2 The First Two ICT Eras

The organization of the global ICT infrastructure shifted dramatically from the mid 1950s through the end of 2000. Technology and policy changes drove the shift.

In the early years, ICT was essentially two markets: a monopoly telecom marketplace and a distinct, concentrated computer and software industry centered on mainframes and mini-computers. During the 1960s and the 1970s, the growth of data networking and nascent competition in equipment and value-added services led to increased value-added competition in both services and equipment as a limited exception to monopoly. As networking matured, it gradually brought these two markets together into an integrated ICT market. This ultimately led to the breakup of AT&T (which occurred on January 1, 1984) and to the introduction of more sweeping competition and subsequent innovations in the market, particularly in the United States. The introduction to the mass market of the first model of the IBM personal computer (in August 1981) accompanied increasing competition and innovation in the telecom markets.[1] The deployment of the PC across the business and consumer landscape fueled the growth of client-server architecture, created new packaged software markets (enterprise resource planning, productivity software) and consumer uses (word processing, graphic design), and defined the architecture for a generation of devices and applications. The network's scope, its performance, and market-based networked applications continued to evolve in response to the growth of the Internet during the 1990s.

This chapter outlines the evolution of ICT markets during two distinct periods since the 1950s. The first period begins during the early postwar years and extends to the breakup of AT&T. The second period stretches from 1984 to about 2000. (The post-2000 period is discussed in chapters 3–5.)

Before plunging into the details, it is useful to mention three long-term trends in the ICT infrastructure. The first trend involves the end points on

the ICT networks: What is their number, scope (ubiquity), and heterogeneity? How many and what type of processors and data sources connect at the edge of the network? Consider the evolution of terminals. First there were voice-only dumb terminals, then there were dumb data terminals, and finally powerful networked PC terminals emerged. The second trend involves the price point for a specific speed or quality of service in ICT markets. This point determines which applications might be usefully deployed across a network. Sometimes performance levels are not available. (During the 1980s there was no commercial 256K data transport.) At other times the main issue is price and therefore widespread deployment. (Home broadband networking was too expensive during the late 1990s for many applications that were emerging in 2005.) The third trend involves the breadth of applications that are supported by the network, as determined by the processing capabilities, the location of the processing and application logic, and interoperability across the network. Mainframes were limited in their processing power and in their ability to run applications that relied on data from multiple systems and resources. Client-server architectures continue to evolve. Cable televisions running on cable networks once mainly relied on dumb data-entry terminals. But as applications increasingly run partly in "the Cloud" and partly on devices at the edge (see chapter 4), additional flexibility and resources both at the edge and in the network will be needed.

Here, two policy elements are highlighted: (1) The ICT industry gradually grew more modular since the 1950s. The 1968 Carterfone decision was especially momentous in extending modularity. It introduced disruptive new rules that allowed firms to connect equipment to the public network so long as it caused no harm.[2] The slow march toward greater modularity continues and may be accelerating. (2) In parallel, governments undertook pro-competitive policies. They increasingly embraced policy interventions that promoted competing infrastructures to enhance service competition and, also pressured competitors to embrace modularity. For example, the AT&T breakup, the IBM plug-and-play intervention, and the Microsoft antitrust case all aimed at limiting the ability of leading firms in important network segments to leverage their positions in one network element into downstream or upstream elements.

Technology and Market Evolution: 1950s–1983

The first phase of convergence of computing, software, and communications began in the mid 1950s and extended through 1983. Except in the

United States and a few smaller countries, the telecom market was characterized by government ownership of monopoly carriers with extensive cross-subsidies to achieve universal voice access. Transmission quality and throughput were optimized for voice networks, which made adding networking capabilities difficult and expensive. Until the mid 1970s, network intelligence was expensive and highly centralized in order to concentrate network switching at the top of a hierarchy of switches.[3] Network transmission capacity was sparse, expensive, and specialized. This meant that intelligence in the network was limited and that expanding intelligence was expensive and physically difficult. Early networking services were geared toward large business users and were slow. Quality voice services required 64 kilobits per second; data rates on these circuits were far slower and less reliable. As a result, networking focused almost exclusively on large business centers. Telecommunications and broadcast required separate transmission networks. Even the introduction of two new broadcast infrastructures, cable and direct satellite broadcast to the home, were dedicated specialized infrastructures. When computer networking took off, issues involving the quality, the speed of transmission, and related technical issues made the traditional networks' practices inadequate for the new data networks.[4]

This era was characterized by limited deployment of low-performance IT. Most systems had limited processing capacity and dedicated linkages across hardware and software elements. Early on, the major IT users were governments and large enterprise buyers. Even after the 1956 IBM "Plug and Play Compatible" antitrust decision partially separated the hardware and software markets, IT was mostly dominated by significant data processing applications for the largest government and enterprise buyers. The 1969 IBM consent decree finally decoupled hardware and software, opening the door open to a stand-alone software industry separate from hardware vendors.

During the 1960s, stresses to this structure emerged as the speed of networks increased. New industries appeared that sought entry into various parts of the market.[5] Rivalry for the terminal equipment on the communications network emerged in the late 1950s as large users sought specialized functions that traditional telephone networks had trouble meeting. More stresses emerged as the speed of networks increased. New industries appeared and sought entry into various parts of the market. The initial introduction of what is now called "modularity" provided the conceptual policy breakthrough that helped address potential conflicts between those intent on connecting equipment to the network and those demanding the

protection of network integrity. It quickly became evident that transparent network standards for interfacing with equipment could allow a variety of manufacturers to supply equipment and evolve new technical features.

On the computing side, the mainframe computing experience produced a growing pool of programmers who could write code independent of the big computer companies. US public policy helped drive this market evolution. Specifically, the government antitrust suit against IBM led to the decoupling of hardware and software. This promoted the take-off of an independent software industry featuring packaged software,[6] a software industry quite different from the one associated with the PC industry. Still, this development started to erode IBM's dominance and contributed to the move toward modularity in computing hardware and software.

