

## 2

### UNIVERSAL SERVICE: A CONCEPT IN SEARCH OF A HISTORY

UNIVERSAL SERVICE entered the vocabulary of American telecommunications in 1907. The slogan “one system, one policy, universal service” was coined by Theodore Vail, the President of AT&T, and propagated in the company’s annual reports from 1907 to 1914.<sup>5</sup> Its appearance came, as we shall see later, at the peak of a fierce competitive struggle between the Bell system and thousands of independent telephone companies. The idea of universal service served as the linchpin of the Bell system's argument for transforming the telephone industry into a regulated monopoly. The emergence of the concept thus marked an important turning point in the history of American telecommunications.

Most historians and policy makers believe that when Vail invoked *universal service* he meant the same thing we mean by it today: regulatory policies to promote the affordability of telephone service through cross subsidies.<sup>6</sup> This book disputes that widely-accepted view. There is, it argues, an important difference between Vail's concept of universal service in 1907 and the conception prevailing now. Understanding that difference is what this book is all about. At stake is not simply a question of historical semantics, but a far-reaching reinterpretation of the history of telecommunications with significant implications for current and future telecommunications policies.

In contemporary discourse, *universal service* policy is synonymous with government policies to promote the affordability of telephone service and access to the network. Sometimes this means direct subsidies to telecommunications construction in remote areas, such as the Rural Electrification Administration loan program. More commonly, it refers to attempts to maintain affordable local rates

---

<sup>5</sup> American Telephone and Telegraph Company, ANNUAL REPORT 17-18 (1907).

<sup>6</sup> See, for example, Herbert S. Dordick, Toward a Universal Definition of Universal Service, UNIVERSAL TELEPHONE SERVICE: READY FOR THE 21ST CENTURY? (Annual Review of the Institute for Information Studies, 1991).

by means of rate averaging and cross-subsidies within the nation's telecommunications system. This might mean, for example, imposing higher charges on long-distance users in order to reduce charges for basic local telephone service for households. Or it might mean charging the same rates for a long-distance call between two small, remote towns in Montana as for calls between Philadelphia and New York, when in fact economies of scale make calls between the latter two high-traffic centers far less expensive than the former. Whatever the mechanism, pushing telephone penetration towards 100 percent is seen as a policy goal of sufficient importance to justify various forms of public intervention in the industry. Underlying all these policies is the assumption that without active government intervention, access to telecommunications would be well below the socially optimal level.

*Universal service* in that respect is an expression of liberal egalitarianism, like universal schooling, literacy, or voting rights. More than just a telephone in every home, it implies that a ubiquitous communications infrastructure can contribute to national unity and equality of opportunity. In debates over the emergence of competition in the telephone industry during the 1970s, 1980s, and 1990s, the concept has become a pillar of the developed world's postal, telephone, and telegraph monopolies.<sup>7</sup> Telephone companies and regulators warned that universal service could not have been achieved without the regulated monopoly structure, and that competitive market forces had to be thwarted or tempered lest those goals be undermined. More recently, advocates of a new "information superhighway" have also drawn upon the concept to promote broadened access to new technologies.

This is the *modern* construction of universal service policy – the definition that has prevailed from about 1975 to the present time. This construction, and particularly its link with the regulated monopoly regime, draws its legitimacy and support from (questionable) interpretations of history. The all-encompassing network is perceived as an historical achievement of public regulation and a monopoly structure. Indeed, the modern idea of universal service comes with a full-blown version of its own historical origins. According to the conventional wisdom, universal telephone service was a public policy mandated by the 1934 Communications Act, and consciously brought into being by regulators acting in conjunction with telephone monopolies. A typical statement of this view appeared recently in an industry trade journal. "Telecommunications public policy crystallized in America with the Communications Act of 1934. Its goal was clear: the provision of universal service to every citizen in the country... Telephones at the time were viewed as a 'social necessity' that should be provided to all."<sup>8</sup> The crowning achievement of this system, so the story goes, was the 92 percent household penetration ratio of the telephone just prior to the AT&T divestiture.

The authors of that claim are not historians and offer no evidence for their claim. But that is precisely my point: the viewpoint they express is so common and so widely accepted that it is assumed to need no substantiation. In making such a statement the authors are merely reiterating something that most business people, academics, and regulators involved in the telephone industry take as a truism. There are minor variations on the theme. The telephone companies, particularly those associated with AT&T, emphasize Vail's formulation of the idea and the telephone industry's contribution to its realization. In their view, the Communications Act simply ratified the far-sighted

---

<sup>7</sup> Nicholas Garnham, *Universal Service in Western European Telecommunications. European Telecommunications Policy Research* (IOS 1989).

