Essays on the Industrial Organization of Telecommunications and Network Industries

Thesis by

Alvaro S. González Staffa

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2000
(Submitted May 1, 2000)

Acknowledgements

Completing this dissertation would not have been possible without the help of my committee. I especially thank Simon Wilkie, my advisor, for the hours he spent encouraging new research ideas—he may be the best around at doing this. David Grether and Mattew Jackson, also members of my committee, had their patience tested time and again and passed with flying colors. Robert Sherman, while not an official member of my committee, is the most altruistic since he receives no official credit for advising me, yet he was often the unfortunate audience to expositions of some of my most embryonic (and silly) ideas.

Douglas Galbi, while not supporting my plan to return to graduate school, was my biggest backer once I was in the thick of it. I would have written a better dissertation had I taken all of his advice. Heidi King provided that intangible support that anyone who has taken on such a solitary and arduous task does not underestimate. Heidi's daughter, Kristin, kept my mind off Caltech (in a helpful way) with her silly poetry, bad jokes, and stories of wrestling matches with her hyper black mutt.

Finally, I thank my mother who was thousands of miles away, in Uruguay, though-out my years at Caltech. While physically apart, her encouragement was with me daily. She will probably not understand a word in these pages, but is likely to believe that they are the best words ever written on the subject just because her own son wrote them.

This document would not have been as pretty without the patient assistance of my friend, Dr. Monica Giannelli.

Abstract

This dissertation is comprised of three essays on the industrial organization of telecommunications and network industries. In the first essay on the economic history of competition between US international cable telegraph and radiotelegraph carriers, we find that between 1919 and 1934 there was accommodated entry, collusive price policies, and strategic investments. The dominant cable telegraph operator faced capacity constraints and higher operating costs than the entrant that used radio technology. Because the entrant had lower costs, the incumbent accommodated entry into the US-Europe telegraph market. Since the main international cable and radio telegraph operators benefited from collusive price policies, they agreed, sustained, and enforced these policies. Lastly, in order to sustain a binding collusive agreement with the dominant radiotelegraph operator, the dominant cable operator invested heavily in excess traffic capacity.

The second essay is an applied econometrics investigation on household choice of long-distance calling plans. We show that risk and uncertainty play significant roles in explaining household long-distance calling plan choices. High variability in calling patterns makes it difficult for callers to predict their future usage and therefore know which plan is cheapest across several months. Our findings show that there exists a bias against paying monthly fees for long-distance telephone service even when these fees are coupled with lower marginal rates. This bias against monthly fees is the opposite to the findings in several studies of consumer behavior in the local telephone service market.

Finally, in the applied theory essay on bargaining mechanisms and network interconnection, we examine pricing strategies of network monopolists seeking two-way interconnection. While firms have monopoly power over the node they control, they will set interconnection tariffs equal to marginal costs when connecting to the network of a symmetric, or nearly symmetric, firm. In the case where firms are asymmetric,

getting firms to set interconnection tariffs near marginal costs is a function of the bargaining protocol used to negotiate tariffs. We design a bargaining protocol that achieves weakly Pareto efficient and individually rational bargaining outcomes. In the empirical section of this essay, we find support for our hypotheses on the relationship between comparative firm size and efficient bargaining. Using interconnection tariff and traffic data for US international telecommunications carriers, we find the efficient bargaining frontier and characterized the firms with which US international telecom carriers are able to negotiate efficient interconnection tariffs.

Contents

List of Figures

List of Tables

Part I

Cable vs. Radio Technology in the International Telegraph Industry (1919-1934)

Chapter 1

Economic History

Summary

Competition between US international cable telegraph and radiotelegraph carriers between 1919 and 1934 is characterized by accommodated entry, collusive price policies, and strategic investments. Because in 1919 the established and dominant cable telegraph operator faced capacity constraints and higher operating costs than the entrant, the incumbent accommodated entry from a large radiotelegraph firm into the US-Europe telegraph market. Since the main international cable and radio telegraph operators benefited from collusive price policies, they agreed, sustained, and enforced these policies. Lastly, in order to sustain a binding collusive agreement with the dominant radiotelegraph operator, the dominant cable operator invested heavily in excess traffic capacity.

1.1 Introduction

L.B. Trioblet [30], in 1929, was the last to chronicle the history of international telecommunications in the US by examining transpacific communications. Since then, no other study on the history of US international telecommunications has been published. As a result of this paucity in historical research, the origins and nature of the firms and institutions that function in international telecommunications are not well known. This study fills the gap in historical research on the economic history of international telecommunications. While this paper covers the history transatlantic traffic, it is not simply a geographical extension of Tribolet's research. Economic, as well as historical, analyses are applied to the question of competition in the early international telecommunications industry.

This paper examines competition between US international cable telegraph and radiotelegraph carriers between 1919 and 1934. This period is of interest because of what took place in the competition over market share for transatlantic message traffic between US telegraph carriers. Issues of accommodated entry, collusive price strategies, and strategic investments characterize the story of international telegraph's early development in the US.

This study begins in the years 1919-20 when the Radio Corporation of America (RCA) entered the international telegraph business. RCA used wireless radiotelegraphy to transmit messages. RCA's entry presented the most formidable challenge to Western Union, the well-established and dominant cable overseas telegraph operator at the time. It is important to note that Western Union took a relatively accommodating stance to RCA's entry. At the time of RCA's entry, Western Union had built a nationwide network in the US and had a well-established foothold in the US-Europe telegraph market. Despite its financial, competitive and leadership position in the market, Western Union showed little signs of resistance to RCA's entrance. The reasons why Western Union accommodated RCA are addressed in this study.

The competitive relationship between RCA and Western Union, once RCA was established in the market, is of interest in this study. Western Union and RCA never changed the price paid for priority class telegraph messages between New York and London for twenty-three years (1923-1946). The price for priority messages remained stable despite the Great Depression and political turmoil in the US and Europe. One would expect that the Great Depression would drive down prices, but it did not. The possible explanations for stable prices between 1923 and 1934, eleven of the twenty-three years, are explored.

Finally, this study addresses excessive investments by Western Union in submarine cable to service the US-Europe link. In 1923, Western Union had capacity in excess of 40 percent of all worldwide cable message traffic. The following year, 1924, Western Union deployed a high-capacity cable and laid two more in 1926 and 1928. By the end of 1928, Western Union had more than tripled its capacity of 1923. If Western Union serviced world points serviced by all US cable companies existing in 1928, Western Union could have carried all of these firm's traffic, plus 45 percent more, on its own. This capacity can be viewed as excessive, wasteful and serving no purpose. On the other hand, these substantial investments in submarine cable could be viewed as strategic in Western Union's relationship with RCA and potential entrants in the international telegraph market. The nature and purpose of these investments are also addressed.

The passage of the Communications Act of 1934 marks the end of this study. This Act established the Federal Communications Commission (FCC) to enforce new regulations on the telecommunications industry. These new regulations changed the nature of competition and collusion in the international telegraph industry. The nature of these changes remains for another study to address.

1.1.1 Scope of the study

The data used in this study are taken from Congressional testimonies, statistical compilations of government-required industry reports, and special studies. Development and competition in the international telegraph industry are tracked by examining trends of three variables: International telegraph prices between New York and London; market share in the form of the number words sent and received between the US and Europe, and levels of investment made by telegraph carriers between 1919 and 1934. Unfortunately, there is less information about RCA's operations than Western Union's. More is known of Western Union's operations as a result of Congressional investigations and its annual reports to its stockholders. There are no annual reports that provide international telegraph traffic data available for RCA,[1] and since most Congressional hearings took place early in the inception of RCA, Congressional records

[1] RCA had made it a practice to report all of its earning in such a way as to make it impossible to separate its international telegraph business from its manufacturing and broadcasting enterprises.

are not as enlightening as they are about cable operations. As a result, much of what is discussed about RCA is inferred from documents, articles, and what was known then about domestic radiotelegraphy. Because there exist data limitations—either data are not available or they are incomplete for some of the variables this study tracks—some arguments are also supported by the use of theoretical economic models.

The market between the US and Europe was important to the early development of the overseas telegraph service market. More generally, Europe was the first continent in the development of an extensive international network of telegraph communications. Moreover, US firms extended their network to Europe first, and invested most heavily in that continent. By 1936, Europe generated nearly 89 percent of in and outbound traffic, measured in words sent, for Western Union. For RCA, Europe generated approximately 70 percent of all messages transmitted. Of all the European countries active in the development of international telegraph, the United Kingdom was the most important. By 1936, 14 cables, or 90 percent of all capacity for all transatlantic traffic was connected to England, and these cables represented well over 60 percent of worldwide message capacity.[2]

1.2 Development of US Domestic and International Telegraph Industry

The development of the international telegraph is a story based in the United Kingdom. The British had an early lead in both cable and radio telegraphy. British dominance in technology and substantial capital resources allowed it to develop its telegraph network worldwide, making it an essential node of any firm that considered entering the international telegraph market. On the domestic front, Western Union's early preeminence in the overseas telegraph business was carried over from its advantage in the domestic market. RCA, between 1919 and 1920, emerged with help from the Federal government and the know-how and resources of General Electric.

[2]See J.M. Herring and G.C. Gross (1936). *Telecommunications: Economics and Regulation*, pp. 26.

The Federal government's help in establishing RCA consisted of consolidating the relevant radio patents that made international radio transmission work and resolving outstanding legal disputes between patent holders. Westinghouse and General Electric held many important patents; without their assistance, RCA would not have been able to get off the ground.

1.2.1 The domestic telegraph

By 1880, Western Union emerged as the largest of 77 telegraph companies, most of them small, that reported to the Census. Western Union had built a coast-to-coast network, allowing it the unique capacity to make seamless, end-to-end domestic connections through the use of its own network. Western Union's main competitor, Postal Telegraph Company, was a distant second in size. Western Union delivered an average of four times as many messages as Postal throughout the 1920s. In 1938, after 50 years of domestic competition, Postal Telegraph held assets worth almost $71 million. In comparison, Western Union's assets totaled approximately $301 million in 1938.[3]

Seamless connections meant that Western Union did not need to rely on competitors to terminate messages, a problem that plagued smaller competitors. Western Union's competitors had to form alliances and cartels in order to furnish the same coast-to-coast service that Western Union provided. These alliances, however, meant that contracts were required to specify responsibility for the delivery of a message from one company's network as it was passed on to another company's network. Assigning responsibility for the timeliness, accuracy, or even final delivery or pick up of a telegraph message was oftentimes difficult [22].

After many years of trying, the courts did not satisfactorily settle the issues of liability involved in complicated interconnection regimes. Western Union found that the most satisfactory way to resolve the domestic interconnection problems was through consolidation. The issue of interconnection would remain important well into the era when RCA entered the market.

[3] US Senate (1941), *Study of the Telegraph Industry: Appendix to the Hearings before a Subcommittee of the Committee on Interstate Commerce.* 77^{th} Congress. S. Res. 95, pp. 242 and pp. 271.

Deploying submarine cables was a costly and risky venture. It took two failed attempts, and the loss of another cable, before the Anglo-American Telegraph Company became the first firm to successfully establish a permanent transatlantic link between Hearts Content, Newfoundland and Valentia, Ireland. These cables were owned and operated by a British company. These cables cost approximately $3,000,000 each, and $150,000 to deploy.

By the time that the British had nearly 15 years of experience in deploying submarine telegraph cable, US firms became involved in deploying transatlantic cables themselves. The first US-owned transatlantic cables connected Canso, Nova Scotia and Penzance, England. These cables were laid by the American Telegraph and Cable Company in 1881 and 1882. Western Union entered the international telegraph market when it leased the two submarine cables deployed by the American Telegraph and Cable Company. Both cables were leased to Western Union for fifty years, and those leases marked the entrance of the largest domestic telegraph carrier into international service. When Western Union entered the international telegraph market in 1881, there were several transatlantic cable companies in existence: the Anglo-American Telegraph Company (British) owned three cables in working order between Valentia, Ireland, and Hearts Content, Newfoundland; the Anglo-American Company also established a cable link between Brest, France, and Saint Pierre, Canada, terminating in Duxbury, Massachusetts, and sold it to the French; the Direct United States Cable Company (British) owned one cable between Ballinskellings Bay, Ireland, and Torbay, Nova Scotia, that terminated in Rye Beach, New Hampshire; la Compangnie Française du Télégraphe de Paris à New York (French) laid a cable between Brest and Saint Pierre and a cable between Saint Pierre and Cape Cod, Massachusetts, a cable between Saint Pierre and Louisburg. The French also owned certain land lines in the US and Canada, and had established a cable between Penzance, England, and Brest.[4] Table 1-1 provides a summary of the year of installation and which firm leased or owned the cable, and the cable's capacity measured in words per minute.

[4] *Cf.* Hearings on S. 6, 71^{st} Congress, 1^{st} Session, Part 8, pp. 501.

Points of connection	Year installed	Leased or owned by	Capacity (words/min)
Valentia, Ireland, to Heart's Content, Newf.	1873	Western Union	150
Sennen Cove, England, to Heart's Content, Newf.	1874	Western Union	150
Penzance, England, to Halifax, Nova Scotia	1874	Anglo-American	150
Brest, France, to St. Pierre Island, Canada	1879	French Cable	300
Valentia, Ireland, to Heart's Content, Newfound.	1880	Western Union	300
Sennen Cove, England, to Bay Roberts, Newf.	1881	Western Union	Approx. 500
Sennen Cove, England, to Bay Roberts, Newf.	1882	Western Union	Approx. 500
Waterville, Ireland ,to St. John's, Newfoundland	1884	Commercial Cable	Approx. 500
Waterville, Ireland, to St. John's, Newfoundland	1884	Commercial Cable	Approx. 500
Kerry County, Ireland, to Fayal, Azores	1888	Direct US Cable	Approx. 500
Kerry County, Ireland, to Heart's Content, Newf.	1888	Direct US Cable	Approx. 500
Valentia to Heart's Content, Newfoundland	1892	Western Union	Approx. 500
Valentia to Heart's Content, Newfoundland	1894	Western Union	Approx. 500
Waterville, Ireland, to Canso, Nova Scotia	1894	Commercial Cable	Approx. 500
Brest, France, to Cape Cod, Massachusetts	1898	French Cable	Approx. 500
Penzance, England, to Fayal, Azores	1900	Commercial Cable	Approx. 500
Fayal, Azores, to New York City	1900	French Cable	Approx. 500
Waterville, Fayal, Azores, Canso, Nova Scotia	1900-01	Commercial Cable	Approx. 500
Brest, France, to Fayal, Azores	1903	French Cable	Approx. 500
Fayal, Azores, to Halifax, Nova Scotia	1903	British Imperial	Approx. 500

Source: F.J. Brown (1930), *The Cable and Wireless Communications of the World: A survey of present day means of international communication by cable and wireless*, pp. 9-11.

Table 1-1: Telegraph Cables Between the US and Europe

At the outbreak of World War I, there were six transatlantic cable companies connecting Europe to the US. Only four telegraph companies transmitted messages through the 6 cable companies' cables. The four that transmitted messages were: the Commercial Cable Company (US) owned five cables; the Direct US Company (British) owned one cable; the French Company also had two cables; and Western Union leased five cables in 1911 from the Anglo-American Company (British) and also leased two cables from the American Telegraph and Cable Company

(US). The German Company owned and operated two cables, but these were cut and re-deployed to France and the UK at the beginning of the First World War and were permanently deployed to these nations under provisions of the Treaty of Versailles.[5]

1.2.2 Transatlantic radio

Before the US government intervened, the scattered possession of important patents prevented US development of radio technology.[6] The US Navy realized the military advantages of establishing a reliable international radio communications system during WWI, and prompted the consolidation of patents needed to establish the US radiotelegraph industry. The US Government also wanted American companies to own wireless telegraph circuits because cables cut during the war had caused undamaged cables to be overloaded. The government-led consolidation of patents was undertaken under the looming threat that the British would otherwise gain an international radio monopoly. So through the creation of the Radio Corporation of America in 1919, the US government promoted an international radiotelegraph industry independent of foreign technology and equipment.[7]

When RCA first entered the field of commercial transoceanic communications, it established offices in New York and later in Boston, Washington, and San Francisco. However, a large part of the total volume of US international communications traffic originated in or was destined to points in the United States other than those in which RCA had offices. To reach points where RCA had no offices, it had to either duplicate the offices of the telegraph companies already there, or negotiate contracts with them for the interchange of traffic.

Between 1920 and 1931, RCA negotiated a contract with the Postal Telegraph system to

[5] *Cf.* J.M. Herring, and G.C Cross (1936) *Telecommunications: Economics and Regulation*, pp. 130.

[6] The British also had a lead in the development of early international radio communications and facilities. However, US industries (Westinghouse, General Electric, and Bell Systems Laboratories) were not far behind the British in developments of radio technology. Together the inventions of De Forest of Westinghouse, Alexanderson of General Electric, and Colpitts of the Bell System solved the problem of efficiently generating and receiving continuous low power energy waves. This made transoceanic radiotelegraphy economical and more reliable.

[7] G.P. Oslin (1992) *The Story of Telecommunications*, pp. 174.

handle some of its transatlantic traffic. This contract was not a satisfactory arrangement from RCA's perspective because RCA suspected that Postal did not pass on all messages. Instead, Postal was suspected of passing on messages to the Commercial Cable Company, Postal's international telegraph ally and RCA's rival. Commercial would instead receive the revenue and business intended for RCA.

RCA's suspicions were not groundless. During the first nine years of its activity, RCA sent to Postal Telegraph ten times as many transatlantic messages for delivery at interior points as it received from Postal for transmission abroad. While RCA was at a disadvantage in collecting outbound message traffic to Europe because it did not have direct contact with customers in many cities, the number of messages coming into the US through RCA stations should have been approximately equal to the volume of traffic originating from the US. The only plausible way to explain the large disparity in traffic between RCA and Postal Telegraph is that Postal was diverting RCA's messages to Commercial.

Obtaining a domestic interconnection contract with Western Union for its transatlantic traffic was impossible for RCA until 1931. By denying RCA the benefits of a reliable domestic interconnection regime, Western Union hindered RCA's ability to compete for business in the European market. Until 1931, Western Union delivered RCA transatlantic messages to domestic points, charging RCA the highest tariff per word in Western Union's price schedule for the distance required to terminate the delivery. As for outgoing traffic, Western Union refused at all its offices in the United States to accept public messages intended to be sent abroad via the RCA's stations. The competitive position of RCA improved under contractual arrangements obtained in 1931 with Western Union.[8]

[8]Herring and Gross (1936), pp. 203.

1.3 Accommodated Entry

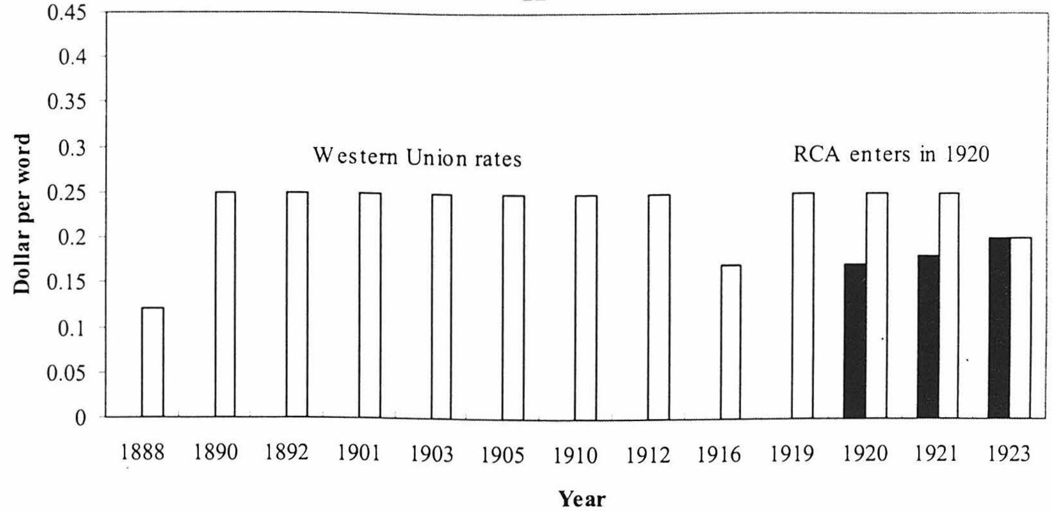
Having acquired a dominant position in the market for telegraph services between the US and Europe, Western Union was confronted with the question of how to best exploit this position for the benefit of its stockholders. This section focuses on Western Union's pricing strategy before and after RCA's entry.

In 1888, Western Union rebounded from a price war that was a result of the substantial deployment of cables by foreign and US companies to service US-European telegraph traffic (see Figure 1-1). Between 1890 and 1923,[9] Western Union chose the pricing strategy of a monopolist; it supplied the market up to its capacity, and left any residual demand for telegraph services to be supplied by its smaller rivals. This monopoly pricing strategy can be seen in Figure 1-1 where prices are relatively higher, from a historical perspective, and steady. Except for a short interruption during the First World War, Western Union charged 25 cents per word for messages transmitted between New York and London, the busiest link until 1923.

The success of the monopoly pricing policy depended on the rate at which rival firms expanded or entered into Western Union's market. The high profits Western Union made by holding prices high should have encouraged entry. However, entry into the international telegraph cable market was difficult. Entry required substantial financing for the deployment of cables, negotiating landing rights with a number of foreign governments, and obtaining cooperative agreements with domestic telegraph firms on both sides of the ocean to pick up and deliver telegraph messages from interior points in respective countries.

Firms planning to enter the international telegraph market had to be as big or bigger than

[9] Foreign telegraph carriers were *de facto* prevented from competing with US telegraph firms. In 1875, President Grant required that any government-owned or operated telegraph monopoly that wished to establish direct cable contact with the US must reciprocate, and allow US firms to connect a cable to shores of that foreign nation. This reciprocity condition stuck and was extended to toll rates; all tolls had to be equally divided by the telegraph carriers transmitting and receiving the message [24]. Furthermore, Western Union's and RCA's size compared to foreign operators, couple with the fact these carriers serviced the largest single market for telegraph service, put Western Union and RCA in the market leadership positions. As price leaders, Western Union and RCA set prices, and foreign carriers followed.



Source: Statistical Abstract of the United States, 1790-1970. Chapter R, Communications, Table 89-92.

Figure 1-1: Telegraph Rates for Priority Messages New York-London (1888-1923)

Western Union, or else they would be unable to exploit the same economies of scale advantage Western Union had already attained. However, no firm had the kind of capital to risk entrance into the international telegraph service market at such a large size. As a result, Western Union faced very little competition from entrants under monopoly pricing, and was able to sustain the high price without losing market share.

Given the formidable barriers to submarine cable technology presented, it is not surprising that when entrants appeared in the international telegraph market, they used radio technology to overcome these barriers. While operation costs were nearly the same between cable and radio technology, cable operators faced substantially higher start-up costs. A submarine cable cost nearly ten times more than a radio transmitter capable of sending and receiving nearly the same number of messages per minute. For example, cables that had a capacity of 300 words per minute were an estimated a $3,000,000 each. On the other hand, transatlantic telegraph radio stations cost, on average, $300,000 to $500,000 each and could carry 200 words per minute under optimal conditions.[10] Even when new cable technologies improved the carrying capacity of cables, the cost of investing in these new technologies increased as well. In 1880, the cost of transatlantic cable was approximately $3,500,000. In 1924, Western Union laid a new type of cable costing $4,800,000 and $150,000 to deploy.[11] While there were substantial gains in efficiency from deploying higher capacity cables, the large capital outlays required to deploy these cables represented a larger portion of operating expenses. Fixed costs determined more of the underlying costs of running a cable telegraph operation than the same costs did for radio.

As a result of the introduction of transoceanic radiotelegraph communication, RCA entered aggressively into the market with a low entry price for telegraph services and was able to sustain this pricing policy because of radio technology's lower fixed costs . In 1920, when radiotelegraph service was established between the US and England by RCA, RCA charged 17 cents per word; the established cable rate was 25 cents per word. The radiotelegraph rate was increased to 18 cents per word on January 1, 1921. On April 15, 1923, the radio rate was raised once again to 20 cents per word. Five days later, Western Union matched that price by lowering its rates by five cents per word (see Figure 1-1). This rate of twenty cents per word was to remain at that level for 23 years. When Western Union matched RCA's rates and sustained them at this level, it was a sign that this was a tariff level at which Western Union would accommodate the new entrant.

1.3.1 Comparing cost functions

In early cable telegraph history, cable operations were more labor intensive. Between 1870-1920, because of the long cables used in transoceanic transmission, many messages required

[10] Herring and Gross [16], pp. 122. Brown (1930: pp. 108) does not provide an exact figure, but under an agreement between nations aimed at coordinating their radiotelegraph transmissions requires that the use of the English beam system be used at a rate of no less than 100 five-letter words per minute. When General J.G. Harboard was President of RCA, he wrote, in a book published in 1928, that traffic speeds had been increased to 200 words per minute (*The Radio Industry: The story of its development*, Chapter II, pp. 91).

[11] The American entrepreneur, Jay Gould, paid the Siemen Brothers $7,000,000 to lay two Atlantic cables between Penzance and Canso, Nova Scotia, in 1881 and 1882 and formed the American Telegraph and Cable Company. [23], pp. 182 and 291.

manual re-transmission at some point along the way to amplify the current, increasing the clarity of the message. Technological developments in cable affected the labor composition of operations in the cable telegraph business. Between 1921 and 1924, regenerators sped up the operation of telegraph cable circuits and decreased the cost of operations by reducing the number of staff required to re-transmit massages. As a result, cable's labor costs were reduced substantially.

In Appendix 1.B, Tables 1.B-1 and 1.B-2, we provide a rough guide to the nature of RCA and Western Union's operating expenses. The expense line detailing the costs of conducting operations for each company shows that RCA had consistently higher expenses in this category than Western Union. While for Western Union conducting operations comprised approximately 60 to 70 percent of total expenses, the same expenditure for RCA is consistently around 85 percent or more of total operating costs.

Wages and salaries were the largest operating cost item for both cable and radio telegraph operators. The reduction of labor costs for cable operators was a substantial contribution to lowering cable's underlying costs. The trend in the reduction of labor costs shows up in the data provided in Table 1.B-3 in Appendix 1.B. In the last column of that table, we calculate the percent wages and salaries comprise of total operating costs. There is a definite downward trend in wage and salary costs, between 1919 and 1934, as a proportion of total operating costs. As a percent of total operating expenses, salaries and wages made up an average of 75 percent of total operating expenses around 1920. By the mid-1920s, the wages and salaries were approximately 56 percent of total operating expenses. A more prominent downward trend is clear in the number of employees hired,from 2,688 in 1919 to 1,894 in 1934—a period when the number of international messages sent and received was growing rapidly; .

Two striking differences are apparent in RCA's labor expenses; the number of employees stayed relatively the same throughout the period and labor costs consistently comprised around 80 to 85 percent (sometimes even 98 percent) of total operating expenses. Examine Tables 1.B-2 and 1.B-4. As we noted above, expenses incurred for conducting RCA's operations comprised approximately 85 percent of total costs of operations. In reviewing Table 1.B-4, we conclude that wages and salaries comprised nearly all of the expenses for conducting operations. Moreover, the proportion of total operating expenses made up of wage and salary costs were never substantially reduced throughout this period.

Labor cost savings enjoyed by Western Union were a direct result of acquiring new regenerator technologies that eliminated the need to re-transmit messages. The majority of personnel hired by RCA or Western Union ran the telegraph offices of these companies in same way. The cost of collection and delivery, no matter what the medium of transmission, was likely to be the same. Neither one of these two firms required anymore workers to take down incoming messages and deliver them, or send the messages to other points. However, cable technology moved much more quickly in eliminating the need for re-transmitting long distance messages than did radio technology (see Appendix 1.A).

Radio was less expensive to repair and maintain. As a proportion of total expenses, maintenance was smaller for radio than for cable. For RCA, maintenance expenses were approximately 4 to 7 percent of all expenses. Maintenance expenses for cable were 23 to 26 percent (see Tables 1.B-3 and 1.B-4 of Appendix 1.B). Fishing trawlers, anchors, chafing against rocks or ice floats, not use, were the principal culprits in damaging cables. Earthquakes or volcanic activity on the ocean floor, comparatively more rare events, were the only causes of damage to cables in the deep ocean. For radio, stress on the equipment from sending high voltage charges into the air caused most of the damage to these stations.

Cable held a slight cost advantage over radio in the rate of depreciation, but the lower levels of investment that cable companies required to operate negated any cost advantage that resulted from lower rates of depreciation . A small portion of annual expenses for cable operation can be attributed to depreciation. Cables were depreciated at a rate of 2.2 percent per year, assuming a 45 to 46 year life span.[12] However, some of the cables laid in the late 1870s were still in

[12]Depreciation for cables was definitively decided in the Treaty of Versailles where a series of experts agreed that the cables laid down before World War I would last 45 to 46 years, and could be depreciated at 2.2 percent per year.

operation by 1930, giving these cable life span of more than 50 years.[13] Radio's lower initial investment costs were partially offset by the shorter life of wireless stations, resulting in a higher rate of depreciation and replacement of radio stations. Depreciation was approximately 4 to 5 percent per annum. However, as we detailed above, a radio station that transmitted nearly as many messages per minute as one cable cost approximately ten times less. The higher cost in depreciation were swamped by the initial larger capital investments for cable operators.

Western Union's strategy after RCA entry was designed to minimize Western Union's loss of market share. The reasoning is the following: Suppose the rate of expansion of RCA, and the rate of entry from other radio telegraph operators, depended on the price that Western Union set for its telegraph service; the higher the price set by Western Union, the more rapid RCA's expansion and the higher the rate of entry by new firms. At a high price, RCA would earn higher profits than Western Union by the fact that it is the low cost producer. RCA would enjoy high growth, but would also be wary of incumbents entering the market and competing away RCA's high profits. RCA would therefore have a vested interest in setting up barriers to entry to deter new firms that could also exploit the cost advantages of radio technology.[14] Once RCA invested in barriers to entry, Western Union would also enjoy the benefits of these barriers. In the end, Western Union would do better implementing the strategies we discussed above than it otherwise would have if it had priced aggressively and ultimately lost a price war against RCA.

There is theoretical support for the arguments that Western Union acted rationally in leaving prices high. Dynamic limit pricing models, such as Gaskin's [11] or Kamien and Schwartz's

---

[13] *Ibid.*, Brown, p. 38.

[14] RCA did not have a legal monopoly of radio technology. Several domestic carriers exploited the business opportunities created by the development of international radio communications. Tropical Radio Telegraph and a number of subsidiaries of International Telephone and Telegraph Company (ITT) entered the radiotelegraphy business. During this early period, various carriers of telegraph communications began to compete with each other using high frequency radio signals and telegraph cables in serving the same overseas points. Moreover, some telegraph carriers added intermodal competition by using high frequency radio circuits as backup facilities. ITT became the most significant company to use both radio and cable facilities, as the result of the acquisitions of Mackay Radio and Telegraph, Commercial Cable Company, and All America Cables and Radio.

[18], substantiate Western Union's economic rationality in foregoing short-run profits for a share of the market in the future. Moreover, it is theoretically plausible that Western Union may have expected RCA to have a vested interest in keeping future entrants out of the market once RCA entered. Bernheim [6] was one of the first to suggest that entrants have an interest in keeping other firms out of the market once they become incumbents. Lastly, Gilbert and Vives [12] present a model in which firms will invest in entry deterrence, though it will also benefit other incumbents that do not invest in deterrence. In the Gilbert and Vives model, firms earn private benefits to their investments since the level of investments will ultimately decide market shares.

Western Union, after the First World War, experienced serious capacity constraints and would not have been able to carry the extra demand for services that would take place when tariffs fall in a price war. Throughout the period of the War, it was too risky to navigate through international waters to repair or replace damaged cables. Because of binding capacity constraints, Western Union could not credibly continue with its short price war against RCA—the price war would have lead to a continued increased demand for cable telegraph services that could not be met. RCA's entry came at a time when Western Union was capacity constrained and therefore least prepared to fight it. At the same time, RCA was unable to carry all of cable's traffic. So RCA could not have carried out its price war much further either.[15]

With respect to RCA's pricing strategy, it is plausible that RCA could have undercut Western Union's tariff in hopes of driving the dominant cable operator out of business. One would expect an aggressive pricing strategy from the firm that holds the cost advantage. As we stated above, when RCA first entered the market, it did price its services lower than the established price, but RCA did maintain this policy, deciding not to fully exploit its cost advantages through aggressive pricing. In April 20, 1923, in response to radio's lower rates, Western Union lead all cable companies by matching RCA's rate, and charged 20 cents per

[15] U.S. Senate (1945) *Hearings before a Subcommittee of the Committee on Interstate Commerce*, 76^{th} Congress, S. Res. 95, Part 1, pp. 472.

word. This rate was to remain unchanged until 1923.[16] (See Figure 1-1.) The reasons why RCA did not undercut prices are addressed in the following section where the benefits of collusion are examined.

