

Internet pricing and the history of communications

Andrew Odlyzko

AT&T Labs - Research

amo@research.att.com

<http://www.research.att.com/~amo>

Revised version, February 8, 2001.

Abstract

There are repeating patterns in the histories of communication technologies, including ordinary mail, the telegraph, the telephone, and the Internet. In particular, the typical story for each service is that quality rises, prices decrease, and usage increases to produce increased total revenues. At the same time, prices become simpler.

The historical analogies of this paper suggest that the Internet will evolve in a similar way, towards simplicity. The schemes that aim to provide differentiated service levels and sophisticated pricing schemes are unlikely to be widely adopted.

Price and quality differentiation are valuable tools that can provide higher revenues and increase utilization efficiency of a network, and thus in general increase social welfare. Such measures, most noticeable in airline pricing, are spreading to many services and products, especially high-tech ones. However, it appears that as communication services become less expensive and are used more frequently, those arguments lose out to customers' desire for simplicity.

In practice, user preferences express themselves through willingness to pay more for simple pricing plans. In addition, there is a strong "threshold" effect to usage-sensitive billing. Even tiny charges based on utilization decrease usage substantially. In a rapidly growing market, it is in the service providers' interest to encourage usage, and that argues for simple, preferably flat rate, pricing. Historical evidence suggests that when service costs decrease, such arguments prevail over the need to operate a network at high utilization levels and to extract the highest possible revenues.

Contents

1 Introduction	3
2 Growth in communications: quantity and quality	9
3 The effectiveness and utility of pricing	12
4 The role of distance dependence in pricing	16
5 The conventional economic argument for flat rate pricing	18
6 The strong public preference for flat rate pricing	20
7 The dramatic effect of usage sensitive prices	21
8 Dynamic effects of flat rates	22
9 Mail	24
10 Telegraph	25
11 Wired voice phone	26
12 Cell phones	28
13 Residential access to the Internet	30
14 Non-Internet data networks	31
15 Dedicated business connections to the Internet	32
16 Software agents and other technological palliatives	33
17 Conclusions	35

Internet pricing and the history of communications

Andrew Odlyzko
AT&T Labs - Research
amo@research.att.com
<http://www.research.att.com/~amo>

1. Introduction

The history of communication technologies, including ordinary mail, the telegraph, the telephone, and the Internet, shows a consistent pattern. Quality rises, prices decrease, and usage increases to produce increased total revenues. At the same time, prices tend to become simpler. Will the Internet follow the same trend?

The Internet has historically treated all packets equally, and pricing has been predominantly through flat monthly rates depending only on the size of access links, not on usage. However, there is a strong momentum towards changing both of these principles. This would go against the historical trend of other communication services. The basic reasoning behind this move was articulated by Pravin Varaiya in the INFOCOM'99 keynote lecture:

Although flat-rate continues to be the predominant form in which Internet access is sold, that form of pricing is unviable. Flat-rate pricing encourages waste and requires 20 percent of users who account for 80 percent of the traffic to be subsidized by other users and other forms of revenue. Furthermore, flat-rate pricing is incompatible with quality-differentiated services.

To properly evaluate Varaiya's claims, it helps to consider historical precedents. For example, in the early days of telephony, local calling around the world was typically covered by a fixed monthly fee. This practice was frequently questioned. An investigation of phone service in New York City in 1905 concluded, in words strikingly similar to those of Varaiya,

that, so far as large cities are concerned, unlimited service is unjust to small users, favors large users unduly, impedes expansion of the telephone business, tends to inefficient service, and that, as a financial proposition, is unsound.

p. 246 of [Stehman]

The technology and economics of early telephony (in particular, the diseconomies of scale caused largely by the need for human operators to set up each call) made the reasoning behind that 1905

conclusion even more compelling than the arguments supporting Varaiya’s call for abandoning flat rate for Internet access. This led most of the world towards metered local phone rates.

In contrast to other countries, unlimited local calling for a flat monthly fee for residential users has persisted in most of the United States throughout the 20th century. It may have seemed unsound in 1905, and most experts still feel it is unsound. Yet if we compare the telecommunications industries in different countries, we find few signs of harm from this “unsound” practice. Table 1.1 shows that U.S. citizens use their phones considerably more than inhabitants of other rich industrialized countries at a cost that is only slightly higher. Thus at least from this superficial view, it appears that both consumers and service providers benefit.

Table 1.1. International comparison of telephone industry revenues and usage in 1997.

country	revenues as fraction of GDP	minutes of phone calls per person per day
Finland	2.52%	16.6
France	1.93	10.6
Japan	2.06	10.6
Sweden	2.05	20.7
Switzerland	2.66	13.0
U.K.	2.51	12.7
U.S.	2.86	36.9

Not only has the U.S. phone industry managed to thrive in spite of its supposedly unsound practice of unlimited local calling, but Germany, Japan, and the U.K. are re-introducing limited forms of flat rate pricing. The pressure for such unmetered plans in other countries is also growing.

Usage-sensitive pricing is effective. The problem is that many of its effects are undesirable. In particular, such pricing lowers demand, often by substantial factors. Fig. 1.1 shows what happened when AOL switched to flat rate pricing in October 1996. Over the next year usage per person tripled. (It took that long only because AOL could not expand capacity quickly enough to satisfy demand.) Further, usage has been increasing ever since at a rapid pace. On the other hand, French Internet users typically have flat monthly rates from their ISPs, but pay by the minute for their local connections. They have on average been spending a constant time online over the last few years, as is shown in Fig. 1.1. (These are Internet users, not Minitel ones, who even in peak years for that service spent under 3 minutes online per day.) The current French usage is similar to that of AOL subscribers before the introduction of flat rates. That the difference in usage is caused primarily by pricing, not by culture, is

shown by an experiment that is taking place right now. In May 1999, Telecom New Zealand introduced flat rates in its XTRA ISP business. The results so far are shown in Fig. 1.1. Under the stimulus of flat rates, New Zealanders' usage has been moving from almost exactly that of French Internet users (and somewhat higher than that of AOL members before the introduction of flat rates in late 1996) towards that of current AOL subscribers.

