5-2006

# Communications

Alexander J. Field
*Santa Clara University*, afield@scu.edu



Recommended Citation
Field, Alexander J. 2006. "Communications," in Historical Statistics of the United States: Millennium edition. Cambridge: Cambridge University Press, volume 4, pp. 977-98.


https://doi.org/10.1017/ISBN-9780511132971.Dg.ESS.01

# CHAPTER Dg
# Communications

*Editor: Alexander J. Field*

## Communications

Alexander J. Field

The communications sector of an economy comprises a range of technologies, physical media, and institutions/rules that facilitate the storage of information through means other than a society's oral tradition and the transmission of that information over distances beyond the normal reach of human conversation. This chapter provides data on the historical evolution of a disparate range of industries and institutions contributing to the movement and storage of information in the United States over the past two centuries. These include the U.S. Postal Service, the newspaper industry, book publishing, the telegraph, wired and cellular telephone service, radio and television, and the Internet.

The communications sector, regrettably, has received relatively little attention from economic historians. None of the three volumes of the *Cambridge Economic History of the United States*, for example, dedicates a chapter to it (Engerman and Gallman 1996, 2000). Volume 1, on the colonial period, contains virtually nothing on the subject, not a single index entry even to the carriage of letters or the publication of newspapers or books. Volume 2 mentions, in passing, the post office, the telephone, and the telegraph, but there is no organized treatment of the origins or economic impacts of these systems. Volume 3 does reference communications in a number of its essays and provides some systematic discussion of late-twentieth-century telecommunications policy. But the technological and regulatory issues we continue to deal with in the first decade of the twenty-first century did not arrive full-blown with the breakup of the Bell System or the explosive growth of the Internet. There is a technological and historical context to these developments, and this essay, along with the statistical series whose interpretation it illumines, attempts to provide it.

Communications industries or sectors satisfy human needs directly by selling or providing goods or services to households. In addition, they provide services to other businesses as intermediate inputs in the production of final goods and services. In business-to-business transactions, communications services have been and remain especially important in implementing capital-saving strategies in customer industries. Such practices save resources by enabling a given stock of physical capital to be used more intensively. Innovative capital-saving strategies, along with labor-saving ones, have also loomed large in the history of communications industries themselves. Where appropriate, the essay calls attention to these strategies.

Transactions in this sector can be roughly divided between those involving one transmitter and one receiver (point to point), and those that involve one transmitter and many receivers (broadcast). Different segments of the communications sector have facilitated different blends of these two main types of service, and the blends have sometimes changed over time. Further subdivisions and identification of intermediate categories are also possible.

In the one-transmitter/multiple-receiver case, for example, we may distinguish narrowcasting from broadcasting. In usages more inclusive than is common, I will characterize the maintenance of a Web site, as well as book, newspaper, and other print media publication, as broadcasting, alongside traditional examples from the radio and television industries, but treat junk mail delivery or email sent to an electronic mailing list as narrowcasting. The distinction, although not hard and fast, is that in the latter cases, audiences have been more selectively targeted by the sender than would result from the simple choice of a medium or channel.

In point-to-point communications, we may distinguish between those in which messages are exchanged in more or less real time (synchronous), and those in which sending and receiving are separated by an interval (asynchronous). Communication between two parties may be simplex (information flows in only one direction), half duplex (it flows in either direction, but only one way at a time), or full duplex (it flows at the same time in either direction). Generalizations about particular parts of the communications sector, although often robust over extended periods of time, are at risk in rapidly changing technological environments.

### Changes in Communications to 2000

The intent of this chapter is to provide an overview and a statistical window into how the nation has moved and stored information for more than two centuries. The tables are organized into four groups: telegraph and telephone; radio and television; U.S. Postal Service and delivery services; and books, newspapers, and periodicals. This organization provides continuity, in the context of newly added series, with the organization of the preceding editions, although much has changed in the communications sector since 1970.

The first group of statistical tables, which now includes the Internet, covers the telegraph and telephone and the infrastructure necessary to deliver all of these services. The telegraph, an industry whose birth, growth, and senescence is chronicled in these pages, predominantly serviced demands for point-to-point communication, the major exceptions being the system for broadcasting stock prices to specialized printing telegraphs (stock tickers) and the technology used by wire services such as the Associated Press to provide raw material for newspapers.

***Acknowledgments***
Alexander Field thanks Peter Temin and Gavin Wright for comments on earlier versions of the essay.

Even more than the telegraph, the telephone network was specialized for two-way point-to-point communication, in spite of the fact that some of Alexander Graham Bell's early promotional stunts envisaged broadcast (for example, playing "The Star-Spangled Banner" in Boston so that an audience of 2,000 could hear it in Providence or having an opera singer in Providence entertain an audience in Boston) (Casson 1910, p. 51). Although telegraphic communication was traditionally asynchronous, the telephone satisfied the demand for synchronous communication. This distinction, however, has been weakening as a consequence of the diffusion of ancillary hardware and software, such as answering machines and voice mail. And certainly some telegraphic queries and responses, such as those encouraged by the TWX (teletypewriter exchange) system introduced by the American Telephone and Telegraph Company (AT&T) in 1932, were sufficiently close together in time that we may view the resulting communication as synchronous in the same way that emails or instant messages are rapidly exchanged over the Internet.

An emphasis on two-way communication as a distinguishing feature of telephony is also arguably vulnerable in the present technological environment, at least if both parties to the conversation are assumed to be individuals. Computerized systems can now provide customized one-way messages for automated reminders of dentist or doctor appointments or let one know that one's child did not show up for school. And customers can communicate directly with computers using voice menus, touch tone buttons, and increasingly more sophisticated voice recognition software. Thus, either the initiator or responder in a two-way exchange may now be a computer rather than another human. How widely these new modalities will ultimately penetrate into the traditional spheres of phone use remains to be seen. And whether the telephone itself, as a transmitting and receiving appliance, will ultimately remain distinct from the computer is an open question.

The Internet, understood best as an evolving set of rules or protocols for interconnecting computers and networks of computers, was barely a year old in 1970. In the last three decades of the twentieth century, these shared rules, in conjunction with rapid technological advance in the manufacture of computing equipment, facilitated a growing variety of cross-machine interchange involving remote access, file transfers, electronic mail, and transmission and receipt of text and graphical information using the hypertext markup language (HTML) format. Time-sharing technology predated the Internet, and many of these features were available prior to 1969 for users of a single mainframe. What is new with the Internet is the ability of users of different central-processing units easily to communicate and exchange data with one another.

Electronic mail (email), the most important use of the Net for its first two decades, satisfies, as did the telegraph to which it is most closely related, the demand for asynchronous point-to-point communication. The invention of the electronic mailing list shortly after the appearance of email, however, quickly enabled narrowcasting. With the opening of the Net to commercial use, more indiscriminate broadcast to email addresses, known as spamming, has engendered protests and resistance from users, as well as actions by Internet service providers (ISPs) to reduce its incidence.

From the standpoint of households, an ongoing social and policy issue will remain the appropriate segregation of commercial and personal communication within these new channels. In the past, lines of separation were clearly if somewhat arbitrarily drawn. Magazines and newspapers contained commercial messages, but books did not. Radio and television were largely supported by advertising, but films were to a much lesser degree. The Postal Service might transmit "junk" commercial messages; the telephone, historically, did not.

In threatening traditional boundaries, changing technologies have led and will continue to lead to new debates and policy choices. Unsolicited narrowcasting over phone lines to fax machines, for example, is now legally restricted, as a result of complaints about the costs of materials required to print out advertisements. How will this play out with respect to voice communication? Recorded political messages narrowcast to thousands of phone lines are an increasingly common campaigning tool, and there is no technological obstacle to the use of this method for the transmission of commercial messages. Will such activity backfire by generating ill will for the sender? If it is acceptable for a human to cold-call and pitch a commercial product, why not a machine?

Conflicts between commercial and private communication space were originally less of an issue on the World Wide Web, as compared with email, fax, or phone, because individuals voluntarily choose access to commercial content, for the most part, whereas they disliked having it thrust before their ears or eyes. Nevertheless, the increasing sophistication and intrusiveness of pop-up and pop-under ads and their variants are softening this distinction. Sites on the Web, because of their interactivity, both broadcast and provide opportunities for point-to-point communication, typically over secure (encrypted) channels if money or confidential information is changing hands. Web sites, identified by their URL (uniform resource locator), are computer servers prepared on demand to transmit and receive material formatted in HTML according to the hypertext transfer protocol (HTTP). These protocols offer greater display flexibility and interactivity and quicker access to related data or sites than was possible in the 1980s with bulletin board systems, and they continue to become more sophisticated. In 1970, of course, there was no World Wide Web (its rules would not be set forth for another twenty-one years), let alone access to computer bulletin boards, and until 1992, the Internet was off bounds for openly commercial enterprise. The 1990s saw enormous growth in the commercial uses of the Web.

Conflicts between freedom of expression and perceived threats to public safety or morality remain active policy issues with respect to the Internet, as they have been with respect to radio and television broadcast and the publication of print or recorded media. Easier (and cheaper) access to some types of specialized information (that involving explosives manufacture or pornography, for example) has not met with universal acclaim, and there continues to be legal and public policy debate about such issues as the use of filtering technology in public access terminals (for example, those in libraries). Debates about the appropriate individual, political, or social control of information flows will undoubtedly persist.

At the end of the 1990s, as the stock market bubble reached its peak, many businesses, desirous of exploiting opportunities to make product data easily available, moved aggressively to take advantage of new ways of "grabbing eyeballs." At the same time, corporations struggled to protect proprietary information behind internal firewalls, and the growing popularity of wireless networks aggravated these problems. In 2000, households provide information about themselves at a much cheaper cost than was previously possible, and they enjoy much more convenient access to a wide range of information, goods, and transactions. But individuals have also had to come to terms with the consequences of inexpensive methods for keeping track of what one has been reading, watching, or buying, and with whom one has been communicating and about what. The transformation of the computer industry and telecommunications in the last three decades of the twentieth century raised, and continues to raise in heightened form, long-standing issues about personal privacy in a democratic society.

The second major group of tables of this chapter chronicles the radio and television industries and their associated infrastructure. Here, the main business model in the United States has entailed broadcasting to final consumers who support the services through advertising costs built into the prices of purchased goods and services, and, to a lesser degree, through direct fees for service, charitable contributions, or tax subsidies. Tables Dg131–161 provide data on advertising revenues for commercial radio and television, and Table Dg172–180 provides details on income sources for public broadcasting. Table Dg117–130 chronicles the growth in the total number of radio and TV broadcast stations, both commercial and noncommercial, along with trends in the number of radio and television sets produced. It also provides information on the growing importance of cable systems and videocassette recorders, each of which has offered customers access to greater diversity of content, and the option, for a price, of consuming entertainment without the interruptions of commercial messages. In 1970, the Public Broadcasting System and National Public Radio were in the process of being created, and home videocassette recorders (VCRs) were unavailable before 1975. At the end of the twentieth century, home VCR technology was a maturing technology, already being supplanted by digital video disks (DVDs) as well as digital video recorders (DVRs).

DVRs and accompanying services, such as those currently provided by TiVo, represent an evolving technology whose potential is only beginning to be appreciated by consumers and other commentators. These services download television listings for a two-week period over a modem, allowing the user to select and record programs far more easily than with a VCR. These programs can also be watched while they are still being recorded. TiVo makes it extremely easy to fast-forward through commercials. Thus, a program that usually runs from 9:00 to 10:00 P.M. with commercials can now effectively be watched starting at 9:20 P.M., or at any time thereafter, with virtually no commercial interruption.

Sophisticated computer users with high-speed Internet access can already swap all varieties of digitized compressed media – not just audio MP3 files but also video and feature-length movies, as well as episodes of popular television shows in MPEG and other formats. They can do so even after the demise of Napster, the free downloading service, through the use of peer-to-peer networks such as Kazaa. It remains to be seen whether legal challenges or the technological development and consumer acceptance of advanced copy-protection systems will extend the life of a commercial paradigm for marketing recorded media that has now endured for more than a century, since Edison's innovation of recorded music, remains to be seen.

In addition to the frequencies reserved for broadcast, portions of the radio spectrum have historically been set aside for "safety and special" radio services that facilitate point-to-point communication in applications where wired connections are impossible or impractical (Table Dg162–171). The technology of radio transmission, indeed, was originally developed as a means of extending the telegraph's and telephone's point-to-point capability over water. Separate bands remain allocated for marine and aviation services and for police and taxicab dispatch, where mobile communication has been of great utility. Amateur radio operators (hams) can obtain licenses to broadcast on certain frequencies, and a small band of forty channels is reserved for short-distance communication by households over citizens band, for which, after 1983, licenses were no longer required (Table Dg162–171).

Unless encrypted, two-way radio communication is inherently less private than that which generally takes place over wires. Sometime this can be an advantage, as in marine communication, where ships can keep their radios tuned to a certain frequency for general postings or distress calls and switch to another if an individual conversation is desired. And in police work, it can be helpful for patrol cars to have an overall sense of the activity and problems in their region. On the other hand, exchanges are accessible to anyone with access to a scanner; one is communicating over the wireless equivalent of a party line.

In the last three decades of the twentieth century, the most important innovations in the area of radio transmission have been in the development of cellular and paging services using previously and newly allocated portions of the spectrum. AT&T first introduced mobile radiotelephones in 1946, but they were expensive and inconvenient. One had to rely on an operator to connect with land lines over the public switched network, and conversations were simplex, only one direction at a time, with a button on the microphone for switching between send and receive. An improved system was introduced in 1964 and further improved in 1969, but the number of users in any metropolitan area was limited because the system relied on one powerful central antenna that monopolized channels over large geographic areas. Thus, aside from the safety and special services, individuals had limited access to mobile communications.

The last two groups of tables in this chapter examine the Postal Service and the broadcast print media, including books, magazines and newspapers, as well as the libraries that store them. The Postal Service, which plays an important role in delivering these media to households, also transmits narrowcast "junk mail," as well as asynchronous point-to-point written communications among and between businesses, government entities, and households. These sectors rely on technologies less fundamentally revolutionized in the nineteenth and twentieth centuries than those covered in this section.

Nevertheless, each medium faces questions about its future in an age in which the costs of storing and sending data have fallen precipitously. The market for printed encyclopedias has been radically transformed, and greatly reduced, by the capability of providing this information cheaply on CD-ROMs, DVD disks, or over the Web. Back issues of a growing range of academic journals are now available through vehicles such as JSTOR, with the print market protected by a blackout window on electronic availability covering the most recent three to five years. On the other hand, the content of current issues of most major newspapers and many magazines and periodicals (but not, in general, back issues) can now be accessed free of charge via the Internet. Finally, an increasing number of books in the public domain are available electronically and can be freely downloaded onto handheld or personal computers as a result of initiatives such as those of Project Gutenberg. The ways in which libraries operate are being altered by the increased use of electronic reference tools, and the future of the Postal Service continues to be debated. It is too early to say how changes in the technologies and economics of communications will ultimately influence the markets for and pricing of print media.

