Death through Success: The Rise and Fall of Local Service Competition at the Turn of the Century

Markus Möbius *
Harvard University

July 2011

Abstract

This paper develops a model of delayed network effects to explore the curious dynamics of competition in the local telephone market between AT&T and the 'Independents' at the turn of the century. In the early years of telephone diffusion, local service competition between these two non-interconnected networks became widespread, but declined rapidly when diffusion rates started to slow down after 1907. The analysis is based on the observation that urban markets subdivide into social 'islands' along geographical and socio-economic dimensions: users are more likely to communicate with subscribers 'inside' their island than with those 'outside' it. A simple dynamic model demonstrates how minority networks can thrive and preserve their market share at a low state of development when islands form essentially independent niche markets. As the industry matures, these niches 'grow' together and standardization occurs. The implications of the model are confirmed using a small panel data set of US cities.

JEL Classification: D43, D46, K23, L12, L96
Keywords: local service competition, interconnection, standardization, delayed network effects, cost of incompatibility

*I would like to thank Daron Acemoglu, Abhijit Banerjee, Peter Temin and seminar participants at MIT and Harvard for helpful comments. I am especially grateful to Glenn Ellison for very detailed comments on the first draft of this paper. All errors are mine. Financial support from the MacArthur foundation is acknowledged.

1 Introduction

The dominant position of AT&T in the American telecommunication industry until its break-up in 1982 masks the fact that the company had once faced a formidable challenge at the turn of the century. After AT&T's loss of its patent protection in 1893, independent telephone companies first entered uncontested rural markets. Soon afterwards they challenged AT&T in the cities. As the Independents interconnected amongst each other but not with AT&T, both systems competed mainly on the local exchange level forcing subscribers to duplicate providers if they wanted universal service. After years of double digit growth, the independent movement controlled almost half of the US market. However, local service competition declined rapidly in subsequent years as diffusion rates declined. By the mid-1920s, AT&T had regained control over the entire US telephone network either directly or indirectly through dependent regional sublicencees.

The existing literature offers two types of explanations for this intriguing competitive dynamics. They are based either on global network effects or predatory strategies pursued by AT&T. Neither theory can fully account for both the rise and the fall of the Independents. In this paper, I propose a new dynamic model of early local telephone competition using a detailed analysis of communication patterns within an urban social network. The model gives rise to delayed network effects that favor the emergence of a single standard only at high telephone penetration rates.

I argue that urban markets subdivide into social 'islands' along geographical and socio-economic dimensions. Communication between agents in the urban social network is characterized by local interaction because users are more likely to communicate with subscribers 'inside' their island than with those 'outside' it. At low telephone penetration rates, these islands are essentially isolated niche markets that can be colonized independently by competing telephone companies. This allows minority networks to thrive and preserve their market share. However, as the telephone usage diffuses, communication between the islands increases and these niche markets 'grow' together giving rise to two types of delayed network effects. First, the lack of interconnection imposes an increasing efficiency loss (especially on business subscribers forcing them to subscribe to both competing networks in order to talk to outside customers). As a result, public support for 'wasteful' competition diminishes and gives way to calls for consolidation and compulsory interconnection of rival networks. Second, duplication by businesses tends to be asymmetric in the sense that business subscribers who originally chose the minority system have a much larger incentive to duplicate than those who chose the majority system. This duplication effect threatens the long-term survival of a minority system because duplication weakens its hold on islands. When the state of development is sufficiently advanced, standardization can arise endogenously. In addition, I find that both effects complement each other: the political economy effect is the strongest when both systems are of similar size, while the duplication effect dominates in the other situation.

The paper revives the role of mass market forces in the selection of telephone standards emphasized by early historians of the telephone industry such as Anderson (1907), Stehman (1925), and Brooks (1975). They adopted AT&T's natural monopoly view of telephone service since the telephone business is about connecting all subscribers with each other. Unconnected networks were therefore considered a major nuisance warranting eventual consolidation. It is worth noting that this global network view is not supported by the empirical evidence as the number of calls per residential telephone did not increase linearly with telephone penetration either in cities or nation-wide but instead barely changed at all. Moreover, the simple network effect story can only explain the decline of local service competition but fails to account for its initial success. The more sophisticated local network model proposed here overcomes this problem since network markets have a tendency to converge towards a common standard only at high penetration rates.

Recent research has emphasized the role of predatory strategies in the decline of local service competition. Gabel (1969), Lipartito (1989) and Weiman and Levin (1994) make the case that AT&T regained its monopoly though a combination of preemptive investment into long-distance telephony, pricing below average cost, strategic acquisitions, sublicensing and regulatory capture. The predation view and the network story developed in this paper are, in my view, complementary. AT&T pioneered its anti-competitive strategies in the South and only extended it to the rest of the country after 1907. Under threat of anti-trust litigation it had to stop its campaign in 1913. A substantial number of independent regional systems survived the onslaught, especially in the highly developed North Central states, and every third US city was still contested by rival, incompatible networks. Nevertheless, telephone subscribers and regulators pressed for further consolidation in the industry and local service competition by and large disappeared until the end of World War I. This suggests that the formation of regional monopolies was unavoidable as the industry matured. AT&T's actions only accelerated this process and ensured control of most of these regional entities in the post-competitive era.

The remainder of the paper is organized as follows. Section 2 discusses the initial success of independent entrants into the market and the growth of local service competition up to 1907. Section 3 draws on various sources to illustrate the presence of 'social islands' within the social network of a city. Taking these communication patterns into account the formal model is developed in section 4. Section 5 analyzes the dynamics of network competition and derives the conditions under which asymmetric duplication can weaken the hold of the minority system on its islands sufficiently to be driven out of the market. The efficiency losses due to competition are evaluated in section 6. It is shown that the resulting political economy effect and the asymmetric duplication effect complement each other. Section 7 discusses what effects compulsory interconnection would have had on the dynamics of competition. Section 8 tests the main empirical implications of the model and analyzes to what extent it can explain the decline of local service competition in the maturing industry. For this purpose, panel data on telephone ownership for a small number of US cities at the beginning of the 20th century has been collected.

2 Competition in Telephony, 1894-1913

The American Bell company (renamed into AT&T after 1899) had a monopoly in the US telephony market until 1893/94 when its principal patents expired. Entry occurred almost immediately with 18 independent commercial systems being established in 1893 and 80 in 1894. By 1900 independent companies were founded at a rate of 500 a year and by the end of 1902 a total number of 3,113 Independents crowded the industry. Competition led to a dramatic decrease in telephone rates of often more than 50 percent even in towns without competition (Weiman and Levin 1994). The lower cost of telephone service accelerated the adoption of the technology dramatically. Between 1893 and 1902 the subscriber base increased tenfold (see table 1) and by 1907 the North Central states achieved penetration rates of 13 percent. The Independents managed to gain market share at the expense of AT&T and controlled half of all the telephones in the US in 1907.

The rival systems followed different expansion strategies. The American Bell company pursued a top-down approach of systematic expansion with the goal of creating a fully interconnected nation-wide network. Local exchanges were built according to uniform high standards in order to incorporate them easily into the nascent long-distance network (Langdale 1978). This left rural areas and small towns often completely undeveloped until competition forced AT&T to extend the network into the smaller communities. After 1894 the company invested even more aggressively into its long-distance network in order to keep up with the rapid growth of local exchanges. Under the slogan "One System, One Policy" AT&T sustained long-distance lines even if they were unprofitable. The Independents on the other hand built their systems from the bottom up. The first wave of entry occurred in the rural areas and small towns with no telephone service.[1] Local service competition became widespread with a second wave of entry around 1900 in the bigger towns and cities when AT&T was challenged on its established

[1] The populous North Central states with their large rural populations became the hotbed of competition. In 1902 of the 3,113 commercial independent exchanges 1,694 were located in Illinois, Indiana, Ohio, Kansas, Michigan, Minnesota, Missouri, Iowa and Wisconsin (Bureau of the Census 1902, table 10).

Table 1: Diffusion of telephony in the US, 1894-1912

	Stations of all systems	Bell system Share	Bell system annual growth		Cities with competition
1894	270,381	100	1885-94:	6.26	2
1897	NA	NA	NA		23
1902	2,371,044	55.6	1895-1902:	22.06	55
1904	NA	NA	NA		60
1907	6,118,578	51.2	1903-1907:	18.11	57
1912	8,729,592	58.3	1908-1912:	9.81	37

Sources: Telephone Census (1912), table 2; Mueller (1997), table 6-2; Warren Report (1938), table 33. Only cities with a population of at least 5,000 are considered. All values except those in the first column are percentage points.

territory. Whereas only 2 percent of all cities had local competition in 1894, almost two third had at least two local systems ten years later (see table 1).

Due to the lack of a coordinating strong center the Independents could not match AT&T's investment in long-distance lines. Independents had to establish exchanges in a critical number of town and cities within a region before they could start to invest heavily in toll lines. Most users, however, had no demand for ultra-long distance calls. Between 1902 and 1912 at least 97 percent of all calls were local (Bureau of the Census 1912, table 19). About 90 percent of long-distance calls originated and ended within a 50 miles radius.[2] Interstate calls were still too expensive for most subscribers and the telegraph provided a cheaper alternative for sending ultra-long distance messages.[3] Independent exchanges could therefore survive by forming regional systems with an interconnecting network of toll lines that would satisfy most of their subscribers' demand for long-haul connections. By 1906, this process became evident in those parts of the country where Independents were most strongly represented, such as the North Central states. Mergers gave rise to independent regional operators which owned exchanges in typically ten to thirty key cities and had exclusive connecting contracts with adjoint Independents. In

[2]In Chicago 90 percent of all messages in and out of Chicago were transmitted within a radius of 100 miles (Committee on Gas, Oil and Electric Light 1907, p. 120). In Indiana 89 percent of all messages originating or terminating at a population center stayed within a radius of 35 miles (*The American Telephone Journal* July 5, 1902: p. 15).

[3]In 1902 a three minute station to station call from New York to Philadelphia cost $0.55 and to Chicago $5.45 (about 5 percent of the *yearly* flat business rate in Chicago) according to the Bureau of the Census (1975, p. 784). The corresponding costs of a 10 word telegraph message on the other hand were just $0.25 and $0.50 in 1908 (p. 790). Long-distance telephony remained expensive until the invention of the electronic repeater in 1915.

Indiana, Ohio and Michigan the independent movement even managed to establish dedicated long-distance operators.[4]

The lack of interconnection during the early competitive era might seem strange in the light of contemporary regulators such as the Federal Communications Commission (FCC) and state utility commissions who routinely mandate open interconnection and determine access fees between competing networks. However, early regulation was fragmented at the city level as telephone companies operated under a municipal franchise which allowed them to use public roads for line construction. City councils primarily regulated entry into the market and local telephone rates, but a legal framework which would have allowed them to impose interconnection was not in place yet. Voluntary interconnection did not occur mainly because AT&T opposed it. The company did not want its competitors to freeride on its heavy investments in long-distance lines, and feared that interconnection would allow local service competition to continue indefinitely. On the other hand, many Independents initially feared that interconnection would slow down the growth of their own networks. Exclusive control of rural exchanges and many small towns gave them a powerful negotiating tool to secure franchises in the bigger cities. Moreover, interconnection would have halted the development of their own nascent long-distance network.[5]

Between 1900 and 1903, the Southern Bell Telephone company pioneered a new strategy as the independent movement continued to prosper in its territory. In order to prevent the formation of regional independent systems, AT&T offered interconnection to independent companies in *non-competed* territories. The so-called sublicencees enjoyed a local monopoly but were not allowed to interconnect with any independent exchange. Sublicensing fragmented the independent movement and made it more difficult for the Independents to build viable regional systems with a satisfactory toll network. The independent movement was thus successfully halted in the South at an early stage of development. In 1907, AT&T extended sublicensing to non-competing Independents in the rest of the country with control of between 55 percent to 60 percent of all independent telephones. By 1913, 87 percent of them were interconnected with AT&T (Mueller 1997, chapter 9). Independent operators in cities with local service competition were isolated and many of them were taken over by their Bell competitor between 1907 and 1913.

Unlike in the South, however, AT&T could not eliminate local service competitions in the rest of the country, especially the North Central states. In 1913, 37 percent of cities with populations over 5,000 still had dual exchanges down from 59 percent at the beginning of 1907, and 34 percent of all independent telephones (75

[4]The Independents managed to handle 28 percent of all US long-distance traffic in 1907 and presumably a much higher share in the North Central states.