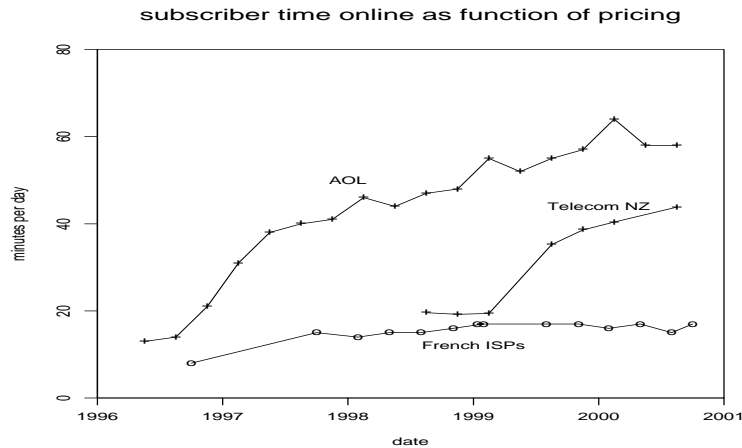


Figure 1.1. Time spent online as function of charging method. AOL and New Zealand Telecom XTRA ISP service introduced flat rate plans in October 1996 and May 1999, respectively, leading to surges in usage. French ISP subscribers pay for each minute online.

The question for service providers and policy makers is whether Internet usage should be encouraged or discouraged. Flat rates are by far the most effective method for stimulating usage. The British and the Japanese have decided that they would like to encourage greater Internet penetration. That is why they are re-introducing flat rates. AOL in the mid-1990s resisted the move to flat rates, correctly fearing the increased network load they were likely to cause. However, just as Dr. Strangelove and The Bomb, AOL has learned to live with and love flat rates. It has decided that its future is in providing more services to its customers. AOL's business plan over the next four years is to triple yet again the time its subscribers spend online [Hansell].

On the Internet, increasing usage is the main imperative for service providers. They do have to make enough money to recover their costs, obviously, but in the long run, they have to encourage their customers to increase usage of the network. Transmission technology is increasing available bandwidth very rapidly. Therefore to avoid ruinous competition like that in long distance voice services today, carriers have to persuade users to take advantage of the new capacity. This is a similar process to that

of the computer industry. The most successful companies are those that manage to sell the latest PCs with the fastest processors, largest memories, etc.

The logic of quality and price differentiation is impeccable. In principle such practices can improve the efficiency of the economy. Unfortunately they conflict with very strong consumer preferences for simplicity, and especially for flat rates. Such preferences are not easy to incorporate into quantitative economic models. What forced AOL to adopt flat rate pricing was pressure from its subscribers, illustrated by the following incident from the fall of 1996:

What was the biggest complaint of AOL users? Not the widely mocked and irritating blue bar that appeared when members downloaded information. Not the frequent unsolicited junk e-mail. Not dropped connections. Their overwhelming gripe: the ticking clock. Users didn't want to pay by the hour anymore.

...

Case had heard from one AOL member who insisted that she was being cheated by AOL's hourly rate pricing. When he checked her average monthly usage, he found that she would be paying AOL more under the flat-rate price of \$19.95. When Case informed the user of that fact, her reaction was immediate. 'I don't care,' she told an incredulous Case. 'I am being cheated by you.'

[Swisher], pp. 160-162

The behavior of this AOL customer is not atypical. A large fraction of U.S. residential users would save if they opted for their ISPs' hourly plans instead of purchasing the \$19.95 per month all-you-can-eat option. Such behavior is invariably treated (when it is treated at all) in works on communications economics as an irrational annoyance that interferes with clever and efficient schemes. For example, here is how one paper on local phone service describes this situation:

... Clearly a movement to a positive per call charge would increase aggregate economic efficiency. Yet nearly all proposals for a move to [usage-sensitive pricing] have met stiff consumer resistance. The reluctance seems to persist even when customers face the prospect of a [usage-sensitive pricing] plan that would, on average, result in a lower monthly bill.

[Panzar]

That paper then goes on to propose a usage-sensitive pricing plan that would hopefully help wean customers from their apparently irrational reluctance to embrace such schemes.

This paper takes a different approach to the problem of pricing. It considers user preferences as a key factor. It presents a view of communications pricing as that of a continuing conflict between the need to optimize and people's reluctance to optimize. The historical evidence shows that, as communication systems have grown and technology has advanced, the balance has moved towards catering to user preferences. The need to extract maximal revenues and to maximize efficiency of the infrastructure have assumed secondary roles.

Quality differentiation and price discrimination strategies are valuable tools, and their use is increasing for good reasons. They are most noticeable in airline pricing, but are spreading to other areas. For example, Coca Cola is experimenting with vending machines that will automatically raise prices when temperatures are high. We can expect such practices to be widely adopted for two main reasons. First, the evolution of our economy is increasing the role of fixed costs in the provision of goods and services. Therefore pricing on the basis of marginal costs is becoming untenable, and it becomes necessary to price on the basis of customers' willingness to pay. That calls for quality differentiation and price discrimination approaches such as those of airlines and Coca Cola. Second, modern information technology is making such practices possible. In the past, Coca Cola might have wanted to price drinks depending on its customers' thirst, but could neither predict the degree of that thirst, nor could it adjust prices in a timely fashion. Now it can do both.

While price and quality differentiation are spreading, in communication services the trend has been towards simplicity. For example, in long distance voice telephony, the most popular plans are the simple ones that are independent of time of day or distance. In the wireless arena, the fastest growth is in offerings such as the AT&T Digital One-RateTM plan, which feature a single payment for a large block of time, and no roaming fees. Even on the Internet, the historical trend so far has been towards flat rates. A decade ago, the Internet was primarily an experimental tool for researchers. The general public was restricted to the mass market online services, such as CompuServe, Prodigy, and AOL. These networks charged not just for minutes of connect time, but even for individual email messages. Email charges were eliminated first, and by the middle 1990s, these services switched to unlimited access for a flat monthly fee. They were forced into this switch by customer complaints and competition from ISPs that offered flat rates. (This was another instance of history repeating itself, since the dominance of flat rates for residential local calling in the U.S. appears to have resulted largely from the competition between phone companies a century ago.) The attempt to move the Internet back towards usage-sensitive charging might thus be regarded, in Samuel Johnson's words, as "a triumph of hope over experience."

The trend towards simplicity noted above is not new. In later sections many more examples will be presented, based on two centuries' worth of data on the evolution of mail, telegraph, telephone, and data services. Users value simplicity, and in particular flat rates. They like best a single uniformly high level of service for a fixed fee. Historically, even when fixed-fee subscriptions were not offered, the trend has been to simplify the rate structure. (For other work that considers user preferences in communications, see [BouchS, BouchSDM].)

How do customer desires for simplicity translate into incentives for service providers to avoid complicated price and quality differentiation strategies? The answer appears to be that as economies of scale and technological change lower unit costs and increase frequency of usage, service providers can collect more money through simple plans. This is explored in more detail in sections 5 and 6. Section 8 discusses additional incentives that service providers can employ to increase usage, in order to benefit from network effects and to enhance the chances of migrating customers to more lucrative services.