1.4 Collusion

The invariability in price for telegraph services between the US and Europe, for the period between 1923 through 1946,[17] is worthy of attention. Under strict government regulation, this invariability in prices could be easily explained away as public sector tinkering. However, international telegraph price stability existed between 1923 and 1934, a period before the Federal Communications Commission and international price regulation was established.[18] Price stability took place in the midst of a worldwide economic depression and two World Wars.[19] The Great Depression surely presented circumstances under which economic conditions should have led to changes in the pricing policies in the overseas telegraph market. The fact that the Depression did not have this effect implies that the underlying macroeconomic conditions had

[16]When direct radio service was established in Norway, Germany, and France, in 1920, the radio rates were generally lower than the rates cable had been charging. To Norway, the radio rate was fixed at 24 cents per word, in 1920, compared to the cable rate of 35 cents per word. The cable and radio rate to Germany was fixed at 36 cents per word in 1920; but was later reduced to 25 for radio and 30 cents for cable. To France, the radio tariff was established at 20 cents per word, whereas the cable tariff had been 25 cents per word for many years. In April 1923, the radio toll was increased to 22 cents per word, and the cable toll was reduced to that same figure. In 1923, radiotelegraphy was established in Italy. The initial price per word between the US and Italy was 26 and the cable price was 31 cents per word. These rates to Italy remained the same throughout the 1920s and 1930s.

[17]See U.S. Bureau of the Census, *The historical statistics of the United States, Colonial Times to 1970.* Chapter R, Communications. GPO: Washington, DC (1975).

[18]Soon after 1934, price collusion became FCC policy. In testimony to the Senate Subcommittee on Interstate Commerce, the Commissioner of the Federal Commutations Commission, Paul A. Porter testified to the fact that the Commission regulated rates so that cable operators set the price and radio operators agreed to match that price and not go any lower (U.S. Senate (1945) *Hearings before a Subcommittee of the Committee on Interstate Commerce*, 79^{th} Congress, S. Res. 187, Part 1, pp. 140).

[19]It is fair to assert that since the price of an international telegraph message had remained the same throughout this period, the price also reflects the charge by the foreign carrier to the US carrier for terminating message, then foreign carriers were also colluding to keep prices high. That may indeed have been the case. We do not address the issue. In either case, whether foreign carriers were or were not colluding with US telegraph carriers to keep prices high, the argument that we present is not affected.

little to do with Western Union's and RCA's pricing policies.

Ironically, turbulent times economic and political times may have in fact aided price stability according to Green and Porter [13]. According to the theory on collusion proposed by Green and Porter, if firms have no perfect way to coordinate the prices they should charge for its services, colluding firms may not be able to perfectly detect whether the other firm is holding up their end of the bargain under the noise of turbulent economic conditions. Under a non-collusive environment where firms are trying to adapt to turbulent economic conditions, we expect firms to change their pricing policies often. However, the price of adaptation in a collusive environment is to confuse your accomplice in collusion, and possibly leading the other firm to believe that your adaptations are actually attempts to undercut prices and rob it of market share. Rather than risk initiating a price war by changing prices in response to changing market conditions and confusing firms participating in the collusive agreement, holding public prices steady may be the only way colluding firms can signal to each other that the deal is being upheld. Firms colluding are caught between the rock of avoiding unilaterally raising their prices because doing so will concede the complete market to the firm that maintained a lower price, and the hard place of not lowering prices because it risks triggering a price war as the other firm responds with its own, deeper, price cuts. In the end, firms that cannot perfectly or credibly communicate their intentions on their price policy will maintain a stable price as a way of avoiding either of the two outcomes that result from changing prices. Paradoxically, they may be more stubborn in maintaining price stability when the market is most turbulent. Below, we will examine Western Union's and RCA's respective market shares and speculate that stability in these allowed these two firms to believe that both parties were upholding the implicit collusive bargain and that no secret deals with big consumers of telegraph services were taking place.

In a situation when firms cannot signal coordinate prices because the economic environment is too noisy for them to send perfectly detectable signals, it is likely that firms will monitor each other by tracking their share of the market. We review Figure 1-2 to determine what was happening to market share. It is clear that while Western Union was growing, RCA was growing faster and increasing its share of the international telegraph service market between the US and Europe. The only time when Western Union experienced a shrinking of its market, it could have attributed this loss to the economic conditions of the Great Depression. However, the continual loss of market share could be expected to have triggered a response from Western Union. We will argue below that particular circumstance in which WU found itself after World War I made it nearly impossible to retaliate against RCA and gain back its share pre-RCA entry. On the contrary, WU's strategy was to accommodate RCA and coordinate its pricing strategy with its main radiotelegraph competitor.

Observers of the international telegraph industry assumed that price collusion might take place. Collusion would explain price stability. In the first edition of his book, F.J. Brown [8], a prominent British telegraph expert,[20] expressed the following opinion on the subject of collusion:

> A war of rates is probably the last thing that either side (radio and cable) desires, and—as responsible representatives on each side have frequently suggested—the sensible thing seems to be a mutual arrangement satisfactory to the competitors and not inimical to the interests of the public. (pp. 76: 2nd Ed.) [Parentheses mine]

How did collusion start? In order to sustain a collusive agreement, firms must resolve many problems. First, an agreement must be reached. Second, there must be a method of detecting cheating. As soon as the price is raised above the noncooperative level, firms have an incentive to cheat on the collusive arrangement. Finally, once cheating is detected, a punishment must be meted out. Was the international telegraph environment conducive to solving the problems of reaching an agreement on and sustaining price collusion? And could both firms, Western Union and RCA, benefit from it? To both questions, the answer is yes.

[20] F.J. Brown was formerly Assistant Secretary of the British Post Office in charge of Cable and Wireless and Director of the International Cable Companies' Association when he wrote that passage.

RCA and Western Union had common interest in obtaining and sustaining a collusive agreement. For Western Union, a collusive agreement implied that RCA, the low-cost producer, would not reduce tariffs that could result in losses for Western Union. For RCA, a collusive agreement allowed it to expand its network, and guarantee more cooperation from Western Union in pick up and delivery of overseas messages destined to, or originating from, points in the parts of the US where RCA did not have offices.

Cheating was easily detectable by both firms. The conditions for detection were present in the international telegraph service market: many buyers, the availability of information about each firm's behavior, publicly posted prices, and multiple interactions between Western Union and RCA.

Lastly, both RCA and Western Union had the ability to punish each other. RCA maintained its cost advantage, even after Western Union applied new labor saving technologies in the mid-1920s. With this cost advantage, Western Union was always under the threat that RCA could lower tariffs to unprofitable levels if Western Union was caught cheating. On the other hand, Western Union handled RCA's transpacific traffic to and from interior points in the US. Western Union traded goodwill in the delivery and pick up of RCA's transpacific traffic for RCA's compliance in the collusive agreement over transatlantic tariffs. Goodwill was required since the courts inability to reliably assign blame to the party that bungled the delivery of a message across networks was well known [22]. Western Union knew that it could lose its leverage against RCA if RCA sought other partners to deliver its messages domestically or if RCA established its own domestic network. As a result, Western Union sought a more permanent threat against RCA through investments in excess capacity. We discuss how excess capacity investments function as punishment threats in Section 1.5.

1.4.1 Agreement on collusion

For RCA and Western Union to agree on colluding, they must have shared an interest in cooperating. The benefits of collusion outweighed the benefits of either competitive pricing, or pricing below a competitor's cost with the goal of driving that competitor out of business. The latter strategy, predatory pricing, is most commonly reserved for the low-cost producer. In the case of the overseas telegraph service market, the low-cost producer was RCA. One would expect RCA to implement a predatory pricing policy.

RCA was a good candidate to win a price war against Western Union; as the low-cost producer, if RCA charged for its services below its own costs, its losses would have been lower than Western Union's. Furthermore, RCA had tremendous financial resources to withstand a fight that entailed losses. The well-heeled General Electric Company held interests in RCA and had several cross-licensing agreements with it as well. General Electric would have likely provided financial resources to RCA in order to protect these licensing agreements and other interests. The extent of RCA's access to financial backing became apparent when it was able to purchase the assets and patent rights of the American Marconi Company only soon after RCA's inception. In the first full year of RCA's operation, it established itself in the most important markets; England, Hawaii, Japan, Norway, Germany, and France.

Throughout the same period, Western Union's fortunes were not as good. It was well known that Western Union suffered from substantial capacity constraints because several of its transatlantic were damaged and temporarily inaccessible for repair. Western Union was unable to navigate, throughout the First World War, the seas where its cables lay because these waters were often occupied by enemy ships. Throughout WWI, Western Union's cables were suffering damage beyond normal wear and tear since enemies of the US often sabotaged submarine cables as a war strategy. By the time the war ended and RCA appeared on the scene, Western Union spent several years sending messages through a limited number of cables. As a result of these circumstances, Western Union was in no shape to fight RCA's entry. If Western Union attempted to stop entry by lowering its prices, it was likely to have generated more telegraph message demand than WU's cables could transmit and receive.[21] RCA did not

[21] US Senate (1946). *Hearings before a Subcommittee of the Committee on Interstate Commerce*, 79^{th} Congress, S. Res. 187, Part 2, pp. 249.

face such binding capacity constraints. Radio technology allowed RCA to increase capacity at a relatively faster pace than it would have taken Western Union to repair and deploy more cables.

As long as RCA was growing, RCA saw no reason to commence a price war. RCA's market expanded under competition with Western Union. Figure 1-2 shows the extent of RCA's growth. Within three years of operation, RCA had a foothold in all of the important European markets, and one Asian one, too—a feat that Western Union could not match. At the same time, radiotelegraph operators increased their share of worldwide message traffic from nothing to over 44 percent of the market between 1920-1936, and RCA carried nearly 80 percent of all radiotelegraph traffic. By the end of 1934, the Radio Corporation consisted of 57 circuits linking the United States and 47 countries together on its radio network. The benefits of an accommodating pricing policy with Western Union were positive for RCA.[22]

In any case, RCA may not have wanted a price war because the costs of fighting Western Union were likely to be steep. Western Union ran a profitable domestic telegraph network, as well as an international one, and had substantial financial resources. In 1926, Western Union's assets for its domestic and international network were worth nearly $300 million, $351 million in 1930, $346 million in 1934, and $302 million in 1938.[23] By comparison, RCA's asset holdings were worth just over $19 million in 1938. With this disparity in size and resources, Western Union was not an opponent that could be easily disposed of; RCA could have expected a long and costly price war.

RCA was also not completely free from Western Union's retaliation if it pursued a predatory pricing policy. Western Union handled RCA's transpacific traffic destined to, or originating from, interior points in the US. Beginning in 1931, Western Union also handled RCA's transatlantic traffic. While Western Union was legally bound to carry out the letter of the contracts that RCA and Western Union had negotiated with respect to message pick up and delivery,

---

[22] Herring and Gross (1936) pp. 85.

[23] US Senate (1941). *Study of the Telegraph Industry: appendix to the Hearings before a Subcommittee of the Committee on Interstate Commerce.* 77th Congress. S. Res. 95, pp. 257 and pp. 271.

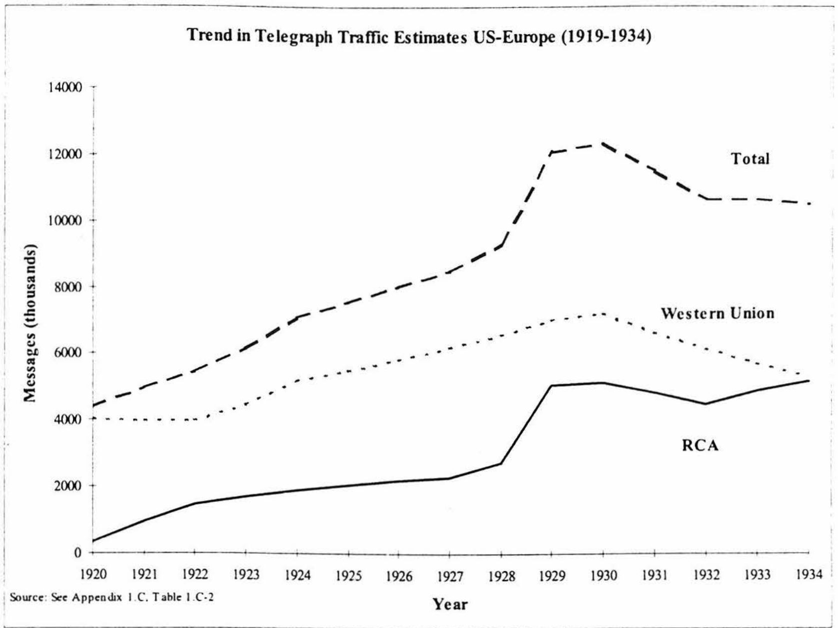


Figure 1-2: Western Union, RCA Market Shares for European Traffic

these interconnection contracts were difficult and costly to monitor. Nonnenmacher [22] writes extensively on the legal problems that domestic telegraph companies had with their interconnection regimes. Legal problems meant that it was difficult to prove when Western Union was negligent in passing on RCA's messages. In order to rely on Western Union's service, RCA would have to depend on Western Union's goodwill. In turn, Western Union's goodwill was likely to be related to RCA's behavior in the competition for transatlantic telegraph traffic.

No collusive agreement between Western Union and RCA could be made public and enforced legally. In fact, there were severe legal sanctions to be had if the US Department of Justice detected RCA, or Western Union, engaging in predatory pricing or other anticompetitive behavior. Just five years before RCA emerged as an international telegraph carrier, the telecommunications industry experienced its own enforcement of the Sherman Act when AT&T

was forced to sell its holdings of Western Union Telegraph Company stock to comply with the provisions of the Act. This enforcement of the Act against AT&T served notice to the telecommunications industry that their behavior was being watched. Therefore, any coordination in pricing between WU and RCA would have to be done in such a way as to avoid government detection.

1.4.2 Detecting cheating

The mere fact that price is above the noncooperative level and that the marginal revenue for a firm is greater than its marginal cost gives each firm an incentive to cheat. The incentive to cheat and garner higher payoffs is most alluring when the cheater cannot be caught by the competing firm. In an environment where catching a cheater is hard, the collusive agreement falls apart. On the other hand, if detection is relatively easy, and not costly, then a collusive agreement can be sustained. Western Union and RCA competed in a market where detection was relatively easy and inexpensive but not perfect.

Collusion is more successful when there are many customers. If the number of bargains that undercut the rival's price is independent of the number of deals previously cut by the cheater, then the number beneficial arrangements will be a function of the number of customers there are in the market to whom bargains are offered. The higher the number of customers, the more bargains offered. The probability of detecting a cheat in price will depend on the probability that the other firm will run into a customer that has received a price cut from the other firm. In the international telegraph service, under the type of message class in which collusion was to have taken place, there are many buyers for this service. The number of messages transmitted in this class of service is in the millions, and the customers who demand these messages are numerous.

When price information is well known, collusion is facilitated. The transaction for the delivery of a priority telegraph message from the US to Europe was initiated by a customer's visit to a telegraph office during business hours. The transaction between customer and telegraph operator took place in the office, and the tariff schedule was known. Prices were publicly posted, and price discrimination would have been easily detected.

Output sold was nearly public information. The volume that radio operators transmitted through their stations was easily ascertained. Radio signals were dispersed in a broad direction where anyone could determine the number of words and messages being transmitted at any time by picking up those signals. In the case of cable, capacity of each cable was also well known, and how many cables each company owned or leased was also part of the public record. Indeed, much of the data obtained from this study on cable capacity and traffic volumes was obtained from Senate hearing testimony and records available to all. It is hard to see why Western Union and RCA would have had to expend tremendous effort or resources in determining each other's market share. Secret deals with large customers, the only ones to whom it is worth providing such deals. would have been detected in increase activity on the cables or in radio waves. In fact, Western Union and RCA shared message re-transmission facility in the Azores.[24] In this building, sat the employees of both companies. These employees regenerated their cable or radio signals to assure that their respective stations would receive the messages clearly. All messages sent or received by these companies, between Europe and the US, passed through this shared facility. Any increased activity resulting from private price cuts would have been detected by the increased number of messages the employees of the company that provided such a deal would be re-transmitting .

Lastly, and most importantly, Western Union and RCA competed in the same market for many years. Multiple and repeated interactions allowed Western Union and RCA to condition their behavior on the past actions that the other took. By doing so, these firms induced cooperative, collusive outcomes. In other words, through multiple interactions, RCA and Western Union were given plenty of opportunities to monitor each other's actions with little cost to them.

[24]See Coggeshall, I.S., *An annotated history of submarine cables and overseas radiotelegraphs: 1851-1934*, Ed. D. De Cogan, School of Informatuion Systems (UEA) (1993).

1.4.3 Punishments

Detection of cheating on the collusive agreement was inexpensive and relatively easy for RCA and Western Union. However, detection is of little value to the offended party if there is no means of retaliation against the cheater. For RCA, its advantage as the low cost producer allowed it to credibly threaten and implement a strategy of price undercutting if Western Union were to cheat on the agreement. Western Union's punishment options were not so straightforward.

All-out competition was likely to be RCA's punishment for Western Union's cheating. RCA would have felt less pain from this punishment strategy. If RCA believed that because of its cost advantage Western Union would make greater losses when RCA priced its services at marginal cost, then lowering its prices was a viable and credible punishment strategy.

In comparison, Western Union held leverage over RCA. While Western Union could affect the quality of RCA's transpacific traffic, this leverage may have been ephemeral. RCA had options that allowed it to substitute Western Union's domestic network, eliminating its dependence on Western Union. For example, Postal Telegraph, Western Union's rival, also carried some of RCA's traffic. While Postal was not as big as Western Union, and therefore did not reach as many points, RCA could have bought Postal and expanded Postal's domestic reach. Deploying its own domestic network was another alternative for RCA. Either one of these strategies would have replaced Western Union's domestic network, and leave RCA free to pursue a more aggressive pricing strategy against Western Union.

In summary, while RCA benefited from a permanent cost advantage with respect to cable, and this advantage was RCA's means of punishing Western Union, Western Union did not profit from any such permanent leverage over RCA. The only source of Western Union's leverage, the handling of RCA's international traffic domestically, depended on RCA not building or buying another domestic network. The possibility that RCA would replace Western Union's domestic connections exposed Western Union to the risk that RCA would at one time shirk on its commitment to the collusive agreement. If RCA acted independently of the agreement, this almost certainly meant that Western Union would lose the European market at one point or another. Since RCA could not be trusted to sustain the collusive agreement indefinitely without a credible punishment, Western Union employed a strategy of excessive investments in submarine cable.

1.5 Excess Capacity and Collusion

Western Union could have used strategic investments to overcome its weaker negotiating position against RCA. These strategic investments were in the form of deployment of cable in excess of expected traffic demand. On the other hand, one could argue that capacity investments were not strategic but based on managers' optimism about Western Union's future. WU managers invested heavily based on expected bullish growth in the demand for telegraph traffic. We test several scenarios of growth for demand to see if a rosy outlook of future demand alone justifies the investments in cable capacity that Western Union made. We find that capacity calculations show that even under the most optimistic scenario, Western Union's capacity was in excess of expected demand.

The question of whether it is economically rational to invest in excess capacity is addressed by examining models of dynamic duopolistic competition. While it has been well established theoretically that investing in excess capacity is a viable competitive strategy in a duopolistic market in general [26], [9], the viability of this strategy also holds in a market characterized by capacity constraints [19], [5]. In an industry characterized by capacity constraints, such as the telegraph industry, excess capacity is used as a discipline device that avoids competitive pricing between firms. In a dynamic version of a capacity constrained duopoly competition game, a collusive outcome is attained.

1.5.1 Theory of investments in excess capacity

The asymmetry in cost structures exhibited by RCA and Western Union makes it harder to sustain a collusive agreement between the two firms. RCA could have cut prices to a level that would have lead to losses for Western Union. The prospect that RCA would cut prices and force losses on Western Union lead to strategic investments by Western Union. By making strategic investments in excess capacity, Western Union hoped to respond, in kind, to RCA's price cuts and cut prices low enough to generate losses for RCA as well. However, cutting prices that low would generate tremendous demand for telegraph traffic, assuming reasonably high price elasticities. Therefore, WU had to convince RCA that it could carry as much traffic as would be generated by lower price levels or otherwise its threats of price cuts were not credible. Western Union's ability to inflict losses on RCA was based on carrying as much of the additional traffic the price war generated.

Western Union's traffic capacity was well known, and so the potential damage that Western Union could inflict was easily ascertained. The transparency of its capacity was both good and bad for Western Union. It was bad because it could not bluff; Western Union could not lie about its cable capacity. It had to make the actual investments in order to be credible. But it was good because RCA knew that Western Union knew that it could not bluff. Western Union's inability to bluff implied that its actions, without doubt, signaled its intentions.

The sunk cost nature of submarine cable deployment also helped Western Union send a clear message that a price war was equivalent to mutually assured destruction. Ocean cable investments were made, they were largely sunk costs. Once deployed, cables could not be easily re-deployed for other uses. Sunk costs in cables were a barrier to Western Union's exit out of the international market. Because Western Union was unable to leave the market without heavy losses, investing in cables signaled to RCA that flight was not a possible outcome of a price war. Western Union would have to fight in a price war since its investments were worthless for other uses.

When Western Union made sizable investments in higher capacity and more efficient cables, these investments sent three signals to RCA: 1. Western Union was making a commitment to stay in the market; 2. Western Union was becoming more efficient, and the losses that RCA could inflict on it with a price cuts were falling with each new investment; and 3. All traffic generated by lower prices would have to shared because Western Union's capacity was able to take the new load. The second signal of Western Union's commitment to cost cutting was also reinforced when it invested in cable regenerators, and other cost cutting technologies. Western Union's commitments, in the form of investments, were designed to signal to RCA that Western Union had no other option than to carry through with competitive pricing, or even a price war, if RCA reneged on the collusive agreement.

Benoit and Krishna [5] (hereafter, BK) addressed the strategic use of investments in a dynamic duopoly model. In their model, pricing strategies employed by firms avoid the competitive outcome—marginal cost pricing. In a dynamic setting of duopolisitc competition, where firms compete repeatedly on price, but make an initial capacity decision, the resulting prices are above the Cournot equilibrium and collusive. With the BK model, it can be shown that firms benefit from installing capacities that they hold in reserve only to punish other firms that become more aggressive in the course of deciding their pricing strategy. The larger the excess capacity a firm holds in reserve, the more severe the punishment that the firm can inflict on the other and therefore the more likely it is to obtain cooperation and reach agreement on a collusive pact.

The BK model postulates that if quantity (capacity) choices are relatively inflexible, firms generally carry excess capacity in equilibrium to ensure that collusive behavior is sustained. Collusion, therefore, is not costless to maintain. In this model, it is costly to make threats credible—the costs of carrying excess capacity. This is exactly the hypothesis that was introduced to explain over capacity in transmission and receiving capacity in the international telegraph industry.

The BK is an extension of the Kreps-Scheinkman [19], so it is instructive to review some of their results. Kreps-Scheinkman introduce a model of Bertrand competition. They incorporate capacity constraints in their two-period model. In the first period, firms choose quantity produced (accumulate capacity), and in the second period, the quantities are fixed and firms choose prices. This model shows that the Cournot outcome holds for capacity-constrained firms competing against each other. Moreover, the equilibrium profits of each firm is affected by the capacity of the competing firm. This equilibrium statement is represented below.

In the game g, there is a unique equilibrium outcome $x_1 = x_2 = x^*(b)$ (the Cournot quantity equilibrium) and $p_1 = p_2 = P(2x*(b))$. Where $x_i, i = 1, 2$, is the firm i/s output, and p_i is the same firm's price. Let $e_i(x_1, x_2)$ be the equilibrium payoffs of firm i in the $g(x_1, x_2)$. Then, $x_2' \geq x_2$ implies that $e_1(x_1, x_2') \leq e_1(x_1, x_2)$ and symmetrically for firm 2. This last statement shows that each firms equilibrium revenues decline as the other's firm capacity level increases. If a firm expands its capacity and produces with it, the Cournot result shows that this capacity can be used to punish an errant opponent.

BK's contribution to Kreps and Scheinkman was to extend the number of periods of play where firms first choose capacities, and then choose prices, repeatedly. The extension to repeated play provides analysis of long-run competition. In long-run competition, in order to attain a perfect equilibrium, it must be assumed that firms do not discount the future too highly. In other words, if interest rates were high in present play, a firm may choose to cheat, put the extra earnings that it obtained through cheating in the bank, and take its punishments in the future, once its cheating is detected and the punishments applied. The interest received from the bank on the earnings from cheating multiplied by the probability of getting caught cheating would have to be higher than the discounted stream of payoffs a firm would receive from cooperating for a firm to be induced to cheat. By discounting the future highly, a firm is willing to forego payoffs then to achieve higher payoffs now.

Assuming that firms do not discount the future highly, $\delta \geq \frac{1}{2}$, where δ is the common discount factor, then there exists a perfect equilibrium path in the infinite horizon game, G, specified by BK. The result of the game, G, except for those equilibria in which firms choose Cournot prices and quantities, are equilibria in which firms choose capacity in excess of what they need to produce in equilibrium, and that both firms charge the same price. The result of a similar price, throughout the competition game, is the collusive behavior described in this paper. Holding of excess capacity provided a threat against RCA. The excess capacity established by Western Union throughout the 1920s was a signal to RCA that it was able to handle the extra load of messages should RCA want to enter a price war and expand supply.

Several aspects of the BK model are applicable to the situation described in the competition game between RCA and Western Union. First, firms are allowed to choose both quantities (capacity) and prices. Capacity related decisions, as is the case in the telegraph industry, are long-run related decisions. It is unreasonable to assume that either Western Union, or RCA, could change their capacity easily. On the other hand, pricing decisions could change in the short run, as the BK model allows.

Competition games that involve investments strategies must be modeled in a dynamic setting. The pricing repeated game employed by BK provide two analytical benefits in describing competition between RCA and Western Union: more realistic results from modeling competition in a dynamic setting—static models of duopoly poorly describe intertemporal competition; and the excess capacity results that stem from the ability of firms to change the rules of the games through their own investment behavior. The dynamic setting describes the multiple interactions environment in which RCA and Western Union competed. In the case of this model, as well as empirical observation, the repeated nature of these interactions lead to collusion.

1.5.2 Western Union's excess capacity

Theory supports the conjecture that it is economically rational for Western Union to have invested in excess capacity in competition for the overseas telegraph service market with RCA. This section presents empirical evidence that also supports the same conjecture. We provide estimates of Western Union's over capacity. These estimates show that Western Union did invest in excess of the most optimistic demand scenario, supporting the strategic investments hypothesis.

In 1882, Western Union leased ocean cable lines from the American Telegraph and Cable Company for fifty years at a price of 5 percent of $14,000,000, and by 1911, all of the Anglo-American cables had been leased by Western Union for ninety-nine years. Through leasing, Western Union had three cables with a 300 per minute word capacity, and four more with approximately a 500 word per minute capacity (see Table 1-1). The oldest of these cables was the one laid in 1873, and the newest was the one laid twenty-one years later, in 1894. Under a conservative estimate that these cables last 45 years, these cables should be expected to have been replaced between 1918 and 1933. As previously mentioned, the First World War made it difficult to deploy cables since the seas were potentially dangerous even for ships suited for naval battle.

In 1924, Western Union laid a new type of cable from New York to Horta. It carried 1,500 letters a minute, compared with 150 to 500 letters of earlier cables.[25] The cable cost approximately $2,000 per mile in 1924. The total cost of the cable was $7,000,000, plus the cost of laying down that cable was $150,000.[26] In 1926, and again in 1928, Western Union deployed new cables across the Atlantic. These cables were an improvement to the original permalloy cable laid in 1924 in that they had a capacity of 2,400 letters per minute. The total cost of all three cables was approximately $25,000,000.[27]

Because European message traffic figures are not available, worldwide cable traffic figures are used to estimate Western Union's expected demand. By using worldwide figures, instead of estimates of European traffic flows, the calculations of Western Union's capacity will be biased against the hypothesis that Western Union was investing in excess capacity and in favor of the conjecture that Western Union was investing under an optimistic scenario. Since Western Union's demand expectations will be calculated not on European traffic, which is a subset of

[25] The high capacity obtained from the new cables was made possible by loading the cable; wrapping around the copper conductor core an alloy tape consisting of nickel and iron. Bell Engineer G.W. Elmen discovered a cable loaded with permalloy tape would keep the signal more sharply defined, that one signal could follow in succession at a greater speed without distortion from signals overlapping each other at the receiving end.

[26] Oslin (1991), pp. 291.

[27] Herring and Gross (1936), pp. 122.

worldwide traffic, its investments based on those expectations will be higher.

An optimistic scenario for the demand for cable telegraph services is postulated in order to provide an upper bound on Western Union's expected demand. This optimistic scenario includes the extreme assumption that all worldwide cable traffic was between the United States and Europe, and that Western Union had 100 percent market share. The actual percentage of European cable traffic to worldwide cable traffic was closer to 90 percent in 1936. At that time, Western Union carried approximately 50 percent of worldwide cable traffic. Using the optimistic scenario, Figure 1-3 shows that the 20 percent per annum line is the one that most closely estimates the cable capacity for the demand for which Western Union invested. The estimated per annum growth rate for European traffic was approximately 9 percent. Because the only way to justify the investments that Western Union made imply that the extreme assumptions of the optimistic scenario must be believed, either the assumptions or the hypothesis must be accepted. The assumptions are indefensible, and do not comply with the available empirical evidence. Therefore, the hypothesis must be accepted. Western Union over invested in submarine cable. The calculations for Western Union's cable capacity are provided in Appendix 1.C.

1.5.3 Peak load

Peak load considerations is a possible explanation for Western Union's capacity investments. Peak load means that demand has periodic high variance and there are short, predictable periods during the day in which demand is high enough as to justify these investments in capacity. By reviewing how much Western Union carried, on average, in excess capacity before RCA entered the international telegraph market we may determine if peak load considerations explain the level of capacity after RCA entered the market. In other words, we seek to determine whether average loads on submarines changed or stayed at the same level before and after RCA entered the market. If average load fell, as we find, it supports our assertion that Western Union held extra capacity to sustain international telegraph service prices higher than they would had been if competition between RCA and Western Union existed.

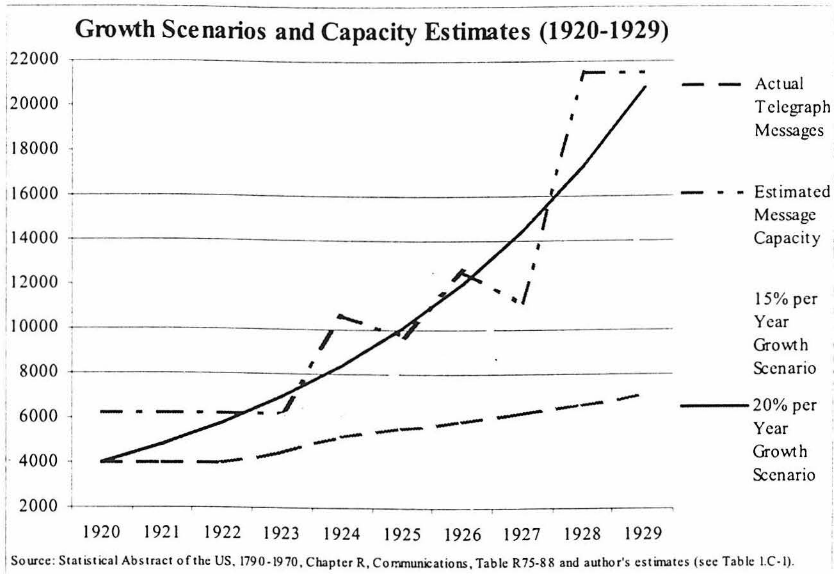


Figure 1-3: Western Union Traffic Growth Scenarios

From available data, we first determine the capacity of Western Union cables for two of the years before RCA entered the market. Unfortunately, reliable traffic data for international telegraphs in the United States are available only for 1907 and 1912.[28] In those years, the number of messages transmitted to Europe from the US was 5,869,000 in 1907 and 5,841,000 in 1912. Assuming that there was no message imbalance between the US and Europe, Europeans sent the US approximately equal number of messages.[29] In sum, there were approximately 12,00,000 messages carried by all ocean cable operators in 1907 and 1912. These message statistics include the activity of 14 other, much smaller, companies that competed with WU for European traffic. Under the best case scenario, WU held 85 percent market share.[30] This

[28]According to the authors of the U.S. Bureau of the Census *The historical statistics of the United States, Colonial Times to 1970,* the First World War affects the count and demand for telegraph traffic.

