1

The Big Picture

The word *telecommunications*, a twentieth-century amalgam of Greek and Latin roots, literally means the art of conveying information "from a distance." For millennia, people had to rely on messengers to perform this task, which was as costly per message sent as it was time-consuming. When the Greeks repelled the Persians at Marathon in 490 BC, the legendary messenger Pheidippides could not shout the good news back to Athens, for it was 26.2 miles away, nor could he call anyone up, for there were no telephones; instead, he had to run. Several hours later Pheidippides arrived in Athens, gasped out the news, and died of exhaustion. There had to be a better way—but for the next 2,300 years or so sending a flesh-and-blood messenger on a trip was the normal method of delivering information from one place to another.

One dramatic break from that convention appeared in postrevolutionary France. In the early 1790s, Claude Chappe invented a system of relaying *visual* messages hundreds of miles across the French countryside over a network of towers spaced about 20 miles apart. For example, someone in Paris would manipulate the mechanical arms at the top of one of these towers to spell out a coded message; his counterpart in another tower 20 miles away would read the message and duplicate it for the benefit of the person manning the next tower down the line, and so on. Weather permitting, this system could be used to transmit a message from Paris to the border of Germany within ten minutes. Other societies

had used visual communications techniques, such as ship-to-ship semaphore signals and such land-based mechanisms as smoke signals and torches. But the French, quickly joined by several other European countries, improved greatly on the idea by developing a nationwide communications *network*. By the Napoleonic era of the early 1800s, the French had developed a sprawling tower system radiating from Paris to such far-flung destinations as Cherbourg, Boulogne, Strasbourg, Marseille, Toulouse, and Bayonne.[1]

Before long, these networks, which could be used only in daylight and good weather, confronted the first revolutionary technology in telecommunications: the telegraph. Developed by Samuel Morse in the 1830s, the telegraph sent encoded messages down copper wires by rapidly opening and closing electrical circuits. The telegraph dominated telecommunications until it too was gradually replaced by the next revolutionary technology: the telephone system, invented by Alexander Graham Bell in 1876 and widely deployed throughout much of the United States within a generation. In the 1890s, Guglielmo Marconi exploited the discovery that the airwaves, like copper wires, could propagate electromagnetic signals, and so "radio" technology was born.

Today, although precise definitions differ, the term *telecommunications* is broadly defined as the transmission of information by means of electromagnetic signals: over copper wires, coaxial cable, fiber-optic strands, or the airwaves. This technology —which underpins radio and television, wireless and wireline telephone service, the World Wide Web, email, instant messaging, streaming video, and every other Internet application—is the sine qua non of contemporary global culture.

Telecommunications is also a uniquely volatile field economically, technologically, and politically. The disputes that arise within and among the different sectors of the telecommunications industry, often in response to these rapidly changing conditions, have triggered some of the fiercest public policy wars ever waged. In the United States, the very structure of

the industry turns on the decisions of various governmental agencies, most notably the Federal Communications Commission (FCC). The policy questions answered at the FCC and elsewhere influence not just how we communicate with one another and what video programming we watch (and how we watch it), but the fate of an industry that in the United States alone accounts for hundreds of billions of dollars in annual revenues and more than a million employees.

Nonspecialists, however, confront a vexing conundrum in trying to learn this field: to comprehend the whole of telecommunications policy, one must first understand its parts; but to understand the parts, one must first comprehend the whole. This chapter aims to overcome these difficulties by covering the major themes of telecommunications competition policy at a high enough level, and with as little jargon as possible, to help nonspecialists understand how each of the policy issues discussed in subsequent chapters fits into the big picture. To this end, the first part of this chapter introduces the peculiar economic characteristics of the telecommunications industry that drive most forms of regulation in the United States. The second part then introduces the market-transforming phenomenon of *convergence*—the competitive offering of familiar communications services through nontraditional technologies, such as the provision of video programming over telephone lines and telephone service over cable TV wires. As we discuss, the key to ultimate convergence is the greatest invention of the late twentieth century: the Internet. You can access the Internet through any number of dissimilar fixed-line and mobile technologies; and, once there, you can exploit virtually every form of communication ever created, including written correspondence (email and instant messaging), phone calls (voice over Internet protocol, or VoIP), newspapers and books (the Web), radio broadcasts (streaming audio), and television entertainment (streaming video).

I. Economic Principles

Why does competition in the telecommunications world—unlike, say, competition in the world of home appliance manufacturing—present public policy issues of such importance and complexity? Answering that question requires a familiarity with the basic economic phenomena that regulators have long cited to justify regulatory intervention in telecommunications markets. At the risk of some oversimplification, we sum up the most important of these phenomena in three concepts: *network effects*, *economies of scale and density*, and *monopoly leveraging*. We address each of these concepts in turn.

A. Network effects and interconnection

Flash back to the infancy of the U.S. telephone industry at the turn of the twentieth century. Different telephone companies often refused to interconnect with one another, and each had its own set of subscribers. Few consumers, of course, wanted to buy several telephones and pay subscription charges to several telephone companies simply to make sure they could reach anyone else they wished to call. Unfortunately, this was the choice many consumers faced.