The history of communication suggests strongly that as services become less expensive and are used more widely, the balance shifts away from the need to segment the market, and thereby to extract maximal revenues and to maximize utilization efficiency of the infrastructure. Instead, customer desire for simplicity becomes dominant. This phenomenon is especially pronounced at the level of individual consumers. The business-to-business market is different from the business-to-consumer market, and the focus of this paper is on the latter, where individual preferences matter the most. There is no sharp dividing line between the two markets, but there is a substantial distinction. For example, McDonald's offers free sugar to its customers, and builds the cost of this service into the price of the coffee. It would surely be more efficient in terms of allocating resources to charge for each packet. It would also be fairer, in that customers who do not use sugar would not be subsidizing those who do. (It would also be healthier for the customers, as consumption of sugar would undoubtedly decrease!) Yet that is not done, and the customer desire for simplicity, as well as McDonald's desire to keep transaction costs low, lead to "wasteful" practices. On the other hand, McDonald's buys its sugar in bulk, and undoubtedly has purchasing experts who play off various suppliers against each other, arrange long-term contracts, and probably even trade on the sugar futures markets. This shows the spectrum of business decisions on how far to optimize. The general tendency for businesses appears to be to optimize when the optimization can be done by dedicated specialists (such as the sugar buyers at McDonald's). When the optimization requires many small actions by large groups of employees, the tendency is to opt for simplicity.

The general conclusion is that we should strive for simplicity, even at the cost of efficiency. That

is how the world of communications has been evolving for the past two centuries, and that is how it is likely to evolve in the future.

Note: The history of communication pricing is a vast subject, and it is impossible to do it justice in the space of this paper. Much more detail is available in the manuscript [Odlyzko5].

A section at the end, between the Acknowledgements and the References, lists sources for the data presented in the tables and figures. It is worth emphasizing here that all the monetary figures are in current dollars, and thus are not adjusted for inflation. Price indices for the 19th century are rather imprecise, and in any case inflation-adjusted costs do not present an accurate picture of the burden of prices back when living standards were much lower. Therefore some prices are stated in terms of hours of work required to pay that price.

2. Growth in communications: quantity and quality

The explosive rise of the Internet is only the most recent chapter in a remarkable history of humanity becoming increasingly connected, the “annihilation of space and time,” in a phrase going back at least to the early 19th century. The volume of messages addressed to an average individual has been increasing steadily, at times extremely rapidly.

This section surveys the long trend of increasing communication. The first part is devoted to quantitative measures, the growth in spending and in volume of transactions. The second part considers the qualitative changes in how we approach communications. The role and evolution of pricing are then the subject of the rest of the paper.

The communications technologies considered in this paper are primarily point-to-point ones, and generally exclude mass media, such as newspapers, book publishing, radio, and TV. One reason is to keep the size of this work manageable. A related one is to avoid the complexities of trying to measure the volume of information delivered by broadcast media. Finally, as is shown in [Odlyzko5], connectivity (as in point-to-point communication) is much more important than content (which is what broadcast communication is about).

The communications industry as a whole has been expanding for centuries. Not only that, but most of the prominent services have continued growing. The tables in this section provide more eloquent testimony to this historical fact than words can. Even the venerable mail, often derisively called “snail mail,” is still growing. The remarkable story of this ancient yet still widely used technology is summarized briefly in Table 2.1, which shows key statistics in the development of the U.S. Postal Service (USPS). (For most of its history it was known simply as the U.S. Post Office.) Although there have been

Table 2.1. Growth of the U.S. Postal Service

year	expenditures (millions)	expenditures as percentage of GDP	pieces of mail (millions)	pieces of mail per person per year
1790	\$0.032	0.02%	0.8	0.20
1800	0.214	0.05	3.9	0.73
1810	0.496	0.09	7.7	1..07
1820	1.161	0.18	14.9	1.55
1830	1.933	0.21	29.8	2.32
1840	4.718	0.28	79.9	4.68
1850	5.213	0.20	155.1	6.66
1860	14.87	0.39		
1870	24.00	0.33		
1880	36.54	0.35		
1890	66.26	0.51	4,005	63.7
1900	107.7	0.58	7,130	93.8
1910	230.0	0.65	14,850	161
1920	454.3	0.50		
1930	803.7	0.89	27,887	227
1940	807.6	0.81	27,749	211
1950	2,223	0.78	45,064	299
1960	3,874	0.77	63,675	355
1970	7,876	0.81	84,882	418
1980	19,412	0.70	106,311	469
1990	40,490	0.70	166,301	669
1998	57,778	0.68	197,943	733

many predictions that mail would decline as a result of competition from the telegraph, the telephone, and more recently from email, it has continued to expand. It went from fewer than a million pieces of mail in 1790 to almost 200 billion in 1998. Prices, discussed in Section 9, have decreased. The decline in prices was dramatic in the 19th century. Even in recent years prices have been decreasing, if we adjust them for increases in earnings.

Postal system revenues are more than three times larger as a fraction of GDP than they were 150 years. In spite of this, mail is no longer the dominant communication technology, as it is dwarfed by the phone industry. Comparison of tables 2.1 and 2.2 shows the telephone passed the mail in revenues by 1920, and today is about four times as large. Adding the contributions of these two industries, we find that their role in the economy, measured as a fraction of GDP, has increased almost 20-fold over the last 150 years, from 0.2% to over 3.5%. That is higher than current military spending!

Telephone service initially was also extremely expensive. For example, a century ago, unlimited

Table 2.2. Development of U.S. phone industry

year	revenues (millions)	revenues as percentage of GDP	phone calls per day (millions)	phones calls per person per day
1890	\$16	0.12%	0.0015	0.00002
1900	46	0.26	7.8	0.1
1910	164	0.46	35.6	0.4
1920	529	0.58	51.8	0.5
1930	1,186	1.32	83.5	0.7
1940	1,286	1.29	98.8	0.7
1950	3,611	1.27	171	1.1
1960	8,718	1.73	288	1.6
1970	19,445	1.99	494	2.4
1980	59,336	2.13	853	3.8
1990	133,837	2.33	1,272	5.1
1998	246,392	2.88	1,698	6.3

local service in New York City cost \$20 per month. That is comparable to \$2,000 per month currently, when one considers prices in relation to average earnings. How many people would be willing to pay \$2,000 per month for Internet connectivity today? That such prices were paid testifies to the value of communication services as perceived by users. The precise economic benefits of communication are a complicated subject beyond the scope of this work. However, they have consistently been high, often demonstrably so. (At a high level, this claim is proved by the statistics in the tables throughout this work. Since nobody was overtly forced to pay for any of these services, yet usage and spending went up, customers must have felt they were getting their money's worth.)

The growth in volume of information has led to a change in how we react to it. Over the last two centuries we have moved from scarcity to surfeit. People used to be information-deprived, and eager for any scraps they could get. The books [John, Kielbowicz] have many illustrative examples of the eagerness with which mail was received in 19th century America. Today, our task is to cope with a flood of information.

