Universal Service in Telephone History: A Reconstruction

by Milton Mueller

Columbia Institute for Tele-Information
Graduate School of Business
Columbia University
809 Uris Hall
New York, NY 10027
(212)854-4222

Universal service in telephone history

A reconstruction

Milton Mueller

The universality of telephone service is generally believed to be an achievement of regulated monopoly and rate subsidies. This paper critically examines the historical claims of what it terms the ideology of universal service. It shows that a ubiquitous telephone infrastructure developed in the USA because of competition between Bell and the independents in the period 1894–1921. Moreover, it shows that it was the refusal of Bell and the independents to interconnect with each other, a phenomenon which is generally ignored or condemned in the historical and economic literature, which propelled both systems into a race to achieve universality, leading to rapid increases in penetration and geographic scope, particularly in rural areas. The phrase universal service, which first emerged in telephone policy debates in 1907, did not mean a telephone in every home or rate subsidies, but the interconnection of the systems into a unified, non-fragmented service.

The author is Assistant Professor at Rutgers University, School of Communication, Information and Library Studies, 4 Huntingdon Street, PO Box 5067, New Brunswick, NJ 08903-5067, USA (Tel: +1 908 932 7910; fax: +1 908 932 1202; Bitnet: MMueller@Zodiac).

'Universal service' is one of the most commonly cited principles of telecommunications policy. Like the words 'democracy' and 'equality', the term has lofty connotations. More than just a telephone in every home, it implies that a ubiquitous communications infrastructure can contribute to national unity and equality of opportunity. Historically the concept has been applied and interpreted in ways that have had a powerful impact on public policy and regulation. More recently it has gained currency in North American, European and Asian policy debates as competition has encroached on traditional communications monopolies.[1]

In contemporary battles over the introduction of competition in telecommunications, universal service, along with its sister concept, natural monopoly, became an ideological pillar of the developed world's postal, telegraph and telephone monopolies. Telephone companies and many regulators warned that the regulated monopoly structure was set up with the preservation and advancement of universal service in mind, and that competitive market forces had to be thwarted or tempered lest that goal be undermined. This claim has succeeded in establishing an opposition between competition on the one hand and universal service on the other. That opposition, however, is largely based on premises and arguments regarding the history of the telephone. The ubiquity of the telephone is asserted to be an achievement of a particular institution: regulated monopoly.

This paper investigates the development of the universal service concept in the USA. It advances revisionist theses concerning the original meaning of the term, the role of competition in the achievement of a ubiquitous telephone infrastructure, and the universal service claims of regulated telephone monopoly. The paper is cast as an exercise in revisionism, but it might be more accurately labelled a historical correction, for it is the prevailing view of universal service which revises and distorts the historical record.

The words 'universal service' entered the vocabulary of US telephone regulation and policy in 1907, when a competitive struggle between the

[1]Nicholas Garnham, 'Universal service in European telecommunications', in *European Telecommunications Policy Research*, Proceedings of the Communications Policy Research Conference, 22–24 June 1988, Windsor, UK, IOS, Amsterdam, 1989; Institute for Information Studies, *Uni-*

continued on page 353

0308-5961/93/050352-18

Bell System and independent companies was at its peak. The paper shows that the universality of the telephone network became an issue at that time because the competing systems were not interconnected with each other. A competitive race between unconnected telephone systems ensued, a phenomenon which I label 'access competition'. The thesis of this paper is that it is the dynamics of access competition, more than any other single factor, which explains why telephone service was extended to rural and small-town America during the 1890s and early 1900s, and why by 1920 US telephone penetration in both rural and urban areas reached levels that were not achieved in other parts of the developed world until the 1970s.

The paper has two primary objectives. One is to counter false but popular and influential ideas about the historical relationship between universal service, regulated monopoly and competition. The other is to call attention to access competition as a historical phenomenon. The role of competition between non-connected telephone exchanges in the development of the US telephone infrastructure is a badly neglected and often misinterpreted topic. The most influential account of the competitive period, the Telephone Investigation of the Federal Communications Commission (1939), devotes only a few negative sentences to it.[2] Its incomplete and inaccurate treatment of the subject has misled two generations of historians. Gabel,[3] Brock[4] and many other policy analysts and economists treat the lack of interconnection as an anti-competitive abuse, a misinterpretation which has had a profound influence on contemporary policy. Lipartito,[5] Langdale,[6] Fischer[7] and other historians with access to primary sources mention it in passing, 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.

The modern definition of 'universal service' and the historical claims of regulated monopoly

The term 'universal service' was first used by Theodore Vail, the President of AT&T, in the company's 1907 *Annual Report*.[8] As I will establish, however, Vail's terminology belongs to a different era, and its meaning should not be confused with the current concept. The purpose of this section is to analyse the modern construction of universal service, and to show that the policy it represents is neither as old nor as venerable as is commonly assumed.

In its common modern construction, universal telephone service means reaching every member of society, no matter how remote or poor. Widespread access to telephone service is seen as a policy goal of sufficient importance to justify rate subsidies, a legal obligation to serve, and other forms of government intervention in the industry. In essence, universal service is equated with ubiquitous geographic coverage, universal household penetration and proactive government subsidies to achieve these goals. Ubiquitous telephone access in this sense is an expression of liberal egalitarianism, like universal schooling, literacy or voting rights.

The real key to the modern construction of universal service, however, was that it linked these political goals to a particular system of economic organization. That system was a protected monopoly which

continued from page 252
versal Telephone Service: Ready for the 21st Century?, Institute for Information Studies, 1991.

[2]Federal Communications Commission, *Investigation of the Telephone Industry in the US*, Arno, New York, 1939.

[3]David Gabel, 'The evolution of a market: the emergence of regulation in the telephone industry of Wisconsin, 1893–1917', PhD dissertation, University of Wisconsin, 1987.

[4]Gerald, Brock, *The Telecommuncations Industry*, Harvard University Press, Cambridge, MA, 1981.

[5]Kenneth Lipartito, *The Bell System and Regional Business*, Johns Hopkins University Press, Baltimore, MD, 1989.

[6]John V. Langdale, 'The growth of long distance telephony in the Bell System', *Journal of Historical Geography*, Vol 4, No 2, 1978, pp 145–159.

[7]Claude Fischer, 'The revolution in rural telephony, 1900–1920', *Journal of Social History*, Vol 21, No 1, Fall 1987, pp 5–26.

[8]American Telephone & Telegraph Company, *Annual Report*, 1907, pp 17–18.

sustained itself via averaged rates and revenue-pooling arrangements. The total revenues of the system sustained the network as a whole, but rates did not necessarily reflect the incremental costs of its individual components. This approach to the economic sustainability of the telephone network I will call the 'system perspective'.

A description of its internal workings can quickly become complicated, but the system perspective had three essential elements:

- The methods which separated local and long-distance service for cost allocation and revenue recovery purposes were determined on an end-to-end basis.[9]
- Long-distance rates were determined by geographic averaging; that is, they were based entirely on distance and not on route density.
- Long-distance revenues were pooled and the settlement payouts to local exchange carriers which originated the call were proportionate to their costs. Thus the surplus from low-cost routes and exchanges helped to sustain higher-cost exchanges and routes. The internal procedures governing these revenue flows and cost allocations came to be known as separations and settlements.[10]

According to the conventional wisdom, universal service was a public policy mandated by the 1934 Communications Act, and consciously brought into being by regulated monopolies organized according to the system perspective. A typical statement of this view appeared recently in a 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.'[11] The crowning achievement of this system, so the story goes, was the 92% household penetration ratio of the telephone just prior to divestiture. A recent book on the topic by academics echoes this claim and reasserts regulators' role in promoting universal service by means of rate subsidies: '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'.[12]

Although this construction of universal service has a powerful grip on the telephone industry, regulatory circles and many academics, its historical claims are questionable. The words 'universal service', for example, never appear in the Communications Act. Suggestions of a link between universal service policy and the Act are directly contradicted by its legislative history. The law was passed in 1934 to consolidate into one agency the communications regulatory functions of the Interstate Commerce Commission and the Federal Radio Commission, not to establish any new goals or policies. The bill's House sponsor, Sam Rayburn, explicitly stated that the Act did not change existing law.[13] There is nothing in the text of the Act which can be construed as mandating or even suggesting a policy of subsidizing telephone penetration.[14]

There is a bit more substance to the alleged link between the system perspective and the pursuit of universal service, but even here the connection is tenuous and misinterpreted. The advent of rate-base regulation in the 1920s led to a long struggle over the board-to-board versus the station-to-station method of separating the costs of local and long-distance service. Supreme Court decisions in 1930 and 1933

[9]A series of court and regulatory decisions beginning with *Smith et al* v *Illinois Bell*, 282 US 133 (1930), and *Lindheimer* v *Illinois Bell*, 292 US 151 (1933), approved the station-to-station separation method over the board-to-board method. This meant that some of the costs of the local exchange plant were recovered from interstate long-distance calls because local facilities were used to make interstate calls.