Such arrangements are quite wasteful in that they misallocate society's scarce resources away from their most productive uses. To be sure, the prospect of extra profits from the successful deployment of a closed (non-interconnected) telephone network may well have encouraged some entrepreneurs to build a better product and reach customers more quickly than they otherwise would.[2] Apart from those incentive effects, however, consumers typically received little added value from multiple subscriptions that they would not have received from one subscription to a single carrier if the various networks were interconnected and exchanged traffic at reasonable rates. For the most part, consumers simply paid more money for the

same thing, which meant that they had less money to spend on purchasing things of value in other markets.

In the absence of any interconnection obligation, virtually every telephone market in early-twentieth-century America reached a tipping point in which the largest network—the one with the most subscribers—became perceived as the single network that everyone had to join, and the rest withered away. The potential for certain industries to slide into monopoly in this manner illustrates an economic phenomenon known as *network effects*. In many markets, individual consumers care very little how many other consumers purchase the same products that they buy. For example, the bottle of shampoo you just bought does not become significantly more or less valuable to you as the number of other purchasers of the brand increases or falls. The telecommunications industry, like several other "network industries," is different: the value of the network to *each* user generally increases or decreases, respectively, with every addition or subtraction of *other* users to the network.

Suppose, for example, that you lived in a midwestern American city in 1900, and two noninterconnecting telephone companies were offering you service. You would have been much more inclined (all else being equal) to select the company operating 80% of the lines rather than the one operating 20% because the odds would have been much greater that the people you wished to call would be on the larger network. The absence of interconnection arrangements among rival networks thus created a cutthroat race to build the largest customer base in the shortest time frame—and then put all rivals out of business by pointing out the limited value of their smaller networks. Economies of scale—a carrier's ability to reduce its per customer costs by increasing its total number of customers—further accelerated this process by permitting larger carriers to undersell smaller ones.

By the early twentieth century, the U.S. telephone market had tipped. In most population centers, the victor was the mammoth Bell System: a corporate family of very large "operating companies" that

provided local exchange (i.e., telephone) service and were eventually bound together by a long-distance network known as "Long Lines." All of the far-flung operations of the Bell System were owned by American Telephone & Telegraph (AT&T), which maintained its own equipment-manufacturing arm (Western Electric) and for a time also held the rights to patented technologies developed by the Bell System's namesake: Alexander Graham Bell.

In the areas AT&T did not control, which typically were the less populous ones, the so-called "independent" local telephone companies vied for market share. In many cases, AT&T sought to coerce these independent companies into joining the Bell System by refusing to interconnect them to AT&T Long Lines, which was then the only long-distance network in the United States. The independent companies were in no position to build a rival long-distance network. Even if they could have cooperated to construct the needed transcontinental facilities (and done so without infringing any remaining AT&T patents), they still could not have used that shared network to send calls through to the increasing majority of Americans who were served by the noninterconnecting Bell System. As a result, without interconnection rights, these independent companies could not provide their customers with satisfactory telephone service—that is, service extending beyond the local serving area—unless they could somehow duplicate the nationwide physical infrastructure the Bell System had built up over several decades of sharp dealing and self-reinforcing good fortune. That was an economic impossibility.

AT&T's coercion of the independent companies ultimately aroused the attention of the antitrust authorities at the U.S. Department of Justice. In the Kingsbury Commitment of 1913, AT&T resolved the ensuing dispute by agreeing to interconnect its Long Lines division with these independent local companies and to curb its practice of buying up independent rivals.[3] In exchange, the government placed its effective imprimatur on AT&T's monopoly control over all U.S. telecommunications markets in which it was

already dominant. This episode provides an instructive contrast to the anticompetitive conduct that ultimately led to the breakup of the Bell System 70 years later into its local and long-distance components. In 1913, AT&T used its control of the *long- distance* market to suppress other *local* carriers. As explained later in this chapter, AT&T would subsequently leverage its control of most *local* markets to suppress the *long-distance* competition that technological advances had made possible by the 1960s.

The network effects phenomenon presents different competitive questions in different industries, and reasonable people can disagree about when the government should require a firm to share access to its customer base. But when such intervention is deemed necessary, the usual solution is an *interconnection* requirement. Suppose you own a telephone network, and one of your subscribers wants to place a call to someone who subscribes to Provider X's network. If Provider X's network is larger than yours, it may have the incentives just described to refuse to interconnect, in which event your subscriber learns that the call has failed and so considers defecting to Provider X. But if the government forces Provider X to take the call onto its network and route it to the intended recipient, your customer remains satisfied, and you stay in business. Interconnection obligations work the other way as well: Provider X cannot preclude its subscribers from reaching yours.

For more than one hundred years, much of telecommunications policy has focused on the rules governing interconnection among the many conventional circuit-switched telephone companies that together make up the *public switched telephone network* (PSTN). Those rules govern, for example, what (if anything) one telephone company must pay another for access to the called party on the receiving end of a call. And they often govern the physical details of the interconnection arrangements between different telephone companies.