When communication opportunities were rare, and transmission slow, it was natural that great care was taken with correspondence. Those who lament the lack of style in current letters compared to the often essay-quality compositions of the 17th or 18th centuries need to realize the different environment we operate in. We do not have weeks to compose a letter, and speed is of the essence. This trend is exacerbated with email. Email messages are often sadly deficient in style, spelling, punctuation, and grammar. Instant messaging is typically even worse. However, when it is necessary to deal with

scores of email messages per day, it is natural to treat them as informal conversations. After all, are we expected to always speak in grammatically correct sentences? The same argument, though, argues that we will not want to worry about pricing details.

For a communication service provider, the task is to make life for the customer as painless as possible. A seamless integration of the Internet into people's lives should be the goal. One of the most successful companies in this area, at least so far, has been AOL. An AOL press release, dated February 2, 2000 [AOL], claimed 21 million members, as well as usage records, as follows:

Among the AOL service's other usage records, underscoring that it is becoming more and more essential to members' everyday lives, are:

...

- ** 110 million e-mails sent daily (50 million a year ago);
- ** 600 million instant messages sent per day (400 million a year ago);
- ** 200 million stock quotes daily (113 million a year ago); and
- ** 5.2 billion Web URLs served a day (2.1 billion a year ago).

It is clear that AOL measures its success by how deeply it is involved in its customers' lives. As we move towards broadband always-on access, it will be possible for carriers to provide even greater services. The question will be how to choose the quality and price combinations to encourage greatest use. Efficiency is likely to be of secondary value.

3. The effectiveness and utility of pricing

For any service to be viable in the long run, it has to recover its costs. Even the supposedly "free" Internet is nothing of the sort, as it is largely paid for through monthly fees by individual users or their employers or universities.

How costs are recovered from users is a matter of pricing. This section is a brief overview of some purposes that pricing can serve, aside from the fundamental one of cost recovery. For example, the choices that users make among different quality and price combinations provide valuable information for the service provider about what the customers wish and are willing to pay for.

There is a wide spectrum of options for pricing communications services. To illustrate, let us consider the regular postal service. At one extreme one could have a scheme in which the mail carrier

would collect letters, and for each one would bargain with the sender, taking into account the destination, the length of the letter, the likely interest the sender has in getting the letter delivered, how heavy the carrier's mailbag is, and so on. Moving slightly away from this extreme towards greater simplicity would be a fixed schedule of charges, covering each possible source-destination pair of houses. Such a schedule might be based on some allocation of postal costs, and would depend on how close to the nearest post offices the sender and receiver live, how much other traffic arrives at or leaves their houses, and how much mail their neighbors originate. A simpler scheme yet is one in which the charges depend just on the distance between the origin and destination. (Among such schemes, a more complicated version would depend on the distance along the actual physical route taken by a letter, a simpler one would depend on just the air distance.) Simpler yet would be a scheme that has a fixed charge, independent of distance. Finally, the simplest system of all would have no charges at all, and would provide a free service, paid for by general taxation.

In addition to the range of options sketched above, there are several others that can be added, almost orthogonal to that one. For example, one can have, for most of the options listed above, time-of-day pricing, or quality differentiation with respect to speed of delivery.

Most of the feasible price and quality combinations have been tried in the past. For example, before postal systems became widespread, getting a letter delivered required finding somebody traveling in the right direction and negotiating for delivery. On the other hand, at least for some time in the mid-1850s, Iceland had mail delivery paid for out of general taxes [Miles]. Historically, from the early 18th century until today, the postal system has moved from charges depending in complicated ways on source, destination, and route, to a uniform fee per piece, up to some weight or size limit that few items are likely to reach. There is service quality differentiation, with several classes of mail, and even different prices for first class mailings depending on the degree to which they are pre-sorted, say. However, most of these gradations in price and quality are aimed at influencing large mailers. Individuals are presented with very simple options.

On the Internet, many pricing proposals call for complicated schemes with charges varying from packet to packet. Such schemes can be proved to have various optimality properties, whether the basic philosophy is to have charges closely related to costs, or to maximize revenue. (For one such proposal, employing a Vickrey auction, which has been tremendously influential in stimulating thinking about Internet pricing, see the paper of Jeff MacKie-Mason and Hal Varian [MacKieMV].) On the other hand, most charges today are through flat rates. There is little advocacy for public funding of the Internet, other than for research networks, although a few calls are still heard occasionally for public

funding of peering points, to avoid the sticky interconnection problems. The general attitude is that communications users should fund the services they consume. A competitive service environment certainly requires that. The question is, how closely should one attempt to link the fee that a customer pays to either the costs that customer's usage generates, or to the value that customer derives from the service?

Historically, the relation between cost of providing a communication service and the price charged for it has been slight, as we will see repeatedly in later sections. Cross-subsidies and price discrimination have been common.

In our society, the word discrimination has a strongly negative connotation. However, price discrimination is a common business tactic that is growing in importance, and serves a useful social purpose. For detailed discussion of the different types of price discrimination, and crucial role that it plays in making the economy more efficient, see [Varian1, Varian2].

In practice, any seller that wants to price discriminate has the difficult task of finding out the customers' willingness to pay and getting them to accept the discrimination. A common way to solve this problem is to impose artificial restrictions that will induce the customers to sort themselves out. Such sorting out of users is the main purpose of the advance purchase and Saturday night stayover restrictions on inexpensive airline tickets. The assumption that carriers make is that business travelers will often need to make last-minute changes in their schedules, and also will be unwilling to extend their time away from home over a weekend. This does work much of the time. It allows airlines to offer inexpensive fares to the vacation travelers who can adjust their schedules to fit the restrictions, and who might not travel at full fares. The result is that often there are two passengers sitting next to each other on the same flight, but one of them has paid more than five times as much as the other one. The system does serve to produce high utilization, decreasing average fares, and also availability of seats on most flights even at the last minute (although at high prices).

The airline yield management system does work, but it is generally disliked by the public. People find the bewildering array of constantly changing prices and conditions unpleasant, exasperating, and unfair. One way to make price discrimination more palatable to the public is through quality differentiation. The airlines practice that also by providing several classes. First and business classes get comfortable seats and good meals, while coach passengers are squeezed into tiny spaces and receive minimal meal service.