[10]For a detailed history and description of separations and settlements procedures see Carol M. Weinhaus and Anthony Oettinger, *Behind the Telephone Debates*, Ablex, Norwood, NJ, 1988.

[11]Barbara J. Farrah and Mike Maxwell, 'Building the American infostructure', *Telephony*, 20 April 1992, p 45.

[12]Entman, 'Introduction', in Institute for Information Studies, *op cit*, Ref 1.

[13]Rayburn said: '. . . 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), cited in D. Horwitz, *The Irony of Regulation Reform*, Oxford University Press, New York, 1989, p 122, fn 136.

[14]Those who read the modern construction of universal service into the preamble's call for 'a rapid, efficient, nationwide and world-wide wire and radio communication service with adequate facilities and reasonable charges' are projecting a modern preconception onto the past. There is little to indicate that this was anything more than boilerplate rhetoric when it was written. 'Adequate facilities and reasonable charges' could be part of the mandate of any regulatory commission of the period. 'Rapid, efficient, nationwide wire and radio service' is simply a statement of a broadly desirable goal with nothing said about how to achieve it.

sanctioned the station-to-station principle, but a uniform, nationwide system of separations and settlements based on this principle was not actually put together until 1947.[15] In the debate over cost separation methods and policies up to 1947, there is no evidence of regulators' intentions to subsidize telephone penetration. One finds instead a complex set of compromises and negotiations among AT&T, state commissions, federal regulators, large independents and small independents designed to solve the problems caused by the application of rate-base, rate-of-return regulation to a network that offered multiple products and spanned multiple jurisdictions.[16] The real issue was not the promotion of universal service in the modern sense, but (1) how to define reasonable rates while ensuring that telephone companies would be adequately compensated for all of their properties, and (2) how to separate the rate base into federal and state jurisdictions.[17]

True, separations and settlements were based on the premise that the network was an integrated whole, not a combination of discrete routes and services; therefore all exchanges and routes had to be taken into account in determining the costs of service for rate-making purposes. This, as we shall see, was an extension of the earlier concept of universal service that developed in the competitive period. But revenue settlements and cost separations were not originally conceived as a mechanism for subsidizing some users or regions at the expense of others. Nor were they part of an explicitly formulated law or policy mandating universal service. They were a set of practices that evolved gradually from the regulated monopoly framework.

Not until the late 1960s and early 1970s is there evidence that regulators began to consciously manipulate the separations and settlements process in order to subsidize residential rates. The crucial change came with the adoption of the Ozark plan in 1970, which shifted ever-larger portions of the local non-traffic-sensitive plant to be recovered from interstate (long-distance) revenues.[18] Ironically, this move to exploit the social policy possibilities of the separations and settlements process came at a time when the justification for such a subsidy was weak, as at least 85% of all US households already had telephone service.[19] Telephones were becoming universal for much the same reason that automobiles and television sets became universal: Americans wanted them and their increasing affluence made it possible for most of them to get them. It should be noted, too, that keeping residential rates artificially low is not synonymous with the promotion of universal service. Residential consumers' desire to pay less for service does not necessarily mean that subsidies were necessary to make service affordable.

Thus the modern notion of universal service, which links high household penetration with the separations and settlements practices of a regulated telephone monopoly, is a very recent construction. It is not a longstanding historical policy with its roots in the Communications Act. A uniform, nationally administered separations and settlements system based on the station-to-station principle was not fully operative until the 1950s. Cross-subsidies to promote household penetration did not emerge until the 1970s. And regulators' perceived need to administer a subsidy of residential users did not come until household penetration was already approaching 'universal' levels.

These observations lead to two important questions. First, if the modern construction of universal service developed as recently as I say

[15]The station-to-station method was not fully implemented until the adoption of the first uniform *Separations Manual* by the National Association of Regulatory Utility Commissioners and the FCC in 1947: see Weinhaus and Oettinger, *op cit*, Ref 10.

[16]Rate-base regulation demands that the rates charged by a telephone company for a particular service be based on the book costs of the physical plant used to provide the service, plus expenses and a reasonable rate of return. Applying this logic to a telephone call that occupies the facilities of two or three companies and crosses three different jurisdictional boundaries is no simple matter, because the same facilities are used for many different services. For example, one uses the same local access line and central office switch for a local call, an intrastate toll call, an interstate long-distance call and an international call.

[17]The record in *Smith* and *Lindheimer* indicates that the dispute revolved around the reasonableness of rates, and more specifically around the methods used to determine the size of the rate base. The Supreme Court ruled that separation of interstate and intrastate plant 'is essential to the appropriate recognition of the competent governmental authority in each field of regulation' and that some part of the local exchange plant should be 'apportioned' to interstate service, otherwise 'the exchange property . . . will bear an undue burden'. There is no indication that regulators were attempting to keep exchange rates low to stimulate telephone penetration, or that the regulators or the Supreme Court recognized subsidization of exchange access to promote universal service as a valid criterion in rate making.

[18]The Ozark plan, concluded in 1970 and implemented in 1971, shifted a growing portion of the local loop's cost recovery to the interstate jurisdiction. See Weinhaus and Oettinger, *op cit*, Ref 10, pp 83–103, for a description and analysis of the Ozark plan. Horwitz, *op cit*, Ref 13, p 235, notes that state regulators' support for the Ozark plan was partly a response to pressure from public interest groups to keep residential rates low.

[19]The FCC's *Statistics of Communications Common Carriers* for the year ended 31 December 1965 reported that 85% of all US households had telephone service; the *Statistics* for 1970 reported that 92% of all households had telephones. Because the method used to measure household penetration at that time is thought to have overstated the actual amount, I have deducted 7% from each estimate, which yields a household penetration percentage of 85% for 1970 and 78% for 1965. Federal Communications Commission, *Statistics of Communications Common Carriers*, FCC, Washington, DC, 1965 and 1970.

Table 1. Statistical abstract of the US telephone system, 1895.

Total US population:	69 471 144
Total number of telephones:	251 994
Overall penetration rate (%):	0.36
Business/residential ratio (%)	90/10

Distribution of telephones by size of city

Population category	% of US population	No of telephones	% of all telephones	Penetration rate (%)
Cities over 50 000	21	143 455	57	1.00
Cities 10 000–50 000	9	71 536	28	1.11
Towns 2500–10 000	8	28 441	11	0.51
Rural areas	62	8 562	3	0.02

Distribution of exchanges by size of city

Population category	No of places	Places with exchanges	% served
Cities over 50 000	72	72	100
Cities 10 000–50 000	294	288	98
Towns 2500–10 000	1297	474	37
Inc places under 2500	7710	259	3

Sources: 1890, 1900 Census; American Bell Telephone Co, Exchange Statistics Book, Bell Labs Archives.

it did, what did Theodore Vail and his contemporaries mean when they used the term back in 1907? Second, if the current meaning constitutes a decisive shift in usage, when, how, and why did the new meaning displace the older term? The answers to these questions have important implications, both for our understanding of the telephone's history, and for contemporary telecommunications policy.

Telephone scope and penetration before and after independent competition

Looking back after 15 years of independent competition, Theodore Vail claimed in 1909 that AT&T had pursued universal service from its inception.