Changes in network performance and the emergence of a stand-alone software industry were important, but the most disruptive development during the 1960s and the 1970s was the rise of computer and corporate networking. Networking opened new markets for firms, sparked new demands from users, and required new policy responses. Policy makers recognized that the status quo was no longer sustainable. On the telecom side, new rules made it easier to attach terminal equipment to the telecom network, liberalized entry into data networking, and allowed private corporate networks for voice and data services. In services, the new entrants slowly undercut AT&T's dominance in long-distance and data transmission facilities and services. Prices responded; service innovations followed.[7] The United States was the exception during this period. Most of the world's markets were dominated by vertically integrated, government-owned firms with close ties to vertically integrated equipment providers.

The following is a summary of what happened from the 1950s through 1983:

- The number, ubiquity, and heterogeneity of network end points accelerated as PC connections to the Internet proliferated and as voice and data mobility spread.
- The price for services of comparable quality and speed declined sharply. The decline in cost structures spanned applications and services.
- The breadth of applications supported by the network increased substantially.

Technology and Market Evolution: 1984–2000

The second phase of convergence of computing, software, and communications began with the breakup of AT&T in 1984 and extended through

2000. The gathering momentum of the microprocessor revolution for personal computing, competition in communications networking, and a second generation of computer networking architecture shifted the market horizon again. By the mid 1980s, the semiconductor industry began to enable deeper network architecture changes and revolutionize ICT devices' power at the edge of the network. Telecommunications switching grew more sophisticated, but this happened more slowly than intelligence could be incorporated in computers and other devices operating at the network's edge. This "flipped" the logic of network architecture even as Moore's Law took hold and the spread of PCs in business and consumer arenas created new demands for networked applications and services.

The telecommunications market was characterized by the gradual but forceful introduction of competition in all infrastructure, hardware, software, and services segments. Two important consequences were the build-out of narrowband dial-up networking in the consumer marketplace and the beginning of broadband to the home. Dramatic improvements in the capacity and cost of lasers and electronics and the explosion of data traffic they prompted led to the build-out of backbone fiber and broadband to more business users. Another result was the beginning of metropolitan fiber networks and broadband consumer networks. Transmission capacity expanded dramatically, from snail-paced modems dripping out data at 128K to the T3 (45 megabits per second) connections that became routine for larger enterprises.[8]

Another major development was the explosive growth of mobile wireless. In developing countries mobile wireless connections rapidly overtook wireline connections when the introduction of second-generation ("2G") systems greatly upgraded capacity and quality while reducing costs. By 2000, mobile communications had emerged as a vertically integrated competitor to the wired network in all market segments except for data.

The Internet and its commercialization also were hugely important. The Internet revolutionized the architecture and underlying capacity of the network. The beginnings of inter-networking dated from the mid 1980s (Cisco shipped its first router in 1986), when companies and network providers began to "inter-connect" their networks. In 1991 US policy changes enabled the commercial use of the Internet. This set the stage for the ICT growth of the 1990s. By 1994, the Internet swamped commercial email services. In August 1995, Netscape went public, igniting the "dot com" boom. In the United States, and to a limited extent elsewhere, new Internet services providers (AOL, MSN) and later large content and e-commerce applications (Yahoo, @Home, eBay) aimed to take advantage of the

network's power and scope. A myriad of smaller, more specialized applications also emerged that built their businesses on powerful, cheaper PCs, broadband networking at the office, and widespread narrowband networking in the home.

The burgeoning PC market, advances in the PC's operating systems, and the growth of networked enterprise computing supported the development of new, packaged, mass consumption, software applications and attracted enormous investment in, and innovation around, PC-based software.[9] Declining price/performance ratios spurred widespread deployment and adoption of vast enterprise software packages to manage enterprise-wide processes and data. Packaged software for PCs opened the way to greater complementarity of software products, particularly between the Microsoft software platform and specialized software applications. This strengthened Microsoft's position by creating a new set of hardware and software industries focused on the PC ecosystem, from mice to games to semiconductors. The emergence of the Internet and in particular a new PC application used to "browse" content and services, reinforced the client-server architecture that dominated enterprise architectures.

In the mid 1990s, serious challenges began to undermine the existing technology, economics, and policy equilibria. Technologically, the growth of Internet standards, data protocols, and Application Programming Interfaces (APIs) outside the control of any single platform vendor created momentum for more open APIs. On the PC, Microsoft defined the APIs that other applications used to interact with Windows. Microsoft's power provoked strong opposition, which led to intense commercial rivalries and disputes. From the Microsoft litigation an important legal right emerged that allowed software developers to reverse engineer software interfaces to create complementary and substitute software.[10] Limitations on the extent of Microsoft pricing flexibility across original equipment manufacturers and the requirement that Microsoft publicly share terms of OEM agreements were related and equally important parts of the Microsoft antitrust settlement. This limited the ability of Microsoft to "punish" OEMs for inclusion of competing software on Windows machines or for shipping non-Windows computers.

The emergence of the Internet provided Tim Berners-Lee with the base from which he launched a suite of software applications—now known as "the World Wide Web"—that further altered these dynamics.[11] HTML, the programming language that enabled the Web, consciously avoided the Microsoft approach and embraced open APIs. Netscape's Web browser and the subsequent inclusion of Microsoft's browser in Windows sounded the

death knell of Internet Service Providers (ISPs) that forced consumers to rely on proprietary software systems to access the Web.[12]