So far we have discussed arguments for price discrimination. On the other hand, there are many cases where the incentives for such discrimination clearly exist, but it is not practiced. For example,

voice services bring in huge revenues. On the other hand, they require low bandwidth. Hence on a broadband network, if transmission of video is to be affordable, and charges are to be based on bytes transmitted, voice should be essentially free. Yet voice services is still what provides the overwhelming bulk of revenues to the communications industry. Hence the argument has often been made that voice transmission should be charged at higher rates than video. This type of price discrimination is appealing on public policy grounds. However, the same argument would argue for high charges for email. After all, email is still the “killer app” of the Internet, even more than the Web, and the bandwidth it requires is modest compared to Web surfing. Moreover, special email charges would (as many people have suggested) help control spam. Yet, historically, charges for email, which used to be common in online services, have been phased out. (Not only that, but we have a multiplicity of services offering free email accounts.) And indeed, with scores of email messages per day for most “wired” people, who would want to deal with the extra complexity of having to decide for each one whether it is worth paying for it?

The general trends in communications, as is documented in later sections, has been towards decreasing price discrimination, and towards simplicity. On the other hand, in the general economy, there is also a perceptible tendency towards complicated pricing, such as yield management strategies, auctions, and the like. This is evidenced not only by the airlines, but also by the initial successes of eBay and priceline.com. My conclusion is that the conventional economic arguments for quality and price differentiation are all valid. However, they conflict with some basic human desires. For inexpensive and frequently used goods and services such desires are strong enough to overcome the economic push for optimization.

It is hard to make categorical statements about human preferences, since people exhibit contradictory traits. They often simultaneously increase and decrease their risks, as when they buy both lottery tickets and insurance. However, generally speaking, people are risk averse, especially in the face of large potential losses. That is why the insurance industry is vastly larger than the gambling industry. In general, people are also averse to varying prices, and are more willing to accept variations in quality than in price, as is illustrated by many examples in [Odlyzko2]. That paper also discussed such economic puzzles as the laws against scalping. Such laws are a puzzle for two reasons. One reason is that according to standard economic doctrine, scalpers provide a useful function, by moving desirable tickets into the hands of those who value them most highly. Thus their business should be encouraged, and not driven underground. Another reason such laws are a puzzle is that from the conventional economic perspective it is hard to explain why they should be necessary at all, no matter how one views

scalpers. Why is there room in the economy for scalpers? Why don't the owners of sports arenas and concert halls adjust prices to clear the market in the first place? The most satisfactory explanation seems to be that both sports arena operators and lawmakers are reacting to human preferences for simple, predictable, and fair pricing. Other examples of similar seemingly irrational behavior are cited in [Odlyzko2]. In most cases, though, such examples serve as illustrations only, and do not provide hard estimates. In general, it appears likely that appearing to follow general notions of fairness will be increasingly important in ecommerce ([Odlyzko1, Zajac]), at least for small value transactions.

4. The role of distance dependence in pricing

The most frequently used communication service, and the one that attracts the most spending, is the phone. Charges for phone use have traditionally depended on the distance of the connection. Local calls were usually the least expensive, and prices of long distance calls rose steeply with distance. The one common communication service whose tariffs did not vary with the distance was the mail. (Even there, foreign letters invariably carried higher fees than domestic ones.) When charges were distance-independent, this was often a cause for comment, and was often referred to as a case of "postalized rates." In particular, the Internet has often been pointed out as unusual in ignoring distance. Occasionally worries have been expressed about the harmful effects of this policy, which ignores the costs of long distance transport. Of special concern have been the expensive trans-oceanic links. In some cases (such as that of the British academic JANET network), charges have been imposed on member institutions for traffic coming from the U.S., but not from Europe.

While costs of long distance transport are a concern to many Internet experts, the historical trend in other communication services has been to increasingly ignore them. For example, in long distance telephony in the U.S., rates became distance insensitive in the mid-1990s. A key role in this development has been played by technology, which has lowered transport costs. However, this paper avoids detailed consideration of internal costs of service providers, and considers primarily the prices consumers have been faced with. Originally, even postal fees were not "postalized." They became distance-independent only in the 19th century, starting with the famous British postal reform of 1840. Since that time, many communication services have traveled either partially or all the way towards abandoning distance sensitivity. This trend in prices is illustrated in tables 9.2, 10.1, and 10.2.

A simple argument says that if prices are distance-insensitive, then the basic network costs (of switching and transmission) must be a small part of the price that is charged. But if those costs are low, then the costs of tolerating various inefficiencies (such as carrying low-value communications) are also

likely to be low. That argues for a single grade of uniformly high quality service, and simple pricing.

An even more convincing case for simple pricing and simple service offerings can be developed based on a detailed historical study of the distance sensitivity of rates. Frequently such distance sensitivity was far greater than was justified by cost differences. It was a form of price discrimination. This practice reached an extreme form in the U.S. phone industry in the 1970s and 1980s, when long distance revenues were used to subsidize local service. The postal rates in Britain before 1840 (as well as those in most countries in the early 19th century) were also examples of price discrimination, charging far more for long distance communication than was justified by costs.

Why would carriers impose charges that increased more steeply with distance than costs? The basic reason is the perception that long distance transmission is more valuable. In U.S. voice telephony, regulators argued that expensive long distance calls were a luxury for the rich, and so could in effect be taxed to provide basic service for the general population. More generally, such practices are a form of value-based pricing. This role has been recognized for a long time. For example, in the discussions preceding the British postal reform of 1840, which did introduce distance-insensitive prices, we find the following claim:

Distance is the chief ingredient—indeed, we might say, the very *sine qua non* and essence—of any and every system of post-office charge... If it were not for distance, there would be no post offices at all. The intrinsic value of the conveyance of a letter... is exactly equal to the time, trouble and expense which is saved to the correspondence—of which the best, if not the only measure is distance; and as the difficulty of private conveyance increases, so much increase, in graduations proportional to the distances, the value of the conveyance. The gods must annihilate both time and space before a uniform rate of postage can be reasonable or just.

J. W. Crocker, *Quarterly Review*, 1839

(as quoted in [Gregory])

Well, gods did annihilate both time and space. The uniform Penny Post was introduced in 1840, and has historically been regarded as a great success.

The wide variation in costs of transmitting messages can be used to justify non-uniform prices. However, it is interesting to observe that many of the price schedules that were not uniform depended on a simple formula that took only distance into consideration. (See tables 9.1 and 9.2, for example.) Yet if a carrier were to do what is theoretically socially optimal, namely have prices close to marginal

costs or to willingness to pay, it would have much more complicated price schedules. Costs are only weakly related to the distance a message travels. The main determinants of costs tend to be factors such as particular layout of a network, and the traffic on individual links. This appears to have been pointed out in a public forum for the first time by Rowland Hill in 1837, in discussions of pricing in the British postal system. A value-based pricing plan would also surely not depend in a simple way just on distance. And indeed, when we examine the competitive spot markets for long distance bandwidth that are emerging, we find that the old tariffs based on distance are disappearing, and prices vary from place to place and from day to day. However, this is for links that cost tens of thousands of dollars per month. In the early days of many mass communication services, such as the telegraph or the telephone, when costs and prices were high, we also find prices varying, depending on particular circumstances. Later, decreasing prices were associated first with a move to simple pricing based just on distance. Such rates thus already represent a considerable concession to the desire for simplicity! As a technology evolves, costs drop, and usage increases, simple distance-only rates tended to get replaced by even simpler distance-insensitive prices. This was true for both government monopolies, private monopolies, as well as competitive carriers.

