite (DBS) Networks ... 66
7-1 The Natural Gas Industry ... 71
7-2 Unbundling of Gas Sales from Pipeline Transportation, 1992 and Beyond ... 73
7-3 Electric Power System and Poolco ... 79

Tables

3-1 Characteristics of Networked Industries ... 12
3-2 Problems Common to Networked Industries ... 13

Executive Summary

Executives and regulators in networked industries all over the world (in particular, communications, transportation, and energy) knew they were entering a new world when market forces were introduced into those industries in the late twentieth century. Everyone knew that things would change, but almost no one forecast the landscape that developed by the year 2000. Few predicted decreases in quality of service (indeed, most had predicted an increase) nor the wave of consolidation that swept each industry. Assumptions about how these networks operate and how they respond to forces from markets and from government apparently need to be revised. Revision may be especially important for those trying to devise business plans and public policy for the newest communications network, the Internet.

Although much about these networked industries makes them different from one another, in their responses to the introduction of competition they have been strikingly similar. Another area where networked industries are comparable is in their initial development and operation. This report examines the early development, characteristics of operation, and operational problems common to communications, transportation, and energy networks—and offers a comprehensive nomenclature to facilitate the discussion of networks across industrial lines—with a focus on the role of cooperation in networked systems and the effect on a network when cooperation is reduced by the introduction of competition.

Chapter One

Introduction to Networked Industries

With the twentieth century drawing to a close, businesses and governments all over the world have been struggling to find models for using and regulating the network of networks known as the Internet. At the same time, both the public and the private sectors have been seeking models for "deregulating" (by which was meant "introducing competition into") the networks in the industries that provide transportation, communications, energy. Even the World Bank, a strong advocate of privatization, expressed doubt that market forces can make these industries successful without a workable regulatory and legal framework in place.[1]

To find clues for model-building, it seems reasonable to look at the nature of networks, which are a particular kind of system, and of networked industries. This report explores networked industries in the transportation, communications, and energy sectors for common elements with which to build a model. The report examines the histories of these networked industries to locate common or repeating patterns and thereby avoid (or at least observe) the "vengeance of history." For example, many economic and political realities of the telegraph network reappear with the Internet.[2] These repeating patterns are not limited to the same industrial network. Some things that appear first in transportation networks appear later in communications networks. For example, at several points in the history of transportation in the United States, established networks of the industry have been threatened by new technologies. In the early 1970s, the airlines and trucking industries underwent extensive regulatory overhaul and liberalization (to ease entry for new competitors). That there are many similarities between the history of railroads and of the Internet has been noted, including the railroad industry's promise of a better society.[3] Thus, transportation networks may offer a glimpse of what may happen to communications and energy early in the twenty-first century.

The broad, inclusive model developed, for the first time, in this report is necessary, because old models (i.e., assumptions about how networked industries operate) have been falling to the forces of the new technologies that have been redefining those networks, allowing industries to encroach on one another. New regulatory models have been made necessary by a new political faith in the allocation of each network's resources by the marketplace. New technology and competition have been introduced into networked systems in the past and have signalled certain results, for example, a spurt of mergers and acquisitions in a particular industry. It is unlikely that

---

[1]Peter Montagnon, "Doubts at World Bank on Infrastructure Sell-Off," *Financial Times*, July 27, 1999, 4.

[2]Tom Standage, *The Victorian Internet: The Remarkable Story of the Telegraph and the Twentieth Century's On-Line Pioneers* (New York: Berkeley Books, 1998).

[3]See Edward Rothstein, "Looking at the Transcontinental Railroad as the Internet of 1869," *The New York Times*, Dec. 1, 1999, [On-line]. URL: <http://www.nytimes.com/library/tech/99/12/biztech/Articles/11rail.html>.

a set of "rules" that could offer precise predictions for these systems will be developed soon, given the enormous (and growing) complexity of each network. But regulation or business planning of any kind requires at least some preliminary *assumptions* about how networked industries will respond to a new economic or regulatory stimulus.

Assumptions are, by their nature, dangerous unless they can be tested frequently against what is actually happening. Assumptions have led to failed business and regulatory schemes when the "facts" upon which they were based were in flux or when they failed to recognize history repeating itself. The threads of ideas teased from this study of networks will, it is hoped, be woven in new ways by engineers, economists, and policymakers and tested against real-world operations of twenty-first century networks.

The following chapter offers a new synthesis of the basic principles of network operation and an indication of what the networks described here all have in common. Although people working in and regulating what used to be called "public utility" networks have always noted that there is much in common, they nevertheless have tended to concentrate on the things that make those networks different. Chapter Three presents ten common characteristics of these networks along with four common problems in their operation. Chapter Four reviews ways those characteristics have played out in government's involvement in these networks, because, inherent similarities notwithstanding, politics still counts. In this report, individual networks are distinguished by their physical characteristics and histories, and such differences may often be as important as their similarities. The goal here is to fashion a comprehensive *starting place* for business decisions and regulations.

To illustrate similarities, Chapters Five, Six, and Seven, respectively, examine in brief the histories and operations of transportation, communications, and energy networks, in a way meant to exemplify, rather than define. Readers familiar with the history of one or another industry may miss much that might have been included, but because this is a report, not a book (and given that readers of such reports have only limited for them), only the barest outline is presented in order to provide pertinent examples.

Chapter Two

Networks and Networking: An Overview

Networks are collection of smaller entities connected with one another in order to function (at least part of the time) as a larger entity. In the human body, networks allow cells to communicate with the brain (the nervous system) and to receive energy and dispose of wastes (the circulatory system). In human societies, networks allow individual people and businesses to combine their resources in order to conduct matters over very long distances which they could not accomplish alone: communicate with one another and deliver goods and services. As networks grow larger and more efficient, they break down the need for local services such as retailing, education, and medicine, allowing such services to be concentrated in hubs.[1]

Networks operate as sets of connections built for the benefit of those connected, and by their nature they are cooperative constructs. Organisms and economic units do not cooperate unless cooperation will allow them to obtain more of a scarce resource than they could obtain on their own through competitive interaction.[2] When some parts of a network perceive that they can get what they want without cooperating with the rest, they are likely to leave and perhaps compete with the old network for scarce resources. This report examines these phenomena as it looks at the introduction of competition into cooperative networks for transportation, communications, and energy. For example, large customers or concentrations of customers, or both, for a resource (e.g., air transport, computation, electricity) can aggregate enough demand to make their network viable and can abandon smaller or more dispersed customers on the network.

Cooperation in networks is necessary also at the organizational level. Central coordination usually is necessary for signalling and scheduling (routing, switching, capacity control) as well as for setting standards to insure interoperability among the parts of the system. Individual parts of the system (sometimes called nodes) cannot perform coordinating functions nor have they any incentive to do so, because such activities (theoretically) will not get them more of the resources available within the system.

Networks may be organized in a variety of ways, depending on their purpose. This report looks at two kinds of networks, point-to-point networks and point-to-multipoint networks. Neither is better than the other; the two are simply different and each works better in different situations.

---

[1]For this hubbing effect in communications networks, see, e.g., Wilson Dizard, Jr., *Old Media New Media: Mass Communications in the Information Age*, 2nd ed. (White Plains, N.Y.: Longman Publishers, 1997), 71-73.

[2]See P. H. Longstaff, *Competition and Cooperation: Biology Meets Business* (Cambridge, Mass.: Harvard University Program on Information Resources Policy, P-98-4, October 1998), 29-37.

Networks that must allow all nodes to connect with one another so that traffic can move from any one node to any other are *point-to-point* networks (see **Figure 2-1**). Examples include roads and telephones. They allow traffic (goods, information, etc.) to move in both directions in the network hierarchy, so that resources can to be sent from anywhere in the system, not just from a central point. This type of network is useful where resources are diffused and must be moved to many different places.

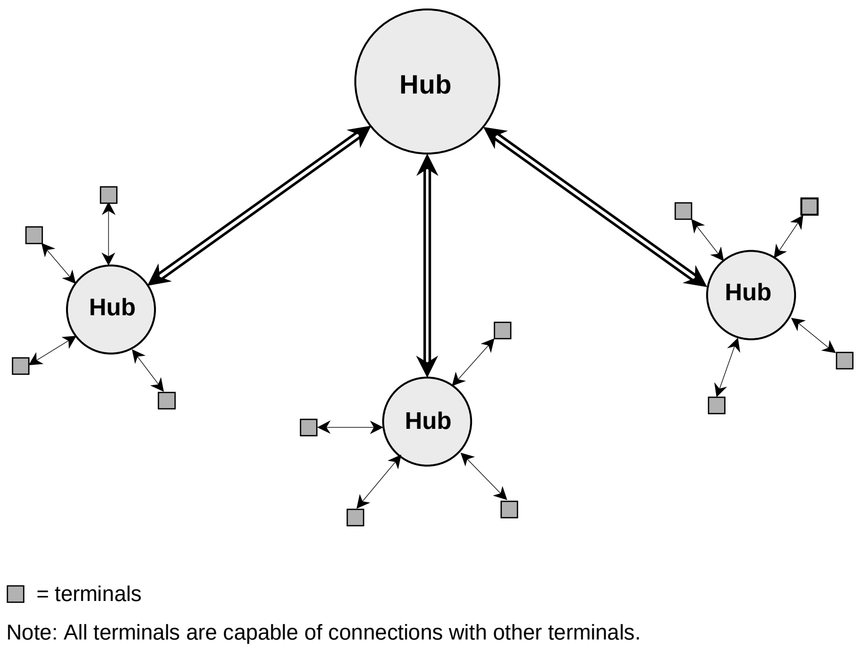


Figure 2-1

Point-to-Point Network

Most point-to-point networks include many point-to-multipoint hubs. In systems with more than a few points, a hub-and-spoke system is more efficient, i.e., all points are not directly connected but are instead connected through the hub. These simple systems can be combined with others to form complex networks in which there are several layers of hubs. Hubs become critical to the operation of the system, because they *switch* traffic from one line to another and for that reason can become *bottlenecks*. The hub's ability to move traffic quickly and efficiently can determine the capacity of the entire network. In networks in which a hub processes the traffic before the traffic is redirected, the quality of that processing will have ramifications throughout the network and not just on the two points that act as sender and receiver. Hubs can facilitate two-way traffic in the system (e.g., communications and transportation networks) or control a one-way flow that must accommodate varying levels of local demand (energy networks).

Point-to-multipoint (or one-to-many) networks only require nodes to connect with a central controlling mechanism, which acts as the system's central supplier of resources (see **Figure 2-2**). Examples include broadcast and cable networks, energy networks (gas and electric), and computer systems that have a central processing unit (CPU). This kind of organization is sometimes referred to as a *command-and-control* network, because its nodes are not directly connected with one another (although they may develop connections outside the network) and because they send information about the local situation (e.g., needs) to a central supplier or controller, which then sends them resources or commands. This kind of network works best where a centrally located resource must be distributed to many places at the same time. Figures 2-1 and 2-2 necessarily represent simplified pictures of point-to-point and point-to-multipoint networks. Modern networks usually have many connections at many levels, some within the network (for example, two lower level hubs of the network that have direction connections to each other) and some connect to other hubs through an outside network (such as airports connected by bus routes).

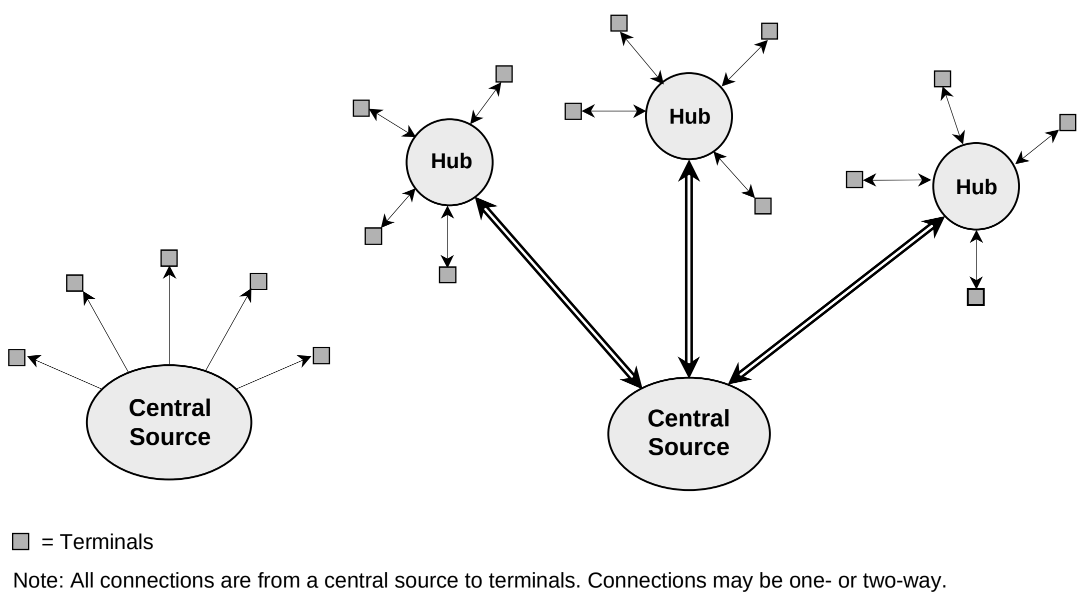


Figure 2-2

Point-to-Multipoint Network

What distinguishes a *networked industry*? As suppliers to and customers of other firms, all firms are parts of networks. They have all been urged to consider networks crucial to their success and to think of themselves as part of a larger, interdependent whole.[3] But in some industries firms cannot exist unless they can connect with other firms that look just like themselves. Usually such firms have a territory that they call their own (or that they share with a few others), and most were once considered *natural monopolies*. Such firms need to move things outside their own territory in a system that is compatible with their own. They have high fixed costs (which act as a barrier to entry) and low marginal costs.

Networked industries generally fall into three categories: transportation, communications, and energy. Public water services and garbage removal share some characteristics of networked industries but tend to have less need for connections with networks outside their own territories.

New networks in transportation, communications, and energy often are developed because they can increase the *speed* of movement in the channel. Increases in speed are no accident, because very often the initial research and development (R&D) used by the networks are conducted or financed by military authorities seeking strategic advantages over opponents through faster access to information and other resources. Speed can be enhanced by increasing the rate at which traffic moves through, enters, or leaves the channel. For example, flying from New York City to Newark, New Jersey, may take less time in the channel (as compared with driving by automobile or bus) but more in the process of loading or unloading and getting to and through airports.

Increased speed in transportation, communications, and energy has engendered many of the attributes of post-industrial societies as well as rapid diffusion of innovations. But the amount of change that people can accommodate within a relatively short span of time, as well as what they will pay for faster service, may have a limit. For example, most travelers have been unwilling to pay higher prices for faster transatlantic jet service on the Concorde.

As technology made it possible for communications traffic to move through airwaves and various forms of wires, large segments of the communications sector were physically divorced from the transportation network but continued to have a profound impact on one another. In the late 1990s, many commentators and scholars have predicted that improved communications (via digital broadband services) will allow more firms and other institutions to travel less while they conduct more business through videoconferencing and computer networks, i.e., that communications will *substitute* for transportation.[4] Others have regarded transportation and communications as *complements* for each other, in that as more people communicate more, they want to meet and vice versa. Both communications and transportation networks extend each

---

[3]See, e.g., Adam M. Brandenberg and Barry J. Nalebuff, *Co-opetition* (New York: Currency Doubleday, 1996); and James F. Moore, *The Death of Competition: Leadership and Strategy in the Age of Business Ecosystems* (New York: Harper Business, 1996).

[4]See, e.g., Vincent Mosco, *Will Computer Communication End Geography?* (Cambridge, Mass.: Harvard University Program on Information Resources Policy, P-95-4, September 1995).

person's (or institution's) range of activities.[5] Large transportation and communications networks allow *new markets to open* for goods, services, and ideas and thereby increase competition (and turbulence) in the newly available territories. Both these networks can change the structure of social institutions, such as the family, when they make it possible for children to leave the vicinity of their parents to look for new opportunities in the hubs that the networks make possible.[6]

Because these networks are often regarded as "public goods," most countries usually have not allowed network facilities to be owned by foreign nationals. This barrier has begun to break down as foreign investment has come to be seen as necessary for growth or even survival in the new competitive environments. The late 1990s was a time of relative world peace, so that the dangers involved in the ownership of public goods by foreign nationals or multinational corporations came to be regarded as worth the risk in return for the improved services deemed critical for world competitiveness. The technological changes this liberalization made possible reduced the costs of many services provided by the networks at the same time that it allowed some previously centralized functions to be *distributed* within the network. Distribution of functions to lower levels of the network hierarchy lets large users form subnetworks and opt out of some of the services formerly provided by The Network (that is, the previous, more cooperative, integrated network).

New networks may also offer greater *flexibility* in how and whom they serve. Increased flexibility allows greater *specialization* in the traffic that can be transported. The communications network known as the Internet allows nearly instantaneous communication in a variety of formats (text, audio, video) in a *point-to-point* network, so that anyone on the network can send messages in these various forms to anyone else on it. Increases in the speed and flexibility of communication allow for a tighter definition (thus, fragmentation) of the markets for entertainment and information traffic. One result may be one of the great paradoxes of the twenty-first century: a faster, more flexible communications network that may eventually slow the large-scale diffusion of innovation and ideas by fragmenting the mass audience that had been critical to the rapid diffusion of information about new communications products and services.[7]

During the twentieth century, communications, transportation, and energy networks saw new technologies make possible the invasion of their territories by other networks that offered a substitute for the original service (*internetwork* competition). Networks were invaded also by similar companies from other "territories" (*intranetwork* competition) after government

---

[5]Arnulf Grubler, *The Rise and Fall of Infrastructures: Dynamics of Evolution and Technological Change in Transport* (Heidelberg and New York: Physica-Verlag, 1990), 254-258.

[6]See, e.g., Thomas C. Cochran, "The Social Impact of the Railroad," in *The Railroad and the Space Program: An Exploration in Historical Analogy*, edited by Bruce Mazlish (Cambridge, Mass.: Massachusetts Institute of Technology [MIT] Press, 1965).

[7]See E. Rogers and F. Shoemaker, *Communication of Innovations* (New York: The Free Press, 1971). For an overview of the literature on diffusion of innovation and the role that communications play, see, e.g., Grubler, 11-69.

implemented policies for liberalization and privatization (selling off government-owned network assets). Internetwork and intranetwork competitors from other countries appeared with the advent of new rules for free trade, such as the Uruguay Round (1994) of the General Agreement on Tariffs and Trade (GATT).

*Intra*network competition generally occurs when two or more firms try to sell the same services (air transport, telecommunications channel access, electric generation, etc.) to the same customers. As in all intraindustry (and intraspecies) competition,[8] the result will be a weeding out of the weakest players. In the late 1990s, many companies emerged to enter newly opened markets for networked services, but few were expected to survive the inevitable cutthroat price competition.[9]

*Inte*rnetwork competition occurs when one network seeks the customers of another network, e.g., where railroads try to take traffic away from canal networks or where satellite services try to take customers away from cable networks. Competition may result in the displacement of one network by another, because introducing a new competitor may not widen the market for the service—the market may not offer enough demand to keep both networks in business. For example, studies have noted a “transportation constant”: people in many cultures will devote no more than 1 to 1.5 hours per day to transportation.[10] There also seems to be a limit on the amount of time that people will spend on communications and entertainment.[11] For example, if households spend (on average) no more than X percent of income or time on a particular activity (transportation, energy, information, or communications), then that percentage is a limited resource for which all providers of the service must compete. Services sold by a new competitor will decrease the services that can be sold by incumbents.

Thus, introducing competition does not guarantee that all the players will stay on the field to compete in a way that will benefit consumers. Individual businesses or industries will almost certainly fail. But old networks do not die immediately (and sometimes not at all). They may instead abandon various markets to new entrants, conceding the battle for traffic of higher value (passengers, information) or lower cost (long-haul) in order to concentrate on commodities (mineral ores, grains) and other traffic that are less time-sensitive or more cost-sensitive (entertainment).[12] Several networks may have the same productivity (speed $\times$ capacity) but with the variables in different proportions. Network capacity can be affected by bigger channels (wider

---

[8]See Longstaff, *Competition and Cooperation*.

[9]See, e.g., John R. Emshwiller and Katherine Kranhold, “California Power Deregulation Isn’t as Open as It Looks,” *The Wall Street Journal*, Feb. 17, 1998, B2.

[10]Grubler, 2-3.

[11]For a discussion of such limits, see, e.g., W. Russell Neuman, *The Future of the Mass Audience* (Cambridge, England, and New York: Cambridge University Press, 1991).

[12]For the response of railways to competition from automobiles and airlines, see Grubler, 115-122.

freeways) or bigger vehicles (larger trucks) in the channel or by more efficient terminals, scheduling, or signalling. For example, although railroads are slower than airplanes, their greater capacity for traffic by certain vehicles (trains) moving in the channel means they may be equally productive.

The introduction of competition has had much the same effect on all the networks studied here. Where once there was a system of impermeable local monopolies (usually protected from invasion by government) that cooperated to form a larger system, there appeared one system in which (*intra*- and *inter*industry) competition increased the difficulty of *intra*network cooperation at the operational level and led network owners to seek growth through *inter*industry cooperation (vertical integration) and through mergers and acquisitions (horizontal integration) to build separate, highly integrated networks.[13] Existing networks that become subject to increased competition retain some of their business by becoming *feeders* to new networks, creating a larger, bimodal system.

Other factors in the competitive battles among networks include *range* (how many possible connections can be made and at what distance) and *quality* (how dependable is delivery and what shape will the traffic be in when it arrives). These factors (and the availability of competitive services) all affect the *price* of the network's services. The consumers' final choice of a particular network will reflect a complex mixture of these factors.

Networks develop either from the bottom up or from the top down. In bottom-up development, small segments of the network are built where there is a specific demand for the new service, i.e., demand pulls a new technology into an area (e.g., railroads that link a mine and a sea port, telephones that link businesses in a local area). In top-down development, the network arrives before demand in hope that its arrival will spur economic development, which, in turn, will create demand, i.e., the technology pulls demand into the area (railroads and communications services sponsored by local government or land developers). The new processing power of computers, which allows networks to track customer use, may create a hybrid development strategy, one planned from the top but actually based on demand at the bottom.[14] Both bottom-up and top-down development strategies require a network to reach "critical mass" before becoming economically viable.[15]

Bottom-up networks depend on cooperative systems to interconnect: they must agree on signalling, security, scheduling, and channel configuration for traffic to move from one part of the

---

[13]The effects of competition and cooperation in such a system are not discussed at length here, but see Longstaff, *Competition and Cooperation*, 20-28.

[14]See Kevin P. Coyne and Renee Dye, "The Competitive Dynamics of Network-Based Businesses," *Harvard Business Review* (January–February 1998), 99-109.

[15]See, e.g., David Allen, "New Telecommunications Services: Network Externalities and Critical Mass," *Telecommunications Policy* **12**, 3 (September 1988), 257-72.

network to another. In top-down networks, agreement is generally made in advance by a central authority. This critical need for cooperation sets networked industries apart from other industries and makes intranetwork competition problematic. Both bottom-up and top-down networks therefore foster uniform procedures and practices among all participants and may change those procedures or practices only slowly, in light of new market or technological conditions.

In theory, an "open network architecture" combines attributes of bottom-up and top-down networks.[16] There is no central authority, but the necessary agreements on network configuration, signalling, and so on, are set as "standards" for the system in advance by government or by agreement among industry players. The advantage of open network architecture is that it allows any service that complies with the standards to operate on the network (thus encouraging innovation), but it also means that capacity (thus the speed of the system) is difficult to control. The standards selected are critical to the success of the network: standards set too high will allow fewer network service providers (as well as senders and receivers) to use the network, while standards set too low may affect the reliability of the system.[17]

---

[16]See "ONA: Unbundling and Competition in Interstate Access Markets," *Telecommunications Policy* **16** (April 1993), 194; and G. J. Mulgan, *Communication and Control* (New York: Guilford Press, 1991), in particular, Chapter Seven, "The Limits of State Control and Deregulation."