[5]The New Long Distance Company of Indiana for example prohibited its member exchanges from connecting to different long-distance carriers for 99 years (MacMeal 1934, p. 174).

percent of them in competed territories) were still unconnected to the Bell system (Mueller 1997, p. 110-12). Many regional systems rooted in larger cities had managed to survive thanks to their extensive toll networks. AT&T's heavy-handed attempt to unify the telephone system was meanwhile coming under increasing attack. In the face of anti-trust legislation, the company was forced to back away from its pursuit of universal service and announced the "Kingsbury Commitment" in 1913. The most significant concession was a moratorium on further acquisitions of independent exchanges in competed cities.

The Kingsbury Commitment did not save local service competition even though the number of non-connected independent telephones stabilized between 1913 and 1917. Consolidation proceeded in the background by swapping territories instead of acquisition. This time the pressure came from the demand side: city councils, federal and state authorities had to agree to those transactions in order to waive the Kingsbury commitment, and in many cases voters expressed their desire to unify the service through a referendum. Therefore, the industry continued to converge towards standardization as regional monopolies replaced local service competition.

3 The Structure of the Social Network

In order to model local service competition, one needs to understand communication patterns of the urban *social network* which I assume to be exogenous.[6] This section describes how the two groups of subscribers, businesses and residential customers, communicated between and amongst each other. Based on these observations, the formal model is developed in the next section.

From the sparse historical statistics on residential communication patterns one can conclude that about 25 percent of all telephone calls were made to various businesses and between 45 and 75 percent of conversations were conducted with friends and acquaintances. In 1909, a Bell manager listened in on a sample of conversations at a residential Seattle exchange. He found that 20 percent of calls were orders to stores and other businesses, 20 percent were made from homes to family-owned businesses, 15 percent were social invitations, and 30 percent were "pure idle gossip" (Fischer 1992, p. 79).[7] These numbers are consistent with the

[6] Various studies such as Kern (1983) and Meyerowitz (1985) claim that the telephone itself changed the social network by weakening local ties in favor of extra-local ones and by eradicating the concepts of space and distance. In the most comprehensive study on this topic, however, Fischer (1992) finds that the telephone did not open any new social contacts. On the contrary, telephony was an anti-modern technology in the sense that it reinforced existing social relationships.

[7] Telephone companies initially fought their residential customers over social calls, which they saw as an unnecessary burden on their networks. Only after World War I did AT&T discover sociability as a selling point for their service.

results of AT&T research which showed that nowadays about 75 percent of all messages are social calls to friends and family.[8]

The bulk of social calls were apparently conducted with a small circle of friends. AT&T research has shown that about half of all calls from a given residence are made to just five numbers. The importance of a small number of social neighbors can help to explain the curious fact, that during the competitive era the number of calls per telephone remained virtually unchanged. This phenomenon can be confirmed both for state statistics as well as city data.[9] If subscribers were to really care about the total number of people they can communicate with, one would expect a steady increase in the number of calls per telephone. The telephone census of 1902 (p. 30) offers the following explanation:

> Experience shows that this does not happen, because of what is termed the "acquaintance factor". In every community each individual is acquainted with and transacts business among a certain limited group; and while such circles of acquaintances overlap and the business increases more rapidly than is indicated by a simple arithmetical ratio, it does not increase quite as fast as the square of the population.

These empirical observations suggest that the structure of the social network of residents resembles a locally finite graph in which each agent has strong links only to a small number of neighbors independent of the size of the total graph. Finite dimensional lattices and random graphs are the two extreme geometries in this class of graphs. This paper uses a lattice structure to model the residential social network. Lattices are better suited to capture the distribution of strong links within the social network of residents because unlike random graphs they exhibit a high degree of 'clustering'. Clustering implies that my neighbor's neighbor is with high probability also my own neighbor. This property, first observed by Granovetter (1973) and confirmed in subsequent studies (see, for example, Granovetter (1974) and Watts and Strogatz (1998)) is a robust feature of real world social networks.

Residential subscribers are differentiated along several dimensions such as social status, place of residence and work, and ethnicity. MIT students are more likely to be acquainted with fellow MIT students rather than Boston University students. Faculty members are more likely to call fellow faculty rather than undergraduates. Subscribers are therefore located in different regions or 'social islands' of the social

[8]When an accident in 1975 knocked out thousands of telephones in New York City, 70 percent of all users declared in a subsequent survey, that they had mostly missed social calls. In a 1985 poll, Californians estimated that about 75 percent of all their calls are for social purposes (Fischer 1992, p. 226).

[9]Between 1902 and 1912, the number of calls per telephone in the US and the North Central states was at about seven calls per day (Bureau of the Census 1912). In 1907, the number of calls per measured rate telephone in the city of Chicago was also seven (Committee on Gas, Oil and Electric Light 1907).

lattice such as the island of MIT undergraduates for example. The friends and acquaintances of a subscriber in the social network are more likely to be found within the same social island rather than outside it. Given the importance of social calls, one would expect residential subscribers within the same island to coordinate their network choice.[10] Brooks (1975, p. 110) reports:

> In Minneapolis, for example - according to the recollection of a survivor of the competitive era there - the Bell exchange, being the longer - established, was the exchange of the socially elite, while the competing Tri-State Telephone Company was for just about everybody else.

As mentioned earlier, an estimated 25 percent of all residential messages were made to stores and other establishments.[11] This highlights the importance of business subscribers in the social network. During the early years of telephony business and industry were in fact the primary users of the new technology. The rapid diffusion of telephony during the competitive era, however, shifted the balance soon towards residential users. By 1905, the share of residential telephones[12] had already risen to about 50 percent[13] in the cities and by 1920 the proportion was close to 70 percent (Bureau of the Census 1975, table R1-12).

How did businesses use the telephone? Valuable insights can be gained by inspecting a break-down of telephone subscribers in Louisville, KY in 1910 given in table 2. There the Cumberland Telephone Company (Bell) was competing against the independent Home Telephone Company. Business users had to communicate with their residential customers and their suppliers. These upstream businesses were usually large-scale establishments which had a duplication rate of close to 100 percent. They were the first ones to order a second telephone, but made up just 1.5 percent of all telephone subscribers in the city. Due to the high degree of duplication amongst suppliers, medium-scale and small-scale businesses presumably

[10]Even at the beginning of the century the telephone was already used by a substantial share of people of each social class. Fischer (1992) finds that in the three Californian cities of Palo Alto, Antioch and San Rafael about 50 percent of managers and professionals had telephones in 1910, but also 20 percent of all commercial white collar households and 10 percent of blue collar workers.

[11]Fischer (1992, p. 361) suggests an estimate for the share of functional calls from residential telephones of between 20 percent and 60 percent.

[12]The share of residential telephones understates the share of residential subscribers. Large businesses often had private branch exchanges and extension telephones. In Chicago, for example, the share of business telephones in 1909 was 52 percent. When I correct for private branch telephones the share drops to 38 percent (*Electrical World* 1909: p. 1072,1504).

[13]In Columbus the Citizens Company reported a 55 percent share of business telephones in 1900 and only 39 percent in 1907 (Johnston 1908, p. 8). In Indianapolis 39 percent of subscribers were business users in 1904 but only 31 percent ten years later (*The American Telephone Journal* 1904: p. 309; Geijsbeek and Lawrence 1914, schedule 1). In Louisville the business subscriber share is estimated to be at most 45 percent in 1905 by dividing the number of business subscribers in 1910 (see table 2) by the total number of subscribers in 1905 (see table 3).

Table 2: Telephone distribution and duplication rates in Louisville, 1910

Size of Business	Both Phones	Home Only	Bell Only	Duplication Rate	Subscriber Rate
Large-scale					
Telegraph	4	0	0	100	100
Gas and Electric	4	0	0	100	100
Fast Freight	11	1	0	92	100
Railroads	21	2	2	87	100
Banks	25	2	2	86	100
Hotels	21	6	0	78	100
Sub-total	86	11	4		
Medium-scale					
Druggists	83	69	3	53	100
Coal dealers	46	42	9	47	100
Insurance	65	46	36	44	NA
Dentists	35	44	3	42	63
Liquor dealers	43	56	18	37	NA
Plumbers	25	45	1	35	74
Attorneys	85	109	90	30	78
Butchers	19	47	7	26	NA
Dry goods	15	36	6	26	21
Groceries	182	466	62	25	NA
Sub-total	598	970	235		
Small-scale					
Billiard/ Bowling	2	10	0	16	NA
Carpenters	11	55	9	14	50
Barber shops	1	6	1	12	NA
Bakers	9	61	9	11	39
Saloons	64	487	19	11	87
Tailors	8	60	9	10	NA
Churches	3	12	14	10	NA
Residences	900	5,449	3, 971	9	20
Sub-total	998	6,140	4,032		

Source: Mueller (1997, Table 7-1). 'Home' refers to the independent Home Telephone Company while 'Bell' denotes the Cumberland Telephone Company, a Bell licensee. All values are percentage points.

made their network choice dependent on the choices of their residential customers rather than their upstream suppliers.

Several features of the Louisville survey support this hypothesis. The presence of single-phone grocers connecting to either the independent or the Bell exchange indicates, that some grocers mainly served subscribers to one of the exchanges. While most single-phone businesses favored the independent company by a ratio of at least 5:1, the Bell company was almost even amongst insurance agents and attorneys. The longer established Bell exchange was likely to serve more upper-class citizens who were amongst the first telephone subscribers in the city before the onset of competition. Business subscribers thus seemed to serve customers in different social islands such that their network choice would depend to a greater extent on the choices of residents within those islands rather than outside them.[14] However, the high degree of duplication amongst medium-scale businesses suggests that some business users either served customers of several social islands with different competing networks, or that they cared sufficiently about customers outside their main customer base and ordered a second phone.

To summarize, I have presented evidence for the local character of the social network that underlies the demand for telephone conversations. Independent entrants had to exploit this structure in order to gain an initial foothold in a town. In the light of the stylized facts developed above, one would expect that telephone entrants should have first targeted a group of close neighbors within a social island and the associated businesses serving that island. The following account from a publication of the New England Telephone and Telegraph Company (1908, p. 11) describes the process of soliciting the initial group of subscribers for a new exchange exactly according to that mechanism.[15]

> The origin of the Independent movement is somewhat typical ... The subscribers to the Bell system are not solicited to subscribe to the new system, at least not at first. A man who has no telephone is asked if he wouldn't like one ... [He] desires to know how many subscribers he can communicate with ... he need not bind himself except conditionally upon a certain number of subscribers being secured. Are there any of his friends with whom he or his family would especially desire to communicate? If names are given, these people are seen and assured

[14] Fischer (1992, p. 180) reports further evidence in his study of three Californian towns. Florists catering to the middle class had phones whereas corner stores serving working class neighborhoods did not have one due to the lower subscriber rate amongst blue-collar workers.

[15] A surprisingly small number of subscribers was necessary to establish the first exchange. The manager of the Eureka, California, exchange for example reported that "my hardest work was getting my first twelve subscribers", which was regarded as the minimum number for an exchange (Fischer 1992, p. 64). It is noteworthy that telephone companies used solicitors mainly during the start-up phase. As soon as demand outstripped supply solicitors were laid off.

> that Mr. A, the first one visited, is going to put in a telephone. It is not difficult in this way to secure a number of signatures to conditional contracts ... The next step is with this list of names to go to some local dealer ... The canny canvasser has ascertained from those whom he has previously visited ... where they do their trading, who is their butcher, baker or grocery man. The dealer sees the possibility of increased orders ... Thus the signatures are secured.

Entry into the telephone market was considered to be easy at the onset of competition when telephone concentration was low and many social islands were yet undeveloped. Theodore Gary, president of the Interstate Independent Association, for example writes in 1907:

> There is a place for two telephone companies in about every community of 10,000 population and up, occupied by a Bell company, because it generally fails to develop a territory. Competition is desirable unless there is a development by one company of one telephone to five or six persons in the community. With full development there is very little room or demand for competition.[16]

4 The Model

I consider a city that consists of an infinite number of social islands.[17] Each island is made up of n residents and a continuum of business users such as grocery, butcher, physician.[18] Every resident has a number of 'friends' within her social island with whom she communicates. Specifically, I assume that residents are located along a circle such that each of them is friends with her two direct neighbors.[19] Every resident is a 'core customer' for each type of business serving that island. Figure 1 shows the resulting network linking residents and businesses of a social island.