5. The conventional economic argument for flat rate pricing

There are strong incentives for service providers to engage in price discrimination, and increasing opportunities to do so. However, there are also arguments for flat rate pricing. A frequently cited one is that of reduced costs for service providers. A more substantial reason for preferring flat rates is that they reduce transaction costs of the users. The predictability of money flows of flat rates is attractive for both sides.

A major advantage of flat rates for service providers, much more important than the ones mentioned above, is that such rates represent a form of bundling. Bundling is the strategy of offering several products or services for a single price, and is common. A frequently cited argument for bundling is that it simplifies life for consumers by providing a single bill and a single point of contact. Such convenience is desirable. If that were the complete explanation, though, bundling might justify extra charges for the added convenience. In practice, though, bundles are almost always priced lower than the sum of prices of their components. For example, the Microsoft Office bundle typically costs about half of what purchasing Word, Excel, Powerpoint, and the other pieces separately would cost. Thus the main justification of bundling has to be something other than convenience.

The main incentive for service providers to bundle is that it allows them to obtain higher revenues.

They do this by taking advantage of uneven preferences among consumers for various components of the bundle, and thereby reduce consumer surplus.

Bundling (which has been studied extensively in economics) is not always more profitable than selling a-la-carte. However, for most reasonable distributions of demands, bundling is better for the service or good provider. (However, mixed bundling, in which the various items are offered for sale separately as well is always better.) In general, the more items there are in the bundle, the greater the gain in bundling (at least for zero marginal costs).

Flat rate pricing plans are a form of bundling. Consider a simple example. Suppose that among the hundreds of millions of Web sites, Alice is interested in just 5, and would like to download 2 MB per month from each. However, she is willing to pay \$1.00 per month for the material from site 1, \$2.00 for that from site 2, and so on up to \$5.00 per month for access to site 5. Thus she should in principle be willing to pay a flat rate of \$15.00 per month for access to all the sites. However, on a per-byte basis she is only willing to pay \$0.50 per MB for material from site 1, \$1.00 for material from site 2, and so on up to \$2.50 per MB for material from site 5. Now suppose that Alice's ISP charges strictly per byte volume. If the price is set at \$1.50 per MB, Alice will pay only for access to sites 3 through 5, and will download 6 MB per month for total revenues to the ISP of \$9.00. Any other price will result in even lower revenues for the ISP. Thus in this example, flat rate pricing produces revenues that are 67% higher than with volume charging.

The example above does not produce conclusive proof that flat rate is better for the ISP than volume charging. The gain of 67% shown there holds only if the cost of carrying the additional traffic in the flat rate scenario is negligible, and if the ISP can guess that \$15.00 is Alice's maximal willingness to pay. If the ISP prices the flat rate service at \$20.00 per month, Alice will not sign up, and revenue will be zero.

Suppose that providing the additional service that flat rate stimulates carries negligible marginal cost. (This is not true for residential modem access to the Internet, since modem maintenance is the most expensive part of running such a business. However, for cable modem or ADSL services that are always on, this can be almost true.) Let us further assume that, as above, users know what utility they derive from each possible bit they might receive or send. Under those conditions, whether a service provider obtains more revenue from a flat rate or a usage-based tariff depends on the distributions of demands among potential customers. This problem is considered in [FishburnOS], where it is shown that for a majority of what seemed to be a large class of realistic demand distributions, flat rate produces higher revenue for the service provider.

The point of the discussion above is that even in the conventional model with simple consumer utilities, it is often more profitable for the service providers to offer flat rate service. As it turns out, there are other factors that make flat rate pricing even more compelling.

6. The strong public preference for flat rate pricing

The frequent willingness to pay extra in order to avoid metered services was noted in the Introduction, in the discussion of how AOL was forced to move to flat rates. This is a general phenomenon. Perhaps the first well documented large scale experiments in this area were those carried out by the Bell System in the 1970s. By that time the United States had had unlimited residential local phone calling for a flat monthly rate for over half a century. (There were exceptions to this rule, for example New York City and Chicago, but most of the country was on flat rate plans.) In the Bell System experiments of the 1970s, customers were usually offered a choice of three plans. The least expensive (in terms of the basic monthly rate) allowed for only a few local calls without additional charges, and then had high fees for additional calls. The most expensive plan provided unlimited local calling. It was a great surprise when most of the people decided to stay with flat rate plans. This even included the majority (typically over 60%) of the people who were making practically no phone calls, and would have benefited the most from the usage sensitive tariffs. The results of these experiments are described in the references to [FishburnOS].

The Bell System studies identified several reasons for the flat rate preference, such as risk avoidance and overestimate of usage. Some of these reasons are part of what Nick Szabo has called “mental transaction costs” [Szabo]. Computer processing power and bandwidth are growing and getting less expensive. However, our personal processing power is not increasing. Human attention is an ultimate scarce resource.

Since the 1970s, much additional evidence has accumulated of the strong preference for flat rate plans. Recently, further detailed evidence of the attractions of flat rate plans has been obtained in the carefully controlled INDEX experiment. This experiment, described in [AltmannRV, EdellV], tested users’ willingness to pay for various Internet access options. Most of the INDEX investigators appear to conclude from the experimental data that differentiated services and usage-sensitive pricing would be better for both ISPs and users (cf. [EdellV]). However, one can also interpret the data as supporting flat rates. One argument in this direction is based on the dramatic effect that metered billing has in decreasing usage, and is presented in the next section. Right now we mention another argument. In some phases of the experiment, participants were offered a choice of paying a flat fee to free themselves

from the constraints of metered billing. It was discovered that on average, these participants were willing to pay considerably more in flat fees than through metered billing [Chu]. Their willingness to do so in a commercial setting would probably be much greater than it was in the INDEX experiment. In INDEX, they had excellent and immediate feedback on their usage (and thus on their real or potential bill), easy control over their options and could change their pricing plan each week. In a commercial setting, where they might not see their charges until their credit card bill arrived six weeks later, they would likely be willing to pay considerably more to avoid unpleasant surprises.

7. The dramatic effect of usage sensitive prices

Fig. 1.1 shows how a switch from metered to flat rates increases usage. It is also known that imposition of metered rates decreases usage. Although there have been few careful studies of this, it was widely known that when various phone systems around the world switched from flat to metered rates for local calls, usage dropped. (See [Kraepelien] for the best listing of data on this subject that I have been able to find.)