[17]See, e.g., Martin C. Libicki, *Information Technology Standards: Quest for the Common Byte* (Boston: Butterworth–Heinemann, Digital Press, 1995); and Cristiano Antonelli, "The Economics of Standards," *Information and Economics Policy* **6** (December 1994 [Special Issue]). The article comprises the whole of the special issue.

Chapter Three

Characteristics of Networks

The networks examined here have ten characteristics in common (see **Table 3-1**). They could not operate as networks without any one of them. The histories of these networks show that these characteristics influence the nature of the network to which they belong, and a change in any one characteristic has consequences for the others, as well as for the network as a whole. Thus, these characteristics are *interdependent* parts of the system. Like the heart, lungs, and brain of the system called the human body, these pieces cannot exist without the others, and calling one more important than another is difficult, if not impossible.

The names used here for these characteristics are drawn from a variety of networks and selected on the basis of their use either by nonspecialists or in more than one network. Specialists may find a familiar function called by another name. The selection of names may seem somewhat arbitrary, but finding nomenclature that crosses old boundaries is critical to a broad-based study of networks.

The examples in the table clearly are not exhaustive but intended to illustrate the many ways in which these networks exhibit the characteristics. Each characteristic is discussed in some depth in the following sections of this chapter. Networked industries share not only these ten characteristics but also a history of common problems (see **Table 3-2**) which they must deal with. Each has evolved similar answers to the problems. The forces working to mold the evolution of the networks include improvements in technology, inter- and intraindustry competition, and government regulation.

3.1 Senders

Senders are the people or entities that put traffic into the network. In point-to-point networks (postal service, telephony, highways), they may also be receivers. They may or may not pay to put traffic into the system (although, as a general rule, the cost is ultimately paid by receivers, directly or indirectly, as part of the cost of traffic received). Senders may continue to own the traffic en route (unless it is already owned by the receiver), but they have little control over how traffic gets to its destination—that is, they cannot control how traffic is routed in the network. Their main concerns are speed, reliability, and price, any one of which can become a tradeoff against the others, e.g., they might pay more for more reliable or faster service. Unless these people or entities are high-volume customers they have little economic power over any of these concerns.

Table 3-1

Characteristics of Networked Industries

Network Components	Function	Examples
Senders	Put the traffic into the system	• Manufacturers • Call placers • Electricity, oil producers
Receivers	Take the traffic out of the system	• Retailers, consumers • Postal recipients • Electricity customers
Channels	Routes through which traffic moves	• Roads, waterways, railroads • Airwaves, telephone lines • Electric lines, pipelines
Transport	Movement of traffic through the channel	Movement of trains, trucks, communication signals, gas flow
Traffic/Payload	What moves through the channel	• Commodities, finished goods • Electrons, oil, gas • Signals (bytes, waves)
Security	Keeps traffic safe from theft and keeps networks safe from disruption	Security procedures at airports, guards at cargo terminals, fences around pipeline terminals
Signaling	Communications within the system to ensure system safety and security	Railroad crossing signs, "addressing" digital signals, flow meters in energy lines
Scheduling Routing, Switching, Capacity	Routes traffic within the channel *or* from one part of the channel to another *or* from one mode of transport to another. Traffic may be temporarily stored during this process	*Routing* or *switching* between lines (e.g., rail, telephone, electric trunk lines), or modes (e.g., between satellite and cable or truck and rail car), *or* scheduling capacity in the channel (e.g., for transponder or rail time)
Terminals	Where traffic is put into and taken out of the channel (although traffic may be temporarily stored)	Railroad depot, grain terminals, telephone customer equipment, gas or oil storage depots
Ancillaries	Other components necessary for the network to develop or function, including energy to move traffic through the system	Fueling stations, tugboats, electric power (for phones and computers)

Table 3-2

Problems Common to Networked Industries

Problem	Description	Example
Bottleneck	Where traffic stacks because capacity is limited or temporarily blocked	An airport that is too small, a telephone switch with insufficient capacity, a damaged railroad bridge
Access	• Physical: Ability to connect • Economic: Ability to pay for connection or use	• Physical: People in an area without phone service • Economic: People unable to pay for phone service
Small vs. Large Traffic	Costs to haul many small units of traffic to many receivers are higher because there are fewer opportunities for economies of scale (marginal costs high) and more signals and scheduling required	Aggregated shipments of rail freight or mail can be handled at lower cost
Short vs. Long Haul	Cost of short haul greater per mile because more signaling and scheduling needed than for long haul	Delivering a package, passenger, or telephone call from New York to Los Angeles may be cheaper than from New York to Newark, New Jersey

3.2 Receivers

Receivers take traffic out of the network. In point-to-multipoint networks, they function also as senders (e.g., receivers of natural gas). Like senders, they are concerned about speed, reliability, and price, and, again like them, they have little economic control unless they can aggregate large amounts of demand.

Both senders and receivers as citizens of democratic regimes pay taxes and cast votes, and to the extent that government has aided the particular network, they can as voters and taxpayers demand a return on their tax dollars. Such *public interests* include protection from perceived price gouging and guaranteed service minimums. Those not served by the particular network (e.g., because their limited use cannot cover the costs of providing the infrastructure) demand that government assure *universal service*, whereby senders and receivers in other parts of the network (or taxpayers who may or may not be part of it) pay extra for their service in order to extend the network to unprofitable customers or areas.

Senders and receivers may see both their use of the channel as crucial to their own economic success and an “affordable” price as a key advantage against competition, especially if the competition is outside the network. Many businesses depend on relatively low-cost energy and communications to remain competitive in world markets.

3.3 Channels

Channels are routes along or through which traffic travels. Routes may involve substantial terrestrial infrastructure, such as pipelines and railroads, or they may be allocations of air space or of the electromagnetic spectrum. New channels often are larger by an order of magnitude than existing ones, because they must compete with established channels by offering more links to potential senders or receivers. For example, since 1800, each new transportation infrastructure has expanded into a network ten times larger than the previous infrastructure.[1]

Channels are often expensive to build, because the owner must acquire *right of way* across property owned by others. For example, no electric power network can be built without needing to erect poles and lines over both private and government-owned property. Often it cannot be done without the power of government to force the sale of private property for public use. Government thereby becomes either the owner or a heavy regulator of the channel and as such can insist that the channel maintain some character of a "public good."

Because rights of way and channel infrastructure are so expensive, they are often used by two or more networks. A new network may be built on the structure of an old one. For example, electric and telephone (and, more recently, cable television) networks share rights of way and sometimes even poles. Such rights of way may be public streets, which also act as channels of transportation that were built over the pathways of ancient peoples and, before those, animal paths. Many of the networks discussed here were arranging new uses for their valuable rights of way at the end of the twentieth century, when building new infrastructure was becoming very expensive. Gas and electric companies, for example, were building fiber-optic networks for telecommunications.[2]

3.4 Transport

The *characteristic* here denoted as transport is not to be confused with *transportation networks*, even though both perform the function of moving things around. The *network characteristic of transport* refers to the function of any network in which something is moved from sender to receiver within the network's channel(s). In some newly competitive networks, this function has been "unbundled" from ownership of the channel, and providing transport is often the new service that the new competitors offer.

Transport often involves several classes of service that are priced according to the priority assigned to the traffic transported. In a network operating near capacity, higher classes of traffic

---

[1]See Arnulf Grubler, *The Rise and Fall of Infrastructures: Dynamics of Evolution and Technological Change in Transport* (Heidelberg and New York: Physica-Verlag, 1990), 185-186.

[2]See, e.g., Bloomberg News, "Williams to Raise up to $750M in Offering of Telecom Unit," *The Boston Globe***,** Nov. 20, 1998, D-3.

will move more quickly than lower classes, such as first-class postage. First-class transport may indicate a higher level of service, such as the extra services given to first-class airline passengers.

3.5 Traffic and Payload

The "stuff " that moves in the channel is the *traffic*. Traffic may be unpackaged, such as gas in a pipeline or pedestrians on a road, or it may move as *payload* in some sort of *vehicle*, such as mail in an envelope or passengers in an automobile.

When traffic moves as a payload, its vehicle becomes ancillary to the system (see section **3.10**) and the vehicle must be compatible with the channel. Train cars, for example, must be the same gauge as the tracks. One way to keep traffic out of a channel is to set standards that make the channel incompatible with or inhospitable to particular traffic or vehicles. Such exclusiveness permits a network to keep out traffic from competitors or from other countries. Some rail interests tried to use their influence to insure that roads were inhospitable (that is, very bumpy) to the new steam-powered automobile in order to suppress this new form of competition.[3]

3.6 Security

The integrity of a network is critical, because a network can be slowed (or stopped) by the creation of a bottleneck or by a signalling failure. Damage to nodes or lines (either directly or indirectly by insertion of a contaminant into the system) can cut capacity, requiring traffic to be rerouted. In wartime, the land-based shipment of munitions can be stopped by blowing up bridges, creating a bottleneck. Jamming the radio frequencies (by inserting noise into the channel) that air traffic controllers use could ground all planes, reducing capacity to zero. Security procedures for all networks involve stopping unauthorized access to network facilities that might (intentionally or unintentionally) cause damage to the network or traffic. Security measures are heaviest at terminals where traffic enters and exits the network. Governments have not ignored either the vulnerability of networks that operate within their territory[4] or the possibility of gaining strategic advantages by breaching the security of their enemies' networks.[5]

Security procedures are undertaken also to avoid damage to or theft of traffic while it is in the channel, because the owner of the channel may be held responsible for this damage. Special systems ensure that air cargo is stored in secure facilities and that passengers do not violate the safety measures designed to protect them. The meters on energy networks measure flow and can

---

[3]See David Beasley, *The Suppression of the Automobile: Skullduggery at the Crossroads* (New York and London: Greenwood Press, 1988).

[4]See, e.g., *Critical Foundations: Protecting America's Infrastructures*, Report of the President's Commission on Critical Infrastructure Protection (Washington D.C.: 1997).

[5]See James Adams, *The Next World War: Computers Are the Weapons and the Frontier Is Everywhere* (New York: Simon and Schuster, 1998).

detect thefts of traffic. If senders and receivers judge the levels of damage or theft unacceptable, they may look for another network for delivery of their traffic.

Communications and transportation networks have an additional security concern: privacy. Airlines protect passengers who do not wish to draw attention to themselves or to have their identities divulged. Many communications networks are required by law not to disclose the identities of senders and receivers or the nature of traffic in their system. Cable TV systems cannot disclose which channels their customers subscribe to or which movies they order. The Internet has created many new opportunities for the collection of information about senders/ receivers and a new level of interest in the concept of privacy in networks. This is made complicated by the lack of certainty about whether senders/receivers or the channel *own* the information constructed for the signalling function.[6]

3.7 Signalling

Signalling is the internal communication function of the network. Signalling tells the channel where and how to send traffic. It lets the parts of a network communicate with one another about scheduling traffic, determining the traffic's final destination (sometimes called "addressing"), and notifying the network about matters of safety or security. A train may be signaled to pull onto a side track to allow a train with a higher priority go through on the main track or in order to avoid a collision. A train can signal the number of cars it carries by the markings on those cars. The contents of each car can be made known to the system so that the right goods get to the right terminal.

Advances in communications technology (and communications networks) have improved the signalling capacity of other kinds of networks. In the mid-nineteenth century, telegraphs allowed railroads to contact trains en route, to give them new routing instructions or to warn them of obstructions on the track. With digital systems and bar-coding, traffic in transportation systems can be monitored closely. Signalling rates are set by the upper limit of the speed of a network. A disruption in the signalling can bring the network to a halt, so signalling is subject to strict security.

The "meaning" assigned to coded signals must be agreed upon in advance by all parts of the network. Agreement is a critical element of network administration. If the air-traffic control system uses English to communicate with planes, all airlines must agree in advance to use English and to employ English-speaking pilots. The signals used by a telecommunications network must be understood by the switches that route calls. Signals may all be in the same "language" or code, or they must be instantly translatable (e.g., analog signals into digital ones).

---

[6] Anne Wells Branscomb, *Who Owns Information? From Privacy to Public Access* (New York: Basic Books, 1994).

3.8 Scheduling

No network can move an unlimited amount of traffic; networks must allocate the limited resource called *capacity* by *scheduling* how traffic moves in the network's channel(s). Scheduling includes *routing* (a determination of the most efficient route within the network), *switching* (transferring traffic from one part of a channel to another or from one channel to another), and *capacity planning*. Traffic may be *stored* temporarily during scheduling. In communications and transportation networks the traffic is specific (i.e., not a commodity, like electricity or gas, and the particular vehicles or messages must get to specific places), so it must include the *address* of the receiver and, in many cases, the address also of the sender.

The most efficient routing is often, but not always, the shortest path between two points. A longer route may be more desirable if the shorter one contains a potential bottleneck, such as insufficient capacity at a natural barrier (a bridge), a switch (a road interchange), or damage to the channel (a burst pipeline). Mathematical solutions developed to determine efficient routes in a complex schedule are generally referred to as the "traveling salesman" models. Computer simulations have created new opportunities to model networks and to develop solutions to congestion problems.[7] Routing decisions for the network usually are not directly supervised by government, unless those decisions adversely affect quality of service.

An efficient switch moves incoming traffic to a new level of the network hierarchy in a minimum amount of time and with a minimum number of errors. This is a *sorting* function. The telecommunications switch directs calls from the level of the local loop to the appropriate local line or, for long-distance calls, the appropriate long-line service. A regional postal processing center takes mail from the many post offices of the region and routes it to the next level of sorting or to another local post office. Technology has changed many switching operations, making them faster, more error-free, and changing the level at which those operations can take place, usually moving them higher in the network hierarchy (e.g., postal sorting, which used to be done by postal carriers, is now largely automated at certain hubs).

Short-term capacity planning involves predicting traffic at peak times and making sure that the channel(s) can handle it without slowing the whole network. An examples of capacity planning is the "Mother's Day" phenomenon. The telephone network assumes demand will spike on Mother's Day. Therefore, when a certain capacity is reached, the system keeps callers out to avoid using resources to find paths for all the calls that would slow the entire system. Some cities keep automobile traffic off freeways (with traffic lights at entrance ramps) when a certain capacity is reached (e.g., rush hour) in order to keep the freeway network moving at an acceptable pace.

---

[7] See, e.g., Peter Weiss, "Stop-and-Go Science," *Science News* **156** (July 3, 1999), 8-10.

If keeping traffic out of the network at peak times is not acceptable, the network may need either to increase its infrastructure (e.g., add fiber-optic lines or more highway lanes) to handle peak loads (leaving them underused during off-peak times) or divert traffic to the channels of other networks at peak times. Increasing infrastructure may prove difficult if it cannot be accomplished incrementally but must be done in large chunks (e.g., by adding another power-generating station). Long-term capacity planning involves predicting how much traffic the network will be asked to carry in the future. This involves assumptions about new uses that consumers will find for the network and potential changes in usage patterns and demographics. Many telephone switches were designed for 3 to 4 minute calls and may be ill-equipped to handle the large amounts of "bursty" data signals of Internet traffic.

If the costs for new capacity must be added to a regulated rate base, the assumptions can be critical and controversial. They can be used to justify more investment in controversial power plants, airport facilities, telephone switches, the costs of which are then passed on to customers for electricity, air travel, and telephony.

If traffic must be stored because of an asymmetry in incoming and outgoing rates (a backlog is created when more comes in than can go out), additional cost is added to the system. Higher capacity switches may reduce these costs, but they may be more expensive than occasional storage (the cost of a new rail-switching yard may be higher than that of storing cars during a backlog).

If a sender can presort or preroute its traffic it will expect the cost-saving to be passed on to itself in a competitive environment (where the network's functions are not protected by a government monopoly). Businesses that presort their mail save the cost of having this scheduling function performed by the postal service and expect therefore to pay lower rates for postal traffic.

Ratemaking (pricing) is an important tool in capacity planning, because it can affect who uses the network, how much they use it, and when they use it. Many of the networks discussed here use or have used one or more of the following practices:

- *Peak and Off-Peak Pricing.* This practice is especially powerful, because it can be used to even out system loads and reduce the highest levels of peak capacity by encouraging a shift to off-peak times in exchange for lower prices. Some peaks are predictable—they occur at certain times of the day, seasons of the year, or points in the business cycle—while others—abnormal weather conditions—are not. Variable rates based on unpredictable peaks would cause customers to assume the risks of cost fluctuations.

- *Interruptible/Noninterruptible Pricing.* These distinctions allow customers to make decisions about risks for unpredictable peaks and about their own willingness to forgo service during peaks, when prices can rise dramatically. If no customers are willing to be interrupted, the network has the option in case of a capacity shortage of reducing service to everyone, which would mean a slowdown of the entire system.

- *Distance Pricing.* The distance between senders and receivers (or between hubs) is usually a cost factor, because the cost of distribution increases with each mile owing to the costs of rights of way and/or energy consumed (although these costs may still be lower than short-haul). This is politically sensitive in rural areas where distances are often greater so that rural residences and industries would pay more for their service.
- *Volume Discount Pricing.* Most network services can be sold at just above marginal cost to heavy-traffic customers. This is crucial in a competitive environment, where large customers are the least expensive to serve and potentially the most profitable. Heavy-traffic customers are also the most likely to bypass the network by building their own infrastructure, reducing capacity demands on the network. Volume discounts allow resellers (those competing for retail sales but which do not have their own local distribution) to compete with the local distribution company by aggregating the demand of many low-traffic customers to obtain lower rates.
- *Long-Term Contracts.* Long-term relationships allow senders and receivers to plan for capacity needs and to make a sufficient investment in the network to develop strong cooperative ties. Such contracts become a negative factor when cooperation begins to look like collusion to defeat the interests of other parties peripheral to the transactions, such as suppliers or customers.

3.9 Terminals

The point where traffic enters or leaves the network is here called a *terminal*. Terminals may be used both to send and receive (as in airports or telephones) or for one or the other (as in broadcasting, where a transmitter is used to send and a TV set to receive). Terminals may also store traffic temporarily off the network (as in oil depots and parking garages).

Like ancillaries (see section **3.10**), terminals must be compatible with the channel and sometimes with several channels. A grain terminal may need to be located on a waterway *and* a railroad *and* an interstate highway in order to schedule its traffic efficiently, but its facilities must accommodate barges, railcars, and trucks.

At the turn of the twenty-first century, debate has arisen on which terminal will ultimately bring digital information and entertainment into the home, the computer or the TV set. Makers of set-top converter boxes hope to supply the means to make either terminal compatible with the channel.

3.10 Ancillaries

Although a network's parts are *interdependent*, no network is *independent* of the world in which it operates. Building a network at any particular time and any particular place depends on existing political and economic conditions. A network is tied to the resources it runs on. Gas networks need gas—and they also need the right regulatory and financial environment, customers who are not "locked in" to another energy source, technical and managerial talent, and terminal equipment that customers find easy to use and economical to convert to, among other matters. (A network cannot require rocket science to convert a home furnace from oil or wood, and the consumer must be reasonably certain conversion will be worth the cost in the fairly short term.)

Ancillaries include the following:

- *Power sources,* e.g., electricity for computer networks and gasoline stations near roads
- *Production inputs*, used by one or more functions, such as paper for mail, light-weight metal alloys for airplanes, and low-resistance wires (usually copper) for electric transmission
- *Other networks* to transport critical inputs: an electric utility depends on the electric network to distribute its product; railroads or barges to bring coal; or a gas network to bring it natural gas; telecommunications networks to keep the system load even and report outages; mail service to send bills to customers; and road systems for trucks to maintain the system.

Ancillaries must be available at a cost that keeps the network competitive with substitute products or services. A sharp rise in the price of paper would hurt the mail service and make it less competitive with substitute communications networks such as telecommunications. On the other hand, if the price of postal service were similarly to increase, the value of paper would suffer owing to reduced demand. Thus, network managers and regulators are well advised to pay close attention to the network's ancillaries (and to the ancillaries of competing networks). The need for control to some extent over ancillaries often leads to the vertical integration of networked industries, even to the point of acquiring a resource outside the scope of "the network," such as when an electric company buys a coal mine.

3.11 Common Problems

3.11.1 Bottlenecks

Bottlenecks occur when one part of a network has lower capacity than the other parts. It may be a temporary problem, caused by a natural disaster, a technical failure, or a failure to build (or maintain) similar capacity in all parts of the system. Scheduling problems may mean that a railroad bridge may be able to handle less traffic per hour than the rail lines that connect with it. Building a bridge big enough to handle all the flow at peak capacity may be too expensive, so when traffic comes to the bridge it slows (even to the point of backing-up and queuing for access—with priority in the queue given to certain classes of traffic). If the bridge is washed out by a flood, traffic must be routed around it and its effective capacity drops to zero. Bottlenecks can be caused also both by the failure of members of a network to cooperate and by equipment upgrades (including upgrades to ancillaries, such as recording devices for TVs).

If bottlenecks become too costly, the network will need to find a way to build capacity at the bottlenecks. Building involves the investment by (or cooperation of) many parts of the network, and some parts may benefit more than others (e.g., those who had to pay a premium to be at the front of the queue). Working out the interests of all stakeholders can be a complex problem, and often government is called in to arbitrate. The introduction of competition into the network only increases the difficulty: players that can use bottlenecks (or their ability to get around them) as a competitive advantage have no incentive to cooperate.

The term “bottleneck” has also been used to describe the inability of new competitors to transport services in order to gain access to the network’s local branches (high-cost/low-return), which often makes building competing infrastructure uneconomical. These issues may be better thought of as issues of *access* (see section **3.11.2**), because adding capacity (either physical capacity or by better scheduling) does not address the real problem.

3.11.2 Access

When government does not own a network, it often finds it necessary to mandate access to privately owned network channels for various public purposes, including defense activities, and for relief from bottlenecks or scheduling problems. Mandated access for competitors is used also to insure the maintenance (or creation) of competition in the network.

In the United States, railroad industry-mandated access is accomplished through Directed Service Orders, which are issued by the federal government. These orders require one line to accept the traffic of another in case of damage to a right of way owing to a natural disaster or other emergency that has snarled traffic. They have been used also when a line has become insolvent to assure continuity until new management takes over. A Directed Service Order can also require one line to use the cars or engines of other lines to keep traffic moving.[8]

U.S. cable TV networks must give broadcasters access to their networks under “must carry” provisions of the law. Congress apparently felt that access to the cable is necessary for anyone trying to reach television receivers; if there are to be a variety of sources for information, cable systems must carry the signals of local broadcasters (their competitors for advertising dollars and viewers).

The price charged for access to a company’s infrastructure can be the most controversial part of these public policies, and government is usually forced to act as a referee in deciding on a price that is “fair” and not designed to keep competitors out of “essential facilities.” Certain network facilities or functions are assumed to be so expensive to construct or operate that they act as barriers to the entry of new product or service providers. The calculation of access pricing attempts to determine which of the incumbent’s costs can be passed on to competitors seeking access, but these calculations are an accounting nightmare and become part of the political

---

[8]See Frank J. Dooley and William E. Thoms, *Railroad Law a Decade After Deregulation* (Westport, Conn.: Quorum Books, 1994), 55.

debate. The price for access to a company's infrastructure remains a difficult theoretical and practical problem for policymaking in many of the networks discussed here.[9]

3.11.3 Small Loads vs. Large Loads

Small loads or batches of traffic are more expensive to handle per unit than large ones, because they require more switching and routing. A carload of books is cheaper to handle *per book* than a carton of books. Delivering 10,000 megawatts of electric power to one customer is cheaper than delivering the same amount to many small customers.