### Squeezing More out of Bandwidth

It is impossible to discuss telecommunications without encountering references to bandwidth. In analog communications, bandwidth refers to the difference between the highest and lowest frequency used by a given communications channel. Thus, because telephony transmits information within a subset of the audible spectrum (about 300 to 3,500 cycles per second, or hertz), the bandwidth of a voice channel is about 3,000 hertz. In digital communications, width is measured in bits per second. Higher bandwidth means speedier data transfer, but it does not actually mean that the individual units of data per se move more quickly. Electrons move over unshielded twisted pair (the standard telephone wiring used in homes, consisting of a pair of insulated copper wires twisted together to reduce crosstalk) at about two thirds the speed of light: 124,100 miles per second. This is approximately the same speed reached by photons over what is called a T3/DS3 fiber-optic cable, which provides the backbone of the Internet. By making the individual on/off signals of shorter duration, among other ways, these "pipes" can be made effectively wider, in the sense that they can handle a larger stream of bits (binary digits) per unit of time.

Whether in wired or wireless applications, communications channels are scarce, and much of the progress in this sector has involved attempts to use them more efficiently. Aside from tricks that speed up rates of "throughput," other economies have been achieved by reducing interference between channels (a major concern in wireless communications), by reducing the amount of time during which no data are effectively flowing through the pipe (through packet switching and channel reallocation), by increasing the information content of a given "bitstream" through data compression and through carrier-wave frequency division "multiplexing" (known as wavelength division multiplexing in fiber-optic applications), and by creating multiple channels in a medium that otherwise would accommodate only one. All of these strategies are discussed later in the essay.

Cellular telephone service represents a new implementation of established transmission technology made possible by advances in computing and multiplexing that facilitate automatic reallocation of channels as they become available, multiple simultaneous use of the same channel, and handoffs from one local cell to another. By lowering transmission power, and thereby allowing the same frequencies to be reused in nonadjacent cells, the coverage of a cellular system can be extended geographically without monopolizing channels over large areas.

The Advanced Mobile Phone (AMP) system is the analog system jointly developed by Motorola and AT&T and introduced in Chicago in 1983. In contrast to other two-way safety and special services, cell phone technology uses one channel for transmission, another for receiving. Advanced multiplexing and, in newer systems, digital compression technologies have vastly expanded the effective number of conversations that may be conducted simultaneously over a given allocation of spectrum. These techniques include frequency division multiple access (FDMA), standard in analog systems like the AMP, and time division multiple access (TDMA). In the former, each user gets a slice of spectrum in a particular region for the duration of a call or until handed off to a new cell, at which point the frequency may be reassigned to a new user. In TDMA, each channel itself is sliced into three very narrow time slots, each of which is used to transmit a separate communication. Thus, one channel can serve to transmit three conversations simultaneously through the use of sequences of short interleaved bursts of data that are then demultiplexed and appropriately recombined at their receiving ends. GSM (global system for mobile communications) uses a variant of TDMA with encryption and is the standard in Europe. American GSM phones, however, operate at a different frequency than their European counterparts.

Code division multiple access (CDMA) is the most complex, using split spectrum technology borrowed from the military. A conversation is digitized and then split apart, with different parts assigned to different frequencies and then reassembled at the receiving end. This requires more expensive switching equipment but makes the most efficient use of spectrum. All of these multiplexing technologies have their origins in nineteenth-century solutions to challenges for squeezing more data through given wires in telegraphy and telephony. Although cellular service relies on radio transmission, data on the cellular phone industry are grouped along with those on wired telephones. Table Dg103–109 includes data on the growing number of employees, subscribers, cell sites, revenue, and capital investment.

Pagers represent another interesting and distinctive use of radio spectrum. Invented in 1949, they were a minor feature of the environment in 1970 but have greatly expanded in use in the intervening years. All pagers use a similar technology. A ground-based tower broadcasts a stream of data on a single frequency. Receivers monitor this frequency for messages containing a special address, or "cap code." Upon receipt of such a code, the earliest machines simply beeped, alerting the wearer to "phone home" to retrieve a message. In the 1970s, tone voice pagers appeared, beeping and playing a short voice message. But requiring considerable bandwidth, they have been less successful than devices, introduced in the late 1980s, that receive and display a short alphanumeric message on a liquid crystal display.

In 2000, many mobile individuals carry both a pager and a cellular telephone. Each has advantages, and to some degree they complement each other. The range of reception and battery life of a pager are superior to those of a cellular phone, but one cannot return a call on it. As in so many other aspects of the communications sector, much is in flux, and some convergence in equipment and services is now occurring with the development of short message systems for cellular telephones.

Finally, in 1970 there were no global positioning system (GPS) satellites, whose navigational beacons are available in 2000 without charge, courtesy of the U.S. Department of Defense. By triangulation involving very minute differences in the time at which radio signals from four or more different satellites are received, this specialized form of information transfer provides individuals with precise information on their longitude, latitude, and elevation above sea level, as well as the correct time, all without the intervention of any human operators. Military and other authorized users can pinpoint location within 22 meters on the ground and elevation to within about 28 meters. A slightly degraded service enables civilians using handheld receivers/computers to fix location within 100 meters horizontally and elevation within 156 meters.

## Historical Perspective

The technological focus of the remainder of this essay is on systems of communication and on those sectors whose technologies have undergone the most radical transformation. Postal networks and the publication of books, newspapers, and other print media predate the nineteenth century, but virtually everything else – telegraphy, telephony, radio, television, and the Internet – were created after 1840 because they were dependent on advances in the generation, storage, and manipulation of electricity.

Better understanding of the past and present of the communications sector is also critical from an aggregate perspective. From the early 1970s through the mid-1990s, both labor and total factor productivity in the United States grew much more slowly than was true in the half century preceding. The limited total factor productivity (TFP) growth we have experienced has been highly concentrated in the telecommunications sector and the closely related manufacture of computers (Gordon 2000; Oliner and Sichel 2000). Productivity (including TFP) growth rates did move up in the late 1990s, but it is not clear in 2000 whether this acceleration will be sustained. The big question remains whether the investment of hundreds of billions of dollars in new computer and telecommunications equipment as intermediate inputs into the production of other goods and services has laid the groundwork for as-yet unrealized revolutions in practice that will boost labor productivity in consuming sectors above and beyond what one would otherwise expect from such capital deepening.

In 2000, the communications sector of the U.S. economy is in the throes of transformative change, the final impact of which remains difficult at this point definitively to discern. This is true both for industries that produce these services and for industries that use them. This essay offers the reader an explanation of what lies behind the historical series, an explanation that may facilitate attempts to peer ahead. It provides tools for thinking about the nature of convergences that will take place among communications modalities traditionally quite distinct, about what will change unrecognizably in the next decades, and about what will not. Some of the most recent developments mentioned here are not yet adequately chronicled in the accompanying historical series, but they almost certainly will be in future editons of this work.

### Postal Service

We begin with the oldest public networks facilitating point-to-point communications. Prior to the development of sophisticated optical and electromagnetic telegraphs in the nineteenth century, detailed messages could move only on written media, and at speeds over land no faster than the fastest horse and over water no faster than the fastest sailing ship. Public-switched postal networks did exist in the classical world and were nowhere more highly developed than in the Roman Empire. But in the Middle Ages, such systems, along with Roman roads, fell into desuetude. Religious, state, and commercial organizations maintained private communication lines, but public systems appeared again only in the seventeenth century. Britain, for example, established a government postal service open to the public in 1635.

In postal matters, as in many other respects, colonial America took its lead from England. British and, more generally, European practice was based on the principle that the post office should be a revenue-generating institution. Like import tariffs, revenues were intended not only to cover expenses but also to help defray costs of other government activities. In 1691, the Crown granted Thomas Neale a franchise for a North American Post Office. In 1707, the Crown reacquired the franchise from Neale's successors, and from then until 1775, colonial mail service, like colonial defense, was provided by a branch of the British government, in this case, the General Post Office in London. Some colonial actions, such as that of the Massachusetts Bay legislature in 1639 designating a particular Boston tavern as a mail drop, affected postal service, but most of the shots were effectively called in the mother country.

The prerevolutionary economy, even more than was the case in England, was oriented externally. Water has always been the preferred means of moving goods in economies with poor internal transport because its low friction enables transit with less expense of energy than is associated with haulage of goods or people over land. Most seventeenth- and eighteenth-century written communication, like most merchandise trade in those centuries, moved over water between England, continental Europe, the Caribbean, and the colonies or, for intercolonial communication, on ships pursuing the coastal trade. The remainder traveled on a limited network of generally north–south post roads.

In 1775, as the American Revolution began, Benjamin Franklin took over operation of the system, but for the next seventeen years, it continued within largely colonial parameters. In 1789, the country had seventy-five post offices, almost all concentrated in coastal port cities, and linked by 1,875 miles of post roads. Service was virtually absent from the hinterland, and no explicit subsidies encouraged the circulation of printed matter.

With independence came increased demands for internal improvements, not only in the movement of goods and people but also in transfers of cash and written and printed materials. Article 1, Section 8 of the U.S. Constitution gave the federal government a monopoly on the movement of mail and gave Congress the right to establish post offices and post roads.

The Post Office Act of 1792 established a foundation for a system that dominated American internal communication for half a century. It led to the construction of a network that, within a few decades, annually carried hundreds of thousands of newspapers and periodicals, millions of dollars of cash payments, and a more limited number of first-class letters. In doing so, it played an important and often underappreciated role in nation building and commercial development within the young republic.

The Act formalized the preexisting exchange privilege allowing printers to swap ownership of newspapers freely, and it provided heavily subsidized carriage of newspapers from printers to subscribing households. These subsidies were a boon to the news industry but came at the expense of merchants, particularly in the Northeast, who had to pay higher rates for first-class service. In the early nineteenth century, the costs of producing a newspaper dropped owing both to breakthroughs in the making of paper from wood pulp (the Fourdrinier process, the basis of most modern papermaking, was invented in 1803) and to progress in putting ink on paper, in particular, the steam-powered rotary press (1814). These complemented cheap carriage and extension of the postal service into the hinterland to create, in the first half of the nineteenth century, a national market for newspapers and the information they transmitted (John 1995, p. 37). As Table Dg253–266 shows, by 1850 there were more than 250 daily newspapers publishing in America.

This national market preceded by several decades the establishment of similar markets in other commodities. Not until the 1850s does one begin to see national distribution of hard and soft goods, and not until the 1880s, with the availability of the railroad, the telegraph, and the refrigerator car, could perishables be marketed across the nation. News, like fish, gets stale quickly, but unlike fish it is the one perishable that does not require refrigeration. Because of this feature and the fact that it had a high value in comparison to the weight of its medium, it could easily be transported in a post rider's saddlebag.

The success of the antebellum American postal system was an administrative achievement, not one enabled fundamentally by new technology. Relying for actual carriage on contracts with stagecoach operators and post riders, and railroads when they became operational in the 1830s, it consisted of a sophisticated hub-and-spoke distribution system operated by an extensive three-level bureaucracy. Top-level management operated out of Washington. Second-level managers ran regional sorting stations. At the third level were thousands of local part-time postmasters. In 1831, the Postal Service employed more than three out of four civilian federal workers, and its corps of 8,764 postmasters was larger than that of the federal army of 6,332 (John 1995, p. 3). Not until the 1870s, with the Pennsylvania Railroad, would a private business enterprise exceed the Postal Service in employment.

Positions of postmaster in the United States were valued because of their steady salaries and the prestige associated with federal service. In 1835, for example, after serving in the state militia, Abraham Lincoln secured appointment as postmaster of New Salem, Illinois. In the antebellum United States, the local post office was the principal, if not the only, manifestation of the reach of the federal government into most Americans' lives.

In the first half of the nineteenth century, the United States became the first and, to that point, only country to abandon the view of the Postal Service as a cash cow for defraying other government expenses. As Congress eagerly designated new routes and offices, revenues that might otherwise have been surplus were effectively plowed back into the system, resulting in a rapid extension of service to all corners of the hinterland. In 1828, the United States had seventy-four post offices for every 100,000 inhabitants, versus seventeen in Great Britain and four in France. Canadian postal service was so poor that merchants routed interprovincial communication through the U.S. system (John 1995, p. 5).

As the data suggest, offices were built ahead of demand in a way sometimes criticized as wasteful. The fact that Congress, rather than the Postmaster General, determined new routes opened the process to political jockeying among communities vying for access to communications links with the outside world. But in conjunction with the subsidy for newspaper carriage, the resulting network enabled white males, regardless of their location, to participate actively in the political life of the nation. The character and extent of the service surely contributed to the high participation rates among eligible voters characteristic of the first century of U.S. history. Alexis de Tocqueville arrived in America expecting to find an information gradient as he journeyed away from the coast. In this he was disappointed, finding the rural backwoodsman in Michigan as well informed as his fellow citizen in coastal cities.

The history of the Postal Service makes its imprint in a number of ways in *Historical Statistics of the United States*. Series Dg181 provides the longest continuous series in this chapter, that for the number of post offices in the United States (see also Figure Dg-A). Beginning with the 75 reported for 1789, the number grew steadily, with a brief hiccup during the Civil War, until 1901, when it peaked at 76,945. Over the next century, the number of offices fell steadily (the series does not include branches and contract offices). The United States now has approximately the same number of post

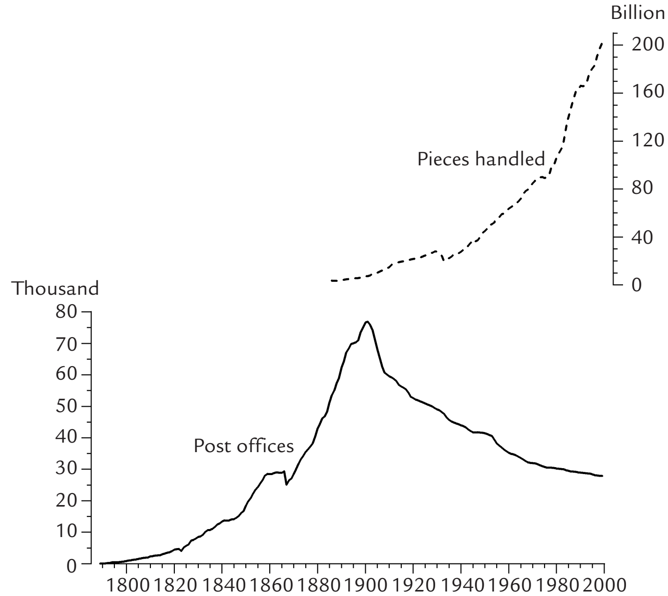


**FIGURE Dg-A Postal Service – post offices and pieces handled: 1789–1999**

***Sources***

Series Dg181 and Dg189.

offices it had during the Civil War, although, of course, they service a much larger population and land area.

Series Dg182–184, showing revenues, expenses, and their difference, demonstrate that although the system has sometimes earned surpluses, it has also incurred deficits and has not, over time, been a source of revenue for the general government. The American polity has not insisted on surpluses and has been willing to tolerate deficits in its postal service, but not on a persisting basis. The Postal Reorganization Act of 1970, triggered by deficits, transformed the Post Office – between 1829 and 1971 a cabinet-level department of the federal government – into a governmentally owned corporation, the U.S. Postal Service (USPS). Taking the longer view, however, the divisive political issue in the United States has been less the degree of subsidization the Postal Service should provide to other government operations and more the question of the nature of cross subsidizations among different classes of mail.