Another major change was the quiet but fundamental transformation of the ICT production system. Traditionally, vertically integrated firms in both telecom and IT delivered complete systems (hardware, software, integration services, and support) to customers. By the late 1980s, international challenges from Japanese electronics vendors and the growth of the software industry created pressures[13] and opportunities for vertical disintegration, commoditization, co-competition in equipment and services, and a dynamic of user co-invention of breakthroughs.[14] The breakup of AT&T began the dynamic vertical disintegration of the telecommunications network into local and long-distance services. In the 1990s, the advent of a new class of specialized fiber-optic transport networks, of which Level 3 was the most prominent example, segmented the market further.[15] Forces for commoditization and competition augmented those of vertical disintegration. Barriers to entry generally declined and global production networks increased the universe of potential entrants in one niche after another.[16] Speed and declining barriers to entry meant that the life cycles of products became shorter and the ability to maintain premium pricing declined rapidly for most products. Demands from sophisticated IT and telecom users also began to set the agenda that new vendors scrambled to meet. They illustrated the forces of co-invention by users of digital technology.[17] The evolution of more flexible and less expensive modular systems made it easier for users to innovate in critical technologies directly or by working intensively with suppliers. The rebellion of the office floor against centralized computing proved a springboard for local area networking of desktop computers. The growth of the Web browser and the Web opened a mass consumer market. Amazon, eBay, and others introduced another set of complementary users and vendors built around e-commerce. This dynamic played out first and proceeded furthest in the United States, but other countries moved down the same path.[18]

Meanwhile, after several fruitless efforts to mandate standards for computer networking, Western Europe reluctantly made plans for wide-ranging competition in the wired network infrastructure. The cost efficiencies and technology and service innovations that occurred in the United States eluded Europe.[19] With the notable exception of Finland, most of Europe did not introduce general wired network competition until 1998. Mobile competition (usually in the form of duopoly) sprang up earlier, but few in Europe believed that this limited competition would have major implications for the wired network.

Broadband networks for households became common in Asia, in Europe, and in North America during the late 1990s, causing many countries to rethink their policies. The crucial point in broadband deployment was the determination of most countries to close the gap with the United States in Internet connectivity (using telephone dial-up modems) and to leapfrog it as broadband deployed. This is precisely what occurred. In mid 2007, the top five world leaders in fast, affordable broadband networks were Denmark, the Netherlands, Switzerland, Korea, and Norway. In June 2007, the US ranked 15th globally in broadband Internet penetration. From June 2006 to June 2007, the number of broadband subscribers in OECD countries grew by 24 percent, from 178 million to 221 million.[20] These shifts were caused as much by policy and politics as by the technological decisions discussed in later chapters. The same dynamics almost certainly will drive broadband for wireless and mobile services. Historic broadband penetration from 2001 to 2007 for the OECD countries as a group and for the major countries is tracked in figure 2.1.

The Political Economy of Marketplace Change in the United States

At the core of our argument about the political economy of markets are political institutions and their impact on the incentives and authority of elected politicians to shape marketplace outcomes to the advantage of specific sets of constituents. In view of the importance of the United States in global ICT markets and the centrality of the American institutional argument for later chapters, this section sketches our institutional argument in the context of the first two ICT eras in the US.

The American political system has three salient features relevant to communications policy: the division of powers, the majoritarian electoral system, and federalism.[21] First, the division of powers in the US government was designed to make it difficult to initiate major policy changes but also difficult to rapidly undo major commitments. The division between the president and Congress (and between the two houses of Congress, one based on population and the other on equal representation for each state) creates many points during the decision process at which an initiative can be stopped.[22] This hampers the passage of major changes in laws that have sweeping geographic consequences and a wide range of winners and losers. Only two major US telecommunications laws were passed during the twentieth century: one in 1932 and one in 1996. Thus, much of the decision making about federal policy resides at the Federal Communications Commission, which is charged with implementing the acts.

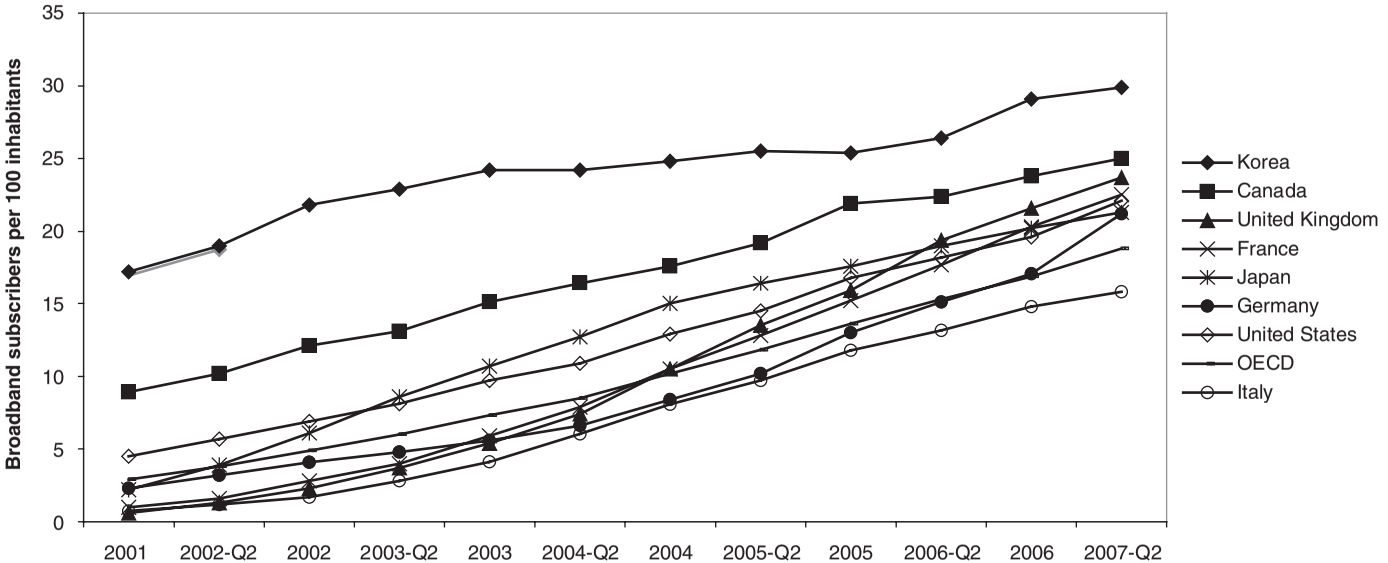

Figure 2.1
Broadband penetration, historic G7 countries plus Korea. Source: "OECD Broadband Statistics to June 2007," Organization for Economic Cooperation and Development, October 13, 2007.