The much touted commerce on the Internet that allows small businesses to sell to anyone anywhere in the world may lead to intense price pressure for many items, forcing those items into commodity status (e.g., flowers or books) and risking the extinction of local retailers. Internet commerce may increase the traffic in transportation networks for small loads over long hauls for delivery to Internet customers and require a massive reorganization and growth in the capacity of transportation and postal and parcel networks.

Consolidators or *forwarders* aggregate traffic to take advantage of the lower rates offered to large customers or for large loads. Traffic can be stored until the optimum amount is ready for transport or until capacity in the channel can be purchased at a reduced cost and resold at a profit, though often a rate lower than that charged by the channel owner. This occurs in telecommunications, freight shipping, and passenger traffic for railroads and airlines.

3.11.4 Short Haul vs. Long Haul

Most networks start as a collection of disconnected short hauls that later connect to form a larger system. The first railroads were built to connect a quarry or mine with a river or barge canal and only later to connect several cities. Interconnection was difficult (it was discouraged by competition between lines), and passengers and freight were necessarily moved from one terminal to another outside the system. Similarly, the first telephone systems connected only local businesses (and local residents). Gas was distributed first only locally, from the facilities where it was processed from coal. The value of being connected with a network rises with the number of people connected, so interconnection is important for wide adoption and, in some cases, has become an important competitive tool.

The economics of short haul and long haul, however, are quite different. Long-haul routes typically are cheaper (per traffic mile) to build and meet less local citizen and political resistance, because they usually go through less populated areas. The costs of acquiring right of way for interstate highways and long-distance telephone lines usually are less than those of right of way

---

[9]See, e.g., James R. Ratner, "Should There Be an Essential Facilities Doctrine?" 21 *UC Davis L. Rev.* 327 (1988); and Alexander C. Larson and Dennis L. Weisman, "The Economics of Access Pricing, Imputation and Essential Facilities, with Application to Telecommunications Policy," *Communications Law and Policy* **3**, 1 (Winter 1998), 1-33.

in metropolitan areas, where the price of real estate may be much higher. Long-haul routes are thought to have more or less constant returns to scale and can maximize efficiency by scheduling traffic for higher capacity use (i.e., sending fewer batches of traffic but making them larger). In some cases, long hauls are more efficient per mile, because the cost of frequent starts and stops is high, e.g., airplanes, which use more fuel on takeoff and landing than in flight, or a network that must pay each time traffic enters or leaves the system at a terminal. Many long-haul operators are less adversely affected by increases in the costs of production inputs (e.g., fuel) or ancillary services (tugboat services) than local or short-haul operators.

Short-haul routes require more costly, individualized scheduling and signalling and are generally less efficient to operate and more expensive to build than long-haul. High sunk costs may make them a “natural monopoly,” because no one would invest the money necessary to set them up without the assurance of the returns made possible by monopoly rents.[10] Short-haul routes have different economic realities from long-haul and can be separately owned or operated parts of a network. Many of them consist of many local companies (or local branches of large companies) that deliver traffic bound to or from locations outside their region.

The short-haul portion of most networks is regulated in the United States by individual states, while interstate long-haul is regulated by the federal government. The impact of competition on former monopoly network elements at both short-haul and long-haul levels may be very different, but, as of late 1998, the parameters of the difference were only speculation. If competition reduces long-haul rates to something near cost, then the price for service on heavily trafficked lines may well go down and activity sensitive to this change in price may concentrate at points along those lines. The concentration would affect the economics of switching as well as other bottlenecks on those lines and might have a profound impact on the geography of economic development and business location patterns. Unfortunately, policy discussions seldom address these possibilities.

[10]For a discussion of this issue in postal and telecommunications networks, see, e.g., John C. Panzar, “The Economics of Mail Delivery,” in *Governing the Postal Service*, edited by J. Gregory Sidak (Washington, D.C.: American Enterprise Institute [AEI] Press, 1994), 1-29.

Chapter Four

Regulating Networks

4.1 General Considerations

The networks discussed here all have used public money or resources to build their infrastructure and are regarded as crucial to public welfare. Economists refer to their services as "public goods," which all citizens use and which therefore should be supported by government. Government either owns or regulates these networks to assure "reasonable" rates and "universal" service. The value of public goods is often expressed as "social savings" and calculated by estimating the costs to an economy if the network were shut down.

There is also a more cynical view of the politics of monopolies created to deliver public goods. As Thomas Jefferson recognized, politicians may pursue the single goal of creating monopoly rents to distribute them to favored constituents and constituencies as a form of patronage. Although the identities of favored groups and the mechanisms for the distribution of monopoly rents has changed over time through a process that may be called dynamic capture, the basic political goal remains the same: Reward those who reward you.[1]

Regardless of motivation, governments soon find that they must protect the monopolies they have created from businesses that would take a portion of the monopoly's revenue or market. The borders of the granted monopoly franchise are always tested by those who see a chance of high profits or low costs. An early example in the United States was in the postal system and involved the definition of a term that has come to have new meaning in modern telecommunications: *packet*.[2] The postal laws of 1792 gave the new postal service a monopoly over the transport of any "letter or letters, packet or packets, other than newspapers."[3] When in 1831 a business in Maine announced that it would deliver commercial documents (thus, saving customers high postal fees), its owner was charged with violating federal law that protected the postal monopoly. The federal court determined that "packet" meant a bundle of letters and that the legal documents at issue in the case were not letters. Thus, not all paper messages moved between cities were subject to the monopoly.[4]

In the twentieth century, postal authorities took the position that their monopoly includes the transport of payroll checks, Walt Disney™ posters, tickets to professional football matches,

---

[1]Richard J. Pierce, "Commentary," in *Governing the Postal Service*, edited by J. Gregory Sidak (Washington, D.C.: AEI Press, 1994), 111.

[2]One method of routing and switching digital messages is to form them into individually addressed *packets.*

[3]Act of Feb. 20, 1792, ch. 7 section 14, 1 Stat. 232, 236.

[4]*United States v. Chaloner* 25 F. 392 (D.-Maine 1831).

computer programs, and documents that electronically transmitted (if these are converted to hard copy when carried from sender to receiver).[5]

With each breach of the fence around a network monopoly, revenues flow out of the network, leaving less to invest in infrastructure or subsidize service. The only customers or services left for the monopoly may be those no one else wants: customers that cannot be served at a profit or those without the political muscle to demand a loophole that will allow them to operate outside the network. New technologies have allowed networks to become more *permeable*, allowing competitors for their services to invade the network and sell services formerly available only from the monopoly provider.

Government involvement in networks at the federal or national level has been justified in the name of military necessity.[6] In time of war, the efficient movement of troops, supplies, and information is critical. National defense (an indisputably federal concern) thus becomes a major legal justification for federal involvement in the building and regulation of networks for transportation, communications, and energy.

Another justification of federal aid and control of transportation systems was delivery of the mail. It was part of the political and economic equation of roads, railroads, and airlines (see **Chapter Seven**). In addition to federal support for the channels through which the mail would go, the government supported many of the vehicles that carried mail through the channels. The earliest form of federal support was generous contract terms given to the stagecoach lines that carried the mail, which were particularly important in the south and west, where a mail contract often meant the difference between life and death for stagecoach businesses.[7]

Government can become a customer of the network at relatively high prices in order to add resources that act as internal subsidies. Government purchases to add resources to the system (or furnishing resources directly) can become controversial in a system of intranetwork competition, because the purchases act as subsidies to some competitors but not to all and can tip the competitive balance. Amtrak is heavily subsidized by the federal government. The fees government pays railroads to use their lines for Amtrak trains have a substantial impact on those railroads' overall profitability. If the fees a railroad charges the government for Amtrak access

---

[5]See Edward L. Hudgins, *The Last Monopoly: Privatizing the Postal Service for the Information Age* (Washington, D.C.: Cato Institute, 1996), 18-21.

[6]This argument became particularly important after the nation's first experience defending its territory in the War of 1812, when many battles were fought in newly acquired territories where there were no efficient transportation routes. Ibid., 31-37, 137-139. This argument provided the major justification for the use of federal money (and control) for the Interstate Highway System. See Tom Lewis, *Divided Highways: Building the Interstate Highways, Transforming America* (New York: Viking Penguin, 1997), 107-08.

[7]Richard R. John, *Spreading the News: The American Postal System from Franklin to Morse* (Cambridge, Mass.: Harvard University Press, 1995), 93-94.

reflect a high proportion of that railroad's total costs, then taxpayers would be subsidizing the railroads chosen for Amtrak service.

Government has helped build networks by acting as something of a marketing tool for the new technology, perhaps by using that technology for public sector business, for education, or for other public goods, such as medical care. Government has directly subsidized demonstration projects to show the public what a network could do. In 1844, the federal government helped to build the first telegraph system, which was between Baltimore and Washington.

The growth of network-based industries has many analogues in the growth of biological networks. Alfred Lotka, one of the first to describe competitive equations in biological networks, noted that his work seemed to apply also to growth curves for the railroad industry. In *Elements of Mathematical Biology* (1924), Lotka observed that a species (or technology) that is fitter (as described by certain factors) than its competitor will eventually replace that competitor if both are competing for the same scarce resource and neither has other sources with which to meet this need.[8] But, as noted in **Chapter Two**, networks that invest large amounts of capital in infrastructure do not die suddenly, because they find traffic to carry for which they are the fittest network available.

Government's role in the ecology of the new competitive networks may have changed from planner and director to arbitrator of disputes among contestants in a larger, more open system. This role may entail just as much regulation as in the old system of central planning but is more concerned with multidimensional controls on price, interconnection rules, technology standards, and access to markets.[9] Policymakers are always asked to create a "level playing field" during the transition to a competitive market, but they often forget that true competitive equilibrium rarely exists either in nature or in markets and, even if it were established, it would last only as long as it would take for the players to adjust to the new rules.

Channels have been regulated in three ways: as common roads, common carriers, and private carriers.

Common roads are channels that accommodate traffic of many sizes and shapes, and the owner of a channel does not control the vehicles that carry traffic through the network. Common roads are generally (but not necessarily) owned by government, which may charge when traffic is put into the network (e.g., toll roads) and for the use of bottleneck facilities, such as bridges or locks, in order to control capacity. Senders choose their routes within the network. Common

---

[8]This is known as the Lotka–Volterra equation. See A. Lotka, *Elements of Mathematical Biology* (New York: Dover Publications, 1924; reprt. 1956); and A. Volterra, *Leçons sur la Théorie Mathematique de la Lutte pour la Vie* (Paris: Redigées par M. Brelot, 1931). For a detailed discussion of biological models for business competition, see P. H. Longstaff, *Competition and Cooperation: From Biology to Business Regulation* (Cambridge, Mass.: Program on Information Resources Policy, Harvard University, P-98-4, October 1998).

[9]G. J. Mulgan, *Communication and Control* (New York: Guilford Press, 1991), 159.

roads are not responsible for the safety of traffic in the system, except insofar as they keep the channel free from obvious hazards. Highways and canals are common roads, as are public-access channels on local cable TV networks. Digitization of information may make it possible for more communications channels to become common roads.

Common carriers carry the traffic brought to them over a channel that they control. They control all aspects of the traffic flow in the channel—routing, switching, transport, and, occasionally, terminals. They do not own the traffic while it is in the channel, and they are not responsible for damage it may cause to others or for a violation of the law. Transportation common carriers are generally responsible for the safety of the traffic while it is in the network.[10] Common carriers have sometimes been given a monopoly to serve a part of a network in order to recover risky up-front investments, but because they were considered heavily invested with "public interest," most of them have been owned by government or regulated as to level of service, duty to accept all traffic brought to them, and rates charged. The trucking industry, railroads, telephone service, and leased-access channels in local cable systems all have been treated as common carriers.

Private carriers own (or control) both the channel and the traffic they carry through it. They control all aspects of the network and are rarely regulated by government (although private carriers may use part of government-owned or -regulated channels, such as common roads or common carriers). Private networks exist in gas pipelines, telecommunications (including corporate and government intranets), package delivery, and air transport.

4.2 Effects of Competition (It's Not "Deregulation")

The pattern that has been set when competition is introduced into networked industries can be described as follows. Although late-twentieth-century political debates often call this introduction of competition into networked industries "deregulation," the relative level of regulation does not go down, but, rather, shifts focus. Networked industries are still regarded as critical to industrial economies, and no one (including new competitors) really trusts the marketplace to deliver a system that will please all stakeholders. Competition seems to mandate new regulations to establish rules of fair play, which can be accomplished by specific rules in each sector or by application of antitrust laws.

Competition often brings both the expected reductions in consumer prices (although large customers benefit far more than small) and the deployment of new technologies, but it can bring unexpected problems. Policymakers and executives do not seem to realize that competition can also reduce cooperation in a network, fragmenting a unified network into several competing ones. The resultant duplication of infrastructure can be costly for consumers and give regulators

---

[10]The law considers them the temporary custodian of the goods or the "bailor." See, e.g., Edwin C. Goddard, *Outlines of the Law of Bailments and Carriers*, 2nd ed. (Chicago: Callaghan Publishing, 1928).

headaches. Given that providers no longer cooperate, some other form of coordination becomes necessary. Usually, coordination means increased government regulation.

Companies that once cooperated to keep The Network efficient by moving one another's traffic, jointly administering scheduling, terminals, and security measures, participating in systemwide pricing which spread costs and profits, and clearing bottlenecks, now, with the introduction of competition (beginning the 1970s, with "deregulation"), fragmented into competing hub systems that offered full service in the attempt to keep customers away from other players. In air transportation, hubbing meant several plane changes for travel between two points at the outskirts of two systems, with flights often many miles out of the way to get to connecting hubs. When deregulation brought down the price of airline tickets to major hubs (high-traffic), the price of tickets to small cities (low-traffic) went up and service was either curtailed or turned over to smaller carriers. In the late 1990s, similar scheduling and pricing patterns appear to be developing in telecommunications, where large companies are "spinning off" rural service to smaller companies.

When government lowers barriers to entry, new players will quite naturally look to markets in which profits may be easiest—long-haul and high-traffic. These markets are the customers that pay the highest rates to subsidize short-haul and low-traffic customers. Incumbent networks then decry taking the revenue these customers had provided out of The Network as "cream-skimming." Eventually, opportunities for entry dwindle as the leading players develop economies of scope and scale, and then those players increase their marketshare. Competing companies sometimes seek refuge from the pain of competition by dividing up markets by market segmentation or territorial division. This practice leads to Congressional investigation and to calls for antitrust enforcement.

Owners of (or those that control access to) the channels are often tempted to use control of access as a competitive weapon, prompting government to move in and mandate access to airport terminals, local telephone systems, and power lines. Competitors often describe their inability to gain access to channels as a "bottleneck" in the system and often ask government to declare the channels "essential facilities" for the purpose of developing and maintaining competition. Government is also asked to regulate the prices that channel owners charge for the use of their infrastructure.

Perhaps the most puzzling result (but, in retrospect, the most predictable) is the wave of mergers and acquisitions that hits networks after "deregulation" occurs. Introducing competition into the systems causes players to do things that may earn them the perceived competitive advantage of being the biggest kid on the block.

4.3 Effects of Competition on Rates and Service

In networks that do not allow competition (i.e., those in which government establishes territories and keeps out competitors), rates are heavily regulated to keep the monopoly service provider from charging monopoly "rents"[11] and to insure customers "fair" prices. Political considerations might require that all customers in the network be charged roughly the same price. But when networks cover large areas and serve all customers in those areas, the economic problems of short-haul and small loads must be compensated for by spreading the costs of these less profitable (or unprofitable) services over prices charged to all the customers in the entire network. Low-cost customers (urban, industrial) end up paying more for service in order to subsidize high-cost customers (rural, residential). Interestingly, consumers often prefer "flat" rates—unlimited access to the network for a set rate per month—even though these rates mean (theoretically) that half the usage (but not necessarily half the users) pays more than their share of the costs. People seem to feel they are getting a better deal by having unlimited access to capacity. This phenomenon may be related to what is known as the "tragedy of the commons,"[12] that is, when use of a resource is not related to the price and easily outstrips capacity.

A system of regulated monopolies requires regulators to approximate costs for operating the network and to oversee the allocation of those costs to various classes of customers at both wholesale and retail levels. Allocation is necessary to favor politically important customers (e.g., residential) and to avoid internal subsidies not approved by regulators. A vertically integrated monopoly provider could allocate more of its costs to one class of service in order to lower them for another class of service where its subsidiary is a customer. A gas company could thus give lower rates to electric utilities in order to favor its subsidiary electric company.

Even when competition is introduced into a network, the need to supervise rates does not seem to disappear entirely. Regulation is often considered necessary during a transition in order to protect customers and keep the competitive playing field fair for new entrants. Regulation of cost allocations becomes particularly controversial during a transition, because regulators need to make sure that rates do not increase dramatically for those previously subsidized—which might prove uncomfortable for some elected officials. The political power for subsidies for rural areas has faded to some extent in the late 1990s as voters increasingly move to urban areas. Such changes in demographics were part of the political calculation after the U.S. Supreme Court in the mid-nineteenth century mandated "one man, one vote" as the operating principle for dividing congressional districts, moving power to where the people were.[13]

---

[11]Economists define this as a rate that exceeds the cost of service and a reasonable rate of return, and it is caused by a lack of competitive pressure to keep rates closer to cost.

[12]See, e.g., N. Gregory Mankiw, *Principles of Microeconomics* (New York: Harcourt, Brace College Publishers, Dryden Press, 1998), 227-233.

[13]See John C. LeGates and John F. McLaughlin, *Forces, Trends, and Glitches in the World of Compunications* (Cambridge, Mass.: Program on Information Resources Policy, Harvard University, P-89-2, May 1989).

The introduction of competition into networks has seldom resulted in building redundant channels for short-haul or local traffic. Instead, competition at the local level usually takes place in the provision of services using existing infrastructure. Competition for transport services requires owners of the channels in the network (who also sell transport services) to provide access to new transport competitors. Regulators are required to continue their review of cost allocations in order to insure competition in the system. And they may need to make sure that a channel owner does not charge competing transporters rates such that they bear more than their share of the costs for operating the channel. To do so would mean that the channel operator could charge its customers lower prices that would eventually put competitors out of business.

Regulators hope that the new competitors in the networks will compete both on price and service, but telecommunications, transportation, and energy services are essentially commodities and the real competition is on price. Reducing prices in order to compete means cutting costs, which sometimes results in a lower quality of service. Dependability suffers, for example, when redundant systems become too costly, as does maintaining excess capacity to handle unexpectedly high traffic loads. By separating ownership of a channel from ownership of the rest of the network, internal subsidies, which once paid for upgrades and maintenance, are lost and these resources are used to reduce prices in search of marketshare. Since competition has been introduced, quality of service has not gone up in railroad, airline, telecommunications, or electric services, but what is rising in each of the networks discussed are customers' complaints about confusing options and fraud.

Chapter Five

Transportation

5.1 Roads, Bridges, and Ferries

Early roads followed trails, often originally animal routes, that had been worn by indigenous populations to facilitate trade or religious rituals.[1] Roads tended to follow river valleys but were on ground high enough to remain passable during floods (thus "highways"). Because the roads usually were the most efficient routes, they became the locations for railroad lines.[2] Before the American Revolution, few roads were built in this country, because England discouraged the movement of goods among colonies in order to foster trade with home country and even at one point such movement was forbidden.[3] After 1776, the river valley roads (bad as they were) brought economic growth into the areas they passed through, because they allowed agricultural commodities and manufactured goods to move from producers to consumers. The new Constitution gave Congress the power to establish "post roads," which connected post offices throughout the country and sometimes crossed state lines. Thus began the federal authority over and funding for communications infrastructure and an uneasy, shifting truce between the states and federal government in the battle over which would have jurisdiction.

Government funding for a network of roads was also the beginning of the political necessity of universal service. Post roads were requested of Congress by local residents, requests that were almost never refused because every member of Congress voted for the roads of all other districts in order to be able to bring home roads for their own district. These roads often were little more than forest paths and, except on main routes, often poorly maintained, but they allowed fairly regular delivery of the mail to most communities.

The connection between transportation and economic prosperity led to a new kind of road, one built outside the old trail routes and with the express intention of bringing economic development to new areas. One early example of such roads was the Cumberland Road, which was built in early the nineteenth century to increase commerce between the east coast of the United States and the newly acquired territory northwest of the Ohio river. The Cumberland Road was one of the first infrastructure projects funded (in part) by the federal government. All the states the road traversed had an interest in its use, but no one state could afford to build the road

---

[1]The information presented in this section is drawn largely from three sources: *An Encyclopaedia of the History of Technology*, edited by Ian McNeil (London and New York: Routledge, 1996); *History of Transportation in the United States Before 1860,* edited by Henry Balthasar Meyer (Washington, D.C.: Carnegie Institution of Washington, 1917); and Tom Lewis, *Divided Highways: Building the Interstate Highways, Transforming America* (New York: Viking Penguin, 1997).

[2]*History of Transportation*, 6-7.

[3]Ibid., 65.

on its own or to gain access to the land of the other states. Federal involvement was critical, because the government could bring all the conflicting interests of the states together to fashion both a larger plan and the political compromises necessary for selecting a route. Washington was able, for example, to make sure that the road built in one state connected with sections built in the others even when the links did not serve particular local interests.

The routes of these new roads were dictated by commercial interests, not by the shortest distance between two points. Routes were bitterly fought over both within and between states as cities jockeyed for the greatest benefit. New York, Tennessee, and Louisiana opposed the entire Cumberland Road project, because it would compete with their waterways (the Erie Canal, the Mississippi River) for traffic. In 1806, after intense pressure both from the new western states, which wanted better access to markets for their agricultural products, and from eastern states, which wanted access to western consumers for their manufactured goods.[4] Because rivers in the eastern United States tend to run west to east (down to the Atlantic), few inland routes were maintained between northern and southern states. The result was a division of commercial and cultural interests and a weakening of any development of shared "national" interests.

As early as 1791, road building had become a "mania" as states became convinced that without roads they would be left out of the country's economic expansion. In 1785, President Jefferson saw them as important to the new "national" life of citizens of the individual states.[5] When many such roads failed to deliver a local economic boom, they were afterward abandoned or taken over by local authorities. In the early 1800s a report on transportation commissioned by Congress warned that roads alone could not bring prosperity. Their success depended rather on other "improvements" related to commerce, such as loading facilities, and connections with other transportation systems, such as canals and rivers, which would make the roads part of a national system and move products to markets within and beyond the United States.[6]

Many of these early roads were "turnpikes," or toll roads, built and maintained by local governments or private entrepreneurs under grants from local authorities. As unpopular as they were among those who paid the tolls, the toll roads were often better maintained and carried more traffic than the "free" ways. But competition from steamboats and railroads led many turnpikes to become neglected. Without the number of tolls they had been accustomed to collect, the operators (private or government) could not afford the costly maintenance necessary. The poor condition of the roads, in turn, only increased calls for more railroads.[7] And "free" roads became an economic (thus, political) imperative as transportation grew into an important part of the overhead for

---

[4]Ibid., 12-18.

[5]Ibid., 51.

[6]Ibid., 135-136.

[7]Ibid., 53, footnotes 2 and 3.

commodities and manufactured goods. Lower transportation costs could make the products of one state very competitive with goods from other states and with imported goods.

In the mid-1800s, lawmakers began to consider public funding of roads important for maintaining competition with the railroads, particularly funding of schemes for “plank” roads made of logs, which allowed wagons to move faster and remained viable during muddy seasons. Plank roads were touted as a basis for *economic development* of new territories, but some legislators saw the lessons of putting the road before the cart. A report prepared in 1848 for the State of Wisconsin noted that “dearly bought experience of the past” had taught “that to thrust improvements in advance of the business to sustain them is futile,”[8] but most citizens and policymakers believed (and, indeed, continue to believe) that infrastructure is necessary to attract development, thus the road should be built first.