4.1 Communication Patterns

Time is discrete and each resident exchanges a total number N of messages per time period through the social network. Communication between the various agents

[16] Theodore Gary in *Telephony* December, 1907: p. 338, quoted by Bornholz and Evans (1983, p. 19)

[17] Assuming an infinity of islands is analytically convenient as it makes the model non-stochastic at the aggregate level.

[18] The model is easiest to present by assuming a continuum of businesses. The results are the same for a discrete version of the model with m types of businesses which is available upon request.

[19] The particular graph mapped out by these bonds can be any graph which exhibits some degree of clustering. My results would still go through qualitatively.

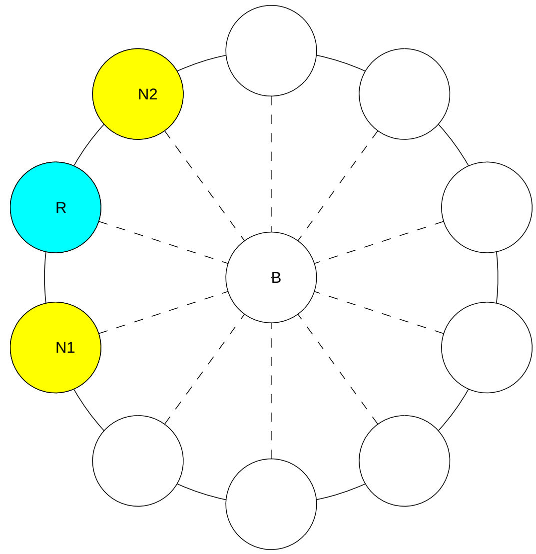

Figure 1: Social island of size 10 - resident R is 'friends' with residents N1 and N2. Business users B serve the entire island.

is assumed to be symmetric, i.e. any two sets of agents send and receive an equal number of messages between each other.[20] I assume for simplicity that the demand for communication is inelastic with respect to the cost of communication: technological innovations such as the telephone will therefore reduce the cost of communication but will not increase demand for it. A share α of those N messages are exchanged with friends and a share $\beta = 1 - \alpha$ with businesses. Businesses differ by size and residents exchange a share $\tilde{\beta}$ with business type $\tilde{\beta} \in \left[\underline{\beta}, \overline{\beta}\right]$. For example, druggists are likely to receive more messages than attorneys such that $\beta_{Druggist} > \beta_{Attorney}$. These businesses follow a frequency distribution $g\left(\tilde{\beta}\right)$. It will be convenient to denote the volume of calls exchanged with businesses of at most size $\tilde{\beta}$ with

$$G\left(\tilde{\beta}\right) = \int_{\underline{\beta}}^{\tilde{\beta}} yg\left(y\right)dy. \tag{1}$$

The total share of calls that a resident exchanges with businesses is

$$\beta = G\left(\overline{\beta}\right) = 1 - \alpha. \tag{2}$$

[20]Businesses tend to receive more calls from customers than they send. However, the results of the model only depend on businesses to prefer networks which offer more connections to potential residential customers. It does not matter if businesses derive utility from their subscription because they can send messages to customers more easily, or because they expect to become more attractive to customers if they can be more easily contacted.

A resident who contacts a type $\tilde{\beta}$ business will deal with probability $1-\gamma$ with a local business and with probability γ with any of the outside businesses. Each type $\tilde{\beta}$ business therefore exchanges $\tilde{\beta}nN$ messages with customers and with probability $1-\gamma$ such a message is exchanged with a core customer. I assume $0 < \gamma < \frac{1}{2}$, e.g. residents do most of their shopping at their local stores but occasionally frequent outside businesses. Blue-collar workers for example might visit upper-class stores if these are geographically conveniently located. A residents from one suburb might have a favorite physician who lives in a different part of the city. Clearly γ will depend on physical characteristics of the city. A large city with a higher degree of suburbanization is characterized by a lower γ than a small town where different socio-economic classes live closely together.

The following technical assumption on the distribution of businesses is utilized here.

Assumption 1 *The ratio* $R\left(\tilde{\beta}, a\right) = \frac{G\left(a\tilde{\beta}\right)}{G\left(\tilde{\beta}\right)}$ *is (weakly) decreasing in* $\tilde{\beta}$ *for all* $a \geq 1$ *and* $\tilde{\beta} \in \left[\underline{\beta}, \overline{\beta}\right]$.

This assumption simplifies the dynamics of the model considerably but does not change our qualitative results.[21]

4.2 Preferences

Communication is costly for residents and businesses. A resident has to spend a total amount of time $T < 1$ for sending all her messages: she has to write a letter to a friend or walk to a store for making an order. Her total time endowment is 1 and she has an income of y in each time period, which she can spend on consumption C or telephone service. Two competing companies A and B offer access to their networks at a rate K per time period.[22] I assume that the subscription rates are fixed for all time periods which allows me to focus on the demand-side when I analyze the growth of both systems: in particular this precludes strategic interaction between the two telephone carriers. While the assumption might seem extreme, rate schedules were in fact changed infrequently because telephone companies generally operated under a franchise from the city council. Such a franchise

[21] Note, that *any* distribution satisfies assumption 1 for $a = 1$ and $a > \frac{\overline{\beta}}{\underline{\beta}}$ because the ratio function $R\left(\tilde{\beta}, a\right)$ takes the values 1 and $\frac{1}{G\left(\tilde{\beta}\right)}$, respectively. Furthermore, the ratio function is always greater than 1 at $\tilde{\beta} = \underline{\beta}$ and equal to 1 at $\tilde{\beta} = \overline{\beta}$, and therefore has to decrease 'on average' over the interval $\left[\underline{\beta}, \overline{\beta}\right]$. Assumption 1 requires the function to decrease everywhere on this interval, and is satisfied by standard distributions such as the uniform and power distribution.

[22] For simplicity, I assume that the marginal cost of sending a message over the phone is 0. Most cities (especially the smaller ones) had in fact unlimited service, but it is straightforward to incorporate a message cost c in my model on top of the base rate K.

was only granted if the entrant offered competitive rates: the subscription rate K was usually calculated as the cost of laying and maintaining a telephone wire between a subscriber's place of residence and the central office of the exchange plus some reasonable rate of return on the investment.[23] Increases in the rates required a time-consuming process of obtaining the permission from the local authorities.

If a resident sends a share z of her messages over the phone she only spends $(1-z)\,T$ of her time on the remaining messages. Her per period utility function $U\,(C,t)$ in consumption and time has the usual properties and the following condition holds:

$$U\,(y-2K,1) < U\,(y,1-T) < U\,(y-K,1) \tag{3}$$

The left inequality suggests that a resident would never duplicate as she could always do better by having no telephone at all.[24] The right inequality ensures that a resident will certainly subscribe to a network if everyone else has a telephone. A threshold m^r can therefore be defined as

$$U\,(y,1-T) = U\,(y-K,1-T+m^rT)\,.$$

A resident will enjoy a higher per period utility from subscribing to a network l $(l = A, B)$ rather than not subscribing at all if $z_l > m^r$. In each time period she can take an action $a_r \in \{0, A, B\}$.

Some noise is added to the system by assuming that each resident becomes a 'chance adopter' with some small probability ϵ in each time period. In this case, her threshold level decreases to 0 for at least two time periods. Chance adopters will buy telephone service earlier than standard adopters. There are many reasons why some residents might get a telephone without waiting for their neighbors to subscribe first. Advertising might convince a resident to become an early adopter. Alternatively, small clusters of residents might coordinate and simultaneously subscribe.

A business user has to pay a cost v_b for every message to a customer which is not sent by phone. This can be the cost of sending a telegram or a messenger. A business chooses an action $a_b \in \{0, A, B, AB\}$ in each time period, i.e. it either uses no phone, subscribes to network A (B) or chooses dual service. Assume that network l gives access to a share z_l of customers and that a business of type $\tilde{\beta}$ sends $M\left(\tilde{\beta}\right) = \tilde{\beta}\frac{nN}{2}$ messages per time period. The per period costs of the

[23] The city council of Chicago for example launched its own investigation on telephone rates and service and forced the Bell monopolist to accept a lower schedule in 1907 (Committee on Gas, Oil and Electric Light 1907).

[24] The average cost of Bell residential service in 1905 was about 5% of a basic manufacturing worker's earnings. For this reason most residential users could not afford two telephones in their homes (Fischer 1992).

various strategies of a business can then be calculated as follows:

$$\begin{aligned}\pi_0 &= M\left(\tilde{\beta}\right)v_b \\ \pi_A &= M\left(\tilde{\beta}\right)(1-z_A)\,v_b + K \\ \pi_B &= M\left(\tilde{\beta}\right)(1-z_B)\,v_b + K \\ \pi_{AB} &= M\left(\tilde{\beta}\right)(1-z_A-z_B)\,v_b + 2K\end{aligned}$$

It will be again convenient to introduce a threshold level $m^b\left(\tilde{\beta}\right) = \frac{K}{M\left(\tilde{\beta}\right)v_b} = \frac{D}{\tilde{\beta}}$ ($D = \frac{2K}{nNv_b}$). A business will then realize a per period gain by subscribing to a network as soon as it can reach at least a share $m^b\left(\tilde{\beta}\right)$ of customers via its network.

4.3 Decision Rules

In each time period residents and businesses choose the action which maximizes their utility and minimizes their communication costs for that period.[25] This *best-response* dynamics is widely used in the evolutionary game theory literature[26] and is a very natural behavioral assumption in my model. Most subscribers were unlikely to be very forward-looking because the costs of switching a carrier were low.[27] It is therefore reasonable to assume that they compared the benefits of the different systems and then decided to subscribe to the network which suited their immediate communication needs best.

4.4 Initial Configuration

At time $t = 0$ the two competing telephone companies enter the city by soliciting a number of initial subscribers. With probability $\mu \le \frac{1}{2}$ a social island is occupied by company A (the minority company) and with probability $1-\mu$ by company B (the majority company). An entering company signs up a connected cluster of $k < n$ residents and all the businesses of the respective social island (see figure 2). The initial penetration rate amongst residents (concentration of telephone ownership) is therefore $x(0) = \frac{k}{n}$ and the market share $s(t)$ of the minority system amongst

[25] I assume that agents stick to last period's strategy in the case of a tie. This convention will not influence the results.

[26] Examples include Kandori, Mailath, and Rob (1993) and Ellison (1993).

[27] Subscribers usually had to sign contracts on an annual basis. There were no one-time connection fees except of mileage fees in a few cities for residents who lived very far away from the central office (Committee on Gas, Oil and Electric Light 1907, p. 172-).

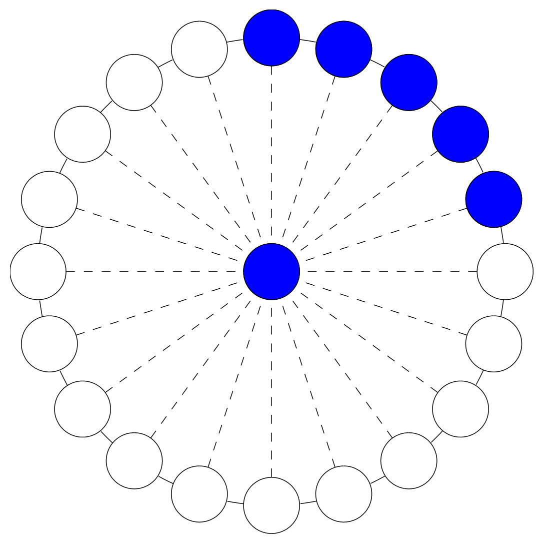

Figure 2: Social island of size $n = 20$ and an initial cluster of size $k = 5$.

residential customers is initially $s(0) = \mu$. After entry, solicitors are laid off, and the system evolves according to the best-response dynamics.

The following set of assumptions make the entry story consistent.

Assumption 2 *The parameters of the model satisfy*

$$\frac{\alpha}{2} + (1-\gamma)\beta > m^r > \beta, \tag{4}$$

$$(1-\gamma)x(0) > m^b\left(\underline{\beta}\right). \tag{5}$$

Condition 4 ensures that a resident will subscribe if the local business and at least one friend subscribe to the same company and will not subscribe if no friend uses the company. Therefore the initial cluster of subscribers will expand along its boundaries. This condition is quite intuitive - otherwise the initial cluster would either shrink instead of expanding and the telephone technology would die out, or all non-subscribers could buy telephone service at time $t = 1$ and the social space could fill instantaneously. Condition 5 allows an entering telephone company to sign up all local businesses after it has secured the conditional contracts from the initial cluster of subscribers.