<sup>8</sup> Barbara J. Farrah and Mike Maxwell, *Building the American Infostructure, Telephony* 45 (Apr. 20, 1992).

vision of a corporate statesman. Academic scholars generally follow in this vein. A recent book on the subject, for example, hails the 1934 Act as the codification of Theodore Vail's vision of universal service<sup>9</sup> and highlights the role of redistributive rate regulation in promoting broader penetration.<sup>10</sup> In short, the modern definition of universal service is based on a holy Trinity comprised of the Communications Act, regulated monopolies, and rate subsidies.

But it is surprisingly easy to refute those historical claims. The words 'universal service' never appeared in the 1934 Communications Act. In preparing the law, Congress filled thousands of pages of the Congressional Record with research and documents about communications companies, but one will search that record in vain for the words 'universal service,' or for any evidence of a policy corresponding to the one described above. No mechanism for subsidizing telephone service was created or authorized in the legislation. Indeed, federal regulation could not have had much impact on the universality of telephone service in the 1930s or 1940s, or even the 1950s. The Communications Act and the FCC had jurisdiction over interstate telecommunications only, and in 1934 less than 2 percent of all telephone traffic crossed state lines.<sup>11</sup> The 1934 Act thus affected only a tiny portion of the overall telephone marketplace. Most of the regulatory action was at the state level and thus was in the hands of state commissions.

If one bothers to examine the text of the 1934 Act, one finds that it is little more than a piece of legislative consolidation. Its stated purpose was to put federal authority over communications into one specialized agency. In order to implement this goal its drafters took parts of the Interstate Commerce Act authorizing the Interstate Commerce Commission to regulate interstate telephone service (passed in 1910) and combined them with the 1926 Radio Act, which created a Federal Radio Commission to regulate broadcasting. The result was a consolidated Communications Act and a single regulatory agency, the Federal Communications Commission. The bill's House sponsor, Speaker Sam Rayburn, explicitly stated that the Act *did not change existing law*.<sup>12</sup>

That brings us back to the point at which this chapter began. If the universal service concept originated not in the Communications Act of 1934 but in the Bell-independent competition of the early 1900s, why did a debate about universality emerge at that time? And if, as I have asserted, Theodore Vail and his contemporaries did not mean by universal service what we mean today, what *did* they mean? As usual, the historical reality is more interesting than the myth. The universality of telephone communications became an issue in the early 1900s because the local telephone exchanges of the Bell system and the independents were not connected to each other. Competition took the form of two separate telephone systems in the same city or town vying with each other for subscribers and for connections to other localities. Subscribers who joined one system could not call the subscribers of the other –unless, as happened about 13 percent of the time, the user subscribed to both systems. Duplicate subscribers (mostly businesses) literally had two separate telephone instruments, Bell and Independent,

---

<sup>9</sup> Herbert S. Dordick, *supra* note 6.

<sup>10</sup> "The goal of having a universal telecommunications service has historically been to keep charges low enough that all but the poorest Americans could afford to make and receive telephone calls..." R. Entman, Introduction, *Universal Telephone Service: Ready for the 21<sup>st</sup> Century?* (Annual Review of the Institute for Information Studies 1991).

<sup>11</sup> *Smith et al v. Illinois Bell* 282 U.S. 133 (1930).

<sup>12</sup> "...the bill as a whole does not change existing law, not only with reference to radio but with reference to telegraph, telephone, and cable, except in the transfer of jurisdiction [from the ICC to the new FCC] and such minor amendments as to make that transfer effective." 78 Congressional Record, 73rd Congress, 2nd Session, 10313 (1934).

on their desks or walls. Even when there was only one exchange in a community, dual service divided subscribers. A Bell exchange could not make connections with the subscribers of competing independent exchanges in other cities, and vice-versa. In effect, telephone users confronted the same kind of barriers to communication as IBM compatible and Macintosh computer users of the 1980s and 1990s. The incompatibility, however, was usually due more to the companies' refusal to deal with each other than to technological incompatibility.<sup>13</sup>

“Dual service” was the contemporary name for competing, non-interconnected telephone exchanges in the same community. Dual service diverges so radically from our current experience with a universally interconnected telephone system that it is hard to appreciate just how widespread and long-lived the phenomenon was. It existed in some form for thirty years, from 1894 to 1924. From 1900 to 1915, at least 45 percent of the U.S. cities with populations over 5,000 had competing, unconnected telephone exchanges. During the peak of the independent movement’s strength, between 1902 and 1910, this percentage remained over 55 percent. Some of the nation’s largest cities had dual telephone systems for more than fifteen years: Cleveland, Buffalo, Minneapolis-St. Paul, St. Louis, Los Angeles, and Philadelphia.