Grassroots political support for government spending on new roads initiated a pattern that would be followed by most of the networks discussed here. Government spending was “encouraged” by a Boston bicycle manufacturer, who organized riders into a lobbying group, financed courses in road engineering at the Massachusetts Institute of Technology, built a short stretch of paved road in the city to demonstrate how the improvements would affect public safety and enjoyment, and got Massachusetts to create a highway commission.[9]

Rivers and other geological features created barriers to long-distance roads. Ferries or bridges were established to retain an all-land route in spite of these natural barriers and to eliminate long detours. Like the first roads, ferries and bridges were often built and operated by local governments or by private interests which had obtained grants (usually monopoly rights) from local government.[10] Local grants often contained reversions to the local government within a period of years after which, presumably, the entrepreneur would have recouped the initial investment and made a “reasonable” profit. Ferries were politically more acceptable than bridges often, because ordinary citizens seemed to resent paying bridge tolls intended to repay invested capital but did not mind paying for the “services” of a ferryman.[11] Toll roads, bridges, and ferries all had different fees for different types of cargo, which allowed local authorities to subsidize some goods at the expense of others.

Ferries (and later bridges) became *bottlenecks* for traffic, that is, places where traffic was forced to slow in order to get through the limited capacity in a channel. From investment in infrastructure and from the prices they charged, ferry operators, like *gatekeepers* of turnpikes, gained the power to speed up or slow traffic. Because the ability of a ferry (or bridge) to cross a

---

[8]Ibid., 303-304.

[9]Lewis, 7-8.

[10]For a history of early ferries and bridges in the United States, see *History of Transportation*, 37-50.

[11]Ibid., 254.

river was a scarce resource, and because ferry owners often had monopoly rights, the prices charged reflected the owners' power. High prices led to building bridges to compete with the ferry or to local regulation of the prices charged by the ferry owners. Were the prices too low, ferry owners could not afford to upgrade equipment to increase capacity when demand increased, so that low prices, too, in effect led to calls for bridges.

New bridges were often hotly contested by incumbent ferry operators but supported by merchants claiming that ferry operators charged too much or could not move goods quickly enough. Ferries operated by states were among the last to get competition from a bridge. The ferry at Albany, the state capital of New York, was founded by local government and later taken over by the state. No bridge was built over the Hudson River until 1856. As in many other cities, a bridge was opposed because it might "interfere with navigation."[12] But even after bridges crossed most rivers, ferries continued to operate where bridges were not possible, and continue still to do so.

As in road building, the locations of new bridges could prove the subject of political controversy. A bridge could insure the economic prosperity of a nearby town. The first bridge over the Potomac to the nation's capital was delayed for several years (1805–1806) while Congress witnessed intense political wrangling between the commercial interests of Georgetown and Baltimore. As late as 1852, citizens of Georgetown demanded relief from Congress for having been "grievously injured for the long period of 45 years by the erection of the Potomac Bridge below the harbor of the town."[13]

The experiences of builders, state and local government, supporters and travelers of early roads were seen again writ large in the 1960s and ´70s, when the Interstate Highway System was built. The federal Highway Act of 1956 authorized construction of the largest engineering project in the United States. The project was supported by taxes on gas and diesel fuel, which went toward the Highway Trust Fund. Federal spending on the project was justified on the basis of military considerations (the evacuation of cities and the movement of troops and supplies in case of potential atomic attack). Public support for superhighways was based on a belief that they would bring economic development to rural areas and reduce congestion in urban areas.

The effect of the interstates on military preparedness and economic development was not exactly as envisioned. Congestion made these highways useless for evacuation or military transport. In networks accessible without charge to either sender or receiver, new capacity creates new demand as new uses for the free channel develop. Traffic engineers call this new demand *induced demand*.[14] These highways made new suburbs possible and brought development in the

---

[12]Ibid., 42-44.

[13]Ibid., 43, at note 4.

[14]Lewis, 291-292.

form of strip malls and shopping centers, which killed downtown retailing and allowed businesses and their employees to move out from central cites. Many interstates were congested soon after being built owing to the effects of induced demand.

In the central cities, the interstates devastated whole neighborhoods by eliminating them, dividing them, or making them too dark or noisy to live in, and even obstructed local traffic. The Highway Trust Fund developed to build these highways drained government attention and resources from mass transit, which was far more appropriate to urban areas. Urban politicians ultimately busted the Highway Trust in 1973 and opened funding up for mass transit.[15] By the 1970s, cities were refusing to build local interstates, and people were disgusted by billboards and the political corruption accompanying the billions of dollars that flowed into road construction. Although opponents of highway "progress" were branded "luddites," highways were clearly not the "one" answer to all the transportation problems in the United States and often brought as many problems as benefits.[16]

Committing government resources to a competing channel can handicap some channels and kill others and create a dependence on one system for critical elements of commerce, such as delivery of manufacturing supplies.[17] People and resources were pulled out of small communities into centers of commerce made possible by better, faster road networks. The roads that took people more quickly to the county seat (and the large stores there) also brought franchises to the town's nearest exit ramp. Local retail could not compete. The roads also hastened the day when the local rail spur was abandoned in favor or road traffic, and small communities were often left to depend on cars for passenger transportation and on the trucking industry to take crops and goods to market.

Access to roads is not possible for those who don't drive (that is, those under age sixteen or those with certain handicaps, or those who can't afford to purchase and insure a car). Where road capacity reaches its limit, access may be denied (via regulated entrance ramps) or allocated by market mechanisms (via tolls). In some cases, the roads have been privatized, and the rights to operate them auctioned by government.[18]

Even for businesses with unlimited access, dependence on one channel means they are threatened not just by the closing of that channel but also by interruptions to the ancillary functions necessary to use it. Ancillaries to roads include fuel facilities (from fodder for horses to

---

[15] Ibid., 221-233.

[16] Ibid., 125-153, 163-169.

[17] Efficient transportation has become increasingly important to manufacturers in the 1990s with the adoption of "just-in-time" systems, which, in order to avoid the costs of warehousing, require the delivery of raw materials just in time to be used.

[18] See Eduardo Engel, Ronald Fischer, and Alexander Galetovic, "A New Method of Auctioning Highways," *Private Sector* [a publication of the World Bank] (June 1997), 8.

gas for trucks) and parking. Ancillaries must be present for a road system to develop and thrive. Road infrastructure, like all networks, must be maintained, but maintenance costs seldom are included in the figures talked about in the planning and building stages. In the 1990s, U.S. interstate highways as well as secondary roads and bridges were in grave need of repair, but government appropriations remain more easily obtained to build new roads than to keep old ones in repair. Much of the damage to roads and bridges is caused by the industry that they helped to create: trucking.

Like railroads and airlines, trucking originally was regulated (at the request of the railroads) as a *common carrier*, i.e., entry and exit from the business was controlled, standards for quality of service were established, truckers were required to accept any cargo brought to them and could charge only the legally allowable limit. Like other modes of transportation, trucking was considered critical to commerce in the communities served. In the 1980s, the trucking industry was deregulated and developed the same pattern of adjustment that followed airline deregulation: many mergers and acquisitions as well as changes in cost structures and profitability.

5.2 Waterways

Water has always been a vehicle for as well as a barrier to transportation. It is the motive force for construction of both boats and bridges. For most of recorded history, access to water routes and bridges over them for land traffic has been a key to the success of both commerce and war. Ancient canals, such as the one constructed in China in 215 B.C. between the Yangtse and the Hsiang rivers, were public works designed to facilitate the movement of troops and war materiel. The Roman empire (and later empires) built many bridges and ships to keep commercial and military interests intact.

In the seventeenth and eighteenth centuries, when most U.S. roads were still paths through the forests, trade and passengers moved along coasts and, later, into the interior on the nation's rivers. Early cities were founded at the sites of good harbors or the junctions of major rivers. With the advent of steam power in the early 1800s, rivers became the major transportation routes, particularly in the southern United States, where plantations had little use for local roads because most of their commerce was conducted with coastal cities and European merchants. As early as 1817, a steamboat company was given a monopoly on all the rivers of Georgia, and the state began improvements on the rivers to accommodate boats through locks and canals and to give the steamboat company access to overland routes in times of droughts and low water. The development of a comprehensive system for steam traffic was complicated by rivalry among Savannah and Augusta, Georgia, and Charleston, South Carolina, each of which saw steamboats as critical to their economic development.[19]

---

[19] *History of Transportation*, 256-260.

Rivers were not always navigable along their full lengths and were the object of the earliest public improvement projects of local colonial governments dredging new channels and clearing obstructions. Many people came to believe that artificial waterways, like European canals, could open up the nation's vast western inland territory without rivers.

The largest and most famous canal project was the Erie Canal in New York State, built in 1818–1825 to connect the Mohawk River with Lake Ontario and the other Great Lakes.[20] The enormous expense of digging and maintaining the canal was justified on the basis of the high cost of shipping goods in the other available channels. At that time the roads in that area were so bad and the tolls so expensive that it was cheaper to ship goods from New York City to Buffalo along the Atlantic coast and up the Mississippi and Ohio rivers than to move them overland only a few hundred miles to the east from the terminus of the Mohawk River. The cost was especially high for heavy commodities, such as the salt and mineral ores coming from upstate New York and Michigan.

Without a less expensive route to the new territories, New York City would have lost its trade base to such cities as Baltimore and Philadelphia. Because the high cost of transportation became part of the cost of the goods shipped, farmers and miners of upstate New York could not compete against the cost of goods from as far away as Michigan and Ohio.[21] The Erie Canal project replaced an earlier public and private partnership to improve the natural waterways of the Mohawk River system. The new Canal was built as a state enterprise, because private credit was not available for a transportation network with such high costs and uncertain profits. Lawmakers justified state investment in the Canal on the ground that benefits that would be widely dispersed to people, such as local landowners and merchants, who did not pay tolls yet profited from the Canal's military significance. Thus, the Erie Canal, like the roads and bridges that crossed it, was thought to be what economists call a "public good."

Yet this was a public enterprise with private competition. The Erie Canal had to be competitive with land-based transportation as well as with other water routes. As local governments improved roads, turnpike owners (and teamsters who worked on those turnpikes) lowered rates to meet competition from the Canal, and, in turn, canal authorities reduced their rates. *Bottlenecks* in the canal system, caused by delays at locks[22] and high fees at loading terminals, made overland transportation more attractive. Like the railroads built later, the Erie Canal needed branches to *feed* local traffic into the main channel. This need led government to

---

[20]See Noble E. Whitford, *History of the Barge Canal of New York* (Albany, N.Y.: J. B. Lyon Co., 1922) and Charles T. O'Malley, *Low Bridges and High Water: On the New York State Barge Canal* (Utica, N.Y.: Diamond Mohawk Pub., 1991).

[21]By 1833, the cost of transportation between Buffalo and New York City had dropped from $100 to $40 per ton. McNeil, 511.

[22]In the mid-nineteenth century, it was not unusual to see sixty to seventy boats waiting to go through a lock. O'Malley, 57.

build more local canals to connect with the Erie Canal in order to create a transportation network that could haul goods from upstate mines and farms to New York City. The local canals (and, ultimately, the Erie Canal, too), however, could not withstand competition from a new transportation network making its way along old wilderness paths: the railroads.

Other actions of local and state government may unwittingly have contributed to the demise of the Erie Canal and the canal system. In an attempt both to limit the power of barge operators and to protect small boatmen, the New York State Legislature declared that no barge corporation with more than $50,000 in stock was permitted to operate on the Erie Canal. The unfortunate outcome was that companies thus could not grow large enough either to deal with the large companies that controlled the elevators and terminals or to operate with the greatest efficiency. The limit on stock was later lifted, but the legislature restricted cross-ownership of canal operators and railroads, discouraging cooperation in bimodal transportation. To diminish the power of terminal owners, the state built public terminals but did not include in them any facilities for grain, an omission that meant grain had to be sent by rail and that took badly needed revenues out of the canal system.[23]

As railroads became an important player, the State of New York tried to force them by regulation to cooperate on the canal system in order to save the state's investment in it and for the benefit of constituents who were paying shipping charges they felt were too high. The regulations were an attempt to give customers of the canal system access to the railroad system by regulating rates on joint rail and water routes (i.e., to force the railroads to prorate goods that came part of the way by water) and to make the railroads allow boatmen to unload near railroad terminals, thus saving the boatmen the cost of transportation by wagon between terminals.

Eventually, New York State decided to upgrade the canal system to accommodate large barges. Its investment, however, became part of the canal's problems and caused dislocations among the businesses that served it. Because the new barge canal did not include tow paths, animal-drawn barges could not be used. The state was forced to provide tugboats for older boats that were not self-propelled, a money-losing proposition that also increased congestion and bottlenecks when the tugs were not efficiently scheduled in the system. The state then bought larger tugs to accommodate larger barges, but this doomed small tug operators tugs working parts of the system under contract to the state, who could not afford the cost of entry into a system with larger, more complex technology.

By 1925, traffic was only 10 percent of that projected, and the expensive state terminals saw little tonnage (in some cases, none) go through.[24] Still, the canal system was defended in the legislature as money well spent, because it acted as competition to railroads. The railroads kept

[23]Ibid., 13.

[24]Ibid.37-39.

the rates charged by the canal so low that the state hemorrhaged money. Like canals in Europe, in the 1990s the Erie Canal is now touted as a tourist attraction and rarely used for commercial shipping.[25]

As the twentieth century closes, the large rivers in the United States are still used for commercial traffic and have benefited from container standardization by the International Standards Organization. Standard sizes allow freight to be offloaded from oceangoing vessels onto inland barges, railroads, or trucks. Much of this traffic consists of commodities and is handled at specialized terminal facilities (e.g., oil, coal, sugar, and grain), which also have facilities for transfer to the rail or trucking system.[26] Like many networks that face stiff competition or new technology, the American waterways have survived where they became part of an intermodal system.

5.3 Railroads

Unlike the Erie Canal and the steamboat routes, railroads began as local transportation vehicles where roads and waterways for transportation were unavailable or too expensive. The first rails were wood over which horses pulled cars loaded with coal or iron ore bound for trans-shipment at port terminals.[27] Other rail lines were built to connect with canal systems. For many years rail lines built to haul many kinds of freight and passengers connected local communities, and no one envisioned a vast interconnected system. In the early years many rail lines were built at different gauges, or track widths, *with the purpose of* frustrating interconnection in order to give a local rail monopoly to the builder and to make shipment to competing cities difficult.

After extensive experimentation with the design of the engines, rolling stock, and rails, railroads became faster, cheaper, and more reliable than other forms of transportation, for both passengers and freight. For rail service to go beyond short hauls on local lines, government authority for trunk lines was needed, which was usually obtained as grants of right of way that allowed a rail company to force the sale of private land. Government intervention was also required to set a standard for gauge and rolling stock, so that freight and passengers need not be transferred to another carrier where a change in gauge occurred. In 1846 in England, a special government commission made the legal standard 4 feet, 8.5 inches, in spite of almost unanimous professional opinion that a broader gauge was better, simply because that width was widely used

---

[25]McNeil, 515-517.

[26]Ibid., 549-550.

[27]The first rail systems were built in 1603 to haul coal in Britain. The first built in the United States was in Pennsylvania in 1809, to carry stone from a quarry to a river; it had wood rails and used horse power. The first railroad to use steam power was built in Baltimore in 1828. See P. J. G. Ransom, “Rail,” in McNeil, 555-558.

then.[28] This standard eventually won out in the United States as well, because before large-scale locomotive manufacturing developed there, it was easy to import standard British equipment.

Interconnected lines meant that a signalling system had to be developed to control traffic both within the railroad system and where a railroad intersected a road. Scheduling allowed many users of a bottleneck—a bridge, a switching yard—to allocate the use of this resource efficiently. Although rail lines were private property, extensive security was necessary to keep out trespassers wanting free rides or to steal the goods being shipped. Because railroads were often used to transport troops and war materials, they became vital to national security in time of war.

In the mid-1800s, after the initial successes of some railroad companies in the United States (especially in New York and Pennsylvania), railroad building (like road building fifty years earlier) became a mania, and tracks were laid to many small cities with dreams of growth. A speculation bubble left many investors with worthless stock, and when, in the early 1870s (notably the Panic of 1873), it burst, local government suffered substantial losses from having subsidized such projects in the cause of economic development. Although the railroad was crucial to the success of many communities that became rail hubs or shipped large amounts of commodities, such as minerals or grain, small communities often learned the hard way that infrastructure does not guarantee development, especially when many small communities used the same strategy to compete for new business and new settlers.[29]

Public financing of railroad projects went through several distinct phases: state loans (1830–1840), municipal aid to private companies (1840–1850), and federal or state land grants (1850–1870).[30] Unlike many countries in Europe, in the United States federal and state governments typically neither built nor operated the railroads. Government investment in private railroads was justified on the same grounds as spending on roads, bridges, and canals: It was a “public good” made necessary by the public need for national defense (transport of troops and supplies) and faster, cheaper delivery of the mail. By 1845, provisions were in place to pay for several *classes* of mail delivery service, depending on the size of shipments, the importance of the traffic, and the speed of delivery. By the early 1900s, many railroads maintained their profitability by depending on government contracts to carry the mail[31] and to haul government personnel and supplies. Other forms of aid included government geographical surveys and elimination of the duty on imported iron for the manufacture of tracks.

---

[28]Ibid., 567-568.

[29]See Lee Benson, *Merchants, Farmers and Railroads: Railroad Regulation and New York Politics* (Cambridge, Mass.: Harvard University Press, 1955).

[30]*History of Transportation*, 553-554.

[31]By 1835, the U.S. Postmaster General declared that rail lines were critical to keep the mail service from sinking into “contempt.” Ibid., 592-605.

Government aid to railroads was fought by road and canal interests but sold to legislatures as an *add-on* to these transportation channels, that is, railroads were to be built where other forms of transportation were not feasible, e.g., where mineral ores were to be carried long distances and where water routes were not available. One early (1808) federal government report predicted that "The sort of produce which is carried to our markets is collected from such a diversity of routes that railroads are out of the question as to the carriage of common articles."[32]

In the early 1800s, state grants of right of way to companies were made on an individual basis and resulted in considerable corruption among the state and local officials who handed out these special privileges. Locally prominent citizens were often made part of the new railroad venture in order to secure the privileges from the legislature.[33] This resulted in such public outcry that in later years the states passed general laws governing all railroad building and regulation.[34]

Much U.S. regulation of railroads was borrowed from Britain and attempted to limit companies to a definite rate of return on invested capital. In many cases, states retained the right to buy a railroad after a certain term or after investors had recouped their original investment plus a reasonable profit. This limitation on profits left railroads in a quandary if they attracted too much new traffic. Since additional freight and passengers added limited marginal costs, the company's incentives were either to reduce traffic after a certain return was attained or to spend the additional income on plant and equipment that they may not really need.

Two visions of railroad economics competed in the early years: a *common road* and a *common carrier.* In the vision of a common road, anyone could put traffic (i.e., trains) into the system, much as the government-built "free" roads that put turnpikes out of business. In the vision of a common carrier, although there might be only one owner of trains, that owner would be forced to carry the goods of anyone who paid the tariff. Common carrier status was said to be "undemocratic," particularly by canal owners, because canals operated as common roads.

According to the theory behind state regulation, the legislature gave railroad companies the right of eminent domain, because railroads were considered a public channel. A railroad company was thus a creature of the state, set up to run as a public utility. In later years the power of eminent domain was not granted if the proposed service would duplicate an existing service, because legislators considered competition wasteful and a danger to the existing service. Most railroad routes, in the view of policymakers, did not have enough traffic to keep duplicate services alive.[35]

---

[32]Comment in *Gallatin's Report of 1808* by Benjamin H. Latrobe, afterward one of the great railroad engineers of the United States. Ibid., 585.

[33]This "investment" by local citizens was used in the 1980s by cable operators trying to get local franchises.

[34]*History of Transportation*, 555-557.

[35]When two firms compete at the same time for the same resource, one will go out of business. In biology, this is phenomenon is known as the competitive exclusion principle. See P. H. Longstaff, *Competition and Cooperation: From*

In the early years, rates were highest in southern states, where traffic was lowest, adding to tensions with the north. Beginning about 1840, all railroad companies began to charge on the basis of *class of service* both for passengers and freight. Rates dropped for freight in the 1840s, as competition with other transportation channels heated up, while those for passenger service did not drop because the railroad remained the fastest, most comfortable mode of passenger transportation. When a system reached full capacity, many companies resorted to *off-peak* pricing to encourage traffic flow for times when the system was underused. The new pricing schemes needed to be approved by the states, because the rates of many older railroad companies were controlled (in theory) by their original charters and because other new railroads became subject to state regulation. Many states insisted on "flat-rate" pricing (similar to rates approved for toll roads and canals), which subsidized distant farmers while making urban consumers pay more than the cost of their transportation (both passenger and freight). This scheme was judged necessary to insure "universal service," and it was politically popular in the nineteenth and early twentieth centuries, when most Americans were still engaged in agriculture.

Long-haul freight had lower marginal costs for railroads and fewer transfer problems. Passing these savings on to customers allowed the railroads to grow very competitive with road and water routes, and eventually the profits subsidized short hauls. Profitability forced the railroads to become a real network, because they were mutually dependent and old rivalries could be set aside to allow new cooperation for interconnection. Cooperation meant agreement on speed, direction of travel, size, connecting devices on cars, a standard for setting times, passenger schedules, baggage and freight transfer, and schedules for bottlenecks, such as switching yards and bridges. Cooperation was extended to price fixing, territory allocation, and other forms of anticompetitive behavior, which brought government deep into railroad affairs.[36]

This new scrutiny by government came just when the railroads were facing the same challenge they had themselves given to the canals and toll roads: technology was making new transportation channels possible. By the 1960s, interstate highways and the airlines began to take some of the railroads' most profitable traffic. In yet another round of a pattern now familiar, government was building (directly or through mail subsidies) new channels in hope both of moving traffic more quickly and less expensively and of using those channels as competition to existing ones in order to bring down prices perceived as too high. But political leaders were loath to allow the railroads to abandon unprofitable lines (particularly in rural areas, at the fringes of the network), because voters in those districts believed such lines were vital to their own economic survival. Not until the 1970s were the railroads finally able to abandon many short-haul routes and refuse small shipments (smaller than carload lots), further increasing the trucking industry's marketshare.

Biology to Business Regulation (Cambridge, Mass.: Harvard University Program on Information Resources Policy, P-98-4, October 1998).

[36]See *History of Transportation*, 565-570.

Railroad rates for *transportation* are defined as *local* (that is, transportation between two points on the local company's system) or *through* (charged by two or more carriers along an intersystem route—either as a *combination* of tariffs or as a *joint* tariff that both carriers have filed together). A carrier may also have rates for *access*, known as *trackage rights*, under which the carrier allows another railroad's trains to use its tracks. Access rights are viewed as a lease agreement under which the lessee gains a right of way rather than as a division of the rates charged for the train.

The "reasonable" rate regulators sought was often a combination of the factors of distance, cost of service, and the relative strength of the railroad. A "just and reasonable" division of through rates allocated more of the rate to weaker companies in order to keep them in business, on the theory that everyone benefits from all parts of a network and that without subsidies from areas of heavy traffic some lines in rural areas would not survive.[37]

Trackage rights (charges for access to the channel) are regulated to keep tabs on all the money in the system, to avoid problems in cost allocations and revenue transfers between carriers in the guise of payments for access. By allowing one carrier to allocate more of its costs to those seeking access, over time, that carrier can lower its prices to customers to the competitive detriment of competitors that need to pay trackage fees. Access fees can create hidden subsidies for other carriers if these are charged less than the reasonable cost of using the line. Regulators may overlook (and even encouraged) such hidden subsidies if a strong line is helping a weak one, such as a feeder line bringing in traffic from rural areas. In such cases, government may act as the ultimate subsidizer by buying access for defense purposes. Trackage rights paid by Amtrak (a "private" passenger service heavily subsidized by the federal government) may act as a hidden government subsidy to many railroads, because Amtrak is the biggest customer of some railroads and makes important profit contributions to many others.[38] Government may impose trackage rights as a condition for railroad mergers or acquisitions, to insure competition on particular lines and out of the belief that this condition will keep prices down.