One critical set of interconnection arrangements, however, remains unregulated: those between the mostly private data

networks that in the aggregate constitute the Internet. For example, to stream movies to consumers, Netflix contracts with a variety of Internet intermediaries such as Level 3 to transmit and store its data content across the globe. In turn, those intermediaries contract with consumer broadband providers such as Comcast and Verizon to receive the data content and deliver it to Netflix's millions of subscribers. No regulator prescribes the terms of such private contracts, known as *peering* and *transit* agreements, or requires any Internet-based network to interconnect with any other. To date, policymakers have expressed confidence that market forces (the mutual self-interest of the Internet's many constituent networks) and other pressures (including social norms and the remote threat of regulatory intervention) will generally produce efficient ways for each user of the Internet to reach every other. Whether regulators *should* play a greater role in supervising these peering and transit agreements is a key subject of chapter 7. Significantly, these Internet interconnection issues are distinct from but closely related to the debate about *net neutrality* rules, which govern whether and when consumer broadband providers may block or discriminate among particular types of Internet content. In chapter 6, we explore the complex relationship between these two regulatory topics, a relationship that the FCC largely overlooked when adopting net neutrality rules in 2010.

Although we have focused so far on physical transmission providers, network effects are endemic to information industries in general. For example, Facebook now has more subscribers by far than other social-networking sites, and it is valuable to its users precisely because so many of their friends and acquaintances are on it. And network effects can sometimes appear in more subtle forms that at first blush do not appear to implicate "networks" at all. For example, the historical dominance of Microsoft Windows in the market for desktop computer operating systems arose from network effects and specifically from what antitrust courts have called the *applications barrier to entry*.[4] At some point in the 1980s, software

designers realized that more users were choosing Microsoft's operating system than the alternatives. In response, more and more applications developers created programs only for Windows, leaving would-be rivals (such as IBM) to sell operating systems that did not have as many programs designed for them and were therefore less popular. As a result, Microsoft won an increasing share of the operating system market, which in turn reinforced the software designers' predictions about the dominance of Windows and their desire to produce applications for it, often to the exclusion of applications for rival operating systems.[5]

In a variety of contexts, policymakers have disagreed about whether network effects create any problems for which the government should offer a solution. The proponents of government intervention argue that monopolization is virtually always an evil to be avoided, reasoning that monopolization of any industry necessarily produces higher consumer prices, less product variety, lower quality, and potentially less innovation.[6] Opponents of government intervention, by contrast, point to a theory of competition, first developed by economist Joseph Schumpeter, that focuses on the "creative destruction" of old incumbents by new insurgents, who are rewarded with monopolies of their own until knocked off their perch by the next round of insurgents.[7] Under this theory, the most significant competition takes place not *within* a market—in the form of price wars or incremental increases in quality—but *for the market itself:* that is, in establishing the next great invention that will displace the old monopoly with a new one. According to modern-day Schumpeterians, monopolies in the digital ecosystem are both temporary and desirable in that the prospect of even short-lived monopoly profits will encourage entrepreneurs to innovate.[8] The Schumpeterians therefore argue for strong intellectual property protection and freedom from both competition-oriented regulation and aggressive antitrust enforcement.

Of course, some monopolies in the digital ecosystem are more

enduring and harmful than others because the barriers to competitive entry are much higher in some markets than others. And policymakers have traditionally concluded that those barriers are particularly high—and the rationale for prophylactic regulation therefore strongest—in the case of physical telecommunications networks that provide last-mile transmission to individual homes and businesses, as telephone and cable television networks do. The reason relates not only to network effects, but also to the topic of our next subsection: *economies of scale and density*.

B. Economies of scale and density

By themselves, interconnection obligations significantly lower the entry barriers posed by the combination of network effects and scale economies because, as discussed, they exempt a new entrant from the need to build a ubiquitous network before competing for the dominant carrier's customers. But interconnection obligations do not eliminate those entry barriers altogether. Although they reduce any advantage that incumbents derive from network effects, they do not ensure that new entrants will benefit from the enormous scale economies enjoyed by a provider with a large, established customer base.*

What are these scale economies? Any telecommunications provider contemplating the construction of a new network faces immense initial costs. For wireline carriers, these include the costs of digging trenches and laying thousands of miles of wires to reach different customer locations. For wireless carriers, the start-up costs can include the price of acquiring spectrum rights, deploying a network of cell towers, and installing (or leasing) wires to connect those cell towers to network switches. These costs are *fixed* in that the carrier must incur them up front before it can provide any volume of service. In many cases, these costs are also *sunk* in that the investment, once made, cannot be put to some other use—a fact that makes the investment particularly risky.[9] In contrast, once the network is up and running, the *marginal* cost of providing service to

each additional customer is often tiny by comparison, particularly for wireline networks. Given these enormous fixed costs and negligible marginal costs, it is often cheaper *per customer* for a carrier to provide service to one million customers than to one thousand customers.

Closely related to such economies of scale are economies of density. The latter are best explained by way of example. Imagine a 1,000-unit beach condominium complex that is both distant from any telephone company switching station and, because of zoning restrictions, isolated from other buildings. If the fixed costs of laying a cable from the nearest switch to that complex were $100,000, a single telephone provider serving the entire complex could spread the recovery of those costs among all 1,000 subscribers for a cost of $100 per subscriber. But if ten providers divided up that customer base equally after laying their own cables to the same complex—each digging up the streets at different times and incurring the same fixed $100,000 cost—the average cost of that tenfold effort would rise to $1,000 per customer, for each provider could spread its $100,000 in costs only over 100 customers rather than 1,000.*

Similar economic considerations explain why fixed-line broadband service is much more costly to provide in rural than urban areas. Suppose your company runs a telecommunications network on a rigidly fixed budget. Assuming the same level of revenues, would you rather build one line to each of 1,000 customers living on widely dispersed farms or 1,000 lines to one apartment building with 1,000 units? Even if the average line length were the same in each example (say, because the apartment building is farther away from your switching station than half of the farms), you would still much rather serve the apartment building because you would only have to dig up the ground once to lay the lines needed to serve those 1,000 units. If you picked the farms option, you would need to dig up the ground many more times to lay 1,000 different cables, and you would have to pay far more to obtain the

rights of way. This radical difference in per customer costs explains why the broadband revolution has come more slowly or not at all to much of rural America—and why policymakers have focused so heavily in recent years on how to encourage greater broadband deployment in rural areas (see chapter 8).