In the first half of the nineteenth century, the politics of subsidizing newspaper delivery were complex. On balance, the printing industry clearly enjoyed the subsidy, but Western printers resisted calls to make carriage entirely free, fearing that if this were done, they would be swept out of business by cheaper imports from the East. Low rates for newspapers, defrayed by high first-class rates, engendered considerable griping among merchants, ministers, and others who felt the need to correspond. Although one could prepay postage in the first half of the century, most postal charges were paid by the recipient, thus giving the service an incentive to see that mail was delivered. But many complained bitterly about their correspondents who sent them letters that were not worth the reading, let alone the postage due.

Although the federal government enjoyed a legal monopoly on mail delivery, it did not move aggressively to prohibit private competitors, with the exception of the heavily trafficked and lucrative northeastern routes. When private mail service threatened the government enterprise there, it defended its monopoly in court. This was paired with a dramatic lowering of first-class rates reflected in the tariffs of 1845 and 1851 (Table Dg209) (John 1995, p. 48). Prior to 1845, mailing a first-class letter was an expensive proposition, costing the equivalent of half a day's wages for a laboring man (John 1995, p. 159).

Postage stamps, first introduced in 1847, represented a major reduction in the transaction costs associated with prepayment. Mandatory prepayment through the use of stamps took effect in 1855. Series Dg186, showing the number of individual stamps sold, has its first entry for 1848, grows very rapidly following the Postal Act of 1851 and its further reduction in first-class rates, and then more than doubles again between 1854 and 1856 as a consequence of the introduction of the prepayment requirement. The 1855 change is also evident in the disappearance of the prepaid category in the 1855 tariffs; see Table Dg209.

The nominal charge for mailing a half-ounce letter anywhere in the continental United States remained unchanged at 3 cents from 1851 to 1958. The 1845 and 1851 acts, which reduced the cost of postage as a deterrent to first-class letter writing, also enabled merchants as well as newspaper publishers to take advantage of low rates. The consequence was the advent of junk mail. Prior to these rate changes, it was relatively inexpensive to make political mass mailings, provided they were classed as newspapers, as illustrated by the abolitionist attempt in 1835 to mail thousands of antislavery papers to Southern slaveholders. But widespread commercial narrowcasting through the mails was not cost effective until the new midcentury rate schedules took effect.

Series Dg188 chronicles sales of the once-ubiquitous stamped postal cards, beginning in 1873. Postal cards appeared first in Austria in 1869 and were initially used in the United States for commercial advertisement. But in 1898, a new law allowed short personal messages to be included on half of the back of a card. Postcards introduced a new style of communication to America: shorter, because the message had to fit on half of one side of the card, and somewhat less intimate, since the message could be read by others.

Although billions of stamps continue to be sold, stamped envelopes, first issued in 1873, represent a product whose time has come and gone (series Dg187). Metered postage, pioneered by the Pitney-Bowes company, first became available in 1920, and in the late 1990s, the USPS approved systems using personal computers and printers that provide similar service. Nevertheless, the failure of stamped envelopes and metered postage to achieve widespread consumer acceptance suggests that the convenience of postage stamps, a nineteenth-century innovation, will likely assure a continued market for them into the future.

The structure of the U.S. postal system bears many similarities to a telecommunications network. The local post office can be thought of as analogous to a local exchange. The system is a switched network because, as is true in the phone system, routes do not directly connect all possible customers. Local depots segregate long distance mail and that remaining within the exchange. Long distance mail is sent, usually by truck, to a limited number of centralized switching (sorting) stations, where it may be sent on to another regional sorting station or to its final destination, usually through contract with commercial air carriers. For local delivery, long distance mail is blended with interexchange communication and distributed over the counter, through post office boxes, or, eventually, through local delivery service. Similarities between postal and telecommunications networks arise because they must address somewhat similar logistical issues and because, in their formative years, both telegraph and telephone systems drew heavily for experienced top management from the Post Office.

Whereas telephone communication occurs synchronously, mail interchange does not. Thus, although reducing delivery time is, for obvious reasons, not a pressing issue in telephony, it has for most of the past two centuries been a central and distinguishing priority of the U.S. postal system. A major innovation in the second half of the nineteenth century, beginning in 1862, was the use of railway cars as mobile sorting rooms. Prior to assuming the presidency of the Bell System for the first of his two terms, Theodore Vail directed this aspect of postal operation. With the growth of the road network in the twentieth century and the reduction in rail service, sorting stations on large trucks performed analogous functions between 1941 and 1969. In 1977, the last railroad post office, between New York and Washington, D.C., ceased operation. Continued reductions in the costs of air transport have returned sorting operations exclusively to stationary offices, as was true when the system began. The first cross-continent airmail moved in 1920; regular service began in 1924, and in 1929, data from this service begin to appear in our statistical tables (series Dg202–204). In 1977, airmail disappeared as a separate rate category for domestic service.

Table Dg190–208 provides a more detailed breakdown of revenues, expenses, and volume by class of mail. More recent data are from a mimeographed publication known from 1969 to 1983 as "Revenue and Cost Analysis" and from 1984 until the present as "Cost and Revenue Analysis." The data on employment can be compared with interest to that in other sectors of communications: In the 1990s, the number of direct employees in the Postal Service was in the same range as the number of employees in the telephone sector (see series Dg3 and Dg208).

The 1970 Postal Reorganization Act gave the system the right to raise its own capital for modernization, and the USPS received its last direct subsidy from Congress in 1982. The system continues to struggle to reduce the costs of routing heterogeneous physical media, principally through an aggressive automation program and incentives for large mailers to barcode and sort materials before drop-off. ZIP codes, introduced in 1963 (ZIP stands for Zone Improvement Plan), and "ZIP code plus four," introduced in 1983, combined with new automatic equipment, have been tools in the effort.

Sorting begins with manual segregation, in which packages, flats, and overseas envelopes are separated from envelopes that can be processed by machine. The latter are aligned (faced) by sensing machines that search for stamps and flip or rotate letters in preparation for automatic cancellation. For correspondence lacking bar codes, multiline optical character readers (MLOCRs) translate addresses into sprayed bar codes in order to facilitate subsequent sorting by automatic equipment. The post office is perceived as low tech, but character recognition software has now improved sufficiently that it can automatically barcode over 30 percent of handwritten letters. Where the machine cannot decipher a handwritten address, the MLOCR signals a remote location where, viewing the letter via video camera, an employee enters the correct addressing information, sent back electronically for bar coding.[1] Sorters read bar codes (evident, upon examination, on most mail received) to separate long distance from local traffic and to sort material into

[1] See the Internet site of the U.S. Postal Service for more information.

correct order for local delivery. At the end of the twentieth century, only 8 percent of letter mail received daily was processed manually, but it consumed half of all the labor costs associated with mail processing.

The postal system is distinguished within our communications sector by the degree to which the implementation of labor-saving technological change dominates the agenda. Automation remains key to cost control and an arena of both progress and challenge. Labor-saving technology aims ultimately to eliminate manual sorting, as telephone switches have eliminated manual switchers (telephone operators). How far the USPS will succeed in this effort without imposing greater uniformity on media and addresses remains uncertain.

Major increases in throughput speed, which, as we have seen, can effectively widen a given physical communication channel and thus save capital, have been much less a priority than in earlier decades, and in some respects, there has been retrogression. The situation with respect to capital costs in particular will worsen if the growth of electronic mail begins seriously to cut into printed correspondence. Series Dg189 suggests that throughout the 1990s, the growth of the Internet did not reduce our propensity to send written or printed communications any more than it has yet reduced the demand for paper in offices (see also Figure Dg-A). As this is written, however, there is evidence that the demand for first-class mail service may be declining, creating an incipient fiscal crisis for the Postal Service. Other uncertainties surround how the USPS express mail and package delivery service will fare in continued competition with private carriers, such as United Parcel Service and Federal Express.

### Newspapers and Books

In contrast with the industries covered in the first two groups of tables of this chapter, those in the latter two groups employ paper as a medium for storing information, and none has depended fundamentally on advances in electrical engineering. This is true for newspapers and books as well as the Postal Service. These services and media were familiar to our Founding Fathers in a way none of those chronicled in the first two groups of tables would have been.

This is not to say that for two centuries, communication based on moving paper has been technologically stagnant. It is simply that the changes here have been less revolutionary. In the newspaper industry, reductions in the cost of making paper from wood pulp continued throughout the nineteenth century, with particular advance associated with improvements in mechanical grinding and the use of sulfites in the late 1860s. The Hoe press, an advanced steam-powered rotary press that printed on both sides of a web or roll of paper, was introduced in 1867 (Lacy 1996, p. 65). The late 1880s saw the introduction of the Mergenthaler linotype and Lanson monotype machines, which dramatically reduced the cost of typesetting. Together, these defined a technological paradigm for newspaper production that endured for a hundred years, until the introduction of computer-based typesetting in the 1980s.

As noted, the first daily newspaper in the United States began operation in 1783. By 1800, 24 papers published daily, with a combined circulation of about 50,000 (Lacy 1996, p. 65). These numbers can be compared with the first information from census data, for 1850, which shows 254 dailies with circulation of 750,000, with circulation increasing to 3.5 million in 1880 and over 15 million in 1900 (series Dg256; see also Figure Dg-B). Newsprint dropped in price from about 12 cents a pound in the 1880s to 3 cents a pound by 1900, and a daily paper at the date cost only 1 or 2 cents, barely covering the cost of paper. Its distribution was enabled by that burgeoning role of advertising.

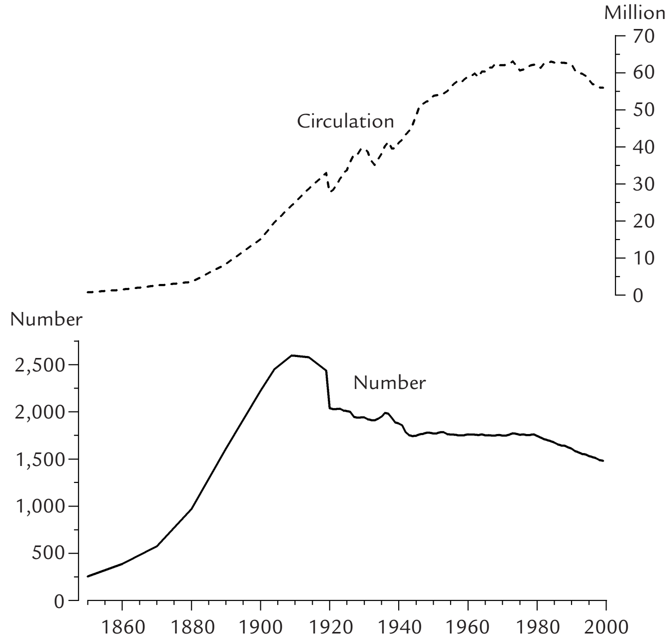


**FIGURE Dg-B Daily newspapers – number and circulation: 1850–1999**

***Sources***

1850–1919: series Dg255–256. 1920–1999: series Dg267–268.

Local newspaper production was further facilitated by postal legislation of 1879 that allowed free distribution within the county in which the newspapers were produced. The Act also provided, for the first time, a low national rate for shipping magazines, enabling periodicals to exploit the expanding rail network and develop a national audience. National magazines helped unify the country culturally and politically in a way analogous to the distribution of newspapers by subsidized carriage in the first half of the century (Lacy 1996, p. 72). The number of magazines in the United States, about 40 in 1800, grew to more than 5,500 by 1900. Series Dg280 picks up the story in 1935, showing at that date 6,546 periodicals.

Many of the technical changes affecting newspaper publishing have, along with improved binding technology, influenced the mechanics of book production. Table Dg225–252 provides data on the total number of books published and, after 1950, their topical distribution. Series Dg225, which includes both new books and new editions, confirms that reports of the demise of the printed book are as yet premature. Note that these totals do not include mass market paperbacks or imports, which are reported, respectively, in series Dg251–252. Thus, to calculate, for example, the share of new fiction books among the total new books published (excluding mass market paperbacks and new editions), one would divide series Dg233 by series Dg227, yielding about 8.5 percent for 1996.

It is remarkable that the publication of new books in the area of economics and sociology has grown so rapidly that it now comprises the largest single category here enumerated, more than 16.7 percent of the total in 1996, up from 11.8 percent in 1982, the first year for which comparable data are available (series Dg247). Perhaps a fifth of book industry sales derive from general fiction and nonfiction, what are called trade books. The remainder, and the vast bulk of the business, consists of textbooks and business, scientific, and technical publications.

### Telegraph

So long as communication beyond the reach of face-to-face interchange was limited to the physical movement of written or printed media, its speed of transmission had binding constraints. With the possible exception of transport by carrier pigeons, it could not move faster than could people or goods. Most early efforts to break out of these barriers involved visual or optical telegraphy; for example, smoke signals mirrors to reflect the sun's light, or light beacons used at night. These primitive telegraphic technologies were the equivalent of narrow pipes: even under good conditions, you could not push much data through them. Developments in both hardware and software were necessary before visual telegraphy could make a major leap forward at the end of the eighteenth century.

Advances in optics in the seventeenth century had made it practical to produce relatively inexpensive telescopes, which significantly expanded the potential reach of semaphoric systems. The most successful of these systems was developed in France by Claude Chappe, with viable long distance service beginning in 1793. Chappe positioned stations at high points on mountains, tall buildings such as church belfries, or specially constructed towers. At each station, two men pointed telescopes in opposite directions. Operators composed signals on a Z-shaped semaphore, with each of the three segments independently rotatable throughout a 360-degree arc by means of hand controls located below. For signaling purposes, each arm could appear in one of eight orientations 45 degrees apart. The system used a basic transmission code of ninety-eight elements, meaning that each signal communicated about 6.5 ($\log_2 98$) binary digits, or bits of information. Carefully worked-out codebooks used these signals to maximum effect, making it possible in good weather, in conjunction with the hardware, to send a detailed message from Paris to Marseilles in about an hour and a half.

Napoleon used Chappe technology to extend his communications reach into French-controlled territory in the Low Countries, Germany, and Italy. But it was after peace in 1815 that the system reached its fullest development. At its peak, a network of 5,000 kilometers, with most of the lines radiating from Paris, made France the world leader in advanced communications technology (Field 1994). The Chappe system was not the only optical telegraph system, but it was far and away the most successful. During the Napoleonic wars, England employed the six-shutter Murray system to link London and Portsmouth, but the line was disbanded with peace. In the United States, semaphores operating over shorter distances were common in commercial cities like New York and San Francisco, where they announced the impending arrival of merchant ships. Their legacy is evident in place-names such as Telegraph Hill.

In the first part of the nineteenth century, Chappe technology was the acknowledged leader in long distance transmission, and its suitability for a coastal line linking New York to New Orleans was discussed in the United States as early as 1807. The project resurfaced periodically for several decades, with the most serious proposal for funding, endorsed by Andrew Jackson's Postmaster General Amos Kendall, put before Congress in 1836.

This proposal was vigorously opposed by Samuel F. B. Morse, a professor of painting and sculpture from New York. Morse had recently patented a new telegraphic technology exploiting electromagnetic principles, and he wanted Congress to support it instead. Although a practical and proven technology, Chappe telegraphs did have serious limitations. In addition to narrow bandwidth, they were always at the mercy of the weather (including fog and heat inversions) and never operated successfully at night. The advent of railroads offered some improvement in the speed at which information could reliably move independently of the weather or time of day, but far less than what people were prepared to pay for, especially in the markets for news and financial data.