The railroad industry did not thrive under the weight of heavy regulation and the loss of markets to new transportation channels. The Staggers Act of 1980 changed many assumptions about railroad economics and regulation.[39] Congress decided that the marketplace, and not the Interstate Commerce Commission (ICC), should regulate the industry. For sellers and buyers to establish a price, more competition was needed, so entry into the system (by way of ICC approval to start a new line) was eased. Although this policy did not result in competition on long-haul

[37]For a comparison of the histories of rate allocation of railroads and telecommunications, see Robert Godbey, *Revenue and Cost Allocations: Policy Means and Ends in the Railroad and Telecommunications Industries* (Cambridge, Mass.: Harvard University Program on Information Resources Policy, P-79-2, July 1979).

[38]See Frank J. Dooley and William E. Thoms, *Railroad Law a Decade After Deregulation* (Westport, Conn.: Quorum Books, 1994), 128-129.

[39]Pub. L. No. 96-448 (1980).

routes, it increased the number of small, independent short lines, many of which were organized to take advantage of unprofitable the lines abandoned by Class I railroads after deregulation. The new short lines serve local agricultural and mineral ore industries and act as feeders for Class I lines. The new companies can profit where the larger ones cannot, because they generally have lower costs for labor (many are nonunion) and low (or no) costs for acquiring abandoned rights of way.[40]

The Staggers Act allowed railroads to act as “contract” carriers, that is, with a big shipper they could have a separate contract that did not follow the tariff filed with the federal government. Thus, the railroad could charge rates as a *contract carrier* with regard to some customers and as a *common carrier* with regard to others. Each contract with large shippers was filed with the government but kept confidential, and the contract could be challenged only if it discriminated against certain ports or terminals or resulted in largely tying up the carriers’ capacity in contract hauls so that not enough was left for common carriage customers.

For the Class I railroads, the new competition and freedom from regulation produced a wave of mergers and acquisitions, abandonment of unprofitable routes, and decreased reliability of service.[41] By the late 1990s, mergers had grown so large that when two railroad networks merged the results were unimaginably complex, sometimes with unanticipated disastrous consequences. Customers called the merger of the Union Pacific and the Southern Pacific a “major debacle”: rail cars were lost in the system, and traffic through major switching yards was slowed to a crawl, indicating there may be a limit to the efficiency of a network as it grows larger.[42] Many customers began to use competitive services or even build their own rail lines.[43] Recognition of a limit to switching capacity (as well as to traffic capacity) led the Union Pacific in 1998 to announce that it would limit the traffic it put into the system. The federal Surface Transportation Board lifted an emergency order that required the Union Pacific to allow competitors access to its tracks in Texas.[44] Sometimes competition can, it seems, interfere with efficient operation of a network.

---

[40]See Dooley and Thoms, 17-41.

[41]See ibid., generally, 43-57.

[42]See Daniel Machalaba, “A Long Haul: America’s Railroads Struggle to Recapture Their Former Glory” *The Wall Street Journal*, Dec. 5, 1997, A-1.

[43]Daniel Machalaba, “Tired of Costs, Delays of Railroads, Firms Lay Their Own Tracks,” *The Wall Street Journal*, Feb, 6, 1998, A-1.

[44]Allen R. Myerson, “Union Pacific to Limit Traffic on Its Tracks,” *The Wall Street Journal*, Sept. 1, 1998, C-9.

5.4 Airlines

The aviation industry was born between the World Wars, just as the railroads were reaching their peak of development.[45] Even after demonstrating that flight was possible, until airplanes were made stable in different wind conditions and more or less reliable, aviation's backers could not show its commercial potential. As was true of the other networks discussed here, resources for R&D to solve the problems of flight came from the military. In World War I, planes were used for reconnaissance, but their limited maneuverability made them ineffective for combat. Government support for the development of better war planes readied them also for commercial duties and for mail delivery after the war. During World War II, government-supported research increased the speed of planes (the jet engine) and led to the development of radar navigation, which reduced the risks of night flying for commercial and mail flights.[46] World War II also led to the building hard-surface runways throughout the world, increasing the speed of take-off and landing and of wing-loading potential.[47]

As in the development of road and railroad systems, government supplied funds for building infrastructure and acted as an "early adopter." Generous contracts for mail delivery made commercial air service possible, ensured the success of early airlines, and convinced governments all over the world that air transportation was a matter of both national security and commercial competitiveness. Most governments made the aviation industry a national asset and ran one airline with a monopoly on domestic passenger and freight traffic.

The United States followed with aviation the pattern it set in the railroad and communications industries: by 1930, privately owned but heavily regulated carriers were given monopolies on certain routes as well as mail contracts amounting to $7 million. The system was put in place by the industry and regulators after the failure of competitive bidding, when heavy competition for airmail contracts resulted in bids at below cost. This "all-or-nothing" approach left the lowest bidder airline without enough resources to perform proper maintenance on equipment (a lack that reduced the reliability of airmail) or purchase equipment necessary for efficient passenger service.

As in the case of the railroads and toll roads, policymakers came to believe that only a monopoly (i.e., not spending resources on competition) would allow companies in the system to amass sufficient assets to build a network. Even that proved inadequate to protect all carriers when the Depression of the 1930s lowered demand. During the Depression large companies

---

[45]Much of the information in this section was drawn from Richard H. Vietor, *Contrived Competition, Regulation, and Deregulation in America* (Cambridge, Mass., and London: The Belknap Press of Harvard University Press, 1994), 23-90.

[46]For a fascinating discussion of night flying and the mails, see Antoine de Saint-Exupéry, *Southern Mail and Night Flight* (London: Heinemann, 1971).

[47]For a discussion of the development of hard-surface runways, and their effects, see J. A. Bagley, "Aeronautics," in McNeil, 609-647.

acquired many small ones, and the remaining companies developed vertical integration strategies known as the "Aviation Trusts," which were later dismantled by federal New Dealers. After that, government regulation of rates grew more intense and more Byzantine, because regulators believed that "excess competition" was bad for the industry (and for its customers), that it would lead to "irresponsible campaigns of mutual destruction." Destruction could not be tolerated in an industry with so much government investment and "public interest." Instead, competition was encouraged in such areas as speed, safety, and aircraft modernization.

Market entry also began to be closely regulated, and any company that attempted to serve a new market had to show how it would serve the public without hurting an incumbent carrier. Protection of carriers in the network became politically palatable during the recession of 1948–49, as carriers saw commercial traffic plummet when wartime traffic (people and equipment being returned to the United States, aid being sent to war-torn countries) almost ended. In the 1950s, however, the economy skyrocketed, and the airlines learned then what many networks had learned before—and would again: when times are good and potential customers can be served at a profit, keeping competitors off the territory is hard. In those boom years, small, "irregular" carriers convinced Congress that they should be allowed to join the fight between small business and giant monopolies (and competition, which lowered prices, could be good only for the constituents of a member of Congress). Competition for routes, policymakers believed, was an incentive for deployment of both new technology and better service.

Competition based on new equipment, however, meant added capacity and added debt, which added costs, drove up prices, and led to even closer rate regulation. Regulators decided, over the objections of the industry, that passengers ought not be burdened with the costs of excess capacity, while the industry claimed it needed the capacity to meet competition, serve peak loads in high travel seasons, and serve small communities that did not generate full plane loads.

The inability to pass on the costs of unused capacity, along with the "stagflation" of the 1970s, led to the death of some carriers saddled with heavy debt and a wave of consolidation in the industry. Consolidation resulted in new antitrust enforcement and forced the industry to expand vertically, rather than horizontally. Airline companies began to buy car-rental firms and hotels to serve their own passengers. In most instances, this strategy proved disastrous because the expected synergy did not develop, and the properties had to be sold off.

In 1978, Congress decided that the time was ripe for full competition in the aviation industry. Deregulation was fought by some of the very companies about to be freed of the regulatory harness. They feared that it might mean another round of the ruinous competition that had initially inspired regulation, which would drive down prices to the point where service would suffer and small communities go unserved.[48] Many of these fears were realized, and many

[48]John R. Meyer and Clinton V. Oster, Jr., eds., *Airline Deregulation: The Early Experience* (Boston: Auburn House, 1981), 1-2.

companies did not survive into the 1990s. From the vantage point of the late 1990s, a similar fate may be seen to await many international carriers as they discover the opportunities and dangers of deregulation.[49]

Some commentators believe the reason the airlines survived the advent of competition lies in two important characteristics of the network at the moment when (1978) competition was introduced: (1) most airlines had a lot of excess capacity and (2) most of them had price-elastic markets that were fairly easy to segment, allowing companies to increase demand by carefully lowering certain fares (mostly for holiday travel) and thereby use what had been merely empty seats.[50] The Air Transport Association calculated that if prices rose by 1 percent then traffic would drop by 1 percent.[51]

As both Congress and the industry predicted, prices took a nose dive on the lucrative transcontinental (long-haul) routes. At one point in the early 1980s, a one-way ticket from New York to Los Angles went for $88.[52] To gain efficiencies, major carriers began to drop unprofitable routes to small communities (short-haul) and reduced costs through scheduling, particularly by “hubbing,” which gave them economies of scale and scope and kept their passengers away from other airlines. In some cases, carriers at the hubs virtually monopolized those airport terminals, which permitted them to raise prices for all passengers and freight entering or exiting the system at those points.

Hubbing, too, had an unintended effect. The shorter hauls in and out of hubs meant that wide-body, high-capacity jets were no longer efficient for most routes and that some of these investments became stranded, literally and figuratively. For trips between small cities and the hubs, new “feeder” airlines out smaller planes into service. The feeders soon developed close ties to hub carriers through joint marketing and sharing gate space, an arrangement that led hub carriers to take ownership interests in the feeders in order to exert greater control and foreclose the entry of other feeders into the market.

Introducing competition held a big surprise for regulators: it triggered a massive round of mergers and acquisitions. In retrospect, this outcome was probably predictable. Increased intraindustry competition usually works to weed out the weakest competitors,[53] and many air carriers were forced out of the government’s protection at the same time that fuel costs

---

[49]See James Ott and Raymond Neidl, *Airline Odyssey: The Airline Industry’s Turbulent Flight into the Future* (New York: McGraw-Hill, 1995).

[50]*Airline Deregulation*, 237-251.

[51]See Ott and Neidl, 95.

[52]This phenomenon is an example of “spite.” Several competitors will reduce rates to below cost in order to gain marketshare and thus deprive competitors of customers. Although the practice hurts everyone, it hurts stronger companies less because they have resources to “waste.” See Longstaff, *Competition and Cooperation*.

[53]Ibid., 26-28.

dramatically soared (the oil shocks of the late 1970s) and interest rates rose to 19 percent. Regulators grew concerned when the carriers tried to protect or enlarge their turf to gain a competitive advantage. Airlines attempted to keep competitors off their turf by denying them landing slots, airport gates, or new, more efficient planes (by buying all those available). Regulators were forced to reexamine assumptions about *contestable markets* in this network.[54]

Competition changed much in this industry: New rules of the game were established for entry and exit from the industry, price and pricing mechanisms, market segmentation, the role of distribution channels, cost structure, operations, demand, service (and safety) and profitability.[55] By the late 1990s, some governments even considered privatizing investments in airport terminals,[56] but few policymakers wanted government to surrender control of this network because they regarded airline service as vital to economic development.

[54]See Dipendra Sinha, "The Theory of Contestable Markets and U.S. Airline Deregulation: A Survey," *Logistics and Transportation Review* **22**, 4 (1984), 405-419.

[55]Vietor, 83.

[56]See Ellis J. Juan, "Privatizing Airports Options and Case Studies," *Private Sector* (March 1996), 17-20.

Chapter Six

Communications

6.1 Postal System

Any contemplation of the future of computer-based communications networks, particularly the Internet, should involve reading the history of postal systems.[1] The parallels are striking. For example, in the early days of the United States postal service, a Congressman declared, "Why, this is a machinery which, in a sense, extends your presence over the whole country."[2] Others similarly exclaimed, "Time and distance are annihilated!"[3] Heralded as the road to true democracy and a boon to merchants who could in this way reach every possible buyer, the postal service not only accomplished both those chores, but it also brought unexpected consequences that reappeared in all the communication systems that followed, especially in telecommunications networks.[4]

The postal system increased opportunities for people to speak on the political issues of the day, but the opportunities also increased politically divisive speech, bringing the First Amendment into question, exacerbating ethnic and religious hatreds, and helping to ignite the American Civil War.[5] The same postal system that brought mail-order catalogs to people in the most remote areas of the country also contributed to the deaths of the businesses of many local merchants that had previously served people in those areas.[6] It brought a reliable channel for newspaper editors but also brought junk mail. In 1829, when Sabbatarians attempted to stop the postal system from operating on the Sabbath, it focussed attention on the relationship between government and religion.[7]

In 1775, the U.S. postal system was taken over from British systems that connected ports along the Atlantic coast and had been used primarily to maintain contact with merchants in

---

[1]This section draws largely on the following sources: Richard R. John, *Spreading the News: The American Postal System from Franklin to Morse* (Cambridge, Mass., and London: Harvard University Press, 1995); Wayne E. Fuller, *American Mail: Enlarger of the Common Life* (Chicago: University of Chicago Press, 1972); and Wayne E. Fuller, *RFD: The Changing Face of Rural America* (Bloomington, Ind.: Indiana University Press, 1964).

[2]John, 10.

[3]Ibid.

[4]John C. Panzar, "The Economics of Mail Delivery," in *Governing the Postal Service*, edited by Gregory J. Sidak (Washington, D.C.: AEI Press, 1994).

[5]In 1853, a mob in Charleston, South Carolina, seized mail sent into the south by abolitionists on the theory that it was an illegal incitement. After that, the New York postmaster refused any similar mail addressed to southern states, forcing the issue of whether states have the right to control information that crosses their borders. See John, 257–280.

[6]Daniel Boorstin, *The Americans: The Democratic Experience* (New York: Vantage Books, 1974), Chapter 14.

[7]John, 169-205.

Britain. Although the entry points into the system and routing of the mail (i.e., in post offices) remained under federal control (but not necessarily government ownership), mail was moved by private contractors who carried it on roads that often had been built or subsidized by the federal government (post roads). Security was important to the postal network, because the network was used also for military communication.

As the flow of settlement and trade moved west, the postal system followed, subsidized by Congress in the belief that the system was critical to economic development, necessary to the education of the masses, and offered a way to strengthen the role of the central government in the lives of the citizens. Often, post offices were established before demand was sufficient for them to be self-supporting; thus, postal patrons in eastern cities subsidized service in the new western settlements. In the earliest days, postal routes were established also by private entrepreneurs and local governments, but these routes ultimately were taken over by the federal government.

The Post Office Act of 1792 established a system by which nearly every citizen was able to receive mail. Ubiquity of access (and very favorable rates for periodicals) provided the basis for the success of the newspaper industry. This was not the last time an industry that wanted to reach a mass audience championed the cause of “universal service” for a government-supported communications channel. By the early 1800s, as much as 90 percent of the traffic consisted of newspapers, although only 3 percent of the population in western states subscribed to a paper (in contrast to 10 percent on the Atlantic coast). A system capable of true point-to-point service was used primarily for point-to-multipoint service. As the twenty-first century nears, the postal service remains the channel for delivery of many newspapers, and rate increases for this service are still hotly contested by the newspaper industry.

The 1792 Act also required U.S. postal authorities to accept mail regardless of content and not to use it for surveillance of citizens. The postal service was thus treated as a *common carrier* that was not responsible for content and must deliver any message paid for (with some exceptions—obscene materials, dangerous packages). Citizens gained confidence in the system, because, unlike the situation in Europe, in the United States the federal government was prohibited from spying on messages.

By the same act Congress mandated that the correspondence of each member of Congress to constituents was to be carried free of charge, a “franking” privilege still in place today. (In the twentieth century the idea of mandating low prices for a congressional candidate’s messages to the electorate found its way into the regulation of broadcast channels.[8]) This right is not inconsiderable; rates for letters are much higher than for newspapers. Prior to the Post Office Act of 1845, the rate for mailing a letter was based on the distance the letter was to go and its weight. A letter from Boston to family on the Northwest frontier could cost the equivalent of a day’s wages for a common laborer. But the sender did not usually pay the postage. It was paid by the

[8] 47 USC Section 310 (a) (7).

receiver, on the theory that a service was not paid for until after it had been performed. (Prepayment became mandatory in 1855.) Because of the high cost, most mail was sent between merchants, whose high rates paid more than the cost of their service and thereby subsidized both mail from Congress and the delivery of newspapers.

The Postal Acts of 1845 and 1851 changed this landscape by reducing the letter rate and eliminating most distance differentials. These changes were supported by most members of Congress from the west (whose constituents wanted to write to family “back east”) and by east-coast merchants, who wanted to reach into the frontier. Some members of Congress from the south held up passage of the law, because they feared reduced rates would mean that small offices would close. Predictably, the new rates instigated a storm of letter writing as well as a barrage of unsolicited mail from local merchants and catalogs from outside the local markets. Cheaper postage stirred competition for local newspapers when papers from larger cities arrived by mail at the new lower rates. It did not take long for country folks to demand that the postal rates be changed to protect them. In 1872, Congress took final steps to eliminate competition for the postal service by giving it a monopoly on delivery of “letters and packets,” although the exact definition of these terms was to be litigated for many years as private enterprises attempted to carve out a niche in the delivery of printed material.[9]

Interest in postal communication opened another business opportunity, one seen again, with the Internet boom of the 1990s: handbooks on “how to” use the new channel. Postal authorities needed to develop ways to deal with the resultant increased demand by increasing efficiency and managing capacity. Capacity planning led to rates based on class of service (as determined by speed of delivery) and to a distribution system based on geographical hubs, where mail was sent to be sorted for delivery outside the local area. In 1800, post offices were divided into two kinds, distribution centers and branch depots.

Until the middle of the nineteenth century, local distribution was accomplished by having patrons come to the post office to pick up their mail, by requiring them to pay a government-supervised “penny post” service that delivered the mail, or by having them subscribe to a private service that picked up and delivered mail to both business and residences. When the local post office served as a terminal for receivers, it became a bottleneck when traffic grew heavy. At peak times, citizens might wait in line for hours, particularly if few windows were open. In New York City in 1850, in the entire city only fifteen windows were available.

---

[9]For an extended discussion of this litigation, see James I. Campbell, “The Postal Monopoly Law: A Historical Perspective,” in *The Last Monopoly: Privatizing the Postal Service for the Information Age*, edited by Edward Hudgins (Washington, D.C.: The Cato Institute, 1996), 9-22.

As the system grew and became more reliable, it became the channel for much of the country's exchange of money and merchandise.[10] As a result, new forms of security were needed to protect letters and packages at peak times. Mail robbery had been a federal crime (punishable by death) since the 1790s.

Also by the early nineteenth century, when postmasters were required to surrender control over any newspapers with which they were associated, control over this channel was separated from control over the messages in it. This separation was thought necessary to avoid favorable treatment for some publications. Questions about the separation of channel owners from message senders arose again, in the middle of the twentieth century, with advent of broadcasting, cable television, and the possibility of delivering entertainment services via telephone lines.

As important as the mails were to economic development of the United States and to building democratic ideals, the U.S. postal system, like the communications systems that followed it, was seen as a purveyor of "bad" messages. While treated as a common carrier and thus not responsible for the content of mail it was required by the 1792 Act to deliver, it was also charged with keeping certain messages—obscene, indecent, scandalous, or seditious messages—out of the channel. Many such laws were struck down by the courts as inconsistent with the First Amendment.[11]

As the twenty-first century dawns, postal systems worldwide are being changed by new technologies and new political support for privatization and competition in the old monopoly networks. Customers can use facsimile (fax) machines, e-mail, and independent package-delivery services to go around the monopoly bottleneck (thus avoiding first-class postal rates and slow, less dependable service). Many such messages once would have been sent as first-class mail, which was the revenue driver for the entire postal network. In the early 1990s, approximately 70 percent of first class mail consisted of business-transaction mail, including bills, payments, and fund transfers,[12] and these are the messages telecommunications companies hope to carry on the Internet.

In the late 1990s, many large users of postal services (such as direct-mail marketers) demanded that the prices for that service be unbundled, that is, that customers should pay only for services actually used. Such large users reasoned that a mailer that either presorted mail or delivered it directly to a hub sorting facility deserved to have its rates lowered by the amount it had saved the Post Office. Allocating such cost savings in a complex network is not so easy as it

---

[10]A contemporary commentator estimated that as much as $100 million moved through the mail—almost double the federal budget in 1855. See John, 54-55.

[11]See William Ty Mayton, "The Missions and Methods of the Postal Power," in *Governing the Postal Service*, 81-83.

[12]W. Clay Hamner, "The United States Postal Service: Will It Be Ready for the Year 2000?" in *The Future of the Postal Service*, edited by Joel L. Fleishman (New York: Praeger Publishers, 1983), 120-121.

sounds, however, and may require regulatory systems that are more adept than any yet discovered as well as new political decisions on who will get the large pieces of the newly sliced pie.

In the late 1990s, both new technologies and demands for unbundling threaten the complex subsidy schemes whereby some mail services subsidize others (to make universal service possible) by taking high-margin services out of the revenue stream. Unbundling poses a threat that includes the possibility that the postal system might be left with a high-level scheduling function (routing and switching at national and international levels) and with high-cost, low-revenue customers, especially residential customers, in short-haul local networks. Another difficulty is the agreements with politically powerful employee unions, which make it difficult for postal authorities to introduce technologies that could increase efficiency (and render their services competitive).[13]

6.2 Telephone System

The telephone system built on a number of nineteenth-century technological innovations, and its development would not have been possible without advances in areas such as materials science, electromagnetism, and acoustics. These advances, many of which initially were made for military purposes,[14] were put together differently in Europe and the United States. The first person (by about two hours) to be granted a patent on a commercially viable system was Alexander Graham Bell in 1876 in the United States. Like many other new ideas, the telephone system's full potential was not understood at that time, and many dismissed it as just another form of telegraph. In Europe the telephone system was operated by postal authorities until the 1990s.

Unlike the telegraph, which was used mainly for time-sensitive, long-distance communication, telephone systems were first built to connect local senders and receivers, particularly businesses in need of fast communication with customers and suppliers. As the price of service fell, local residential customers were added to the system. In the early days, customers were connected with one another through local systems franchised by the Bell Company; but after the Bell patents expired in the 1890s, new entrepreneurs entered many markets to provide equipment and service, especially the high-traffic markets of commercial centers outside the largest cities. Competing telephone services sprang up in more than half of all U.S. cities with populations over 5,000. To make a call, a sender needed to be on the same system as a receiver. In

---

[13]See ibid. and Joel L. Fleishman, "A Candid Assessment of a Decade of Postal Reorganization," in *The Future of the Postal Service*; see also Michael A. Crew and Paul R. Kleindofer, "Pricing, Entry, Service Quality, and Innovation Under a Commercialized Postal Service," in *Governing the Postal Service*.

[14]This section draws largely on the following sources: Irwin Lebow, *Information Highways and Byways: From the Telegraph to the 21st Century* (New York: Institute of Electrical and Electronics Engineers Press, 1995); and *An Encyclopaedia of the History of Technology*, edited by Ian McNeil (London and New York: Routledge, 1996), Chapter 15, "Information," 686–758; Peter Temin, *The Fall of the Bell System* (Cambridge, Eng., and New York: Cambridge University Press, 1987); Richard H. Vietor, *Contrived Competition: Regulation and Deregulation in America* (Cambridge, Mass.: The Belknap Press of Harvard University Press, 1994).

cities with several competing systems, a business had to subscribe to all of them in order to reach all local phones.