Of course, high fixed costs and low marginal costs lead to large scale economies in many industries, from auto manufacturing to applications software production, and most such industries have never been subject to pervasive schemes of prescriptive economic regulation. The difference is one of degree. In most settings, scale economies do not increase "over the entire extent of the market"[10] because at some point average costs stop declining with each incremental unit. In other settings, however, scale economies keep increasing until a provider is serving all customers in the market. In that context, because a single firm can serve the whole market (however defined) with lower overall costs per customer than can multiple firms, the market is said to be a *natural monopoly*.[11]

The government has traditionally addressed such a market by awarding a monopoly to a single firm and heavily regulating its rates, on the theory that rate regulation is the best way to keep consumer prices low. Although our discussion to this point has focused on the local telephone market, the government applied the same natural-monopoly premise to the cable TV market, too. As Judge Richard Posner once explained,

> You can start with a competitive free-for-all—different cable television systems frantically building out their grids and signing up subscribers in an effort to bring down their average costs faster than their rivals—but eventually there will be only a single company, because until a company serves the whole market it will have an incentive to keep expanding in order to lower its average costs. In the interim there may be wasteful duplication of facilities. This duplication may lead not only to higher prices to cable television subscribers, at least in the short run, but also to higher costs to other users of the public ways, who must compete with the cable television companies for

> access to them. An alternative procedure is to pick the most efficient competitor at the outset, give him a monopoly, and extract from him in exchange a commitment to provide reasonable service at reasonable rates.[12]

These considerations led regulators for many years to conclude—somewhat controversially in hindsight—that telephone and cable television markets were each a natural monopoly in this sense and that the "alternative procedure" Posner described would be the optimal means of ensuring dependable service at low rates in any geographic area.

This natural-monopoly premise provided a convenient solution to the problem of network effects as well. Because (the thinking went) there was no reason to allow a second or third provider into the same geographic market to begin with (as that would only dilute the incumbent's economies of scale), there was no need to worry about forcing the incumbent to interconnect with competitors. The principal exception, illustrated by the Kingsbury Commitment, seems almost trivial in this light: different geographic regions would be served by different monopoly providers of local service, and the government would ensure simply that neighboring monopolists interconnected with each other for the exchange of calls between their respective regions and that the national monopoly provider of long-distance service (AT&T) allowed all of these monopolies access to the rest of the country.

Relying on this natural-monopoly premise, many regulators not only refused to order interconnection among potential rivals in the same geographic market, but straightforwardly prohibited new market entry by granting exclusive franchises to the monopolists. In part, policymakers resisted competition not just because they believed in the economics of natural-monopoly theory, but also because they relied on regulated monopolies to advance various social policies, most notably *universal service*—the promotion of affordable telecommunications for all communities. For example, regulators deliberately kept prices for business customers high

(compared to the underlying cost of serving them) as a means of cross-subsidizing affordable rates for other users, such as residential customers in rural areas where economies of scale and density are low and per line costs are therefore high.[13] As we discuss later, this scheme can work over the long term only to the extent that rival providers are barred from competing for the customers who pay the above-cost rates that subsidize low rates for others.

For many years, regulators acquiesced when AT&T's Bell System invoked universal service concerns to justify suppressing competition in all telephone-related markets, including equipment manufacturing as well as local and long-distance services. AT&T's long-lived regulatory success in this respect provides a classic case study in *public choice theory*—the economic analysis of relations between market participants and the government officials they seek to influence.[14] Public choice theory holds that private economic actors will exploit regulatory schemes to obtain or protect "rents"—that is, special benefits that arise from political influence rather than economically valuable contributions to social welfare. Successful rent seeking need not and usually does not take the form of outright bribery. Instead, private actors look for ways to match their own pecuniary interests with regulators' political goals. In the case of telephone regulation, the suppression of competition in the name of "universal service" gave both AT&T what it wanted—formally protected monopoly status—and the regulators what they wanted—an opaque scheme for underwriting low residential rates that avoided all the political costs presented by a more explicit and taxlike system.[15] The victims of such Faustian bargains are consumers, who in the long run are generally better off, at least in the aggregate, when regulators make the hard political choices necessary to remove barriers to competition.

Starting in the 1970s, policymakers began questioning the natural-monopoly assumptions that had been conventional wisdom almost since the inception of the industry.[16] This process followed a

predictable pattern, as we discuss in chapter 2. After the FCC adopted rules allowing competition in the provision of telecommunications equipment, the markets that next fell prey to competition were the ones in which overall call volumes were so huge and the incumbent's retail prices were so far above economic cost that a competitor could efficiently build a rival network and earn large profits even though it had only a small share of the total customer base. The first such market was for business-oriented long-distance services between major cities, a market that MCI and other firms entered in the 1970s and 1980s with the help of both microwave technology and the courts. The second was the market for so-called *access* services: the high-speed links between local networks and long-distance networks. In each case, the companies that owned the core natural-monopoly assets—the local exchanges, with their "last-mile" connections to every home and business in a given calling area—tried to thwart this nascent competition by (among other things) refusing to interconnect with the upstarts or by making interconnection unnecessarily burdensome. In each case, the U.S. government stepped in and mandated nondiscriminatory interconnection.