> The Bell system was founded on the broad lines of 'One System,' 'One Policy,' 'Universal Service,' . . . This is no recent or new idea or theory. It is co-existent with the business. In fact the theory was evolved and developed before the business and [the business] has been developed on that theory.[20]

What did Vail mean by 'one system, one policy, universal service'? To read the modern construction into a statement made more than 80 years ago is obviously ahistorical and misleading. What Vail really meant by 'the theory' of universal service can be clarified by looking at the Bell system's actual behaviour.

Table 1 is an abstract of telephone development in the USA in 1895, after 18 years of Bell monopoly. There were 252 000 telephone subscribers in the country, a penetration rate of 0.0036. Their number was growing by only 5% annually, a rate at which it would have taken many centuries to achieve universal household penetration. More significant still is the geographic distribution of this number: 57% of the subscribers are in the 72 largest cities, although these cities account for only 21% of the total population. Ninety-seven per cent of the incorporated communities with less than 2500 people, and at least half of the cities with a population between 2500 and 10 000, had no telephone exchanges at all. Although 62% of the US population lived in rural areas in 1895, rural areas accounted for only 3% of the telephone subscribers. Equally significant, 90% of the users were businesses.

This pattern cannot be explained away by pointing to the alleged

[20]AT&T, *Annual Report*, 1909, pp 18–19.

Table 2. Statistical abstract of the US telephone system, 1920.

Total US population:	105 710 620
Total number of telephones:	13 411 400
Business/residence ratio (%):	45/55
Overall penetration rate (%):	12.69
Household penetration rate (%):	30.00
Penetration rate in farm households (%):	38.70

Telephone penetration in farm households in selected states

State	Penetration (%)
Iowa	86.10
Kansas	77.90
Nebraska	76.40
Illinois	73.20
Indiana	66.40
Missouri	62.20
Ohio	62.10
Minnesota	62.00

Sources: 1920 US Census; 1920 Farm Census; AT&T-Bell Labs Archives.

higher costs of serving smaller towns. At that time large urban exchanges were the most expensive and difficult to operate. Telephone service in the manual switching era was characterized by diseconomies of scale.[21] In large systems signalling was more complex, maintenance more expensive and labour less productive. The small-scale telephone switchboards needed by small towns and rural areas, on the other hand, were easy to manufacture and inexpensive to operate.

Nevertheless, Vail was not prevaricating; the Bell System *was* pursuing 'universal service' from 1878 to 1895 by its own lights. Far from pursuing social ubiquity in the modern sense, however, the Bell System in the 1880s was modelling itself after the telegraph system of the 1870s. The telegraph was a nationwide, 'universal', business-oriented message communications network linking terminals in all the principal commercial centres. It started in the largest cities and gradually spread to smaller ones, but it never reached households or rural areas. 'One system, one policy, universal service' meant a nationally interconnected, centrally coordinated monopoly like Western Union. This explains the Bell System's emphasis on long-distance development, often to the detriment of local and short-haul toll development.[22] The urban, long-distance and business-user bias of the Bell System was not a product of economic or technical limitations. It was a deliberate business policy.

Western Union achieved its dominance of the industry by being the first to develop a nationally interconnected network. It used its leverage over interconnection to isolate and destroy its rivals.[23] Bell planned to follow in its footsteps. When Vail claimed that Bell's concept of universal service preceded the telephone business he meant it quite literally – the concept was drawn from his experience in and observations of the telegraph business.[24]

Until the intervention of the independents, then, the telephone in the USA was on the same slow and restrictive developmental trajectory as in Europe. Fortunately for the USA, the expiration of Alexander Graham Bell's basic telephone patents in 1893 and 1894 allowed independent equipment manufacturers and exchange service providers to enter the market. The data in Table 1 make an interesting comparison with the statistics in Table 2. Table 2 shows the state of telephone development in 1920, at the end of the competitive era. A dramatic change in the social role of the telephone is evident. Residential users already outnumber business subscribers by a substantial margin. Not

[21]Milton L. Mueller, 'The switchboard problem: scale signaling and organization in manual telephone switching, 1878–1898', *Technology and Culture*, Vol 30, No 3, July 1989, pp 534–560.

[22]See Ref 34 for evidence of Bell's policy of placing intercity long-distance connections at a higher priority than local and short-haul toll connections.

[23]Robert L. Thompson, *Wiring a Continent: The History of the Telegraph Industry in the United States, 1832–1866*, Princeton University Press, Princeton, NJ, 1947; Brock, *op cit*, Ref 4.

[24]Vail's biographer supports these claims, observing that Vail worked as a telegrapher for Union Pacific in the 1860s. During negotiations with Western Union over the right to develop the telephone, Vail insisted on Bell's right to own and operate long-distance lines. Vail's own testimony in the 1918 antitrust case also strongly reasserts AT&T's intention to 'control the business' by controlling long-distance connections just as Western Union had done. See Brock, *op cit*, Ref 4, p 102.

Table 3. Telephone penetration growth (%), 1895–1912.

	USA	Europe
1895	0.36	0.25
1902	2.30	0.30
1912	8.80	0.70

Source: American Telephone & Telegraph Co, 'Telephone statistics of the world', 12 May 1912.

[25]See O.E. Noel, President and General Manager, East Tennessee Telephone Company, to C. Jay French, General Manager, American Bell Telephone Company, December 1894, Box 1066, AT&T-Bell Labs Archives.

[26]See C.A. Nicholson, Central New York Telephone Co, to C. Jay French, American Bell Telephone Co, 6 April 1898, Box 1166, AT&T-Bell Labs Archives.

[27]Common carrier law prohibited discrimination against members of the public and was construed to require competing telegraph companies to exchange messages. But the courts' reading of the common carrier obligation did not require interconnection of competing telephone companies. The most important legal precedent was the Supreme Court decision in the railroad 'Express' cases, 117 US 601 (1886), which held that common carriers were required to serve the *public* indiscriminately, but this did *not* mean that they had to be a 'common carrier of common carriers'. See also *Postal Telegraph Cable Co* v *Hudson River Telephone Co* 467 Supreme Court (1887). For application of these precedents to telephone cases, see opinion of Judge Siebacker, *Dane County Telephone Co* v *Western Union Telegraph Co*, Box 1298, AT&T-Bell Labs Archives; *Syracuse Standard*, 2 July 1898, Box 1166, AT&T-Bell Labs Archives.

[28]See F.R. Colvin to President Hudson, 8 April 1896, Box 1298, AT&T-Bell Labs Archives, a report by a Bell spy on a meeting of the Ohio Independent Telephone Association. One of the independents had initiated litigation to force Bell to connect with it, but the association unanimously asked the company to withdraw its suit.

[29]The FCC *Telephone Investigation*, *op cit*, Ref 2, p 133, characterized it as 'wasteful from the viewpoint of investment and [a] burden on both the telephone operating companies and the rate payer'. Stehman expresses a similar view: G. Warren Stehman, *The Financial History of American Telephone and Telegraph Company*, Houghton-Mifflin, Boston, MA, 1925. An otherwise thoughtful and well-researched treatment of the competitive era by Lipartito, *op cit*, Ref 5, contains no discussion of the refusal to interconnect and its consequences for the competitive struggle.

[30]John Wenders, *The Economics of Telecommunications: Theory and Policy*, Ballinger, Cambridge, MA, 1987, pp 171–183; D. Evans and J. Heckman, 'The early

continued on page 359

only was there an enormous expansion in the number of rural and small-town exchanges and a rapid rise in telephone penetration generally, but rural households have, on average, the *highest* levels of telephone penetration. In 1920, 38.7% of US farms had telephones, versus 30% of all US households. Farm households in states such as Ohio, Indiana, Illinois, Kansas and Nebraska, where independent telephony was strongest, reported subscription rates of 60% and 70%. The most surprising statistic relates to Iowa, where 86% of the 213 439 farms reported telephones in 1920.

In other words, only after the competitive period do we see the kind of geographic and social penetration that can support the modern notion of universal-service-as-social-ubiquity. If telephone development in the USA by 1920 is contrasted with Europe, where with the exception of certain Scandinavian countries independent competition was non-existent, the uniqueness of the US experience is even more evident (see Table 3).