[29]Note that the technology to send and receive messages simultaneously had not been implemented at this time.

[30]According to a Senate report[32], Western Union held around 80 percent of the market at the end of World

means that Western Union sent and received up to 10,200,000 messages in years 1907 and 1912. This translates into average per minute usage of 40 messages per minute or 2000 words per minute. Western Union had cables with the capacity to send and receive a total of 5,000[31] words per minute. Using the same calculations outlined in Appendix 1.C, we determine that the average per minute capacity of these cables is approximately 3,500 words. On average, therefore, Western Union's used their cables at 57 percent of their full capacity. Given the liberal assumption that we made about Western Union's market share in this period, it is likely that the average load was closer to 50 percent in order to meet peak load demand. In comparison, at the end of Western Union's investments in cables between 1919 and 1930 and RCA's entry, Western Union used its cable at approximately 32 percent of full capacity. Note, moreover, that the calculations for the period after RCA's entry were made assuming that all over Western Union's worldwide traffic passed through its European cables. Therefore, it is likely that average capacity was much lower than 32 percent. Unless it can be shown that variance was much higher during in peak load demand after RCA entered the market, there is no explanation for the shift to Western Union's change to very low average usage in cable capacity. We conclude that evidence supports the hypothesis that this much excess capacity was held in case Western Union needed to punish RCA, by lowering its own prices, for defecting from their implicit collusive agreement.

1.6 Conclusion

Examination of the issue of entry leads to the conclusions that Western Union could not prevent RCA's entry into the US-European communications market because of capacity constraints, the financial strength of RCA, and RCA's initial cost advantages. Between starting a price war that Western Union might lose, or accommodation, Western Union's best strategy was to

War I. Assuming a higher market share will bias calculations in support of the argument, making it more likely to show that excess capacity was needed to satisfy peak load demand.

[31]See Table 1.1.

accommodate RCA's entry. Under the circumstance that existed at the time of RCA's entry Western Union was better off allowing RCA to whittle away at Western Union's market share, and surviving the competition RCA presented.

On the issue of pricing policies, this study concludes that the nature of international telegraph industry was conducive to sustaining collusion, and that both parties benefited from this arrangement. Under the settled collusive price, RCA was allowed to grow and expand its network, and Western Union was under the assurance that RCA would not undercut its price and use its cost advantage to drive Western Union out of business.

Lastly, this study concludes that Western Union invested heavily in excess traffic capacity in order to sustain a binding collusive agreement and to prevent further entry into the US-Europe telegraph market. The argument that an economically rational firm, such as Western Union, would find it profitable to carry excess capacity to obtain and sustain collusion is supported by dynamic duopoly theory. When firms decide on investments in capacity and pricing sequentially and repeatedly, the result is that excess capacity is indeed used as a mechanism to achieve and sustain collusion, and the higher price set by the collusive price remunerates the extra cost of carrying unused, excess capacity.

There are several empirical and theoretical extensions that to the research presented in this study. On the issue of collusion, the question remains as to what took place with respect to foreign telegraph carriers while Western Union and RCA were colluding on price. These foreign carriers also exacted costs for interconnection to their national networks as foreign carriers terminated messages from the US. By varying their interconnection charges, foreign carriers introduced price variation that caused uncertainty. The costs exacted by foreign operators added to the uncertainty of RCA and Western Union's cost function, making collusion harder to sustain. The issue of uncertainty, through the presence of price variation, in collusive or cartel environment is not well addressed in the theoretical literature and almost non-existent in empirical research.

Furthermore, the international regulatory environment requires further investigation. Any research effort that incorporates regulatory institutions into the type of analysis provided here will be able to achieve richer and more robust explanations for the motives and actions of firms in the international telecommunications industry.

The two topics presented above—discussion of collusion or cartels in an uncertain environment and regulatory regimes in an international environment—have policy implications. As the US contemplates deregulation of the international telecommunications market, there is concern over the type of alliances that may form between national monopolies. Some argue that there exists too much uncertainty in the international telecommunications markets for stable alliances between firms in different countries to form. In contrast, others believe that the possibility is real that international telecommunications firms will be able to overcome these uncertainties and form stable international alliances that will extend national monopoly powers internationally. The question of how to regulate international monopolies with no international regulatory body is the second concern presented here.

Studies in economic history, such as this one, do shed light on how firms will act in the nearly unregulated international arena. The tools and analyses provided by economic historians is only one set in many that should be applied to these issues. There is much work to be done on issues of international communications for many disciplines.

Bibliography

[1] Abreu, D., D. Pearce, and E. Stachetti, *Optimal cartel equilibria with imperfect monitoring*, Journal of Economic Theory, **39** (1985), pp. 251-269.

[2] American Telephone and Telegraph Company, *Events in Telecommunications History*, AT&T Archives (1992).

[3] Archer, G.L., *The History of Radio*, New York: The American Historical Company, Inc. (1938).

[4] Archer, G.L., *Big Business and Radio*, New York: The American Historical Company, Inc. (1939).

[5] Benoit, J.P. and V. Krishna, *Dynamic duopoly: Prices and quantities*, Review of Economic Statistics, **54** (1987), pp. 23-35.

[6] Bernheim, D., *Strategic deterrence of sequential entry into an industry*, Rand Journal of Economics, **15** (1984), pp. 1-12.

[7] Brock, G.W., *The telecommunications industry: The dynamics of market structure*, Cambridge, MA: Harvard University Press (1981).

[8] Brown, F.J., *The cable and wireless communications of the world: A survey of present day means of international communication by cable and wireless*, London, England: Sir Isaac Pitman & Sons, Ltd. (1930).

[9] Dixit, A., *The role of investment in entry deterrence*, Economic Journal, **90** (1980), pp. 95-106.

[10] Friedman, J., *A non-cooperative equilibrium for supergames*, Review of Economic Studies, **28** (1971), pp. 1-12.

[11] Gaskins, D., *Dynamic limit pricing: Optimal pricing under threat of entry*, Journal of Economic Theory, **3** (1971), pp. 306-322.

[12] Gilbert, R. and X. Vives, *Entry deterrence and the free rider problem*, Review of Economic Studies, **53** (1986), pp. 71-83.

[13] Green, E. and R. Porter, *Non-cooperative collusion under imperfect price information*, Econometrica, **52** (1984), pp. 87-100.

[14] Haas, A. and D. Sarnoff, *The radio industry: As told by the leaders of the industry to the students of the Graduate School of Business Administration, George F. Baker Foundation, Harvard University*, Chicago and New York: A.W. Shaw Company (1928).

[15] Harlow, A.F., *Old wires and new waves: The history of the telegraph, telephone, and wireless*, D. Appleton-Century Company, Inc., New York and London (1936).

[16] Herring, J.M. and G.C. Gross, *Telecommunications: Economics and regulation*, McGraw-Hill Book Company, Inc., New York and London (1936).

[17] Ireland, N.J., *Concentration and growth of market demand*, Journal of Economic Theory, **5** (1972), pp. 303-315.

[18] Kamien, M.I. and N.L. Schwartz, *Limit pricing under uncertain entry*, Econometrica, **39** (1971), pp. 441-454.

[19] Kreps, D.M. and J. Scheinkman, *Quantity precommitment and Bertrand competition yield Cournot outcomes*, Rand Journal of Economics, **14** (1983), pp. 326-337.

[20] Marland, E.A., *Early electrical communication*, London, New York: Abelard-Schuman Ltd. (1964).

[21] Mason, E., *Price and production policies of large-scale enterprises*, American Economic Review, **29**, pt. 2 (1939), pp. 61-74.

[22] Nonnenmacher, T., *Law, emerging technology, and market structure: The development oj the telegraph industry, 1838-1868,* University of Illinois: Working Paper (1996).

[23] Oslin, G.P., *The story of telecommunications*, Mercer University Press: Macon, Georgia (1992).

[24] Schreiner, G.A., *Cables and wireless: Their role in the foreign relations of the United States*, 2^{nd} Ed., The Stratford Company Publishers: Boston, Massachusetts (1924).

[25] Sobel, R., *ITT: The management of opportunity,* Truman Talley Books:New York (1982).

[26] Spence, A.M., *Entry, capacity investments, and oligopolistic pricing*, Bell Journal of Economics, **8** (1977), pp. 534-544.

[27] Squier, G.O., *Telling the world*, The Williams & Wilkins Company: Baltimore, MD (1933).

[28] Thompson, R.L., *Wiring a continent: The history of the telegraph industry in the United States, 1832-1866,* Princeton University Press: Princeton, NJ (1947).

[29] Tirole, J., *Theory of Industrial Organization*, MIT Press: Cambridge, MA (1994).

[30] Tribolet, L.B., *The international aspects of electrical communications in the Pacific Area*, The Johns Hopkins Press: Baltimore, MD (1929).

[31] U.S. Bureau of the Census, *The historical statistics of the United States, Colonial Times to 1970.* Chapter R, Communications. GPO: Washington, DC (1975).

[32] U.S. Congress, Senate, Committee on Interstate Commerce, *Study of the telegraph industry: hearings before a Subcommittee of the Committee on Interstate Commerce, U.S. Senate, 77th Congress, 1st Session, Pursuant to S. Res. 95—76th Congress (extended by S. Res. 268)*, GPO: Washington, DC (1941).

[33] U.S. Congress, House, Committee on Interstate Commerce Consolidations and Mergers of Telegraph Operations, *Hearings before a Subcommittee of the Committee on Interstate and Foreign Commerce, House of Representatives 77th Congress, 2nd Session, Pursuant to S. 2598, July 21-23, 1942*, GPO: Washington, DC (1942).

[34] U.S. Congress, Senate, Committee on Interstate Commerce, *Consolidations and mergers of telegraph operations: Hearings before a Subcommittee of the Committee on Interstate and Foreign Commerce, House of Representatives 77th Congress, 2nd Session, Pursuant to S. 2598, July 21-23, 1942*, GPO: Washington, DC (1942).

1.A Appendix

Chronology of Events in International Telecommunications History

1854

The Anglo-American Telegraph Company, a British holding company, obtains exclusive rights to land cables from Newfoundland, the eastern most stretch of land in the North American continent. This means that other entrants would be at a cost disadvantage by having to lay longer cables to connect North America with Europe. This exclusive right expired in 1904.

1855

After a series of successful mergers with smaller, independent telegraph carriers, the Mississippi Valley Company acquired Ezra Cornell's lines. The Mississippi Valley Company renamed itself Western Union.

1858

August 5 - After two previous attempts, the first on August of 1855 and the second on July of 1857, the first transatlantic telegraph cable was completed. After transmitting 732, the cable failed on October 20. The voltage used to transmit messages was too high.

1865

August - American entrepreneur with the financial and technical backing of a British company, Telegraph Construction and Maintenance Company (TC&M), manufactured a cable with four layers of gutta-percha and spiral armor of ten iron wires, making it three times the size of the 1858 cable. Approximately 1,200 miles of this cable was laid before it broke and sank to the bottom of the ocean.

1866

June 12 - The American Telegraph Company, which owned the landlines handling cable traffic, was bought by Western Union. This purchase was made a month and a half before the first successful cable was laid. The cable was operated under a cartel of United States and foreign in which Western Union was the U.S. agent for the receipt and distribution of international messages. July 27 - Through the financial backing and technical expertise harnessed by the Anglo–American Telegraph Company, a British holding company, the second transatlantic telegraph cable opened. Permanent communication by wire between the U.S. and Europe was established. On September 8, another venture salvaged the old cable laid down in 1865, establishing two cable connections between the U.S. and Europe. These British owned operated at eight words per minute at a price of $100 in gold or $150 in greenbacks for twenty words or less, including address, date and signature; additional words at five dollars gold.

1868

The French Atlantic Cable Company laid a telegraph cable, the French Duxbury Cable, connecting the U.S. to France directly. The cable reached Duxbury, Massachusetts, on July 23, 1869. Because the Anglo-American Company had exclusive landing rights, the Massachusetts Legislature granted a charter to the Ocean Telegraph Company for the French cable.

1870

Lord Kelvin developed a siphon recorder. With this device, a fine ink mark was made down the center of a moving tape, when positive and negative impulses of electricity passed through the wires. A coil swung the siphon above or below a zero line, making mountain and valley markings on paper. This allowed the messages to be permanently recorded on paper. The messages were indecipherable to those without training, so that they required a trained operator to translate.

1872

The 1866 cable failed, leaving two cables under the operation of the Anglo-American Telegraph Company.

1873

The French, Ocean Telegraph Company, was bought out by the Anglo-American Telegraph Company. The Anglo-American Telegraph Co. deploys another cable (its third) between Ireland and Newfoundland.

1874

The Anglo-American Telegraph Co. deploys another cable (its fourth) between Ireland and Newfoundland.

1877

The Direct United States Cable Company, a rival to the Europe-US cable route monopoly held by the Anglo-American Telegraph Company, was added to the cable cartel formed by Anglo-American.

1879

November 10 - An agreement between Western Union and AT&T established a merger between telegraph and telephone service.

1881

Jay Gould, rebuffed by the Anglo-American Telegraph Company for his requests for Atlantic connections, encouraged the Siemens Brothers to lay two cables between Penzance, England, and Canso, Nova Scotia, for $7,000,000 and formed the American Telegraph and Cable Company (AT&C). All-America Cables, Inc., was incorporated in New York. It operated submarine cables between the continental U.S. and the West Indies, Central America and South America. In 1938, All-America Cables, Inc., was bought by American Cable and Radio (AC&R). Its name was changed to All-America Cable and Radio, Inc. AC&R also owned all of the capital stock of Mackay Radio & Telegraph and Commercial Cable Co. IT&T owned 58.17 percent of AC&R by 1959.

1882

May 12 - Western Union, a U.S. telegraph operator established in 1851, leased ocean cable lines from the American Telegraph and Cable Company for fifty years at a price of 5 percent of $14,000,000. This allowed the largest telegraph company, Western Union, to enter into international service. Western Union went on to buy AT&C for $2,030,000 years later.

1884

The Commercial Cable Co. completed and opened for traffic two cables between Canso, Nova Scotia, and Waterville, Ireland. The editor of the New York Herald was instrumental in organizing Commercial Cable since the press was a heavy user of telegraph services. The reason for organizing a competing company was to bring international telegraph message rates down. Commercial went on to buy Postal Telegraph, established in 1881, to extend its network domestically. Commercial never rivaled Western Union's network completely.

1889

Cable relays were developed to permit automatic operation, eliminating the need for the workers who retransmitted messages across the Atlantic Ocean. Because of the characteristics of cable transmission, however, direct service was not possible over long routes that involved several relays.

1899

A drum relay to repeat an arriving message into another cable was introduced.

1901

December 12 - Wireless expert Guglielmo Marconi signaled Morse code for the letter "S" across the Atlantic from England to Newfoundland.

1902

December 21 - Guglielmo Marconi sent a wireless telegraph message from Glace Bay, Nova Scotia, to England.

1909

December 20 - American Telephone and Telegraph Company (AT&T) acquired control of 30 percent of Western Union Telegraph Company stock by purchase for $30,190,000. Western Union had assets of nearly $500 million, but it had weathered badly the depression of 1907-1908. As a result of the AT&T purchase of Western Union, Theodore N. Vail was president of both AT&T and Western Union.

1911

By this year, all of the Anglo-American cables had been leased by Western Union for ninety-nine years.

1912

Dr. Lee DeForest perfected the three-element vacuum tube to operate as a true amplifier. Dr. DeForest also converted the vacuum tube into a generator of high-frequency currents, creating an oscillator. This oscillator made possible to use the tube in generating or sending of continuous radio waves—whether over air or over wires. The invention of the vacuum tube modulator soon followed. This solved the problem of enabling low power voice energy to control considerably higher power waves required for radiotelephone transmitting. This was achieved by Dr. E.F.W. Alexanderson of General Electric and E.H. Colpitts of the Bell System. These two inventions were crucial to the development of the radiotelephone. E.S. Heurtley's magnifier introduced to increase the power of incoming telegraph signals sent through ocean cable.

1913

The Justice Department, under the provisions of the Sherman Antitrust Act, demanded that AT&T and Western Union be divorced. It accused Northwestern Bell of violating the Antitrust Act by not connecting with independent telegraph and telephone companies. Justice suspected the formation of a formidable communications monopoly. December 19 - A letter from Nathan C. Kingsbury, vice-president of AT&T, to the Attorney General of the U.S., committed AT&T to dispose of its telegraph stock. This promise lead to the termination of the antitrust suit against AT&T.

1914

Advances in vacuum tube technology during this year enabled engineers to begin developing carrier systems for telephone wire circuits and radiotelephone systems. March 19 - AT&T sold its holdings of Western Union Telegraph Company stock to comply with its promise to the Attorney General. The effect of this action was to permanently establish a distinction between telegraph and telephone service. To this day, telegraph and telephone service cannot merge under one company since it would violate the Sherman Antitrust Act.

1915

May 18 - The transmission of a telephone conversation using both radio and wire technology is achieved. August 26 - Beginning of radio trials to transmit speech across the Atlantic. October 21 - AT&T accomplished the first transatlantic radiotelephone transmission between the Eiffel Tower in Paris and a transmitting station in Arlington, Virginia.

1916

G.W. Elmen, an engineer in the Western Electric Research Department, sought a material that could be magnetized more easily and maintain its magnetism longer than iron when he developed an alloy between iron and nickel. The permalloy that he invented was more efficient than the iron cables used by overseas telegraph operators. Western Union would take advantage of this new technology in 1924, laying down a transatlantic permalloy cable that year.

1918

During the previous year, radio technology was developed for military purposes. July 24 - President Woodrow Wilson issued a proclamation assuming control of the telephone and telegraph systems of the United States, placing them under direction of the Post Office Department. On July 30, 1919, return of control was granted to the private owners of AT&T and the telegraph company owners.

1919

The Federal government fostered the organization of the Radio Corporation of America (RCA) in 1919, and forced resolution to the pending patent tangles that the Navy did not resolve during the War. The goal was to keep the U.S. radio communications industry free of foreign control. Prior to its incorporation, the British Marconi (radiotelegraph) Company attempted to obtain exclusive rights to the Alexanderson generator. Negotiations between Marconi and General Electric were nearly completed when the Navy Department indicated its objection to the ownership of this and other American radio patents by foreign interests. At the Navy's behest, President Wilson was asked to intervene and force General Electric to cancel the deal. Wilson did so and General Electric agreed. RCA was organized instead. As soon as RCA was established, it purchased the assets and patent rights held by the American Marconi Company controlled by British interests.

1920

July 1 - A cross-licensing patent agreement was set up between AT&T and General Electric Company, at the request of the United States government, to promote the development of radio communication and make inventions covered by the patents, especially those developed during the war under military direction, generally accessible. This agreement was later extended to include the Radio Corporation of America, Western Electric, and Westinghouse Electric Company. RCA inaugurated radiotelegraph service between the United States and England. Later in the same year, Service was established with France, Germany, and Norway.

1921

Regenerators were developed to allow operation of telegraph cable circuits direct from the United States to Europe, thus speeding up the service and decreasing the cost of operations. The use of regenerators eliminated manual relay and permitted the operation of direct cable service between New York and London.

1923

Another improvement came in 1923 when, in order to increase its traffic-carrying capacity, the Commercial Cable Company, an American firm, laid a line with a copper conductor double the size of previous conductors.

1924

The Bell system announced development of a high permeability alloy made of a new permalloy that Elmen discovered in 1916. Western Union laid a cable of this entirely new design between Hammels, New York, and Horta, Azores. At the Azores, this cable exchanges traffic with cables owned by an Italian company and a German company. This new cable was of the continuous loaded type, and consisted of a copper conductor and a separately wound metallic tape made of permalloy. The capacity of this cable was in excess of 1,500 letters per minute. This capacity compares with the capacity of less than 500 letters obtained with the original type of transatlantic cable and with a capacity of about 1,000 letters per minute developed by the Commercial Cable Co. laid in 1923.

1925

The Commercial Cable Co. deploys a new cable between the U.S. and Europe consisting of a very heavy conductor with the capacity of approximately 1,000 letters per minute. The cost of laying down this cable was in excess of that incurred in laying down the 1923 Commercial cable.

1926

Mackay Radio & Telegraph Co. was incorporated in Delaware, completely owned now by American Cable & Radio (AC&R) which in turn is partially owned by International Telephone & Telegraph Co. Western Union lays down a cable with a capacity of 2,400 letters (or 400 words) per minute.

1927

January 7 - The beginning of transoceanic radiotelephone service on a commercial basis. Commercial service was established between the British Post Office and the American Telephone and Telegraph Co. (AT&T), inaugurating radiotelephone service between New York and London. Later this service was extended to include, on the U.S. side, all of the United States; and on the European side, all of the UK and, in successive steps, most of the countries of continental Europe and other countries accessible through the use of facilities terminating in London.

1928

Western Union lays down another cable with the capacity for 2,400 letters per minute.

1929

RCA Communications, Inc., was organized as a subsidiary of Radio Corp. of America to take over the latter's radiotelegraph operations.

1930

Press Wireless, Inc., was formed to handle press traffic, avoiding what they considered exorbitant prices by other telegraph carriers.

1931

Inauguration of the teletypewriter exchange service (TWX) by AT&T. Teletypewriters had been in use on private lines since 1915. The new service established central switching exchanges through which any subscriber could communicate by teletypewriter with any other subscriber on the service.

1933

By the end of this year, transatlantic service was in operation for most of North America and the principal countries of Western Europe. Ten direct radiotelephone circuits had been established.

1934

The Communications Act of 1934 established the Federal Communications Commission (FCC) charged with regulating, among other things, international communications by telephone and telegraph, whether wire (Canada and Mexico), ocean cable, or radiobroadcast and other forms of radio services. The Act was designed to promote reasonable rates, protect consumers, and maintain adequate public communications facilities.

1.B Appendix

Western Union Company Income from Operations
(thousands)

	1929	1930	1931	1932	1933	1934
Operating Revenues from:						
1. Foreign message transmission	11,725	10,314	8,175	6,339	6,490	5,919
2. Other revenue	361	344	341	311	309	302
TOTAL OPERATING REVENUES	12,086	10,658	8,516	6,650	6,799	6,221
Operating expenses from:						
3. Maintenance and repairs	1,522	1,502	1,361	1,068	1,006	1,124
Percent maintance to total expense	25.7%	24.7%	25.8%	25.3%	22.7%	24.4%
4. Conducting operations	4,008	4,061	3,780	3,287	3,049	3,134
5. General, administrative, taxes	402	508	137	-126	379	346
TOTAL OPERATING EXPENSES	5,932	6,071	5,278	4,229	4,434	4,604
Percent of expenditures to revenue	49.1%	57.0%	62.0%	63.6%	65.2%	74.0%
Operating income	6,154	4,587	3,238	2,421	2,365	1,617
Other charges[1]	2,122	2,567	1,571	1,581	1,697	1,599
Total income from operations	**4,032**	**2,020**	**1,667**	**840**	**668**	**18**

Source: Federal Communications Commission's Annual Reports and carrier's records.

[1] Includes rents for plant leased from Anglo-American Telegraph Co., Ltd., and net rents (rent less certain dividends received from lessor) for plant leased from American Telegraph & Cable Co. and International Ocean Telegraph Co.

Table 1.B-1: Western Union Income from Operations

R.C.A. Communications Income from Operations

	1929	1930	1931	1932	1933	1934
Operating Revenues from:						
1. Foreign message transmission	4,598	4,053	3,592	3,503	4,013	4,285
2. Other revenue[1]	286	339	351	466	451	591
TOTAL OPERATING REVENUES	4,884	4,392	3,943	3,969	4,464	4,876
Operating expenses from:						
3. Maintenance and repairs	168	176	120	143	99	162
Percent maintance to total expense	5.0%	5.5%	4.8%	5.4%	3.8%	5.6%
4. Conducting operations	2,943	2,516	2,126	2,190	2,129	2,293
5. General, administrative, taxes	226	525	278	292	344	413
TOTAL OPERATING EXPENSES	3,337	3,217	2,524	2,625	2,572	2,868
Percent of expenditures to revenue	68.3%	73.2%	64.0%	66.1%	57.6%	58.8%
Operating income	1,547	1,175	1,419	1,344	1,892	2,008
Other charges	269	226	201	186	216	233
Total income from operations	**1,278**	**949**	**1,218**	**1,158**	**1,676**	**1,775**

Source: Federal Communications Commission's Annual Reports and carrier's records.

[1] Continental message transmission revenues for the years prior to 1940 are net after certain commissions paid to Western Union on domestic traffic

Table 1.B-2: RCA Income from Operations

				Employees		
	Western Union					
Year	Telegraph messages (thousands)	Operating revenues (thousands)	Operating expenses (thousands)	Number	Wages and salaries	% wages & salaries/total op. expenses
1919	**581**	*8,078*	*6,356*	**2,688**	**4,612**	72.6%
1920	**4,037**	*8,414*	*6,370*	**3,062**	**4,882**	76.6%
1921	**3,987**	*8,764*	*6,370*	**3,111**	**4,283**	67.2%
1922	**3,992**	*9,129*	*6,315*	**2,603**	**3,902**	61.8%
1923	**4,465**	*9,508*	*6,259*	**2,349**	**3,459**	55.3%
1924	**5,198**	*9,904*	*6,203*	**2,340**	**3,463**	55.8%
1925	**5,520**	*10,316*	*6,148*	**2,352**	**3,559**	57.9%
1926	**5,861**	**10,745**	**6,094**	**2,309**	**3,469**	56.9%
1927	*6,224*	*10,304*	*6,040*	*2,332*	*3,395*	56.2%
1928	*6,609*	*9,881*	*5,986*	*2,299*	*3,392*	56.7%
1929	*7,018*	**12,086**	**5,932**	*2,300*	**3,393**	57.2%
1930	**7,227**	**10,658**	**6,071**	*2,319*	**3,422**	56.4%
1931	*6,591*	**8,516**	**5,278**	*2,213*	**3,265**	61.9%
1932	*6,011*	**6,650**	**4,229**	*2,111*	**3,115**	73.7%
1933	*5,482*	**6,799**	**4,434**	*2,015*	**2,972**	67.0%
1934	*5,000*	**6,221**	**4,604**	**1,894**	**2,794**	60.7%

Figures in **Bold** are data obtained from the Statistical Abstract of the United States, 1790-1970, Chapter R, Communications.
Figures in *Italics* are the author's estimates and taken from other sources such as company records or FCC Reports.

Table 1.B-3: Western Union Operating Revenues and Expenses (1919-1934)

				Employees		
	RCA					
Year	Telegraph messages (thousands)	Operating revenues (thousands)	Operating expenses (thousands)	Number	Wages and salaries	% wages & salaries/total op. expenses
1920	**350**	*4,861*	*1,511*	*918*	*1,331*	88.1%
1921	**960**	*4,317*	*1,743*	*1,161*	*1,684*	96.6%
1922	**1,445**	*4,103*	*1,865*	*1,236*	*1,793*	96.1%
1923	**1,770**	*3,861*	*1,992*	*1,198*	*1,737*	87.2%
1924	**1,890**	*4,709*	*2,243*	*1,381*	*2,002*	89.3%
1925	**2,060**	*5,222*	*2,178*	*1,476*	*2,140*	98.3%
1926	**2,195**	*4,574*	*2,287*	*1,487*	*2,082*	91.0%
1927	*2,300*	*4,492*	*2,246*	*1,523*	*2,208*	98.3%
1928	*2,750*	*4,796*	*2,398*	*1,691*	*2,367*	98.7%
1929	**5,092**	**4,884**	**3,337**	*1,870*	**2,136**	64.0%
1930	**5,151**	**4,392**	**3,217**	*1,679*	**2,478**	77.0%
1931	**4,863**	**3,943**	**2,524**	*1,602*	**1,976**	78.3%
1932	**4,497**	**3,969**	**2,625**	*1,667*	**1,926**	73.4%
1933	**4,909**	**4,464**	**2,572**	*1,642*	**1,878**	73.0%
1934	**5,177**	**4,876**	**2,868**	*1,654*	**1,943**	67.7%

Figures in **Bold** are data obtained from the Statistical Abstract of the United States, 1790-1970, Chapter R, Communications.
Figures in *Italics* are the author's estimates and taken from other sources such as company records or FCC Reports.

Table 1.B-4: RCA Operating Revenues and Expenses (1920-1934)

1.C Appendix

This appendix presents the methods and results of calculations on Western Union's cable capacity.

The total capacity of Western Union's cables was approximately 7,800 words per minute at the end of 1928, if three of the cables deployed in the late 1890s are counted as functional (these cables each had a capacity of 500 words per minute). However, according to Brown (1930), thirty percent of capacity is held in reserve in case a cable breaks. So actual cable capacity is approximately 5,400 words per minute. Furthermore, a deduction in the number of words counted must be made for the telegraphic acknowledgment of receipt, for requests for the repetition of doubtful words and for the repetitions themselves, and for other similar unpaid correspondence. This means that to actual traffic reported as revenue words, words for which revenue was collected from a customer, there is an additional 20 percent more words that the cables must carry, but are not reported in the data used. Deducting 20 percent from the remaining capacity, Western Union's capacity at its peak was approximately 4,300 words per minute.

Because the traffic is in units of messages, not words, capacity must be calibrated to reflect messages, not words. An average of 20 words per message is postulated. For every message sent, there were 30 or so words for the preamble that Western Union required on every message, stating the time of handing in, the number of words, the route (where necessary), and the serial number and class of the telegram. Message capacity, at its peak in 1928, was 86 messages per minute, assuming that each message contained an average of 50 words. If an 18 hour workday is assumed, along with a business year of 240 days, Western Union had the capacity to send over 22,000,000 messages per year. In 1928, all cable carriers, for all markets, sent 14,812,000 messages. If Western Union controlled 85 percent of the whole market, then Western Union sent approximately 13,000,000 messages in 1928. The assertion that Western Union invested in over capacity is proven, but it must be remembered that Western Union, under the expected life of the older cables, was about to lose 1,500 words per minute of capacity. By 1933, Western Union's capacity was expected to be 14,256,000 per year. Western Union transmitted approximately 8,600,000 messages that year. Traffic fell between 1928 and 1933. However, Western Union could not have predicted the fall in traffic since it was a result of the Great Depression.

The next task is to check if Western Union was planning its investments under overly optimistic scenarios, or planning excess capacity as a strategic move in order to keep RCA in check, as is asserted here. Western Union's investment decisions were based on the firm's expectations of what capacity it required. Because submarine cable was expensive, and there existed a substantial lag between time of planning for a new cable and deployment, a four-year lag is postulated. This means that Western Union made investments in 1920, to satisfy demand in 1924. However, Western Union based its expectations of the future on its past. Therefore, the data used to determine investments for 1920, to satisfy demand in 1924, were based on the previous years of operation. For cable, reliable demand data is available up to 1916. A four year retrospection rule is postulated.

The following describes how European traffic was estimated for Western Union. In 1936, European cable traffic is estimated at 90 percent of all cable traffic. Western Union carried 65% of Europe's traffic. According to the Census figures for the period between 1902 and 1912, the rate of growth for the demand for international telegraph messages was approximately 9 percent per annum. The 1917 through 1922 rate of growth is not used because of its unreliability. The per annum growth rate between 1917-1922 are not used because: 1. It is difficult to defend any growth estimates of demand for traffic based on actual messages sent and received during the First World War. 2. The Federal government had control of many telegraph facilities until July 30, 1919. The telegraph industry was operating, up till then, under wartime conditions and government control. The managers of Western Union could not have expected demand conditions to be the same in the future as the demand for telegraph services experienced under government ownership. 3. The method of not using the most current figures when these estimates were influenced by war conditions was already established by the Federal government in the estimate of the growth of agricultural support programs.