Finally, in the Telecommunications Act of 1996, also known as the "1996 Act," Congress seemed to dispense with the natural-monopoly premise altogether. It abolished all exclusive franchises, ordered all telephone companies to interconnect with any requesting carrier, and declared all "local exchange" markets—in addition to the long-distance and "access" markets—open for competition. But Congress could not repeal the laws of economics. In many settings, it remained commercially infeasible for new wireline competitors to build brand-new telephone networks bridging the last mile to all of their subscribers' buildings. The main exception to this rule lay in some local exchange markets—such as densely populated, downtown business districts—where high volumes of voice and data traffic enabled new entrants to exploit fiber-optic technology by building telecommunications networks all the way to their customers. In less densely populated areas, however, such as many

suburbs and most rural areas, call volumes could not support the efficient construction and operation of wholly duplicative telephone networks replete with thousands upon thousands of fixed connections to all homes and businesses.

As discussed in chapter 2, Congress attempted to address that concern in part by granting new entrants qualified rights to *lease capacity* on the facilities owned by the incumbent telephone company at regulated rates, thus enabling them to "participate" in the incumbent's economies of scale by availing themselves of the same low per unit costs. For the ensuing decade, the telecommunications industry was consumed with bitter arguments about how best to implement those leasing rules.

Those debates now seem almost antiquated because they have been largely overshadowed by the rise of technological alternatives to traditional wireline telephone service. Few residential consumers today buy telephone service from competitive providers that lease last-mile lines from incumbent telephone companies. As discussed later in this chapter, however, roughly three in ten American households have "cut the cord" by relying entirely on their mobile phones for voice communications. Tens of millions more subscribe to VoIP services offered by cable companies and Internet-based providers such as Vonage and Skype. Meanwhile, satellite providers such as DISH and DirecTV and telephone companies such as Verizon and AT&T now provide subscription video services in competition with traditional cable television companies.

In short, the natural-monopoly rationale underlying traditional regulation appears to have succumbed to commercial and technological developments, at least in retail consumer markets within most population centers. That does not mean, however, that these natural-monopoly concerns have disappeared altogether. For example, as discussed in chapter 2, regulators still view incumbent telephone companies (telcos), in some contexts, as monopolists in their provision of certain business-line services known as "special access." And as discussed in chapters 5 and 6, the FCC now tends to

view most regional markets for fixed-line broadband Internet access as duopolies, dominated in a typical area by the legacy telco and cable incumbents. From a competitive perspective, a duopoly is better than a monopoly, but it is hardly optimal.[17]

Some industry observers fear that these consumer broadband markets may become even more concentrated over time, ultimately veering in many areas toward a new cable monopoly for broadband pipes into the home.[18] They predict that as consumers come to expect ever-faster speeds over their broadband connections, cable incumbents, with their higher-capacity pipes, will begin pulling away from telcos for market share. And they fear that the telcos, if they can hope at best for a fraction of the customers in any given neighborhood, will be unable to justify the multi-billion-dollar network investments needed to supply a viable competitive alternative to cable. These observers cite the example of Verizon, the only major U.S. telco to deploy on a wide-scale basis fiber-optic lines all the way to individual homes. Dampening hopes for greater cable TV competition, Verizon announced in 2010 that it would begin winding down new deployments of that immensely high-capacity but costly infrastructure, known as Verizon FiOS.[19] A year later Verizon's majority-owned wireless affiliate entered into a cross-marketing arrangement with cable companies, which the Justice Department and the FCC approved in 2012 with various conditions.[20]

Critics viewed these developments as evidence that, at least in many areas, telcos will put up only tepid resistance to growing cable dominance for fixed-line broadband connections.[21] But that conclusion is subject to sharp debate. It is also by no means clear that telcos actually need to deploy fiber to the home in order to meet consumer bandwidth demands, at least in the near-to-intermediate term.[22] And, of course, consumers consider a range of service features when choosing a broadband provider, including not only speed but also price. In short, predictions about the future intensity

of residential broadband competition remain highly speculative.

C. Information platforms, monopoly leveraging, and net neutrality

So far, we have addressed the regulation of *horizontal* relationships within the telecommunications industry: the relationships between competing providers of substitutable services. Now we introduce the equally complex set of issues presented by *vertical* relationships between providers of communications-related goods or services in complementary markets. Vertical relationships arise across the economy: for example, between wheat farms and bakeries or between bakeries and grocery stores. Vertical integration by a firm across adjacent markets is often desirable because it can produce significant *economies of scope*: cost efficiencies obtained by producing several products at once. In most industries, moreover, competition in each of the adjacent markets liberates these vertical relationships from the need for heavy governmental oversight. To the extent the government gets involved, it is typically through ad hoc enforcement of the antitrust laws.