The phenomenon of access competition

The role of competition in the early 1900s in promoting the growth of telephony is widely acknowledged by historians. A crucial part of the story has been overlooked, however. What makes the Bell–independent rivalry particularly interesting is the unique form the competition took. In sharp contrast to the telecommunications competition of the present day, which relies upon interconnection of the competing networks, the Bell System and the independents refused to interconnect with each other. This form of exchange competition was known at the time as 'dual service'. I refer to it as 'access competition'.

Access competition was the outcome of several business strategy and court decisions made between 1894 and 1898. From the beginning, the national Bell organization refused to authorize its licensee companies to connect with the toll lines of 'opposition' companies.[25] It also suppressed attempts by local exchange competitors to subscribe to the Bell exchange and hook up the Bell line to an independent company central office.[26] Efforts to legally compel interconnection were prevented by the prevailing interpretation of common carrier law.[27]

Ultimately the most important barrier to interconnection was that by 1898 the organized independent movement itself stopped seeking interconnection and lobbied against efforts by state legislatures to require the two systems to interconnect.[28] In the first four years after the patent expiration, the independents rapidly established a presence in the small towns and rural areas neglected by Bell. Interconnection with the Bell System would have taken away their exclusive control of connections to these areas. The independents came to believe that they could beat the Bell System and had no need to join it. In combination, these decisions ensured that competition would take the form of rivalry between separate, unconnected systems.

The phenomenon of access competition has attracted little attention and analysis from historians. When it is commented upon at all, it is typically dismissed as a perverse and destructive tactic.[29] Economists generally characterize it as an anti-competitive tying agreement or refusal to deal which raises antitrust concerns.[30] In contemporary telecommunications competition, competing carriers are legally required to interconnect their systems. No network is allowed to obtain a

competitive advantage from its 'bottleneck' control of access to subscribers.

The contemporary bias towards interconnection has blinded most historians to one of the central features of the competitive era: the peculiar dialectic regarding the universality of telephone service created by the lack of interconnection. Access competition at once restricted and promoted the ubiquity of the telephone. It restricted the system's universality because it fragmented telephone users into two groups. In cities with competing exchanges, larger business and government offices had to subscribe to both exchanges so that everyone could call them.[31] Users served by only a Bell exchange could not call users in nearby towns served only by an independent system.

At the same time, however, access competition rewarded the pursuit of universality by the telephone companies. A telephone system with more people on it is, *ceteris paribus*, more valuable than one with fewer subscribers. Each system becomes more valuable to its subscribers and gains a competitive advantage over its rival as it extends telephone service to more users and more locations. When competing systems are not connected, each system is permitted to fully appropriate the value of its subscriber universe by excluding its competitors from access to its network. This phenomenon, known as the 'network externality', has been explored in depth in recent economic theoretical literature.[32]

If the value of a telephone system increases as the number of subscribers (technically, its scope) increases; if neither network can grow by means of interconnecting with a rival system; and if the bulk of the market for telephone service is not yet developed, then access competition creates three powerful incentives to pursue universality: (1) it rewards the first to establish telephone exchanges in unserved areas; (2) it creates pressure to make the price of access as low as possible, so as to attract new subscribers and draw away subscribers from the other system; (3) it rewards those who interconnect local exchanges with toll lines as quickly and as extensively as possible. In sum, access competition places the highest possible premium on the scope of a network. This premium was the driving force behind the Bell–independent rivalry of the early 1900s.

Access competition was not a minor episode in the history of the US telephone system. It lasted for 25 years (longer than our current experience with long-distance competition). At its peak, around 1904–14, more than 55% of the US population lived in cities or towns where there were two unconnected telephone exchanges (see Table 4).[33] As late as 1924 dual service still existed in several major cities.

Access competition in the making of universal service

This section goes into some detail about the progress of telephone competition in the early 1900s. Although the material may seem familiar, a detailed recounting is necessary to prove that it was the absence of interconnection, and not simply competition as such, which promoted the telephone companies' pursuit of universality. The points of departure are the three incentives identified in the preceding section.

Reaching unserved areas

The independents achieved their initial successes by establishing exchanges in the medium and small-sized market towns Bell had ignored.

continued from page 358
history of competition in the telephone industry', in *Breaking Up Bell*, North-Holland, New York, 1983, pp 32–33; Gabel, *op cit*, Ref 3, p 354.

[31]In a typical city with evenly matched competing exchanges, approximately 12% of the users took duplicate subscriptions. The vast majority of the duplicate subscribers were businesses, and the probability of duplication increased with the scale of the business. Cumberland Telephone & Telegraph Co, acquisition of Central Home Telephone & Telegraph, Kentucky, 1910. Box 39, AT&T-Bell Labs Archives.

[32]W. Brian Arthur, 'Competing technologies and lock-in by historical events', *Economic Journal*, Vol 99, 1989, pp 116–131; Paul A. David, 'Understanding the economics of QWERTY: the necessity of history', in W.N. Parker, ed, *Economic History and the Modern Economist*, Basil Blackwell, Oxford, UK, 1986; Joseph Farrell and Garth Saloner, 'Competition, compatibility and standards: the economics of horses, penguins and lemmings', in H.L. Gabel, ed, *Product Standardization and Competitive Strategy*, North-Holland, New York, 1987; Michael Katz and Carl Shapiro, 'Network externalities, competition, and compatibility', *American Economic Review*, Vol 75, No 3, 1985, pp 424–440; Michael L. Katz, 'Technology adoption in the presence of network externalities', *Journal of Political Economy*, Vol 94, 1986, p 822; Jeffrey Rohlfs, 'A theory of interdependent demand for a communications service', *Bell Journal of Economics and Management Science*, Vol 5, No 1, Spring 1974, pp 16–37.

[33]The most thorough documentation of the scope of dual service competition is the *Telephone and Telegraph Atlas*, a complete map of Bell and independent exchanges and toll lines compiled by AT&T between 1910 and 1912. The *Atlas* is in the AT&T-Bell Labs Archives, Warren, NJ.

Table 4. The growth of dual service, 1894–1909.

	A No of cities	B Per cent	C Population	D Per cent
1894	22	2	399 000	1
1898	249	30	6 189 000	23
1902	449	55	14 617 000	54
1906	466	57	15 263 000	57
1909	451	55	15 085 000	56

Only cities over 5000 in population counted. Column A is the number of cities 5000 or over in population with competing, non-connected exchanges. Column B is the number of cities with dual service as a percentage of all cities with a population over 5000. Column C is the total population of all cities over 5000 in population with dual service. Column D is the total population of cities with dual service as a percentage of all cities with a population of 5000 or more.

Sources: *Telephony* 1894–1909; *Telephone and Telegraph Atlas*, AT&T-Bell Labs Archives; Bell and independent exchange rates. 1912–13, Box 29. AT&T-Bell Labs Archives.

Unlike Bell, which concentrated on constructing a grand national system, the independents concentrated on connecting their exchanges with short-haul toll lines.[34] Frequently they made a conscious effort to cover territories Bell had missed in order to increase their leverage in the rivalry.[35] As the number of independent telephone subscribers in the surrounding countryside grew, the independents obtained the leverage to establish a competing exchange in the urban hubs. This pattern of the periphery advancing on the centre – the reverse of the Bell strategy – was repeated in numerous cities and suburbs.[36] In order to avoid losing the system rivalry, Bell was forced to build thousands of new exchanges and to greatly extend its short-distance toll network. New types of service more suitable to small towns were made available: Class F party lines were offered to link residents in thinly populated regions near a Bell exchange; 'Petersham' town service established public toll stations in places too small or remote for exchanges.[37]

By 1907 commercial independents had established 10 109 public exchanges – 10 times the number Bell had established after 18 years of monopoly – and the smaller farmer and rural organizations had established 17 702 small-scale telephone systems.[38] Prodded by competition, the Bell System opened approximately 4500 new exchanges in cities with fewer than 10 000 population during the same period.[39] A 1909 statement by a Southwest Bell representative confirmed that the expansion was a product of access competition. 'We have scraped along for the past ten years,' he said, 'building exchanges and toll lines that we ought not to have constructed *except for the purpose of causing the service to be more valuable than that of our adversary*.'[40]

Pricing access as low as possible

From 1894 to 1900 the average monthly rate for local exchange service dropped by more than half. This did not occur simply because competition drove monopolistic station equipment prices to cost. Telephone prices generally consist of two parts: a charge for access, and a charge for usage. Pricing after 1894 was deliberately constructed to minimize the access cost barrier in order to encourage large numbers of new subscribers to join. Both Bell and the independents established exchange service in areas at unprofitable rates in order to enlarge the scope of the network and increase toll usage revenues.[41] It was not unusual for Bell operating companies to temporarily set their rates at $1/month, or even to provide service for free, in cities where an independent exchange had taken away many of their subscribers.[42]

[34]A Bell subscriber in Quincy, IL, in 1894 could call Peoria (132 miles away), Springfield (102 miles away) and Chicago, but there were no Bell exchanges or toll lines connecting Quincy to the rest of its own county, nor were there any lines to the farmers and merchants in neighbouring Brown, Hancock and Pike counties. 'The story of the telephone in Quincy, Illinois', Theodore N. Vail Chapter, Telephone Pioneers of America (Illinois Bell Telephone Co, 1948).