Seventeenth-century scientific advance had also led to improved understanding of electricity and magnetism, providing foundations for a different solution to the problem of long distance communication, one that relied on fast movement of electrons along wires, as opposed to photons between eyeballs. Absent early nineteenth-century progress in producing and storing low-voltage high-amperage electricity, however, proposed implementations of electric telegraphy could not work over any but the shortest distances.

Prior to the beginning of the nineteenth century, scientists could store electrons by accumulating static electricity in a Leyden jar, a primitive capacitor. Such a power source appeared to be promising as a basis for long distance telegraphy. In 1753, Charles Morrison proposed running twenty-six wires from source to destination. In front of each he would hang a piece of paper suspended from a string, one corresponding to each letter of the alphabet. As he sent static electric charges down particular wires, they would cause displacement of the appropriate paper scrap (pith balls, in later proposals). Similar explorations used electric charges sent over wires to generate gas bubbles through chemical reactions in tubes of liquid. The main problem with all these proposals was reliance on static electricity: high voltage, low amperage, and subject to severe attenuation in signal strength, particularly in the presence of atmospheric disturbances.

Morse's ability to develop a commercially successful telegraph was predicated on Alessandro Volta's advances in battery technology (1800), which made available, as an alternative to static electricity, a source of lower-voltage, higher-amperage current whose signal strength was more reliable over distance. Morse also benefited from Hans Christian Ørsted's discovery in 1820 that electric current would deflect a magnetized needle. Charles Wheatstone's telegraph, developed in England and patented in 1837, exploited this discovery directly. Morse's device required the utilization of this knowledge to make an electromagnet. Building on Ørsted's work, Michael Faraday and Joseph Henry, of whom more later, provided Morse with this component.

Having contributed to the demise of the proposal for a U.S. long distance visual telegraph, Morse finally wrangled $30,000 from Congress in 1843 to build a demonstration line from Baltimore to Washington using his technology. Even though visual telegraphy was a well-understood technology, experimentation in electric telegraphy had about it a questionable penumbra owing to its attraction of a large number of charlatans and quacks. The bill providing for Morse's subvention passed in the House by the small margin of eighty-nine to eighty-three, with seventy abstentions, in a debate marked by sarcastic references to, and half serious proposals to fund, experiments in mesmerism (Standage 1998, p. 46).

The apparatus in this line used Joseph Henry's electromagnet, when activated by a transmitted electric signal, to move an armature holding a pencil, thus intermittently marking a paper-covered drum rotated by a clock mechanism. On May 24, 1844, Morse transmitted his first message, "What hath God wrought?" This question, which might better have been posed when the telephone was invented thirty-two years later, was rhetorical. The American public understood immediately what Morse had wrought, and to what uses it could be put, even if the initial financial success of the Washington-to-Baltimore line was disappointing (this was one consideration leading the U.S. government to spurn Morse's offer to sell the system to it).

Nevertheless, even before Morse's successful demonstration, telegraphs (irrespective of their mode or prospective mode of operation) were considered hot technologies in the United States. This is reflected not only in the early interest in importing Chappe systems but also in the forty-five newspapers that had by 1820 incorporated the word in their titles (John 1995, p. 87). Morse's achievement thus took place within a receptive political and commercial environment.

The next two decades were a period of rapid and at times chaotic growth of the telegraph industry under private ownership (Thompson 1947). The U.S. telegraph system grew rapidly and alongside of the railroad, with which it often shared rights-of-way and to which it provided logistical support. In this period, systems were set up based not only on the Morse patents but also on those of Alexander Bain, Royal E. House, and David E. Hughes. Some conservative investors stayed away because of fears that the government might change its mind again and step in to build or acquire its own system. Many of the fledgling firms faced difficulties raising cash, and thus acquired assets of other companies using stock, laying the foundation for later complaints about "watering." After a decade of growth involving fragmented systems using sometimes incompatible technology, consolidation began. The New York and Mississippi Valley Printing Telegraph Company brought together a number of smaller systems under its leadership, emerging in 1856 as the Western Union Telegraph Company, data from which appear prominently in our tables. Eventually, in a process culminating in 1866, shakeout occurred with consolidation under one firm.

In 1861, eight years prior to the driving of the golden stake at Promontory Point, the company linked San Francisco by telegraph to the East, receiving $400,000 from Congress and subventions from the California legislature to encourage the investment (Goldin 1947). And after a three-month abortive experiment with a poorly insulated cable in 1858, the United States and Europe were permanently linked by telegraph in 1866. Transmission speeds were initially low, limited to 2 words per minute (wpm), but rose to 20 wpm in 1870 with an improved receiving mechanism.

Congress watched the consolidation of the telegraph industry with growing unease. Whether or not Western Union was a *natural* monopoly, there is little question that it exercised monopoly power over national telegraphic traffic for at least two decades. This period extends roughly from 1866 to 1885, when the Postal Telegraph Company (subsequently part of the International Telephone and Telegraph Company, or ITT)[2] entered as a competitor, and the telephone arguably began to serve as a competing technology within some heavily trafficked intercity routes, particularly the Boston–New York–Philadelphia corridor.

Between 1870 and 1896, at least twelve proposals emerged from congressional committees for government participation in the sector. The U.S. Postal Service run by the federal government, as well as European systems in which both post and telegraph systems were government owned, provided examples that fueled political pressure for nationalization, for government entry as a competitor, or for government subsidy of a competitor.

Complaints levied against Western Union echoed many of those raised against railroads at the time. High rates, of course, figured heavily. Average message rates in the United States fell from $1.09 in 1867 to $ .30 in 1898, although the economy-wide deflation over this period would have accounted for about half of this decline. Whatever one's views of the legitimacy of Western Union's rate structure, however, there is consensus that as a result of it, much more so than was true in Europe, the telegraph in the United States remained predominantly an instrument of commercial rather than individual household use.

Other complaints included poor service (errors of transmission, delayed messages, and violation of secrecy); the ability of stock speculators to obtain preferential access to the telegraph; free telegraph privileges granted to public officials to influence their votes; discrimination among regions and communities with respect to rates; long-term contracts with press agencies, such as the Associated Press, and with hotels and railroads that created barriers to entry for potential new entrants; and unfairness to workers (long hours, low wages, poor conditions) and to stockholders (stock watering). Perhaps the most damning complaint was the lack of technological progress, particularly in comparison with some of the European systems.

In addition to better insulation for undersea cables, advances in telegraphy in the second half of the nineteenth century included new input and output devices, as well as innovations that saved capital by increasing the rate of data flow. Morse's initial line was a single uninsulated wire, with the return part of the circuit through the ground, an implementation that remained standard throughout much of the remainder of the century. Underwater, however, cables had to be insulated; water (particularly salt water), unlike air, conducts electricity relatively easily. Gutta-percha, a nonstretch latex covering perfected in 1847, reliably insulated the first England-to-France telegraph cable in 1851. Gutta-percha and rubber remained the preferred insulation for undersea cables until 1947, when they were replaced by polyethylene sheathing.

In 1856, Morse replaced the marking device with a sounder, with telegraph operators decoding the sequences of sounds directly into letters, a system that unexpectedly speeded up the service. With the introduction of the sounder, U.S. telegraphic technology completed a framework that would remain in place throughout the remainder of the century, although there was some technological advance in peripheral equipment. Thomas Edison, who began his career in telegraphy, made his first major contributions with an improved stock ticker, a specialized printing telegraph used for broadcast that formed half of the telecommunications backbone of a national stock-trading technology that endured for the better part of a century, from the 1870s to the late 1960s (Field 1998). The other half of the system consisted of leased wires from brokerage houses across the country to New York for the transmission of orders and confirmations of execution.

[2] Postal Telegraph got its name because of repeated government proposals to strike a deal offering space in post offices for a private telegraph system competitive with Western Union. The deal was never consummated, but Postal Telegraph liked the name, even though the firm was as private an entity as Western Union.

Elsewhere I have described how the telegraph was used to implement capital-saving innovation. In customer industries, it permitted, for example, single-tracked railroad systems, higher rates of fixed capital utilization in manufacturing, and faster rates of inventory turn in wholesale and retail distribution (Field 1992). Capital-saving innovation was also critical in telegraphy.

For a quarter century, all telegraphic communication in the United States was simplex. The default position for a telegraph key was in closed position. Opening the key broke the circuit, alerting stations along a line to an imminent transmission. To interrupt an incoming message for an urgent outgoing transmission, a telegrapher opened his key. All sounders along a line then went dead, and the original transmitter would cease hearing his keystrokes echoed in his sounder. Ceasing transmission, he placed his key in closed position and awaited the incoming, presumably urgent, message. Only one message at a time could be transmitted.

Duplex technology, which enabled messages to be sent in both directions along one wire, originated in Germany and was improved upon in the United States by J. B. Stearns in 1871. In 1874, Thomas Edison patented a quadruplex system that enabled *two* messages to be sent in both directions at the same time.

In the area of multiplexing, considerable credit must also be given to the French engineer Jean-Maurice-Emile Baudot. Baudot brought us into the modern information age not only with a sophisticated time-division multiplexing technology in 1874 but also by developing the first true digital code. The Baudot multiplexing system consisted of a stationary face plate covered with five concentric copper rings, with the plate divided into six sectors, each of which was assigned to an individual user. At any moment of time, each sector of each ring either did or did not have an electric potential applied to it. A rotating armature with electric brushes swept the rings, producing bursts of on or off sequences, which were then demultiplexed at the receiving end by a synchronized device that sent signals to six separate receivers. Baudot's work on time-division multiplexing enabled 90 wpm transmission, and it laid foundations for the engineering triumphs of the late twentieth century that have permitted the rapid development of our ability to expand bandwidth by moving more information through channels in given periods of time. Baudot's work also facilitated the development of the teleprinter and teletypewriter in the twentieth century.

In visual as well as electromagnetic telegraphy, software was as important as hardware in expanding the utility of the systems. Morse code expanded in 1851 to include diacriticals (such as the *accent aigu* or *grave* in French) and became International Morse Code, which survives to this day. Although telegraphic transmission involves sequences of making and breaking a circuit, International Morse is not a true digital code because it involves five rather than two fundamental signals. The dot is the basic unit, with dashes defined as three times the duration of a dot. Dot-length pauses separate elements of a character code, triple-dot pauses separate characters, and seven-dot pauses separate words. Morse code requires skilled telegraph operators acting as coders and decoders. Because of the different lengths of dots and dashes and the different lengths of pauses, automated coding and decoding is cumbersome.

The five-bit Baudot code, patented in 1874, is the first true digital code because it relies almost entirely on sequences of two signals: on or off. With five bits, corresponding initially to the five concentric rings of his multiplexer, thirty-two ($2^5$) different symbols can easily be transmitted. To expand flexibility of his code, Baudot also pioneered by making two of them escape sequences, signaling the shift to a different symbol set, such as numbers. Baudot's multiplexing equipment, subsequent advances such as Donald Murray's punch-tape stored-message transmission system (1903), and teletypewriters all used it or a modified version.

Morse's first devices printed code, and the cumbersome House machine, with which they initially competed, could print text directly. With the switch to sounder technology in 1856, however, the United States became conservative compared with Europe in its adoption of printing telegraphs (the exception being the specialized stock ticker). As early as 1855, for example, the American Hughes had invented a practical printing telegraph using a piano-style keyboard. These were used extensively in Europe, but not in the United States.

It was not until well into the twentieth century in the United States that teleprinter technology began to spread beyond its application in the area of finance. In 1915, the Associated Press began switching its sounders to teleprinters for receiving broadcasts "over the wire." The five-bit Baudot code, slightly modified, drove teletypewriters that could both send and receive keyboard or paper-tape input, which could be outputted along with print. Linked teletypewriters in corporate networks with administrative message switching provided the first crude electronic mail service. Paper tapes were used to store and forward messages on switched systems; these networks were known as torn tape systems. Beginning in the 1930s, the Telex and TWX (teletypewriter exchange) services made two-way point-to-point teletypewriter service available to individual businesses (Table Dg22–33).

Teletypewriter service was commercially successful for about fifty years, from the 1930s into the 1980s. TWX was introduced by AT&T in 1932; Telex began in Europe. Western Union brought Telex service to the United States in 1962 and bought TWX from AT&T in 1970 (the two systems could not talk directly to each other). After 1980, the technologies became virtually moribund, at least in the United States, replaced in their functions by fax (facsimile) and email, but their half-century heyday represented an important transitional step in the development of more sophisticated and flexible data transmission equipment, protocols, and peripherals. In 2000, everyone with a networked computer has the equivalent of a teletypewriter on the desktop, but its use has required an expanded basic symbol set, both to control output devices and, for example, to enable the use of both upper- and lowercase letters.

ASCII, the American Standard Code for Information Interchange, was first defined by the American National Standards Institute in 1968 (standard X3.4), and can be thought of as a direct descendant of Baudot with a larger number of basic symbols. Its original implementation was as a seven- rather than a five-bit code, thus enabling 128 ($2^7$) distinct symbols. With escape codes for shifting into different character sets, ASCII has subsequently been made even more flexible.

Critical to the ability of networks to handle increased quantities of digital data at faster speeds have been innovations in error-checking technology. An important example is the invention of the parity bit in 1948 by Richard Hamming of Bell Labs. In transmitting seven-bit ASCII code, for example, an eighth bit can be added such that the sum of all eight is either even or odd (one or the other convention must be selected). If even parity has been chosen, then a sequence of eight bits with an odd number of 1s indicates an error, and a request can automatically be made to retransmit the byte.

Data-compression algorithms, along with error-checking technology, have been critical in increasing the ability of given data "pipes" rapidly and reliably to move large quantities of information. Data compression, for example, makes it much easier to send large quantities of data, particularly graphics, over phone wires. A one-page fax transmission, for example, must provide data on more than 1.5 million pixels. Much of a printed page, however, consists of white space. By using run-time codes, one transmits the number of sequential white spaces, instead of using a code for each individual space. Transmission times are dramatically reduced, and the same physical infrastructure is capable of much larger flows of data. Both error-detection and data-compression software are capital-saving techniques: they enable more data to be reliably pushed through a given pipe in a given period of time.

The tables presented here encompass the rapid rise, maturation, and eclipse of telegraph technology, the first true electronic transmission medium. Railroad and telegraph service developed symbiotically in the nineteenth century. But the nature of the demands for both changed in the twentieth, and with competing technologies, in particular the telephone, the economic rationales for joint production faded. Series Dg20 identifies the string of losses suffered by Western Union beginning in 1932 and extending through 1952.

In 1949, the company ended its long-standing agreements with railroads, giving up all claims to ownership of poles, wire, or equipment. But whereas railroads survived competition from automobiles, trucks, and airplanes by specializing in the movement of bulk freight, Western Union was never quite able to reinvent itself or find a sustainable market niche. It limped along for another four decades, trying a number of new business plans, including specialization in the now moribund teletypewriter technology.