5 The Dynamics of Network Competition

After both competing systems have entered all social islands at time $t = 0$ they both grow by expanding in their respective islands. However, greater telephone penetration decreases the market share of the smaller network because niche markets disappear as the market becomes saturated. It is in this sense that local service competition harbors the seeds of its own destruction because competition decreases subscriber rates and accelerates diffusion of the technology.

Theorem 1 *There exists a critical initial market share* $0 \leq \mu^* < \frac{1}{2}$ *for the minority system such that its market share* $s(t)$

a) stays constant for $\mu > \mu^*$.

b) is decreasing over time for $\mu < \mu^*$ *and converges to the long-run market share* $s_\infty = \lim_{t\to\infty} s(t)$ *which is strictly smaller than* μ. *Furthermore,* $\lim_{n\to\infty} s_\infty = 0$ *and* $\lim_{\epsilon\to 0} s_\infty = \mu$.

The proof of theorem 1 proceeds in several steps. I first characterize the evolution of both networks under the assumption that every new residential subscriber chooses the established system in her island. I then analyze the decision of business subscribers to duplicate, and its effect on new residential adopters. Because the two competing systems are of unequal size business users are duplicating in an asymmetric fashion - minority businesses are more likely to duplicate than majority businesses. This *duplication effect* makes minority islands susceptible to invasion by the majority system. For this reason, the market share of the minority system is always weakly decreasing over time.

A number of simulations at the end of the section illustrate these theoretical insights. The model is also calibrated for a typical early 20th century US city with local service competition by performing a back of the envelope calculation on the Louisville data. This exercise suggests that the duplication effect could easily apply for reasonable parameter values of the model.

5.1 Diffusion of the Telephone

In each time period, the two residents at the boundary of an *existing* cluster will always choose the established system in their island because of assumption 2. For the moment, also assume that all chance adopters choose the established system.

The initial cluster of each island grows as residents along its boundary get a telephone. Its size at the end of period t is therefore $k+2t$. Chance adopters speed up the diffusion of the telephone because they provide 'seeds' for new subscribers. The penetration rate $x(t)$ at time t can be derived easily.

Lemma 1 *The concentration of telephone ownership in the city at time* t *satisfies*

$$x(t) = 1 - \max\left[\frac{n-k-2t}{n}(1-\epsilon)^{t^2}, 0\right]. \tag{6}$$

Proof: The probability that a single household has no telephone at time t can be calculated as

$$\frac{n-k-2t}{n}(1-\epsilon)^t \prod_{i=1}^{t}(1-\epsilon)^{2(t-i)}$$

The first two terms of this expression denote the probabilities that the household is not part of the expanding original cluster and that the household has not been a chance adopter up to time t. The remaining terms describe the probabilities that the two neighbors at a distance i of the household have not been chance adopters up to time $t-i$. This expression can then be simplified to obtain formula 6. QED

Note that both systems can preserve their market share as long as all new subscribers choose the established system in their island because both minority and majority islands grow on average at the same rate.

5.2 Asymmetric Duplication

The spread of telephone ownership gives rise to duplication amongst business users. Those with greater communication needs (larger $\tilde{\beta}$) will duplicate first. Assuming that all residential users in minority islands subscribe to the minority system one can define the threshold business size $\tilde{\beta}_A(t,\mu)$ at which a minority business is just indifferent between duplicating or subscribing only to the minority company:

$$m^b\left(\tilde{\beta}\right) = \gamma(1-\mu)\,x(t) \quad \Leftrightarrow \quad \tilde{\beta}_A(t,\mu) = \frac{D}{\gamma(1-\mu)\,x(t)} \tag{7}$$

It is possible that this threshold lies below or above the support $\left[\underline{\beta},\overline{\beta}\right]$. In this case either all or none of the minority businesses duplicates. The threshold size $\tilde{\beta}_B(t,\mu)$ at which majority businesses duplicate is defined analogously:

$$\tilde{\beta}_B(t,\mu) = \frac{D}{\gamma\mu x(t)} \tag{8}$$

The first observation is that duplication is always *asymmetric* because minority businesses are more likely to duplicate than majority businesses at any state of development:

$$\frac{\tilde{\beta}_B(t,\mu)}{\tilde{\beta}_A(t,\mu)} = \frac{1-\mu}{\mu} \geq 1. \tag{9}$$

This asymmetry is greater the more unequal both systems are, i.e. the smaller the market share μ of the minority system. It will be convenient to work with an alternative measure for the asymmetry of business duplication based on the ratio

$R_D(t,\mu)$ of the volume of calls made by non-duplicating majority businesses to the volume of calls made by non-duplicating minority businesses

$$R_D(t,\mu) = \frac{G\left(\tilde{\beta}_B(t,\mu)\right)}{G\left(\tilde{\beta}_A(t,\mu)\right)} > 1 \tag{10}$$

The ratio $R_D(t,\mu)$ is a measure of how well the minority system holds on to its base of business subscribers in comparison to its rival. A larger ratio indicates the relative loss of exclusive business subscribers to the minority system.

Duplication tends to become *more asymmetric* in the sense that the ratio $R_D(t,\mu)$ increases over time, i.e. relatively fewer calls are made by non-duplicating minority businesses. This claim is obviously true if the minority system is very small because no majority business would ever duplicate. However more and more minority businesses do so as an increasing number of their outside customers have a telephone. For an intermediate market share of the minority system, the claim is true 'on average': at a low state of development few businesses duplicate on either side and the ratio R_D is close to 1. At later stages most (possibly all) minority businesses duplicate and the ratio R_D becomes large. Our technical assumption 1 ensures that the ratio $R_D(t,\mu)$ (and therefore the degree of asymmetry) increases at all time periods.

5.3 Chance Adopters and the Duplication Effect

The network choice of chance adopters depends solely on the share of business calls that can be made through each system because chance adopters have no 'friends' who possess a telephone. It is easy to show that for chance adopters in minority islands the minority system is always the dominant choice in the absence of duplication (note, that $\gamma < \frac{1}{2}$):

$$(1-\gamma)\beta + \mu\gamma\beta > (1-\mu)\gamma\beta \tag{11}$$

An analogous condition holds even more strongly for chance adopters in majority islands.

Duplication can change this tradeoff between both systems for chance adopters in minority islands (asymmetric duplication makes the majority system even more attractive for chance adopters in majority islands). Condition 11 becomes:

$$\begin{aligned}(1-\gamma)\beta + \mu\gamma\beta &+ (1-\mu)\gamma\int_{\tilde{\beta}_B(t,\mu)}^{\overline{\beta}} \tilde{\beta} g\left(\tilde{\beta}\right) d\tilde{\beta} > (1-\gamma)\int_{\tilde{\beta}_A(t,\mu)}^{\overline{\beta}} \tilde{\beta} g\left(\tilde{\beta}\right) d\tilde{\beta} \\ &+ (1-\mu)\gamma\beta + \mu\gamma\int_{\tilde{\beta}_A(t,\mu)}^{\overline{\beta}} \tilde{\beta} g\left(\tilde{\beta}\right) d\tilde{\beta}\end{aligned} \tag{12}$$

The inequality can be simplified:

$$\frac{1-(1-\mu)\,\gamma}{(1-\mu)\,\gamma} > \frac{G\left(\tilde{\beta}_B\,(t,\mu)\right)}{G\left(\tilde{\beta}_A\,(t,\mu)\right)} = R_D\,(t,\mu) \tag{13}$$

If both systems have equal market share ($\mu = \frac{1}{2}$) the ratio R_D always equals one because minority and majority businesses duplicate at the same rate. Condition 13 then reduces to condition 11 and chance adopters in minority islands have no incentives to ever defect to the majority system. However, the more unequal both networks are (the smaller is μ), the harder it is for the minority system to hold on to its business subscribers because duplication becomes ever more asymmetric and increases over time. If condition 13 is violated, the minority system can cease to be the dominant choice for chance adopters in those islands. This phenomenon is referred to as the *duplication effect*.

As long as the duplication effect holds chance adopters always defect to the majority system and form new seeds for majority clusters inside minority islands. Note, that residents at the boundary of such an 'invading' cluster will subscribe to the majority system if they do not own a telephone yet, or switch to the dominant system if they do (see figure 3). Due to condition 13, they prefer the majority system for communication with businesses and they are at most indifferent between both systems for residential calls because at least one of their two neighbors subscribes to it. The minority system will therefore gradually lose 'invaded' islands and hence market share to the majority system, even though it might continue to grow for a while in terms of an absolute number of subscribers.

The duplication effect is most easily generated when the minority system has a market size close to $\mu = 0$. No majority business ever sees a need to duplicate while more and more minority businesses do so as telephone ownership spreads. A necessary and sufficient condition for the duplication effect to apply at some state of development is the violation of condition 13 under full development:

$$\frac{1-(1-\mu)\,\gamma}{(1-\mu)\,\gamma} < \frac{1}{G\left(\frac{D}{\gamma(1-\mu)}\right)} \tag{14}$$

For example, the duplication effect will always apply if all minority businesses duplicate under full development.

If condition 14 does not hold, the duplication effect will never apply, and the minority system will hold onto its initial market share at all stages of development (this is the special case $\mu^* = 0$ in theorem 1). Otherwise, there will be some critical $\mu^* > 0$ such that the duplication effect does not apply for market shares of the minority system above this critical level, but will apply at some point in time for all market shares below it.

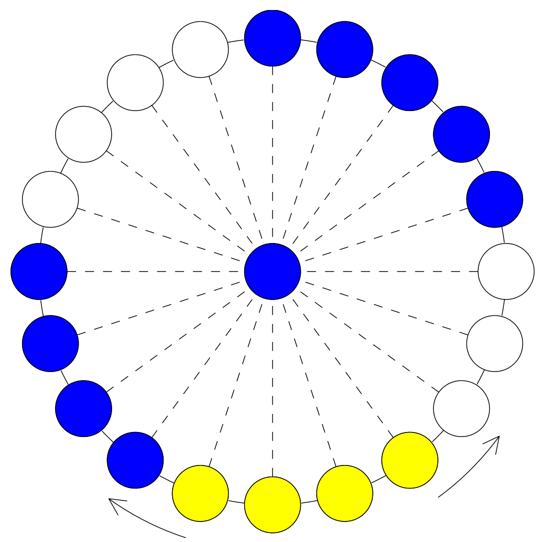

Figure 3: Social island of minority company has been invaded by the majority company (lightly shaded): majority cluster expands at the expense of the minority company

Assumption 1 guarantees that the duplication effect is always monotonic, i.e. if it applies at some time t_1 it applies at any time $t_2 > t_1$. However, the duplication effect has a natural tendency to persist even if assumption 14 does not apply. The invasion of minority islands makes duplication even more (less) attractive for minority (majority) businesses. Therefore, duplication becomes even more asymmetric and the duplication effect can continue to apply even though condition 13 is not violated (this condition was derived under the assumption that all subscribers choose the established system in their island). In this way the duplication effect feeds on itself.

5.4 Proof of Theorem 1

The first part of theorem 1 simply states that the minority system will not lose market share if the duplication effect does not hold even at full development.[28] This will certainly be the case if the initial market share of the minority system is sufficiently close to $\frac{1}{2}$. The second part of the theorem considers the opposite case where the duplication effect applies after some time period $t_D < t_F$ before full development has been achieved. From then on, all chance adopters in minority

[28]Note, that full development is achieved at period $t_F = \left[\frac{n-k}{2}\right]^+$ time periods, where $[x]^+$ denotes the smallest integer greater or equal than x. After this interval the initial cluster has expanded over its entire island even if there has been no chance adoption in the island.

islands defect to the majority system and every 'infected' social island eventually falls to the majority system. Because some islands will always become infected, the long-run market share of the minority system has to be strictly less than μ. The minority system can only survive in those islands where no chance adoption takes place between periods t_D and t_F; such an island is referred to as isolated. As the size of social islands increases, the probability of an island to be isolated tends to 0 because $t_F - t_D$ increases. The long-run market share of the minority system therefore tends to 0. On the contrary, every island tends to stay isolated as chance adoption becomes rare. The minority system can then hold on to its customer base.[29] QED

Note, that the long-run market share s_∞ of the minority system is always (weakly) smaller than its initial market share μ even if the duplication effect is non-monotonic (assumption 1 does not apply).