[35]In 1896 the Secretary of the Ohio Independent Telephone Association wrote a letter to every independent exchange urging them to accelerate 'the construction of toll lines connecting towns so small as not to be reached by the Central Union [Bell] Co'. F.R. Colvin to President Hudson, 8 April 1896, Box 1298, AT&T-Bell Labs Archives.

[36]Buffalo, NY, Kansas City, MO, and Philadelphia, PA, all developed independent networks in the rural and suburban areas which then helped to sustain successful exchanges in the cities.

[37]Party Line Development, 1898–99; Telephone Service for Small Exchanges, 1894, Box 1258, AT&T-Bell Labs Archives.

[38]Telephone Census, 1907.

[39]*Ibid.*

[40]*Telephony*, 30 January 1909; emphasis added.

[41]Doolittle Letter Book 12, p 331 (1896), AT&T-Bell Labs Archives. Thomas Doolittle's advocacy of the demand for interdependence of local exchange and long-distance service influenced American Bell President Fish, who wrote in 1902: 'it is at least worth considering whether or not cheap exchanges in the small towns do not add enough to the toll business to make them a proper investment, even if there is no profit in the small exchanges'.

[42]Lipartito, *op cit*, Ref 5, p 120; Gabel, *op cit*, Ref 3, pp 88–97.

Since the value of the Bell exchange diminished as its subscriber base shrank, Bell felt compelled to retain subscribers at practically any cost.

The need to maintain a large subscriber universe also affected the structure of the technology. Both contestants began to offer inexpensive four-party, eight-party and even ten-party lines to increase their subscriber universe.[43] The object was to get as many subscribers onto the system as quickly and as cheaply as possible.

Interconnecting with other exchanges

Interconnection with exchanges in other locations proved to be a rapid and relatively inexpensive way for a telephone company to enlarge its subscriber universe. The independents formed state associations to facilitate coordination. In contrast to the longstanding myth that the independent companies were exclusively local, they succeeded in establishing commercial long-distance companies that were regional in scope.[44] Bell was forced to respond by expanding its toll network and rationalizing its operator practices involved in transferring long-distance calls. Eventually, competitive pressures forced Bell to liberalize its no-interconnection-with-independents policy, and it began to develop connecting arrangements with independent exchanges and farmer lines in areas where it had no presence.

Previous historical work has not recognized the extent to which rural telephone development was driven by the imperatives of access competition. By 1902 about 6000 farmer lines and rural mutuals had been established, and 15 598 rural lines were being run on a commercial basis.[45] Independent and Bell alike took note of what came to be known as 'the farm line proposition', ie negotiations over which system the farm lines would choose to interconnect with. Bell's New York and Pennsylvania operating company developed two special rural line contracts, one to establish a small switching station in the farm houses the other to connect farm lines to a toll station on Bell System lines.[46] To farmers who built and maintained their own lines, Bell offered to interconnect for only $2 per year, compared to the $10 per year offered by the independents.[47]

One Bell manager who was particularly active in urging his local managers to go after the farmers said, 'I say to you managers that whenever you have the farmers tied on to your exchange you have got the merchants where you want them.'[48] Another Bell manager, decrying the lack of rural development of the Bell System in the Rocky Mountain area, warned that if the independent got the farmers 'he has anchored his exchange'.[49] Thanks to access competition, the once-neglected farmer became a highly sought-after prize.

By far the most important and effective policy Bell adopted in response to access competition was sublicensing of independent exchanges. Sublicensing referred to an interconnection agreement between Bell and an independent exchange located in a community unserved by Bell. As part of the contract, the independent agreed not to connect with independent exchanges or toll lines operating in direct competition with Bell. In turn, Bell agreed not to establish a competing exchange in that community.[50] Sublicensing could be an attractive option to both parties: the independent's customers could call over Bell toll lines and access Bell exchange subscribers in the region; Bell expanded its scope at the expense of the independent opposition movement.

[43]Party Line Development, 1896–99, *op cit*, Ref 37.

[44]The United States Telephone Company of Ohio, the Kinloch system of St Louis, the Kansas City Home Telephone Company and at least 15 other independent long-distance networks offered competitive toll service of up to 300 miles in length. See Pickernell to Hall, 12 May 1909, Box 1376, AT&T-Bell Labs Archives; 1908 *Annual Report* of the US Telephone Company, Box 36, AT&T-Bell Labs Archives.

[45]Telephone Census, 1902.

[46]General Order 34, 14 February 1900, Box 1330, AT&T-Bell Labs Archives. See also Lipartito, *op cit*, Ref 5, p 152.

[47]*Telephony*, Vol 17, No 13, 27 March 1909.

[48]*Cumberland Telephone Journal*, Vol 10, No 1, 15 January 1904, p 12. AT&T-Bell Labs Archives.

[49]*Ibid.*

[50]General Managers Letter Book 632, 31 October 1901, AT&T-Bell Labs Archives.

The sublicensing policy was first adopted by the national organization in 1901 and was further liberalized in 1908. The internal Bell debate over sublicensing independent exchanges underscores the fact that service to all the country was not part of the original conception of the business as Vail claimed. AT&T's Chief Engineer in 1900 lamented that the Bell System had not appreciated the need for subcontracting with independent operators earlier: '[If] it could have been foreseen what an extensive development of the telephone business would be required to meet the needs of the people, and the amount of capital involved, it would have been good policy . . . to have encouraged [Bell licensee companies] to sublicense to local people the right to furnish service in country districts and villages and towns.'[51] Clearly, Bell had not anticipated the universal demand for the telephone.

In the South the percentage of Bell System telephone connections provided by independent sublicensees grew to 41% by 1909.[52] In Ohio, Indiana and Illinois, hotbeds of independents competition, Bell in 1907 owned and operated only 310 exchanges representing 188 000 telephones, while independent sublicensees accounted for 777 exchanges and 192 000 telephone subscribers.[53] In Missouri and Kansas sublicensed telephone stations outnumbered Bell-owned telephones by two to one.[54] Theodore Vail's policy memoranda of this period state explicitly that 'unremunerative' rural and small-town areas should be left to independent sublicensees.[55]

The progress of sublicensing has been documented well in Langdale[56] and Lipartito.[57] But its significance in the context of access competition and its implications for the universal service myth have not been fully appreciated. Despite Bell's later claims that universal service (in the modern sense) was its policy from the beginning, Bell ultimately obtained most of its access to small-town and rural America through sublicensing agreements with independent companies. Moreover, its decision to 'reach out and touch' the rural areas was not a product of its own commitment to universal service, but a policy forced upon it by the rigours of access competition.

Each of the three sections above demonstrates how access competition promoted a universal telephone infrastructure by placing a premium on a network's scope. Had the competitors been interconnected, on the other hand, the incentives to pursue universality would have been greatly weakened. Independent competitors would have found it much easier to establish service in the urban areas already developed by Bell, and could have concentrated on simply undercutting Bell's price. The Bell System might never have undertaken the massive capital investments required to enlarge its exchanges in outlying areas and its network of long-distance lines, as these investments would not have given it a competitive advantage over the less extensive networks of the independents. Likewise, the independents would have had no incentive to construct alternative toll networks to connect independent exchanges. Incentives to restructure the technology to cheapen the cost of access would have been less powerful. Neither Bell nor the commercial independents would have needed to be in any hurry to reach out to the rural areas and smaller towns, because with interconnection it would not have mattered which system reached them first.