Peak message traffic is recorded for the last year of the Second World War, but series Dg8, Dg12, and Dg14 chart the subsequent slide in messages handled, number of telegraph offices, and employees thereafter. Figure Dg-C paints a similar picture. In 1988, the Western Union Telegraph company sold off its international private-line business to a Swiss company, its Westar satellite to GM Hughes Electronics, and its business services group (for example, teletypewriters) to AT&T. Subsequently the Western Union corporation focused almost exclusively on money transfers, its original businesses eclipsed by rapidly changing technologies. But although the telegraph industry is effectively dead, its spirit lives on in the digitized information flows used for long-distance telephony and in the packet-switched Internet.

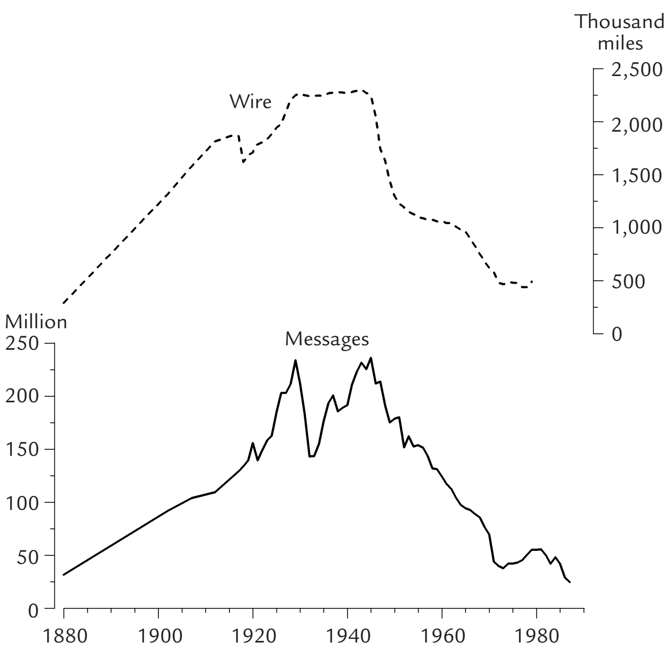


**FIGURE Dg-C Domestic telegraph industry – messages handled and miles of wire: 1880–1987**

***Sources***
Series Dg8 and Dg10.

## Telephone

The telegraph made it possible to transmit signals at close to the speed of light. In that sense it, rather than the telephone, marks the entry into the modern communications age. The main use of the telegraph was for long distance communication, a market in which, throughout the first decades of the twentieth century, the telephone effectively offered no competition. The telephone's initial success lay in satisfying a demand, then only latent, for convenient local communication. Because of the different nature of the service it provided, and because its use did not require the assistance of a skilled operator, telephone stations were installed in individual households and businesses, a rare event in telegraphy. (A telephone station is an installation capable of receiving and transmitting voice. Thus, a connected handset is a telephone station, and charges for calls without operator assistance are referred to as station-to-station rates.) As a consequence, the telephone gave rise to a range of advances in areas, such as switching and signaling, that eventually revolutionized and led to the convergence of technologies for transmitting data and voice.

Both Elisha Gray and Alexander Graham Bell were searching in the 1870s for capital-saving procedures that could move more than one message simultaneously over a single wire. Their experiments with harmonic telegraphs represented early attempts at frequency-division multiplexing. The idea underlying a harmonic telegraph was to use different tones to transmit different message streams at the same time, demultiplexing at the receiving end. The idea was a good one, although it could not then be practically implemented, and it was in multiplexing that astute investors saw the potential for big returns. Bell's financial backer and future father-in-law, George Hubbard, counseled him against working on a machine for voice transmission, "which would never be more than a scientific toy," and urged him to continue research on the harmonic telegraph, which, if successful, would "make him a millionaire" (Casson 1910, p. 25).

Bell's patent application described an improvement in telegraphy, but in fact the electronic foundations of telephony were different. The essence of Bell's innovation was the principle of transmitting voice over a *continuously completed* circuit, with fluctuations in sound waves generating corresponding fluctuations in direct current as a result of changing resistance in a medium that was produced as speech caused a diaphragm to vibrate.

Gray, a distinguished engineer whose telegraphic equipment firm Gray and Barton became Western Electric when brought under the control of Western Union, experimented with metal reeds tuned to specific frequencies that, when vibrated, induced fluctuating current in coils. The current, when transmitted to a similarly designed receiver, could make similarly tuned reeds vibrate. In 1874, he showed that a steel diaphragm placed in front of an electromagnetic coil (a forerunner of a loudspeaker) could reproduce any of the tones, but he was unable at the time to figure out a way to design a similarly flexible transmitter.

Bell also experimented with musical or multiple telegraphs. His breakthrough, which led to a device for transmitting speech, was in conceiving of the use of a membrane for transmission as well as reception, and it was on this basis that he filed a patent application for a telegraph and telephone system on February 14, 1876, just hours before Gray filed his intention (caveat) to patent similar devices.

Gray's caveat described a transmitter that attached a membrane to a metal rod inserted in a metal cup containing acid. Electric current from a battery flowed between the cup and rod, and variations in the distance separating them, as the diaphragm vibrated in the presence of speech, caused variations in resistance and, consequently, current flow. Bell used a similar mechanism three weeks later, when, spilling some acid on himself, he asked his assistant, Thomas Watson, to come to his assistance. March 10, 1876, marks the first successful telephonic voice transmission.

Bell took a selection of his apparatus to the centennial exposition in Philadelphia in the same year. Exhibited in a back corner amidst electric light bulbs, Gray's musical telegraph, and 60-wpm printing telegraphs from Western Union, Bell's creation, at first ignored, eventually became the hit of the exposition, admired by such pioneers as Joseph Henry and William Thomson – later Lord Kelvin – the chief engineer of the Atlantic cable. The public demonstrations in Philadelphia, which dispensed with both beaker and battery, showed a diaphragm attached to a magnet within an electric coil (magneto) that, in the presence of speech, induced current flows used to reproduce the sounds at the receiving end.

In spite of the favorable reception from scientific notables in Philadelphia, the telephone's reception by the public can charitably be described as tepid. Indeed, the press heaped ridicule on the scientific toy or, in some cases, branded it an outright hoax. Bell persevered with further tests, with simplex communication demonstrated over distance in Ontario in August 1876, and the first duplex (two-way) communication in October. May 1877 saw the first commercial application of the system, for a burglar alarm business in Boston. Bell's advertisements at this time stressed three advantages over telegraphy: (1) no skilled operator was required; (2) one could transmit 100–200 wpm versus 15–20 wpm using a Morse sounder; and (3) it required no battery. Bell's early telephones all used magnetos for transmission and did not require battery power until 1882.

With patents as its only capital, Bell, Hubbard, and Thomas Sanders organized the first Bell telephone company, giving themselves each 30 percent of the shares and Watson 10 percent. As Herbert Casson put it, "The four men had at this time an absolute monopoly of the telephone business, and everybody else was quite willing that they should have it." In the fall of 1876, Bell's financial backers, fearful about the commercial potential of their device and their ability to defend the patent, offered to sell their rights to Western Union for $100,000. The chief executive officer of Western Union turned them down: "What use," he said famously, "could this company make of an electrical toy?" (Casson 1910, p. 59). Western Union did not foresee the demand for communication over local networks, even though some telegraphic implementations in New York were evolving in that direction.

Western Union changed course rapidly when employees at one of its subsidiaries, the Gold and Stock Exchange Company, ripped out some of their printing telegraphs and replaced them with telephones. Western Union responded as might the software giant from Redmond, Washington, quickly organizing the American Speaking Telephone Company, capitalized at $300,000, with three well-known and competent engineers, including Edison and Gray, as technical staff. Western Union ignored the Bell patent, assuming that it would triumph in court on the basis of Gray's caveat, and introduced, on its own, important technical advances, in particular, Edison's carbon transmitter.

Edison had exploited an unusual feature of carbon: its variability to conduct electricity depending on how much it is compressed. He situated granularized anthracite coal between two electrodes covered with a steel diaphragm. Vibrations in the diaphragm compressed the granules, moving the electrodes closer together, and changing the resistance to the electric current that passed between them. The component was easy to manufacture, inexpensive, and long lasting, and it remained the standard in telephone handpieces until supplanted in the 1980s by electret devices. The principle remains the same, although in these components, a metal-coated plastic diaphragm creates fluctuations in an electric field across an air gap between diaphragm and electrode, causing a varying electric current to move down the wires. Transistor amplifiers are needed to strengthen the signal variations with this technology.

Telephony requires not only the arrangement of a voice circuit between sending and receiving stations but also a means of signaling: ringing the called party or ringing the operator in a switched system. The solution was to use bursts of higher-voltage current to trigger ringers. Initially ringer current was generated by hand-cranked magnetos, but with the switch to central-office batteries, the current was supplied at the local exchange. Dial telephones used pulses of this higher-voltage current to trigger the movement of automatic electromechanical switching equipment.

The extent to which managers in the nineteenth century moved from and between the U.S. Postal Service, the telegraph industry, and the telephone industry is notable. In 1845, Morse hired Amos Kendall, Jackson's Postmaster General, to help him set up commercial telegraph service. Theodore Vail, whose cousin Alfred also worked closely with Morse, headed the government's railway postal service, and was responsible for introducing the system whereby mail was sorted in moving railway cars. In 1877, he began the first of two highly successful but nonconsecutive terms as president of Bell. Western Union's aggressive move into telephony actually rejuvenated the Bell company by legitimating a technology that many still were not taking seriously. The company was able, in exchange for stock rather than cash, to buy a transmitter as good as Edison's from an inventor named Francis Blake.

Meanwhile, Western Union was confidently attacking the Bell patent, declaring Gray the true inventor of the telephone. But after a year of litigation, Western Union's chief patent attorney concluded that Bell would win. In the peace treaty, Western Union agreed to admit that Bell was the inventor of telephony and that his patents were valid and to withdraw from the telephone business. Bell agreed to buy the fifty-five-city Western Union telephone system. In 1880, the American Bell Telephone Company was created, with capitalization of $6 million. By 1882, the firm booked over a million dollars of gross earnings, and on February 6 of the same year, it acquired Western Union's equipment subsidiary, Western Electric. In that year, Bell stock soared to over $1,000 per share.

In 1893 and 1894, the original Bell patents expired, and a multitude of independent local telephone companies entered the business. In 1897, these independents formed the National Telephone Association, the genesis of the United States Independent Telephone Association, known after the breakup of the Bell System as

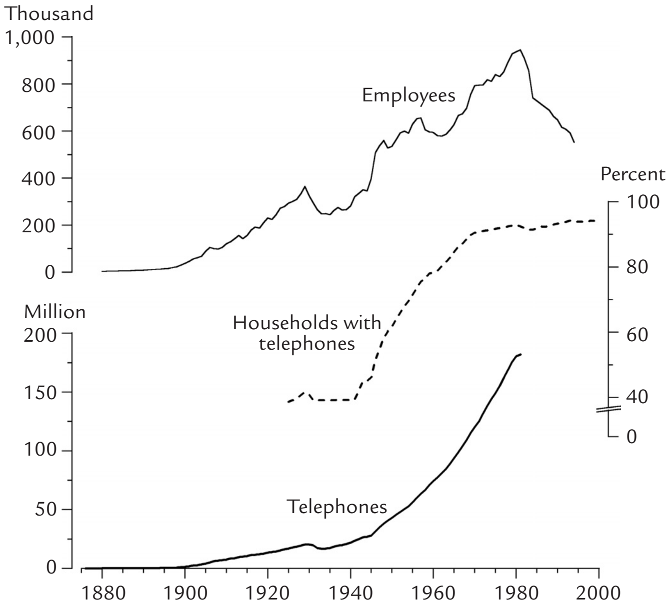


**FIGURE Dg-D Telephone industry – telephones and employees: 1876–2000**

***Sources***

Telephones: series Dg34 and Dg37. Employees: 1880–1974, series Dg41, which covers the Bell (AT&T) System; thereafter, series Dg40, which covers the entire U.S. telephone system.

the United States Telephone Association, and today as the United States Telecom Association. This organization is the main non-Bell source for the historical data recorded in the tables presented here. As series Dg34 and Figure Dg-D show, telephone use grew explosively. In New York City alone, the number of users jumped from 56,000 in 1900 to 810,000 in 1908. In 1879, the New York City telephone directory was one card with 252 names. By 1908, it had grown to an 800-page volume published quarterly.

Although there had been some exploration of switching technology within telegraphy, particularly in a network developed by a set of New York law offices (Fagen 1975, p. 729), switching had simply not been a major bottleneck or challenge in most applications. Most telegraph lines were single-wire ground return circuits, with many stations along them. Essentially, they were party lines. The bulk of telegraph use was long distance. Where routing required that a message switch from one line to another, it was simply decoded, recoded, and retransmitted on its next leg, a procedure replicated in the torn-tape teletypewriter networks of the twentieth century.

These retransmission lags were inconsequential given the asynchronous nature of telegraphic communication and the time required to move the message the "final mile." The telegraph system relied on messengers to deliver messages from the local office to the receiving household. Because the telephone did not require a skilled operator to encode and decode outgoing and incoming messages, and because it could service a demand for local communication much more conveniently than telegraphy, it made economic and technical sense to locate telephone stations in households, which had simply not been the case with telegraphs.

As the number of local users in a city grew, it immediately became apparent that economies could be garnered by connecting multiple stations to a central exchange, rather than trying to connect each directly to all of the others. To illustrate the effect, consider that five subscribers require ten lines to connect each to the other $(4 + 3 + 2 + 1)$, only five "loops" if connected to a central switch. Ten subscribers require forty-five connections for station switching, ten if connected to a central switch, and the economies become greater as the number of users increases. The development of central switches was another important capital-saving innovation, as was the continued use of party lines (multiple stations on a single loop) in many sparsely settled regions of the country.

The telephone exchange originated in New Haven in 1878. The first central office exchange using a common battery system (eight or more storage cells) began operating in Lexington, Massachusetts, in 1893. Common battery systems gradually replaced the bulky and potentially dangerous power sources that home users had had to put up with from 1882 onward. The use of lead acid batteries, recharged by line voltage generators, and with backup internal combustion engine generators, remains, remarkably, standard practice to this day. The system produces reliable, clean (hum-free) direct current, and telephone service that often remains uninterrupted, even when a household's electrical power supply fails.

The construction of switchboards was a very serious technological challenge. As the numbers of subscribers in a local exchange began to rise above 500 or 1,000, the original boards became overwhelmed. Charles Scribner developed the multiple board, capable of serving up to 10,000 customers, which formed the upper limit to the number of lines that could effectively be switched manually in a local exchange.

When a customer picked up a handset, current flow began, and a small light lit on the switchboard. The operator inserted a plug into a jack (socket) below the light, obtained the desired destination from the customer, and used a key switch to send ringer current to the called customer. Upon answer, the operator plugged a second cord into that party's jack, connected these cords together, and was removed from the loop. With the use of these boards and experienced operators, local calls could be completed in a very short time. While the central exchange saved an enormous amount of wire, boards themselves were capital-using (they could cost more than a third of a million dollars) but labor-saving (they saved up to ten seconds a call) devices (Casson 1910, p. 147).

Initially, a three-letter alpha denoted the exchange and the last four numbers the line within it, reflecting the practical maximum of 10,000 local stations per exchange. Later, as the number of exchanges grew, the two–five system was adopted. The first three digits still identified the exchange, but the last of these could be any number, increasing the number of possible identifiers. Seven-digit phone numbers were introduced in 1958.