	Telegraph Messages	Telegraph Messages	Message Capacity	Growth Scenario	Growth Scenario	in Excess of Demand
1920	4598	4037	*6242*	*4037*	*4037*	54.6%
1921	5046	3987	*6242*	*4643*	*4844*	56.5%
1922	5546	3992	*6242*	*5339*	*5813*	56.4%
1923	6288	4465	*6242*	*6140*	*6976*	39.8%
1924	7230	5198	*10596*	*7061*	*8371*	103.8%
1925	7732	5520	*9725*	*8120*	*10045*	76.2%
1926	8218	5861	*12628*	*9338*	*12054*	115.5%
1927	8694	6224	*11177*	*10739*	*14465*	79.6%
1928	9546	6609	*21482*	*12349*	*17358*	225.0%
1929	12352	*7018*	*21482*	*14202*	*20830*	206.1%
1930	12626	7227	*22353*	*16332*	*24996*	209.3%
1931	11757	*6663*	*22934*	*18782*	*29995*	244.2%
1932	10853	*6144*	*23224*	*21599*	*35994*	278.0%
1933	10835	*5714*	*23224*	*24839*	*43193*	306.5%
1934	10700	*5314*	*23224*	*28565*	*51832*	337.1%

Source: Statistical Abstract of the United States, 1790-1970, Chapter R, Communications, Table R75-88.

Notes: *Italics* signify estimates.

[1] WU traffic was estimated from data on traffic from all cable operators, assuming that WU maintained 85% market share with Europe representing 70% of the total market.

Table 1.C-1: Comparison of Acutal and Estimated Western Union Traffic (1920-1934)

Year	Total	Western Union	RCA
1919	581	**581**	na
1920	4387	**4037**	**350**
1921	4947	**3987**	**960**
1922	5437	**3992**	**1445**
1923	6165	**4465**	**1700**
1924	7088	**5198**	**1890**
1925	7580	**5520**	**2060**
1926	8056	*5861*	**2195**
1927	8524	*6224*	**2300**
1928	9359	*6609*	**2750**
1929	12110	*7018*	**5092**
1930	12378	**7227**	**5151**
1931	11526	*6663*	**4863**
1932	10641	*6144*	**4497**
1933	10623	*5714*	**4909**
1934	10491	*5314*	**5177**

Figures in **Bold** are data obtained from Statistical Abstract of the United States, 1790-1970, Chapter R, Communications.
Figures in *Italics* are estimated from data obtained from other sour

Table 1.C-2: European-US Telegraph Traffic (1919-1934).

Part II

Empirical Study of Household Choice Among Long-Distance Calling Plans

Chapter 2

Applied Econometrics

SUMMARY

While there exist several studies of the firms that compete in long-distance telephone service market, little is known about the consumers who demand these services. This study fills this gap of knowledge by examining consumer behavior in the long-distance telephone service market. We use the Bill Harvesting II Call Detail and Aggregate Databases to study 1,939 households' long-distance telephone service choices and usage for May and June of 1995. Following what is known about consumer behavior in the local telephone service market under uncertainty, risk, and imperfect information, we also examine the role of risk, uncertainty and imperfect information in shaping choice of long-distnace calling plans among those offered by AT&T, MCI and Sprint.

We find that risk and uncertainty play significant roles in explaining household long- distance calling plan choice. We also find that households that call long-distance more than most could save the most by choosing the right calling plan. However, frequent callers are not any more likely than others to choose the cheapest plan. Frequent long-distance callers have high variability in month-to-month usage. High varaibilty makes it difficult for these callers to predict their future usage and therefore know which plan is cheapest across several months. Lastly, our findings show that there exists a bias against paying monthly fees for long-distance telephone service even when these fees are coupled with lower marginal rates. This bias against monthly fees is the opposite to the findings in several studies of consumer behavior in the local telephone service market.

2.1 Introduction

This is the first study, produced outside of industry, to examine consumer behavior in the long-distance telephone service market.[1] While there exist several comprehensive studies of firms and firm behavior in the long-distance telecommunications market,[2] there are few studies on consumer behavior. The ready availability of data to study firms, in contrast to the lack of it for consumers, may be one reason for this lopsided focus on firm behavior.[3] Of the few studies on consumer behavior in the long-distance market that do exist, these are commissioned by the long-distance carriers themselves and their results are considered business secrets not to be shared with the public.

Long-distance telecommunications services represent a growing and substantial monthly expense for many households. US consumers spent approximately $86 billion in long-distance service in 1995 and more than $90 billion in 1998. Beyond being a large market, long-distance telecommunications provide the backbone upon which many emerging information technologies may be built. Similar issues over how to price consumer access to and use of the Internet, for example, are found when examining how consumers choose access and consume long-distance telephone services. As a result of emerging telecommunications technologies, there is an increased focus by economists to understand how consumers react to different pricing schemes to pay for these services.[4]

Many households pay more for their long-distance services than they should even when less expensive options are available to them. For example, of the 7,716 households in our data set

[1] The closest to a study of the one we undertook here is found in a study by Hartman et al.[7]. In that study, the authors found that demographics (especially age), among other factors, are significant in explaining the choice of long-distance telephone company. However, Hartman et al. do not address calling plan choice which is the topic of this study.

[2] See Huber et al. (1993), MacAvoy (1994) and Knittel (1997) to name three of the more recent studies.

[3] Long-distance carriers file financial and operating data with the Federal Communications Commission (FCC). The FCC compiles and produces an annual publication, *Statistics of the Communications Common Carriers*, in which these data are made available to the public. However, no such comprehensive nor regular public compilation of data on residential consumers of long-distance telephone services exists.

[4] See MacKie-Mason, J.K. and H.R. Varian (1995), Pricing the Internet *In Public access to the Internet,* ed. B. Kahn and J. Keller, Cambridge: MIT Press, pp. 269-314.

that made long-distance phone calls throughout an eight month period, we find that households spent \$3,370,825 on these phone calls and could have saved \$1,468,949 had they chosen the plan that charged the least for calls made. We calculated this 43.6 percent in savings by repricing the same phone calls on the plan that was the least costly for that month. While this is a naive calculation, it does provide a rough measure of how much models that assume complete information and no transaction costs would have left to explain of this sub-optimal behavior.

Not all households seem to make as big a mistake in choosing their long-distance telephone service plans as others. The bulk of these savings (approximately \$1,255,000) come from households that made between 16 to 30 or more than 30 long-distance calls per month. Households with the largest calling volume (more than 30 calls per month) could have saved an average of 47.0 percent on their monthly long-distance bills. In comparison, households making fewer than 8 calls per month could have saved approximately 25.8 percent per month. The fact that there exist so many large consumers of long-distance services who could be saving large sums if they chose the right plan is surprising. One would expect that the opportunity for large savings provides incentives for these consumers to become aware of their calling volumes, options, and prices and to take advantage of this information to choose the least costly plan. These large volume callers seem to be either unaware of their options or unwilling to change plans.

In defense of consumers who do make calling plan choice mistakes, the following observations about long-distance services may explain their errors. First, there exist many prices for the same long-distance service, depending on the calling plan, so that search cost may be prohibitively high for a household that may want to inform itself about all long-distance services. Second, long-distance telephone services are priced in complicated and confusing ways so that knowing what price one is paying for a phone call, even after a plan is chosen, is difficult. Third, households must make plan choices before they actually know what their consumption will be. All consumers will make mistakes when having to guess about the future.

Our task is to determine how households cope with complicated pricing schemes, imperfect information and uncertain demand when there is some risk involved. More specifically, we test three hypotheses about consumer behavior and choice in the long-distance calling plan market:

- Uncertainty, imperfect information, and risk are significant in explaining household choice of long-distance telephone service and subsequent usage;
- Households use their past calling patterns as an informational shortcut in order to cope with uncertainty and imperfect information; and
- Some households are better informed about calling plans and their calling patterns because they could obtain substantial savings by reduced uncertainty and imperfect information.

In order to structure our understanding of how households cope with uncertainty, risk and imperfect information, we build on the insights of studies on consumer behavior in the local telephone service market. In the local market, most telecommunications companies offer residential customers several service options for local access. The first is commonly called flat rate service, whereby a household is able to, for a fixed fee per billing period, make an unlimited number of calls within a specified geographic area. The second is generally referred to as measured service, under which the household pays a lower fixed fee per billing period, but incurs a measured marginal expense for each additional call made. A tariff, in the context of local calling, is an algorithm for determining the associated cost of consuming a particular service option, which, in the case of measured service, will be individual specific. The situation whereby several tariffs are offered, from which the customer is able to select the tariff under which consumption is billed, is known as a self-selecting tariff. Economic theory suggests that a rational consumer, faced with a self-selecting tariff, will choose the option that maximizes the personal satisfaction derived from consuming a bundle of goods and services, of which the choice of service option is a (possibly trivial) component. However, to date, empirical and theoretical models of consumer behavior under self-selecting tariff are plagued by apparent inconsistencies, much like the ones we found in our data set and listed above. The main area of concern is the current indication that many subscribers (often a majority) behave irrationally under self-selecting tariffs, and select flat rate service when, in fact, a measured service would be cheaper.

Empirical studies suggest that of those customers who self-select flat rate service, more than half could benefit by switching to a measured service. Conversely, of those people subscribing to measured service, a lesser but nevertheless significant proportion of customers could benefit by switching to a flat rate service [13].

The existence of this flat rate bias in the local telephone service market points out the important role of uncertainty, imperfect information, and risk aversion in shaping consumers' choice of telephone services. For example, the argument that risk aversion plays a role in flat rate bias is based on the view that choosing the flat rate option eliminates bill variation so that flat rates have an insurance quality against exogenous factors that may induce out of the ordinary use of telephone services [13], [25]. However, fluctuations in local telephone bills are so small relative to income that it is hard to imagine the marginal utility of income being anything but constant over the relevant range [3]. So, it is difficult to imagine that households are so worried about insuring themselves against unexpectedly high local telephone bills. However, long-distance bills may be another issue; a household's long-distance bill can be hundreds of dollars per month. They may be more likely to seek an option that guarantees that a long-distance telephone bill won't bust the budget from one month to the next. In addition, consumers may be risk-neutral but overestimate their demand [17]. Overestimating demand may be linked to a past occurrence when the household consumed substantially above their average because of outside and unpredictable circumstances. If such overestimation is common, then consumers may be seen to choose a plan that is inappropriate for their current levels of telephone service consumption.

Another explanation for flat rate bias may also present insights to choice of service behavior in the long-distance calling plan (hereafter, LDCP) market. Some propose that aversion to risk is not necessary in order to account for flat rate bias [3]. An explanation can be found simply by assuming an *ex ante* choice of rate plan, such that the choice of plan be made before usage for the period is known with certainty. Srinagesh [22] recognizes the interpretation of higher willingness to pay as an option value arising from bill stability, while effectively dismissing the claims of the existence of risk aversion on the grounds of credibility. Srinagesh argues that the hypothesis of risk aversion cannot explain the choices of customers whose use is too high to justify a measured service. Indeed, in the long-distance market, there exist plans that exploit this type of preference for bill stability.[5]

Many of the papers cited had the advantage of panel data where households are observed to consume telephone services under different pricing incentives. Having the advantage of such panel data allows the researcher to effectively address the simultaneity problem. The simultaneity problem refers to the difficulty the researcher has in being able to distinguish between the effects of plan choice and plan usage in panel data of telephone usage. Plan choice is presumably a function of a household's expected usage patterns and, at the same time, the plan itself affects a household's expected calling patterns. The studies above either had data where households were observed consuming under one LDCP and then under another, or the author used simulation techniques to obtain such data. We do not have data in which we observe a large enough set of households using long-distance telephone services under one type of plan and then another. We address the simultaneity problem with our own formulation.

The following section begins by summarizing some aspects of supply of long-distance telephone services. The purpose of this summary is to give the reader an idea of the kinds of choices a consumer faces in the LDCP market. The following section covers some of the literature on demand for local telephone access and local calling plans relevant to the formulation of our own empirical model. We conclude with the results of our models.

The results of our study are succinctly summarized below:

- Uncertainty and risk is significant in explaining long-distance calling plan choice;
- Households use average long-distance telephone consumption from previous periods and

[5] The MCI Any Time charges a very high fixed fee of $11.50 per month. With this plan the consumer gets one hour of calling, anywhere and at any time. If the consumer can predict his or her calling pattern relatively well, so as to guarantee that they will not consume beyond the one-hour allocation, then this is a very economical plan.

the variance of past usage as way to cope with uncertainty and imperfect information;

- There exists a bias against subscribing to plans that have a fixed fee component, implying that households may underestimate their long-distance telephone usage; and
- The uncertainty model for the lowest income households, where average usage was used as a variable to minimize uncertainty and imperfect information, performed nearly as well as the certainty, perfect information model.

The last result may seem surprising. We would assume that higher income households have the most incentives to garner accurate information about the prices they are paying for their calls.[6] However, higher income households also have the highest variance in their calling patterns so they are relatively more uncertain about the future usage. Therefore, they are more likely than lower income groups to use average calling volumes from previous periods as a cue of how much they will consume in the future. Given the fairly steady pattern of long-distance telephone service that lower income groups consume, they are relatively more certain of their future usage.

2.2 Supply of Long-Distance Telephone Services

A long-distance phone call is made from one Local Access Transport Area (LATA) to another; i.e., it is an inter-LATA call. A LATA is a geographic unit of administration of local telephone service. A LATA may contain several Area Codes, as is the case in a very urbanized area like Manhattan or Los Angeles, or can share an Area Code with another LATA, as was recently the case in Montana. The charge for long-distance, or toll service, call is made on a per-message basis; thus it is referred to as "message toll service" (MTS).

[6]It is true that a household with a higher income will also have higher opportunity costs and they may be less likely to search for lower prices for their calls. However, given the large sums that could have been saved, it seems reasonable to suspect that opportunity costs cannot explain the whole story for their mistake in not choosing the least expensive plan. See Knittel [11] for search cost arguments in the long-distance calling market.

The price of a long-distance phone call is defined by four dimensions: duration in minutes, time of day and day of the week the call was placed, and distance between originating point and termination point. Duration is calculated in increments of minutes where a call a fraction of a minute is rounded up so that the customer is charged the full minute. Prices are lower for calls placed on weekends, Saturdays and Sundays. Weekdays, between 8 AM and 5 PM are charged the highest prices.

Calls that travel longer distances over telephone wires are priced higher per minute than those of shorter distances. Toll rates increase generally as distance increases, but less than proportionately at longer distances. Increases in toll rates with greater lengths of haul reflect the tendency of the cost per message to increase with distance as a result of greater length of circuits, the additional facilities required to secure satisfactory transmission, and the lower volumes of traffic over the average route. In this study, we will be concerned with Direct Distance Dial (DDD) long-distance calls. DDD means that these are long-distance telephone calls placed by dialing 1+Area code+ number and do not require the assistance of an operator.

The differences in charges for calls based on different days of the week and different hours of the day in which a toll call is placed are partially explained by practice of peak load pricing. Traditionally, peak load pricing was designed with three goals in mind:

- To enable the telephone company to sell a volume of service that will maximize utilization of its network;
- To achieve thereby the lowest average costs of operation; and
- To afford the telephone company the highest possible profits as a result of the first two goals being achieved.

Long-distance telecommunications carriers attempt to achieve full utilization of their network by offering reduced rates to attract, during hours which otherwise would have a comparatively light volume of traffic, business which would otherwise not be obtained at full rates. At the same time, long-distance carriers realize that telephone demand is relatively more inelastic for business customers than for the residential one. As a result, telephone companies charge higher toll rates during business hours and weekdays, when the business customer is most likely to make a call, than for weekends and evenings.

A calling plan is a set of prices that a long-distance carrier offers to a household to determine the charges for calls made. The Call Detail Database reports subscription to eighteen domestic LDCPs. Of these eighteen plans, seven are offered by AT&T, five by MCI, and the remaining six are offered by Sprint. These calling plans can be classified into three types; flat-rate calling plans, volume discount calling plans, and plans that offer discounts to telephone numbers or area codes that the subscribed customer calls most. We have enough information about the pricing of the calling plans, or sufficient observations in our data set, to study household choice for 11, out of the 18, plans.[7]

Flat rate plans. Flat rate LDCPs require the customer to pay a monthly fee to receive an allotted period of calling minutes to be used at off-peak periods—weekends and after 5 PM on weekdays. Generally, after the customer has used up his or her allotted minutes, she/he will be charged for each additional minute at discounted rates. AT&T's *Reach Out America* and *Any Hour Saver* plans are all flat rate offerings. MCI's *Easy Rate* and *Prime Time* plans also fall into the category of flat rate plans. Sprint offered the *Day Plus* as its flat rate plan.

A flat rate plan can be good for a customer with predictable calling patterns. Predictability requires that the customer has limited usage (up to 60 minutes per month for most plans) and can restrict usage to those periods when the plan allocates the "free" minutes of calling.

Volume discount plans. While MCI did not offer a volume discount plan, AT&T offered *True USA* and *Simple Savings* and Sprint offered the *Plus* calling plan. Volume discount

[7] We obtained descriptions for 10, out of 11, calling plans from Dr. Robert Self's *Long Distance for Less Updates* (February, 1996), Market Dynamics, Inc., of Bethesda, MD. Tracy Waldon, of the Federal Communications Commission's Common Carrier Bureau, Industry Analysis Division, provided us with a description of AT&T's True USA plan.

plans offer steeper discounts the more the customer uses the phone to make long-distance calls. These are geared to heavy users of long-distance calling services. Residential customers with heavy use (above $20 per month) in the evenings and weekends that call several area codes benefit most from these plans.

Plans with discounts to calling circles. All three major long-distance carriers provided discounts to calls made to phone numbers or area codes to which the subscribed customer called the most. With AT&T's *Select Saver*, the customer chooses the area code to which the discounts are applied. Under *The Most* calling plan, Sprint provides discounts to the number called the most that month. Finally, MCI's *Friends and Family* is a calling circle plan that offers even bigger discounts if the party called is another MCI customer that is also subscribed to the *Friends and Family* plan. The customers that can receive the greatest discounts with these plans are those that have limited (or predictable) calling circles.[8]

2.3 Models of Local Service Plans

We are not aware of any studies on plan choice and usage in the long-distance telephone service market. However, related issues appear and are addressed in published and unpublished studies of the local calling market.[9] The aim of much of the empirical literature on local calling plan and usage was twofold:

- To determine the extent that household demographics could explain calling patterns, and in turn if these patterns determined calling plan choice; and

[8] If the reader is interested in comparing several long-distance calling plans for their own use, they should use several monthly bills from the recent past, and go to the Telecommunications Reasearch and Action Center's (TRAC) website (www.trac.org/webpricer/form.html), to price several plans. TRAC is a non-profit consumer group based in Washington, DC.

[9] Taylor (1994) notes that his reason for writing his survey book of the telecommunications industry was to ameliorate the problem of inaccessibility of this literature. The papers that exists on telecommunications industry are not widely available to the public. These papers are either in industry studies that firms guard as company secrets, or hidden away in testimony of cases dealing with relevant telecommunications issues.

- To determine the price elasticities of demand in the local calling market.

While several of these studies have varied in methods, data sets used, and assumptions, there is some consensus that the demographic variables, namely income, race, and age composition of a household, are significant variables in explaining variations in call patterns observed in local telephone usage data. The more concrete consensus is on price elasticities. Taylor [24] cites an industry standard of approximately -0.5 as the toll elasticity of the local telephone service. Taylor comes to this conclusion after surveying various studies. Moreover, race, age composition and size of a household, income, and recent residency in an area influence local calling plan decisions and calling patterns [2].

2.3.1 Two-stage model of plan choice and usage

The standard model commonly used to describe the consumer's subscription decision for telephone service is the two-stage model, pioneered by Squire [21] and Littlechild [14], and well formulated in Taylor [24] and Train [27]. Within the model, consumers must initially decide whether or not they are going to subscribe to the service being offered. Assuming the consumer decides in the affirmative, then the second stage requires a decision regarding how much of the chosen service to use. If the consumers' surplus associated with the optimal level of usage in the second stage exceeds or equates the subscription price, then the first stage decision is to subscribe [13]. Conversely, if the consumers' surplus associated with the optimal level of usage in the second stage falls below the subscription price, then the first stage decision is to not subscribe.

The formal representation below draws largely on Kridel et al. [13], which presents the consumer's problem in the context of a simple two good economy. Consider a representative consumer with utility function $U(x, m)$, where x is the consumption of telecommunications services and m represents all other goods. The two stages of the consumer's problem described above are solved in reverse order, such that the consumer, given access to the service, will choose to

$$\max U(x,m) \ s.t. \ px + m = y \tag{2.1}$$

where y is income and p is the usage price of the telecommunications service chosen.

This describes the second stage decision problem for a consumer who already has access to the network. Assuming the second order conditions are satisfied, then the first order conditions yield the ordinary demand curve, $x(p,y)$, from which consumers' surplus can be derived. The consumer's first stage decision problem is solved such that the consumer should:

$$subscribe = \begin{cases} 1 \text{ if } \int_p^\infty x(t,y)dt > K \\ 0 \text{ if } \int_p^\infty x(t,y)dt < K \end{cases}$$

where K is the subscription price. This two-stage decision process forms the backbone for most applied telecommunications demand analysis.

In using this model to investigate consumer choice and usage, we must assume that consumers are certain of their future level of usage and well-informed about all relevant prices. We also must assume that consumers faces no transaction costs, and can switch calling plans instantaneously, and they know all relevant prices, and therefore face no search costs.

We can adapt the two-stage model to fit our needs in explaining consumer behavior in the long-distance telephone service market by assuming that consumers are already provided this service and that their first choice is to determine to which long-distance calling plan they will subscribe. The consumer's problem with respect to choice of service is then to maximize utility choosing the plan that minimizes expenditures on calls given their known level of future usage. We provide the following example in which consumers in the long-distance calling plan market face a nonlinear (two-part) tariff for these services:

1. There exist two alternatives, $\{i,j\} \in C$;

2. The price per unit of alternative i is p_i and the price per unit of alternative j is p_j.

3. Alternatives have fixed charges $F_i,\ F_j \geq 0$.

The utility, U_i, from consuming alternative i is $U_i = y - (F_i + p_i x)$. Similarly, for alternative j, utility is $U_j = y - (F_j + p_j x)$. We assumed that the level of consumption, x, does not change before the consumer chooses a plan. In comparing the two plans, a representative consumer will choose alternative i when $U_i > U_j$. In other words, plan i is chosen when:

$$F_j - F_i + (p_j - p_i)x > 0$$

If we consider the price of all other goods and services, p_0, and the consumer's income, y, the equations for the demand of a plan are:

$$U_j = f(\tfrac{p_i}{p_0}, \tfrac{p_j}{p_0}, \tfrac{F_j}{p_0}, \tfrac{y}{p_0})$$

$$U_j = g(\tfrac{p_i}{p_0}, \tfrac{p_j}{p_0}, \tfrac{F_j}{p_0}, \tfrac{y}{p_0})$$

$$j = h(\tfrac{p_i}{p_0}, \tfrac{p_j}{p_0}, \tfrac{F_j}{p_0}, \tfrac{y}{p_0}) = \begin{cases} 1 \text{ when } j \text{ is chosen} \\ 0 \text{ when } i \text{ is chosen} \end{cases}$$

Functions f, g, and h are functions of calling plan rates, the prices of other goods and services, and income.

This model predicts that, if a consumer chooses to consume the same number of minutes under either of two LDCPs, that consumer would choose the cheapest plan. This result is a product of the fact that the model neglects other significant influences on consumer decisions.

There are problems in implementing this model. To correctly estimate consumer choice behavior, there needs to exist sufficient price changes in the data and consumers must be observed consuming under all plans. If the number of tariff elements associated with all available options exceeds the number of price changes, it is not possible to use the above model to estimate demand and say anything about the representative consumer's preferences. With the large number of available alternatives in the long-distance telephone service market, we may run into problems in estimating choice behavior if there is not enough price variation in our data set. Restrictions on the pattern of cross-price elasticities across alternatives offered by economic theory would help only marginally if many calling plan alternatives are relevant. If the number of price differences between plans is small, the number of alternatives that can be considered in the analysis is likewise small.

Subscription to one alternative or another is inherently a discrete variable. Alternatives calling plans are mutually exclusive, where usage in any one plan will be zero when another plan is chosen. All of the models we discuss use discrete choice models of plan choice. In using a discrete choice model, individuals are assumed to maximize their net benefit or utility. The probability of choice of any alternative is assumed to be a function of the net benefit or utility associated with that alternative. In order to know the net benefit of each alternative, the costs savings associated with each alternative must be measured. If the benefits are a function of the tariff elements alone, alternative specific constants in the models absorb all the influence of there tariff elements. Instead, analysts use the costs associated with each option to compare alternative plans, and each cost is calculated using some customer-specific measure of usage. If a new alternative is being offered, usage prior to the introduction of the new plan can be used to calculate the cost of each alternative.[10]

2.3.2 Option value model of plan choice

The observed skewed preference for fixed rate plans in the local telephone service is much of what drives economists to examine the nature of this choice of local toll plans. Many economists argue that uncertainty and risk aversion play a key role in call plan decisions. Uncertainty is an implicit characteristic of future demand for telecommunications services, which, in many cases, invalidates the two-stage decision approach outlined above. If we do not take into

[10] Indeed, this is exactly what we did to calculate the savings of our data set that we cited in the Introduction.

consideration uncertainty in consumer choice, we impose the assumption that the consumer knows, in advance, whom she will call, when she will call, and for how long they will converse. In the presence of uncertainty, the two-stage decision model requires modification such that usage demand is replaced by expected usage demand. Consumers cannot easily tell what their typical calling patterns might be, and this greatly complicates decisions about plan choice.

By explicitly modelling an option value into consumer demand, we can address some of the uncertainty problems inherent in calling plan choice. In the context of local calling options, measuring consumer benefits associated with particular options while incorporating perfect foresight is, in a theoretical sense, relatively simple. Changes in consumers' surplus provide an appropriate measure of a particular service options' contribution to consumer welfare. However, in reality future prices and outcomes are largely uncertain, and there is a strong case for concluding that measuring expected consumers' surplus alone is inadequate. Option value is the adjustment to expected consumers' surplus that is required when uncertainty exists in the environment in question.

Weisbrod [28] pioneered the concept of option value, which was initially used as an explanation of how market failure may occur in circumstances where usage uncertainty, combined with a failure to collect the value of reducing uncertainty about the availability of supply, distorts perceptions of true market conditions. He argued that consumers, faced with uncertain demand for a good or service, would be willing to pay some premium such that the total willingness to pay exceeded the expected consumers' surplus. The premium is paid by the consumer in order to maintain an option to consume the good or service in future periods, where the premium, commonly called the option value, represents a certain payment, independent of usage sensitive factors.

The following is a description of the Kridel et al. [13] study of extended area service in the US, and a summary of its findings. The authors use and measure the option value to address the problem of uncertainty over the choice of different local telephone services.

Extended Area Service (EAS) is a route specific calling plan, offering EAS customers unlimited calling both into and from the Dallas local calling area for a fixed fee of \$19.85 per month. Those who do not subscribe to EAS pay a usage sensitive fee accumulated according to an \$0.11 marginal usage per minute price. The study concluded that neither of the existing consumer choice models did a good job of predicting the observed subscription levels.

Kridel et al. utilize observed purchasing and usage data from a sample of households to estimate a probabilistic choice model, which includes both continuous and discrete components. The discrete portion of the model is employed to predict the probability associated with EAS subscription for a given individual, from which the market penetration prediction can be derived using a sample enumeration technique. The continuous portion of the model can be used to predict EAS usage. To facilitate estimation, a semi-log demand function of the form $X = Ae^{-ap}y^{\beta}$ is adopted, where X denotes usage, p is price, and y is income. Usage benefits derived by the subscriber from purchasing EAS will be a measure of the change in consumers' surplus that results when the service is purchased. This is found by integrating the demand function over the price range $[0, \rho]$, which yields the expression:

$$\Delta CS = \frac{A}{\alpha} y^{\beta}(1 - e^{-\alpha p}) = \frac{X_N}{\alpha} - \frac{Xo}{\alpha}$$

where X_N is the customers expected usage under EAS and X_O is the usage under toll. Given N uncertain states of the world, indexed $i = 1, ..., N$, we wish to consider ΔCS_i, and then calculate expected consumer surplus:

$$ECS = \sum \pi_i CS_i$$

where π_i represents the probability that the i^{th} state of the world will occur, and CS_i is the corresponding level of consumers' surplus. The two-stage decision model will predict EAS subscription if ECS exceeds the subscription price. However, we are interested in the total benefit derived from EAS, the value of which, for any individual, may be greater than, equal to, or less than ECS. Total benefits, which are unobservable and therefore approximated by the subscription price, K, are given by:

$$TB = OV + \Delta CS + \varepsilon$$

where OV is option value, ΔCS is as defined above, and ε is a well-behaved error term. The probability that a customer will purchase EAS is therefore given by:

$$\Pr(EAS) = \Pr(\varepsilon > K - OV - \Delta CS),$$

which may be rewritten:

$$\Pr(EAS) = \Pr(\varepsilon^* > \gamma(K - OV - \Delta CS))$$

where $\varepsilon^* = \frac{\varepsilon}{\sigma}$ and $\gamma = \frac{1}{\sigma}$ is a choice parameter, while σ is the standard deviation of household usage over the sample period prior to EAS introduction. By utilizing actual EAS subscriber usage, it is recognized by Kridel et al. that further refinement of parameter estimates is possible by comparing expected usage and actual usage under EAS.

Using this framework the authors conclude that expected consumers' surplus cannot explain at least 30% of the customers' decisions (40% of the 2/3 subscribing to EAS + 9% of the 1/3 non-EAS subscribers). Expected consumers' surplus does not seem to provide the best measure of welfare derived from local telephone usage, implying that models utilizing expected consumers' surplus alone may not accurately explain consumer choice among self selecting tariffs. However, the above analysis, in line with normative models of economic theory, views EAS subscription as rational decision making under uncertainty, although empirically this may not be the case. Colton's [4] contribution suggests that consumers may simply be misinformed about the services available, thereby inhibiting the rational decision making process. Kridel et al. tentatively allow for this possibility, but note that steady increases in EAS subscription over time, combined with low drop rates, suggest that error is not a viable explanation. They also express concern regarding sample size, but find no justification for why a longer time period would systematically work to increase predicted subscription rates. Finally, it is noted that the model may be misspecified in a way which systematically underpredicts subscription. However, in their defence, the authors maintain that viewing EAS subscription as rational decision making under uncertainty, and explicitly modelling this uncertainty through option value, significantly improves the estimation.

2.3.3 Dynamic stochastic model of plan choice

Srinagesh [22] presents a general dynamic stochastic model of choice by introducing a random component into the decision making process. In this model, individuals once again face a two-stage decision problem, although in this case it is assumed that calling volumes for individuals vary substantially across periods. The variation Srinagesh identifies is partly stochastic, reflecting the reality that calls are often initiated to handle unforeseen developments, and partly deterministic, reflecting seasonal variation in schedules and lifestyles. He further hypotheses that rational individuals, when choosing between service options, possess some knowledge of the underlying probability distribution associated with calling volumes over a planned horizon, and that individuals seek to optimize over this horizon.

The model assumes that there exist I consumers, indexed $1, ..., I$, and that consumer i has a planning horizon consisting of T_i billing periods, indexed $1, ..., T$. Q_{it} is defined as the number of local calls made in billing period t, while Y_{it} is the amount of a Hicksian composite consumed in period t. The stochastic component is represented by θ_{it}, which describes the shocks that affect the i^{th} individuals taste for local calls in period t, while F_{it} is the corresponding probability distribution function. The model is such that an individual must choose a service option *ex ante*, although consumption is determined *ex post* (once θ_{it} is observed), with the i^{th} customer's preferences being modelled by the period t utility function:

$$U_{it}(Q_{it}, Y_{it}, \theta_{it})$$

It is assumed that there exist $k = 1, ..., K$ billing options available to the individual, and that each option generates a (possibly nonlinear) budget set, denoted by B_{itk}. The budget set describes the (Q_{it}, Y_{it}) combinations attainable by the i^{th} individual in period t given the k^{th} billing option. Once again the decision problem is solved in reverse order, such that the individual maximizes expected utility (with respect to F_{it}) in each period in the horizon, subject to the particular tariff and the realization of θ_{it}. The first stage decision then consists of selecting the plan that yields the highest maximized expected utility.

The second stage problem can be formally expressed as:

$$\max_{Q_{it}, Y_{it},\, t} \sum U_{it}(Q_{it}, Y_{it}, \theta_{it}) \; s.t. \; (Q_{it}, Y_{it}) \in B_{itk}$$

which yields the solution (Q^*_{itk}, Y^*_{itk}). This solution will be dependent on B_{itk} and θ_{it}. The first stage problem then becomes:

$$\max_{k,\, t} \sum U_i(Q^*_{itk}, Y^*_{itk}) dF_{it}$$

The outcomes of each of the respective decision processes will reveal a subscription rate for each option and a usage distribution in each period for the entire population. Srinagesh hypothesizes that data pertaining to both the usage distribution in any given billing period and the option selected by each consumer are generated by rational individuals who solve the very complex problem described above. Further, he provides a restricted model that can be used to gain insight into the relationship between the choice of service option and actual usage in any given billing period.