The government has long treated the communications marketplace differently. To govern vertical relationships in the telecommunications industry, policymakers have relied not only on after-the-fact antitrust enforcement, but also on prescriptive regulation. The rationale for that policy choice has its roots in the same natural-monopoly premise discussed above. If there is only one provider of a given communications service in a particular locality—one local telephone company for local voice service, one cable operator for multichannel video service—then, the thinking went, the provider would have strong incentives to harm competing providers of complementary services, such as independent long-distance companies (in the case of local phone companies) or unaffiliated cable channel programmers (in the case of cable operators), in order to help its own business or that of its vertically related affiliate.

Two developments now challenge that traditional premise. First, as the natural-monopoly premise has yielded to technological breakthroughs, competition among rival providers has reduced (but not eliminated) concerns that any given provider will harm consumers by denying them efficient access to complementary applications. Consider, for example, a cable operator that offers consumers both broadband Internet access and video-programming services. Depending on complex economic factors noted below, that cable operator might have incentives to impede its customers' access to unaffiliated Internet-based video services (such as Netflix streaming video or peer-to-peer file sharing) if it fears that those services will reduce consumer demand for its own lucrative video-programming services, such as video on demand. But it will be less likely to *act* on those incentives if it faces sufficient competition from a rival broadband/video provider (such as Verizon FiOS) because frustrated customers could respond by switching to the rival. As discussed in chapter 6, a key element of the *net neutrality* debate concerns whether competition between rival broadband providers is (and will remain) sufficient to keep each broadband provider from engaging in harmful discrimination against unaffiliated Internet-based applications providers that rely on unimpeded access to broadband platforms in order to reach consumers.

Second, quite apart from the degree of competition in the broadband market, many economists today, influenced by the Chicago School of antitrust economics, take a more skeptical view of "vertical leveraging" claims than did policymakers throughout much of the twentieth century.[23] That skepticism has roots in the *one-monopoly-profit* principle first developed by nineteenth-century economist Antoine Cournot. As Cournot observed, the total profits a monopolist can earn if it seeks to leverage its monopoly in one market by monopolizing an adjacent market are often no greater than the extra profits it can earn anyway simply by charging more for the monopoly product itself. As a result, even a monopoly

provider of platform services may lack strong incentives to harm unaffiliated applications providers because doing so would create no extra profits but would deprive the platform of value-enhancing applications and thus weaken the platform monopoly.

There are, however, important exceptions to this general principle. One arises where the platform provider fears that unaffiliated applications providers pose a threat to the underlying platform monopoly. The turn-of-the-millennium Microsoft antitrust case illustrates this exception. Although the Microsoft Windows platform dominated the market for desktop operating systems, Microsoft had generally cooperated with unaffiliated developers of complementary applications even when those applications competed with Microsoft's own applications. But federal antitrust authorities successfully argued that Microsoft had tried to crush Netscape in the 1990s—not because Netscape had designed an ordinary Internet browser that could run on top of Windows (and thus enhance its value), but because Microsoft feared that the Netscape browser might develop into a rival platform in its own right and thus devalue the underlying Windows monopoly.[24]

Another exception to the one-monopoly-profit phenomenon arises where the platform service is subject to price regulation. If so, the provider may well have incentives to discriminate against firms in adjacent markets because it will be unable to recoup all otherwise available monopoly profits from the sale of the platform service itself and will need to extract them instead from those other markets. This exception is sometimes called *Baxter's Law* in honor of William Baxter, the Justice Department official who cited it in the early 1980s as a reason for breaking up AT&T's Bell System. As Baxter understood, AT&T had a strong incentive to leverage its (price-regulated) monopoly in local markets to suppress competition in the adjacent long-distance market (see chapter 2). Unlike traditional telephone service, however, broadband Internet access is not subject to price regulation, and Baxter's Law is therefore generally inapplicable in that context. We return to these complex

economic issues in chapter 6, which explores the legal and economic underpinnings of the net neutrality debate.

II. Technological Convergence and Statutory Obsolescence

Until recently, most forms of electronic communications fell neatly into one of two general categories: *point-to-point* communications and *broadcasting*. The first category describes the traditional world of telephone companies: the transmission of content from a person or machine to a discrete recipient. Examples include ordinary telephone calls and fax transmissions. Broadcasting involves the transmission of content to a broad community or at least anyone who cares to watch or listen. Examples include television and radio.

For most of the twentieth century, people closely identified each of these categories of service with a particular medium of transmission. In particular, they assumed that commercial point-to-point voice services (telephony) would be conveyed over the copper wires of the telephone system and that radio and television broadcasting services would be provided over the airwaves. The Communications Act of 1934 was originally written with this assumption in mind. Congress designed Title II of the Act to govern wireline "common carriers"—that is, the companies that provided telephone service indiscriminately to the public at large. And it designed Title III to govern "radio communications," a category that grew to encompass both radio and television broadcasting. Under Title III, the FCC licensed radio and television stations to broadcast programming "in the public interest" over the airwaves.

And so the world remained until the 1960s, when something peculiar happened: companies increasingly began to transmit television signals not over the airwaves, but over wires. For a long time, such "cable television" service provided no new programming; it was designed only to transmit stronger signals of conventional

broadcast programming to people whose homes were too far away from a transmission tower to receive clear pictures (or any pictures). Even so, the seeming anomaly of wires being used for broadcasting threw the regulatory world into tumult, for it raised questions about how the FCC could legally follow through on its expressed intent to regulate this new creature and preempt contrary state and local regulation. After all, Title II addressed common carriage, not broadcasting, and Title III addressed use of the airwaves, not wires.