In rural and other low-density, low-traffic areas, many of new systems were built by local municipalities or locally organized cooperatives, because the Bell system was not interested in going there. Like the other networks discussed in this report, the telephone system was seen as crucial to the economic development of every city and village in the country. Telephones were thought necessary to public safety and for rapid commercial transactions. Rural politicians made sure that these systems benefited from government subsidies in the form of low-cost loans and tax breaks (and, later, from transfers from more profitable systems through a government-imposed "universal service fund").

Because a long-distance call required a sender's local system to be connected with a receiver's system, both had to be connected with the Bell System's long lines, because Bell controlled important patents for long-distance service. Bell refused to connect with new competitors or the many rural systems that were not its direct competitors.[15] A customer desiring long-distance service had necessarily to subscribe to a Bell system.

But this strategy of denying competitors access came just as the federal government began (at the beginning of the twentieth century) to question the power of large corporations on the people's pocketbooks, fearing that, in the absence of competition, the prices these corporations would charge might grow exorbitantly. Bell's denial of connection and its efforts to acquire smaller companies threatened to embroil it in the antitrust activities then breaking up similar giant companies in oil, steel, and railroads. To avoid antitrust problems, Bell became more cooperative with independents and began to buy up competitors to merge them into the Bell system. By 1913, twenty-six states had made interconnection compulsory. Even so, Bell did not avoid the notice of the antitrust enforcers, and in that same year it agreed to the "Kingsbury Commitment," by which it agreed to give up ownership of Western Union, stop buying competing phone systems, and allow interconnection with its interexchange and long-distance services. But government insistence on competition soon gave way to a need for cooperation.

During World War I the Postmaster General was given control of wartime communications (wired and wireless) to facilitate and coordinate military communications. Predictably, cooperation—not the competition urged by antitrust regulators—became the order of the day, and the telephone system was treated like the postal system, that is, as an integrated system in which traffic can flow from any one place to any other place. The Postmaster General made an effort to make prices conform to rate systems set up for postal services and to bring them closer to a

---

[15]Many small companies also opposed interconnection, because they feared it would preclude their banding together to challenge the Bell system. For a discussion of the history of interconnection, see Milton Mueller, *Universal Service: Competition, Interconnection, and Monopoly in the Making of the American Telephone System* (Cambridge, Mass.: MIT Press, 1997).

national uniformity on a per-mile basis. The wartime experience inaugurated an era of cooperation, in which, in the minds of many regulators, the need for efficiency began to outweigh the benefits of competition.

Deciding that competition was not in the best interests of telephone subscribers, Congress in the Willis–Graham Act of 1921 permitted the consolidation of competing phone companies (many of them failing). Theodore Vail, who had been president of AT&T (the Bell Company's long-distance subsidiary), is often credited with convincing government to eliminate competition in order to insure universal connectivity (not to be confused with universal service) and efficiency in a system based on regulated monopolies. This regulatory scheme allowed the Bell system to buy more local systems and increase its marketshare. It also led to compatible and highly cooperative systems for signalling, scheduling, and terminal equipment, all of which increased the system's efficiency and, for customers, its ease of use.

By mid-century, telephone monopolies were firmly in place, and regulators (state commissions and the Federal Communications Commission [FCC]) had instituted extensive regulation of service (including capacity planning) as well as rate-making principles to accomplish public (and political) goals. Low-cost services (large-traffic, high-density, and long-haul) charged prices much higher prices than the cost of their service in order to subsidize high-cost services (small-traffic, low-density, and short-haul).

- *Value of Service Pricing.* The value of service pricing assumed that commercial customers of telephone service got more "value" from the service, because they used it to make money. Given that such customers were often located in urban (high-density) areas and used the service more often (large-traffic), the cost of providing local service was lower but the prices charged were higher. The policy operated as an indirect corporate tax to subsidize residential service.

- *Distance Uniformity Pricing.* It was difficult for the telephone industry to explain to the political leadership and to customers (who were also voters) why some telephone companies had higher cost structures than others and thus charged more per mile for a call than others—to customers, it looked like one big system, just as the post office did. If some customers had to pay more than others in the next county or state, the situation seemed "unfair." Nationwide averaging (applied to interstate and interexchange services) increased subsidies to high-cost or inefficient services from low-cost, efficient ones.

- *Low-Cost Local Service.* In the early days (1921–40) of telephone rate-making the population of the United States was widely dispersed, on farms and in small towns (low-density). People lived near their families, shopped locally or by mail, and had little need of long-distance telephone service. To encourage homeowners, farmers, and small-town businesses to get on the telephone network, local rates had to be kept low—even though local (short-haul) service was the most expensive to build and maintain. Long-distance rates were calculated on a "board-to-board" basis, that is, they were based on the assumption that long-distance included only costs associated with moving calls between two local exchanges or "boards."

In the years 1930 to 1947, a series of Supreme Court cases and FCC rulings[16] yielded rate-making on a "station-to-station" basis, that is, all the costs of moving traffic from the sender to the receiver (including local costs for the call) were included in the long-distance rate. The rulings mandated that more long-distance revenues were to be paid to local exchanges for access to their systems. But the new rate-making scheme required the costs of long-distance and local service to be allocated, or separated. To accomplish this gigantic accounting task, a formal system of separations was developed by 1947. Separations never allocated the costs to everyone's satisfaction but did so well enough that, in combination with other give-and-take mechanisms, local and long-distance providers were reasonably happy—until cracks appeared.

They appeared in the last quarter of the twentieth century, when both technological changes and deregulatory fervor swept the United States (and the world), forcing *the system* to open. Such forces upset the political and economic equilibrium, leaving policymakers scrambling to find new ways to achieve the public goals, such as universal service, set for telephone service.[17]

Military research and development during World War II brought technical changes. Bell Labs focussed most of its resources on wartime projects that led to breakthroughs later spun off to the civilian sector and which changed the face of the world economy, in particular, the use of computers and microelectronics. Technical innovations also changed the economics of telephony by increasing the efficiency of scheduling (routing, switching, and capacity control), signalling, and terminal equipment, which enabled greater distribution of those functions and decreased dependence on a hierarchical system. As a result, the system grew permeable, allowing entrepreneurs from outside it to connect and offer services not burdened by rate subsidies, so large customers could purchase the terminal and scheduling equipment that allowed them to build their own local networks and perform some of the activities formerly provided by the telephone company at the mandated high rates. This new capacity to perform some functions of the network led to calls for unbundled rates, allowing consumers to pick and choose the parts of the network they wanted to use (and pay for). At the same time, pressure to change the system was coming from companies that discovered that telecommunications services were a growing share of their cost of doing business.

Computers (in which digital coding replaces analog) and microelectronics made private networks and competing long-distance service economical. These technical advances also made possible new services such as data transfer and the Internet. Entrepreneurs challenged telephone companies' procedures and charges for access and put pressure on government to open the system in the late 1990s, just as the political winds were changing and competition was, once again, seen as a way to make communications services better.

---

[16]Starting with *Smith v. Illinois Bell* 282 US 133 (1930). See Carol L. Weinhaus and Anthony G. Oettinger, *Behind the Telephone Debates* (Norwood, N.J.: Ablex Pub. Corp., 1988).

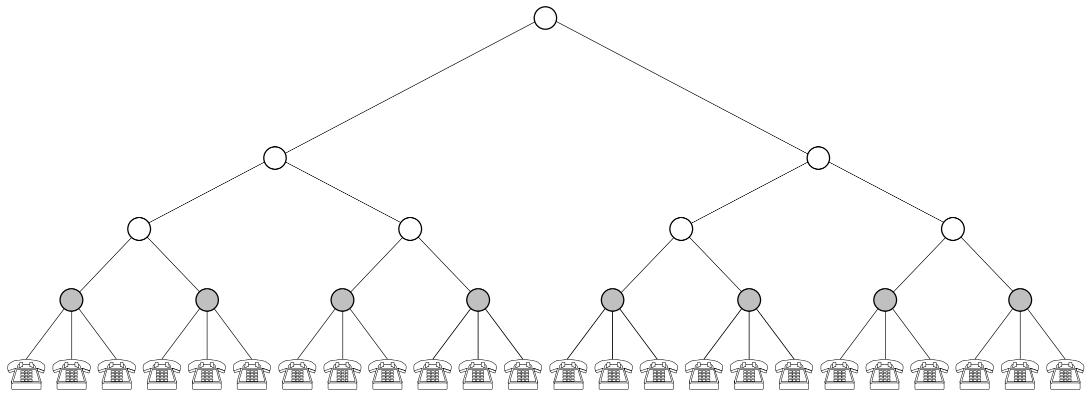
[17]See P. H. Longstaff, *Telecommunications Competition and Universal Service: The Essential Tradeoffs* (Cambridge, Mass. Harvard University Program on Information Policy Research, P-96-2, May 1996).

Communications policymakers believed that competition would (as in other sectors) lower prices and increase services for consumers. But this view ignored several factors: that some consumers (low-density, low-traffic, short-haul) already received service below cost; that new competitors were not interested in those customers at their current rates; and that even the most competitive system could not deliver services below cost for long. Price competition was possible only for consumers paying above the cost of their service (commercial, high-traffic, long-haul, high-density). There is little evidence to suggest that that policymakers considered the possibility that competition might create problems as well as opportunities.

Competition in the telephone network replayed many scenes of airline deregulation in the 1970s and ´80s. The Network evolved into several competing networks that tried to serve all their customers' needs, including local, long-distance, and Internet access. Smaller (usually rural) companies seemed poised to become aggregators or feeders into the competing networks. Although prices for long-distance (long-haul) service fell in the highly competitive price wars of the 1990s, the price of local (short-haul) service rose as the subsidies that had kept it low (relative to costs) were drained out of the system. Prices for business services remained somewhat higher (thus more profitable), because demand was less elastic.

Competition caused many companies to seek the competitive advantages made possible by increased economies of scope and scale through unprecedented rounds of mergers and acquisitions in the late 1990s, when the already large telephone companies grew even larger. Small companies also merged, picking up rural telephone exchanges that large companies had found uneconomical, although in some states accounting rules made small telephone companies more profitable as independents than as part of large companies or systems. Lack of cooperation among the new competitors led to a confusing variety of service for consumers, and the dependability of service declined, along with investment in redundant systems. Dependability suffered also from difficulties in capacity planning, formerly a function of rate setting, as those putting traffic into the system no longer were part of the rate-making procedures. Politicians were worried that rural areas might not get the new services of the much-touted information age when investment rushed to high-traffic metropolitan areas.

In the 1980s and ´90s, the technical innovations and regulatory changes in the United States broke down barriers to entry into telecommunications (e.g., by wireless, satellite, electric, and cable companies), blurring the boundaries of the telephone industry. The boundaries between many communications services became permeable, bringing telephone companies seeking new sources of revenue into markets for entertainment and information services. Interindustry competition, which had been expected to lower prices for consumers and make new services available, instead brought political and industrial turmoil, reducing the likelihood that The Network and the public policy goals it addressed could be put back together.



Source: Carol L. Weinhaus and Anthony G. Oettinger, *Behind the Telephone Debates—1, At the Heart of the Debates: Costs, Control, and Ownership of the Existing Network* (Cambridge, Mass.: Harvard University Program on Information Resources Policy, P-85-6, April 1985), Figure 17, 48.

Figure 6-1

The Telecommunications Network

The telephone or telecommunications system is a point-to-point network over which any sender or receiver can reach any other sender or receiver with access to the network. The network is generally hierarchical: calls go from customer premises equipment (terminals) to a local hub, which switches calls made within the local calling area. If a call goes outside the local area, it may be routed through several layers of higher level switching centers before getting to the receiver's local hub. This entire system is sometimes referred to as the public switched network (PSN),[18] although in a era of competition that designation is clearly misleading. The infrastructure of the PSN is privately owned, and, given that new providers often are in direct competition with an incumbent provider, which owns the infrastructure and controls signalling and scheduling, the price charged for access has become a competitive issue. Many new competitors act as aggregators of demand by residential and business customers in order to secure reduced rates for high-traffic customers from the incumbent local service provider. Here, as in the other networks discussed in this report, long-haul and high-traffic customers are the least expensive to serve, intensifying competition for long-distance, large business, and metropolitan customers.

The channel through which telecommunications traffic travels can be a wire or, if the traffic is sent over the air, microwave or other high-frequency signals. Some new satellite systems avoid the "public" network by directly linking senders and receivers through their own infrastructure and by using their own systems for signalling and scheduling. Thus, The Network may evolve into several networks, which may or may not be connected. Competitive pressure can cause the various channels to resist interconnection and attempt to develop exclusive domains in which their customers can be kept from their direct competitors and from substitute products and

---

[18]Sometimes also called the public switched telephone network (PSTN).

services.[19] Quality of service for all customers appears to be one of the casualties of competitive telecommunications services as telephone businesses seek to squeeze as many costs out of their systems as possible in order to meet competitive pressure on prices—customer service and backup systems often are among the first to be reduced.[20]

6.3 The Internet and Computer Networks

It is difficult to compare computer networks to all the other networks discussed in this report, because the history of computer networks is so short and so recent and because they are undergoing rapid evolution. Yet parallels exist.

Individual computers linked via the telecommunications network are not, strictly speaking, in a network but are simply different types of terminal equipment (which use digital coding, instead of analog) to put traffic into the telecommunications channel. True computer networks connect computers without using the telecommunications channel; they have their own channels (wire or wireless) and their own signalling and scheduling systems. Computer networks can be used for point-to-point communication (e.g., e-mail) or point-to-multipoint (e.g., World Wide Web pages or for messages "broadcast" to many receivers at once).

Digital coding has enabled many innovations in signalling and scheduling, because digital messages do not require setting up a circuit between sender and receiver (circuit switching). Traffic can, for example, be sent in "packets" that do not all go to the receiver by the same route but do arrive at roughly the same time. Like pieces of postal, or "snail," mail, each digital packet is addressed individually and can be switched into whichever channel is most efficient at the instant the packet enters the system. "Packet switching" is used for Internet traffic in the telecommunications network—messages do not need to arrive at exactly the same time in order to set up a Web page or deliver an e-mail message—and for that reason it is often called the Internet Protocol (IP). When local area networks (LANs) are connected with one another they are said to become a network of networks, in which the signalling and scheduling systems must be compatible and a higher level of scheduling must be capable of settling conflicts between systems, for example, when traffic from more than one system enters the same channel at the same time. Lack of centralized traffic control for packet networks may change the very nature of the telecommunications network and make it operate more like highway networks, in which each packet (or car) makes its own way to its destination.

---

[19]For more on related problems in telecommunications policy, see Russell W. Neuman, Lee McKnight, and Richard Jay Solomon, *The Gordian Knot: Political Gridlock on the Information Highway* (Cambridge, Mass.: MIT Press, 1997).

[20]Rebecca Blumenstein and Stephanie Mehta, "Lost in the Shuffle: As the Telecoms Merge and Cut Costs, Service Is Often a Casualty," *The Wall Street Journal*, Jan. 19, 2000, A-1.

Protocols for actions and standards for hardware and software are critical to the operation of computer networks.[21] Protocols govern the rules for signalling and switching in the many layers of computer networks, while standards are set for terminal equipment, ancillaries (such as power supplies), and the software to be shared across or within networks. Standards and protocols are developed through industry trade associations, government-sponsored initiatives (often in the name of international competitiveness), and in the marketplace, where consumers decide, for example, which hardware or software will "win" by becoming the industry standard.

A very large network that is not a collection of smaller networks is called a wide area network (WAN). WANs may look like early telephone networks: they connect members of a community (for example, many units of a multinational company), but they do not directly connect that community with the rest of the world. The possibility exists that competition in telecommunications may cause large-traffic customers to build their own WANs, so that the integrated telecommunications network could evolve into a series of private or semiprivate networks, as in the airline industry.

Capacity planning can be complex for computer networks, because peak loads and high-volume users slow the system for everyone else, as in an electrical brown-out. If no charge or penalty can be levied for use at peak times or for bursts of heavy-volume traffic, users have no incentive to participate in capacity management. Network managers are forced to build capacity for the highest amount of traffic that can be accommodated at acceptable rates of operation. This situation resembles the one faced by traffic engineers, who need to plan for rush-hour traffic so it can move at an acceptable (if not the maximum) speed. In the late 1990s, telecommunications channel capacity was still a scarce resource, but government regulation for allocating it was not favored by "deregulators" in Congress. This led to calls for resource allocation based on market principles: charging not only for access but also for the number of bits sent; charging more for messages sent during periods of heavy traffic[22]; or enabling a "smart market" to allocate access through an electronic bidding process each time a sender seeks access to the network.[23]

---

[21]See, for example, Andrew S. Tanenbaum, *Computer Networks*, 3rd ed. (Upper Saddle River, N.J.: Prentice-Hall, 1996); and Martin C. Libicki, *Information Technology Standards: Quest for the Common Byte* (Boston: Butterworth–Heinemann, Digital Press, 1995);

[22]In 1998, Sprint announced it will begin to market its Integrated On-demand Network™ (ION), which requires a "hub" in each residence or business to coordinate traffic for all phones, faxes, computers, and so on, and to convert the traffic into bits, format it into packets, and put it into the telecommunications network. Sprint promises customers virtually unlimited bandwidth but will charge only for bits sent. This service creates many policy dilemmas, because it redefines the nature of the telecommunications service. See Ira H. Goldman, "Technology Will Kill Telecom Taxes," *The Wall Street Journal*, Aug. 10, 1998, A14.

[23]See, for example, Jeffrey K. MacKie-Mason and Hal R. Varian, "Some Economics of the Internet," *Networks, Infrastructure, and the New Task of Regulation*, edited by Werner Sichel and Donald Alexander (Ann Arbor: University of Michigan Press, 1996), 107–136.

Perhaps no other networked industry owes so much as computer networks to government subsidies and Department of Defense (DOD) research. The basic principles of computers were developed as part of research on radar and computing firing coordinates for large guns. The Internet was developed as a communications system for defense and university-based research when, in the 1960s, a decentralized system was judged more secure than a centralized one. Like other improvements to communications initially developed for the military, computers made information move from sender to receiver faster, an important advantage in war.

Early computer networks (outside the military) were built for business communications and other computing functions. Business networks generally were built without direct government subsidies, even though government was often an important first purchaser. Initially, they were not intended to connect with other business networks, but the connections were set up after the benefits of communications services such as e-mail became known outside the military and academe. Eventually, the price of network access (i.e., of computer terminal equipment and local telephone Internet access providers [IAPs]) fell enough to attract residential customers. As with all the other industries discussed in this report, a new network became the "must have" for economic development, and government looked for ways to help build it—by subsidizing access for "have nots," thus becoming an important first purchaser (in order to deliver other public goods, such as education and medical care) and by convincing investors of the potential demand. Government also subsidized the huge infrastructure (e.g., computers to schedule traffic) that uncoordinated capacity planning seemed to mandate. As the Internet has grown, some central coordination clearly has become necessary to maintain reliability for increasing traffic and unexpected peak loads, but whether it will be a government enterprise or a cooperative effort by often fractious (and highly competitive) players at all levels of the system remains unclear.[24]

The other focus of public policy, predictably, was access to the various network channels for competitors. Because telecommunications (or cable or satellite) networks are the channels for most computer communications, the prices charged for access are critical to the economics of the computer industry. Policies to encourage competition created controversies about the price of access to the lines or airwaves that incumbent channel owners charged new competitors. Competition (at least in the short term) could lower access prices for high-traffic customers, such as IAPs, but telecommunications companies have demanded the right to charge IAPs enough to allow the channel owner to keep up with the capacity demands the IAPs make on the system. This controversy is especially tricky because IAPs *are* competitors of telephone companies. They offer signalling and scheduling services that allow customers to send long-distance messages by e-mail or by "Internet telephony" without additional payments to the telecommunications company

---

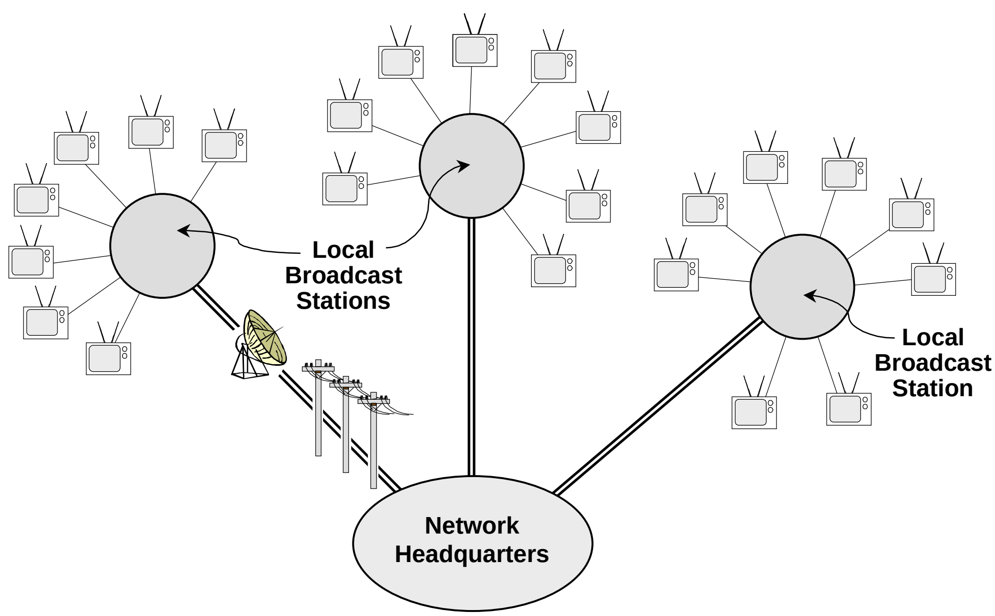
[24]See Rebecca Quick, "Is the Internet Outgrowing Its Volunteer Traffic Cops?" *The Wall Street Journal*, Sept. 12, 1997, A6.

providing both transport and channel.[25] “Hubbing” began to develop in late 1999 as various senders organized around portals where customers get access to the system.[26]

As in all networks, some external, human cooperation is critical to the security of traffic and access. Although digital traffic can be encrypted to avoid eavesdropping (so that no one can understand the message without knowing the encrypted algorithm—theft of information by breaking into local networks remains sufficiently commonplace to be a serious problem that can consume much of a system manager’s and law enforcement’s time and energy. The problem demands a cooperation among everyone at every layer of the network that may prove difficult to accomplish in a highly competitive industry, where one system’s security may be a marketing advantage over competitors.

6.4 Broadcast, Cable, and Satellite

Like the other networks discussed in this report, broadcast, cable, and satellite services (see **Figures 6-2**—**6-4**) owe much of their technological advances to military needs and funding. The

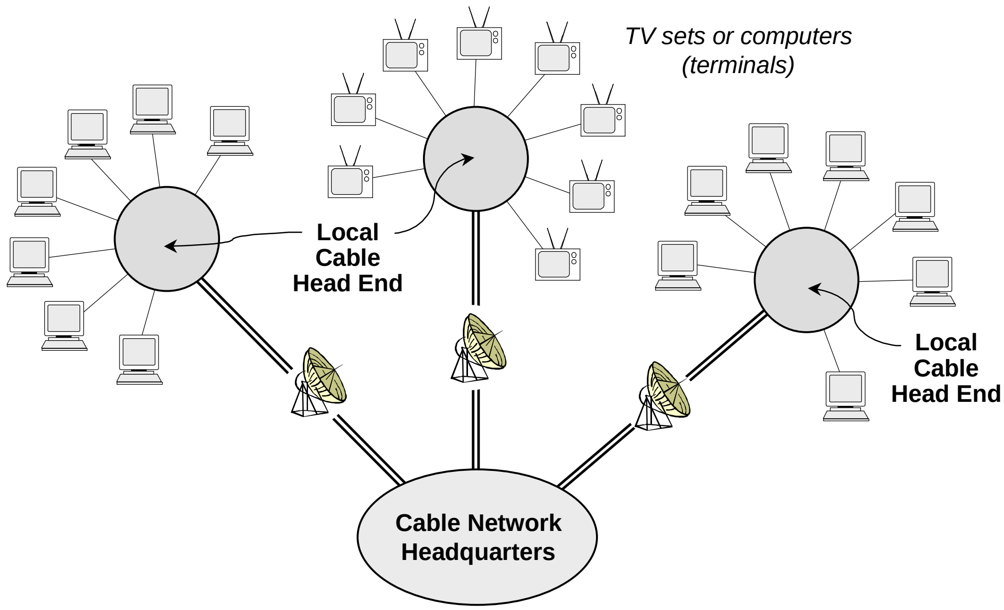


Note: Connections between network headquarters and local broadcast stations are by satellite signals or telecom lines.