The inherent conflict between the executive and legislative branches means that Congress is less willing to grant the kinds of discretion to executive bureaucracies that are found in parliamentary democracies, where the division between the executive and the legislature is more blurred.[23] In some areas Congress recognizes the need for a substantial amount of expert bureaucratic authority. Thus, the FCC is allowed to deal with complicated issues where many of the detailed political costs and benefits are difficult to determine. Congress then uses a variety of devices to delegate power to the bureaucracy with specialized controls.

Congress confirms FCC commissioners nominated by the president and stipulates a 3–2 split of commissioners, with the president's party holding the majority and the chairmanship. The political sensibilities of the major parties thus are replicated in the FCC's majority voting system. (Commissioners are political appointees, but usually are more technocratic and less ideological than members of Congress.) Congress also uses the power of the purse by threatening to use its budgetary powers to instruct the FCC on certain issues—for example, barring the FCC from using public funds to create rules to auction communications satellite spectrum. Similarly, it mandates elaborate FCC procedures to ensure transparency in decision making so that all interested parties will have access to the decision process. Members of Congress can observe the process with an eye to politics, and can try to influence FCC if there is a compelling political interest.[24] These complexities constrain the FCC's flexibility and complicate its ability to focus on competition problems when they arise. Thus, when such problems are identified, the FCC relies more on elaborate "ex ante" rules than on ad hoc solutions. The net result is that the FCC responds to presidential and congressional politics but is legally empowered to make important discretionary policy. It is subject to judicial review for its adherence to a reasonable reading of the underlying law. It bases its decisions on its analytic judgment, the evidence on the public record developed in each procedure, and an instruction to use this discretion to serve the public interest. These expert and transparent but politically informed decisions influence market dynamics.

A second feature of the US political institutions is that presidential and congressional elections are based on winner-take-all voting. Analysts of electoral systems have shown that this voting system builds a strong interest in "brand identity" for political parties. Despite the role of lobbying and campaign contributions, parties invest to develop a reputation with policy initiatives on broad issues that they believe will mobilize elite and mass electoral support.

Telecommunications policy traditionally influenced the high technology industry and research communities. It achieved broad political salience to the voting public in two ways: (1) To promote equity, there was continuing sensitivity to telephone pricing, and now broadband pricing. (2) It was part of the broader debates over economic policy, including the debates over "deregulating" the American economy and the creation of the "new" or "Internet" economy to stimulate growth. Thus, the Clinton administration highlighted its telecommunications policy to polish its reputation as pro-competition and pro-innovation.[25] It bragged about early US leadership in the mass adoption of the Internet. Similarly, the George W. Bush administration worried about the potential embarrassment of America's lagging position on deployment of broadband.

A third feature of the institutional context is federalism, the division of authority between the federal and state governments. The US Constitution reserves all powers not given explicitly to the federal government for the states. Moreover, each state is allocated two senators, regardless of its population. This increases the power of low-population farm and mining states at the expense of heavily populated, urbanized states. Federalism matters for telecommunications policy directly and indirectly. It directly impacts the subsidy of rural communications users and providers, which powerfully constrains all regulatory policies influencing pricing and competition. Federalism indirectly provides a foundation for strong competition policy. State authorities used competition policy to shelter local competitors from dominant national competitors that held various advantages over them and smaller firms would enlist the support of their senators. The pivotal role of rural states in the Senate also heightened interest in competition rules that emphasized consumer welfare because those states have less interest in industrial policy favoring national champions.[26] The result was an economy with broad geographic scope for its competitive firms and far less concentration in its major industries, including telecommunications and electronics, than its counterparts in other major countries.[27] The US also had a telecom market whose behavior was skewed by a pricing structure that bore little relationship to efficient costing. The implications for telecommunications policy were profound.

The Political Economy of the First Era (1950–1983)

As was demonstrated many times between the 1950s and 2000, even with divided powers, policy evolution can move quickly if economic interests and political institutions are aligned. In this era, the United States was by far the largest ICT market and its economy had a continental scope. As it

became evident that there was strong political support for policies strengthening competition, the scale of the American market allowed pro-entry policies to open the market to new entrants while simultaneously retaining market scope for incumbents. New services became available for large enterprise users that were deploying ICT to help enable new production and service processes that they needed to stay competitive nationally and internationally. Large enterprise buyers also began using long-distance telephony to increase branch coordination.[28] This produced a group of large potential customers concerned with the market's organization among sophisticated firms and guaranteed an environment favorable for political entrepreneurship. Thus, policy changes unfolded faster in the United States than elsewhere.

The role of large users mattered because they were transformed by ICT and intensified their policy advocacy. Eventually, ICT became more than a cost factor for US-based multinational firms. In response to rising competition, US financial institutions and many manufacturing firms evolved into information analysis companies that deliver information in the form of, for example, a global financial or engineering product. Global banks no longer focus mainly on checking or mortgages. Their edge comes from complex and ultimately riskier financial products that rest on high levels of computing and global information operations that are rolled out quickly on a global scale over their ICT infrastructures. Multinational manufacturers understand that the cost and quality of production are important, but the information intensive, global design and service functions are their critical edge. Boeing executives sometimes joked that an airplane is a flying software stack because there is more value added in the complex programming than in the sophisticated airframe.[29] This fundamental shift in the strategic use of ICT persuaded these firms to become committed advocates for changes in ICT markets.