The FCC ultimately invoked the general enabling language in Title I of the Act to assert what it called *ancillary authority* to regulate this and other new technologies that substantially affect the explicit subjects of its regulatory authority. In 1984, long after the Supreme Court upheld the FCC's initial rules exercising this strikingly open-ended regulatory authority (in 1968),[25] Congress stepped in and added a new Title VI to the Communications Act to govern federal, state, and local regulation of cable television services.

A similar need for statutory reform arose in the 1980s, when "cellular" wireless technology gave consumers an altogether new means of placing telephone calls. This technology used the radio spectrum—long the province of specialized broadcasts by taxi dispatchers and police officers in addition to television and radio stations—for regular communications among members of the public at large. That development produced another anomaly unanticipated in the structure of the 1934 Act: the use of the *airwaves* to provide point-to-point common carrier services. Congress eventually patched this legal hole by adding provisions to Title III to govern the regulation of this new service.[26]

The use of radio signals to carry telephone calls and of wires to carry broadcast programming are examples of *technological convergence*: the use of different technologies to provide similar services. But the examples of convergence just discussed are tame in comparison to the upheavals triggered by the Internet. By placing a

"call" over your broadband Internet connection to a distant website, you can listen, along with the citizens of Prague, to the broadcast of a Czech radio station. Or you can log onto Pandora, an Internet-based service that keeps track of the music you like and sends you—and you alone—a personalized stream of songs. Or you can use software to chat with a friend across the world through instant messaging or VoIP. Or you can cancel your cable or satellite TV service and watch video entertainment over the Internet from streaming-video services such as YouTube, Netflix, and Hulu.

Here is the critical point: you can do all these things no matter what type of broadband connection you use to reach the Internet—whether you use your telephone line, your television cable, or a mobile broadband device such as a smartphone, tablet, or wireless-enabled laptop. The Internet supports every type of communications service ever invented. And as a general matter, you can run each of those services over any high-performance fixed or mobile broadband platform, irrespective of the platform's underlying technology.

The Internet owes its birth and explosive development to digital technology, which came of age commercially in the 1980s and 1990s. As chapter 5 discusses in depth, digital technology provides concise mathematical representations of the world in the form of 1s and 0s. The software inside your computer or mobile device decodes and converts those 1s and 0s into everything from voice conversations to photographs to documents to Prague radio broadcasts. The Internet, in turn, is a conceptual aggregation of many individual networks, most of them privately owned, that use a common protocol and addressing scheme— the *Internet protocol* (IP)—for transporting packets of 1s and 0s among computers and other smart devices. The computers on each end of a data session do not "care" what physical conduits link them together so long as the packets are delivered quickly enough for the relevant software programs to run properly. And, for the most part, the Internet's physical infrastructure has not traditionally "cared" what software

programs those packets are associated with; it just delivers the packets and lets the computers do the rest. In part because "a bit is just a bit" in this sense, the Internet severs any strong logical or practical link between communications *services* and the physical *media* over which they are transmitted to consumers.

Technological convergence has uprooted some of the most basic premises of telecommunications policy. Take, for example, the case of telephony, the transmission of ordinary voice conversations between two or more people. For many decades, this was the exclusive province of heavily regulated wireline telephone companies operating on the PSTN, the interconnected universe of telephone networks that rely on legacy circuit-switched technologies. Because of that historical pedigree, wireline telephone companies remain subject, in a gradually declining number of areas, to monopoly-era *dominant-carrier* regulations that prescribe the rates they can charge consumers for conventional telephone service. And as part of their universal service requirements, such companies, unlike their competitors, are often compelled to provide telephone service to rural areas where per line costs far exceed regulated revenues.

These PSTN-centric regulatory obligations are increasingly anachronistic because conventional circuit-switched wireline networks are becoming steadily less relevant to modern communications. The percentage of American households that have cut the cord to their local phone company and rely exclusively on their cellphones for voice service has surged from about 6% in 2004, when this book's first edition was written, to 31.6% in 2011.[27] And even one-third of the nation's *fixed-line* residential voice connections—28.2 million out of 89.8 million total as of mid-2010—were provided by cable companies and other nonincumbent telephony providers.[28]

These alternative voice providers have already integrated, or within a few years will have integrated, all of their services onto

unified IP platforms. And some incumbent telephone companies have followed suit, offering their own triple-play bundles of voice, video, and broadband Internet access over converged IP platforms as well. Data traffic used to constitute a tiny percentage of the signals flowing over a voice-centric telephone network; soon, however, voice traffic—VoIP—will become just a small percentage of the bits flowing over a data-centric network.

If the Internet ecosystem is to function properly, this comprehensive technological convergence must now be matched by equally comprehensive reforms to a regulatory regime still marked by the preconvergence assumptions of the original Communications Act of 1934. First, except where lingering natural-monopoly conditions make one provider dominant in a particular market, like services should generally be regulated alike, no matter what physical medium is used to provide them. For example, it is an increasingly questionable policy to single out wireline telephone companies for dominant-carrier regulation in their provision of retail voice services in areas where they now may have only a minority share of voice subscribers. As cross-platform competition increases, a regulatory regime that still treats substitutable platforms differently will distort the marketplace by, among other things, creating artificial regulatory advantages for one set of competitors over another.