The hub-and-spoke system that was used to economize on physical capital in local exchanges also underlay the eventual construction of regional and long distance connections (Weiman 2002). And within firms and organizations, private branch exchanges (PBXs) replicated the structure at a lower level of the hierarchy. Calls within an organization were switched internally, those requiring an outside line routed to loops connected to a phone company's local exchange.

As noted, telephone switching was labor intensive, leading to potential economies from more automatic devices. The Strowger switch, patented in 1889, was an electromechanical device that arrayed 100 contacts in a ten-by-ten curved cylindrical matrix. An electrical brush stepped up or down within the cylinder or rotated to reach any of the 100 contacts. Callers drove movements of the brush directly by pulses of ringer current. In 1896, Almon B. Strowger devised a dial that would automatically generate the pulses. Although independent systems pioneered in introducing this innovation, the Bell System resisted it, claiming that it did not work well in urban settings, and often removed steppers in independent systems it bought. Not until 1919 did Bell begin installing automatic switching equipment itself.

The first crossbar switch, which permitted up to ten simultaneous connections, was installed in 1938. The Number 5 model, introduced in 1948, became standard in the Bell System and, in 1978, accounted for the largest number of automatically switched lines in the world. In 1965, Bell installed the first all-electronic switching system, the 1ESS. These early electronic switches still used metallic contacts – reed switches – that were opened or closed by application of a switched magnetic field. All-digital switches, essentially specialized computers, were introduced in 1974. Built on advances in semiconductor technology, the new switches did away with electromechanical switches in the same way that semiconductor technology would shortly render obsolete electromechanical desk calculators. Touch-tone dialing was introduced in 1963, in anticipation of moves to electronic switches, particularly in long distance communication.

Today, an unshielded twisted pair of copper wires drops down or up from one's phone to an entrance bridge. These wires run either directly to the phone switch (local exchange) in the area or to a digital concentrator, a refrigerator-size box that converts the analog signal into digital form (in a sense returning us to the old make-or-break methods of the original telegraph system) from which the digitized signal runs along with others through coaxial or fiber-optic cable to the main switch or local exchange.

When a caller picks up a handset, the telephone goes off hook, and direct current begins to flow to the receiver from the local exchange. The central office switch detects current flow and sends a dial tone back along the wires to the station. After the caller dials a number, the called number is located, and the central office switch sends a powerful surge of current at a specific frequency down the wire to it, triggering the ringer or alerter. When the called phone goes off hook, a connection is made between the two telephone stations, and in phone parlance, the call is completed. The basic protocols have been the same for a century. Indeed, if one takes a telephone from the 1920s and updates the plug, it will work flawlessly in the current environment.

Telegraph companies typically strung a single uninsulated steel wire on metal poles with glass insulators, with the return circuit through the ground. Such lines worked poorly for telephone transmission because interference degraded voice intelligibility much more easily than it could a telegraphic signal. The remedy was to run a twin-wire system, first demonstrated between Boston and Providence in 1883. Although this resulted in a major improvement in the quality of service, it was an expensive capital-using solution.

The second advance involved changes in the materials used to make wires. Steel wires gave rise to rapid signal attenuation. Copper was better, but brittle, and its weight, given the diameters of wire originally used, was incompatible with the typical 100-feet span between two poles. Hard-drawn copper, first introduced in 1884, solved this problem. Cross talk – interference among two or more simultaneous conversations – was addressed by twisting pairs of wires together, but areas of dense telephone demand, such as New York City, soon found their skies clouded with a web of wires, exposed to oxidation, ice, sleet, and other elements.

Putting telephone cables underground, however, required careful experimentation with insulators. At first, wires were wrapped in gutta-percha and soaked in oil (to eliminate moisture) in the manner of an undersea cable. Eventually, dry-core cables, with each wire wrapped in paper twine (so as to preserve some air as insulation) and then shrink-wrapped in lead, became the standard. By the turn of the century, more than 1,200 cables were often bunched in one sheath, accessible through manhole covers and at switching boxes where individual wires ran to customers (Casson 1910, p. 134). The cabling in a New York skyscraper at the turn of the century could weigh as much as fifty tons.

Copper wire was four times as good as steel for telephony, but it was also four times as expensive, and the immense technological payoffs to multiplexing in both telegraphy and telephony were driven by attempts to save on this capital – to increase its throughput or utilization rate. Another way to save capital was to use thinner wires.

The initial use of telephones, and their predominant use today, remains local calling, with residential demand figuring prominently. But from the very outset, Bell had conceived of a grand system in which the phone company linked local exchanges via long distance trunk lines, thus bringing major commercial centers into voice contact. When long distance telephony became possible, businesses rather than households drove demand for it. This was so because, unlike the telegraph, the telephone permitted quick, real-time message exchange involving nonstandardized information that was critical to the conduct of wholesale trade. For this type of negotiation, face-to-face communication was the preferred modality, but if that was not possible, the telephone was vastly preferred to the telegraph. Thus, the telephone substituted for railroad transportation as much as for telegraphic communication (Weiman 2002).

The development of long distance calling required technical advances beyond those necessary for local service. Better cabling, loading coils, and improved amplifiers were critical in the effort to deal with signal attenuation. Attenuation was doubly problematic because it differentially weakened signals of different frequencies, leading to distortion and degradation of intelligibility. The problem could be alleviated somewhat by increasing the cross section of the copper wire, but because of the physics of transmission, the cross section had to increase not proportionally but by the square of the distance multiple if one wished to extend range. Thus, to double the maximum range of intelligibility, one had to quadruple the cross section of the wire. The unsuitability of thick copper wires for aerial installations has already been noted. Because of the physics and economics of transmission, telephony was effectively limited to a maximum of thirty miles until the development of the loading coil and the solution to the mathematics of how it should be deployed. This range was adequate for the development of local exchange business but inadequate as a means of linking major commercial centers, such as New York and Chicago.

A loading coil is a metallic doughnut around which the copper wires going in both directions on a telephone line are wound. These coils were first tested in 1900, and by 1925, more than 1.25 million were in use. Introduced every several miles in a line, they reduced signal attenuation and distortion and increased the maximum range of telephony, in the absence of human repeaters, to about 1,000 miles. Equally important, they permitted the use of thinner wire, thus saving as much as $40 per mile of cable (Casson 1910, p. 139). Lighter wire also meant that aerial cables could be strung over longer distances using a smaller number of spans and, thus, fewer poles. The power of this capital-saving innovation is reflected in a conversation with an AT&T vice president that was reported by the inventor of the coil, Michael Pupin, in his autobiography. The executive told him that over twenty-two years, coils had saved AT&T more than $100 million in capital costs (Pupin 1923, p. 339). Pupin himself made a fortune from the sale of his patent rights to the company. A central part of his contribution was working out mathematically the optimal spacing of these devices (Wasserman 1985).

Although no longer used in long distance service, loading coils continue to be found in local exchange cabling. After almost a century of use, however, they are now being rapidly removed. Loading coils block high-frequency signals, improving the clarity of a phone message but rendering copper wires unsuitable for digital subscriber line (DSL) service. DSL is a technology for providing broadband access to the Internet over twisted pairs at frequencies above those used for voice; it thus multiplexes the final mile and permits continuous data access without tying up a phone line.

Loading coil technology and mechanical amplifiers (using the weakened signal to drive a loudspeaker placed near a carbon microphone) allowed telephony over substantial distances, as in the New York-to-Chicago service introduced in 1904. But mechanical coupling introduced distortion because not all frequencies were amplified proportionally. No more than three could practically be connected in series; therefore, such amplification could not enable a transcontinental line. For that, one needed Lee de Forest's vacuum tube. De Forest's patents, issued in 1907 and 1908, were for a "Device for Amplifying Feeble Electric Currents." The tube, which he called an audion, was also good at picking up modulated radio waves, and the inventor was principally interested in wireless applications (Fagen 1975, p. 260). It took five years for their potential in long distance wired telephony to be appreciated.

Using triodes as electronic amplifiers, AT&T opened the first transcontinental line, between New York and San Francisco, on January 25, 1915. Built with one-sixth-inch copper wire weighing 870 pounds per loop mile, it required 2,500 tons of wire on 130,000 poles, three vacuum tube repeaters, and loading coils at intervals of eight miles. Even so, upon inauguration, it allowed only one two-way conversation. The initial and subsequent costs of calls on this line, along with those for other long distance routes, are reported in Table Dg56–63.

In 1918, building on modulation technology developed for wireless broadcasting , Bell introduced carrier-wave frequency-division multiplexing over wires, which quadrupled the number of simultaneous phone conversations per line. The same technology could allow up to twenty-four simultaneously transmitted telegraphic signals on a single wired circuit. These improvements are again examples of the relentless search for capital-saving innovation in this sector.

More repeaters were added to the New York–to–San Francisco line in 1915 and 1918, and in 1920 all of the loading coils were removed. AT&T addressed the concomitant problem of signal loss instead by using twelve improved vacuum tube repeaters. In 1950, Western Union installed the first undersea vacuum tube repeaters, and in 1956, AT&T used them in the first transatlantic telephone cable. This use of tubes, which operated flawlessly for ten years, was an engineering triumph. But their days were numbered. Indeed, it was the Bell System's desire to find alternatives to the use of tubes as repeaters (amplifiers) that drove the program of research in solid-state physics, culminating in 1947 in the invention of the transistor at Bell Labs (Mowery and Rosenberg 2000, p. 878). A transistor can perform the same functions as a vacuum tube, but it consumes less power, generates less heat, and is less prone to failure.

In telephony, advances in multiplexing have continued to be critical for getting more bandwidth out of existing pipes, be they copper, fiber-optic, coaxial, or microwave. So-called T1 service was developed in 1957 and widely introduced in 1962. T1, which is a transmission and not, per se, a cabling standard, digitally encodes twenty-four analog signals and multiplexes them into a 1.544 megabit per second (Mbps) signal that can be sent over copper wire. Digital transmission is based on pulse code modulation, a technology invented by Claude Shannon and others in 1948 at Bell Labs. Pulse code modulation encodes a voice or audio source by sampling it 8,000 times a second and representing each sample by eight bits of data. In other words, the height of the audio signal, which represents a complex combination of many sine waves with different phases, is measured at discrete time intervals as one of 256 ($2^8$) levels, or quantiles, and so encoded using eight binary digits.

Using time-division multiplexing, twenty-four different signals are then interleaved, along with a single synchronization bit, producing a 193-bit frame [$(24 \times 8) + 1$]. Sending 8,000 of these frames per second gives rise to the T1 bandwidth of 1.544 megabits per second ($8{,}000 \times 193$). Each individual channel is thus associated with a bitstream running at one twenty-fourth of this rate, or 64 kilobits per second (Kbps). Increasing bandwidth is thus enabled, not by sending individual bits faster, but by making each one of them shorter. T1 was designed to exploit the shortest bit duration and thus the maximum number of bits per second that could reliably be transmitted along the twenty-two-gauge copper wire common in most metropolitan systems. Pulse code modulation is the basic technique used to convert music into digital data on a CD-ROM, except that the sampling rate is higher (44,100 times per second for each channel), and a sixteen-bit rather than eight-bit value is created for each sample.

Bell had tried and failed in 1877 to make a telephone call over the Atlantic cable. Although a telephone cable was successfully introduced between San Francisco and Oakland in 1884, it would be more than seventy years before cabled telephone service reached across the Atlantic. Regular service by radio to Europe utilizing very powerful, long radio waves was introduced in 1927. Long waves, unlike the shorter microwaves, could bounce off the earth's ionosphere. Capacity was very limited, however, and the initial cost was $75 for a three-minute call (see Table Dg56–63).

The first transatlantic telephone cable in 1956 (actually, twin cables) used coaxial and tubed repeaters and was able to handle thirty calls. AT&T launched the Telstar satellite in 1963 with the intent to supplement the cable for transatlantic telephony. But because of the distances involved (these satellites travel in orbit at 22,300 miles above a fixed point on the earth), geosynchronous satellites introduce a total of about a half-second delay between send and receive. Signals must travel over 45,000 miles in each direction on a transatlantic call; electromagnetic radiation, such as radio waves, travels at about the speed of light in the ether: 186,000 miles per second. Experience has now shown that customers will not tolerate, without complaint, a delay of more than about a tenth of a second. Where satellite links are used today, the standard is to have conversation go in one direction by satellite and return by fiber-optic undersea cable. Geosynchronous satellites, which use very high frequency radio signals with high bandwidth channels, are ideally suited for one-way television and data transmission, such as sending a newspaper from one printing plant to another across the country. The commercially ill-fated Iridium system utilized a network of low-orbit satellites, which, because of their lower altitude, introduced less delay.

Over land, copper, fiber-optic, and coaxial cabling is supplemented with microwave links based on line-of-sight transmission from one station to another, situated at high points about thirty miles apart. They are the closest modern equivalents to the old Chappe telegraph stations. Although analog signals travel all or part of the way from a customer to the local exchange, today's long distance links consist exclusively of digitized data.

As a conclusion to this section on the telephone, a few words are in order about regulatory changes. Governmental concern about the monopoly power of AT&T was an issue periodically throughout the twentieth century. With the expiration of the original Bell patents in 1893 and 1894, and the explosion of independents, conflicts arose because of Bell's refusal to allow local systems to interconnect with the Bell long distance lines. And in 1909, as mentioned, Bell acquired control of Western Union, giving it a stranglehold on long distance and much local communication, but also providing convenience to consumers, for example, by making it possible for the first time to order a telegram over the telephone. Four years later, under government pressure, however, the company, in what is known as the Kingsbury Commitment, agreed to let go of Western Union, to connect its long distance lines to independent operators if certain conditions were met, and not to acquire an independent if the U.S. Department of Justice objected.

In 1974, Justice brought an antitrust suit against AT&T that was finally settled in 1982, with an historic divestiture agreement (the modified final consent decree) implemented in 1984. AT&T assets shrank at one fell swoop from $150 billion to $34 billion, as twenty-two regional telephone systems were grouped into seven Regional Holding Companies (RHCs), also known as Regional Bell Operating Companies (RBOCs), each with about 60,000 employees and $8 billion to $9 billion of assets (Vietor 2000, p. 1004). As a result of mergers, four remain today. AT&T kept long distance service, Bell Labs (later spun off as Lucent Technologies), and equipment manufacturing (Western Electric), with the RBOCs blocked from these sectors, but AT&T was to stay out of the provision of local service. Series Dg76–78 provide data on the revenues and net incomes of local exchange carriers.

The Telecommunications Act of 1996 relaxed many of the restraints of the modified final settlement, allowing new entrants (including AT&T) to provide local service and permitting RBOCs to offer long distance service. AT&T has since established significant ownership stakes in the cellular phone business (by purchasing McCaw Cellular Communications in 1994) and in cable television through its acquisition of TCI (Telecommunications, Inc.) and has continued to acquire cable systems, with the intent of providing local service through cable telephony. As this is written, the success of the strategy appears in doubt, as AT&T apparently intends to break itself into four smaller companies.