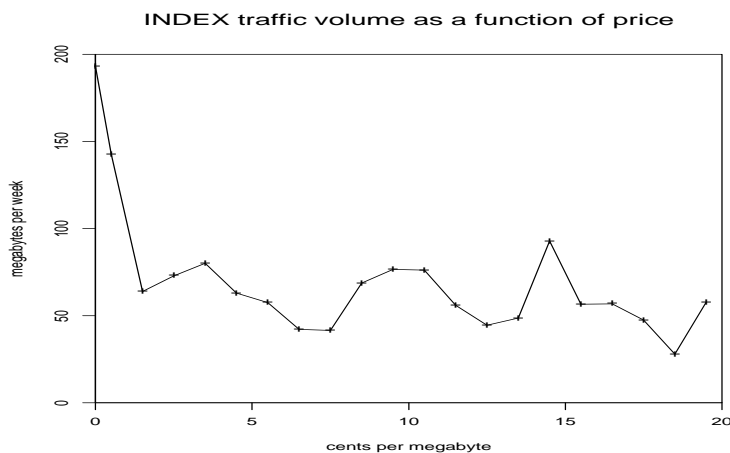


Figure 7.1. Data received by participants in the per-byte charging part of the INDEX experiment as a function of price.

The decreases in usage that are caused by metered rates can in principle be explained in standard economic terms. All that happens is that the transactions with low value to the user are not carried out, which increases aggregate welfare if that value is less than the cost to the network. The problem with this conventional explanation is that many of the observed decreases are considerably larger than one

would normally expect. That this is indeed a problem, and that a psychological explanation is more appropriate, can be seen in the data from the INDEX experiment, mentioned in the previous section, and described in more detail in [AltmannRV, Chu, EdellV]. During one phase of that experiment, participants were subject to straight per-byte charges. In the initial week, all 56 participants in that phase had free service, and downloaded an average of 193 MB each. During each of the following six weeks, they had to pay for all traffic to their homes, with prices chosen according to the principles of statistical experiment design, and taken from the range of \$0.001 to \$0.20 per MB. (It is worth noting that residential modem customers typically download about 60 MB per month. Thus, if they are charged \$20 by their ISP for unlimited service, they are effectively paying about \$0.33 per MB. Cable modem or DSL customers typically download 5 to 10 times as much, but usually pay between \$40 and \$60 per month, for an effective price of between \$0.07 and \$0.20 per MB. Thus the prices in the INDEX experiment were not very high.) As is shown in Fig. 7 in [EdellV], even low prices led to a big drop in usage compared to the week with no charges. However, that bar chart aggregates data to such an extent that it is hard to tell what the effect of very small charges is. The INDEX experimenters have kindly supplied me with the raw data for that chart, and I have processed it to produce Fig. 7.1. (The interpretation and presentation of the data are strictly mine.) The entry for the price of 0.5 cents per MB is the average of the subject-weeks where the subject faced a price of between 0.1 and 1 cents per MB, the one for a price of 1.5 cents per MB is the average for prices between 1.1 and 2 cents per MB, and so on. Each data point corresponds to an average of 14 subject-weeks, so that experimental error was substantial, and accounts for the jagged line. (There was clearly nothing special about prices between 14 and 15 cents per MB to induce a large jump in usage!) The main point is that even prices below 1 cent per MB led to a giant drop in usage (to an average of 143 MB per week from the 193 MB during the week of free service). Prices between 1 and 2 cents led to a further drop to an average of 64 MB per week. Further increases had a much smaller effect on usage. Yet even at 2 cents per MB, the usage of 200 MB/week that was observed during the free week would have cost just \$4 per week. It seems more reasonable to attribute this behavior to a psychological aversion to metered rates than to a utility function that has an unusual shape.

8. Dynamic effects of flat rates

Whether to offer flat or metered rates in a competitive environment is a business decision. One way to select the preferred course is through a traditional accounting analysis. Find out how much potential customers are willing to pay extra for the flat rates. (Either run an experiment like INDEX,

or do market trials, focus groups, and other standard methods of marketing.) If this amount is large enough to compensate for the inefficiencies introduced by a move away from metered pricing, give the customers what they like, namely flat rates. If the willingness to pay for flat rates is too low, offer only measured rates. While such an analysis seems appealingly rational, it is likely to be seriously misleading in an environment such as that of the Internet, where rapid growth is the rule. The problem is that this analysis is static, and ignores both network effects and temporal evolution. To properly evaluate the choice between flat and measured rates, one has to take into account the indirect effects of those rates on other users, as well as on future usage of a particular customer.

The value of a network is based not just on the number of users in it, but, even more, on the intensity with which it is used. If you are connected to a network such as the Internet, you can be reached by others, and you may contribute content to the network. You are also more likely to try new services and products if you are online often. Thus your presence makes the network more valuable to others. Hence some of the costs of your usage can justifiably be spread over other participants. Thus the loss in efficiency that flat rates cause is smaller than it might seem at first, and the benefits from increased usage that flat rates stimulate are greater.

A related reason that the static analysis of willingness to pay for flat rates is likely to be misleading is that it neglects the dynamic aspects of the Internet connectivity market. Network service providers have every incentive to persuade customers to move up to higher quality connections. The first step is to move them from dial modems to cable modems and ADSL links. The next step will be to persuade them to pay for fiber connections, on which bandwidth can be increased step by step. Given the progress in technology, providers are forced to plan on persuading their customers to trade up. A fixed link is too likely to be commoditized, and unable to offer the services that people come to expect. Who is most likely to want to move to a higher bandwidth connection? Is it an American modem customer who pays flat rates for local connectivity and Internet access and is online about an hour per day (Fig. 1.1)? Or is it a French modem customer who pays by the minute, and is online for 17 minutes per day? Or is it a French Minitel customer of a few years ago, who paid by the minute and was online for under 3 minutes per day?

The above two effects cannot be quantified yet, with our current state of knowledge, but they appear substantial. They certainly did play a role in the evolution of pricing of various communication services, as we will see later, and they argue strongly in favor of flat rates.

9. Mail

Postal systems are ancient, and there is a voluminous literature on their history, technology, politics, and economics. This literature all by itself illustrates the conflicting drives for price discrimination, efficiency, and simplicity. that have shaped the evolution of communications. Here I present just a few highlights from this literature. For a fuller treatment, see [Odlyzko5].

The earliest postal systems were primarily or even exclusively for government use. With the growth of commerce, and the need for economy in operation, these systems were soon extended to carry private letters. The mail business was often extremely profitable, and was often used as a form of taxation.