Figure 6-2
Television Networks

[25]For access issues, see, e.g., “Cable Internet Unbundling: Local Leadership in the Deployment of High-Speed Access,” *Federal Communications Law Journal* **52**, 1 (1999), 211-238.

[26]See, for example, “Entertainment.com Debuts: Warner Bros. Online Collects Animation, Video Online; First of Several ‘Hub’ Sites,” *Broadcasting and Cable* (Nov. 29, 1999), 43.



Note: Connections between cable network headquarters and local cable head ends are by satellite signals.

Figure 6-3
Cable Networks

first use of the term "broadcast" came during World War I, when the U.S. Navy took over wireless radiotelephony to send messages to many ships at once. The DOD, through specific funding and through coordination of critical patents during both world wars, has been responsible for much of the R&D that made radio and television possible. Satellite services would not have been possible without defense research into (among other areas) jet propulsion, materials science, and astrophysics. Most broadcast, cable, and direct broadcast satellite (DBS) services are organized as point-to-multipoint networks that operate in a hierarchical fashion but typically do not send traffic two ways.[27] In the future, as customers interact with program providers (e.g., to order merchandise or predict the next football play), all of these networks may become "interactive," which may mean becoming multipoint-to-point or point-to-point services by which any sender can send traffic to any receiver on the system (e.g., interactive audio or video over cable). But is a cable provider that offers point-to-point audio still a cable network, or is it a telephone network? Big regulatory constraints and opportunities hinge on the answer to this and other puzzling questions presented by the hybridization of technologies, but at the turn of the century no politically acceptable answers are on the horizon.

---

[27]This section draws largely on the following sources: Sydney Head et. al., *Broadcasting in America: A Survey of Electronic Media*, 7th ed. (Boston: Houghton Mifflin, 1994); and Joseph R. Dominick, Barry L. Sherman, Gary Copeland, and Gary Copeland, editors, *Broadcasting, Cable, and Beyond: An Introduction to Modern Electronic Media* (New York: McGraw-Hill, 1993).

For many years, the channel for traffic between local broadcasters and television networks was AT&T's long lines; indeed, TV network broadcasting overlapped with the telephone network. Since the 1970s, telephone service for network traffic has been replaced in many cases by satellite distribution for both broadcast and cable. Although local stations or cable operators can send video of breaking local news of national interest to network headquarters, most traffic is from the top down: the central hub of the TV or cable network provides programming, a scarce resource that local stations and cable operators cannot produce on their own because of their small number of viewers or listeners (level of traffic). Thus, there are only small economies of scope and scale. TV and cable networks act as hubs that allow aggregation of demand for programming and advertising.

Even though most traffic comes from the top down, broadcast and cable systems generally were built from the bottom up. Initially set up to serve local communities, they were only later linked (long-haul) to "networks" that supplied the traffic they delivered (short-haul) to the local system. Broadcast uses the electromagnetic spectrum "owned" by the public and allocated by the FCC to local radio and TV stations. Because cable operators must string their cables over or under city streets, they must obtain a franchise from local municipalities. Regulation of broadcast and cable is based on their use of public facilities to build their channels.

Traffic in broadcast and cable channels has generally been controlled by those with control over access to the channel, which has been seen as potentially damaging to the flow of important ideas (with which channel owners might disagree or which they might not see fit to broadcast).

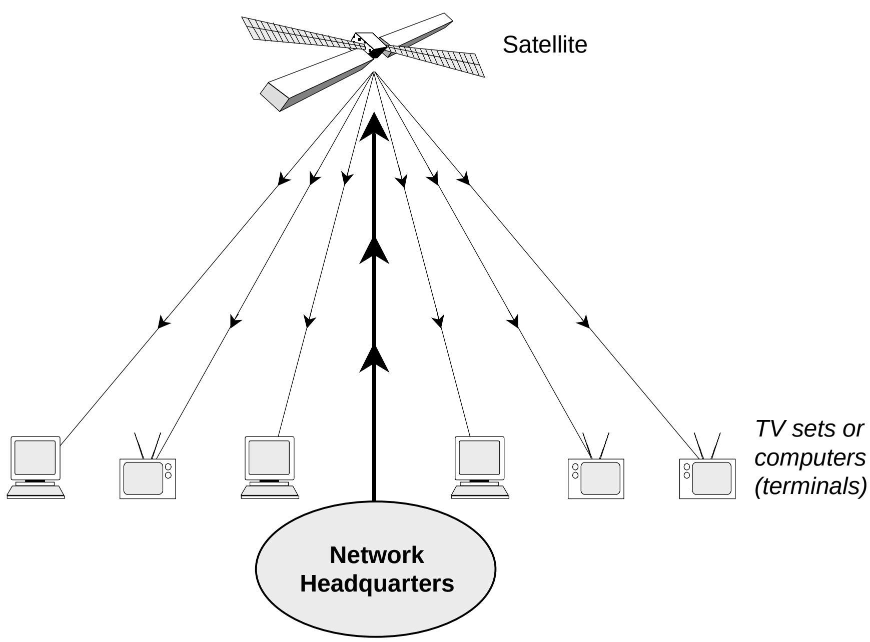


Figure 6-4
Direct Broadcast Satellite (DBS) Networks

Government therefore mandated that broadcasters must devote part of the capacity of their channel to programming for children, even if the broadcasters cannot make money on this service. They and cable operators must give access to those personally attacked on their programs and to political candidates, in the name of the competition of ideas. Access has been mandated also in the name of furthering economic competition. Under "must carry" laws, for example, cable channels must give access to local TV traffic. Satellite operators cannot be denied the opportunity to purchase programming (traffic) developed or controlled by cable operators.

Broadcast and cable system face the same operating problems as other networks with respect to the economies of scale and scope that make distribution of traffic more profitable from long-haul or high-density situations, where more traffic can be aggregated to cover costs. Local broadcast stations and cable operators incur all the costs associated with the channel at the local level, including costs for transmission facilities and staff and to maintain the license or franchise to operate. In discussing broadcasting and cable systems, the terminology becomes confusing: the program providers (senders) are known as "networks" (the American Broadcasting Company [ABC], the National Broadcasting Company [NBC], Fox), even though they do not operate the channels (except for the local stations they own).

Unlike their local affiliates, broadcast and cable "networks" are not licensed, because they do not have their own channels (they use their affiliates for local distribution and carriers, such as telephone lines or satellites, for long hauls). Broadcast and cable network companies have shared some of their profits with their local stations, in the form of direct payments, but in the 1990s, as the competition for viewers heated up and revenues declined, such payments have been reduced and several broadcast networks have talked of eliminating them altogether. The TV networks have instead asked their local stations to pay them for very expensive programming, such as professional sports.

As of the late 1990s, the prices charged for access by advertisers (except political advertisers) is not set by regulation but based on the anticipated number of viewers of a particular show. Given there are only twenty-four hours in a day, each channel has a limited capacity, and access for advertisers is allocated by peak-viewing-time pricing (i.e., higher prices for prime time) and variable pricing for preemptible (interruptible) service (i.e., a preemptible advertiser will lose its spot if a higher paying customer wants it).

Local operators must cooperate at a national level to set standards and procedures for scheduling, signalling, security, and ancillaries (e.g., frequency allocations and standards for equipment for sending and receiving)—a cooperation ordinarily facilitated or mandated by the FCC.[28] Congress charged the FCC with promoting communications services nationwide, and the

[28]The Radio Act of 1927, which gave the federal government the power to license broadcasters and allocate frequencies, was requested by broadcast interests when sales of radio receivers declined because of the enormous

FCC has sought ways to provide something like universal access to news and entertainment services, although the introduction of competition has made it difficult to maintain for low-traffic channels in rural areas or for small communities.[29] Competition from cable, satellite, and Internet services affected the weakest broadcast companies (usually the smallest) first and started a tide of cooperation in the form of Local Marketing Agreements (LMAs) and other agreements to share facilities as companies tried to survive by reducing costs.

New levels of competition have made cooperation difficult, to say the least, within and among various communications channels for such elements as transmission standards. The FCC finally declined to set standards for digital TV and will, instead, let the market decide whether a computer or broadcast standard will be adopted by consumers. This procedure has the advantage for consumers of giving them choice, but it also means that those choosing the wrong receiving equipment (TV set or computer) will be left with junk when the signal is sent in the other format.

Scenarios being played out in the other networks discussed in this report were evident also in entertainment networks. In the 1990s, broadcast, cable, and satellite networks all saw increased mergers and acquisitions, foreign direct investment, and breakdowns in cooperation. As a result, large networks evolved into smaller ones, program providers (including TV networks) became "feeders" for large (cable and satellite) networks, and competition and investment increased in long-haul services (national and international programming) and high-traffic (metropolitan) areas, where demand can be aggregated, lowering marginal costs. At the same time, investment and competition have been reduced (or they disappeared) from short-haul (local programming, except for news and sports) and low-traffic (rural) areas.

A significant difference between these and other communications networks is that the entertainment networks have always been highly competitive and the pricing of media products has always been based on the level of exclusivity—for example, a broadcaster pays more for a program to which it has exclusive rights in its own market. This concept is new to those growing up in the telecommunications industry: as the Internet becomes increasingly competitive, and comes more and more to resemble "media" networks, the concept of value through exclusive dealing may well invade a formerly cooperative and egalitarian system.

---

interference problem owing to lack of coordination. The new order in the system allowed the predictability necessary for large investments in the technology.

[29]See Longstaff, *Telecommunications Competition and Universal Service: The Essential Tradeoffs*.

Chapter Seven

Energy

Many generalizations here about networks that provide natural gas and electricity apply also to other forms of energy that do not require specialized channels or government rights of way (for transmission lines) to provide transport. Coal, for example, uses common carriers such as roads, barges, and railroads as channels, thereby avoiding direct regulation of price and service.

Energy networks look like multipoint-to-multipoint systems because they have many senders and receivers, but they do not carry specific cargo to specific receivers. The original sender of any particular unit of energy is unknowable once the unit is in the system, because units are undifferentiated commodities, with little variation in quality. Retail customers do not send energy to one another, and they seldom communicate with senders except to indicate (through monitoring devices) how much energy they use in the aggregate. The flow in energy networks goes nearly all in one direction, from suppliers to customers. Several companies may transport the energy product, but generally they are invisible to receivers. Retail customers ordinarily deal only with a local distributor.

But energy networks, like all networks, need systemwide agreements on signalling and scheduling in order to control capacity, deal with bottlenecks, and monitor security. The economics of long-haul traffic are different from those of short-haul, just as serving large customers is different from serving small ones. Participants in an energy network need to work together to set standards to ensure that terminals and ancillaries are compatible systemwide (or else they will battle it out in the marketplace, with the standard chosen by consumers eventually taking the market). In energy networks, as in transportation and communications networks, new technologies allow greater distribution of functions, unbundling many previously centralized (particularly for the generation of energy). Unbundling has allowed some large customers to generate their own power, bypassing the network and thus reducing the resources available to subsidize service to short-haul and low-traffic customers. With “deregulation” in the 1990s, and with the encouragement (if not assurance) of competition, came the need to develop rules for access to the network by competitors (who, owing to costs, cannot build competitive local distribution channels). Many new entrants aggregate demand from small customers, then buy capacity from incumbent energy suppliers at the wholesale rates available to large customers, and so make a profit on the difference between wholesale and retail without needing to make substantial investments in infrastructure. Independent brokers also have emerged to bring buyers and sellers together.

7.1 Natural Gas

Natural gas is found at oil deposits and often requires considerable transportation to reach retail customers.[1] Unlike solid fuels, natural gas cannot economically be stored or transported in trucks or rail cars. Cost-efficient transport requires a special pipeline (channel) between the well (senders) and the customers (receivers). Gas networks (see **Figure 7-1**) are unique in that they are based on a depletable resource with known reserves, a situation with important implications for their economics and regulation. For residential and industrial uses that require the application of heat (e.g., heating water and living space, cooking, and industrial processes such as smelting), gas competes with electricity. And gas can be used to produce electricity.

"Natural" gas was not the first gas industry to build networks. At the turn of the twentieth century the natural gas network took over the infrastructure and customers already developed for "town gas," which, in most cities in the United States at the end of the nineteenth century, was manufactured from coal. Because no technology then existed for long pipelines, natural gas companies were always local. Most were investor-owned, and they operated under a franchise from local municipalities. When arc welding and seamless pipes made long pipelines possible and when large compressors supplemented pressure at the wellhead to move the gas, natural gas became a viable competitor to town gas outside the oil fields.[2] Its use required customers to convert terminal equipment (e.g., stoves, furnaces). Natural gas eventually won, because it has higher British thermal units (BTUs) than coal gas and because it is cleaner and cheaper, all of which made it attractive to large industrial customers. When coal gas was no longer available even reluctant residential customers were forced to convert.

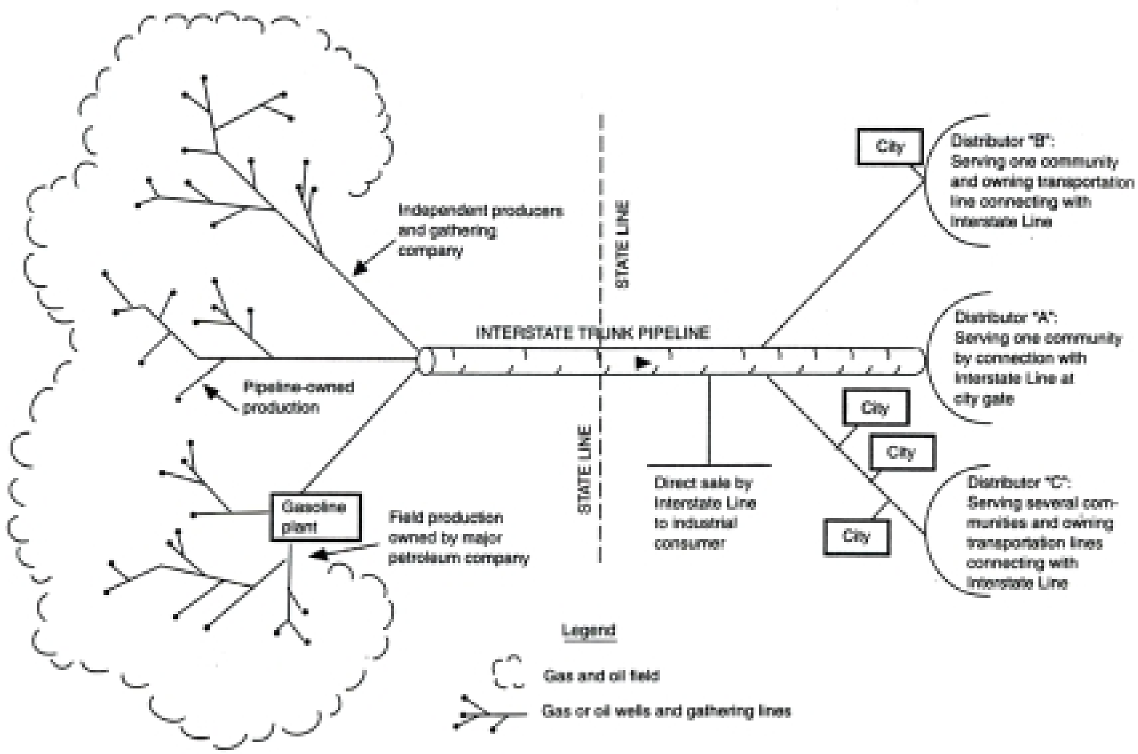
Like other industries, gas went through a period of consolidation, becoming in the early twentieth century a series of vertically and horizontally integrated holding companies with interests in many energy sources, including gas's major competitors, oil and electricity. This industrial structure reduced competition and raised prices for consumers, eventually pushing Congress to act. Several years after the stock market crash of 1929 (which made the power of large industrial organizations suspect), the holding companies were broken up by the Public Utility Holding Company Act of 1935. After having supplied service for almost four decades without government regulation, the natural gas industry was found by Congress to require extensive regulation, because it was a truly integrated business, i.e., production, transportation and distribution were inseparable and so needed to be regulated as a whole.[3] By the 1990s, the notion of integration as a basis for regulation was turned on its head as policymakers demanded

---

[1]This section draws on the following sources: Anrej Juris, *The Emergence of Markets in the Natural Gas Industry* (Washington, D.C.: The World Bank, 1998); and Richard H. Vietor, *Contrived Competition: Regulation and Deregulation in America* (Cambridge, Mass.: The Belknap Press of Harvard University Press, 1994).

[2]When only natural pressure at the well was available, distribution was limited to about 200 miles.

[3]Federal Trade Commission, *Utility Corporations*, Federal Trade Commission Report, Part 84-A (1936), 80.



Source: Richard K. Vietor, *Contrived Competition and Deregulation in America* (Cambridge, Mass., and London: The Belknap Press of Harvard University Press, 1994), Figure 13, 102. (Permission pending.)

Figure 7-1
The Natural Gas Industry

that the industry unbundle its services to create separate competitive markets in production, transportation, and distribution.

Regulation of the energy industry included price regulation of the production sector, where many producers compete to sell to pipeline companies. The regulated "just and reasonable" rate for producers consisted not of the market price for gas (the market could not be trusted to produce a "fair" price) but of the producers' depreciated costs plus a "reasonable" return on investment. This rate-making scheme reduced the incentive for suppliers to explore for new gas and kept the regulated price far below the market price, driving up demand while further reducing supplies. The shortages that resulted in the 1970s spelled the beginning of the end for producer price regulation and the beginning of the end also of regulation of a fully integrated system.

But before deregulation, the government attempted to increase supply through subsidies for R&D, to develop new sources of energy such as coal gasification (the old "town gas") to act as consumer substitutes. When prices finally were deregulated, the price of natural gas rose enough to justify new exploration, increasing supply and making government-subsidized substitutes both unnecessary and uneconomical.

Even after production had been left to market forces, the transportation and distribution parts of the system remained heavily regulated. Regulation had never been easy or simple, however, because of the difficulty of allocating costs between intrastate (prices set by state authorities) and interstate services and of allocating fixed and variable costs among customers. Policies that attempted to ignore market realities such as peak-load pricing had unsuspected consequences (distorting supply and demand) and became closely tied to political considerations. Before 1952, for example, gas prices included a demand charge (to cover the fixed costs of providing peak-load capacity) and a commodity charge (to cover the volume actually purchased). This pricing system became politically untenable when it became clear that large customers could order interruptible service (that is, during peak periods they could switch to another power source), thus avoiding paying high-demand charges and moving more of the fixed costs to small customers. The difficulty was "corrected" by attributing half the fixed costs to the commodity charge, but this solution often reduced the price for gas at peak periods to below average cost, encouraging use during peak periods and disrupting logical capacity planning.

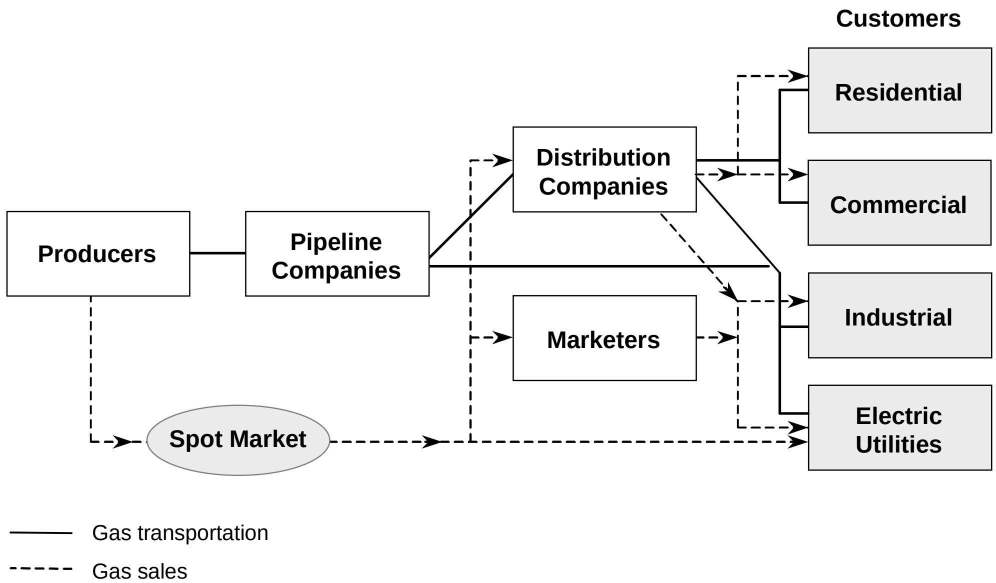
In the 1970s, when rising prices led large customers to seek lower cost, alternative fuels, these careful political balances were upset. As customers left the gas network, more of the costs had to be borne by small customers that lacked such flexibility. The flight of large, low-cost customers illustrates a dark side of interindustry competition in networked industries: customers that can leave the system to seek lower prices elsewhere will do so, bequeathing to those that cannot afford to convert (small business and residential customers) the burden of paying the network's costs and, thus, eventually encouraging more customers to leave.

The combination of deregulation of wholesale gas prices and continued regulation of the other parts of the network brought unanticipated problems.[4] Market prices are rarely stable, and when a market-driven system (production prices) meets a regulated system (transport and distribution prices), something must give. Those at the fault line are regulators and policymakers. To reach the final customers (receivers), the competing producers (senders) need access to the channel at "reasonable" rates. The gas transmission and distribution companies did not welcome the new freedom for senders (producers) brought by deregulation, because the senders were also seeking access to compete with them for retail customers. Producers saw the transmission and distribution companies as bottlenecks in the network, which, by setting high transport charges, could keep them from reaching customers. High access fees led producers to call for regulation, to facilitate competition at the retail level. They maintained that rules for access must create "fair" competition between those that own the channel and those that would compete with them. The

[4]For an extensive discussion of the early deregulation of the gas industry, see Vietor, 124-157.

producers also asserted that rules would have to incorporate antitrust law principles forbidding cross-subsidies and discrimination by the owners of these "essential facilities."

After many political tradeoffs and legal debates,[5] the provision of natural gas was unbundled and left to develop separate markets for product, transportation, and distribution, with the structure of the industry itself left to change in response to the new relationships (see **Figure 7-2**).[6] The transition had substantial costs—its actual cost has been estimated at $10.2 billion as of 1995[7]—as old contracts and assets were revalued. Until the apportionment of those costs was worked out, competition could not be implemented successfully.



Source: Andrej Juris, "Development of Competitive Natural Gas Markets in the United States," *Private Sector* (March 1998), Figure 3, 23.

Figure 7-2

Unbundling of Gas Sales from Pipeline Transportation, 1992 and Beyond

The new structure of the industry encouraged mergers and consolidations to gain economies of scope and scale intended to improve competitive advantage and create large, integrated companies. Rates for transport could then take into account both the fixed costs of the pipeline and the varying charges designed to assist in capacity planning, such as "firm" and "interruptible"

---

[5]Culminating in Federal Power Commission Order No. 436 of 1985, and Federal Energy Regulatory Commission Order No. 636 in 1992.

[6]For analyses of the impact of new regulatory frameworks, see, e.g., *New Horizons in Natural Gas Deregulation,* edited by Jerry Ellig and Joseph P. Kalt (Westport, Conn.: Praeger Publishing, 1996).