The basic assumptions of the model remain intact, such that the choice of service option is made by individuals only once, and usage decisions are made after uncertainty is resolved in each billing period. It is additionally assumed that customers face a single period planning horizon, with tastes represented by the uncertain demand function $Q_i = X_i \exp(-\beta_i p)$, where X_i is distributed according to a gamma distribution with shape parameter 1 and mean μ_i. We assume for simplicity that $\beta_i = 1$, and that μ_i follows a gamma distribution with shape parameter 1 and mean 120. These assumed distributions reflect stylized facts and are therefore broadly consistent with observed usage distributions.

Srinagesh's analysis is primarily concerned with explaining why cross-sectional data reveals a "bias" toward fixed rate service in customer choice. His simulations illustrate the apparent bias displayed by rational individuals with uncertain demands, based on the assumption that consumers make their choice decisions *ex ante*, and usage decisions *ex post*. Srinagesh stresses that there are two interactive factors generating the bias in the data. The first is that the critical level, the level of usage at which a measured service and a fixed rate service are equally expensive, is close to the inflection point, implying that the customer distribution density is much higher to the right of the critical point than it is to the left. The second recognizes that variance of use increases with mean usage, implying that a larger interval of customers to the right of the critical level can have realized usage below the critical level.

Overall, the model above provides a dynamic stochastic framework for analyzing customer choice among self-selecting tariffs, while successfully providing an explanation for the apparent bias associated with choice among service options in the telephone market. The model also suggests a general framework for econometric analysis of usage distributions, although there is nothing in this approach that limits its application to fixed rate service and measure service, or even to telecommunications.

Other models as Hobson and Spady [8] specify uncertainty by assuming that each household has a taste parameter for long-distance phone calls, θ,that is random every month. The household's choice of calling plan and telephone usage could be conditional on the θ parameter:

$$\phi(X, C|z) = \int f(X|\theta, C)h(C|\theta, z)g(\theta|z)d\theta$$

By writing $g(\theta|z)$,where z is a household's demographic characteristics, they indicate that each household has their own value of θ, and that there is a population distribution corresponding to each value of z. The distribution of actual demand, $f(X|\theta, C)$, where C is the calling plan choice, and is relevant since it will determine the price the household will pay for long-distance calls. Actual demand, X, is therefore conditional on the choice and demographic characteristics. Hobson and Spady specify that $X \neq x^*$; actual usage may not be equal to optimal usage. The authors give many reasons as to why actual usage may deviate from the optimal—vacations, and unusual number of call initiations from others, even taste variations. The main point they make by modeling plan choice and subsequent use in this way is to have these choices depend on a household's future (uncertain) utilities.

Kling and Van der Ploeg (hereafter, KP) [10] studied household local option plan choice and usage for Michigan Bell. They assumed a consumer surplus function where the net benefit of two alternatives depended on the difference in the fixed monthly fees, the value of the change in usage across alternatives, and the value of the call allowance. KP made assumptions about the stochastic patterns of the daily call behavior of individual households, the distribution of calls across households, and class of service choice. Using eight months of daily call data for 427 local measured and flat rate service customers in Michigan, they estimated a joint model of class of service choice and usage. The results suggested that a high proportion of those who did not choose measured service, but could have saved by doing so, were unaware of the option.

KP use a stochastic demand model, where there exists a distribution over prices, to generate results of choice and usage under uncertainty. They calculate how consumer surplus is affected by choice of plans that minimizes these price distributions but actually cost more upfront; i.e., these plans have higher monthly fees. This is the insurance argument—relatively more risk averse households are willing to choose plans that have less variation in calling prices as a way to insure itself from exorbitantly high telephone bills in months when circumstances induce higher long-distance usage.

KP begin by assuming a semi-logarithmic demand function:

$$q_i = \theta_i \exp(\beta p)$$

where

$q_i \equiv$ local calls demanded by the i^{th} household

$\theta_i \equiv$ satiation usage for household i

$\beta \equiv$ a parameter that measures the price response

$p \equiv$ marginal price of a local call

The semi-logarithmic function is defined where $p = 0$ and therefore allows for a satiation level of usage under flat-rate service, namely θ.

With income and demographics taken into account, the demand function becomes:

$$q_i = \theta_i \exp(\alpha z + y^{\gamma} \beta p)$$

where y denotes income and z denotes a vector of demographic characteristics for that household. By including income as a non-linear, multiplicative term with the marginal price, KP estimate usage elasticities from a model of call of service and usage demand for the residential local call market.

KP use a Polya-Aeppli distribution to parameterize calling pattern distributions. We cite these authors' reasons for choosing this distribution, namely:

- There appears to be a contagious element to call where one call leads to another call; and

- Many of the days in the data set contain no calls (zero demand) for a large number of households.

We employ the Polya-Aeppli distribution in our statistical model of the daily calling behavior of individual households:

$$f = \begin{cases} \Pr(x=0) = \exp(-\lambda) \\ \Pr(x=k) = \exp(-\lambda)\Gamma^k \sum_{j=1}^{k} \binom{k-1}{j-1} \left[\frac{\lambda(1-\Gamma)}{\Gamma}\right]^j, \text{ where } k = 1,2,..,n. \end{cases}$$

This distribution is an appropriate model of the distribution of calls per day for an individual household when: (a) the number of different purposes for communications has a Poisson distribution with parameter λ and (b) the number of calls required for any purpose has a geometric distribution with parameter Γ.

2.4 The Model

The model we present differs from both the stochastic choice models and the option value models in important ways. First, we use a utility maximization approach, as opposed to that of a maximization of consumer surplus, in determining calling plan choice.

Second, while we use a similar framework to that found in KP and Hobson and Spady, but we do not model consumer uncertainty in the same way. Our formulation is closer to that found in the options value where use the mean and variance of previous usage to get a measure of expected future usage. More specifically, we model consumer uncertainty as the average of past usage plus a well-behaved error term. We also use the variance in call patterns as a measure of uncertainty. KP and Hobson and Spady made assumptions about the stochastic patterns of the daily call behavior of individual households and the distribution of calls across households to model uncertainty.

Third, the KP and Hobson-Spady models estimated a joint model of class of service choice and usage to get at the extent of information consumers had about the market and price elasticities. These models had the benefit of panel data in which households were observed to make calls under various pricing plans. We do not have the benefit of such data. As a result, the models presented above can directly estimate elasticities for each household. In our formulation, we use the variable that determines price elasticity as a search parameter that maximizes our proposed maximum likelihood function. We use a probit, discrete choice model to estimate a household's satiation level, usage and price elasticity and determine the extent of uncertainty in this market.

In the following section, we show that consumer utility theory can be used to create the econometric model to explain LDCP choice. We begin by assuming that the consumer is completely informed about the rate structures of all alternative calling plans, and can accurately estimate their future monthly long-distance call consumption. We relax these assumptions once we examine the implications of the fully-informed, certainty model.

2.4.1 Utility function

We begin by assuming that consumers maximize utility U, where U is a function of the quantities of all services consumed. For simplicity, all other goods consumed in the economy will be lumped together as income minus long-distance telephone bill expenditure. Income and all prices will also be deflated by the price of non-telecommunications service. The usage price for a particular LDCP j is p_j. If there exists a flat fee with LDCP j, we represent that fee with $F_j \geq 0$. The usage associated with that plan is x_j. Income is represented with y. The variable θ represents the consumer's satiation point where he would consume θ minutes of long-distance service if $p_j = 0$, where $\theta \in [0, \infty]$. Total non-telecommunications goods and services are represented as $y - \sum_{j=1}^{n}(p_j x_j + F_j)$. Let $B_j = p_j x_j + F_j$, where B_j is total expenditure on LDCP plan j. Then for given level of income, LDCP rates, and flat fees, utility U_j will be a function of quantities (minutes) consumed:

$$U_j(x_j : y, p_j, F_j) = \frac{1}{y}[y - \sum_{j=1}^{n} p_j x_j + F_j + \sum_{j=1}^{n} \frac{x_j}{b}(\ln \frac{\theta_k}{x_j} + 1)] \tag{2.2}$$

When utility is maximized for this form of utility function, at the point where marginal utility is zero of U_j, it can be shown that:

$$x_j^* = \theta \exp(-b\frac{p_j}{y}) \tag{2.3}$$

We know that $-b\frac{p_j}{y}$ is the consumer's demand elasticity with respect to price since $\frac{\partial \ln x^*}{\partial \ln p_j} = -b\frac{p_j}{y}$. For the law of demand to hold, b must be positive. In this model, demand elasticity is a function of income. In our formulation of the utility function, we assume that long-distance telephone calls are luxury goods where more is consumed of this service as income increases, *ceteris paribus.* Note we also assume separability of utilities between phone calls. The separability assumption implies, for example, that calls of different times of day are not substitutable.

One of the features of the semi-log utility function used here, and by KP [10] and Hobson and Spady [8] previously, is that this function is consistent with risk aversion. Absolute risk aversion R is defined as the ratio between the negative of the second derivative over the first derivative of utility with respect to income [19]. If this ratio is positive, then the agent is said to be risk averse. The above utility function yields:

$$\begin{aligned} R &= -(\frac{\partial^2 U}{\partial y})/(\frac{\partial U}{\partial y}) \\ &= \frac{2}{y} > 0 \end{aligned} \tag{2.4}$$

Usage variations result in fluctuations in telephone bills over time. Risk aversion, in this context, suggests that customers with higher bill variation will tend to subscribe to rate plans with lower usage rates and higher monthly subscription fees more frequently than customers with low bill variation, given similar level of average bill savings.

Note that with this utility function, the extent of risk aversion is inversely proportional to income. Therefore, higher income households are relatively less risk averse than lower income ones.

2.4.2 Bill minimization and certainty model

One key hypothesis to be tested is whether consumers minimize their bills or maximize utility in their choice of rate plan. While bill minimization is a simpler theory, it may fail to predict subscription of plans that have a substantial flat rate component—i.e., high fixed fee with comparatively lower marginal costs for the customer.

In the previous section, we posited the following utility function:

$$U_i = (x_i; y, p_i, F_i) = \frac{1}{y}(y - p_i x_i - F_i) + \frac{x_i}{b}(\ln \frac{\theta_k}{x_i} + 1) \tag{2.5}$$

For positive prices for long-distance calling service i, x_i will be less than θ_k, and $\ln \frac{\theta_k}{x_i}$ will be nonnegative. We assumed that the consumer's problem in deciding which calling plan to choose could be solved by maximizing the consumer's utility with respect to a budget constraint y and determining which of these plans provided the consumer with the highest utility. Let U_{ik} be consumer $k's$ total utility associated with choosing plan i and let U_{jk} be the same consumer's utility in usage for choosing plan j. Consumer k will choose plan i and usage x_i if $U_{ik} > U_{jk}$. The condition for choosing i and usage x_i, assuming utility maximization is:

$$U_{ik} - U_{jk} > 0$$

By substituting, we get the following relationship:

$$\frac{1}{y}(y - p_i x_i - F_i) + \frac{x_i}{b}(\ln \frac{\theta_k}{x_i^*} + 1) - (\frac{1}{y}(y - p_j x_j - F_j) + \frac{x_j}{b}(\ln \frac{\theta_k}{x_j^*} + 1))$$

If we substitute for x_i^* and x_j^* in expressions $(\ln \frac{\theta_k}{x_i^*} + 1)$ and $(\ln \frac{\theta_k}{x_j^*} + 1)$, respectively, and factor out, we get:

$$U_{ik} - U_{jk} = \frac{F_j - F_i}{y} + \frac{x_i^* - x_j^*}{b} \tag{2.6}$$

The above expression formally states the prediction of utility theory. Namely, utility theory predicts that consumer k prefers LDCP i over j when the value from a change in plans is nonnegative. The usage levels under plans i and j for the same consumer will differ because the price for the last minute of usage will vary between the two plans. Actual usage associated with plan i rather than plan j would therefore depend not only on the marginal usage price under plan i but also on the subscription fees of both plans and the price of the last minute of use under both plans.

Equation 2.6 shows that consumers decide on their LDCP by comparing the fixed monthly fee across plans and the cost of usage on each plan. A plan with a higher monthly fee will be chosen if the difference between monthly fees is offset by higher usage levels induced by lower marginal prices. The above specification also suggests that a simple discrete choice model of plan choice could be directly assumed fully taking into account usage variation across plans.

The above equation also appears to exclude the size of bill savings across alternatives from utility comparison. The size of bill savings is central to plan choice if households were solely bill minimizers. Given that this utility maximization condition is not the same as bill minimization, we ask how does this model compare to positing the consumer's choice problem as bill minimization. Our result differs because we posited a utility function that has the property of risk aversion—the well-known duality result applies to risk neutral agents.[11] We next show that we do not need to drop risk aversion to get pure bill minimization from this utility function.

Given the apparent differences between bill minimization and utility maximization, we next examine under which conditions our model satisfies pure bill minimization. Economic theory suggests that a pure bill minimization algorithm in choosing a plan implies that price elasticity, b, would approach zero. We show next that a pure bill minimization can be obtained by examining the above equation as b goes to zero.

The price elasticity variables appear only on the second part of the above equation, so we can restrict our examination there. So, we can restrict the problem to determining the relationship:

$$\lim_{b \to 0} \frac{x_i - x_j}{b}$$

By L'Hôpital's rule, we know that the above expression is equivalent to the following:

$$\lim_{b \to 0} \frac{x_i - x_j}{b} = \lim_{b \to 0} \frac{dx_i}{db} / \frac{db}{db} - \frac{dx_j}{db} / \frac{db}{db}$$

The usage that maximizes the consumer's utility when b approaches zero is $-\frac{p}{y}\theta_k$. So we can substitute for the above expression:

$$-\frac{1}{y}(p_i - p_j)\theta_k$$

Therefore, as b approaches zero, we get a pure bill minimization algorithm for LDCP choice:

$$U_{ik} - U_{jk} = -\frac{1}{y}[(F_i - F_j) + (p_i - p_j)\theta_k]$$

[11]See Varian (1992) p. 113.

The above analysis shows that bill minimizing and utility maximization are the same under the assumption that the consumer's price response is close to zero and therefore choose calling plans by repricing at their satiation usage across alternatives, choosing the calling plan that costs the least.

We can estimate the bill minimization model, to test if this is the algorithm that households use to choose plans. If we posed the choice problem in terms of expenditure minimization, we would find that our model underpredicted choice for plans that had zero marginal costs and positive fixed fees. Under a pure bill minimization hypothesis, the consumer will choose plan j (where $C_j = 1$) when:

$$C_j = \begin{cases} 1 \text{ if } \frac{(p_j\theta_k+F_j)-(p_i\theta_k+F_i)}{y} < 0 \\ 0 \text{ if } \frac{(p_j\theta_k+F_j)-(p_i\theta_k+F_i)}{y} > 0 \end{cases}$$

We can test the bill minimization hypothesis by using the relationship found in Equation 2.6, namely:

$$C_j = \begin{cases} 1 \text{ if } \frac{F_j-F_i}{y} + \frac{x_i^*}{b} - \frac{x_j^*}{b} < 0 \\ 0 \text{ if } \frac{F_j-F_i}{y} + \frac{x_i^*}{b} - \frac{x_j^*}{b} > 0 \end{cases}$$

If we substitute $\theta \exp(-b\frac{p_i}{y})$ and $\theta \exp(-b\frac{p_i}{y})$ for x_i^*, x_j^*, respectively, we can implement the following econometric model:

$$\Pr[C_j] = \begin{cases} 1 \text{ if } \alpha_1 + \beta_1(\frac{F_j-F_i}{y}) + \beta_2(\exp(-b\frac{p_i}{y}) - \exp(-b\frac{p_j}{y})) < 0 \\ 0 \text{ if } \alpha_1 + \beta_1(\frac{F_j-F_i}{y}) + \beta_2(\exp(-b\frac{p_i}{y}) - \exp(-b\frac{p_j}{y})) > 0 \end{cases} \quad (2.7)$$

Note that coefficient $\beta_2 = \frac{\theta}{b}$. The α_1 is the intercept term and is expected to be zero, and the β_1 is the coefficient that measures net utility of the fixed fees between the two plans. Since the default no plan choice has no fixed fee ($F_j = 0$), β_1 is the coefficient for $\frac{-F_i}{y}$. The model predicts that $\beta_1 = 1$. Since the β_1 term is an extra free term in our discrete choice model, we divide both β_1 and β_2 by β_1 so that $\hat{\beta}_1$, the estimate of β_1, is equal to one. When neither of the two plans that we compare in our discrete choice model charge fixed fees, then $\hat{\beta}_1$ cannot be estimated.

If bill minimization theory is correct, then we would expect that b would be close to zero and that the coefficient β_2 would be extremely large. Furthermore, the econometric formulation in Equation 2.7 can also be used to explain plan choice under assumptions of certainty about future usage and perfect price information.

The reader should note that the variable b is an unknown. Instead of directly estimating this variable, we use it as a search parameter to maximize the log- likelihood function of our discrete choice model. We set the discrete choice problem as one of choosing between no plan and the plan in which we are interested. For example, for estimates of the coefficients for the AT&T *True USA* calling plan, we compared households that chose this plan to those that did not choose any plan. We ran a probit discrete choice model with increasing values for the variable b and stopped when the subsequent value of b yielded a log-likelihood function more than log-likelihood function that resulted from the previous value of b.[12] We use the value of $\hat{b}$ that yields the maximum log-likelihood function of our probit model as the estimate of b.

An adjustment to the variables p_i and p_j were required since these were not always directly observed. For example, when a household that subscribes to MCI Friends & Family calling plan made a fifteen minute call to another state on a Monday between 5:45 PM and 6:00 PM, the per minute charge for that call can be determined by knowing the characteristics of that plan that allowed us to calculate those charges. However, a majority of the time in a 24 hour day, no long-distance calls are made. It is difficult to determine what the effective price of a possible call would be if the phone had been used instead of left on the hook. We assumed that when we have minutes of non-usage, the relevant price associated with these observations is the distribution of prices that could have been charged had a call been made. With 11

[12] We can guarantee ourselves of a global maximum because of the concave properties of our net utility function.

distance bands and 3 time bands (or ratebands), there are up to 33 per minute prices possible. Rather than associating each non-use observation with a 33×1 vector of prices, we represent this vector by the mean of all possible prices. In other words, the price associated with any non-use observation is the mean of the distribution of possible prices. The reader should note that the mean of possible prices may be a function of the household's usage pattern since many callplans offer discounts based on the household's consumption.

2.5 Uncertainty Model

In this section we address our modelling of uncertainty.[13] The stochastic commodity model of Curien and de Fontenay [5] sheds significant light into optimization decision of telecommunications users under uncertainty. Under the stochastic commodity model, consumers are assumed to choose between alternatives based on an expected portfolio of calls varying duration, length of haul, and time of day. The expected portfolio governing consumer decisions is assumed to be determined taking into account all relevant rates, but actual usage may differ from expected usage due to random factors. Over time, the expected portfolio would be heavily based on prior experience, but usage in individual month could be very different from expected usage. The optimizing consumer would change the expected portfolio of usage in response to new prices and alternatives. Because consumers expect their usage to vary over time, they take into account the likely variability of usage and the financial risk associated with that usage variation.

Building on the ideas of Curien and de Fontenay, we assume that consumers examine several past telephone bills to ascertain their pattern of telephone usage. In order to cope with uncertainty, households use the average of their past usage as rule of thumb measure for future usage. Households are also aware of the variability in their usage pattern from month-to-month, and take usage variance into account when assessing the risk of having an unusually high telephone bill from one month to the next.

[13]See 2.A Appendix for a brief discussion of the imperfect information model. There we show that the uncertainty model and the imperfect information model, under our formulation, are essentially equivalent.

The rule of thumb measure is mean usage $\bar{x}_j$of previous periods. The household's guess is inaccurate by one plus a random variation in usage ε_j,where $\varepsilon_j \sim N(0, \sigma^2)$ and is assumed to vary independently of x_j. Expected utility under uncertainty can be written:

$$E[U_j] = \int f(\varepsilon_j)[\frac{1}{y}(y - p_j\bar{x}_j(1+\varepsilon_j) - F_j) + (\frac{\bar{x}_j}{b})(\ln\frac{\theta}{\bar{x}_j(1+\varepsilon_j)} + 1)]d\varepsilon_j \quad (2.8)$$

The term, $f(\varepsilon_j)$, is taken to be a frequency distribution for ε_j. All terms of $x_j \int \ln(1+\varepsilon_j)d\varepsilon_j$ are approximately zero in expectations when $E[\varepsilon_j] = 0$. We let:

$$\tau_j = \int \varepsilon_j f(\varepsilon_j) d\varepsilon_j$$

and

$$\delta_j = \int \varepsilon_j^2 f(\varepsilon_j) d\varepsilon_j$$

where δ_j is the variance of $\varepsilon_j \forall j$.

Optimal usage under the uncertainty model is:

$$\bar{x}_j^* = \theta \exp(-b\frac{p_j}{y}) - \frac{\tau_j}{1+\tau_j} - \frac{\delta_j}{1+\tau_j}$$

We can compare the choice between two plans under uncertainty from the equation derived above:

$$\begin{aligned} E[U_j] - E[U_i] = \; & \tfrac{1}{y}[(p_j\bar{x}_j + F_j) - (p_i\bar{x}_i + F_i)] + \\ & \tfrac{\bar{x}_j}{b}(\theta - \ln\bar{x}_j + 1)(1+\tau_j) - \tfrac{\bar{x}_j}{b}(\theta - \ln\bar{x}_i + 1)(1+\tau_i) - \\ & \tfrac{1}{y}[p_j\bar{x}_j\tau_j - p_i\bar{x}_i\tau_i) - \tfrac{\bar{x}_j}{b}(\tau_j + \delta) + \tfrac{\bar{x}_j}{b}(\tau_i + \delta)] \end{aligned}$$

We can implement the following econometric model of uncertainty:

$$\Pr[C_j]=\begin{cases} 1 \text{ if } \alpha_2+\beta_3(\frac{F_j-F_i}{y})+ \\ \beta_4[(\exp(-b\frac{\bar{p}_i}{y})-\frac{\tau_i}{1+\tau_i}-\frac{\delta i}{1+\tau_i}) - (\exp(-b\frac{\bar{p}_j}{y})-\frac{\tau_j}{1+\tau_j}-\frac{\delta_j}{1+\tau_j})] < 0 \\ 0 \text{ if } \alpha_2+\beta_3(\frac{F_j-F_i}{y})+ \\ \beta_4[(\exp(-b\frac{\bar{p}_i}{y})-\frac{\tau_i}{1+\tau_i}-\frac{\delta i}{1+\tau_i}) - (\exp(-b\frac{\bar{p}_j}{y})-\frac{\tau_j}{1+\tau_j}-\frac{\delta_j}{1+\tau_j})] > 0 \end{cases} \tag{2.9}$$

The uncertainty model above is similar to the certainty model except for the addition of the δ, τ, and the use of average minutes. Like the certainty model, there are two coefficients, β_3 and $\beta_4 = \frac{\theta}{b}$, and an intercept term, α_2, included in the model. The model predicts that the estimate of β_3, $\hat{\beta}_3$, be equal to one and that $\alpha_2 = 0$.

The reader should note the similarities of our way of modelling uncertainty to that found in the options value literature. The most significant difference between this model and the options value one is the fact that the mean and variance of past usage is a multiplicative term in our utility function. In the options value model, variance appears as an additive term and is therefore independent of actual usage.

2.5.1 Data

We use long-distance telephone service usage data found in the 1995 Call Detail Database. This database provides information on all long-distance calls at the household level. The database is a catalogue of all calls made by approximately 10,000 households for a period between October 1994 and June 1995. These households turned in their monthly telephone bills to PNR and Associates, and the data in these bills comprise the *Bill Harvesting II* Call Detail Database.

In order to estimate our own model of household choice of long-distance plan options, we use a subset of this large data set, where we have both demographic data and telephone usage data for 1,939 households. We have information for the 33,048 outgoing long-distance calls these 1,939 households made in the months of May and June of 1995. We use only these months because there exist a noticeable, and unexplained, drop in participation from households before and after these months. We choose May and June of 1995 because participation rates are most stable in these months.

Each observation in the Call Detail Database is a long-distance phone call. We specified our model on the basis of per minute net utility calculations. In other words, we assume that the consumer is reviewing his or her per minute charges and marginal utility from phone usage and deciding whether to initiate, continue or discontinue using the service. In order to implement our model, we weighted each telephone call observation by the number of minutes consumed on that call. We also divided yearly income by a factor of twelve so that monthly service fees and income were analyzed using the same units of time (in this case a month). Non-usage of long-distance telephone service was measured on a per minute basis. No usage is presumably a decision members of the household make based on expected returns from using the phone. When we observed that a household did not use the phone for a day, we weighted that non-use observation by 1,440 which is the number of minutes in a 24-hour day.[14] When a household used the phone to make a long-distance call lasting 30 minutes, we weighted use by a factor of 30. The rest of the day, where the phone was not in use, zero usage was weighted by a factor of 1,410.

We obtained the income data on the households for which we have long-distance telephone use information from *Bill Harvesting II* Aggregate Database.[15] Income information was collected from a randomly chosen subset of households that participated in the collection of monthly telephone usage data used to construct the *Bill Harvesting II* Call Detail Database. Most of this data was collected in the months of April, May and June of 1995.

[14] We divided these 1,440 into three ratebands. For example, in a weekday there are 540 minutes in the day rateband, 360 minutes in the evening rateband, and 540 minutes in the night rateband.

[15] This data set was also provided by PNR and Associates, Inc., of Jenkintown, PA.

2.6 Results

We conclude that the bill minimization hypothesis can be rejected outright. Recall that for this hypothesis to hold the variable that determines elasticity, b, would have to be close to zero and $\hat{\beta}_2$ would be a large number. Neither one of these predictions held. The results from the certainty model (Table 2-1) and the uncertainty model (Table 2-2) appear below.

Calling Plans	$\hat{\alpha}_1$	$\hat{\beta}_2$	$\hat{\beta}_1$	Elast. $(\frac{\hat{b}}{y})$	R^2
AT&T True USA	-0.643** (-17.976)	4.760** (18.110)	n.a.	-3.52	.0764
AT&T Reach Out	-1.959** (-11.810)	0.076* (1.737)	111.87** (2.695)	-4.27	.0297
AT&T Simple Savings	-1.211** (-24.547)	5.540** (7.731)	n.a.	-4.45	.0424
AT&T Select Saver	-1.428** (-5.065)	3.075** (2.434)	51.28 (0.426)	-1.41	.0150
AT&T Any Hour	-2.093** (-5.340)	0.1407 (0.364)	57.746* (1.839)	-4.65	.0162
MCI Friends & Family	-1.201** (-24.499)	5.038** (10.415)	n.a.	-3.86	.0633
MCI Easy Rate	-2.308** (-10.720)	2.810** (1.964)	25.84 (0.211)	-4.79	.0013
MCI Prime Time	-2.077** (-10.892)	0.102 (0.528)	22.71 (0.614)	-4.39	.0033
Sprint The Most	-2.016** (-6.500)	0.157 (0.425)	n.a.	-3.66	.0010
Sprint Plus	-1.974** (-24.557)	18.456** (4.479)	n.a.	-5.39	.0286
Sprint Day Plus	-1.736** (-9.284)	9.572** (4.137)	457.42** (1.858)	-4.95	.0738

Note: ** indicates that coefficient is significant at the 95% confidence interval. ** indicates that coefficient is significant at the 90% confidence interval. z scores appear in parentheses below coefficient estimates.

Table 2-1: Empirical Results of Certainty Model

Using the Pseudo R^2 as a measure of fit for discrete choice models, the uncertainty models for the 11 callplans perform better than the certainty ones. The Pseudo R^2 is higher for all of the uncertainty probit models for each of their respective certainty model probits. Given that the weighting made sample sizes large, there is concern as to how robust the significant estimates are. For this model, and the uncertainty one that follows, we found that changing magnitude of the variables by plus or minus 17.5% did not change the probability of choice for most of the models in which the coefficients are significant at the 95 percent confidence levels. We conclude that the significance levels of our coefficients are not the result of large sample size.

Calling Plans	$\hat{\alpha}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Elast. $\frac{\hat{b}}{y}$	R^2
AT&T True USA	-0.400** (-14.213)	0.080 ** (22.066)	n.a.	-1.62	.1771
AT&T Reach Out	-1.921** (-11.618)	0.050** (6.588)	100.31** (2.689)	-1.51	.0417
AT&T Simple Savings	-1.141** (-24.799)	0.019** (14.423)	n.a.	-1.50	.1238
AT&T Select Saver	-1.990** (-12.228)	0.008** (9.948)	6.92 (0.050)	-1.81	.0630
AT&T Any Hour	-1.490** (-13.922)	0.040** (10.439)	19.81* (1.255)	-2.71	.0785
MCI Friends & Family	-1.017** (-23.992)	0.080** (16.426)	n.a.	-1.21	.1384
MCI Easy Rate	-2.260** (-10.742)	0.002** (6.122)	33.63 (0.278)	-2.56	.0424
MCI Prime Time	-2.020** (-13.002)	0.040** (7.526)	32.71 (0.864)	-1.38	.0673
Sprint The Most	-2.112** (-22.272)	0.067** (6.442)	n.a.	-2.50	.0868
Sprint Plus	-1.926** (-24.178)	0.016** (9.240)	n.a.	-1.55	.0878
Sprint Day Plus	-1.580** (-8.459)	0.102** (8.500)	545.63** (2.148)	-1.11	.1180

Note: ** indicates that coefficient is significant at the 95% confidence interval. ** indicates that coefficient is significant at the 90% confidence interval. z scores appear in parentheses below coefficient estimates.

Table 2-2: Empirical Results of Uncertainty Model

For each model, we test how well they predict the take of a plan correctly. Using this measure, we find that the uncertainty model performed better. Only for those certainty models that yielded the highest Pseudo-R^2s, *AT&T True USA, MCI Family and Friends* and *Sprint Day Plus*, the model predicts take correctly 61.7, 58.9 and 62.8 percent of the time, respectively. These are the only three plans for which correct predictions were well above 50 percent for the certainty model. In comparison, the uncertainty model has several plans for which it correctly predicts above the 60 percent mark and only two for which prediction falls under 60 percent.

The certainty model yielded some unreasonable predictions compared with the uncertainty model. The average elasticities predicted by the certainty models are high (Examine Table 2-1, column 5) and beyond the expected range.[16] In comparison, the uncertainty model estimates average elasticity within the expected range of -1.0 to -3.0. A comparatively unreasonable range for predicted elasticities that the certainty model yields is another reason to favor the uncertainty model specification.

Furthermore, estimates of per minute satiation levels are well above 1, the limit, for 7 out of the 11 calling plans in the certainty model. The model predicts that θ should be well below 1. In contrast, estimates of θ in the uncertainty model are never above .102 (See Table 2-2).

The uncertainty model provides significant estimates for all 11 of the β_4 coefficients compared to only 7 significant β_2 coefficients for the certainty model. However, the uncertainty model does not perform any better than the certainty one with respect to number of significant estimates for the β_1 and β_3 coefficients. We conclude that the net utility gained from the difference in monthly fixed fees is insignificant in explaining the pattern of choice of calling plans.

Finally, our models predict that the intercept term should be zero. Note that the estimates for the intercepts, α_1 and α_2, are significant in both models for all calling plans. While

[16] Above we cited Taylor's industry standard of -0.5 price elasticity for local service. We expect that this average provides a floor for price elasticity in the long-distance service market, but estimate of elasticity above -3.5 seem very high.

the intercept term is generally lower for the probits of the uncertainty model, the intercept is consistently negative and within a well-specified range between -0.400 and -2.308. We interpret this finding as a bias in favor of the no plan default. In our probit model, the dependent variable was the plan to which the household was subscribed and on the right-hand side of the equation we had a net utility calculation; where utility from the default plan (choice=0) was subtracted from utility from the plan of the plan other than the default (choice=1). In setting up the probit model to equal zero when the no plan default was taken and to equal one when the other plan was chosen, we were able to determine that there exists a bias among households in favor of the no plan option. The consistently negative, and significant, intercept term implies that there is always a tendency to choose zero even when net utility calculation yields positive benefits. No fixed monthly fee is the consistent feature in the default plan, and its absence is likely to be the explanation for the no plan bias. We conclude that households are averse to paying a monthly fee, unlike the local telephone plan market, in the long-distance calling plan market.