The political institutional legacy of the American market structure shaped the way that emerging interests played out. No firm legal basis for AT&T's national long-distance monopoly existed, and many smaller telecom carriers remained in small states and rural areas. This lacuna arose because it always proved politically difficult to craft legislation to authorize (or, later, to preserve) a monopoly. In addition, previous antitrust actions created huge electronics firms that rivaled AT&T and lusted to supply equipment to American telecommunications networks. In 1956 their limited antitrust victory forced Bell Labs to license its technology to them at little or no cost. Meanwhile, federal power sharing with the states with regard to telecom pricing and a sympathetic Senate restricted AT&T's ability to lower

long-distance costs so that the transfer of funds to support smaller rural carriers could continue.[30] AT&T offered special discounts to large corporate customers, but could not offer true cost-related pricing. Thus, large customers continued to seek market change.

The growth of computer networking, especially by IBM's smaller rivals, created another powerful set of motivated allies that were unhappy with AT&T's dominance.[31] An IBM "plug compatible" industry grew up that targeted the networking market. This led directly to the formation of a "corporate competition coalition" made up of computer companies that wanted to create customized computer networks or feared AT&T's entry into the computer equipment market. The computer companies were joined by large corporate clients, smaller electronics equipment vendors, would-be resellers of basic phone services, and government agencies, all seeking better deals.[32]

Increasingly, governance was guided by a new principle: modularity. It became common to distinguish among "basic phone services" provided over the general public network, the equipment that enabled it, and new advanced communications and equipment functions made possible by new electronic and computing technologies. Momentum grew to competitively deploy new "value-added" services and equipment.

Four important norms emerged that enabled greater modularity. The roots of the first norm came in 1956, when the limited liberalization of attachment of terminal equipment was allowed. Twelve years later, the Carterfone decision opened the way toward full freedom of competition in equipment attached to the network by creating the first norm to implement modularity. The FCC held that new equipment attached to the network was acceptable if it did "no harm to the network." The FCC recognized that the demands for computer networking required less restrictive equipment markets.[33]

The "no harm to the network" norm implied a freedom of choice that grew into a second norm, technology neutrality that resonated with US political and market institutions. The US rarely picks civilian technology champions. Its diverse economy usually does not generate political agreement on a single technology path for any market. Further, by the 1980s US policy makers questioned whether they could readjust their direction if they chose the wrong technology path. For these reasons, neutrality seemed a sound policy norm with political merit.[34]

At the same time, the FCC lurched toward allowing competition in the provision of networked computer data services. In 1959, AT&T Long Lines established a discount rate for its largest corporate and government

customers that was a reasonable proxy for a wholesale price for the leasing of transmission capacity to the new data networks. The FCC embraced this benchmark when it forced AT&T to lease capacity at a wholesale price to computer networks.[35] In doing so, the FCC embraced for narrow purposes what eventually became a third norm of modularity. The government mandated that network operators with market power must share their network capacities with competitors. Eventually a fourth norm developed that held that the deployment of value-added, competitively provided services should not undermine the basic pricing and service structure of the general public network. By linking prices for sharing the network to an established rate, the FCC laid the basis for skirting the political issues raised by monopoly pricing.

When a monopoly exists, government pricing intervention is complicated because it is difficult for the monopolist to differentiate prices among different classes of customers whose elasticity of demand varies. It is also a political swamp because every interest group makes special claims about rates. Politicians wanted pricing favorable to household consumers, especially in rural and low-income areas. This clashed with the network's cost structure because costs were higher in rural areas where longer cable transmission distances supported fewer customers. Moreover, in view of the large common costs of networks, such as billing systems, the attribution of costs to different services and areas involved creative, albeit government dictated, accounting. In general, the pricing formulas caused denser urban areas to subsidize rural areas, long-distance customers to subsidize local service users, and businesses to subsidize individual users.

Anchoring the AT&T price for computer networking to existing pricing for large customers was politically reassuring for political leaders because data services were added to a pre-existing rate compromise that AT&T had promised would not upset consumer pricing.

Sidestepping major pricing reform also opened the wedge for allowing "private networks" to connect geographically far-flung firms' offices with capacity leased from AT&T at wholesale rates. MCI applied for permission to provide specialized corporate services over its own microwave network in 1962 and in 1969 won approval for its first link, between Chicago and St. Louis. When the FCC generalized this decision in 1971, only about 3 percent of the total Bell system revenue was at stake.[36] The FCC also allowed private line carriers to interconnect with AT&T facilities. Predictably, the battle over the terms of interconnection led to MCI, and later to Department of Justice suits that culminated in the decision to divest AT&T, which took effect on January 1, 1984.[37]

Overall, the first era of policy change introduced the principle of modularity. At first it helped create limited competition in "value-added" information services, private network services, and competition in terminal equipment. The entry of the computer industry into the telecom policy realm was a difficult transition. Computer vendors and their largest customers wanted to network expensive mainframe computers to allow more efficient cost sharing and operations. Major technology suppliers and large network users pushed for policy change. The consolidation of these two discrete industries was helped along by the debates over the terms for equipment competition that began in the late 1950s and over the leasing of network capacity for the new computer networks and over private corporate services that emerged later.

The Political Economy of the Second Era (1984–2000)

The breakup of AT&T and the introduction of competition in the long-distance services and network facilities markets was the breakthrough event that sparked the global reorganization of the telecommunications industry and then revolutions in computing and broadcasting. The emerging "managed entry" governance rested on the idea that incumbents often might use essential bottleneck facilities to manipulate the market to the detriment of competition. Regulating the shared use of a monopoly infrastructure seemed complicated and unlikely to create innovations in infrastructure that might emerge from a networked designed from scratch. Still, no entrant was likely to roll out a national network quickly, thus diminishing the value of network externalities (more connections make a network more valuable) for its customers. Thus, a second organizing principle for market governance emerged: Encourage the emergence of competing network infrastructures by removing legal barriers to their creation and by forcing the dominant incumbent to share its network with its rivals. This turned into a governance system of managed market entry. In the United States this meant extensive regulation of "dominant" carriers; in other countries it often took the form of controlling the number and qualification of entrants.