5.5 Numerical Simulations

In order to illustrate the duplication effect I simulate the model for two cities with social islands of size $n = 100$ and $\epsilon = 0.01$. The minority network has initial market share of $\mu = 20$ percent and the initial telephone concentration is $x(0) = 5$ percent. I assume that each island is occupied by only a single type of business and that a minority business starts to duplicate in city 1 and 2 when the telephone penetration reaches the threshold $x_A^* = 30$ percent and $x_A^* = 50$ percent respectively. In both cities majority businesses would never want to get a second phone at any state of development because the threshold penetration rate for majority businesses would be $x_B^* = \frac{1-\mu}{\mu} x_A^* > 1$. Therefore, all chance adopters in minority islands will defect to the majority system as soon as the penetration rate has reached the critical level x_A^* and the duplication effect holds.

Figure 4 shows how the market share of the minority company is initially stable. As soon as minority businesses start to duplicate, minority social islands get invaded. Although both systems still grow in absolute terms the market share of the minority system is decreasing as new subscribers are more likely to join the majority company. Eventually the minority company will suffer a net outflow of subscribers as their users will increasingly defect to the majority company. Although most islands will get invaded a few minority islands will be spared such

[29] It is instructive to compare the role of noise (chance adopters) in my model with its use in evolutionary game theory where it has become a common tool for selecting between multiple equilibria. Kandori, Mailath, and Rob (1993) and Young (1993) introduced the concept of a stochastically stable equilibrium by analyzing the behavior of the system in the limit when $\epsilon \to 0$ while keeping the 'size' n of the system constant. For my model stochastic stability is of no interest because the 'undisturbed' model has a unique equilibrium. I am rather interested in the precise medium- and long-run dynamics of the system for the case where the noise is small but not negligible.

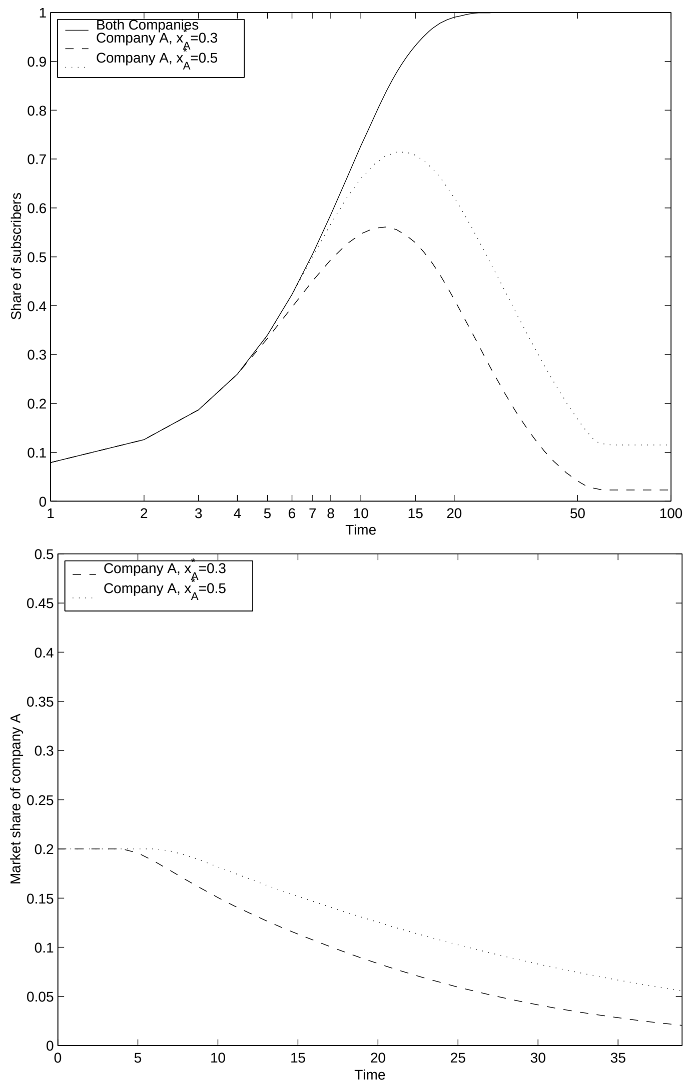

Figure 4: Average share of subscribers to both telephone systems in a minority social island when minority businesses start to duplicate at a penetration rate of $x_A^* = 30$ percent and $x_A^* = 50$ percent (top with log time axis); evolution of market share of minority system (bottom): $\epsilon = 0.01$, $n = 100$, $x(0) = 0.05$, $\mu = 0.2$

that the long-run market share s_∞ of the minority system is bounded away from 0.

It is particularly interesting to observe that a marked slowdown in the diffusion rate of the telephone approximately coincides with the absolute decline of the minority company. This phenomenon is very intuitive: once the social islands fill up, the minority company can no longer offset the defection of its customers to the majority company by wooing first-time subscribers. While the majority company initially expands through the growth of the entire market, it subsequently grows by increasing its market share at the expense of its rival.

5.6 Calibrating the Model for the City of Louisville, KY

The Louisville data in table 2 is sufficiently detailed to perform a simple back of the envelope calculation and calibrate the model for a typical US city with local service competition.

The Louisville survey reveals that the market share of the minority company (here the Bell system) amongst residents and small-scale businesses was about 40 percent. Now assume, as done in the model, that each social island has a similar number of associated type $\tilde{\beta}$ businesses. This implies that about 40 percent of all medium-scale businesses should be minority businesses. Of the 1,803 medium-scale businesses 833 (46 percent) subscribe to the Bell system either exclusively or as a second phone. This suggests that most duplicating businesses were minority businesses with a duplication rate of about 70 percent. In contrast, few majority businesses seem to have duplicated.

Plugging these parameters into condition 13 one can calculate an upper bound for the share of inter-island communication γ which would be required for the duplication effect *not* to apply and chance adopters in minority islands to choose the Home Telephone Company (w_A and w_B are the raw duplication rates of all minority and majority businesses):

$$\begin{aligned}\gamma &< \left[(1-\mu)\left(\frac{G\left(\tilde{\beta}_B\right)-G\left(\tilde{\beta}_A\right)+G\left(\tilde{\beta}_A\right)}{G\left(\tilde{\beta}_A\right)}+1\right)\right]^{-1} \\ &\leq \left[(1-\mu)\left(\frac{1-w_B}{1-w_A}+1\right)\right]^{-1}. \end{aligned} \tag{15}$$

For the second inequality the following fact is used:

$$\frac{G\left(\tilde{\beta}_B\right)-G\left(\tilde{\beta}_A\right)}{G\left(\tilde{\beta}_A\right)} = \frac{\int_{\tilde{\beta}_A}^{\tilde{\beta}_B} \tilde{\beta} g\left(\tilde{\beta}\right) d\tilde{\beta}}{\int_{\underline{\beta}}^{\tilde{\beta}_A} \tilde{\beta} g\left(\tilde{\beta}\right) d\tilde{\beta}} \geq \frac{(w_A-w_B)\,\tilde{\beta}_A}{(1-w_A)\,\tilde{\beta}_A}.$$

The exercise suggests that the duplication effect was likely to be relevant unless the social islands were fairly isolated from each other ($\gamma < 0.38$). Note, that Louisville is not an ideal candidate for the duplication effect to apply because the two competing networks had similar size. Nevertheless, the model can generate the duplication effect even in this example for reasonable parameter values.

6 The Cost of Incompatibility

Lack of interconnection between competing networks generates efficiency losses as subscribers cannot call users on the rival system without buying telephone service from both companies. These losses are particularly high for business subscribers: businesses in minority islands, for example, cannot reach a share $\gamma(1-\mu)$ of their customers with a telephone while residential users miss out on no more than a share $\beta\gamma(1-\mu)$ of their calls. Furthermore, the efficiency costs of competition increase over time for business users but decrease for residential subscribers because an increasing number of businesses duplicate.

In this section two phenomena are analyzed: the specific efficiency costs of competition $C^b_{\tilde{\beta}}(t,\mu)$ to business subscribers of type $\tilde{\beta}$ and the average efficiency loss to all business users defined as:

$$C^b(t,\mu) = \int_{\underline{\beta}}^{\overline{\beta}} C^b_{\tilde{\beta}}(t,\mu)\, g\left(\tilde{\beta}\right) d\tilde{\beta} \tag{16}$$

This *cost of incompatibility* has to be compared with the price mark-up which a monopolistic telephone company would charge. The difference between the cost of incompatibility and the cost of monopoly serves as a proxy for the level of public support for competition. I do not derive the pricing policy of a monopolist explicitly in this paper. Instead, I argue that changes in the level of support for competition can be inferred from changes in the cost of incompatibility. Demands for compulsory interconnection, and an end to 'wasteful competition' are likely to become more vocal if the efficiency costs of competition increase over time. This phenomenon is called here the *political economy effect*.

The main results are summarized in the next theorem and proved in appendix A. The first part of the theorem tells us that the efficiency loss due to competition is largest if the competing networks have similar market shares. This is easy to see in the absence of business duplication when the efficiency loss can be expressed as

$$C^b(t,\mu) = 2\gamma\mu(1-\mu)\, x(t)\, \beta \frac{nN}{2} v_b \tag{17}$$

which is strictly increasing in the market share μ of the smaller system. Duplication complicates this expression considerably but the proof in appendix A shows that the basic result goes through. In the case of both networks being of similar size the second part of theorem 2 also tells us that the efficiency loss is increasing over time.

Theorem 2 *The total cost of incompatibility* $C^b(t,\mu)$ *is strictly increasing in the market share* μ *of the smaller system. If the duplication effect does not apply* ($\mu > \mu^*$) *then the type specific cost* $C^b_{\tilde{\beta}}(t,\mu)$

1. *is increasing in both the market share* μ *of the smaller system and over time* t.

2. *is bounded below by* $2\mu K$ *if type* $\tilde{\beta}$ *businesses have started to duplicate.*

Some remarks on theorem 2 are in order.

1. The political economy effect and the duplication effect complement in each other in working against the minority system. Support for competition is likely to be the lowest if both systems have similar market share, and it grows as telephone ownership spreads. On the other hand, the efficiency losses due to competition are low if the minority system has a small market share. The reverse is true for the duplication effect: it never applies if both systems have similar share. However, if the minority system is sufficiently small the duplication effect tends to decrease the market share of the minority system at advanced penetration rates.

2. Despite potentially large efficiency costs of competition in mature markets businesses are likely to have welcomed competition at low states of development. First, the initial market share of the entrant is likely to be small because it could only enter a few free social islands at once. Second, telephone penetration is low which implies a small efficiency losses even if the independent system manages to build up a substantial market share. The largest business users with the highest duplication rate are likely to be the first ones to oppose competition. This seems to have been indeed the case in most cities.

3. Residential users, in particular subscribers to the majority system, are able to conduct an increasing share of their communication over the telephone due to increasing duplication. If they support competition at time $t = 0$ they should certainly support it at later time periods. While initially both residents and businesses share the (small) cost of competition, the burden shifts more and more to duplicating business subscribers over time. This suggests that plans for the consolidation of competing exchanges under a regulated monopoly would have to 'buy off' small-scale users, and offer them rates comparable to those under competition.

4. The duplication effect can make the type specific cost of incompatibility non-increasing in both the market share of the smaller system and over time. Defecting subscribers in minority islands cause a greater efficiency loss to non-duplicating business subscribers in minority islands than efficiency gains for non-duplicating business subscribers in majority islands (as $\gamma < \frac{1}{2}$). However, the

average efficiency losses are lower for more unequal systems because the duplication effect benefits non-duplicating majority businesses and does not hurt duplicating minority businesses (the reverse case is ruled out by asymmetric duplication).

6.1 Estimating the Cost of Incompatibility in Louisville

Louisville is a good candidate for estimating a fairly large efficiency cost because the market share of the minority system is relatively large at about $\mu = 40$ percent. A duplication rate for minority businesses is estimated to be about 70 percent, and close to 0 for businesses in majority islands. Theorem 2 therefore suggests that the average cost of competition to medium-scale businesses in Louisville was at least $2 \times 0.40 \times 0.7K = 0.56K$ (assuming that the duplication effect did not apply and $\mu > \mu^*$).