In Table 2-3, we show estimates of per minute usage, by calling plan, for both the certainty and uncertainty models and compare these estimates to actual (average) per minute usage. The limit of per minute usage should be one, but the estimates for usage are above one for some plans in the certainty model since the satiation points the model predicted were also substantially over one. While the uncertainty model does better than the certainty model in estimating actual per minute usage, it consistently underestimates.

Calling Plans	Certainty Model	Uncertainty Model	Average Actual Usage
AT&T True USA	2.013	0.013	0.020
AT&T Reach Out	0.056	0.003	0.015
AT&T Simple Savings	1.840	0.012	0.018
AT&T Select Saver	0.107	0.005	0.018
AT&T Any Hour	0.001	0.008	0.019
MCI Friends & Family	1.603	0.011	0.020
MCI Easy Rate	0.062	0.006	0.015
MCI Prime Time	0.025	0.011	0.016
Sprint The Most	0.091	0.009	0.020
Sprint Plus	2.299	0.010	0.020
Sprint Day Plus	0.033	0.006	0.012

Table 2-3: Per Minute Estimates both Models and Actual Usage

In a previous section where we classified plans into three categories—flat rate, volume discount, and calling circles plan. Neither model performs better explaining the pattern of choice in any one of the sets of plans. For a majority of the plans that charged a fixed fee, the coefficient estimated for these components of net utility β_1 and β_3 were insignificant in both models. One might expect that the fixed fee component of a calling plan may be insignificant in explaining the pattern of choice when the fees are relatively small. Low fees are charged for the AT&T *Select Saver* ($1.90 per month) or AT&T *Reach Out* ($3.00 per month) plans and the estimated coefficient $\hat{\beta}_1$ and $\hat{\beta}_3$ are insignificant for both plans for both models, as might be expected. However, the estimate for these two coefficients are also insignificant in explaining choice for MCI's *Prime Time* and *Easy Rate* where the customer is charged $11.50 and $5.00 per month for subscription to these plans. We would expect that the consumer would take these fees into consideration since they are substantial in relation to the average monthly bill. This does not appear to be the case from the results of our model.

2.6.1 Bounded rationality versus inertia

Given that the uncertainty model did reasonably well in predicting take of calling plans, we may want to ask whether households saved substantially by using the rules of thumb implied by the uncertainty model. If households chose well using boundedly rational reasoning, are there other households like themselves, with similar calling patterns, that do not use the same rules of thumb for choosing their calling plan? If there are households such as these, does the inertia of staying put with the plan they have keep them from switching to more economical plans? We attempt to answer these questions in this section.

At the introduction of this paper, we provided some statistics on savings households could have obtained had they chosen the optimal cost minimizing plan. As an upper bound in savings, we stated that the sum of the savings from all 7,716 households was $1,468,949 out of $3,370,825 spent on long-distance calls throughout the period the survey. That amounts to approximately $38 per household per month for the 5 month period. This is an upper limit in savings since it required perfect information, instantaneous switching from call to call, and no transaction costs in order to achieve these savings. If we allow for uncertainty, and choose plans that can achieve the highest savings for an extended period of time, instead of from call to call, we find that household savings are substantially lower; approximately $7.82 per household per month. However, the per month savings figure, allowing for uncertainty, has a very high standard deviation and a large range; a standard deviation of $17.03 and an upper bound of $167.52. This means that the rules of thumb modeled by the uncertainty model specification do not always work well in maximizing overall savings across several periods. When we examine savings between households with no plans and households with plans, under uncertainty, we find that the levels of average savings across several months is not very different. However, the distribution of savings between the two groups is substantially different. Examine table 2-4. The distribution in savings supports the original hypothesis postulated by the uncertainty model; namely, households with higher variance in their calling patterns choose a calling plan because they are better off doing so.

Calling Plan	Mean	Std. Dev.	Frequency
No	6.09	5.17	2,703
Yes	8.18	24.43	5,013

Table 2-4: Average Monthly Savings of Households With and Without LDCPs

We explore further whether uncertainty plays a part in the choice of choosing a calling plan by specifying the same uncertainty discrete choice probit model that we used to examine the choice between no plan and a specific option plan. In this case, the choice will be between no plan and any option plan. We present the results of this probit model in Appendix 2.B. It should be of no surprise to find that the uncertainty model performs well under this specification since the data used to run this model is the aggregate of the sample data used previously to test each option plan choice. This test lends more credence to the assertion that uncertainty over future long-distance telephone calling consumption may lead some households to choose a calling plan to reduce variance in their monthly telephone bills.

While the previous test supports the assertion that uncertainty leads some households to subscribe to calling plans, there is another test which can help us determine whether inertia plays any roll at all in the choice of a LDCP. One way to do this is to examine the AT&T customer population and compare to the rest. If there are customers that would suffer from the inertia of their past, it would have to be these AT&T customers. As the established incumbent, it is likely that AT&T had to do nothing to get present set of customers to which it provides service. On the other hand, MCI and Sprint had to attract customers away from AT&T by enticing them away from what they were accustomed to doing. Since MCI and Sprint subscribers had to assert themselves to move away from AT&T, these customers are more likely to experiment with calling plans, study their calling patterns, compare bills from one company to the other, and use the rule of thumbs that were modeled by uncertainty. In other words, the uncertainty model would misspecify the behavior of AT&T customers, since inertia played a role in how they chose their plans, but would be appropriate for all others. In order to find out whether inertia, in place of uncertainty, plays more of a role in calling plan choice with AT&T customers than with customers of MCI and Sprint, we implement the same discrete choice probit model between plan and no plan but divide our data between AT&T customers and everybody else.

The results of the uncertainty model for AT&T customers and all others can be found in the second table of Appendix 2.B. If the hypothesis that AT&T customers are more inertia bound than others is true, then the uncertainty model would not have been the correct specification for choice between no AT&T plan and some AT&T plan for this company's customers. The results show no significant difference between AT&T customers and all other groups. Therefore, the hypothesis that AT&T customers are more inertia bound in their choice of plans, when compared to others, does not hold up in these data. We know turn to a look at the data by income groups to determine whether higher income groups are better informed, and therefore more likely to choose the correct plan, than lower income groups.

2.6.2 Income and uncertainty

In Appendix 2.C, we show the results of our test on the relationship between income and uncertainty. We hypothesized that income is negatively related to uncertainty. Higher income households, making more phone calls than others, would have greater payoffs in choosing the most optimal plan since they would have the greatest absolute savings in doing so. Because the correct calling plan choice may save a higher income household more money, they have incentives to be well informed about the calling plan options and to learn about their pattern of telephone usage. At the same time, higher income households also have higher opportunity costs of searching. Competing long-distance carriers' subscription campaigns may reduce higher search costs for heavy users of long-distance telephone services and therefore make the search cost argument less compelling.

In order to test our hypothesis about the relationship between income and demand uncertainty, we begin by dividing the households in our data set into four income groups—up to $17,500, between $17,501-$37,500, between $37,501-$52,500, and more than $52,500 per year. These groups contain approximately 475 households each. For each group, we ran the certainty and uncertainty models as we did before for all households. In order to make this task manageable, we limited ourselves to examining these two models, by income group, for only those plans for which the companies charged a monthly fixed fee. Therefore, we limit ourselves to six plans and four income groups for a total of 24 probits.

The results of the probits are unexpected. While we might expect that the uncertainty model would explain the pattern of plan choice for lower income households relatively well (and it does), our expectations were that the same would not be the case for the two higher income groups. However, uncertainty was significant in explaining the pattern of plan choice for both higher income groups. In examining the results of our probit models, we see that for Income Group 3 the uncertainty models perform significantly better (in terms of significant coefficients) in comparison to the certainty models. The same is also true for the models for Income Group 4. The coefficient that measures net utility from the difference in fixed, monthly fees, β_3, while insignificant in calling plan choice for the lower income groups, is also insignificant for many of the plan choices under certainty model for the higher income groups, but not so for the uncertainty model. This finding, on the pattern of significance of the β_3 coefficient in the uncertainty models of the higher income groups, lends some weight to the arguments that plans with fixed fees may be chosen as a way to minimize bill variability.

The data shows that higher income households have higher variability in their calling patterns, and this variability is likely to be the source of this group's uncertainty. Heavy long-distance service users, as defined above average consumption of minutes per week on the phone, also have the highest variance in their usage pattern (see Table below). We divided the data set into four groups based on their weekly usage categories. In examining the table, we see that the variance in weekly usage increases as the number of minutes increases as well. This indicates that large consumers of long-distance service one week are likely to consume substantially fewer minutes the next week. At the same time, households that use the telephone for a few minutes per week are likely to sustain that pattern.

Calls per Week	Mean	Std. Dev.	Frequency
1-7	9.55	5.90	24,149
8-16	33.62	7.99	22,737
17-29	70.21	14.00	22,583
30 or above	185.50	109.52	23,104

Table 2-5: Statistics of Minutes Consumed by Call Frequency Group

The findings contained in the Table above are consistent with the results of our uncertainty model.[17] Namely, incorporating uncertainty into the econometric model allows for more significant explanation of the pattern of calling plan choice among higher income groups (heavy phone users). Uncertainty is a significant explanatory factor in calling plan choice because heavy users of long-distance telecommunication services may find it harder to predict their usage pattern because of their high variance in week-to-week usage.

2.7 Conclusion

This study represents the first steps into research on consumer choice in the long-distance calling plan market. We began by examining the literature on plan choice in the local telephone service

[17] In 2.B Appendix, we show that high income households are the heaviest users of long-distance telephone services.

market. As in the local telephone service market, consumers are expected to make a plan choice before they consume services. Making choices *ex ante* introduces uncertainty based on the fact that consumers may poorly estimate their future demand. In the long-distance calling plan market, uncertainty may be exacerbated by the fact that long-distance bills can be substantial. A wrong calling plan choice, for those that use these services the most, can be a costly mistake.

While this research benefited from previous work in the local telephone service market, it was clear that none of the models presented could be applied directly to consumer choice in the long-distance market. The role of uncertainty of future demand and imperfect information of relevant prices feature prominently in explaining local telephone service options and we set out to determine whether the same was true for long-distance service.

We first reviewed option value models where consumers must guess their future consumption. Economists who use these models recognize that applying standard consumer surplus analysis to customer subscription decision-making will fail to capture *ex ante* willingness to pay some premium to guarantee a future level of consumption. The existence of an option value implies that a rational individual may choose to select a tariff even though the subscription price exceeds calculations of expected consumer surplus. Perhaps the most intriguing example of seemingly irrational consumer behavior is the overwhelming flat rate preference exhibited by local telephone users worldwide. This bias prevails despite the availability of apparent bill savings and consumers' surplus benefits, which would derive from the adoption of an alternative tariff choice.

We also reviewed local telephone service choice option and dynamic stochastic choice models. These models either simulated usage by agents under various calling options, or had the benefit of data in which households used telephone services under various calling plans. By having the benefits of this type of data, other papers were able to address the difficult problem of simultaneity that is found in telephone usage panel data. Simultaneity refers to the problem that telephone usage is a function of the calling plan to which the household is subscribed, and, at the same time, the choice of calling plan is presumed to be a function of prospective telephone usage. Therefore, it is difficult for the researcher to determine demand elasticities for a particular household unless the data set contains observations where the same household is consuming telephone services under various plans. We proposed a discrete choice model in which one of the variables was not directly estimated but used as a search parameter to search for the maximum of the log-likelihood function of the discrete choice model.

We incorporated aspects of the option value model and dynamic stochastic models to formulate our own model of service plan choice. We postulated a discrete model of long-distance calling plan choice as a net utility maximization problem. We had usage and demographic data for 1,939 households for the months of May and June of 1995. In our model, we did not have usage data of households using telephone services under various calling plans. As a result, we could not estimate some parameters of the net utility maximization problem directly, but rather had to use a search technique over the parameter that determined price elasticity to estimate the parameters of our probit model. We examined calling plan choice under the assumptions that consumers had full information about prices and demand, and compared this model to one of imperfect information about prices and uncertainty in a household's future demand for long-distance telephone services.

We conclude that uncertainty is important in explaining the choice of long-distance service calling plans. When comparing a discrete choice probit model containing variables in which uncertainty was accounted for, the uncertainty model performed substantially better than the model in which full information and demand certainty was assumed. Furthermore, the econometric models showed a bias on the part of households to choose the no plan option. By choosing not to enroll in a call plan, it is likely that households are discouraged by the monthly enrollment fees that most plans carry. As a result, some households underestimate their future long-distance telephone usage and choose sub-optimally recalculating the monthly bill under a cheaper plan. Finally, we showed that uncertainty may be as prevalent in the demand for long-distance telephone service among all income groups. We had expected that higher income groups, since they would gain most from choosing correctly, would invest more in reducing uncertainty. However, high income households have the most variability in their calling patterns, making it comparatively more difficult to reduce uncertainty.

Bibliography

[1] Atherton, T., M. Ben-Akiva, D. McFadden, and K. Train, Micro-simulation of local residential telephone demand under alternative service options and rate structures, in A. DeFontenay, M. Shugard and D. Sibley (eds.), *Telecommunications Demand Analysis: An Integrated Approach*, North Holland: New York, NY, pp. 137-163 (1990).

[2] Brandon, B.B. (Ed.), *The Effects of the Demographics of Individual Households on their Telephone Usage*, Cambridge, MA: Ballinger Publishing (1981).

[3] Clay, K. D. Sibley and P. Srinagesh, *Ex post vs. ex ante pricing: Optional calling plans and tapered tariffs,* Journal of Regulatory Economics **4**(2), pp. 115-138 (1992).

[4] Colton, R. *Consumer information and workable competition in telecommunications markets,* Journal of Economic Issues **27**(3), pp. 775-792 (1993).

[5] Curien, N. and A. DeFontenay, Stochastic commodity model, in A. DeFontenay, M. Shugard and D. Sibley (eds.), *Telecommunications Demand Analysis: An Integrated Approach,* North Holland: New York, NY, pp. 119-136 (1990).

[6] Federal Communications Commission, *Statistics of Communications Common Carriers,* FCC: Washington, DC (1996).

[7] Hartman, R.S. and Z.F. Naqvi, *Estimation of household preferences for long distance telecommunications carrier,* Journal of Regulatory Economics **6**, pp. 197-220 (1994).

[8] Hobson, M. and R. Spady, *The demand for local telephone service under optional local measured service,* Bellcore Technical Memorandum, TM-ARH-011423, Summer (1988).

[9] Huber, P.W., M.K. Kellogg, and J. Thorne, *The Geodesic Network: 1993 Report on Competition in the Telephone Industry,* US Department of Justice/Antitrust Division, GPO: Washington, DC (1993).

[10] Kling, J.P. and S.S. Van der Ploeg, Estimating local call elasticities with a model of stochastic class of service and usage choice, in A. DeFontenay, M. Shugard and D. Sibley (Eds.), *Telecommunications Demand Analysis: An Integrated Approach,* North Holland: New York, NY, pp. 119-136 (1990).

[11] Knittel, C.R., *Interstate long distance calling rates: Search costs, switching costs, and market power*, Review of Industrial Organization **12**, pp. 519-536 (1997).

[12] Kridel, D.J., *A consumer surplus approach to predicting extended area service (EAS) development and stimulation rates*, Information Economics and Policy **3**, pp. 379-390 (1988).

[13] Kridel, D.J., D.E. Lehman, and D.L. Weisman, *Option value, telecommunications demand, and policy,* Information Economics and Policy **5**, pp. 125-144 (1993).

[14] Littlechild, S.C., *Two-part tariffs and consumption externalities,* The Bell Journal of Economics and Management Science, **6**(2), pp. 661-670 (1975).

[15] MacAvoy, P.W., *The Failure of Antitrust and Regulation to Establish Competition in Long-Distance Telephone Services,* Cambridge, MA: MIT Press and AEI Press (1996).

[16] Maddala, G.S., *Limited-Dependent and Qualitative Variables in Econometrics,* Econometric Society Monograph No. 3, Cambridge, UK: Cambridge University Press (1994).

[17] Mitchell, B.M. and I. Vogelsang, *Telecommunications Pricing: theory and practice,* Cambridge, UK: Cambridge University Press (1991).

[18] Park, R., B. Wetzel, and B. Mitchell, *Price elasticities for local telephone calls*, Econometrica **51**, pp. 1699-1730 (1983).

[19] Pratt, J.W., *Risk aversion in the small and in the large*, Econometrica, **32**, pp. 122-136 (1964).

[20] Schmalensee, R., *Option demand and consumers' surplus: Valuing price changes under uncertainty,* American Economic Review, **62**, pp. 813-24 (1972).

[21] Squire, L., *Some aspects of optimal pricing for telecommunications,* The Bell Journal of Economics and Management Science, **4**(2), pp. 515-525 (1973).

[22] Srinagesh, P., *A dynamic stochastic model of choice,* Bellcore Technical Memorandum, TM-ARH-021364, Summer (1992).

[23] Tardiff, T.J., *Effects of presubscription and other attributes on long-distance carrier choice,* Information Economics and Policy, **7**, pp. 353-366 (1995).

[24] Taylor, L.D., *Telecommunications Demand: A survey and critique, Second Edition,* Ballinger: Cambridge, MA (1994).

[25] Train, K.E., D.L. McFadden, and M. Ben-Akiva, *The demand for local telephone service: A fully discrete model of residential calling patterns and service choices*, Rand Journal of Economics **18**, pp. 109-123, (1987).

[26] Train, K.E., *Optimal Regulation: The Economic Theory of Natural Monopoly,* Cambridge, MA: MIT Press (1996).

[27] Train, K.E., *Self-selecting tariffs under pure preferences among tariffs,* Journal of Regulatory Economics **6**(3), pp. 247-264 (1994).

[28] Weisbrod, B.A., *Collective Consumption Services of Individual Consumption Goods,* Quarterly Journal of Economics, **78**, pp. 471-477 (1964).

[29] Wilson, R.B., *Nonlinear Pricing,* Oxford University Press: Oxford, UK (1993).

2.A Appendix

Imperfect information model Given the complexity of how long-distance telephone calls are priced and packaged in plans, consumers of long-distance calling plans usually have imperfect about the relevant rate for these telecommunications services. Households do not have detailed rates and tariffs available to compute the exact choice of plan that minimizes their telephone bill. Another possible issue is whether consumers respond to average price or marginal price of services. The difference between average and marginal prices may be quite large in principle. Because consumers may not know which tariff applies to a particular call, many of them respond to the average cost across all calls rather than to the true cost of each call.

Average prices may be a poor choice on which to base a decision on a plan. For example, if consumers takes their monthly long-distance bill and divides this by the number of minutes that they consumed that month, the average price of a call will be negatively correlated with the consumer's level of usage, especially when the plan has a high monthly fixed fee.

By including the average cost across all calls as a price variable, it is possible to take into account the extent of a consumer's imperfect information about the true price of a call under the consumer's chosen calling plan. Like uncertainty, we can use a rule of thumb measure that a household may use to determine the marginal price of their long-distance service. The rule of thumb measure could be mean price calculated by taking the monthly bill and dividing it by the total number of minutes consumed. The average price $(\bar{p}_j)$ of previous periods, and the household's guess is inaccurate by one plus a random variation ε_j,where ε_j is assumed to randomly vary independently of $\bar{p}_j$. The term, $f(\varepsilon_j)$ is taken to be a frequency distribution for ε_j. Then expected utility can be written:

$$\int f(\varepsilon_j)[\frac{1}{y}(y - \bar{p}_j(1+\varepsilon_j)x_j - F_j) + (\frac{x_j^*}{b_j})(\ln\frac{\theta}{x_j^*} + 1)]d\varepsilon_j$$

It can easily be shown that this formulation is equivalent to the one where we assumed a rule of thumb measure of uncertainty over usage. Once we substitute x_j^* for the relationship in Equation 2.3, we can see that $\bar{p}_j$ always appears as its multiplicative so it is like multiplying $(1+\varepsilon_j)$ by $\bar{p}_j$ is equivalent to multiplying it by x_j which is how we formulated the uncertainty case.

2.B Appendix

$\hat{\alpha}$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\frac{\hat{b}}{y}$	R^2
-1.052^{**} (−13.32)	0.027^{**} (15.148)	53.25^{*} (1.487)	−1.45	.1981

Table 2.B-1: Choice Between Plan and No Plan

Customers	$\hat{\alpha}$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\frac{\hat{b}}{y}$	R^2
AT&T	-1.903^{**} (−12.76)	0.083^{**} (17.774)	63.25^{**} (2.356)	−1.86	.1430
All Others	-2.155^{**} (2.356)	0.072^{**} (9.517)	147.22^{*} (1.702)	−2.11	.1321

Table 2.B-2: Choice Between Plan and No Plan Among Customer Groups

2.C Appendix

The table below supports our assertion that long-distance telephone use is correlated with income. The cells in the table below denote the percent of observations found in group define by income group, the rows, that made a certain number of calls per week (the columns). For example, in row 1 column 1 we find that households with a per year income of up to $7,500 made between 1-7 long-distance phone calls per week. The row total shows the percent of all households, out of all households for that category of telephone usage. The Pearson chi^2 statistic is significant at 9 degrees of freedom and tells us that there is no independence between the distribution of usage categories and income, i.e., there exists a correlation between income and usage. From this table it is clear that it is a positive correlation.

Income per Year	1-7	8-16	17-29	≥30	*Total*
Up to $17,500	30.92	25.29	23.12	20.67	100.00
$17,501-$37,500	28.26	25.16	23.58	22.99	100.00
$37,501-$52,500	24.66	25.49	24.05	25.79	100.00
More than $52,500	19.88	21.98	27.16	30.98	100.00
Total	26.09	24.56	24.39	24.96	100.00

Pearson chi^2(9)=1202.2550 Pr=0.000

Table 2.C-1: Minutes per Week by Frequency of Calls and Income

Part III

Bargaining in Networks

Chapter 3

Applied Theory

SUMMARY

We examine pricing strategies of network monopolists seeking two-way interconnection. We design a bargaining protocol that allows firms to negotiate reductions in interconnection tariffs down, from monopoly levels, to levels where tariffs are set equal to the marginal cost of interconnection. The bargaining protocol achieves weakly Pareto efficient, individually rational bargaing outcomes. Getting firms to bargain to the effiicient point, where tariffs are equal to marginal costs, is a function of the relative sizes of retail markets that monopolists serve.

Next, we empirically test our hypothesis that more efficient bargaining is achieved when firms are symmetric. Using interconnections tariffs and traffic data for US international telecommunications carriers, we find that the data supports our hypothesis. Furthermore, we find the efficient bargaining frontier and characterized the firms with which US international telecom carriers are able to negotiate efficient interconnections tariffs.

3.1 Introduction

In recent years, thinking on how to regulate network monopolies changed dramatically. In the past, regulation focused on monitoring and directing all functions and components of network industries. Today, there is growing consensus that instead of regulating the complete network, it is best to disentangle those components in networks that are monopolistic and must remain regulated, from those components that are competitive and can be left unregulated.[1]

In order to successfully implement this new thinking in regulatory policy, economists and regulators must address the fact that some components of the network may remain under the control of a monopolist, undoing the benefits of promoting competition in other parts of the network. Monopoly control over a component of the network will result in inefficient pricing of the good or service provided to the newly formed competitive components of non-integrated network industries. For example, deregulation of the energy sector must cope with the monopolistic control of the distribution system. Energy distributors can charge high prices to firms competing in power generation because they control access to businesses and residences that buy energy. International telecommunications provides another example of monopoly controlled bottlenecks. While competition may be present among international telephone service providers domestically, interconnection services required to terminate the call in a foreign country are most likely sold by monopolists. High, inefficient pricing of interconnection services in the international telecommunication sector are substantial, and payments to compensate for interconnection services represent a large portion of a US carrier's operating costs. In 1996, US telecoms paid foreign carriers $5.6 billion for terminating international calls. These payments for termination services represent approximately 61 percent of the $9.2 billion in total revenue that US carriers received for providing international telecommunications services.[2]

Interconnection charges are set high in most foreign countries because these services are most often provided by government-owned monopolies. Government-owned monopolies in foreign countries can get away with charging exorbitant fees partly because the US cannot impose regulations to bring monopoly charges down. Given the jurisdictional problems, the solution to bringing foreign monopoly prices down must rely on a decentralized mechanism that induces these monopolies to reduce interconnection charges on their own and not regulation.

[1] Witness the US Telecommunications Act of 1996.

[2] Federal Communications Commission (1998), *Trends in the U.S. International Telecommunications Industry*, Industry Analysis Division, Common Carrier Bureau (June). p. 65.

The purpose of this paper is to determine whether network industry monopolists can be induced not to charge the monopoly prices for interconnection to the network that the monopolist controls. Interconnection replaces the need to establish and maintain complete, stand alone networks. In particular, we are interested in the case of two-way interconnection where firms in a network buy interconnection services from each other. The international telecommunications is an example of two-way interconnection. For industries that use networks as means of distribution, two-way interconnection is an effective way for all firms to extend their offerings to markets beyond each firm's local market. In the international telecommunications example, the use of two-way interconnection means that US telecommunications carriers do not have to build and maintain their own telephone system in foreign countries in order to provide international telephony to homes and businesses in the US.

Our paper is based on the findings and conclusions from two sets of literature on network interconnection regulation and two-way interconnection. The first set of papers in the literature address how retail price regulation affects monopoly pricing of interconnection services. The second set of papers focuses on incentives of interconnection pricing of two-way interconnection between monopolists. The results of retail price regulation are pessimistic; if regulations attempt to apply downward pressure on profits in the retail market, this pressure will result in markups on interconnection tariffs to make up for losses. Laffont et al. [12] show that identical firms can coordinate to set higher interconnection charges, using two-part tariffs, when regulation is applied to both firms at the retail level. Carter and Wright [4] use linear interconnection tariffs and achieve the same conclusions. Citing Laffont et al.'s negative results, Armstrong [1] looks at the optimal role of the regulator in attempting to reduce monopoly tariffs, and also finds that identical firms can circumvent regulators by agreeing on high interconnection charges.

Two conditions facilitate coordination in setting the appropriate mark up in response to regulation: a. the existence of identical, symmetric, firms, and b. two-way interconnection. Carter and Wright [4] wrote a paper on the incentives that symmetric monopolists have to eliminate two-way interconnection charges amongst themselves. In their model, as in ours, identical demand in the retail market defines symmetry between firms. Carter and Wright [4] showed that if symmetric monopolists can coordinate their interconnection tariffs, they will agree not to charge each other for interconnection services. The results for identical monopolists raises the questions as to whether the mutual incentives to eliminate interconnection tariffs also hold for non-identical, or asymmetric monopolists, and what kind of coordination is required to achieve the elimination of tariffs altogether.

We build on Carter and Wright's results that monopolists have incentives to reduce interconnection charges in two-way interconnection when they are symmetric. We begin our study by examining the relationship between relative firm size and incentives to reduce interconnection tariffs. We provide an extension of the Carter and Wright results and find that the elimination of interconnection charges is not a Subgame Perfect Equilibrium for asymmetric firms. In particular, it is not incentive compatible for smaller firms to eliminate interconnection tariffs.

The question then becomes whether there exists a mechanism that can induce, as much as is individually beneficial, smaller firms to reduce tariff levels. We model structured bargaining as a mechanism to coordinate agreements on reducing or eliminating interconnection service charges. A structured bargaining process, where rules determine how firms reach agreement on the level of interconnection charges, allows us to achieve outcomes that reduce tariffs and improve overall profits. We find that even with structured bilateral bargaining between asymmetric firms, it is impossible to obtain the elimination of interconnection service tariffs altogether.

We enhance our analysis of bargaining by introducing more than two monopolists, and these monopolists bargain for fully interconnected networks. When networks are fully interconnected, firms can access a point outside their own network either directly or indirectly through a combination of connections with other firms. Monopolists may want to compete in the interconnections market by offering lower cost indirect interconnection services by reducing monopoly interconnection tariffs on a particular path in the network. However, the question of what rule the monopolist uses to price indirect interconnection services is unanswered. We examine two proposals that policymakers at present discuss as alternative rules for pricing indirect interconnection services in the telecommunications industry; bill and keep defended by Brock, and the efficient component pricing rule of Baumol and Willig.

The results of this paper can be summarized succinctly below:

- The level of efficiency, i.e., how close firms get to eliminating tariffs from bilateral bargaining between asymmetric monopolists, depends on the extent of the asymmetries between monopolists. Bilateral bargaining between symmetric monopolists will lead to the most efficient outcome, the elimination of interconnection tariffs. However, bilateral bargaining between sufficiently asymmetric monopolists will not lead to the elimination of tariffs and is therefore a less efficient outcome.

- Empirical evidence supports the claim that the differences in firm size determines the level at which interconnection tariffs are set. We provide an analysis of traffic flows and accounting rates (interconnection tariffs in the argot of the international telecommunications industry) and find that US international telecommunications carriers were able to negotiate lower tariffs with their foreign counterparts when the telephone traffic flows were more even between countries; and

- The larger firm must be able to propose take-it or leave-it offers in order to achieve the highest level of efficiency that makes neither of the asymmetric monopolists worse off. The take-it or leave-it offer is characterized by the larger firm's elimination of interconnection tariffs in the smaller firm's market, and a positive interconnection tariff that the larger firm must pay.

In the following section, we introduce the model that Carter and Wright developed. We formally model the bargaining process suggested by their paper and extend their results to asymmetric firms. In Section 3, we extend the Carter and Wright framework by introducing bargaining between two monopolists, provide empirical predictions based on these extensions, and test these predictions using data from the US international telecommunications industry. Section 4 contains conclusions and policy recommendations. We also provide several suggestions for further research (empirical and theoretical extensions) that are not only tenable, but interesting and important theoretically and for policy purposes as well.

3.2 Background

The first part of this paper summarizes work by Carter and Wright [4]. We briefly present their model and results, and use these as a basis for extensions. In our extensions we do three things:

- Formally model the bargaining process suggested by Carter and Wright;
- Introduce the solution to the bargaining game over interconnection tariffs for symmetric and asymmetric firms; and
- Establish bargaining rules that result in more efficient negotiation outcomes between asymmetric firms.

Carter and Wright begin by assuming that two monopolists sell each other two-way interconnection services. These monopolists must determine how much to charge each other for these services and how much they will charge consumers for the final product. Firms have full information about demand in all markets.

Firms are involved in a two-stage, sequential game. In the first stage of the game, monopolists set tariffs for interconnection services. Since payment to the other firm in the form of interconnection service tariffs is the only cost in this model, the setting of these tariffs is equivalent to having each monopolists set the other's marginal cost. Once the two firms take into account interconnection service charges, both monopolists set retail prices.

We present modified **Assumptions** that Carter and Wright made in their own paper:

A 1. Tariffs and prices are restricted to be linear where t_{ij} is the tariff charged by Firm i to Firm j, and p_{ij} is the retail price Firm i charges for the good requiring interconnection services from Firm j;

A 2. Each firm's demand for the interconnection services depends only on consumer demand for final goods in the market supplied by that firm;

A 3. Consumer demand, q_{ij}, is linear in b, that is $q_{ij} = (a_{ij} - bp_{ij})$. The variable b determines price elasticity of demand. The slope of demand schedules (b) is assumed to be the same for all final goods markets. The variable a_{ij} determines the size of the market for the final good.

A 4. The cost of providing interconnection services is zero.

A 5. Firms are monopolists in interconnection services and final goods markets.

From now on, final products will be referred to as *goods* and interconnection services will be referred to as *services*. Conversely, the charge for goods is referred to as a *price* (p) and service charge will be known as a *tariff* (t).

The firms' profit functions can be written as:

$$\pi_i = p_{ij}(a_{ij} - bp_{ij}) + t_{ij}(a_{ji} - bp_{ji}) - t_{ji}(a_{ij} - bp_{ij}) \text{ for all } i, j,\ i \neq j. \tag{3.1}$$

The first term of the profit function for Firm i is revenue received from the sale of the final good. The second term represents revenue from the sale of the interconnection services to Firm j. Finally, we subtract the cost of purchasing interconnection services from Firm j to obtain the firm's profit.

When the firm maximizes profits with respect to tariffs and prices, this optimization problem has in the following solution:

Definition 1 *The Subgame Perfect Nash Equilibrium for Firm i is defined by the pair* $(t^*_{ij}, p^*_{ij}(\mathbf{t}))$

where:

$$p^*_{ij}(\mathbf{t}) \in \arg\max_{p_{ij}} \pi_i(p_{ij}, t_{ij}, t_{ji}) \textit{ is defined for } \mathbf{t} = (t^*_{ij}, t^*_{ji})$$

and

$$t^*_{ij} \in \arg\max_{t_{ij}} \pi_i(p^*_{ij}(\mathbf{t}), t_{ij}) \textit{ for all } i, j, \textit{ where } i \neq j.$$

A Subgame Perfect Nash equilibrium means that prices and tariffs are optimal at any stage of the game and comprise the best overall response to the other firm's price and tariff setting strategy. Because prices are set after tariffs are decided, prices are a function of tariffs. In practical terms, this means that the prices consumers pay reflect the level of tariffs that each firm sets; costs are passed on to the consumer.