Why did change occur as it did between 1984 and 2000? Why did change appear first in the United States, and what were the implications for global arrangements? If competition was driven mostly by technological and market forces, why did it unfold so differently in the US? Why did other industrial countries resist and lag behind?

In the United States the combination of slow economic growth and high inflation in the late 1970s raised deregulation of public utilities onto the

presidential and congressional political agendas. Political parties strive to be national policy entrepreneurs. Democrats and Republicans both saw deregulation as a way to show their commitment to revive the American economy.[38] The political economic interests of the corporate competition coalition reinforced their enthusiasm for deregulation.

Renewed antitrust action during the Carter administration set the stage for the breakup of AT&T during the Reagan administration. The decision reflected American political institutions. First, the courts followed established antitrust law that arose from the US political economy and favored a consumer-welfare standard. America's federal system produced this legal approach and a court system with the latitude to back it up. Second, the president and Congress cannot easily take decisive legislative action to steer an industry because it can be blocked at many points. This structural factor sidetracked AT&T's attempt to legislatively assert it monopoly, repulsed increasing pressure from MCI and other upstarts, and convinced a generation of entrepreneurial politicians that identification with AT&T's critics was politically advantageous. Even the Democratic Party, predisposed to supporting organized labor and therefore a likely ally of AT&T and its union members, spawned a generation of "Atari Democrats" critical of monopoly. Economic conservatives in the Republican Party joined them. This coalition sufficed to block pre-emptive legislation to preserve the phone monopoly. Third, although the president and many in Congress were wary of the AT&T antitrust decision, they did not try to overturn it because they saw it as politically risky to favor monopoly.[39] Fourth, the settlement made sense because it could withstand political pressures to protect incumbents before and after the AT&T breakup. The long-distance competition by the new AT&T and monopoly phone services for the new regional Bells mandated by the court protected both local and rural telephone service pricing. The FCC and state public utility commissions could mandate cross-subsidies from long-distance carriers to local phone monopolies and still allow competition to improve services and lower long-distance pricing. Lower long-distance prices appealed to the middle class that tended to vote more than other Americans. Because it did not unwind local subsidies quickly, network competition also appealed to the corporate competition coalition by providing a strong, politically sustainable competition platform.[40]

The outline of a new managed-entry regime that would dominate the United States and then prevail globally emerged from the struggle over the fate of AT&T. The principle of favoring competitive network infrastructures led to the extension of the earlier norm that forced dominant networks

that controlled essential bottleneck facilities to share their capabilities with new rivals to promote rapid industry entry. This required detailed FCC supervision of interconnection by dominant carriers. The challenge was to police against bottlenecks in a way that allowed market forces to rationalize costs, staffing, and prices. Tools such as price caps and dominant carrier regulation were designed to foster pro-competitive interconnection with new entrants and allow pricing rationalization.

Three norms supplemented this competition principle and made it politically practical. First, regulators should adjust prices of local services, without allowing rapid price escalation. Competition had to be reconciled with this goal. Second, to cash in on the political promise of competition, regulatory reforms should promote technological and service innovation for ICT, including lower prices. Economic theory argued for maximizing consumer welfare. This norm clarified what political leaders meant by "consumer welfare." Third, policy makers should be sensitive to employment effects. They could allow labor staffing to decline in dominant incumbents, but needed to cushion job losses by encouraging the entry of new companies which might offset the downsizing of old incumbents.

This mixture seemed politically successful. Prices for long-distance and data services decreased significantly. Service innovation climbed. Initially, computer networking rose and then soared as the importance of the Internet spurted. Politicians could boast that new entrants helped revive American fortunes in the computer and computer networking equipment markets. But trouble was brewing.

The push for technological and economic efficiency ultimately raised two issues. The first issue was purely a product of technological innovation: How should the potential for mobile wireless networks be used to boost competitive network infrastructures? Second, what role should the Bells play? Why should they be barred from entering the long-distance market when their entry might further reduce prices? But how could complete network innovation and competition be achieved in the absence of contestable markets for local communication services and infrastructure? This huge market still wore a regulatory straitjacket.

Originally, mobile services were offered as a duopoly in the United States and most other industrial countries; invariably the local phone company received one of the licenses.[41] The introduction of second-generation wireless services in the 1990s permitted more competition. More competition promised the political benefit of better prices and services for consumers. The largest telecom carriers and equipment suppliers sought lush new growth markets. And, to the hidden relief to all involved, mobile services

still seemed a luxury good that would not significantly substitute for wired voice services. So mobile seemed an area ripe for political profit and innovation. But how?

Policy experts in both political parties favored auctions to more efficiently assign spectrum licenses. The principle was that a properly organized system of market bidding provided more accurate cues for assigning and valuing a scarce national resource (government-controlled spectrum) than discretionary decisions by government officials. This option was embraced because auctions would be easier for new entrants, which might be less connected politically before the auction, but would be grateful afterwards. National spectrum auctions also promised to reduce the federal budget deficit by raising large sums of money. This was a goal of both political parties, of the president, and of Congress.

When the FCC designed its auction system, it envisioned obtaining four to six competitors in every American market.[42] The FCC reasoned that if four to six competitors each had enough spectrum to support a significant network and service offerings, none could dominate. Although continued scrutiny of the interconnection of wireless with wired networks might be necessary, regulators expected that the interconnection rules for wireless networks could be much lighter than those for wired networks. Uniquely, the FCC mandated very low wireless-wire interconnection charges. Only the United States had a multi-firm market and low fees. Other nations slowed the growth of wireless by imposing high fees wireless paid to wire. (The EU imposed high fees, but offset them with limited wireless competition that let wireless carrieers flourish financially.) These differences mattered when there were few wireless customers and almost all their calls went to wire. Now wireless talks to wireless, and this starter move matters less than it once did.[43] Thus, wireless presented a glimpse of what ICT markets after the end of dominant control of bottleneck facilities might achieve.