Unless business users in Louisville expected a 56 percent rate increase from a monopolistic telephone provider, they would have been likely to reject dual service competition in favor of consolidation in the industry. Such large increases were hard to push through, however, because telephone companies operated on a franchise from the city council. In Louisville both networks in fact consolidated in 1910, when the Bell subsidiary bought the Home Telephone Company. After the Kingsbury Commitment in 1913 rate hikes were even harder to achieve as consolidations could only proceed if all relevant user groups consented.

7 Saving Local Service Competition through Interconnection

The model makes grim predictions for the long-term survival of local service competition. If the competing operators are of similar size, the cost of incompatibility becomes unacceptably high and business subscribers will prefer a monopoly. Otherwise the duplication effect can drive the minority system towards extinction. In this section, I show that a requirement for compulsory interconnection by regulators could have preserved dual service competition in the long run by making duplication unnecessary.

Assume that all subscribers can place a call on the rival network by paying a small interconnection fee c per message. Minority businesses then never duplicate as long as the cost of a second telephone exceeds the total expense on interconnection charges, e.g.[30]

$$K > \gamma (1 - \mu) M \left(\tilde{\beta}\right) c. \tag{18}$$

[30]Note, that the total interconnection charge will be largest for minority businesses when the market share μ is close to 0.

Because the fixed rate K includes the cost of maintaining an extra wire for each subscriber while interconnection just requires a link-up of the central offices of the two competing networks through a trunk line this condition presumably held.

Without duplication each chance adopter minimizes the interconnection charges by simply adopting the same system as all other residents in the island. Interconnection therefore reduces both the cost of incompatibility and eliminates the duplication effect. The market share of the minority system does not decrease over time and its long-term survival is ensured.

The model therefore suggests that the lack of interconnection was a crucial factor for explaining the decline of local service competition. This observation makes AT&T opposition to interconnection entirely rational. However, it also suggests that regulators acted too-little-too-late in order to save local service competition. The fragmented nature of regulation at the city level in the early stage of competition and the absence of a legal framework made compulsory interconnection difficult. The states began to form regulatory commissions only after 1909, and the Federal Communications Commission was only created in 1934 (Mueller 1997). By that time local service competition was already in terminal decline.

8 Empirical Evidence

After the Kingsbury commitment in 1913, antitrust authorities allowed further consolidation only if there was was strong local support either from the city council or through a local referendum. In order to waive the Kingsbury commitment the AT&T subsidiaries offered swaps of territories with independent operators which eliminated local service competition but left the number of independent telephones unchanged.

During this late stage of competition public support for unifying telephone service was high. The number of locations with local service competition declined from more than 1,800 in 1914 to roughly 1,000 at the end of world war I (Mueller 1997). In 1921 Congress passed the Willis-Graham Act which nullified the Kingsbury Commitment and allowed further consolidation under the supervision of the Interstate Commerce Commission. By the mid-1920s local service competition had been eliminated.

The dynamics of local competition between 1893 and 1921 is consistent with our model of 'delayed network effects' which allows both for unconstrained competition at low telephone penetration rates, and an increasing tendency towards standardization in more mature markets. This section tests how well the model performs at closer inspection. In particular, I look at the following predictions of the theory.

1. The duplication rate amongst business users increases over time.

2. Minority businesses duplicate at a higher rate than majority businesses.

3. In cities where one system dominates, the market users converge to this system at later stages of development (duplication effect).

4. Public support of competition decreases over time due to increasing efficiency losses of non-interconnected networks (political economy effect). This tendency is particularly strong in cities where both systems are of comparable size.

8.1 The Evolution of Business Duplication

Table 3 documents the evolution of local service competition between the AT&T subsidiary and an independent rival operator in eight Northern cities over at least two time periods. The table summarizes the shares of exclusive subscribers and dual service users. All cities with the exception of Louisville experienced rapid annual growth of the telephone networks which significantly increased the concentration of subscribers in the urban population.

The total rate of duplication defined as the share of duplicates in the total number of subscribers increased for half of the cities over time, remained approximately constant for three cities and decreased slightly for one (Kansas City). However, for the purpose of the model it is the rate of business duplication that matters. Appendix B describes in detail the construction of a *projected duplication rate* for business subscribers which is also listed in table 3. Essentially, this calculation is based on the assumption that half of all subscribers in 1905 were medium and large scale businesses, and that most subsequent subscribers were residents and small-scale businesses.

The rapid increase in the projected duplication rate suggests that the business duplication rates increased strongly as a result of increasing telephone penetration. For four out of eight cities the share of duplicate business users exceeds 50 percent by 1907.[31]

8.2 Evidence for Asymmetric Duplication

In order to generate delayed network effects the rate of duplication amongst business subscribers not only has to increase over time, but it also has to become sufficiently asymmetric for the duplication effect to apply.

[31] These numbers seem plausible when compared to the detailed telephone census of Louisville in table 2. Although this city experienced only very low telephone growth the duplication rate amongst medium-scale businesses was already between 30 percent and 50 percent.

The strongest evidence for asymmetric duplication can be found in the 1910 telephone census of Louisville (see table 2). In the back of the envelope calculation for Louisville in section 5.6, I argued that most of the duplicating business subscribers were minority businesses while few of the majority businesses seem to have duplicated.

For other cities no breakdown of exclusive and duplicate subscribers into businesses and residents is available. However, attention can be restricted to the sample of cities in table 4 where the market share of the minority system was at most 35 percent and therefore smaller than the market share of the minority system in Louisville.[46] The model suggests that the minority businesses in those cities are at least as likely to duplicate as minority businesses in Louisville while the reverse is true for majority businesses. This suggests that most of the duplicating businesses in these cities were also minority businesses. In half of the cities of table 4 the share of duplicating independent subscribers is above 50 percent and in 12 out of 14 cities the share is above 40 percent in 1905. This translates into duplication

[32]The annual rates are given for business/residential unlimited single-party service respectively. Although the Bell rates in Louisville, Kansas City, Minneapolis and St. Paul seem a lot higher than the independent rates, subscribers could choose two-party and four-party service at rates comparable to independent single-line rates. The Independents on the other hand did not offer multi-party lines in those cities.

[33]The projected duplication rate (in percent) has been calculated by dividing the number of duplicates by the reference value B_C. The reference value for the number of 'businesses' in a city has been defined as 50 percent of all subscribers in July 1905.

[34]Telephone concentration refers to telephones per inhabitant. The population data for 1900 and 1910 is taken from the 1910 census. Intermediate values were interpolated.

[35]*Telephone Magazine* January 1903: p. 6

[36]All data for the year 1905 is taken from the Supplemental Telephone Report (Merchants' Association of New York 1905, p. 3-7).

[37]Anderson (1907, p. 28)

[38]Johnston (1908, p. 7f, 24)

[39]The Indianapolis 1903/1904 data is taken from *The American Telephone Journal* May 14, 1904: p. 309.

[40]For data on the number of exclusive and duplicate subscribers to the New company see Anderson (1907, p. 14). Anderson also mentions that the Bell subscriber base is *about* as big as the independent one. In 1907, however, Bell had only 9,800 subscribers and the New company 11,500 (Committee on Gas, Oil and Electric Light 1907, p. 192). I therefore scaled down the number of Bell subscribers by 20 percent.

[41]The 1907 data is taken from the Chicago report(Committee on Gas, Oil and Electric Light 1907, p. 172-).

[42]Mueller (1997, p. 82)

[43]Instead of the 1905 rate schedule for Minneapolis and St. Paul and the 1903 schedule for Indianapolis I used 1902 rates quoted in *The American Telephone Journal* April 12, 1902: p. 229.

[44]Anderson (1907, p26)

[45]Rate schedule for 1906 quoted from *Soundwaves* November 1906: p. 433.

[46]The maximal market share of the minority system is the sum of the duplication rate and the share of exclusive independent subscribers.

Table 3: Panel data documenting the evolution of duplication in eight US cities

City *Bell* *Independent*	Year/ Month	Total sub- scribers	Market share (in %)			Proj. dupl. [33]	Conc. (in %) [34]	Rates (in US$) [32]		
			Bell only	Dupli- cations	Indep. only			Bell	Indep.	Dual
Cleveland	1903/1 [35]	21,000	52.4	*14.3*	33.3	20.3	4.8	48/-	48/36	96/-
Cleveland Tel.	1905/7 [36]	29,560	43.4	*11.9*	44.7	23.8	6.2			
Cuyahoga Tel.	1906/9 [37]	42,000	45.2	*21.4*	33.3	60.8	8.4	84/54	72/48	156/102
Columbus, O	1905/7	12,904	42.8	*10.2*	47.0	20.4	8.3	72/42	40/24	112/66
Central Union	1908/7 [38]	21,340	46.1	*10.1*	43.8	33.4	12.3	54/27	40/24	94/51
Citizens Tel. Comp.										
Indianapolis	1903/10 [39]	11,980	42.6	*11.8*	45.6	21.8	6.2	40/24	40/24	80/48
Central Union	1904/4	12,716	39.5	*10.4*	50.1	20.4	6.5			
New Tel. Comp.	1905/7	12,965	29.5	*12.0*	58.4	24.0	6.3			
	1906/3 [40]	17,000	41.2	*11.8*	47.1	30.9	8.1	54/24	40/24	94/48
Kansas City	1905/7	18,078	46.1	*15.9*	38.0	31.8	8.6			
Missouri Tel. Comp.	1907/1 [41]	32,920	49.0	*13.7*	37.3	49.9	14.8	96/36	54/36	150/72
Home Tel. Comp										
Louisville, Ky	1905/7	14,343	43.8	*17.6*	38.6	35.2	6.7	96/30	48/24	144/54
Cumberland T&T	1910 [42]	16,263	37.9	*18.0*	44.1	40.8	7.3	96/36	48/24	144/60
Home Tel. Comp										
Minneapolis	1905/7 [43]	20,147	50.6	*15.1*	34.4	30.2	7.8	84/60	48/30	132/90
Northwestern Tel.	1907/1	24,998	43.2	*16.9*	40.0	41.9	9.2	90/54	48/30	138/84
Tristate Tel. Comp										
St. Paul, Minn	1905/7	11,124	67.7	*13.3*	18.9	26.6	5.8	84/60	48/30	132/90
Northwestern Tel.	1907/1	14,834	50.5	*19.2*	30.3	51.2	7.4	90/54	48/30	138/84
Tristate Tel. Comp										
Toledo, O	1905/7	9,301	25.0	*10.2*	64.8	20.4	6.1	54/27	48/26	102/53
Central Union	1906/9 [44]	13,300	24.8	*25.6*	49.6	73.2	8.5		52/32 [45]	
Home Comp.										

Table 4: Sample of cities with a dominant system in 1905

City	Market share (in %)			Share of duplicating independent subscribers
	Bell only	Duplications	Indep. only	
Atlanta, Ga	72	18	10	64
Buffalo, NY	66	18	16	53
Columbus, Ga	71	21	8	72
Elgin, Ill	74	7	19	27
Fall River, MA	68	16	16	50
Harrisburg, Pa	70	13	17	43
Mobile, Ala	78	9	13	41
N. Bedford, MA	72	13	15	46
Norfolk, Va	64	18	18	50
Philadelphia, Pa	78	16	6	73
Pittsburgh, Pa	69	14	17	45
Portsmouth, Va	80	11	9	55
St Paul, Minn	68	13	19	41
Syracuse, NY	70	10	20	33

Source: Merchants' Association of New York (1905, p. 5-6): Supplemental Telephone Report. The table includes all the cities from the original source where the minority system had a market share of less than 35%. In all 14 cases the Bell exchange had the commanding lead.

rates of between 80 and 100 percent amongst minority businesses in those cities. These estimates suggest that duplication amongst minority businesses was high enough in cities with a dominating system to make minority islands vulnerable to invasion.

8.3 User Convergence to the Dominant System

After 1907 the annual diffusion rate of the telephone was about half of the rate in the previous years indicating a gradual saturation of the market (see table 1). According to the model delayed network effects are most likely to decrease the market share of minority systems during these late stages of competition.

A nice example for user convergence to the dominant system in western New York during the late competitive era is described by Mueller (1997, p. 137-140). The independent Federal Telephone Co. owned 35 independent exchanges including system in Buffalo, Rochester and Jamestown. In 1916, the manager of the Buffalo-based company, Burt G. Hubbell observed a tendency among subscribers in dual towns to gravitate towards the larger of the two systems. This problem was particularly acute in Buffalo, where Bell outnumbered the independent by a ratio of three to one and where the subscriber list of the Independent was shrinking. Hubbell noted:

> The natural tendency of the public to patronize the company with the largest number of subscribers ... has led to a segregation into telephone districts in each of which one of the two competitors has usually acquired a great predominance of subscribers.