Given the sequential nature of the game proposed, we solve the monopolist's profit maximization problem through backward induction; for prices first and tariffs last. In particular, the subgame perfect equilibrium in tariffs and prices is:

$$t^*_{ij} = \frac{a_{ji}}{2b} \text{ for all } i, j, \ i \neq j. \tag{3.2}$$

$$p^*_{ij} = \frac{a_{ij}}{2b} + \frac{t^*_{ji}}{2} \text{ for all } i, j, \ i \neq j. \tag{3.3}$$

Note that firms set tariffs at the monopoly level (t^*) and retail prices at the double marginalization level (p^*).[3] The double marginalization level occurs when, first, an upstream monopolist sets the interconnection service tariff at the monopoly level. A downstream monopolist passes

[3] If the game were structured differently, where prices are determined first and tariffs last, the optimal retail prices would be the monopoly price and the tariff would be zero.

on the cost of interconnection directly on to the representative consumer by incorporating this tariff into the final price. Second, the monopolist at the retail level sets a monopoly price for the final good in addition to passing the upstream monopolist's interconnection tariff. The representative consumer is charged the high cost of the inputs tariffs set at the monopoly level, and monopoly prices are added to this at the retail level.

While there is reason to be pessimistic for consumers when monopolists are involved in upstream and downstream production, Carter and Wright suggest that in the case where monopolists are identical and require two-way interconnection, there exists an equilibrium where tariffs are zero and prices are set at the standard monopoly level. Formally, the solution with no tariffs is:

$$p'_{ij}(\hat{\mathbf{t}}) \in \arg\max_{p_{ij}} \pi_i(p'_{ij}, t'_{ij}, t'_{ji}) \text{ for } \hat{\mathbf{t}} = (t'_{ij}, t'_{ji}) = 0$$

In the linear case,

$$p'_{ij}(\hat{\mathbf{t}}) = \frac{a_{ij}}{2b} \text{ for all } i, j,\ i \neq j.$$

The elimination of interconnection tariffs is individually rational. Identical firms demand the same level of interconnection services from each other. Whatever Firm i pays Firm j for interconnection services, Firm i will receive in tariff revenue for service provided to Firm j. However, because prices are set at the double marginalization level, the quantity demanded is relatively low in both markets. Market demand is lower than it would be if prices were reduced. In addition, prices are set at the double marginalization level and are set at the elastic part of the demand curve, where a reduction in price will actually generate more revenue than the previous level. Therefore, if both monopolists agree to reduce tariffs, this reduction would lower final good prices, induce greater quantity of the goods demanded, and result in an increase in each firm's total revenue. An increase in total revenue is assured because all prices are set at the elastic part of the demand curve. That means that revenue losses that result from lower retail prices are compensated by larger increases in the quantity demanded that result from price reductions. The reductions in tariffs result in higher revenues from the final goods market, so tariffs will be reduced until they are eliminated, and each firm will earn higher profits.

The elimination of interconnection tariffs is also Pareto efficient. Pareto efficiency is maximized when tariffs are zero because this outcome eliminates double marginalization, resulting in higher profits for firms and lower prices for consumers. The identical firm case is interesting because the point that defines optimal global efficiency also defines what both firms would like to obtain. The goal of the social planner, as well as goal of profit maximizing firms, converge on eliminating tariffs. As a result of this convergence in goals, almost any coordinating mechanism that allows bargains in which tariffs are eliminated will result in the most efficient agreement.

We demonstrate the point that there exists convergence between individual rationality and Pareto efficiency in the symmetric firm case with an illustration of profit possibilities frontier in Figure 3-1. In this figure, let the schedule of profits $\Pi^{'}$ denote the profit possibility frontier where the firms i and j can choose any prices while tariffs are set to zero. Let Π^{*} denote the profit possibility frontier where the firms choose any tariffs while prices are chosen noncooperatively given the tariffs. In other words, the second profit possibility frontier maps Nash equilibrium prices in the profit space. Let B be the point where joint profits are maximized on the profit possibility frontier where tariffs can be set at any level. In the case of symmetric firms, Carter and Wright show that both profit possibility frontiers are maximized at the same point, B, where a hyperplane on the profit space is tangent to each of the profit frontiers. The point N, in Figure 3-1, is the Nash equilibrium double marginalization outcome that lies in the interior of the profit frontier schedule. The line extending from N to B shows the set of points that are Pareto improvements, meaning increases in profits for both firms. Any point above N, along that line, is preferred by both firms. Because B also lies on the schedule of Nash equilibrium prices, it is individually incentive compatible for each firm to set tariffs at levels where tariffs are zero if it could be coordinated.

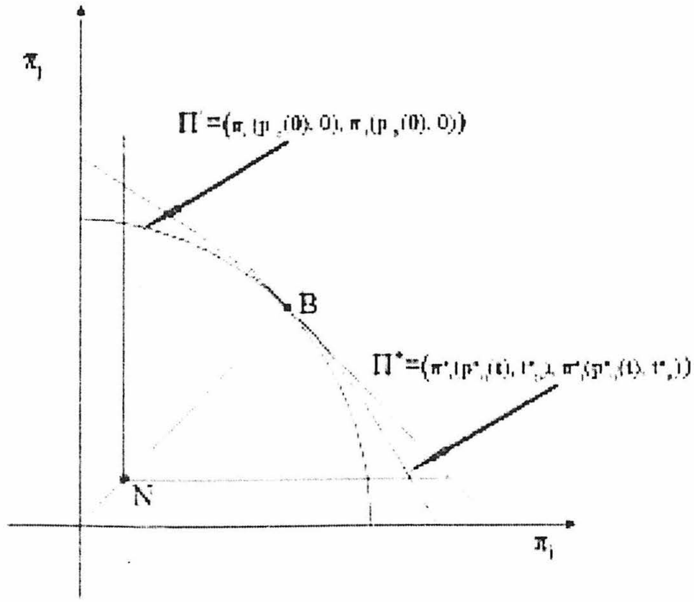


Figure 3-1: Profit possibilities for symmetric firms

The profit possibility frontiers for asymmetric firms are illustrated in Figure 3-2. We know that there exist linear tariffs that imply the joint profit maximizing prices (with independent pricing) and achieve the same total profit, although with a different distribution between the firms. Hence points M and E must lie on the same hyperplane, though in general they do not coincide. It follows that the profit possibility frontiers under regimes Π^* and Π' will only partially overlap as is shown in Figure 3-2. While both firms would benefit from a zero tariff regime, interest in the two alternative regimes conflict. In other words, the point N may be better or equally profit for either one of the firms in comparison to profits at point M. For example, if one firm is very much smaller than the other, it might do better by exploiting the bigger firm's demand at the Nash equilibrium N than cooperating over tariffs at point M.

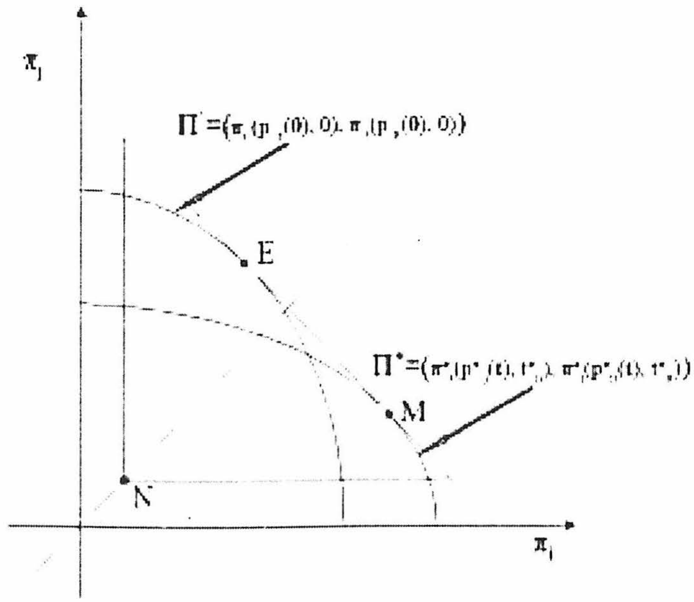


Figure 3-2: Profit possibilities for asymmetric firms

Carter and Wright's results show that firms have incentives to eliminate double marginalization in two-way interconnection, but their results hinge on two assumptions. The first assumption is that firms have to be identical in size in order to want to eliminate interconnection tariffs. The second assumption is that firms can coordinate, somehow, to mutually reduce tariffs. We now drop the first assumption and model bargaining as the coordinating mechanism.

3.3 Bilateral Bargaining

Carter and Wright established that identical monopolies have incentives to negotiate on reductions in tariffs for interconnection services. However, they did not model the bargaining process per se. In this section, we explicitly model the bargaining process and examine the case of bargaining over interconnection service tariffs when there exists asymmetries between firms.

We define asymmetry by referring to the relative sizes of the market for final goods that each firm services. The size of the market is represented by a_{ij} which determines demand for the final good. Asymmetry implies that $a_{ij} \neq a_{ji}$. We sometimes use the shorthand "symmetric firms" when we mean that the demand schedules for final goods are the same for both firm. Conversely, we use "asymmetric firms" in the opposite case.

3.3.1 Bargaining

Formally, we define the two-firm bargaining problem consisting of a pair (F, v) where F is the set of *feasible payoff allocations* (profits), and v represents the *disagreement payoff allocation.* In the two-firm bargaining game, F is a closed convex subset of $\Re^2$, and $v = (v_i, v_j)$ is a vector in $\Re^2$. The bargaining game, Γ, can be formally represented by:

$$\Gamma = (N, (C_i)_{i \in N}, (\pi_i)_{i \in N})$$

Where $i, j \in N$ and $i \neq j$, and N is the set of firms. The variable C_i is the set of strategies for Firm i and, consistent with its use above, $\pi_i : C \rightarrow \Re^2$ is the profit function for Firm i. In this game, the set of strategies are comprised of pairs of tariffs and prices, $(t_{ij}, p_{ij}(t_{ji}))_i \in C_i$ which satisfy the equilibrium conditions of the game. For the two-firm bargaining game, we can write:

$$\Gamma = (\{i, j\}, (C_i)_{i \in N}, (C_j)_{j \in N}, (\pi_i)_{i \in N}, (\pi_j)_{j \in N});\ i, j \text{ where } i \neq j$$

Disagreements points are defined by the minimax value of the set of possible profits which are the Subgame Perfect Nash Equilibrium profits. Formally we write disagreement points in the following way:

$$
\begin{aligned}
v_i &= \min_{(t_{ij},p_{ij})_i \in C_i} \max_{(t_{ji},p_{ij})_j \in C_j} \pi_i((t^*_{ij}, p^*_{ij}(t^*_{ji}))) \\
v_j &= \min_{(t_{ji},p_{ji})_j \in C_j} \max_{(t_{ij},p_{ij})_i \in C_i} \pi_j((t^*_{ji}, p^*_{ji}(t^*_{ij})))
\end{aligned}
$$

There are many bargaining protocols that could be imagined to solve the bargaining problem in such a way as to achieve:

- **Individual rationality** - The bargaining outcome where each firm is at least as well off under the Subgame Perfect Equilibrium as it would be if bargaining breaks down; and
- **Weak Pareto efficiency** - The bargaining outcome where negotiated tariffs are lower than those set by the Subgame Perfect Nash solution.

In proposing a bargaining solution, we follow the Nash program and provide an axiomatic approach. The axioms used to determine the proper solution to the bargaining problem will satisfy individual rationality and weak Pareto efficiency constraints. We provide definitions to the individual rationality and weak efficiency axioms:

Axiom 1 *An allocation, $\phi(F,v) \geq v$, is individually rational where $\phi(F,v) = (\phi_i(F,v), \phi_j(F,v))$.*

This axiom states that each firm will strictly prefer a bargaining outcome that yields higher profits to one that yields the same level or lower.

Axiom 2 *An allocation, $\{(\pi_i, \pi_j) | \pi_i > v_i$ and $\pi_j = v_j\} \in F$, is weakly Pareto efficient iff there exists no $\pi'_i \in F$ such that $\pi'_i > \pi_i$ and there exists a $\pi'_j \in F$ such that $\pi'_j \geq \pi_j$.*

For weak efficiency to hold, the sum of both firms' profits increase with a drop in tariffs. When tariffs are reduced, consumer welfare goes up as well. Knowing that reductions or increases in consumer welfare follow a fall and rise in the sum of firms' profits, we can use firm profits as a measure of welfare for the economy as a whole. We therefore use increases in firms' profits as a proxy measure for Pareto efficiency because of the monotonic relationship between profits and consumer surplus (CS). Most importantly, a measure of the sum of firm profits allows us to rank bargaining outcomes and determine which tariff-setting deal is the most efficient.

We formally present the arguments about the relationship between firms' profits, consumers' welfare, and tariff levels with the following claim and its corollary:

Claim 1 $\Pi(\hat{t}) \geq \Pi(t')$ *if and only if* $CS(\hat{t}) \geq CS(t')$.

Where we define $\Pi(\cdot)$, the sum of firms' profits, as:

$$\Pi(t_i, t_j) = \pi_i(t_i, t_j) + \pi_j(t_i, t_j);\ i, j = 1, 2,\ i \neq j$$

and the $CS(\cdot)$, the sum of consumers' welfare in each market, as:

$$CS(t_i, t_j) = cs_i(t_i, t_j) + cs_j(t_i, t_j);\ i, j = 1, 2,\ i \neq j$$

Corollary 1 $(t_i^*, t_j^*) \in \underset{t_i,\, t_j}{\arg\max}\ \Pi\ (t_i, t_j)$ *if and only if* $(t_i^*, t_j^*) \in \underset{t_i,\, t_j}{\arg\max}\ CS(t_i, t_j)$

We provide formal proofs for these in Appendix A. Given the result stated in Corollary 1, from the point of view of Global Pareto efficiency, the strongly efficient outcome is defined where tariffs for both firms are set at zero.

Definition 2 *Strong efficiency is defined where* $t_{ij}^* = 0$ *for all* i, j *where* $i \neq j$.

We argue that allocations achieved by a bargaining solution with some egalitarian principles married with individual interests such as the Weak Constrained Egalitarian Allocation proposed by Dutta and Ray[6]. In this paper, Dutta and Ray proposed an egalitarian solution concept of transferable utility games that coupled commitment for egalitarianism and promotion of individual interests in a consistent manner. Dutta and Ray's weak egalitarian allocation solution, one agent is at least as well off as before they bargained and at least the other is strictly better off. In order for the bargaining solution to satisfy the egalitarian constraint, the allocation must be the Lorenz-maximal element. That is, the allocation cannot be one where one agent takes all of the benefits of the bargain and leaves nothing or epsilon for the other. The Lorenz-maximal elements requires that each agent in bargaining get the highest allocation in bargaining, satisfying individual rationality.

For our purposes, the choice of this solution concept is practical, in that it narrows the solution among infinite possibilities found in the Nash Bargaining Solution. The use of the Weak Constrained Egalitarian Solution also satisfies the sense of what is fair remuneration for services a network firm provides through the use of its network. We will find that the most efficient bargaining protocol will be to allow the larger of two asymmetric firms to announce the take-it or leave-it offer to the smaller firm. Without restricting the possible allocations to those that satisfy these egalitarian constraints, we would end up consider allocations where the larger firm proposes to take all of the gains in bargaining and leave the smaller firm no better off than it was before. Because interconnection services of the type we envision in this paper have aspects of principal agent problem, it may be to the larger firms benefit to seek bargaining solutions where it demonstrates goodwill to the smaller firm so as to gain higher quality interconnection services from its smaller counterpart.

Our next task is to see how this protocol applies to the interconnections model we proposed.

Bilateral bargaining with symmetric and asymmetric firms

Consider the bargaining protocol proposed above when two identical firms negotiate. We find that the set of tariffs that are the solution to this negotiation result in, as Carter and Wright's paper suggested, strong Pareto efficient outcomes that are individually rational. In comparison to weak Pareto efficiency discussed above, strong Pareto efficiency requires that all firms' profits are higher as a result of negotiations. We present the solution to the identical firms' bilateral bargaining problem in the following lemma:

Lemma 1 *The Bargaining Problem is well defined. In particular, if $a_{ij} = a_{ji}$, there exists a tariff pair $(t'_{ij}, t'_{ji}) \in C$ such that $\pi_i(t'_{ij}, t'_{ji}) > \pi_i(t^*_{ij}, t^*_{ji})$; for all i, j, $i \neq j$.*

The proof showing that a bargaining problem with symmetric firms satisfies the strong Pareto efficiency can be found in the proof for Carter and Wright's [4] Proposition 2. In that proof, Carter and Wright show that the first derivative of a firm's profit function with respect to its own tariff is always zero, satisfying the profit maximization condition. At the same time, the first derivative of a firm's profit function with respect to the other firm's tariff is always negative.[4] Therefore, in a bargaining situation, both firms would have incentives to make offers, simultaneously asking for the elimination of all tariffs. That is, the mutual elimination of tariffs between identical monopolists increases profits for *both* firms. This leads us to the question of whether all bargaining solutions in this bargaining problem result in strong Pareto efficient outcomes. We find that efficiency of outcomes in the bargaining solutions set is determined by the relative sizes of the firms negotiating. We present the following proposition.

Proposition 1 *A Bargaining Solution can be strongly efficient for the two-stage game of tariff and price setting if and only if $\sqrt{\frac{2}{3}}a_{ij} \leq a_{ji} \leq \sqrt{\frac{3}{2}}a_{ij}$.*

Proof. Step 1: Show that a Bargaining Solution can be strongly efficient if $\sqrt{\frac{2}{3}}a_{ij} \leq a_{ji}$. In order to show the necessary condition, assume that $\sqrt{\frac{2}{3}}a_{ij} \leq a_{ji}$ and $\pi^*_i(p^*_{ij}(v_j), v_i) > \pi'_i(p'_{ij}(t'_{ji}), t'_{ij})$. Consider the Table 1-1 below where profits for Firm i, π_i, under disagreement point (v_i, v_j) is defined by $\pi_i(p^*_{ij}(\mathbf{t}^*), t^*_{ij}) = \frac{a^2_{ij}+2a^2_{ji}}{16b}$ and under bargaining, where $t'_{ij} = 0$ and $t'_{ji} = 0$, by $\pi_i(p'_{ij}(t'_{ji}), t'_{ij}) = \frac{a^2_{ij}}{4b}$. By substitution, $\frac{a^2_{ij}+2a^2_{ji}}{16b} > \frac{a^2_{ij}}{4b}$. This implies that $\sqrt{\frac{2}{3}}a_{ij} > a_{ji}$. This is a contradiction. Therefore, when $\sqrt{\frac{2}{3}}a_{ij} \leq a_{ji}$, $\pi^*_i(p^*_{ij}(\mathbf{t}^*), t^*_{ij}) < \pi'_i(p'_{ij}(t'_{ji}), t'_{ij})$.

[4]We provided a verbal explanation of Carter and Wright's proof in the section discussing why identical firms benefit from complete elimination of interconnection service tariffs.

Step 2: Show that if there is a strongly efficient Bargaining Solution, then $\sqrt{\frac{2}{3}}a_{ij} \leq a_{ji}$. A strongly efficient solution requires that $t'_{ij} = 0$ and $t'_{ji} = 0$. Consider Equations 3.1 and 3.3:

$$\pi_i = p_{ij}(a_{ij} - bp_{ij}) + t_{ij}(a_{ji} - bp_{ji}) - t_{ji}(a_{ij} - bp_{ij}) \text{ for all } i \text{ and } j,\ i \neq j;$$

and

$$p^*_{ij} = \frac{a_{ij}}{2b} + \frac{t^*_{ji}}{2} \text{ for all } i \text{ and } j, i \neq j,$$

respectively. By substituting $t'_{ij} = 0$ and $t'_{ji} = 0$, we obtain that $\pi_i = \frac{a^2_{ij}}{4b}$. Profits $\frac{a^2_{ij}}{4b} > \frac{a^2_{ij}+2a^2_{ji}}{16b}$ only when $\sqrt{\frac{2}{3}}a_{ij} \leq a_{ji}$. ■

	Large firm (i)	Small firm (j)
Profits at disagreement point: $\pi^*_i(p^*_{ij}(\mathbf{t}^*), t^*_{ij})$ for all i, j, $i \neq j$	$\frac{a^2_{ij}+2a^2_{ji}}{16b}$	$\frac{a^2_{ji}+2a^2_{ij}}{16b}$
Profits under bargaining: $\pi'_i(p'_{ij}(t'_{ji}), t'_{ij})$ for all i, j, $i \neq j$	$\frac{a^2_{ij}}{4b}$	$\frac{a^2_{ji}}{4b}$

Table 3-1: Profit Allocations at Agreement and Disagreement Points

The proof showed that in bargaining it is individually rational for Firm i to prefer $\pi'_i(p'_{ij}(t'_{ji}), t'_{ij})$ compared to $\pi^*_i(p^*_{ij}(\mathbf{t}^*), t^*_{ij})$ when $\sqrt{\frac{2}{3}}a_{ij} \leq a_{ji}$ for all i, j, $i \neq j$. Using the same techniques, it is straightforward to show that the large firm (i) will always prefer the no tariff solution. At the same time, when $a_{ij} > \sqrt{\frac{3}{2}}a_{ji}$, the small market Firm j will prefer to stick to the disagreement point where tariffs are set at the monopoly level and prices are set at the double marginalization level.

Corollary 2 *When* $a_{ij} > \sqrt{\frac{3}{2}}a_{ji}$, *there exists a solution to the Bargaining Problem where* $t'_{ij} = 0$ *for* i, *and* $t''_{ji} > 0$ *where* $t'_{ji} \in \arg\min_{t_{ji}} \max \pi_j(p'_{ji}(\hat{\mathbf{t}}), t''_{ji})$ *for* j, $i \neq j$ *and* $\hat{\mathbf{t}} = (t'_{ij}, t''_{ji})$.

The corollary defines the most efficient, individually rational set of tariffs resulting from bargaining between two, asymmetric firms. With asymmetric firms, it is individually rational for the larger Firm i to offer to eliminate tariffs altogether, and for the smaller Firm j to accept a reduction, but not eliminate, its tariffs. The intuition for this result is clear if we examine how bargaining takes place between Firm i and Firm j. We explain this result by first assuming that tariffs are set at the Subgame Perfect Nash, monopoly level. We note that tariff revenue for each firm is the tariff (t^*) multiplied by the quantity demanded for the final good in the other firm's market. Tariff revenue for the smaller Firm j is greater than tariff revenue for the larger Firm i for two reasons:

- The tariff for Firm j, t^*_{ji}, is positively related to size of Firm i's final good market (a_{ij}) (see Equation 3.2), so $t^*_{ji} > t^*_{ij}$; and
- Since the slope of demand is the same in both markets, the quantity demanded from consumers in Firm i's market is always greater than in Firm j's market.

In negotiation, firms will trade off the revenue losses from the reduction in tariffs (a first order effect) with the revenue gains that result from lower retail prices and greater quantities demanded as a result of these lower prices (a second order effect). When bargaining begins, Firm i will ask j to reduce its tariff by offering j the extra revenue generated by i's reductions in tariff. As long as Firm i can continue to match any losses in revenue Firm j experiences as a result of Firm j's reduction of tariff t_{ji}, then bargaining can continue. Bargaining will continue until Firm i has reduced its tariff to zero. Remember tariff t_{ij} started at a lower level because Firm j's demand is smaller than Firm i's demand. At this point, tariff t_{ji} will be greater than zero. The above proposition states Firm j's tariff will remain above zero if demand for final goods for Firm j is in the order of $a_{ij} > \sqrt{\frac{3}{2}}a_{ji}$ compared to Firm i.

The following corollary defines the allocation outcome from the most efficient, individually rational bargain that could be achieved between Firm i and Firm j.

Corollary 3 *When* $a_{ij} > \sqrt{\frac{3}{2}}a_{ji}$*, the allocation,* $(\pi_i, \pi_j) \in F$*, resulting from bargaining is weakly Pareto efficient.*

The corollary states that when the asymmetries between firms are large, the individually rational bargaining solution yields a (weakly) Pareto efficient outcome. In this case, the sum of the two firms' profit went up. This increase in the sum of profits results from the fact that the larger firm (i) is made better off as a result of the bargain and the smaller firm (j) is neither worse nor better off. When the smaller firm, Firm j, lowers tariffs, it reduces large double marginalization inefficiencies in Firm i's market. When the larger firm, Firm i, reduces tariffs, it eliminates the same inefficiencies of a smaller magnitude in Firm j's market. The reduction in inefficiencies in Firm j's market are directly offset by the loss of tariff revenues[5] that j no longer collects from i. Hence, the smaller firm j is neither better nor worse off, while Firm i benefits.

Given the finding that a weakly Pareto efficient, individually rational bargaining outcome will result in the elimination of interconnection tariffs by the larger firm and the reduction of tariffs by the smaller firm, we propose a bargaining protocol where the larger firm is given the option of making take-it or leave-it offers to the smaller firm. This protocol will lead to welfare improvements because it is the larger firm that has the individual incentives to reduce tariffs. The larger firm will propose to reduce tariffs, but the smaller firm is indifferent between the lower level of tariffs and the Subgame Perfect Nash monopoly tariff levels. By providing the larger firm the option of making a proposal on tariffs, and forcing the smaller firm to be allowed to accept or reject the offer, tariffs will be reduced overall. In extensive form, the proposed protocol would be implemented in a sequence of stages:

[5]Note that tariffs work as side payments from one firm to the other. Our bargaining allows for side payments through the setting of tariffs, but it does not for unlimited side payments since we restricted tariffs to be non-negative.

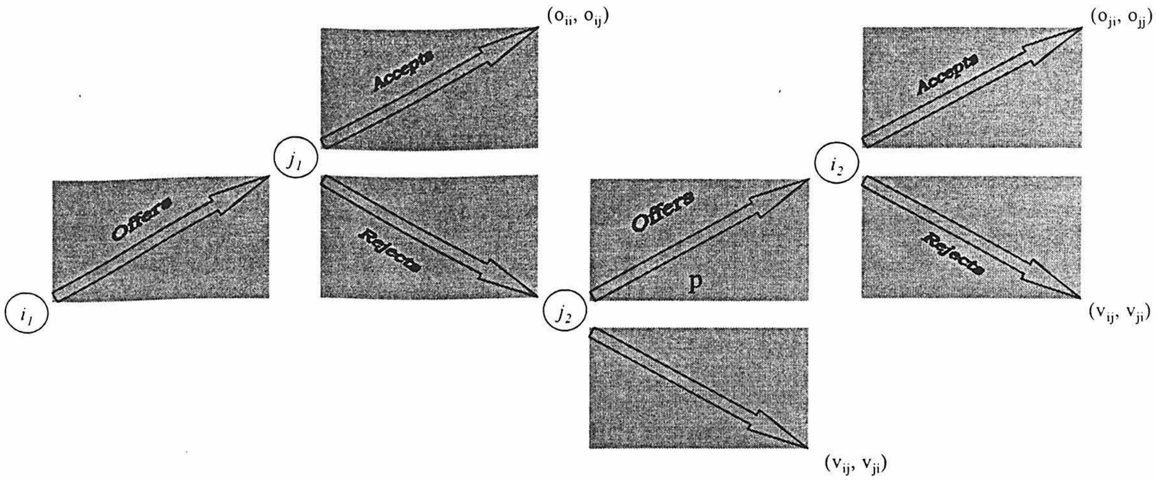


Figure 3-3: Extensive Form Game of Bargaining Protocol

1. The larger firm, Firm i, makes an offer to Firm j on the level of interconnection tariffs for both firms.

2. Firm j can accept or reject the offer. If Firm j accepts, the game is over and tariff levels are set according to Firm i's offer and payoffs are (o_i^i, o_j^i). The superscripts denote the firm that made the offer and the subscripts denote the firm that receives the payoff.

3. If Firm j rejects the offer, then it gets a chance of $(1-p)$, $0 \leq p \leq 1$, to make a counter offer or a p chance to have tariff levels set at the Subgame Perfect Nash monopoly levels and payoffs are (v_i, v_j). The probability, p, is set by the social planner whose aim is to maximize welfare by having all firms set the lowest tariffs possible.

In the state where Firm j is allowed to make an offer, Firm j presents a final take-it or leave-it offer to Firm i. Firm i will accept this offer if it is greater than the profits it would receive had tariffs been set at the disagreement, Subgame Perfect Nash monopoly levels.

A social planner decides on the value of p. In order to achieve the highest level of efficiency, the objective is to make the smaller firm, j, indifferent between accepting the offer made by Firm i in the first round, foregoing the opportunity to make a counter offer that is more favorable to Firm j but less efficient as well. The appropriate p will be set at a level that satisfies the following inequality:

$$(1-p)v_j > p(o_j^i)$$

Solving for the probability, p, the probability satisfies the following inequality:

$$p < \frac{v_j}{v_j + o_j^i}$$

If the social planner set the p according to this inequality, Firm j's expected allocation is higher in accepting the offer than taking its chances on rejecting and receiving the disagreement point payoffs. Firm j will therefore accept Firm i's offer in the first round.

3.3.2 Empirical evidence

In this section, we empirically test one hypothesis and examine one aspect of negotiations over interconnection tariffs between monopolists selling two-way interconnection services to each other. First, we begin by testing the hypothesis that there exists a negative relationship between the ratio of market sizes and the interconnection tariff agreed upon through negotiations. Second, we determine whether the data of international telecommunications shows that firms reached efficient bargains in setting interconnection tariffs. We begin by testing our hypothesis.

Corollary 1 states that the larger the asymmetries between firms, the higher the interconnection tariff the smaller firm will charge the larger firm. Another way to state this relationship is by thinking in terms of a ratio; the ratio where the smaller firm's (j) demand a_{ji} is divided by the larger firm's (i) demand a_{ij}. Our model predicts a negative relationship between this ratio and the small firm's negotiated tariff (t_{ji}). We predict that as the ratio of market sizes gets closer to one, the negotiated tariffs get asymptotically closer to zero.

We use data from the international telecommunications industry to test our hypothesis. In this data, we find outcomes of bilateral bargains over interconnection tariffs between the US (the large market) and foreign telecommunications carriers (the smaller markets). In the telecommunications industry, interconnection tariffs are known as accounting rates. We have accounting rate data for 1997 for the US and 173 countries[6] with which US telecommunications carriers negotiated interconnection contracts. We also have market size data in the form of minutes of telephone interconnection between US telecommunications carriers and foreign carriers. The US is the largest market for international telephone message market.

The rules of the bargain that determined these accounting rates are slightly different from those we used to model our bargaining game. However, none of the actual rules are sufficiently different from our model to cause us to change our model or its predictions. First, the data contains accounting rates resulting from bilateral negotiations between monopolists over two-way interconnection services. In our data, bilateral bargaining took place between a cartelized group of US international telecommunications carriers and, most often, a single foreign monopolist. In theory when firms form a cartel, they act as a single monopolist would so this allows us to consider the four US international carriers as one monopolist. On the other side of the negotiations table, we find that nearly all foreign international telecommunications carriers are monopolists or are also cartelized when there is more than one international carrier in the country. Furthermore, international telecommunications require two-way interconnection. Given these features of how accounting rates are settled between international carriers, the data describe the same problem we proposed to examine in our model—bargains between

[6]We deleted 30 outliers. These outliers included: Andorra, Antartica, Bhutan, Comoros, Congo, Cuba, Equatorial Guinea, French Guiana, Guinea-Bissau, Kiribati, Kyrgyzstan, Laos, Madagascar, Maldives, Mali, Midaway Atoll, Montserrat, Myanmar, Norfolk Island, North Korea, Niue, Palau, Reunion Islands, Sao Tome and Principe, Seychelles, St. Helena, St. Pierre, Vanuatu, Vietnam, and Yemen. They were determined to be outliers because of their large incoming to outgoing minutes—a ratio of over 50:1—with the US market.

two monopolists negotiating over two-way interconnection services.

The bargaining mechanism used to determine outcomes allowed for the splitting of differences from each firm's ideal point.[7] Moreover, the regime governing the accounting rate system requires that interconnection charges be uniform on both sides of the link. The rules for the determination of accounting rates that compel firms to apply the same tariff to each other's telephone message traffic is unlike the model we introduced. In our model, interconnected firms apply different tariffs if they are asymmetric. However, as we will see below, this does not present a problem for our model.