The other important wireless choice involved technology policy. As with computing and terminal equipment for wired networks, on wireless the FCC adopted a norm of technology neutrality. The deployment of multiple architectures resulted. Although the timing varied by market segment, the cost of diverse architectures caused some confusion and delay in deployment of features requiring mass scale. This tracked exactly earlier computer industry developments. Originally, the United States trailed other countries in this field. Eventually, after a shakeout, US reforms led to increased technological innovation and experimentation with equipment, software, and service application mixes and some closing of the gap on wireless with

Europe and Asia. Pricing is much lower and volume usage for voice and data much higher in the US than in the European Union, for example. But penetration remains lower.

Meanwhile, all agreed that the Internet and the Web would lead the next boom in communications and IT investment. The major corporate players wanted to be ready. The bargain leading to the 1996 Telecommunications Act was struck between a Republican Congress and a Democratic White House, each of which had reasons for wanting to reach an agreement.

Predictably, the politically muscular regional Bells, which operated in every congressional district, wanted permission to compete in all markets. Republicans sided with the Bells because their strength was greatest in the West and the South, where the Bells were especially influential, and because the Republicans had won control of Congress in the 1994 election. Most Democrats, including the president, depended on a strong urban base and lined up with the long-distance carriers that had cultivated ties to large urban users and the computer industry.[44] The long-distance companies recognized that pressures for Bell entry were enormous, but they counted on the Clinton administration's support on the terms for their cross-entry into local services. The White House did so; however, Democrats also were re-branding themselves as the pro-competition champions of the information economy, and they did not want to oppose allowing the Bells to compete.[45]

During the legislative bargaining, the Bells rejected a deal that guaranteed them entry into the long-distance and data markets 3 years after passage of the act. Instead, they opted to meet a "check list of obligations" that allowed them fully into long-distance and data only after they demonstrated that their territories were open to local service competition. They made this choice because they believed, wrongly it turned out, that congressional pressure on the FCC would help them gain entry in less than 3 years. However, the Democratic FCC, with strong White House support, interpreted the act to call for strong interconnect obligations for the Bells at long-run incremental costs. This formula enraged the Bells and the Republican Congress.

Many economists, wary of major government regulation, worried that the FCC's terms for interconnection might discourage investment by the Bells and induce inefficient, subsidized entry that rested on the Bells' unrealistically priced facilities.[46] The Bells launched a full-scale legal counterattack on FCC rules. Because American administrative bureaucracies enjoy less latitude than their counterparts in parliamentary democracies, court challenges tied up portions of the interconnection regulation. Still, market bargains were struck because the Bells wanted to claim that they had fulfilled the 1996 act's checklist.

In the later 1990s the emergence of new ICT competitors and the Web bubble led to a huge investment boom in fiber-optic networks. By late 2001 the boom had fizzled and most Web start-ups had crashed, but across the United States and under the Atlantic and Pacific Oceans a huge new installed infrastructure remained. This infrastructure helped kill the pricing structure for traditional long-distance carrier's voice and data transport offerings.[47] It also prompted the US government to exempt Internet "backbone" traffic from regulatory scrutiny, thereby creating an international controversy. Only a proposed merger of MCI and Sprint, then two of the largest backbone providers, prompted regulatory intervention to ensure that competition in the backbone market was not substantially curtailed.

The introduction of infrastructure competition in telecommunications raised worries that incumbents might leverage their control of bottleneck facilities and led to more detailed governance to manage market entry. The same concerns soon extended to the mass consumer infrastructure for networked information technology when email and the Web emerged as a high-profile focus of technology politics and policies.

Until 1994, the ICT infrastructure relied on proprietary email systems (such as MCI Mail and AOL) and computer network formats (such as IBM's System Network Architectures protocols). There was some grumbling about the lack of interconnection of these proprietary, "value-added" services, but this was still a small market for large institutional users and a relatively small technophile population. The proliferation of the Web escalated the commercial stakes and attracted political attention. The Web proved transformative because its simple path to interconnecting all networks quickly overwhelmed existing formats underlying uncompetitive "walled gardens" for data networking and email.[48]

The story was different at the intersection of networking and desktop computing. The Internet also transformed computing and software strategies in the marketplace thereby focusing attention on the logic of market governance built on vertical integration and the control of market power from bottleneck facilities. Thus, in theory Microsoft might leverage its PC operating system (a bottleneck facility) to unfairly enhance its competitive Internet position at the expense of competition and consumer welfare. Worries increased that Microsoft would use its Internet browser packaged with Windows to promote its own software add-ons and content. The political economy logic of the Microsoft antitrust action tracked the history of US electronics policy. Many rivals located outside the Pacific Northwest began a campaign to capture the attention of state and federal authorities. The same issues were raised over the AOL-TimeWarner merger.[49] Although

in that case the operational remedy was restrained, the basic issue of leveraging bottleneck facilities was identical because the US political logic then favored such government decisions.

Change in the broadcasting arena preceded more slowly during the first two policy shifts. Traditionally, broadcasting embraced separate networks (optimized for point-to-multi-point transmission). It was subject to specialized regulation that was shaped by vigorous lobbying (broadcasters controlled congressional members' access to their district's airwaves) and by fierce voter demands for television services at "affordable" prices. The political and cultural sensitivity of broadcast content reinforced the level of arbitrary regulation. The United States maintained a public interest standard for broadcasting that purportedly protected the public interest no matter how difficult that was to define or enforce.[50] Other countries had the added burden of broadcasting rules that tried to protect national culture through various content quotas.