Hubbell had failed to stop the migration of subscribers to the Bell system even after heavy advertising campaigns. The Bell system agreed to consolidate but required the approval of a majority of telephone users in order to apply for a waiver under the Kingsbury Commitment. The Buffalo Chamber of Commerce agreed on a new rate structure which left rates in the middle and the bottom of the communication hierarchy unchanged, the competitors swapped territories and thereby eliminated local service competition. Bell acquired the Buffalo exchange while the Independent gained a local monopoly over Rochester and Jamestown.

8.4 Growing Public Opposition to Local Service Competition

A second type of delayed network effects is the increasing cost of incompatibility and the associated loss of political support for competition. According to the model this effect should be strongest in cities where both systems had roughly similar size. Our sample of cities in table 3 fulfills this requirement. The total business duplication rate $w(t, \mu)$ (i.e. the projected duplication rate) in each city can be used as a lower bound for the average cost of incompatibility to business users:

$$C^b(t, \mu) \geq w(t, \mu) K \tag{19}$$

If separate information on the duplication rates of minority and majority businesses respectively were available, a stricter bound using theorem 2 could be calculated.

Nevertheless, increases in the cost of incompatibility are stark exceeding 50 percent for four of the eight cities. Businesses are likely to have opposed competition in those cities unless they feared enormous rate increases due to consolidation. Public opinion indeed gradually shifted by 1907. Town councils which had welcomed the Independents in the early competitive era became more careful in granting new franchises. Herbert Casson (1910, p. 191), an early historian of the telephone industry, explained the change in public opinion:

> Most people fancied that a telephone system was practically the same as a gas or electric light system, which can often be duplicated with

> the result of cheaper rates and better service. They did not for years discover that two telephones in one city means either half service or double cost.

Municipal governments in Kansas City, Cleveland and Indianapolis began their own investigations of the telephone situation and recommended local interconnection. In Cleveland, the city council described duplication a "nuisance" in 1908 (Mueller 1997, p. 121). This is unsurprising given that at least 60 percent of medium-scale businesses subscribed to both networks by 1906 already (see table 3). In January 1907, motions were introduced in the state legislatures of Nebraska, Missouri, and Kansas for compulsory local interconnection between companies doing business in the same city or village. Twenty-two states finally passed interconnection bills between 1910 and 1913 empowering utility commissions to order interconnection if telephone users of some locality demanded it.

Enforcement was cumbersome, however, and AT&T's opposition to interconnection without consolidation in competitive service cities meant that these laws were almost never applied. On the other hand, support of an outright merger was often tempered by the fear of large rate increases. The Kingsbury Commitment made consolidation a more attractive option for competitive cities as it gave city councils and users enough leverage to negotiate favorable rates after consolidation.

A case study of post-Kingsbury consolidation in Southern California illustrates particularly nicely the political dynamics which worked against dual service in a region where both networks divided the market roughly equally (Mueller 1997, p. 140-43). By 1916, Bell operated 11 exchanges in Los Angeles with 67,000 stations, while the independent company had 60,300 subscribers on 14 exchanges. Both networks had extensive toll connections in the region. In 1910 political agitation against duplicate service began and in 1915 a municipal referendum was held. A 5:1 majority of voters expressed their support for compulsory interconnection which would have left competition between both companies intact. This option turned out to be too expensive as the rival systems were technically incompatible.[47] The city council then decided to insist on consolidation under rate regulation and the companies merged in 1917. A survey conducted by an economics student at the University of Southern California in 1916 found that business users rather than residents were the driving force behind calls for consolidation as predicted by the theoretical model: 100 percent of business subscribers in his sample were troubled by being unable to reach people on another network, but 34 percent of housewives were never troubled at all (Mueller 1997, p. 140-43).

[47] This does not imply that an early general regulatory framework requiring interconnection for the telephone industry would have been expensive as the rival systems would have been technically compatible from the start. The refusal of AT&T to even sell telephone equipment to the Independents led to the development of different technical standards in both networks.

9 Conclusion

This paper argues that local service competition was bound to be a transitory phenomenon in the absence of interconnection between the rival networks because of delayed network effects. When both systems were of similar size, the efficiency loss due to incompatibility was greatest and rising over time. This reversed initial public support for competition and increased the pressure on the Independents to consolidate. On the other hand, if one network dominated the market the minority system tended to lose control of its niche islands since users would eventually gravitate towards the larger system. Both types of delayed network effects promoted the rise of regional monopolies.

This insight allows for a re-interpretation of the predatory policies which AT&T pursued before the Kingsbury Commitment. Rather than destroying local service competition, AT&T secured itself the dominant position in the post-competitive era by taking control of the majority of regional monopolies. In particular, AT&T weakened the independent movement to an extent that it never managed to build a rival long-distance network. When technological progress made long-distance telephony affordable to most subscribers after World War I, AT&T could reap the full benefits of defending its monopoly in this market.

The model causally links the eventual demise of competition with the lack of interconnection between competing networks. It suggests that AT&T's refusal to interconnect rather than predatory pricing was most damaging to competition in the telephone industry.

The modeling approach introduced in this paper can be applied to the analysis of delayed network effects in modern network battles. Two pertinent examples include competition between cellphone operators, and the battle for dominance between the Windows operating system and the Apple Macintosh. In the case of cellphones one might worry whether regulators should permit operators to discriminate between calls which stay within their networks and calls which are made to (or arrive from) outside their networks. Price discrimination creates artificial network effects which might be inconsequential at low cellphone penetration rates but might drive smaller systems out of the market at advanced stages of development. In the case of operating systems, one can observe that Apple was able to hold on to roughly 15 percent of the PC market between 1985 and 1995. However, in subsequent years the gradual saturation of the PC market coincided with a dramatic drop in Apple's market share and the company became increasingly dependent on some well defined niche markets such as education where it enjoyed a dominant position.

References

Anderson, G. (1907): "Telephone Competition in the Middle West and Its Lessons for New England," prepared for New England Telephone and Telegraph Company.

Bornholz, R., and D. S. Evans (1983): "The Early History of Competition in the Telephone Industry," in *Breaking Up Bell: Essays on Industrial Organization and Regulation*, ed. by D. S. Evans. North-Holland, New York.

Brooks, J. (1975): *Telephone*. Harper and Row, New York.

Bureau of the Census (1902): *Telephones and Telegraphs*. Government Printing Office, Washington.

——— (1912): *Telephones and Telegraphs*. Government Printing Office, Washington.

——— (1975): *Historical Statistics of the U.S., Colonial Times to 1970*, II. Government Printing Office, Washington.

Casson, H. N. (1910): *The History of the Telephone*. A.C. McClurg & Co., Chicago.

Committee on Gas, Oil and Electric Light (1907): "Telephone Service and Rates," Report to the City Coucil of Chicago.

Ellison, G. (1993): "Learning, Local Interaction, and Coordination," *Econometrica*, 61, 1047–1071.

Fischer, C. S. (1992): *America Calling: A Social History of the Telephone to 1940*. University of California Press, Berkeley.

Gabel, R. (1969): "The Early Competitive Era in Telephone Communication, 1893-1920," *Law and Contemporary Problems*, pp. 340–69.

Geijsbeek, J. B., and P. Lawrence (1914): *Report on the Denver Plant of The Mountain States Telephone and Telegraph Co.* Denver, Colorado.

Granovetter, M. (1973): "The Strength of Weak Ties," *American Journal of Sociology*, 78, 1360–1380.

——— (1974): *Getting a Job*. Harvard University Press, Cambridge.

Johnston, G. R. (1908): *Some Comments on the 1907 Annual Report of the American Telephone and Telegraph Company.* International Independent Telephone Association, Chicago.

Kandori, M., G. Mailath, and R. Rob (1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, 61, 29–56.

Kern, S. (1983): *The Culture of Time and Space, 1880-1918.* Harvard University Press, Cambridge.

Langdale, J. V. (1978): "The Growth of Long-Distance Telephony in the Bell System: 1875-1907," *Journal of Historical Geography*, 4, 145–59.

Lipartito, K. (1989): "System Building at the Margin: The Problem of Public Choice in the Telephone Industry," *Journal of Economic History*, 49, 323–36.

MacMeal, H. B. (1934): *The Story of Independent Telephony.* Independent Pioneer Telephone Association, Chicago.

Merchants' Association of New York (1905): *Supplemental Telephone Report: Further Inquiry into Effect of Competition.* Merchants' Association, New York.

Meyerowitz, J. (1985): *No Sense of Place: The Impact of Electronic Media on Social Behavior.* Oxford University Press, New York.

Mueller, M. L. (1997): *Universal Service: Competition, Interconnection, and Monopoly in the Making of the American Telephone System.* MIT Press.

New England Telephone and Telegraph Company (1908): "Competition in Telephony," .

Stehman, J. W. (1925): *The Financial History of the American Telephone and Telegraph Company.* Houghton Mifflin, Boston.

U.S. Federal Communications Commission (1938): *Proposed Report: Telephone Investigation.* Government Printing Office, Washington.

Watts, D. J., and S. H. Strogatz (1998): "Collective Dynamics of 'Small-World' Networks," *Nature*, 393, 440–442.

Weiman, D. F., and R. C. Levin (1994): "Preying for Monopoly? The Case of Southern Bell Telephone Company: 1894-1912," *Journal of Political Economy*, 102(1), 103–126.

Young, H. P. (1993): "The Evolution of Conventions," *Econometrica*, 61, 57–84.

A Proof of Theorem 2

Some helpful notation is first introduced. Telephone diffusion in a given island depends on how many chance adopters have chosen service up to time t, as well as their location on the island and the order in which they buy service. A configuration $\eta(t)$ of the social island is then defined by an n-tuple $(t_1, t_2, .., t_n)$ where $0 < t_i \leq t$ denotes the time at which the resident i of the social island has become a chance adopter and $t_i = 0$ indicates that no chance adoption has taken place yet. The set of configurations $S_\eta(t)$ is finite, and each configuration $\eta(t)$ occurs with some probability $p(\eta(t))$. The penetration rate in each island is denoted with $x(\eta(t))$ and the average rate $x(t) = \sum_{S_\eta(t)} x(\eta(t))$ is given by lemma 1. Due to the duplication effect telephone subscribers in minority islands might fall into two groups: a share $x_A(\eta(t), \mu)$ of residents continue to subscribe to the minority system while a share $x_B(\eta(t), \mu)$ of subscribers have defected to the majority system. Note, that $x_A(\eta(t), \mu)$ is increasing in μ and that $x(\eta(t)) = x_A(\eta(t), \mu) + x_B(\eta(t), \mu)$. The average penetration rates $x_A(t, \mu)$ and $x_B(t, \mu)$ of both systems within minority islands are defined as $x_A(t, \mu) = \sum_{S_\eta(t)} x_A(\eta(t), \mu)$ and $x_B(t, \mu) = \sum_{S_\eta(t)} x_B(\eta(t), \mu)$.

A business-type specific measure for the cost of incompatibility is first defined. It captures the *average* loss to type $\tilde{\beta}$ business customers in minority and majority social islands from not being able to communicate to some of their customers. The calculation is done using the assumption that a unified system would offer universal service at the same fixed access rate K under competition.[48]

$$C^b_{\tilde{\beta}}(t, \mu) = \sum_{S_\eta(t)} p(\eta(t)) \{\mu \min [K, (\gamma(1-\mu) x(t) + \gamma\mu x_B(t) \tag{20}$$

$$+ \quad (1-\gamma) x_B(\eta(t), \mu)) M\left(\tilde{\beta}\right) v_b\Big]\Big\} + (1-\mu)\min\left[K, \gamma\mu x_A(t, \mu) M\left(\tilde{\beta}\right) v_b\right]$$

This expression is greatly simplified if the duplication effect does not apply (i.e. $\mu > \mu^*$):

$$C^b_{\tilde{\beta}}(t, \mu) = \mu \min\left[K, \gamma(1-\mu) x(t) M\left(\tilde{\beta}\right) v_b\right] + (1-\mu)\min\left[K, \gamma\mu x(t) M\left(\tilde{\beta}\right) v_b\right] \tag{21}$$

The total cost of incompatibility $C^b(t, \mu)$ for all business users can then be calculated by summing over all business types (see equation 17).