Universal service as Bell's antidote to access competition

By 1907 the USA was a house divided telephonically. The independents

[51]Joseph Davis to President Fish, 23 October 1901. AT&T-Bell Labs Archives.
[52]Lipartito, *op cit*, Ref 5, p 134.
[53]Central Union Co, *Annual Report*, 1907. AT&T-Bell Labs Archives.
[54]Pickernell to Hall, AT&T, 12 May 1909. Box 1376, AT&T-Bell Labs Archives.
[55]Vail's Circular Letter of 10 February 1908 urged licensee companies to 'pursue vigorously the policy of sublicensing' in the part of their territory which was 'more or less unremunerative'. Box 1364, AT&T-Bell Labs Archives.
[56]Langdale, *op cit*, Ref 6.
[57]Lipartito, *op cit*, Ref 5.

controlled 49% of the nation's telephones, and dual exchanges existed in about 57% of the cities.[58] It was at the juncture that Theodore Vail, newly re-installed as President of AT&T, began to promulgate the philosophy of universal service. The term and the doctrine never appeared explicitly before that time, although of course there had been debates about the merits of a competitive, fragmented service versus a monopolistic, unified service.

Vail articulated the universal service doctrine in the company annual reports from 1907 to 1914. The reports were as much political pamphlets as business documents: they were sent to thousands of newspapers and opinion leaders as well as the company's stockholders. In the reports Vail hammered away at the thesis that only a system that was 'universal, interdependent and intercommunicating' could realize the telephone's potential.

What did Vail mean by universal service? The primary thrust of Vail's doctrine was not a commitment to put a telephone in every home or an exchange in every community. Rather, Vail conceived of universal service as an integrated monopoly that could *interconnect all telephone users*. Implementing this vision required eliminating access competition. Indeed, it is impossible to understand the thrust of Vail's arguments unless it is seen as a critique of, and alternative to, access competition.

The following statement from the 1910 *Annual Report* contains the essence of Vail's conception of universal service:

> [The Bell System] believes that the telephone system should be universal, interdependent and intercommunicating, affording opportunity for any subscriber to any exchange to communicate with any other subscriber of any other exchange within the limits of speaking distance.[59]

Contemporary readers can easily misinterpret Vail's references to universality as a commitment to social ubiquity. Vail did in fact make rhetorical jabs in that direction, although they were notable for their vagueness.[60] The uniqueness of Vail's vision, however, lay not in any alleged commitment to extend service everywhere and to everyone. At this juncture no one disputed either the desirability or the inevitability of the telephone's rapid diffusion. Indeed, the independents far outstripped the Bell System in their commitment to extend telephone service to previously unserved areas. What set the Bell policy apart was its commitment to interconnect all telephone users into one big, integrated system.

Vail's doctrine of universal service had three basic components. The first was that the value of telephone service grew as the number of subscribers grew. 'A telephone without a connection at the other end of the line is not even a toy or a scientific instrument. It is one of the most useless things in the world. Its value depends on the connection with other telephones – and increases with the number of connections.'[61] Vail's acute recognition of the network externality provided the basis for his critique of access competition. Competing exchanges fragmented the telephone calling universe, thus diminishing the value of the service. Those who subscribed to one system, he said, received 'a partial value [which] cannot be satisfactory', while 'important users' were forced to take out duplicate subscriptions.[62] To this unwelcome predicament Vail contrasted his alternative:

> The fundamental idea of the Bell System is that the telephone service should be universal, intercommunicating, and interdependent; that there are *certain*

[58] Telephone Census, 1907.
[59] AT&T, *Annual Report*, 1910, p 43.
[60] '[The Bell System] believes that some sort of a connection with the telephone system should be within the reach of all': *ibid*. Just what 'sort of a connection' and the meaning of 'within the reach of' are left unspecified.
[61] AT&T, *Annual Report*, 1908, p 21.
[62] AT&T, *Annual Report*, 1907, p 17.

people with whom *one communicates frequently* and *regularly*; there are a *certain few* with whom one *communicates occasionally*, while there are *times* when it is *most necessary to get* communication with *some other one*, who, until the *particular necessity* arose, *might have been unknown and unthought of. It is this necessity, impossible to predetermine, which makes the universal service the only perfect service*.[63]

Clearly, the network externality implies that a continuous broadening of telephone penetration would be beneficial to users. But just as clearly, Vail's reference to 'the universal service' in this context was not a commitment to extend service to everyone; it simply meant that those who did have telephone service should be accessible to each other and not fragmented into competing exchanges. If the growth of penetration *per se* had been the primary issue, Vail's argument against telephone competition would have lacked any force, for no one disputed the stunning increases in telephone diffusion that had occurred because of competition.

The second pillar of Vail's argument was the claim that universal intercommunication required centralized control and coordination; that is, service should be provided by, or under the control of, a single firm:

> The Bell system was founded on the broad lines of 'One System,' 'One Policy,' 'Universal Service,' on the idea that no aggregation of isolated independent systems not under common control, however well built or equipped, could give the public the service that the interdependent, intercommunicating, universal system could give.[64]

Here again interconnection, not social or geographic ubiquity, is the basic issue being addressed. Unless the network developed under the guidance of a single firm, Vail contended, telephone users' ability to make connections with exchanges in other locations would be thwarted by a lack of coordination and by technical incompatibility. This commitment to system compatibility cannot be equated with a commitment to social ubiquity, although compatibility is of course a precondition of social ubiquity.

The third element of the universal service doctrine was the proposition that monopoly, and not interchange of traffic among the competing systems, was the best way to achieve universal service. From 1907 to 1914 compulsory interconnection became an increasingly common demand among utility regulators. Vail condemned interconnection of competitors as unfair, because it allowed smaller competitors to share in the benefits of the Bell System's larger access universe. Such competition would parasitize the larger system, and amounted to legalized confiscation of its property.[65] Interconnection would also create a messy, heterogeneous telephone system which would lack the technical integrity and coordination of a single system.[66] Having made the case for an integrated monopoly, Vail indicated that he was willing to accept commission regulation of rates and service.

Vail's vision infused the Bell System with a new coherence. 'Universal service' became a competitive strategy, a political slogan and a catchy advertising term all in one. Bell's ability to offer connections to more locations than its rival independent exchanges was its greatest competitive advantage. Instead of fighting to eliminate all independents, it would absorb them into the 'universal' system by making them non-competitive feeders through sublicensing.[67] Advertisements played up the larger scope of the Bell System.[68] Above all, universal service was the spearhead of Vail's drive to achieve political support for the

[63]AT&T, *Annual Report*, 1910, p 39; emphasis in original.
[64]AT&T, *Annual Report*, 1909, p 18.
[65]AT&T, *Annual Report*, 1910, pp 44–46.
[66]*Ibid*, pp 46–47.
[67]See Ref 50 above.
[68]Full-page ads were placed in the *Bulletin of the League of Municipalities* in 1912 and in other magazines. The ads compared the Bell System to the root system of a tree and to the Nile river and its tributaries and claimed 'because they are connected and working together, each of the 7 000 0000 telephones in the Bell system is an integral part of the service which provides the most efficient means of instantaneous communication'.

elimination of competition and the establishment of regulated monopoly.[69] It provided an appealing rationale for the consolidation of competing exchanges that could be used to counter growing antitrust challenges to Bell's dominance.

At that time dual service and universal service both commanded powerful support. From 1910 to 1920 state regulatory commissions, which generally favoured unification and regulation of telephone service, were gradually gaining greater authority over telephony. Many business user groups also supported universal service because of the perceived expense and inconvenience of duplicate subscriptions.[70] Independent competitors and their allies in state legislatures, on the other hand, actively blocked consolidations with the aid of state antitrust laws. They were backed by federal antitrust authorities. The support of the general public was up for grabs: they loathed monopolies and remembered the high prices and unresponsive service prior to independent entry, but were sometimes inconvenienced by fragmentation.