The term universal service emerged from within this environment. It was put forward by the Bell system as the policy alternative to dual service. Universal service meant consolidating the competing telephone exchanges into a local monopoly so that all telephone users could be interconnected. It did not mean a telephone in every home, or government policies to subsidize telephone penetration or affordability. (That argument, which is vehemently disputed by historians whose work has been supported by the Bell system, is documented at length in chapter 8, and also in chapters 9 and 11.) After 1907, the Bell-independent business rivalry was transformed into a political and ideological struggle between two opposing principles of industry organization: dual service vs. universal service. Vail and other Bell spokesmen challenged the fragmentation and duplicate subscriptions caused by the presence of competing telephone exchanges. Independents defended it as a small price to pay for the price restraints, service improvement, and innovation promoted by competition. It is impossible to understand the historical meaning of the term *universal service* without grounding it in the context of dual service competition. One goal of the book is to reconstruct that original meaning and explore its implications for our understanding of telephone history and policy.

There is another reason why the universality of telephone service was implicated in the competitive struggle between Bell and the independents. It was true that dual service competition restricted universality by fragmenting telephone users. But, paradoxically, such competition also *rewarded* the pursuit of universality by the telephone companies themselves in a way that regulation and monopoly have never been able to do. A telephone system with more people on it is *ceteris paribus* more valuable than one with fewer subscribers.<sup>14</sup> Competing systems which are not connected to each other gain a competitive advantage over their rivals as they extend service to more users and locations. This dynamic was the driving force behind the Bell-independent rivalry of the early 1900s. Dual service propelled both systems into a race to wire all parts of the country and attract as many subscribers as

---

<sup>13</sup> Technological differences did play a role, however, as independents often used automatic switching during these years whereas the Bell system was still relying on manual switching.

<sup>14</sup> This is known as the “network externality” in economics. A more formal analysis of its properties and its implications for early telephone competition is conducted in chapter 3.

rapidly as possible. Penetration and geographic coverage in the United States, particularly in rural areas, made the most rapid gains in that period. The other goal of the book is to construct a systematic argument that dual service competition did more to promote the universality of the telephone than the later policies and practices of regulated monopoly. Together, the two arguments challenge orthodox views of telephone industry history. The reconstruction has both historiographical and policy import.

Perhaps most importantly, recasting that period of telephone history leads to a fundamental reinterpretation of why the telephone system became a monopoly. Odd as it may seem, after three anti-trust cases and scores of journal articles, the monopolistic character of telephone service is still a subject of intense historical and theoretical controversy. Traditionally, economists explained the structure of the telephone industry as the most efficient form of organization due to the presumed existence of supply-side economies of scale and scope. This explanation is grounded in the theory of natural monopoly. Many historians and economists have rejected the natural monopoly explanation, however, insisting that monopoly resulted from abusive and predatory actions of the Bell System. Whether monopoly was a product of “natural” economic forces or market manipulation is a matter of some consequence in a country governed by antitrust laws, and until now these two positions have defined the spectrum of opinion on the subject.

This book rejects both views. A different explanation, which might be called the *universal service theory of monopoly*, is advanced. That theory portrays telephone monopoly as a product of a conscious, publicly mediated policy decision to “unify the service;” i.e., to eliminate the user fragmentation created by dual service. In chapter 3, I characterize that outcome in economic-theoretic terms as an attempt to realize *demand-side economies of scope*. That represents a new theoretical position, in that it shifts the explanation for the efficiency of monopoly from the supply side to the demand side and from economies of scale to economies of scope. It is also a distinct historical position in that it stresses that the elimination of dual service was the product of a political consensus rather than a unilateral product of the Bell system. (The existence of a policy consensus, however, does not necessarily mean that the right decision was made, nor does it mean that there was no cleavage between losers and winners. In chapter 11, I explore the political dimension of this choice by examining in detail the consolidation of competing telephone exchanges in several major urban centers after 1915.)