[7]Andrej Juris, "Development of Competitive Natural Gas Markets in the United States," *Private Sector* (March 1998), 21-28.

service, with a secondary market for transport developed by those with firm transportation rights. Most transactions could then take place at hubs established at the intersection of one or more interstate pipelines. Hub operators could then offer a variety of services, including transportation, storage, processing, and trading, and the hubs could operate as a spot market to set the price of natural gas on a continuous basis.

By the late 1990s, retail prices for natural gas had fallen, although not uniformly. Large customers saw drops of as much as 31 percent, while small and residential customers saw on average reductions of only about 12 percent.[8] Competition for residential customers has been slow to emerge.

7.2 Electricity

Before there were gas and electric networks, all power supplies (e.g., wood, coal) had to be delivered to individual customers using the transportation networks discussed in **Chapter Five**.[9] The short-haul portion was the most expensive, and a more efficient way to distribute energy at the local level was necessary for the increasingly urban and industrialized societies of the late nineteenth and early twentieth centuries. The use of electrical power allowed oil, coal, and other fuels to be delivered to power plants, which then converted the fuel into electricity which was then delivered to individual customers by wires over multipoint-to-multipoint networks. In modern electric networks, electricity flows from power-generating facilities by a grid of transmission lines to a local power-delivery service, which brings the electricity to homes and businesses. The industry thinks of itself as divided into three distinct operations: generation, transmission, and distribution, within a model similar to that suggested here. It follows the model suggested here, consisting of senders, receivers, channel, transport, traffic, etc., but with one important difference: the nature of electricity.

Unlike traffic in transportation and communications networks, electricity flowing along connected lines cannot be directed. Electricity generated at one power plant merges with power produced by all other power plants connected by the same system, so receivers cannot know (nor do they care) where their power comes from. This integration means that taking power out at one point of the system decreases flows throughout, and that if power is blocked at some point, the

---

[8]Ibid., 25

[9]This section draws on the following sources: Brian Bowers, "Electricity," in *An Encyclopedia of the History of Technology*, edited by Ian McNeil (London and New York: Routledge, 1996); Richard J. Gilbert and Edward P. Kahn, "Competitive and Industrial Change in U.S. Electric Power Regulation," in *International Comparisons of Electricity Regulation*, edited by Richard J. Gilbert and Edward P. Kahn **(**Cambridge, Eng. and New York: Cambridge University Press, 1996); and *Reshaping the Electricity Industry: A Public Policy Debate* (Cambridge, Mass.: Harvard Electricity Policy Group, Harvard University, 1996); and *A Shock to the System: Restructuring America's Electricity Industry*, edited by Timothy Brennan, Karen L. Palmer, Raymond J. Kopp, Alan J. Krupnick, Vito Stagliono, and Dallas Burtram (Washington, D.C.: Resources for the Future, 1996).

flow in the system will back up, rather like water in pipelines.[10] Because electricity cannot be stored in large quantities, the flow in the network must be constantly monitored and balanced. The nature of electricity and the frequent cross-linking and radial linking of the parts of electrical networks creates a great need for cooperation in signalling and scheduling traffic in the network and, once competition is introduced, makes policymaking complex.

The first electric generators were built to provide energy to large commercial establishments as a replacement for other fuels, usually coal. The first generator built to offer service to the public was in Surrey, England, in 1881: a local mill operator offered his waterwheel to drive a municipal electrical generator in exchange for free lighting for the mill. A Siemens generator was installed, but the network was not a commercial success. Only eight or ten households converted to electric lighting before the system shut down. A technical lesson learned from this experiment was that power was lost between the generator and the receiver premises, allowing the conclusion that generators could never be more than half a mile from their most distant customers and that electric networks would always be local. A year later, in 1882, in New York, the first electric generating station in the United States opened. By 1902, 3,620 electric services were running in cities and towns across the United States. These were local generation facilities and distribution networks. Some large cities even had competing electric services, as companies fighting for a foothold in high-traffic areas sometimes ran electric lines down both sides of streets. In Chicago, for example, forty-eight companies tried to offer service in various parts of the city.[11]

Military interest in electricity dates from a high-voltage “hydro-electric machine” for detonating mines.[12] Because electric power was not portable—that is, it could not be carried onto the battlefield—military efforts to subsidize its use have concentrated on the development of batteries and other storage mechanisms. (This research has also found its way into the transportation networks, in efforts to develop electric cars to meet air pollution standards.) R&D for ancillaries also had a impact. Research on jet engines, for example, funded by the DOD, made possible the gas turbine engines with which large industrial customers generate their own electric power.

In the early twentieth century, government subsidies were used to develop street lighting and electrically powered mass transit (the “gondolas of the people”) as public works. Government subsidy was necessary, because a large user was important to attract investors worried that electricity might not unseat gas as the preferred energy source and because customers would need

---

[10] A new technology, known as FACTS (flexible alternating current transmission system), which can increase or decrease power flow on particular lines and may help alleviate the back-up problem. See *Electricity Technology Roadmap: Powering Progress, 1999 Summary and Syntheses* (Pleasant Hill, Calif.: Electric Power Research Institute [EPRI], 1999), 29-30.

[11] Christopher Flavin and Nicholas Lenssen, “The Electric Industry Sees the Light,” *Technology Review* (May–June 1995), 44.

[12] Bowers, 351.

to replace gas appliances with electric ones. The purchase by government of electric streetlights, tramcars, and underground subways provided the demand industry needed to build large and efficient generators and to convince investors that electricity was indeed the wave of the future.

Like the other networks discussed in this report, electric power was seen as an engine of economic development. Many communities offered incentives to attract electric power companies, and those not successful in attracting them went to the federal government for inexpensive loans, tax breaks, or direct subsidies in order to form local electricity cooperatives or municipally owned systems. Government involvement was common in low-traffic (mostly rural) areas.

In the United States, as early as 1898, some industry leaders, including Samuel Insull, one of the cofounders of the electric industry, sought regulation to avoid "debilitating competition."[13] In 1907, New York and Wisconsin began to regulate the industry, and the other states soon followed.

As technology facilitated the transmission of electrical power over longer and longer distances, small companies were consolidated or combined to build large generation facilities, which offered better economies of scope and scale. This concentration continued until, by the 1920s, many local businesses had been merged into giant holding companies similar to those taken over the railroads and many other industries. During the Depression the stock of electric utility holding companies collapsed amid many allegations of fraud, which ultimately attracted the attention of the federal government. The Public Utility Holding Company Act of 1935 restructured the electric power industry by limiting the scope of each company's operations geographically (usually limiting operation to one state) and economically (limiting operation to the utility business and forbidding certain contacts among affiliated utilities). That same year, Congress gave the Federal Power Commission (FPC) broad power to regulate interstate transmission of power, rates for wholesale power contracts, mergers and acquisitions, and other contacts among utilities that might lead to collusion. Two years earlier, in 1933, with the establishment by Congress of the Tennessee Valley Authority (TVA), the first federally operated hydroelectric facility, government entered as a player into the electric power business.

But electric companies could not go back to being only local networks. They needed to cooperate for signalling and scheduling, and, over time, local networks were connected with large networks, or grids, where energy was shared to meet fluctuating local demand and production took place far from central cities. By 1964, the electric power system had grown so complex that the FPC warned that lack of coordination might soon hurt its reliability. After the huge blackout of the northeastern United States in the following year, the industry acted to establish the voluntary North American Electric Reliability Council (NERC), comprising ten local councils

[13] *A Shock to the System*, 23.

that coordinate service for three highly coordinated grids spanning the United States, Canada, and parts of Mexico.

In the late twentieth century, government has offered subsidies for the development of alternative or renewable sources of energy, such as wind and geothermal energy, to reduce dependence on foreign oil and to lower the pollution from the use of fossil-fuel power generators. Alternative generators often have higher costs than traditional power plants, but, on the theory that buying from alternative generators will allow new technologies to attract the capital necessary for development, regulators forced electric utilities to buy from them. Buying this more expensive power artificially increased the wholesale cost of electricity and caused problems when the electric industry joined other "deregulated" market-driven networks that were keeping costs low. Regulators soon found subsidies for other public policy goals such as alternative energy and universal service hard to mandate.

So long as electric generators were local operations, the need for industrywide standards was limited; but as soon as they were connected with one another (and to millions of homes and businesses with electric appliances), a choice had to be made between alternating current (AC) and direct current (DC). Adjusting voltages by using transformers was easy with AC, but generators cannot operate well in parallel. DC can use batteries to store power for times of low demand and can be sent over long distances more easily than AC. Because terminal equipment (commercial equipment and home appliances) had to be compatible with the current, customers wanted to buy what would be useful in the most places. The battle between competing technologies for dominance led to the now familiar scenario in which only one standard survives.[14] It was fought between Thomas A. Edison, who favored DC, and George Westinghouse, who favored AC. Edison lobbied several state legislatures to ban AC, because, he asserted, it was more "dangerous" than DC.[15] Eventually, AC won the most marketshare and all customers converted from DC. But the marketplace does not set standards overnight, especially when customers have invested a lot of money in the "loser." For example, England did not completely switch to AC until 1985, when the last customer (a Fleet Street newspaper) made the change.

After the competitive era had faded into history with the advent of New Deal regulation, the electric industry in the United States cooperated in the development of new technology through the Electric Power Research Institute (EPRI). The new competition mandates of the 1990s, however, have put EPRI's future in doubt, because the pressure on companies to cut costs may make dues to EPRI seem unnecessary overhead. Interest in collaborative research may again

---

[14]See, e.g., Martin C. Libicki, *Information Technology Standards: Quest for the Common Byte* (Boston and London: Digital Press, Butterworth–Heinemann, 1995); and Cristiano Antonelli, "The Economics of Standards," *Information and Economics Policy* **6** (December 1994, Special Issue).

[15]See Thomas P. Hughes, *American Genesis: A Century of Invention and Technological Enthusiasm 1870–1970* (New York: Viking Press, 1989).

emerge as companies gain in experience with the kinds of technology information that can be shared with competitors. New technology is also one of a limited number of ways by which electric companies can differentiate their products, so companies are not likely to cooperate in technology that either can be deployed at the retail level or will reduce the costs of production. The effect on U.S. electric companies (and on companies that make ancillary equipment) may be disastrous, especially if innovations from other countries were to be introduced before the industry could marshal R&D to compete.

As with the other networks discussed here , many countries are now introducing competition into the electric industry.[16] Renewed faith both in the ability of the marketplace to regulate prices and in new technology (particularly the gas turbine engine) has allowed greater distribution of the generation of electric power. As a result, large customers can bypass the network by generating their own power, with the additional incentive for them that local electric utilities are required to buy any of their excess power.

As of the late 1990s, whether lower prices can be accomplished by competition at the wholesale level or whether they must include sales at the retail level remains unclear. Some commentators are still convinced that, because of the high costs of building infrastructure, local delivery of electricity is a "natural" monopoly.

As in telecommunications, some policies call for granting retail competitors access to an incumbent's local channel at a "fair" price, which would permit them to act as a reseller or aggregator of demand for local distribution. Given the undifferentiated nature of electric service, forced access may have less appeal for long-term regulation of price, because either the incumbent or the competitor will not survive a long price war. Policymakers supporting retail competition believe price can be combined with auxiliary services, such as security systems or telecommunications services, to form a basis for sustainable competition.[17] Regulators at the state level seem to assume that in most states local service will not become subject to market forces and that oversight by local government will still be needed to assure "integrated resource planning."[18]

Competition is complicated also at the wholesale level. It involves forcing competitors to cooperate to keep the system operating. Some proposals promote competition at this level through increased competition for bilateral contracts between power generators and local distribution

---

[16]See, e.g., *International Comparisons of Electricity Regulation*; and Elliott Rosen and Anil Malhotra, "The Dynamics of Independent Power," *Private Sector* (March 1996), 21-24.

[17]See *Service Opportunities for Electric Utilities: Creating Differentiated Products,* edited by Shmuel S. Oren and Stephen A. Smith (Boston: Kluwer Academic Publishers, 1993).

[18]See, e.g., Rodney Stevenson and Dennis Ray, "Transformation in the Electric Utility Industry," in *Networks, Infrastructure, and the New Task of Regulation*, edited by Werner Sichel and Donald Alexander (Ann Arbor: University of Michigan Press, 1996), 89–105.

systems. Others propose creating a complex organization, called a Poolco (see **Figure 7-3**), to act as an independent coordinating body, to solicit bids from generators and transmission companies on behalf of local distributors and to operate a spot market for short-term power supplies. A Poolco can fulfill many signalling and switching functions necessary to keep the power in the system balanced. Most proposals retain regulation of the transmission function (and of companies that operate the transmission grid), because in a competitive system access to the grid is critical to both senders and receivers.[19]

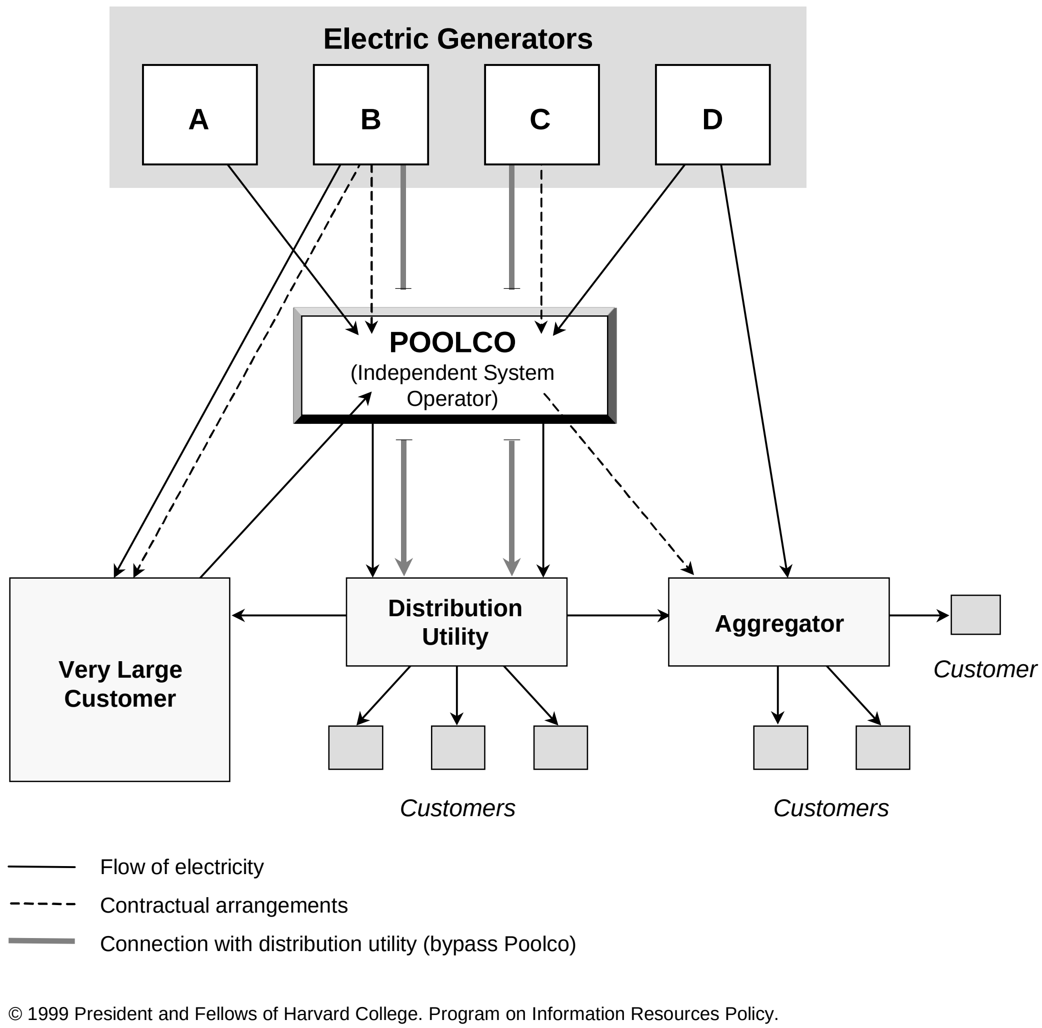


Figure 7-3
Electric Power System and Poolco

---

[19]For a discussion of power pools and their efficiency, see Richard J. Gilbert, Edward Kahn, and Matthew White, "The Efficiency of Market Coordination: Evidence from Wholesale Electric Power Pools," in *Networks, Infrastructure, and the New Task of Regulation*, 37–58.

Electricity rates vary widely across the United States, for many reasons, such as subsidized federal power (from the TVA and other similar projects) in some areas, high local costs for fuel or its transportation, as well as variations in state regulatory policy. Because electricity is necessary to the operation of many industries, its cost can lead manufacturers to move to an area of lower cost, thereby disrupting established local economies. As of the late 1990s, retail rates will probably remain a critical factor in the ongoing policy debates.

Pricing factors seen in networks that need to be taken into account in those debates include the following:

- *Peak/off-peak pricing* is an especially powerful factor, because electrical networks can use it to even system loads, reducing peak capacity by encouraging a shift to off-peak times in exchange for lower prices. Power produced at peak times is the most expensive, because it ordinarily involves putting more generating facilities (usually those that are less efficient) on line or buying power in the spot market. *Interruptible/noninterruptible distinctions* allow customers to make decisions about risks for unpredictable peaks and about their willingness to forgo power rather than pay the spot price. Such distinctions allow an industrial facility to pay less per kilowatt hour in exchange for the right of the power supplier to interrupt or reduce power during peak times. But interruptible pricing may not be feasible for many industrial or residential uses. If most customers are unwilling to have power interrupted, in a shortage power to everyone can be reduced, as in a "brown out."
- *Distance* becomes a factor, because the cost of distribution increases with every mile, owing to costs of rights of way and increased loss of power. Distance pricing is politically untenable in rural areas, where distances from the power source may be greater and farmers and rural industries would pay more for power.
- *Volume discounts* allow power suppliers to sell to heavy-traffic customers at just above marginal cost. The ability to offer discounts is crucial in a competitive environment, where large customers are the least expensive to serve and therefore potentially the most profitable. Discounting is important also because heavy-traffic customers can use new technology to generate their own power and thus bypass the network entirely. Volume discounts also allow resellers (those competing for retail sales but without having their own equipment) to compete with the local distribution company by aggregating the demand of many low-traffic customers to obtain lower rates.
- *Long-term contracts* allow senders and receivers to plan for capacity needs and to invest in the network sufficiently to develop strong cooperative ties. But when used to defeat the interests of competitors or customers, these contracts can raise antitrust questions and regulatory eyebrows.

As competition was being implemented in the late 1990s, it was unclear whether consumers, even large industrial ones, were prepared for the effect of market pricing on the prices they paid for electricity. In June 1998, competition in the wholesale market, along with the emergence of independent power-trading companies (which bought and sold power contracts), created a volatile market response to a Midwestern heat wave. In one week, the market price of electrical energy skyrocketed from $30 per megawatt hour to $7000.[20] Industrial customers that had agreed to interruptible service found their service drastically reduced. Some had the option of continuing service at the market price, while others could not get power at any price. Faced with commitments for power that they could not fill at the market price, some independent power traders simply defaulted, causing some large customers and some local electric utilities to demand greater federal regulation of this newly "deregulated" part of the industry. Customers complained that the utilities had conflicts of interest in such cases, because their new power-trading subsidiaries stood to profit from price fluctuations.[21]

Some analysts have predicted that small business and residential customers also will face increased power costs after deregulation, especially in states that previously had low electric rates. Customers without competitive service must stay with the incumbent provider and now have the smaller profit potential from low-traffic and short-haul service, and they may be stuck paying for all the stranded costs associated with the unused capacity created when high-traffic customers leave for other providers.[22]

Even as the electric power industry struggled to cope with new intraindustry competition, some in the industry prepared to venture outside its traditional business to compete with telecommunications and cable companies in delivering telephony and Internet services. Several electric companies claimed to have developed technology that allowed them to provide these services over standard electric wires.[23] But these new ventures could further diffuse their resources and leave them more vulnerable to competition or reduce investments to insure reliability of their services, or both. By the beginning of the year 2000, the reliability of the electric grid in the United States was already being questioned by both the industry and regulators, and many were backing proposals to give regulatory powers to a quasi-government organization known as the North American Reliability Organization.[24]

---

[20]Katherine Kranhold, "Midwest Firms Cry Foul on Outages" *The Wall Street Journal*, July 10, 1998, A2.

[21]Ibid.

[22]Michael Sel, "Electric Utility Deregulation Is Seen as Costly for Small Business Owners," *The Wall Street Journal*, Oct. 3, 1996, B2.

[23]Gautam Nail, "Electric Outlets Could Be Link to the Internet," *The Wall Street Journal*, Oct. 7, 1997, B6.

[24]John J. Fialka, "Electric-Power Grids' Reliability Erodes: Energy Department Cites Demands of Economy, Utilities Deregulation," *The Wall Street Journal*, Jan. 13, 2000, A-2.

Chapter Eight

Conclusions

Fascinating similarities exist generally in the development and operation of transportation, communications, and energy networks which planners may do well to take into account. While certainly no claim is made here that all networks operate in exactly the same way or that predictions are possible for all (or any) of them, the experiences of one network may at least be interesting in relation to other networks and may offer possible scenarios, even if these cannot be relied on for predictions about the others.

Immediately relevant patterns for business planning and policymaking appear to surface, however, with the introduction, or reintroduction, of *intra*- or *inter*industry competition into a networked industry. In the United States, starting with airline "deregulation," newly competitive networks (including the Internet) underwent the following experiences:

- New entrants that aggregated demand for long hauls and large loads were later mostly absorbed by incumbents or went out of business owing to economies of scale and scope enjoyed by incumbents;
- A vast wave of mergers and acquisitions, which occurred as players attempted to develop further economies of scope and scale;
- Foreign direct investment, as players looked for resources to upgrade infrastructure in order to fend off competition;
- Less cooperation among parts of the network, which resulted in problems of scheduling and security;
- The development of separate networks (hub-and-spoke configurations, developed by each competing network) to keep customers out of competing networks;
- The development of "feeder" lines for each network;
- Price decreases in price-elastic segments of the market, and price increases in inelastic segments;
- Increased competition and lower consumer prices for long-haul routes and high-density areas in the network, but decreased competition and higher consumer prices (and investment) in short-haul and low-density portions; and
- Decreased quality of service.

Such occurrences may not take place in all the networks discussed here, but they are certainly possible. To increase the likelihood of desired outcomes of business strategy and public policy, scenarios may be anticipated and planned for. For builders of the newest networks—computers—the lessons of these experiences are worth noting, particularly in such areas as interconnection, capacity planning, cooperation for security, the impact of ancillaries and, of course, government relations.

Both business planners and policymakers may profit from considering the nature of networks when making assumptions about the effect of competition in and on them. With the new-found interest in the benefits of competition, cooperation appears also to offer distinct public (and private) benefits that may be dangerous to ignore.

Acknowledging that networks have much in common and much to teach one another may be overdue, and a new interdisciplinary study of networks (one that would include scholars from the biological sciences) may help tease out even more threads of information which can then be interwoven to reveal patterns for the development and regulation of networked industries in the next century.

Acronyms

AC	alternating current
AEI	American Enterprise Institute
BTU	British thermal unit
CPU	central processing unit
DBS	direct broadcast satellite
DC	direct current
DOD	Department of Defense
EPRI	Electric Power Research Institute
FCC	Federal Communications Commission
FPC	Federal Power Commission
GATT	General Agreement on Tariffs and Trade (Uruguay Round, 1994)
IAP	Internet access provider
ICC	Interstate Commerce Commission
ION	Integrated On-demand Network™ (Sprint)
IP	Internet Protocol
LAN	local area network
LMA	Local Marketing Agreement
NERC	North American Electric Reliability Council
PSN	public switched network
R&D	research and development
TV	television
TVA	Tennessee Valley Authority
WAN	wide area network

PPPHLONG

ISBN 1-879716-60-7