The emergence of cable television as a rival platform also was of great significance for the ICT infrastructure. Cable began as a series of locally granted franchises and quickly won legislative favor as a way of delivering television to rural areas or urban areas where there were reception problems. The industry profited from the same antitrust legacy that shaped telecom policy when, in 1953, the Department of Justice forced RCA, the dominant equipment supplier for cable, to divest itself of network holdings. (The Department of Justice made an ownership share into a condition of supply.) Finally, in 1984, as access to cable became a popular grassroots issue in both Republican and Democratic districts, Congress passed a bipartisan Cable Act that systemized the terms on which towns and cities could grant cable franchises, ended local price regulation, and banned the Bells from purchasing cable systems. This propelled the growth of cable operators around the country, but especially in the West. The legislative leader was Representative Tim Wirth, a Democrat from Colorado.[51]

As cable became a powerful industry with revenues far exceeding those of the three large broadcast networks, it also sparked consumer ire. When prices climbed rapidly and service was undependable, two-thirds majorities in the House and the Senate passed the Cable Rate Act of 1992 and overrode President George H. W. Bush's veto. The act capped cable rates, insisted that cable make its programming available to its broadcast satellite competitors, and stipulated that cable had to pay for retransmitting broadcast programming. (The cable operators often "paid" broadcasters by agreeing to carry their new cable networks.) Despite this setback, cable's technical infrastructure had the potential for providing broadband to the home but

needed massive capital investment to upgrade it. The cable industry's entrepreneurial leadership mixed financial acumen with a poker-playing style as it played off major IT companies (e.g., Microsoft) and telecom companies (e.g., AT&T) to fund its investment model. (AT&T's investment in TCI ended up costing it dearly.) Eventually, cable emerged as a rival network platform for home data services. This revitalized the industry.

The growth of US cable television and satellite broadcast networks also began to fragment the broadcast markets into numerous specialized channels and market niches. Mass audiences began shrinking. This set the stage for a restructuring of the content industry after 2000.

Parallel Changes around the World

As US policy change progressed, parallel changes were underway elsewhere. Usually changes originated first in the United Sates, but not always. A significant exception, discussed in chapter 8, was the takeoff of the mobile wireless infrastructure more rapidly outside of the United States.

The analysis of trade policy in chapter 7 examines the critical role of US bargaining with the world over the introduction and consolidation of ICT transformation in the late 1990s. The US sought two global changes. In the first era it wanted to extend internationally the competitive provision of value-added networks and the creation of private corporate networks (internal communications). It also promoted policies similar to Carterfone to allow modularity for terminal equipment attached to the telecom network. Germany and the US had spirited, sometimes bitter negotiations over these goals. The US also began pressing Japan to open its international value-added networks to greater competition, a crucial wedge for US multinational firms operating there. The idea was that value-added competition in Japan would boost IBM and other US computer firms that were struggling against a Japanese industrial policy determined to overtake America's lead in semiconductors and computing. Eventually, these bilateral and regional (e.g., NAFTA) negotiations moved to the multilateral trade level.

After the decision to break up AT&T, the US government began to preach the virtues of facilities-based competition.[52] This caused stakeholders elsewhere to revisit their own political economic calculus.

We call the period 1984–2000 an era of "managed competition" because during that time the United States allowed unlimited entry in long-distance but, until the Telecom Act of 1996, retained a monopoly on local phone services. Even then, it micro-managed its markets by implementing detailed regulations that addressed the market power of the dominant

carriers (AT&T and the Bells). (The FCC declared AT&T dominance on long distance to be over in 1996.) They were forced to share their networks so that new entrants could rent detailed technical elements.

Other countries introduced their own competition schemes, but few of them went as far as unlimited entry doctrines of the United States. The timing varied substantially. The United Kingdom, Canada, Australia, New Zealand, Hong Kong, Singapore, Japan, and South Korea followed closely on the heels of the US. Except for mobile wireless, until the official liberalization of European telecommunications on January 1, 1998, the general policy of many EU members was to experiment with value-added competition. Even the countries that allowed facilities-based competition approached managing competition in different ways. For example, Japan limited the number of entrants in network facilities to minimize "disruptive" competition that might endanger the incumbent.[53] This was a popular solution elsewhere too, as in Britain's late 1980s duopoly policy. Many countries also divided the domestic and international markets, and Japan and some other countries maintained elaborate controls on pricing to make sure that all major players showed profits. Others remained committed to active industrial technology policies even after introducing competition.

In addition to the policy changes on market entry and pricing in the 1980s and the 1990s, many advanced economies began separating government from market supply, by fully or partially privatizing their telecommunications industry. Slowly, they also began to substitute arms-length government rule making for management of the market by the former monopoly carrier. In the negotiations that led to the 1997 WTO Basic Telecom Agreement, countries that had recently adopted such changes—worrying that this process easily could go wrong—enshrined the creation of independent regulators in the WTO accord.

The changes in telecom were far more sweeping than those in broadcasting. In broadcast, most advanced economies allowed limited entry for broadcast satellite services, but there was no generalized entry policy. The fate of cable television franchises was uneven. Both satellite and broadcasting changed the economics of media markets by creating options that expanded and fragmented the broadcast channel universe. A more profound change occurred in countries with extensive growth of cable television because it could be upgraded to handle other services, especially broadband data and telephony. Cable emerged as the only credible local infrastructure platform for wired networks fighting entrenched phone companies. During the 1990s, a major divide in national networking emerged between countries that evolved a relatively ubiquitous cable

platform and countries that promoted the entry of new broadcast satellite systems for television. The United States did both. In Britain, and in a few countries where policy makers actively encouraged cable television-telephony competition, a robust cable television infrastructure emerged.

Summary

Market governance for "managed entry" began with the breakup of AT&T in 1984 and was reinforced in the mid 1990s by the emergence of the Internet and the Web. Initially government intervention helped facilitate this new approach to data communications. The idea was to build market competition by controlling legacy and new essential facilities. This was most straightforward in communications where a former monopolist controlled hard-to-duplicate facilities that new entrants wished to rent so they could compete. In broadcast and cable television the usual practice was to license only one competitor. Selective entry was introduced in market segments such as broadcast satellite and cable. No generalized entry policy was implemented in some important market segments, especially broadcast. But in the newer digital ICT age, more market developments were tied to major antitrust cases (IBM, and later Microsoft and Intel) involving the control of an important technology platform.