Another historical issue revisited by this text is the role of interconnection (and the absence of interconnection) in the development of the American telecommunications infrastructure. This is a badly neglected and often misinterpreted topic in the historical literature. The most influential account of the competitive period, the *Telephone Investigation* of the Federal Communications Commission (1939), devotes only a few dismissive sentences to dual service competition.<sup>15</sup> Its incomplete and inaccurate treatment of the subject has misled two generations of historians. Lipartito (1989), Langdale (1978), Fischer (1987, 1993), and other historians with access to primary sources mention it, but fail to draw the crucial linkages between the lack of interconnection, the pursuit of universality by the competing telephone companies, and the demand for complete interconnection as the rationale for the choice of regulated monopoly as the institutional form for the telephone. Since dual service has not been taken seriously by historians, data about its nature and extent has not been systematically collected. Previous

---

<sup>15</sup> Duplication of telephone exchange service is dismissed by the FCC Report as “wasteful from the viewpoint of investment and burdensome to the subscriber.” Federal Communications Commission, *Investigation of the Telephone Industry*, 133 (GPO 1939).

studies have never adequately quantified the phenomenon. Aside from that, the existence of fragmented telephone service raises a number of intriguing historical questions. To what extent did the division of subscribers into two systems correspond to other social divisions, such as social class or ethnic groups? How frequently were users unable to reach desired parties due to competitive fragmentation? Which user groups perceived subscriber fragmentation as the most serious inconvenience? To what extent did the availability of long-distance connections affect the choice of a local subscription? How many and what type of users took out duplicate subscriptions? This book explores these economic and communicative features of dual service in detail.

When the phenomenon of non-interconnection is not simply overlooked, it is often misrepresented. Policy analysts and economists who have written about the early competitive period generally treat the lack of interconnection as an anti-competitive abuse. The interpretation that has generally arisen from those works is that it was Bell's refusal to interconnect with the independents which ultimately defeated them. The truth, as subsequent chapters will show, is very different. Until 1910, the independents were as uninterested in interconnecting with Bell as vice-versa. Further, Bell's refusal to interconnect utterly failed to stop the independents from proliferating throughout the country. Conversely, Bell's decisions from 1901 to 1908 to aggressively *interconnect* its toll lines to noncompeting independent exchanges was a damaging blow to dual service and the most powerful method of promoting its concept of universal service. Furthermore, the Kingsbury Commitment of 1913, which is almost unanimously represented by historians as the "end" of the dual service era and mistakenly counted as the beginning of universal interconnection, has been completely misinterpreted.

If revisiting the dual service era leads to substantial revisions in the way we understand and periodize telephone history, it also has important implications for current and future telecommunications policies. The policy relevance of history is often neglected by decision makers in business and government. Their attention consumed by an uncertain future and a complex, demanding present, they tend to assume that historical analysis can contribute little to their understanding. But upon reflection it is apparent that conceptions of the telecommunications industry's past can and do play a decisive role in policy thinking. The origins myth linking the Communications Act and universal service policy did not come about for nothing. In chapter 13, I show how and why this myth was created in the 1970s, when AT&T was beset by new competition once again. Likewise, current conceptions about the competitive consequences of interconnection and the need for "equal access" are derived mainly from interpretations of telephone and telegraph history. Telephone history has policy implications, which is why it so often has been, and will probably always remain, a contested area. The real issue is whether decision makers will be guided by accurate history or inaccurate history.

This book attempts to reframe the ongoing debate about universal service. If the standard historical assumptions about regulated monopoly's role in the creation of universal service are true, then developing countries and other nations considering competition and liberalization must control and limit competitive forces to preserve universal access. If, on the other hand, dual service competition played a critical role in the development of a ubiquitous telephone infrastructure, and that experience accounts for the tremendous U.S. lead in the extension of telecommunications service, then very different policy conclusions can be drawn.

More broadly, interconnection of competing networks is a critical issue – perhaps *the* critical issue – in contemporary telecommunications policy. The Bell-independent rivalry of the early 1900s provides a reservoir of empirical data about the effects of various approaches to interconnection. True, the social and technological conditions of that period are far different than today’s. The economic principles, however, are the same. A historically accurate portrayal of the evolution of interconnection arrangements in the U.S. telephone industry from the 1890s to the 1920s can only be helpful. The implications of that evidence for current policy are discussed in chapter 14.