The December 1913 Kingsbury Commitment, which has long been misrepresented as the beginning of Bell–independent interconnection, was actually a short-lived and completely ineffective attempt to find a middle ground between dual service and universal service. The Kingsbury Commitment was intended to establish the basis for long-distance interconnection while preserving dual service at the local exchange level. Insofar as it accomplished anything, it represented a victory for the dual service advocates. It prevented Bell from financial acquisition of competing independents and explicitly exempted from interconnection all exchanges that operated within a 50-mile radius of each other.[71] On the other hand, there is no evidence that a single independent ever availed itself of its costly and non-reciprocal toll interconnection arrangements.[72]

In the end the concept of a unified telephone network won the support of the industry, the public and the regulators. As the nation became more urbanized and integrated, many telephone users, particularly small and medium-sized businesses, found a divided service to be intolerable. Consolidation of the telephone systems at either the state or the municipal level increased in frequency.[73] Unification of the service after 1914 was generally a deliberate, publicly mediated process involving city councils, state legislatures, state regulatory commissions and in some cases even statewide public referendums.[74] The federal Willis-Graham Act of 1921 removed the last of the legal obstacles to consolidation by suspending the Kingsbury Commitment and exempting telephone companies from the Sherman antitrust act. In contrast to most of the economics and utility regulation textbooks written decades afterwards, telephone monopoly never emerged because of supply-side economies of scale. It emerged because of the demand-side economies of scope created by universal interconnection.

The Willis-Graham Act is generally treated as the official close of the competitive era. Historians' blind spot with respect to access competition, however, often makes them overlook the fact that the bill's author, Senator Graham, stated explicitly that the main rationale for the law was the elimination of fragmentation caused by access competition.[75] Stehman,[76] Herring and Gross[77] and all other utility texts of the period also observed explicitly that 'unification of the service' was the rationale for the choice of regulated monopoly in telephony. To regulators and politicians as well as users and the telephone companies, 'universal

[69]Bell's positioning of itself as the universal system successfully concealed its own refusal to eliminate fragmentation by interconnecting with its independent competitors. Bell strategically withheld the benefits of a unified service from the public and the independents until it had succeeded in winning support for regulated monopoly as the industry structure.

[70]For evidence of business support for telephone service unification see Delos F. Wilcox, *Municipal Franchises*, Gervaise Press, Rochester, NY, 1910, pp 240–241; Chicago City Council hearings, 4 November 1907, pp 2023–2024.

[71]Subsection (3) of the Kingsbury Commitment text limited Bell–independent interconnection to an exchange 'which is more than fifty miles distant from the exchange in which the call originates'. Kingsbury Commitment, 19 December 1913, p 2.

[72]For accounts of unsuccessful attempts by independents to connect to Bell under the terms of the Kingsbury Commitment see B.G. Hubbell, Federal Telephone Co, to N.C. Kingsbury, 8 October 1914; and W.H. Bassett, Kinloch Telephone Co, to N.C. Kingsbury, 3 July 1917, Box 16, AT&T-Bell Labs Archives.

[73]Major consolidations in Kansas City (Box 17), Ohio (Box 1357), Louisville (Box 39), Indianapolis (Box 36), Los Angeles (Box 18) and Buffalo (Box 25) took place between 1912 and 1925. (AT&T-Bell Labs Archives.) Each of these cases was a negotiated and publicly mediated choice of monopoly in which the end of subscriber fragmentation was the decisive factor.

[74]Kentucky, for example, held a statewide referendum in 1918 to amend its state constitution to permit telephone consolidation. The federal government passed the Willis-Graham Act in 1921 explicitly to exempt telephone companies from antitrust laws which were viewed as preventing unification of the service. Box 39, AT&T-Bell Labs Archives.

[75]'There is nothing more exasperating, nothing that annoys the ordinary business man or the ordinary person more than to have two competing local telephone systems, so that he must have in his house and in his office two telephones, on neither one of which he can get all the people he wants to be in communication with.' 67th Congress, 1st Session, *Congressional Record*, 1 June 1921, p 1966.

[76]*Op cit*, Ref 29.

[77]James M. Herring and Gerald C. Gross, *Telecommunications: Economics and Regulation*, McGraw-Hill, New York, 1936.

service' meant a unified, interconnected monopoly, not regulatory subsidies to promote household penetration.

The legacy of access competition had an important impact on the ensuing era of regulation. Many of the economic characteristics classically associated with regulated monopoly, such as low access rates and toll-to-exchange access 'subsidies', had their origin in the competitive period. System competition prodded both Bell and the independents to take a system approach to revenue recovery; both Bell and the independents established service in locations that may have been unremunerative on a stand-alone basis, but nevertheless contributed to the value of their overall networks by increasing the subscriber universe and stimulating toll usage.

As the telephone companies contemplated regulation, they looked for ways to ensure that it would not penalize them for extending their networks. A detailed discussion of this problem was written by a manager of the Chicago Telephone Company in January 1912. Anticipating rate regulation based on traditional norms of 'reasonableness', he drafted a memo entitled 'A telephone property must be considered as a whole in determining the reasonableness of any rate'. The memo clearly reveals the correlation between Bell's adoption of a system perspective, its pursuit of universal service, and the threat of competition:

> With the telephone stations and lines, all do not and cannot in the very nature of things pay their way, yet they must be continued as a necessary part of the whole system. [T]he elimination of one, lessens the service and economic value of the part that remains . . . a company which does not meet the demands for service and extend [service] beyond the bounds of the local community into the rural districts then to the next community and so on, fails or is forced to give way to an enterprising rival which will provide such an extensive and comprehensive service.[78]

In the past the primary motive for extending the service had been that failure to do so would lose the business to an 'enterprising rival'. Now that competition was waning and the company was faced with both restrictions on market exit and commission regulation of their rates, it wanted to make sure that the concept of 'reasonable rates' took into account all its properties, including the less profitable ones in outlying areas. Regulators, who were also committed to universal service, developed their separations and settlement procedures accordingly.

The retrospective construction of an ideology

As noted before, in the late 1960s regulators began to use their control of separations and settlements to subsidize household rates. The modern redefinition of universal service occurred when these new cross-subsidy practices were threatened by competition in the 1970s. Competition struck at the heart of the system perspective by targeting the routes and services which were overpriced and relying on interconnection with the monopoly network to access connections which were subsidized. The challenge of new entry forced the system to develop an explicit rationale for the system perspective in order to defend itself in the political arena. In the struggle the concept of universal service was reconstructed and linked to the practices of regulated monopoly. Regulated monopoly and its cross-subsidies were retrospectively credited with making telephone service universally available and affordable.

[78] H.O. Seymour, Chicago Telephone Company, 'A telephone property must be considered as a whole in determining the reasonableness of any rate': memo; cover letter dated 26 January 1912. Telephone Pioneers Museum, San Francisco, CA.

One milestone in this reconstruction can be clearly identified: a report submitted to Congress by Eugene V. Rostow on behalf of AT&T in 1975.[79] Rostow, the former chair of President Johnson's Task Force on Communications Policy, had been retained by AT&T to support its legislative efforts to protect itself from new competition. It was AT&T, via Rostow, which first aired the specious claim that a monopoly system devoted to universal service was part of the mandate of the 1934 Communications Act.