[48]The first term captures the loss to the average minority business. The loss is at most K because the business can duplicate. If it does not do so it cannot communicate with outside customers in majority islands and with its inside customers who defected to the majority system. The second term captures the loss to the average majority business. Because residents in majority islands never defect the lack of interconnection affects these businesses only in their communication with those outside customers who are both located in minority islands and have not yet defected.

Assume that the duplication effect kicks in at time period $t_D(\mu)$ (possibly infinite). The function $t_D(\mu)$ is a step function with discontinuities at $\left\{\mu_1^D, \mu_2^D, ..\right\}$.

For $t > t_D$ the marginal duplicating business type in a minority island will be affected by the configuration $\eta(t)$ and the formulae 7 and 8 for the threshold size of the marginal duplicating minority and majority business have to be adapted accordingly. The true marginal duplicating business type in a minority island is denoted with $\hat{\beta}_A(\eta(t), \mu)$ which is greater or equal $\tilde{\beta}_A(t, \mu)$. Similarly, the true marginal duplicating business type in a majority island is denoted with $\hat{\beta}_B(t, \mu)$ which is less than or equal to $\tilde{\beta}_B(t, \mu)$.

The total cost of incompatibility $C^b(t, \mu)$ can be rewritten as

$$C^b(t, \mu) = \mu C_A(t, \mu) + (1 - \mu) C_B(t, \mu), \tag{22}$$

where $C_A(t, \mu)$ and $C_B(t, \mu)$ are the average costs of incompatibility to minority and majority businesses, respectively, and $D = \frac{nN}{2}$:

$$\begin{aligned} C_A(t, \mu) &= \sum_{S_\eta(t)} p(\eta(t)) \left\{ K \int_{\hat{\beta}_A(\eta(t), \mu)}^{\overline{\beta}} g\left(\tilde{\beta}\right) d\tilde{\beta} + G\left(\hat{\beta}_A(\eta(t), \mu)\right) D \times \right. \\ &\times \left. \left[\gamma(1 - \mu) x(t) + \gamma \mu x_B(t, \mu) + (1 - \gamma) x_B(\eta(t), \mu)\right] \right\} \\ C_B(t, \mu) &= K \int_{\hat{\beta}_B(t, \mu)}^{\overline{\beta}} g\left(\tilde{\beta}\right) d\tilde{\beta} + G\left(\hat{\beta}_B(t, \mu)\right) D \gamma \mu x_A(t, \mu) \end{aligned} \tag{23}$$

All functions are continuous and differentiable in the minority market share μ if it falls within the interval $\left(\mu_i^D, \mu_{i+1}^D\right)$. This allows the following derivation:

$$\begin{aligned} \frac{\partial C^b(t, \mu)}{\partial \mu} &= \left[C_A(t, \mu) - C_B(t, \mu)\right] + \gamma x_A(\mu) D \times \\ \times \sum_{S_\eta(t)} p(\eta(t)) &\left[(1 - \mu) G\left(\hat{\beta}_B(t, \mu)\right) - \mu G\left(\hat{\beta}_A(\eta(t), \mu)\right)\right] \end{aligned} \tag{24}$$

This expression is positive because $C_A(t, \mu) \geq C_B(t, \mu)$, $\mu < \frac{1}{2}$, and $\hat{\beta}_B(t, \mu) \geq \hat{\beta}_A(\eta(t), \mu)$.

The cost of incompatibility is discontinuous at μ_i^D: at $\mu = \mu_i^D$ the duplication effect sets in at time $t_D\left(\mu_i^D\right)$ while for $\mu = \mu_i^D + \epsilon$ the duplication effect sets in at a later time $t_D\left(\mu_{i+1}^D\right)$. The resulting 'jump' ΔC^b in the cost function $C^b(t, \mu)$ has to be shown positive.

The proof proceeds by making the jump continuous in the following way. Companies assume that their external environment is stochastic. With probability $1 - \theta$ the duplication effect sets in at time $t_D\left(\mu_i^D\right)$ and with probability θ it sets in at time $t_D\left(\mu_{i+1}^D\right)$. The resulting cost of incompatibility is denoted with $C^b(t, \theta)$. By varying the probability θ from 0 to 1 obtain:

$$\Delta C^b = C^b(t, 1) - C^b(t, 0) \tag{25}$$

The function $C^b(t,\theta)$ has to be increasing in θ. As before, decompose this cost into:

$$C^b(t,\mu,\theta) = \mu C_A(t,\mu,\theta) + (1-\mu) C_B(t,\mu,\theta), \tag{26}$$

Rewriting equations 23:

$$\begin{aligned} C_A(t,\mu,\theta) &= \sum_{S_\eta(t)} p(\eta(t)) \left\{ K \int_{\hat{\beta}_A(\eta(t),\mu,\theta)}^{\overline{\beta}} g\left(\tilde{\beta}\right) d\tilde{\beta} + G\left(\hat{\beta}_A(\eta(t),\mu,\theta)\right) D \times \right. \\ &\times \left. \left[\gamma(1-\mu) x(t) + \gamma\mu x_B(t,\mu,\theta) + (1-\gamma) x_B(\eta(t),\mu,\theta)\right]\right\} \\ C_B(t,\mu,\theta) &= K \int_{\hat{\beta}_B(t,\mu,\theta)}^{\overline{\beta}} g\left(\tilde{\beta}\right) d\tilde{\beta} + G\left(\hat{\beta}_B(t,\mu,\theta)\right) D\gamma\mu x_A(t,\mu,\theta) \end{aligned} \tag{27}$$

The marginal duplicating minority business $\hat{\beta}_A(\eta(t),\mu,\theta)$ is increasing in μ while the marginal duplicating majority business $\hat{\beta}_B(\eta(t),\mu,\theta)$ is decreasing. Note, that the cost function $C^b(t,\theta)$ is continuous and differentiable in θ. In order to obtain the derivative $C^b(t,\theta)$ with respect to θ, first note that

$$\frac{\partial x_B(\eta(t),\mu,\theta)}{\partial\theta} = x_B(\eta(t),\mu,1) - x_B(\eta(t),\mu,0) = \Delta x_B(\eta(t),\mu) < 0 \tag{28}$$

Similarly, obtain:

$$\frac{\partial x_B(t,\mu,\theta)}{\partial\theta} = x_B(t,\mu,1) - x_B(t,\mu,0) = \Delta x_B(t,\mu) < 0 \tag{29}$$

Further conveniently write:

$$\begin{aligned} \frac{\partial C^b(t,\mu,\theta)}{\partial\theta} &= \mu \sum_{S_\eta(t)} p(\eta(t)) G\left(\hat{\beta}_A(\eta(t),\mu,\theta)\right) D \times \\ &\times \left[\gamma\mu\Delta x_B(t,\mu) + (1-\gamma)\Delta x_B(\eta(t),\mu)\right] - \\ &- (1-\mu) G\left(\hat{\beta}_B(t,\mu,\theta)\right) D\gamma\mu\Delta x_B(t,\mu) \end{aligned} \tag{30}$$

Since $\hat{\beta}_A(\eta(t),\mu,\theta) \leq \tilde{\beta}_A(t,\mu)$, while $\hat{\beta}_B(\eta(t),\mu,\theta) \geq \tilde{\beta}_B(\eta(t),\mu)$, and $G()$ is an increasing function:

$$\begin{aligned} \frac{\partial C^b(t,\mu,\theta)}{\partial\theta} &\geq \mu \sum_{S_\eta(t)} p(\eta(t)) G\left(\tilde{\beta}_A(t,\mu)\right) D \times \\ &\times \left[\gamma\mu\Delta x_B(t,\mu) + (1-\gamma)\Delta x_B(\eta(t),\mu)\right] - \\ &- (1-\mu) G\left(\tilde{\beta}_B(t,\mu)\right) D\gamma\mu\Delta x_B(t,\mu) \\ &= \mu\Delta x_B(t,\mu) D \left[G\left(\tilde{\beta}_A(t,\mu)\right)(1-\gamma+\mu\gamma) - \right. \\ &- \left. G\left(\tilde{\beta}_B(t,\mu)\right)(1-\mu)\gamma\right] \end{aligned} \tag{31}$$

The duplication effect applies at some time $t_D(\mu) \leq t$. Furthermore, it is monotonic by assumption 1. Therefore, condition 13 is violated and $\frac{\partial C^b(t,\mu,\theta)}{\partial \theta} \geq 0$. This completes the proof of the first part of the theorem.

For the second part recall that the business-type specific cost of incompatibility for $\mu > \mu^*$ can be expressed as (see equation 21):

$$C^b_{\tilde{\beta}}(t,\mu) = \mu \min\left[K, \gamma(1-\mu)x(t)M\left(\tilde{\beta}\right)v_b\right] + (1-\mu)\min\left[K, \gamma\mu x(t)M\left(\tilde{\beta}\right)v_b\right]$$

This expression is clearly increasing in t. As the penetration rate $x(t)$ increases with t, three separate cases have to be examined. In the first case, neither majority nor minority businesses duplicate. The specific cost is increasing in μ because $\frac{\partial \mu(1-\mu)}{\partial \mu} > 0$. In the second case, only minority businesses duplicate. Therefore, $C^b_{\tilde{\beta}}(t,\mu) = \mu K + (1-\mu)\gamma\mu x(t)M\left(\tilde{\beta}\right)v_b$ which is clearly increasing in μ. In the third case, all businesses duplicate and $C^b_{\tilde{\beta}}(t,\mu) = K$ such that the claim is trivially fulfilled. Also note, that in the last two cases $C^b_{\tilde{\beta}}(t,\mu) \geq 2\mu K$ because $(1-\mu)\gamma x(t)M\left(\tilde{\beta}\right)v_b \geq K$. QED

B Construction of Business Duplication Rate in Table 3

The number of 'businesses' B_C in a given city is defined as 50 percent of all subscribers in that city in July, 1905. Unfortunately, a detailed breakdown of subscribers into residences and businesses is available only for certain cities such as Louisville (see table 2). However, for three out of the eight cities in the sample the share of business subscribers was between 40 and 50 percent in 1905 (see n. 13). The projected duplication rate amongst businesses is calculated by expressing the number of duplicates in all previous and later years as a share of B_C. This methodology allows me to compare duplication rates across time and cities by introducing a common reference point.

Some comments are necessary in order to justify this construction.

1. The vast majority of the duplicates were indeed business telephones. On the basis of its survey of 39 cities in 1905, the Merchants' Association of New York (1905, p. 3) found that almost 85 percent of dual subscribers were businesses. Most of the residential duplicates were in fact physicians and other professionals who would be classified as businesses in my model rather than as residents.[49]

2. Most subscribing businesses after 1905 were either residents or small-scale

[49] In Kansas City almost all residential duplicates were physicians' phones (Committee on Gas, Oil and Electric Light 1907, p. 194).

businesses with low communication demand (i.e. small $\tilde{\beta}$). Therefore B_C should include all large-scale and medium-scale businesses which subscribed early and received the bulk of all business communication from residents.

3. A 'business' in my model serves a particular social island - medium and small scale stores fit that description but not large-scale businesses. I will neglect the latter group in the data. Large-scale businesses had a duplication rate of close to 100 percent but made up only a small share of duplicates: in Louisville, for example, only 7.5 percent of dual users were large-scale businesses in 1910 (Mueller 1997, p. 82). As many of those establishments were up-stream suppliers of medium and small-scale businesses they did not exchange many messages with residents. Due to the extremely high rate of duplication they would not have influenced the network choice of residents in any case.

4. In the context of the model an increase of business duplication is causally linked to the state of development in the city (i.e. the telephone concentration $x(t)$). For this interpretation to be correct, other factors that can explain this phenomenon have to be excluded. In particular, a decrease of telephone rates over time would make duplication cheaper and the business duplication rate might increase even in the absence of network growth. Information on the rate schedules of the various companies for unlimited business and residential service is therefore provided in table 3. With the exception of Columbus, the cost of dual service in fact increased over time.[50] Therefore my business duplication rate is downward biased by not controlling for the subscription rate which works in favor of the model.

[50]Both Bell and Independent operators reported that the telephone industry exhibited decreasing returns to scale. Larger exchanges had to lay their cables underground as the poles could no longer support the increasing weight of the wires. The number of switchboards increased more than proportionally to the number of subscribers as a company had to provide for possible connections between *all* subscribers.