We begin by testing the hypothesis that the US, being the larger market, will attempt to push accounting rates to zero. However, according to our model US carriers will be less successful when bargaining against monopolists servicing smaller markets. The US firm can achieve its best outcome against the smaller firm by making the smaller firm indifferent, allowing tariffs to be set at a level where profits are at least as high as when the smaller firm set them at the monopoly level. In other words, we assume that the larger firm (i) will be able to negotiate the smaller firm (j) tariffs down to $t'_{ji} = t^*_{ji} - t^*_{ij}$. Furthermore, since the same tariff is applied to telephone traffic going both directions, the splitting of the difference between each firm's ideal point is a reasonable way to calculate the accounting rate upon which firms can agree. According to the rules of negotiation that determined accounting rates, we split the difference between the ideal point for the larger firm and the point the US firm was able to convince the foreign firm to take. By doing this, we get the equation that determines interconnection tariffs that must be applied to both markets:

$$\mathbf{t} = \frac{t^*_{ji} - t^*_{ij}}{2}$$

Substituting for t^*_{ij} and t^*_{ji} using the following equalities, $t^*_{ij} = \frac{a_{ji}}{2b}$ and $t^*_{ji} = \frac{a_{ij}}{2b}$, we get the

[7] This way of determining outcomes is consistent with the Nash Bargaining Solution where agents split the surplus from bargaining.

following relationship between accounting rates and market sizes:

$$rates = \frac{a_{ij} - a_{ji}}{4b}$$

From the above equality, we expect an asymptotic relationship between the accounting rates and the ratio of market sizes. For example, as the limit of the ratio goes to zero, that is when the US market is comparatively larger in size with respect to the foreign market, accounting rates should go to infinity. In comparison, when the foreign market is more similar in size to the US market, and the ratio goes to 1, the accounting rate should converge to zero. Based on these relationships, we test the following econometric model:

$$rates = \alpha_0 + \beta_1 \left(\frac{a_{ij} - a_{ji}}{4b} \right)$$

The dependent side of the equation is the accounting rates—the levels of which are determined through bargaining. The independent side of the equation contains an intercept term and the ratio term. Because the ratio is calculated by dividing the minutes that foreign countries called the US over the minutes the US called that foreign destination, we would expect the ratio to be bounded by one at the lower end. If the predictions of our model are correct, we would expect that the intercept term, α, would be zero and that the coefficient to the ratio variable would be negative.

We propose to use the square root of the ratio in our model. There are two arguments for using the square root transformations of the ratio term. First, in looking at Figure 3-4, a graph of accounting rates against market size ratios, it is clear that there is a curve, instead of a negatively sloped line, which fits best the relationship between accounting rates and ratios. A curve, that is asymptotic to both the y and x axes as accounting rates and ratios increase, respectively, is what we expect from our model as well. Second, in attempting to fit the model above, our plot of predict values against the residuals of the model produced a funnel shaped

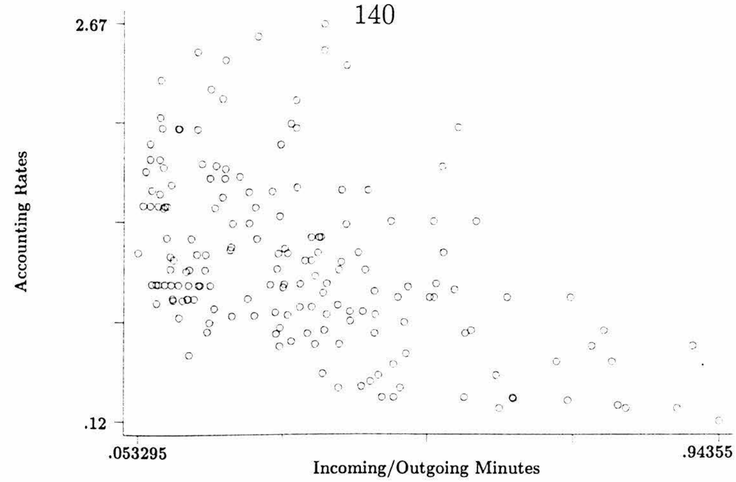


Figure 3-4: Accounting Rates Against Market Size Ratios

cloud of points where the funnel expanded as the ratio of market size became closer to one (see Appendix 3.A). The funnel shape cloud means that the assumption of homoscedasticity that ordinary least squares imposes, where there should exist constant error variance, is violated and that our estimates will be inefficient in that the variances of the estimated parameter is not the minimum variance. One solution to this problem is to use the square root of $\frac{a_{ij}-a_{ji}}{4b}$ instead. Therefore, we propose the following econometric model:

$$rates = \alpha_0 + \beta_1 \left(\sqrt{\frac{a_{ij} - a_{ji}}{4b}} \right)$$

We argue for one more refinement to our model. A dummy term should be added to the model to reflect the differences between bargaining that takes place between the US and a developing country and bargaining between the US and a developed country. By using the dummy term, we compensate for two effects that make developing and developed countries different. The effects are:

- Developing country telecommunications monopolies, since they are mostly state-owned, may not be maximizing profits as we stipulate in our model, but rather they are maximizing foreign exchange revenues. Developing countries have a high shadow price for foreign exchange [5]. A high shadow price means that increases in foreign exchange lead to greater increases in overall welfare in developing countries than in developed nations. For our purposes, this translates to the fact that developing countries are more likely to maximize the amount of foreign exchange they receive from US carriers by charging high interconnections tariffs and sacrificing consumer welfare in their international telecommunications market. Our model does not incorporate foreign exchange maximization arguments based on high shadow prices for foreign exchange in developing countries; and

- Telecommunications carriers in developing countries may experience higher costs in termination and originating international telephone calls. The model we present here does not include cost variables, and our estimation may be biased by the exclusion of these variables for developing countries. By excluding developing countries from our sample, we avoid the problems of biased estimation. Implicitly, we assume that telecommunications firms in developed countries use technology with similar costs than that used by US carriers.

Given our arguments for including the dummy in our specification, we test our model for the following functional form:

$$rates = \alpha_0 + \beta_0 \begin{cases} 0 \text{ if undeveloped} \\ 1 \text{ if developed} \end{cases} + \beta_1 \left(\sqrt{\frac{a_{ij} - a_{ji}}{4b}} \right)$$

Our hypothesis predicts that the coefficient of the dummy term, β_0, would be significant if there are differences between developed and developing countries. Furthermore, our model predicts that the β_0 would be negative since accounting rates are generally lower for developed countries than for developing countries. The results of the regression do not contradict our model's predictions. The sign for the coefficients are what we expected and the estimated coefficients are significant to at least the 95 percent confidence level (we provide the t-statistics below the estimates):

$$rates = \underset{(9.612)}{2.163} - \underset{(-6.123)}{0.705} \begin{cases} 0 \text{ if undeveloped} \\ 1 \text{ if developed} \end{cases} - \underset{(-4.030)}{1.295} \sqrt{\frac{a_{ij} - a_{ji}}{4b}}; \quad R^2 = 0.365$$

Given the empirical verification that there are differences between developed and developing countries, we ran regressions for each group separately. We obtained the following results:

Regression for 22 developed countries

$$rates = \underset{(4.101)}{1.458} - \underset{(-2.310)}{1.402} \sqrt{\frac{a_{ij} - a_{ji}}{4b}} \qquad R^2 = 0.208 \tag{3.4}$$

Regression for 151 developing countries

$$rates = \underset{(9.325)}{2.182} - \underset{(-3.623)}{1.291} \sqrt{\frac{a_{ij} - a_{ji}}{4b}} \qquad R^2 = 0.080 \tag{3.5}$$

While the prediction that the intercept terms should be zero is not verified in our models, the intercept term is significantly smaller for estimates of accounting rates for developed countries in comparison to developed ones. Not surprisingly, given our model using the dummy variable, the sign of the coefficients are what we expected.

Examine Figure 3-5 where we plot the predicted values for the models. In this figure, the curve that appears above traces predicted values for accounting rates for developing countries. The curve below traces predicted values for developing countries using the coefficients for the model in Equation 3.5. Note the asymptotic relationship for both axes that we discussed above. The tendency toward infinity when the market ratios become smaller is pronounced. However, the curve tracing out predicted values for developing countries is nearly parallel with the x-axis as the ratios get larger.

We extend our empirical analysis by examining whether the accounting rate data shows that

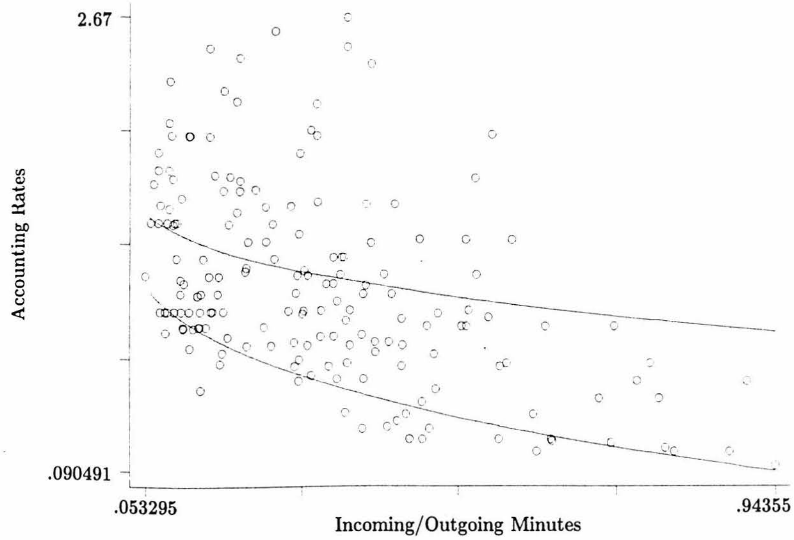


Figure 3-5: Predicted Accounting Rates Based on Market Size Ratios

agreed upon rates are near efficiency. In other words, given the relationship between market sizes and accounting rate levels we established above, are accounting rates as low as in the most efficient bargains struck between the US and other markets? One way to answer these questions is to find the lower bound between accounting rates and relative market sizes in our data. This bound estimates the lowest accounting rates possible given differences in market sizes. Once we estimated this lower bound, we can determine whether there is a systematic deviation away from the estimated lower bound in the accounting rate data. If we find that many of the accounting rate bargains struck are close enough to the lower bound, a measure that will be made specific below, then we can conclude that the US and its foreign counterpart were able to achieve comparatively efficient bargains. Given the FCC's dissatisfaction with the present level of accounting rates, we may expect to find that accounting rate deals do not conform to the efficiency standard.[8]

[8]On January 31, 1996, the FCC announced its new approach to the settlement of accounts between US and

In our analysis of estimating lower bounds of the relationship between market ratios and accounting rates, we follow Sutton's[9] [15] analysis in which he estimated a lower bound for the relationship between advertising and market size data. In Sutton's analysis he used the Weibull distribution corresponding to the errors in the lower bound model. Weibull, or other one-sided distributions, is an appropriate assumption in models of this kind[10] because all error terms will be positive—above the lower bound. These methods involve modelling the distribution of residuals between the observed values $y_k(z_k)$ and the bound $b(z_k)$, as a Weibull distribution. Formally, we write that $x_k = y_k - b$ should be distributed according to:

$$F(x) = 1 - \exp\left[-\left(\frac{x-\mu}{s}\right)^{\omega}\right] \qquad \omega > 0,\ s > 0$$

on the domain of $x \geq \mu$.

We use maximum likelihood methods to fit the lower bound of our accounting rate data. We begin by first parameterizing the lower bound $b(z_k)$. The relationship for a lower bound should minimize the sum of the residuals, $y_k - b(z_k)$, subject to the constraint that all residuals are non-negative. The second step in our analysis is to determine whether the residuals demonstrate a pattern of distribution consistent with the Weibull distribution. The k observations where the residuals are zero are deleted and the rest are used to see if they fit the Weibull distribution.

The decision on the parameterization of the lower bound is based on examination of the relationship between market size ratios and accounting rate levels (see Figure 3-1). In examining this relationship, we determine that a reasonable family of candidate schedules is:

foreign carriers for the provision of international service. In effect, the FCC set benchmarks for accounting rates between the US and its foreign counterparts. If a foreign telecommunications carrier does not negotiate an accounting rate below that benchmark, the FCC will punish that foreign carrier by holding back US settlement payments to compensate for interconnections services provided overseas. See IB Docket No. 96-261, FCC Order No. 97-280.

[9] See Sutton (1994), Chapter 5.

[10] See *Journal of Econometrics* **46**, Vol. 1/2, 1990. This issue is dedicated to Frontier Analysis in production and utility functions. These type of analyses also use one-sided error distributions.

$$\sqrt{rates} = \alpha_1 + \frac{b}{\sqrt{ratio}}$$

We use the square root transformation of accounting rates for two reasons:

The procedure used to estimate lower bounds rests on the assumption that the distribution of residuals is identical at all values of the independent variable (in this case, $\sqrt{ratio}$). This is clearly unreasonable since accounting rates have a lower bound of zero. So we take a square root transformation and use this transformation in our estimates of the lower bound.

The fitted values for this equation are shown in Figure 3-6. The associated set of residuals fits the Weibull distribution well for developed countries, but not for developing ones. The hypothesis that $\mu = 0$ cannot be rejected, therefore, for the developed countries, but it can be so for developing countries. Finally, none of the estimates values of ω are below 2, so maximum likelihood is the correct approach. Smith [14] showed that for $\omega > 2$, the usual asymptotic results of maximum likelihood hold; namely, consistency, asymptotic efficiency, and an asymptotically normal distribution.

The table below contains the actual estimates of the lower bounds (standard errors in parentheses). The values of μ corresponds to the mean of the three-parameter Weibull distribution. We test whether the mean of the residuals is significantly close to zero by comparing the log likelihood values for a restricted model, where we set $\mu = 0$, with likelihood value of the unrestricted model. If the negative log likelihood value for the model where we set $\mu = 0$ is significantly different from the negative log likelihood value without this restriction, then we can conclude that our error terms deviate significantly from the estimated lower bound. In our data, we find that the change in the negative log likelihood value ($\Delta NLLV$) is not significant for the models with and without this restriction for developed countries. However, the restriction on the mean yields a significantly different negative log likelihood value for the developing country model without the restriction. We conclude that our results do not contra-

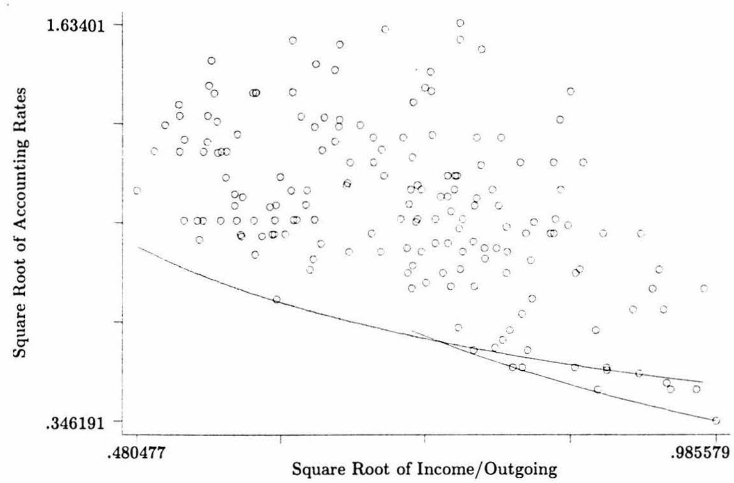


Figure 3-6: Scatter diagram of the square root transformation of *rates* against $\frac{1}{\sqrt{ratio}}$. The fitted bound for developed countries is -0.453+0.572$\left(\frac{1}{\sqrt{ratio}}\right)$. The fitted bound for developing countries is 0.040+0.420$\left(\frac{1}{\sqrt{ratio}}\right)$. The lower bound is that for the 22 developed countries.

dict the hypothesis that developed countries and the US bargain close to the estimated lower bound, but we can reject the same hypothesis when examining bargaining between the US and developing countries. From Figure 3-6 above, the shorter, lower line is the boundary estimate for the developed country set. From this figure, we see that developed countries have markets closer to the size of that in the US; i.e., the market size ratios are close to one. Given the near parity in market sizes, our model predicts that efficiency should be relatively easier to achieve since it is individually incentive compatible. The estimates confirm what we proposed in our model; incentives for symmetric firms involved in bargaining over interconnection tariffs are compatible with achieving the strong Pareto efficient outcome where all tariffs are set at zero.

	α_1	μ	ω	$\Delta NLLV$
Developed Countries	-0.453	-0.16 (0.09)	2.09	0.47
Developing Countries	0.040	-1.38 (2.37)	2.01	1.77

Table 3-2:Weibull Maximum Likelihood Results

In practical terms, the result of the estimation of lower bounds for developing and developed countries means that telecommunications carries in developing countries and the US international telecommunications firms are not reaching efficient bargains. In comparison, when US international carriers negotiate accounting rates with firms from developed countries, it is likely that a comparatively more efficient rate is achieved. These results are consistent with the observation that the Federal Communications Commission has come down hard on developing countries to reduce their accounting rates, but at the same time the Commission hardly mentions developed countries in their efforts to reduce the cost of calling internationally.

Implications

The empirical tests are consistent with the hypothesis that the relative sizes of firms in a bilateral bargain over interconnection tariffs explain some of the variation in the tariffs we see in one network industry. We developed theoretically, and found empirical support for the hypothesis that when asymmetric firms bargain, the larger firm has the incentive to negotiate reductions or elimination of interconnection tariffs. Welfare will be maximized when the larger firm is given greater negotiating power over the smaller. One way to distribute negotiating power in favor of the larger firm is to allow it to a make take-it or leave-it offers to smaller firms that set tariffs at the larger firm's ideal point. However, as we have shown before, setting interconnection tariffs to meet the larger firm's ideal tariff levels is not incentive compatible for the smaller firm. The smaller firm is unlikely to agree with the offer. Knowing that the smaller firm will not accept the larger firm's offer, we can assume that the larger firm will announce tariff levels that are the lowest possible while making the smaller firm indifferent between these and those tariffs set at the monopoly level. From now on, we assume a bargaining protocol where the larger firm makes all take-it or leave it offers, leaving the smaller firm at least as well off if it accepts the larger firm's offer.

3.4 Conclusions

In this paper we showed that bilateral bargaining between symmetric firms will lead to strong Pareto efficient outcomes that are individually rational. This outcome is defined by tariffs set at zero. Because the elimination of tariffs benefit both firms equally, there was little need to induce firms to negotiate. Moreover, because the most efficient outcome coincided with the individually rational outcome, there was no need for rules to guide these firms toward the most efficient outcome. In bargaining between asymmetric firms, our results were different. Because positive tariffs present a greater inefficiency to the firm servicing the larger retail market, the larger firm will have greater incentives to reach a settlement with the smaller firm. Furthermore, because the most efficient, individually rational point in a bilateral bargain between asymmetric firms made the bigger firm better off and left the smaller one indifferent, we proposed a bargaining protocol that induced the smaller firm to agree to this point for the sake of greater efficiency.

For policy makers on regulation of network industries, this paper can be seen as what would take place if full liberalization methods, such as open skies for the airline industry and ownership of telecommunication facilities in foreign countries, are not implemented. Competition for termination will achieve the elimination of monopoly power. However, caution is warranted since the pricing rule used to determine charges for terminating indirect traffic will affect the efficiency achieved throughout the network.

Future research in this are should aim to determine whether the bargaining protocol proposed here would achieve efficient networks when bargaining between more than two monopolies of extreme size differences. We believe there is reason to be optimistic and that there exists an avenue for further research in this finding. Our optimism is based on the observation that the inequalities that sustain the no efficiency result are less likely to be satisfied if we add more firms to interconnected networks. Larger firms, frustrated by bargaining with a smaller firms that will not reduce tariffs, may do better by adding a medium-sized firm to the bargaining table. Implicitly, when we add more firms to bargaining, we formalize what may actually take place in multilateral bargaining. The medium sized firm may be able to get a better deal from the smaller firm. With indirect interconnection, larger firms may be able to interconnect though the medium-sized firm and avoid the high tariffs of the smaller firm altogether. An alliance between the medium-sized firm and the larger one may be beneficial to both.

Given the difficulty network firms may experience in bringing down interconnection tariffs, alliances may be an efficient and appropriate coordinating mechanism. Whether good or bad, alliances are an increasing feature of the landscape in network industries. In the international telecommunications industry, for example, alliances include ten of the twenty largest international telecommunications operators, accounting for approximately 55 percent of all international telephone traffic. In the airline industry, alliance formation has also taken off. In 1998, there were four international alliances forming that covered five continents. In total, there are over 500 alliances in the airline industry.[11] This large number of alliances is dramatic when one realizes that at the start of the decade there were nearly no alliances in the airline industry.[12] More work should be done to explain if alliances are used as a way to coordinate interconnection tariffs and mitigate the costs of having to negotiate with several network firms in piecemeal fashion. In conclusion, we believe that it is worth investigating whether firms benefit when dealing with interconnection multilaterally and with firms of different sizes in coalitional structures like alliances. It may be that the most interesting, policy relevant avenues in which to extend research are in investigating how alliances take shape by using the bargaining in networks approach. The bargaining in networks approach can be an effective way to undertake research on alliance formation. There are several empirically testable hypotheses proposed here. We provided one test for the hypotheses that there exists a negative relationship between the level of the interconnection tariff negotiated between two firms and the relative disparity in the size of the firms involved in bargaining.

[11] The Economist; September 28, 1998. pp. 68.

[12] *Business Week*, "Taking a Whack at Airline Alliances," March 16, 1998, p. 106D-106G.

Bibliography

[1] M. Armstrong, *Network interconnection in telecommunications*, The Economic Journal **108** (1998), pp. 545-564.

[2] M. Armstrong, C. Doyle, and J. Vickers, *The access pricing problem: a synthesis*, Journal of Industrial Economics **44** (1996), pp. 131-150.

[3] M. Carter and J. Wright, *Interconnection in network industries*, Review of Industrial Organization, forthcoming.

[4] M. Carter and J. Wright, *Symbiotic production: the case of telecommunications pricing*, Review of Industrial Organization **9** (1994), pp. 365-378.

[5] J. Drèze and N. Stern, The theory of cost-benefit analysis, in A.J. Auerbach and M.S. Feldstein (eds.), *Handbook of Public Economics, Vol. 2, Chapter 14*, Elsevier Science Publishers: Amsterdam, Netherlands (1991), pp. 909-985.

[6] B. Dutta and D. Ray, *A concept of egalitarianism under participation constraints*, Econometrica **59** (1989), pp. 615-636.

[7] B. Dutta and D. Ray, *Constrained egalitarian allocations*, Games and Economic Behavior **3** (1991), pp. 403-422.

[8] Federal Communications Commission, *Trends in the US International Telecommunications Industry*, **June** (1998).

[9] D. Galbi, The implications of by-pass for traditional international telephony, in J.K. MacKie-Mason and D. Waterman (eds.), *Telephony, The Internet, and the Media: Selected Papers from the 1997 Telecommunications Policy Research Conference,* Lawrence Erlbaum Associates, Mahwah, NJ (1998), pp. 53-68.

[10] D. Galbi, *Cross-border rent shifting in international telecommunications*, Information Economics and Policy, forthcoming.

[11] M. Jackson and A. Wolinsky, *A strategic model of economic and social networks*, Journal of Economic Theory **71** (1996), pp. 44-74.

[12] J.J. Laffont, P. Rey, and J. Tirole, *Network competition: overview and nondiscriminatory pricing*, Rand Journal of Economics **29** (1998), pp. 1-37.

[13] R. Myerson, *Game theory: analysis of conflict,* Cambridge: Harvard University Press (1991).

[14] R. L. Smith, *Maximum likelihood estimation in a class of non-regular cases,* Biometrika **72** (1985), pp. 67-92.

[15] J. Sutton, *Sunk costs and market structure: Price competition, advertising, and the evolution of concentration,* Cambridge: MIT Press (1992).

3.A Appendix

Total welfare is formally defined below:

$$\mathcal{W}(t_i, t_j) = \pi_i(t_i, t_j) + \pi_j(t_i, t_j) + cs_i(t_i, t_j) + cs_j(t_i, t_j);\ i, j = 1, 2,\ i \neq j \tag{1}$$

where π_i, π_j represent profits for firm i and j, and cs_i, cs_j represent consumer surplus for consumers in markets i and j. To simplify notation, we define total consumer surplus:

$$CS(t_i, t_j) = cs_i(t_i, t_j) + cs_j(t_i, t_j);\ i, j = 1, 2,\ i \neq j \tag{2}$$

where $cs_i(t_i, t_j) = \int_{p_i^*}^{a_i} q(p_i)dp_i$. Conversely, because of linear demands we can write:

$$cs_i(t_i, t_j) = \frac{1}{2}q_i(p_i)(a_i - p_i);\ i, j = 1, 2,\ i \neq j \tag{3}$$

Substituting into Equation 2 from Equation 3, we get:

$$CS(t_i, t_j) = \frac{1}{2}q_i(p_i)(a_i - p_i) + \frac{1}{2}q_j(p_j)(a_j - p_j). \tag{4}$$

We determine that the sign of the first order condition with respect to the arguments in Equation 2 is negative. We show that here:

$$\frac{dCS(t_i, t_j)}{dt_i} = \frac{dcs_i(t_i, t_j)}{dt_i} + \frac{dcs_j(t_i, t_j)}{dt_i} \tag{5}$$

Note that $\frac{dcs_i(t_i, t_j)}{dt_i} = 0$. Taking derivatives and rearranging Equation 5, we get:

$$\begin{aligned}\frac{dCS(t_i,t_j)}{dt_i} &= \frac{1}{2}\frac{\partial q_j}{\partial p_j}\frac{\partial p_j}{\partial t_i}(a_i - p_i) - \frac{1}{2}q_j\frac{\partial p_j}{\partial t_i} \\ &= -\frac{b}{2}\frac{\partial p_j}{\partial t_i}(a_i - p_i) - \frac{1}{2}q_j\frac{\partial p_j}{\partial t_i}\end{aligned}$$

By assumption $a_i > p_i$, $q_j > 0$, and $b > 0$ and from our model know that $\frac{\partial p_j}{\partial t_i} \geq 0$, so that

$$\frac{dCS(t_i,t_j)}{dt_i} < 0;\ for\ i,j = 1,2,\ i \neq j \tag{6}$$

We define the portion of total surplus that is from firm profits as:

$$\Pi(t_i,t_j) = \pi_i(t_i,t_j) + \pi_j(t_i,t_j);\ i,j = 1,2,\ i \neq j \tag{7}$$

Claim 2 $\Pi(\hat{t}) \geq \Pi(t')$ *if and only if* $CS(\hat{t}) \geq CS(t')$.

Proof. i.) We prove that if $\Pi(\hat{t}_i,\hat{t}_j) \geq \Pi(t'_i,\hat{t}_j)$, then $CS(\hat{t}_i,\hat{t}_j) \geq CS(t'_i,\hat{t}_j)$.

From the Proof of Proposition 2 in Carter and Wright [4] (Appendix, pp. 376), we know that $\frac{d\pi_i^*}{dt_j} < 0$ for $i,j = 1,2,\ i \neq j$. From the definition of Nash Equilibrium, we also know that $\frac{d\pi_j^*}{dt_j} = 0$. At Nash equilibrium, $\frac{d\Pi(t_i,t_j)}{dt_i} \equiv \frac{d\pi_i^*}{dt_j} + \frac{d\pi_j^*}{dt_j} < 0$. Note that this implies that $\hat{t}_i \leq t'_i$. From 6, we know that $\hat{t}_i \leq t'_i$ must be true in order for $CS(\hat{t}_i,\hat{t}_j) \geq CS(t'_i,\hat{t}_j)$.

ii.) We prove that if $CS(\hat{t}_i,\hat{t}_j) \geq CS(t'_i,\hat{t}_j)$, then $\Pi(\hat{t}_i,\hat{t}_j) \geq \Pi(t'_i,\hat{t}_j)$. From Equation 6, we know that $\hat{t}_i \leq t'_i$. If $\hat{t}_i \leq t'_i$, and because $\frac{d\Pi(t_i,t_j)}{dt_i} \equiv \frac{d\pi_i^*}{dt_j} + \frac{d\pi_j^*}{dt_j} < 0$, then $\Pi(\hat{t}_i,\hat{t}_j) \geq \Pi(t'_i,\hat{t}_j)$. ■

Corollary 4 $(t_i^*,t_j^*) \in \arg\max_{t_i,\,t_j} \Pi\ (t_i,t_j)$ *if and only if* $(t_i^*,t_j^*) \in \arg\max_{t_i,\,t_j} CS(t_i,t_j)$

Proof. i.) We show that $(t_i^*,t_j^*) = (0,0)\ \in \arg\max_{t_i,\,t_j} \Pi\ (t_i,t_j)$.

We know that

$$\frac{d\Pi(t_i, t_j)}{dt_i} = \frac{d\pi_j^*}{dt_i} < 0$$
$$\frac{d\Pi(t_i, t_j)}{dt_j} = \frac{d\pi_i^*}{dt_j} < 0$$

by Proof of Claim 1. This implies that minimum values of t_i, t_j are the solutions to the maximization problem, where $t_i \geq 0$ always. From Equation 6, it follows that $(t_i^*, t_j^*) = (0, 0)$ solves the consumer surplus maximization problem as well.

ii.) We show that $(0,0) \in \arg\max_{t_i,\, t_j} CS(t_i, t_j)$.

By Equation 6, we know that $\frac{dCS(t_i,t_j)}{dt_i} < 0$; $for\ i, j = 1, 2,\ i \neq j$. This implies that minimum values of t_i, t_j are the solutions to the consumer surplus maximization problem where $t_i \geq 0$ always. It follows from Proof of Claim 1 that $(t_i^*, t_j^*) = (0, 0)$ solves the overall profit maximization problem as well. ■

3.B Appendix

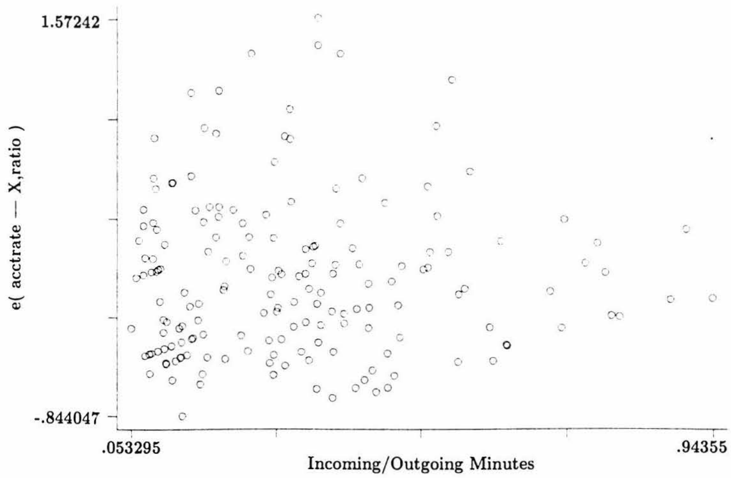


Figure 3.B-1: Residual versus predicted plot using residuals from the following model:

$$rates = \alpha_0 + \beta_0 \; ratio.$$

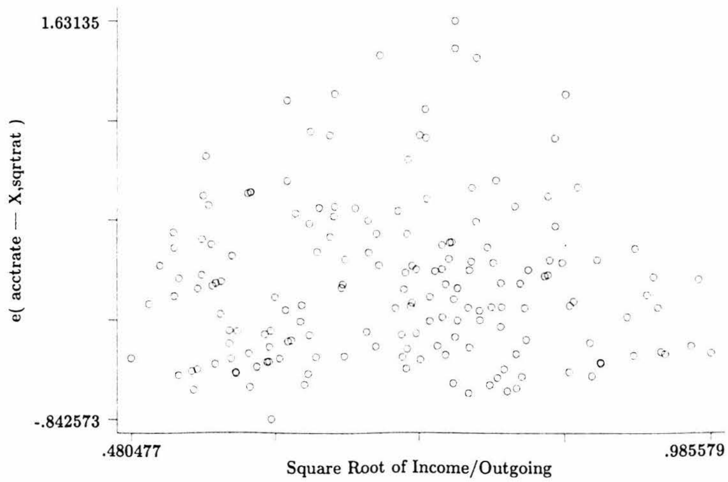


Figure 3.B-2: Residuals versus predicted values plot using residuals from the following model:

$$rates = \alpha_0 + \beta_0 \; \sqrt{ratio}$$