Second, policymakers should aggressively follow through on plans to replace the traditional PSTN-based universal service system, which focuses on extending narrowband telephone service, with a competitively neutral program for supporting affordable *broadband Internet access* to otherwise unserved communities. The need for this transition is the central topic of the FCC's influential *National Broadband Plan*, issued in 2010, and of a sweeping 2011 FCC order that makes broadband deployment the new focus of federal subsidies for rural and other "high-cost" areas.[29] As we discuss in chapter 8, however, that reorientation of universal service goals is easier to embrace in principle than in practice, and it poses a

host of legal, political, and practical challenges.

Third, Congress and the FCC need to work together to devise a unified and stable legal framework for addressing the policy issues raised by the new IP-centric communications ecosystem. As discussed, the Communications Act is divided up into "titles" corresponding to different categories of communications services: Title II for telephone companies and other providers of "telecommunications services" (i.e., common carrier services); Title III for broadcasters; and Title VI for cable television operators. When, in 1996, Congress last enacted major revisions to the Act, it did not clearly foresee the rise of broadband Internet access services, let alone their eventual centrality to all forms of electronic communications. And it therefore did not specify how such services should be classified within the Act's existing structure or how, if at all, they should be regulated.

Given this lack of statutory direction, the telecommunications policy world has become mired in repeated controversy since 1996 about whether broadband Internet access services are properly classified as "telecommunications services" subject to common carrier regulation under Title II or, as the FCC determined in 2002, "information services" subject only to minimal regulation under the Commission's Title I ancillary authority.[30] That debate assumed new prominence when in early 2010 a federal appellate court invalidated the FCC's initial rationale for asserting Title I authority to police net neutrality violations.[31] In response, the FCC floated—and then, in the face of political opposition, all but abandoned—a proposal to repudiate its 2002 decision and reclassify broadband Internet access as a Title II service, a reversal that itself would have been subject to legal challenge (see chapter 6). Broadband Internet access service thus remains what it was in 2002: a Title I information service. But because the contours of the FCC's Title I ancillary powers are highly uncertain, so is the Commission's authority to adopt broadband-related policy, from the net neutrality rules it adopted in late 2010 (under a new Title I theory that remains

pending on appeal at press time) to the broadband subsidy regime the Commission created in 2011. That uncertainty is unlikely to be eliminated anytime soon unless Congress steps back into the fray.

Finally, the rise of convergence and broadband competition creates a new long-term aspiration for telecommunications policy: creating the market conditions needed to phase out most forms of public-utility-style regulation while supporting collaborative industry-led initiatives to develop basic norms of good Internet citizenship. As the *National Broadband Plan* observes, the U.S. broadband market structure is "relatively unique in that people in most parts of the country have been able to choose" among different fixed-line broadband platforms, whereas most foreign consumers have not.[32] The *Plan* notes that this "competition appears to have induced broadband providers to invest in network upgrades" and that "[c]onsumers are benefiting from these investments" in the form of steadily improving broadband performance.[33] Nevertheless, although competitive conditions vary, the typical American has a choice of only two fixed-line broadband providers: the local telephone company and the local cable TV company.

This is where mobile broadband services and spectrum policy come in. Americans are turning increasingly to their wireless devices for broadband Internet access, from smartphones to tablets to laptop cards. Indeed, wireless broadband services are so explosively popular that bandwidth demands, fueled by streaming video and audio, are creating what the FCC has labeled a "spectrum crunch." The federal government oversees who can use particular frequency bands and for what purposes, and there is broad consensus that far too little of it is allocated to meet surging mobile broadband needs. As FCC Chairman Julius Genachowski explained in 2011, "If we do nothing in the face of the looming spectrum crunch, many consumers will face higher prices—as the market is forced to respond to supply and demand—and frustrating service—connections that drop, apps that run unreliably or too slowly. The result will be downward pressure on consumer use of wireless

service, and a slowing down of innovation and investment in the space."[34]

Mobile and fixed-line broadband services may never be full substitutes: fixed-line services will never be mobile, and mobile services may never support the same high-bandwidth usage as fixed-line services at the same price points. That said, freeing up additional spectrum for mobile broadband services not only will make them faster and better in their own right but also might well impose additional competitive discipline on fixed-line broadband services (in the form of a "good enough" broadband service).[35] Like universal service reform, however, spectrum reform is easier to propose than to accomplish, and it confronts enormous political obstacles erected by incumbent spectrum holders, ranging from local broadcasters to the military. As we discuss in chapter 3, navigating those obstacles may be the single most important ambition of telecommunications policymakers in the early twenty-first century, and meeting that policy challenge is a key to resolving many others.

* Although our discussion treats network effects and scale economies as two separate phenomena, they are in fact closely related. Each describes a characteristic of markets in which, all else held constant, increasing the scale of a firm's operations improves the ratio of (1) the value of the firm's services to each customer and thus the revenues the firm can obtain from that customer to (2) the per customer cost to the firm of providing those services. Network effects improve this ratio by increasing the value of the service to each customer, whereas scale economies improve it by decreasing the per customer cost of providing that service. In the absence of regulation, each result can help favor larger-scale telecommunications firms over their smaller rivals.

* Economies of density can be roughly conceptualized as scale economies within a particular geographic area, such as the condominium complex in our example. For ease of exposition, we use the term *scale economies* broadly to include these economies of density.