## Radio and Television

By the 1890s, telecommunications engineers knew how to send data or voice signals over wire. What they did not know how to do well was send such information through the air. Semaphoric systems could be used only for data, were vulnerable to the weather, and by necessity were land based. Semaphore enthusiasts did talk seriously about floating stations anchored eight to ten kilometers apart, but the logistical challenges of provisioning these stations, let alone maintaining clear lines of sight between them, were insurmountable. And although telegraphic data could be sent via the undersea cable, Bell had failed in his attempts to use it to transmit voice. None of these alternatives, moreover, was suitable for mobile communication, particularly that associated with communications between ships and between ships and shore. For practical purposes, telegraph and telephone transmissions were tethered to wire.

James Maxwell's theoretical work on electromagnetic waves paved the way for the first generation and measurement of radio waves by Heinrich Hertz. Between 1884 and 1888, Hertz demonstrated their existence, showed practically how to generate and detect them, and measured their velocity and wavelength. His generator used a battery-driven capacitor, which accumulated electricity until it was discharged across a spark gap connected to short lengths of wire, his broadcast antenna. His detector was even cruder: a circular piece of wire with a very small spark gap in one section. With this device, he generated and detected radio waves across short distances and, in principle, demonstrated the possibility of wireless telegraphy.

A number of scientists in France, England, and Russia contributed to the development of an improved detector based on an insulated tube filled with metal filings (Edouard Branly in France called it a coherer). In the presence of radio waves generated by sparks, the resistance of the filings changed, and it could be measured by a galvanometer or used to set off a bell. Guglielmo Marconi built on these investigations to create a commercially viable system. By 1896, using spark gap and coherer technology, he was able to transmit Morse code without wires over distances of up to two miles. He moved to England and set up the British Marconi Company, and on December 12, 1901, situated in Cornwall, he used a very long antenna raised by a kite to detect transmissions from a station in Newfoundland, more than 1,700 miles away. By 1907, Marconi had established a commercial service, focusing on ship-to-shore applications, which faced no competition from undersea cable. It was in this market that he had his greatest success in the first decade of the twentieth century, facilitated in part by the International Ship Act stipulating the presence of a wireless transmitter on every ocean-going vessel. These devices played a role in the *Titanic* tragedy, summoning assistance to the doomed vessel, although system incompatibility meant that some potential rescuers within thirty miles could not hear the distress signal.

The generation of radio waves by sparks, however, was wasteful of both energy and the spectrum. The average power of a transmission burst was less than the peak power, and the transmission was not closely concentrated in any one frequency, making tuning difficult and raising problems of signal interference as more users competed in a given geographical area. One could not simply treat the ether as one big wire; what was needed was a means of dividing the spectrum into different channels, or frequencies, so that users would not interfere with one another. In other words, one needed frequency-division multiplexing of the radio spectrum.

Figuring out how to do this also opened the way for radio telephony. Sound waves audible to humans vary in frequency from about 20 to 20,000 cycles per second (Hz). Telephones transmit sounds in the 300- to 3,500-Hz band (adequate for voice communication). A transducer, such as a microphone, converts these cycles to fluctuations of similar frequency in direct current, which can easily pass through a wire to the receiver. Sound waves per se cannot be broadcast very far.

Prior to Hertz's work on radio waves, Bell had actually demonstrated the transmission of speech over a beam of light and received a patent for a "photophone." But like the Chappe and all other optical systems at the time, any long distance transmission by light operated at the discretion of the weather. The big advantage of radio waves, much lower in frequency than light, is that their use is almost entirely unaffected by weather.

The trick that made wireless telephony (and the modern radio industry) possible is the use of a carrier wave of a specific frequency whose amplitude or frequency may be modulated in response to an audio signal, and then demodulated at the point of reception. To do this, one needed to generate a radio wave of steady frequency for use as a carrier. Early techniques for doing so were the rotary alternator, championed by Nicola Tesla and Reginald Fessenden, and the electric arc, favored by Valdemar Poulsen. But there was no practical way of modulating waves produced by these devices.

It was de Forest's invention of the triode in 1907 that formed the technical foundation for both broadcast and two-way radio communication as we know it, and also, as we have seen, advances in long distance wired telephony. Radio broadcast requires thousands of kilowatts of electric power, as opposed to the few milliwatts necessary for wired telephony. The vacuum tube could be used to generate carrier waves, to modulate them in response to an audio signal, to amplify the modulated signal so that it could be broadcast over an antenna, to detect the radio signals in a receiver, to extract the audio signal through the process of demodulation, and to amplify it so that it could drive a loudspeaker.

Edison had stumbled upon a diode by inserting an additional electrode into a light bulb, but he did not realize its potential use. In 1904, Ambrose Fleming, working for Marconi, invented the first vacuum tube. A diode (which he called a valve) is an evacuated bulb containing a metal cathode and anode. If an electrical potential is applied between them, current will flow across the gap between them. A diode can serve as a rectifier, turning alternating current (AC) into direct current (DC), and also as a radio wave detector (current flow fluctuates in the presence of such waves), which is what Fleming intended it for.

In 1907, de Forest patented the triode, a tube that inserts a grid between the cathode and the anode. A changing voltage applied to the grid, even at low power levels, will cause fluctuations in the current flow from the cathode to the anode, effectively allowing amplification of the power of a weak audio signal. De Forest's tube, which ushered in the age of modern electronics, made radio telephony possible and, as we have seen, enabled a vast improvement in long distance wired telephony. The triode could be used both as a detector and demodulator of radio waves and as a modulator, essential for wireless radio broadcast.

Building on de Forest's invention, Bell engineers in 1915 succeeded in using long radio waves to transmit speech from Arlington, Virginia, to Panama, Hawai'i, and Paris (Fagen 1975, pp. 368–9). It was essentially this line of research, delayed temporarily by the First World War, that underlay the inauguration in 1927 of radiotelephone service to Europe, the only vehicle for transoceanic voice communication until 1956. Except over oceans and in mobile services on land, wireless two-way point-to-point telephony could not compete with wired implementations, a situation that remains only slightly altered today. Cellular telephone systems, obviously, have successfully demonstrated their potential to serve as an alternative to local wired service, but their appeal depends in part on their interconnection with the wired system, particularly for access to long distance calls.

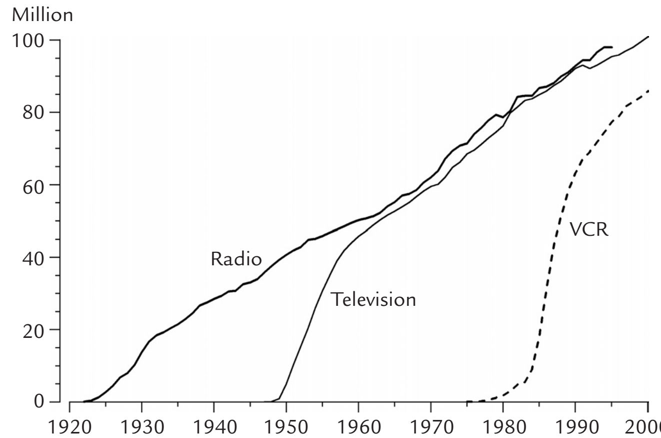


**FIGURE Dg-E Households with radio sets, television sets, and videocassette recorders: 1922–2000**

***Sources***

Series Dg128–130.

***Documentation***

Nearly 100 percent of households had radio sets as of the early 1950s, and so the plot for radios defines an approximate upper bound (for example, compare series Ae79 and Dg128).

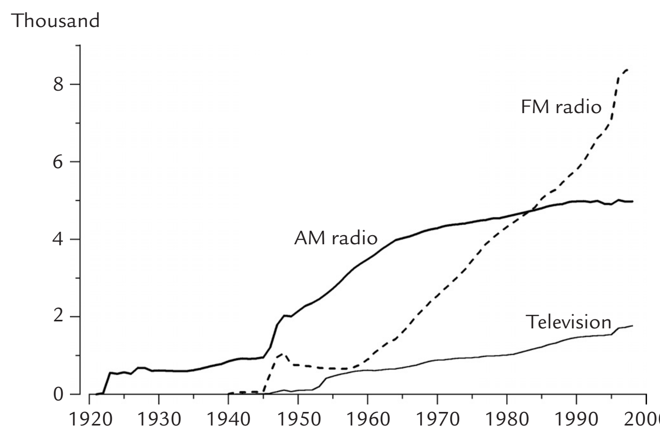


**FIGURE Dg-F Operating radio and television broadcast stations: 1921–1998**

***Sources***

AM radio: series Dg117. FM radio: series Dg118–119. Television: series Dg120–121.

***Documentation***

The graphs for FM radio and television include both commercial and non-commercial stations.

Commercial broadcasting – the industry whose growth is charted in Table Dg117–130 and Figures Dg-E and Dg-F – began in 1921, but no one knew at its birth exactly how it would evolve. As noted, Bell had used telephonic broadcasts as a publicity stunt, and the modality resurfaced with the availability, thanks to de Forest, of powerful audio amplifiers. In 1919 in New York, AT&T used a system of a hundred loudspeakers over a three-block system in a war bond campaign, and on November 11, 1921, 150,000 people in

Arlington, Virginia, New York, and San Francisco, using open telephone lines, amplifiers, and loudspeakers, listened to the dedication of the Tomb of the Unknown Soldier (Fagen 1975, p. 425).

In the event, the radio and television broadcast industry developed with wireless transmission used for local and regional distribution of content and land lines used to interconnect local transmitters in order to generate national radio networks. In the telecommunications industry, however, no generalization should ever be taken as final. Television broadcast now arrives in American homes not only via traditional local broadcast stations but also over cable and via satellite. Radio and video are also increasingly available over the packet-switched Internet, a system that depends largely on land lines. So Bell's suggestion that broadcast might arrive over the telephone network is perhaps rising from the ashes, even if the modalities are far different from those he anticipated.

Unlike newspapers, magazines, or books, broadcasting via radio waves is regulated by the federal government. The traditional story is that in the absence of licensing for transmission on specified frequencies, wattages, and hours of operation, chaos developed because of interference from overlapping signals. Others argue that over the 1920–1926 period, priority-in-use rules derived from common law were adequate to address problems of interference (see Hazlett 1990). What is less in dispute is that the Radio Act of 1927, which created the Federal Radio Commission, had the consequence of permitting industry pioneers to obtain at low or no cost the equivalent of secure property rights to broadcast at certain frequencies, power levels, and times of day. An alternative to the actual regulatory structure adopted, one vigorously supported by Ronald Coase in a classic article in 1959, would have been for the government to auction off spectrum rights to the highest bidder. Coase's proposals were, in fact, implemented in the 1990s in auctions of cellular telephone frequencies, although the "give away in the public interest" principle appears to have informed the recent grant of frequencies for high definition television broadcast to the networks, valued by some estimates at roughly $70 billion.

The Federal Radio Commission became, in 1934, the Federal Communications Commission, one of whose functions remains the assignment of different users to different bands of frequencies. For example, AM (amplitude modulation) radio utilizes a band from 535 kilohertz (kHz) to 1.7 megahertz (MHz); VHF (very high frequency) television, 54–72, 76–88, and 174–216 MHz; and FM (frequency modulation) radio from 88 to 107 MHz. CB (citizens band) radio uses a band of forty channels in the 27-MHz range; garage door openers operate at 40 MHz; cordless phones have bands between 40 and 50 MHz and also 900 MHz; radio-controlled cars, 72 MHz; radio-controlled airplanes, 75 MHz; and so forth. The intent of regulatory allocation is to assign frequencies most appropriate for their use and to prevent signals from interfering with one another within specified time intervals or geographical areas.

The basic principles of the television broadcast industry represent an extension of those developed for radio. A carrier wave is used to send a diplex modulated signal with input from both a video and an audio source. The video signal is carried through AM and the audio via FM. A difference between radio and television is that in radio, the transducer is a microphone, and in television it is a device that converts fluctuations of light into fluctuations of electric current.

The first completely electronic system for transmitting and receiving images over the air was demonstrated in 1927 by Philo T. Farnsworth. Vladimir Zworykin, working for the Radio Corporation of American (RCA), developed an image iconoscope in 1933, claiming priority for the invention of television on the basis of a 1923 application for a patent that was never awarded. David Sarnoff, head of RCA, was determined not to pay royalties to Farnsworth and filed a patent infringement suit against him. RCA's case was shaky, however, and after protracted litigation, Farnsworth prevailed, although his victory was Pyrrhic. World War II delayed the commercial exploitation of the medium, and by the late 1940s, many of his patents had expired. During the 1930s, RCA worked on refining what Farnsworth had invented, eventually developing the image orthicon tube, which was widely used in the early days of television (the name of the Emmy award for television programs is a corruption of "Immy," shorthand for these tubes). Video transducers today include those employing charge-coupled devices, integrated circuits capable of converting light to electricity.

Television requires a visual image to be scanned and broken down into pieces of information that can be transmitted electronically and then reassembled. Just as pith ball telegraphy had anticipated separate wires for each letter, so early schemes for television anticipated separate wires for each picture element. The breakthrough of sequentially scanned elements that rely on the persistence property of human vision to put them together in a moving picture allowed the possibility of transmission over a single circuit. Initial scanning technologies were mechanical, such as the Nipkow disk, a sheet of perforated metal in a spiral pattern that, when rotated, scanned a series of progressively smaller rings covering the picture. Light passing through these apertures struck a photoelectric substance, such as selenium, generating fluctuating current, which, when transmitted to a receiving site, could drive a light with varying intensity, such as a neon gas discharge tube, behind an identical and perfectly synchronized disk, thus regenerating the picture.

The problem was that mechanical scanning technologies could not scan at resolutions higher than about 200 lines. The image disector, invented by Farnsworth, used electromagnets to control a scanning electron beam that traced line-by-line patterns over a signal plate consisting of a mosaic of photoelectric substances acting as small capacitors. As the beam struck them successively, the stored charge was released, generating fluctuations in electric current that drove the video signal.

It was not until the 1939 World's Fair that television broadcast was introduced to a broad audience. Commercial television broadcasting began in the United States in 1941. During World War II, the production of all commercial television sets in the United States was banned, although, in contrast with Britain, a limited broadcast schedule did continue. Only after the war did television begin to take the American public by storm, as can be seen in Table Dg117–130. In 1947, Kodak and the National Broadcasting Company (NBC) introduced the kinescope, a motion picture camera adapted for filming a live television broadcast, thus permitting image storage and rebroadcast at a later date. Kinescopes remained the only available storage technology until the introduction of videotape machines in 1956, and the films made with them are our only record of the early years of television broadcast.

The first coast-to-coast broadcast took place in 1951. That same year, the Columbia Broadcasting System (CBS) began color broadcasts using a mechanical scanning technology related to Nipkow disks, but at the end of the year, all color television broadcast was suspended owing to the outbreak of the Korean War. In 1953, the FCC approved a different RCA-pioneered all-electronic system, known today as NTSC (National Television System Committee). Broadcasts using this system began in November of 1953, and the first production model sets were released in 1954, meeting an initially cool reception from the public. The development of NTSC color television defined the basic technological paradigm for television that persisted through the end of the twentieth century.

Television broadcast has an effective range of about seventy-five miles. Network broadcast is enabled by the use of microwave relay stations about thirty miles apart or (today) by satellite broadcast. The delivery of signals to homes is increasingly over cable (AT&T invented coaxial cable in 1935) or by direct satellite transmission, as well as by local broadcast.