We have only to look at the context of the Rostow report to understand the function of the new universal service ideology. The fateful antitrust suit had just been filed by the Department of Justice in 1974. MCI had just invaded switched long distance with its Execunet service in 1975, a development which not only undermined the cross-subsidies of the Ozark plan but threatened to subvert the whole end-to-end philosophy underlying separations and settlements practices. The company was in the thick of an all-out attempt to persuade Congress to pass a law to preserve the classical monopoly arrangements – the so-called 'Bell bill' of 1976.[80]

During the battle over the Bell bill and the ensuing years of antitrust proceedings, 'universal service' became the rallying cry of AT&T and the other defenders of regulated monopoly. Just as Vail had used the term to fend off access competition from 1907 to 1920, AT&T attempted to use the same term, albeit with a different meaning and in a very different context, to renew the nation's commitment to the regulated monopoly structure Vail had helped to establish. The modern reconstruction of universal service, however, was not an accurate description of a historical policy, but a retroactive rationalization for the institution of regulated monopoly.[81]

As a revised ideology of 'universal service' was pressed into the service of telephone monopolies in the 1970s and 1980s, its meaning changed in ways that obscured what it had meant when it was coined in 1907. A confusion between its contemporary and historical usage has made it difficult for modern scholars and policy makers to appreciate the significance of the earlier universal service debate. And the universal service claims of regulated monopoly have unfairly eclipsed the earlier contribution of competition to the development of a ubiquitous telephone infrastructure.

Conclusions

The analysis presented above leads to several revisionist conclusions about the history of the telephone and the origins of universal service.

The meaning of universal service

The meaning of the term universal service has changed significantly since it was coined, and projecting contemporary definitions backwards is misleading. The modern definition stresses rate subsidies administered by a regulated monopoly. Clearly, this had nothing to do with the original concept, and in fact this subsidy-oriented definition did not emerge until quite recently. From 1907 until about 1975 universal service meant the interconnection of all localities and telephone users into a single system. From a regulatory standpoint it meant that the cost of service was defined from the standpoint of the system as a whole, not as a collection of discrete, stand-alone routes or network components.

[79]Eugene V. Rostow, 'The case for congressional action to safeguard the telephone network as a universal and optimized system', paper based on the memorandum prepared for AT&T for use in the November 1975 hearings before the Subcommittee on Communications of the US House of Representatives Committee on Interstate and Foreign Commerce.

[80]The Consumer Communications Act of 1976 quickly acquired the 'Bell bill' label because of AT&T's sponsorship and all-out lobbying effort on its behalf. For an account of its fate, see Peter Temin with Louis Galambos, *The Fall of the Bell System*, Cambridge University Press, Cambridge, UK, 1987.

[81]In the case of European PTTs the retroactive nature of universal service claims is even clearer. European monopolies adopted the same averaging and cross-subsidy practices as the US telephone companies without attaining anything near the penetration levels of the USA, but nevertheless made 'universality' one of their defences against the onslaught of new competition in the 1980s. As Garnham, *op cit*, Ref 1, has shown, officially proclaimed universal service goals in Europe often coexist with low penetration and large regional disparities in access to the telephone.

True, the diffusion of telephone service was hailed as a desirable thing. Trade journals and the popular press at the turn of the century marvelled at its rapid penetration of farm areas and residences, and interpreted this as a sign of the inexorable progress of the industrial age.[82] Where the 1880s and early 1890s saw the telephone as a specialized commercial device, no one in the 1900s or 1910s would have disagreed with the assertion that eventually there would be a telephone in every home. But this progress was seen as something that would occur naturally as industrialism increased wealth, lowered prices and improved technology. Universalism in this sense posed no special policy issue, required no government action. The real policy issue was whether the telephone would develop under the guise of separate, competitive systems or as a unified monopoly.

Competition and universality

Far from being a policy imposed on the market by enlightened regulators, universality was avidly pursued by the telephone companies at the turn of the century because of the pressures of competition. Ironically, it was the refusal of Bell and the independents to interconnect with each other, a phenomenon which is ignored or condemned in the historical literature, which propelled both systems into the race to achieve universal geographic coverage and universal penetration. The telephone network grew to embrace most of the country because of business responses to a competitive environment that placed a premium on a telephone system's scope. Competition was abandoned, moreover, not because of predatory tactics by AT&T nor even because of the economic exhaustion of the independent movement, but in order to eliminate the fragmentation caused by access competition.

Bell System claims

The Bell System's claim to have pursued universal service from its inception can now be evaluated more accurately. If by 'One System, One Policy' Vail meant that Bell intended to establish a centrally coordinated monopoly, and by 'Universal Service' he meant nothing more than that Bell aimed at a physically integrated system whose subscribers could all talk to each other, then AT&T had indeed always pursued universal service. Nevertheless Vail's claim that the Bell System was founded on the principle of universal service in the modern sense – meaning service everywhere, to everyone – is a half truth at best. It came from looking at Bell System organization retrospectively, in the light of 20 years of independent competition. By that time the scope and usage of the telephone had been transformed so profoundly that the concept of a universal system had taken on a meaning far different from what Vail had originally meant. The Bell System simply failed to foresee what an extensive development of the telephone business would be required. Never in their wildest dreams did the pre-competition Bell managers think that telephone service could be demanded by, and profitably extended to, as many people as turned out to be possible. As it was, even after its massive geographic expansion of the early 1900s, Bell still relied primarily on independent companies to obtain access to small towns and rural areas.

Regulated monopoly and universal service

The role of regulation in the achievement of universal service also seems

[82]Commenting on the growth of residential subscribership in New York City, the *Electrical Review* (Vol 31, No 15, 13 October 1897, p 180) wrote, 'it will not be long before no moderately well appointed residence will be considered completely equipped if it is not connected to the telephone system'. For similar expressions of confidence in the inevitability of the spread of the telephone, see 'The farmer and the telephone', *Electrical Review*, Vol 31, No 11, 15 September 1897, p 126; and 'Making [social] calls by telephone', *Electrical Review*, Vol 30, No 13, 31 March 1897, p 146.

less critical in the light of this historical evidence. Although from 1912 to 1925 both regulators and telephone companies came to favour unified service over competitive fragmentation, regulators and rate subsidies had little to do with the initial extension of service to rural areas. Indeed, rural penetration declined in the late 1920s and 1930s after access competition ceased and the Great Depression struck. Rural Electrification Administration (REA) loans to rural telephone companies after 1949 were a weak substitute for the strong economic incentive to connect the country characteristic of the competitive era. Furthermore, had the competitive period not led to the rapid occupation of rural areas by thousands of independent companies there would have been little for the REA to lend money to. The gradual rise in penetration to 92% after the seccond world war probably had more to do with the doubling and tripling of household income during that period than with separations and settlement practices, for the doubling of local service rates since 1982 has had a negligible effect on overall penetration levels.[83] In retrospect, regulation looks more like an inertial, conservative force than a constructive and creative one.

There are elements of truth in both AT&T's and the regulators' constructions of history. AT&T's vertical integration and commitment to long-distance development did create the backbone of a nationally interconnected network. Basic subscription rates were kept low and the economic health of many small rural systems was boosted by REA loans and the settlement policies of regulated monopoly. But these partial truths have been advanced at the expense of a more fundamental fact about the telephone's history in the USA. The most important historical factor contributing to extensive coverage and high penetration in the USA was 20 years of intense rivalry between telephone systems that were not connected to each other.

Contemporary policy implications

The historical facts about access competition have important policy implications for developing countries. If the standard historical assumption about regulated monopoly's role in the creation of universal service is true, then developing countries should stay with regulated monopolies to develop their infrastructure before experimenting with competition. This is, in fact, the position advocated by the World Bank's telecommunications specialists.[84] If, on the other hand, access competition played a critical role in the developmental stages of the US infrastructure and this experience accounts for the tremendous US lead in the extension of telecommunication service, then a very different policy conclusion can be drawn. Conditions in developing countries, which have low penetration and a stagnant monopoly, often closely correspond to the conditions in the USA prior to independent competition. A policy of open entry and systems competition could have similar effects, although of course there are many differences in conditions.

The promotion of universal service via systems competition also has policy relevance for advanced countries. Most liberalized policies promote competition via equal access interconnection arrangements. Liberalized interconnection encouraged competitors to rely on the 'bottleneck' local facilities of the established network. While these policies make competitive entry easier, they also hinder (or may even prevent forever) the development of alternative, truly universal local infrastructures.

[83]Federal Communications Commission, *Common Carrier Statistics*, 1991.

[84]The July–August 1991 issue of *Transition*, the newsletter of the World Bank's Socialist Economies Unit, contains a statement from telecommunication specialists Nikola Holcer and Tim Nulty claiming that 'the superlative US network' was built 'during the period from 1946 to 1971', and that 'it is the period that is most relevant for the Central European economies'.