There were many technological improvements in the design of both broadcast and receiving equipment, including the replacement of tubes with transistors, and, in the mid-1970s, the introduction of home videocassette recorders. An analog version of high definition television (HDTV) was introduced in Japan in the early 1990s. Requiring an enormous bandwidth, the system has yet to be licensed in the United States, and at the beginning of the new century, only limited broadcasts of all-digital HDTV have begun in the United States.

Table Dg162–171 deals with the use of radio as a point-to-point communication medium. There is no corresponding table for television because, even more so than radio, TV has been almost exclusively a broadcast technology. Closed-circuit systems have been used for surveillance purposes from the inception of the technology, but these were simplex, not designed for two-way communication. AT&T introduced the duplex Picturephone service at the New York World's Fair in 1964, but it was never a commercial success because it would have tied up multiple telephone lines. High bandwidth requirements today are only slowly being overcome through compression techniques and drops in transmission costs. It is not clear that residential customers want video links, at least for local calling. For long distance, inexpensive cameras now permit crude point-to-point "television" over the Web. Video conferencing as an alternative to air travel is available to large corporations and universities, as well as to businesses that use off-site office services. The extent of the long-term demand for these services remains uncertain.

### Internet

The Internet is, at its most fundamental level, a set of protocols for interconnecting networks of electronic devices, particularly computers. Its novelty lies not in the use of a binary code to represent information, nor in innovations in the data transmission pipes themselves, developed initially for telephony and telegraphy. Rather, it involves an entirely new set of procedures, or protocols, for moving data over this infrastructure, the most important element of which has been packet switching.[3]

The theory of packet switching was first adumbrated in a 1961 doctoral proposal by Leonard Kleinrock at the Massachusetts Institute of Technology. Bell engineers had long understood that multiplying the number of lines on a trunk circuit increased its effective capacity by more than the multiplicative factor because the law of large numbers means that less toll business would be lost owing to peaks of congestion. Kleinrock's work envisaged an extension of the queuing theory developed for single-node systems to those involving multiple nodes, that is, situations where there was more than one possible route between source and destination. Packet switching involves sending a message not as a continual stream of data but as a bunch of packets, each with an address and sequencing label, and each of which might travel (depending upon congestion) by a different route to be assembled again at the endpoint.

Unlike a phone conversation, which ties up a completed circuit for its entire duration regardless of the long silences that may intersperse a difficult conversation, packet switching uses circuits only when there is data to transmit, and reduces congestion by reallocating packets on the fly to less congested routes. It is another of a long line of capital-saving innovations in the history of telecommunications. Most voice communication, and the modem links whereby the majority of home users at the end of the twentieth century connected to their ISPs, continue to be circuit switched. Some estimates suggest that if all voice communication were converted from circuit to packet switching, the effective capacity of the existing communications infrastructure could be doubled.

The principle of digitizing data is not new, going back at least to Baudot's innovations in the late nineteenth century. Claude Shannon's work on pulse code modulation in 1948 enabled the digitization of voice. Using analog-to-digital-to-analog conversion, long distance telephone conversations in the 1960s moved increasingly in digital form over T1 lines in the middle part of their journey. The development of mainframe computers in the 1950s and 1960s, and the desire to connect remote terminals to them via the telephone network, had posed the reverse problem: the need to convert digital to analog to digital. The solution was the development of modems, which enable digital data to move over part of its journey in analog form before being demodulated at the endpoint into digital form. But whether analog or digital, voice or data, telephone networks ultimately move information from one endpoint to another by establishing a continuous end-to-end circuit between the transmitter and the receiver. This is what packet switching dispensed with.

The Internet can be dated from 1969, with the interconnection of the first four hosts: the University of California at Los Angeles (Kleinrock was by then a professor there), the Stanford Research Institute (where the team included Douglas Englebart, who later invented the mouse), the University of California at Santa Barbara, and the University of Utah. The achievement lay in linking four physically separate computers produced by different manufacturers using different operating systems. Linkage was made possible by a piece of hardware called an interface message processor (IMP) attached to each machine and interconnected by leased phone lines. The IMP, whose modern descendant is the router, was a special kind of switch built under contract for the Pentagon by Bolt, Beranek and Newman, a Cambridge, Massachusetts, consulting firm. Designed for packet switching, and intended for enabling the first wide area network, IMPs also, in an unanticipated fashion, enabled the first local area networks (LANs) by permitting different computers on the same sites to communicate with each other. Series Dg110 charts the explosive growth in the number of Internet hosts.

With the time-sharing technology of the 1960s, terminals connected by leased wire or via modems over phone lines could access a mainframe from across town or across a country. But while connected, a terminal could talk only to a particular mainframe. An important impetus in developing the Internet was the facilitating of remote access to multiple mainframes, both to distribute computing

[3] See the Internet site of the Internet Society (ISOC).

load over time and to share data. This again reflected the search for innovations that could increase the utilization rates of scarce capital situated in geographically fixed locations.

Few anticipated that the main use of the network by researchers would be to send personal messages to each other via different mainframes. The protocols enabling computer networks to talk to each other ended up putting the equivalent of a souped-up printing telegraph on the desk of everyone who had access, and the value of this access has increased as more people have connected. In a manner totally unanticipated, electronic mail provided convenient, point-to-point asynchronous communication most analogous to that made possible by the telegraph in the mid-nineteenth century.

But the Internet did more than simply reinvent the telegraph. The convenience, ease of access, and low marginal cost made the system not just quantitatively but qualitatively different from anything that existed in the nineteenth century. Users of corporate teletypwriters sometimes had access to limited within-network electronic mail capability. And in time-sharing environments, different remote users of the same mainframe had been able to send messages to each other. But email among users of different machines, available on demand from desktops at zero marginal cost, was simply not possible before 1972. Most likely, the Defense Department would not have funded development of a network had it anticipated its main use; at the time, the telephone seemed perfectly adequate for desktop person-to-person communication.

At the heart of the Internet today are the rules embodied in the TCP-IP protocols. The Internet protocol (IP) defines the standard for addressing packets. The Transmission Control Protocol (TCP) governs the disassembly of messages into packets, attachment of address and sequence headers, selection of alternate routes through the network depending on congestion, reassembly (using the sequencing headers) at destination, and recovery from lost or degraded packets.

The original IP address system anticipated a maximum of 256 nodes. The assumption was that the Net would connect a limited number of large systems, rather than hundreds of thousands of smaller ones. The domain name system, whose growth and composition is charted in Table Dg110–116, was developed at the University of Wisconsin and first introduced in 1984. Twelve-digit IP addresses, which must be unique to each host, correspond to telephone numbers in telephony. Although easy to interpret by computers, they are difficult for humans to remember. Domain name servers solve the problem by translating domain names into twelve-digit IP addresses.

Because interoperability is a main goal of Internet protocols, and indeed a positive attribute of any communications network, the achievement of widely shared standards has been critical for the evolution of telecommunications. Although railroads can and did develop with different track widths, there are many advantages, particularly in the realm of saving capital, of standard gauges. Differences also exist in telecommunications. Standards-setting bodies have been and are important in its development. These include, for example, the International Telecommunications Union, responsible for standards that allow interoperability of national telephone and telegraph systems; the American National Standards Institute (ANSI), which has played a role in formalizing the ASCII code; and the Electronics Industry Association, which defined such conventions as the RS232C serial interface.

What is interesting in this regard is how much the rules of the Internet have *not* developed under the aegis of a standards-setting body. This is particularly the case with respect to the explosive growth of the World Wide Web in the 1990s. In 1992, Tim Berners-Lee at CERN (the European Organization for Nuclear Research) developed the protocols that define Web-based communication. These included addressing conventions (URLs), which also defined the type or scheme of Internet communication, as well as two sets of rules that defined the formatting and transmission of Web pages: HTML (hypertext markup language) and HTTP (hypertext transfer protocol). Berners-Lee's protocols ensure interoperability in the transmission and receipt of Web pages, which can ultimately include text, graphics, audio, and video imagery, all with hyperlinks – automatic connections to other pages, a system arguably anticipated by Vannevar Bush in an *Atlantic Monthly* article in 1945.

The Web protocols operate within the more encompassing TCP-IP conventions that permit other Net applications, such as email and file transfer, which can be integrated and made easily and conveniently accessible from Web pages. In January 1993, the first browser, Mosaic, was released, which led to Netscape and its main competitor, Microsoft's Internet Explorer. This software made possible easy access to HTML pages from distant servers. Because it is an open standard, various plug-ins are now available that permit HTML pages to provide audio and video display, for example.

The four original Internet hosts were connected by leased 45-Kbps telephone lines. Most business and organizations move data in and out using dedicated T1-size pipes to an Internet service provider. The Internet backbone – high bandwidth trunklines devoted to data traffic – was originally provided by the Defense Department through its Advanced Research Project Agency network, ARPANET, and for a time by the National Science Foundation (NSF) through its own network, NSFNET. The intent was to provide remote access to a fixed number of supercomputer facilities. In the event, NSF's motivation of achieving higher utilization rates for these centers was undercut by the rapid drop in the cost of computers: workstations became so powerful and inexpensive that the rationale for the NSFNET disappeared. In 1992, NSF got out of the business, and the backbone is now the responsibility of a number of commercial providers. Data over the backbone, which consists of the longest and fastest links in the packet-switched network, move over T3 links: full duplex fiber-optic pipes twenty-eight times larger than a T1 (about 45 Mbps, as opposed to 1.544 Mbps). T3 lines consist of two cables (one to receive and one to transmit), coaxial into and out of an organization and fiber-optic over most of their length. Data are formatted according to the DS3 standard, and the pipes are sometimes referred to using that designation.

## Conclusion

This essay has shown that even prior to 1969, there were tendencies toward convergence, for example, among the technologies underlying telegraphy, telephony, and radio broadcast. The development of the Internet has accelerated these tendencies. It has put a printing telegraph on the desk of the millions who have access to it; now offers tolerably good and some would say superior substitutes for printed newspapers, journals, magazines, and books; provides decent alternatives to broadcast radio, if not television; and is closing the gap in telephony. The packet-switched technology of the Internet will, no doubt, continue to accentuate the blurring of traditional boundaries separating different communications sectors.

Yet we should be cautious in concluding that everything is, or will be, new. The Internet builds on modalities and technologies developed in the past. It is striking, in looking back over the sector's history to date, how revolutionary, technical change has led so frequently to relatively stable paradigms that endured for decades and even longer. To understand the lines along which the network will develop, and how it will influence modalities of moving and storing information that preceded it, we must be cognizant of its history and the infrastructure and media upon whose foundations it rests.

Much was written in the late 1990s, for example, about the contribution that interconnection would make to a world of distributed intelligence. But the theme is older, dating back at least to the mid-nineteenth century: "by means of electricity, the world of matter has become a great nerve, vibrating thousands of miles in a breathless point of time. . . . The round globe is a vast brain, instinct with intelligence." These words were penned by Nathaniel Hawthorne in *The House of the Seven Gables*. Here, Hawthorne has one of his characters describe the imagined impact of the electromagnetic telegraph (1851, Chapter 17). Major advances in communications technology, from Chappe's optical telegraph through the opening of the Atlantic cable through the development of television in the twentieth century, have frequently triggered powerful, sometimes utopian, or almost millenarian visions of their potential. With the benefit of hindsight, we can put some of these enthusiasms in perspective, and we will undoubtedly do the same looking back at the Internet hysteria of the late 1990s. At the same time, the collapse of stock valuations in this sector at the end of the twentieth century may lead to a reverse bias: underestimating the still only partially realized potential of these technologies to facilitate economic growth and improve human welfare.

## References


Bush, Vannevar. 1945. "As We May Think." *Atlantic Monthly* (July): 101–8.

Casson, Herbert. 1910. *The History of the Telephone*. 1st edition. A. C. McClure.

Coase, R. H. 1959. "The Federal Communications Commission." *Journal of Law and Economics* 2 (October): 1–40.

Engerman, Stanley L., and Robert Gallman, editors. 1996, 2000. *The Cambridge Economic History of the United States*. 3 volumes. Cambridge University Press.

Fagen, M. D., editor. 1975. *A History of Engineering and Science in the Bell System*. Bell Telephone Laboratories.

Field, Alexander J. 1992. "The Magnetic Telegraph, Price and Quantity Data, and the New Management of Capital." *Journal of Economic History* 52 (June): 401–13.

Field, Alexander J. 1994. "French Optical Telegraphy, 1793–1855: Hardware, Software, Administration." *Technology and Culture* 35 (April): 315–47.

Field, Alexander J. 1998. "The Telegraphic Transmission of Financial Asset Prices and Orders to Trade: Implications for Economic Growth." In *Research in Economic History,* volume 18. JAI Press.

Geddes, Rick. 1998. "The Economic Effects of Postal Reorganization." *Journal of Regulatory Economics* 13: 139–56.

Goldin, H. H. 1947. "Governmental Policy and the Domestic Telegraph Industry." *Journal of Economic History* 7 (May): 53–68.

Gordon, Robert J. 2000. "Does the 'New Economy' Measure Up to the Great Inventions of the Past?" *Journal of Economic Perspectives* 14 (Fall): 49–74.

Hawthorne, Nathaniel. 1851 [1883]. *House of the Seven Gables*. Houghton Mifflin.

Hazlett, Thomas. 1990. "The Rationality of U.S. Regulation of the Broadcast Spectrum." *Journal of Law and Economics* 33 (April): 133–75.

John, Richard. 1995. *Spreading the News: The American Postal System from Franklin to Morse*. Harvard University Press.

Kleinrock, Leonard. 1961. "Information Flow in Large Communication Nets." Proposal for a doctoral dissertation, Massachusetts Institute of Technology. Available at Kleinrock's University of California at Los Angeles Internet site.

Lacy, Dan. 1996. *From Grunts to Gigabytes: Communications and Society*. University of Illinois Press.

Mowery, David, and Nathan Rosenberg. 2000. "Twentieth Century Technological Change." In Stanley L. Engerman and Robert Gallman, editors. *The Cambridge Economic History of the United States*, volume 3. Cambridge University Press.

Oliner, Steven D., and Daniel E. Sichel. 2000. "The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?" *Journal of Economic Perspectives* 14 (Fall): 3–22.

Pupin, Michael. 1923. *From Immigrant to Inventor*. Scribner's.

Standage, Tom. 1998. *The Victorian Internet*. Walker.

Temin, Peter, and Louis Galambos. 1987. *The Fall of the Bell System.* Cambridge University Press.

Thompson, Robert Luther. 1947. *Wiring a Continent: The History of the Telegraph Industry in the United States.* Princeton University Press.

Vietor, Richard H. K. 1989. "AT&T and the Public Good: Regulation and Competition in Telecommunications, 1910–1987." In Stephen Bradley and Jerry Hausman, editors. *Future Competition in Telecommunications*. Harvard Business School Press.

Vietor, Richard H. K. 2000. "Government Regulation of Business." In Stanley L. Engerman and Robert Gallman, editors. *The Cambridge Economic History of the United States*, volume 3. Cambridge University Press.

Wasserman, Neil. 1985. *From Invention to Innovation*. Johns Hopkins University Press.

Weiman, David. 2002. "Building 'Universal Service' in the Early Bell System: The Co-Evolution of Regional Urban Systems and Long Distance Telephone Networks." In Timothy W. Guinnane, William Sundstrom, and Warren Whatley, editors. *History Matters: Essays on Economic Growth, Technology, and Demographic Change*. Stanford University Press.