The extraction of profits from postal services may well have reached its highest recorded level in Britain in the early 19th century. That country was desperately searching for money to pay for the Napoleonic and earlier wars and the debts from those wars. The mail turned out to be a bountiful source of funds. In 1839, total revenues of the monopoly government postal service were \$11.95 million. The costs of providing the service were \$3.78 million, for an enviable profit margin of 68%.

The high profit margins of the British postal service were accompanied by a complicated pricing structure illustrated in Table 9.1. This structure did not correspond to the cost structure of providing the service. Interestingly enough, the steepness of the distance dependence of British postal rates increased rather steadily throughout the 18th century, as prices were raised to increase profits. The “single letter” of Table 9.1 (as well as of Table 9.2, which shows a few snapshots in the evolution of pricing of U. S. postal services) refers to a single sheet of paper, the standard unit for pricing mail services then.

Table 9.1. British Post Office rates for domestic “single letter” in 1837

distance	price
up to 15 miles	\$0.083
16-20 miles	0.104
21-30 miles	0.125
31-50 miles	0.146
51-80 miles	0.167
81-120 miles	0.188
121-170 miles	0.208
171-230 miles	0.229
231-300 miles	0.250
301-400 miles	0.271
401-500 miles	0.292

The complicated rates of Table 9.1 were replaced in 1840 with the uniform price of a (British)

penny, approximately \$0.02, for any letter up to half an ounce anywhere in the country. This reform, justly credited to Rowland Hill, is often misrepresented. It is frequently claimed that it was a triumph of common sense. It is said that once Hill pointed out the high costs of computing charges based on distance, charges supposedly much higher than those of transport, the system was changed with a universal gain for everybody. The truth is considerably more complicated, and in fact there are serious debates going on about the effects of Hill's reform. A summary with pointers to the literature is presented in [Odlyzko5]. At this point I will only say that the two points on which there is universal agreement is that the 1840 reform was extremely popular with the public, and that it did lead to substantial growth in mail usage in Britain. Hill's reform was also extremely influential in reforms of other countries' postal systems (including that of the U.S.), leading to almost universal dominance of distance-insensitive rates.

The postal service of the United States was not (except for a brief period during the War of 1812) operated to make a profit. It was a key instrument in the efforts by the federal government to knit the country together. Still, the pricing structure in the first half of the 19th century was complicated, as is shown in Table 9.2. This (as is explained in more detail in [Odlyzko5]) was again not because of costs depending on distance, but rather because of attempts to maximize revenue from letters, so as to be able to subsidize newspaper delivery more effectively. Under intense pressure from the public, rates were lowered and simplified in the middle of the 19th century, becoming independent of distance.

10. Telegraph

The electric telegraph is a fascinating subject to study, not least because its influence was so similar to that of the Internet. It is unusual among communications services in that it did eventually vanish.

A key point in considering the history of the telegraph industry is that it was never large. The high cost and inconvenience kept the telegraph from ever rivaling the mail, for example. However, the trend towards simplicity in pricing is evident here as well.

Tables 10.1 and 10.2 show the evolution of pricing of telegraph services. The 19th century and early years of the 20th century show a rapid decrease in prices. However, later years show an increase. This is one way that the telegraph deviated from the pattern of other communication services. There was an even more basic difference. The telegraph is the only major communication technology to fade away. The price increases in its last stages were part of a vicious spiral. Increasing prices led to decreased demand, which led to more price increases, and so on. Businesses whose costs are largely fixed benefit tremendously from increases in scale, but have difficulty coping with declines in

Table 9.2. U. S. Postal Service rates for first class mail

year		price	hours of work
1799:	single letters		
	no more than 40 miles	\$0.08	0.8
	41-90 miles	0.10	1.0
	91-150 miles	0.125	1.25
	151-300 miles	0.17	1.7
	301-500 miles	0.20	2.0
	over 500 miles	0.25	2.5
1845:	single letters		
	no more than 300 miles	0.05	0.3
	over 300 miles	0.10	0.6
1863:	first half-ounce	0.03	0.2
1885:	first ounce	0.02	0.1
1999:	first ounce	0.33	0.02

volume of operations. Still, prices did tend to become simpler with time, with distance dependence of tariffs diminishing or disappearing entirely.

11. Wired voice phone

The telephone attained its present preeminent position among communication services relatively slowly. For example, even in the United States, it was not until the 1910s that the revenues of the phone industry exceeded those of the postal system. In most countries it took much longer to reach that stage.

Table 10.1. International telegraph rates from New York City (per word)

year	London	Tokyo
1866	\$10.00	-
1868	1.58	-
1880	0.50	\$7.50
1890	0.25	1.82
1901	0.25	1.00
1924	0.20	0.50
1950	0.19	0.27
1970	0.23	0.31

Table 10.2. Telegraph rates (for up to 10 text words) from New York City

year	Philadelphia	Chicago	San Francisco
1850	\$0.25	\$1.55	-
1866	0.25	1.85	\$7.45
1870	0.25	1.00	5.00
1883	0.15	0.50	1.50
1908	0.25	0.50	1.00
1919	0.30	0.60	1.20
1951	0.60	1.00	1.60
1960	1.10	1.45	1.90
1970	2.25	2.25	2.25

The main reason was that the phone had to build its own infrastructure. It also had to compete with the telegraph, which was almost ubiquitous.

Table 11.1. International telephone prices. Standard rate for a 3-minute call from New York City to London

year	current dollars	hours of work
1927	\$75.00	200
1928	45.00	120
1930	30.00	80
1936	21.00	56
1944	21.00	40
1945	12.00	20
1969	12.00	6
1970	9.60	5
1974	5.40	1.6
1980	4.80	0.9
1986	4.83	0.7
1991	3.32	0.3
1995	2.40	0.2
1999	0.30	0.02

Perhaps the greatest early limitation of the telephone was its cost. It was an extremely expensive technology. Phone service was primarily for businesses and the rich. However, with time technology improved, and economies of scale started working towards lower costs. Table 11.1 is a dramatic illustration of the decreases in telephone prices.

Long distance phone service in the U.S. shows an interesting pattern. Initially it was priced uni-

formly for everyone. The prices were high, and traffic was light. For example, in 1930, the Bell System carried around 160,000 toll calls per business day, compared to about 300 million carried just by AT&T alone these days (and more than twice as much by the entire U.S. long distance phone industry). The high early prices did depend heavily on distance, but were the same around the clock. The first time of day pricing was introduced in 1919, with three rates, the highest from 4:30 am to 8:30 pm, a lower rate from 8:30 pm to midnight, and the lowest from midnight to 4:30 am. The time bands argue strongly that the main motivation was not to shift demand and thus lower peak period traffic, but price discrimination, to allow private calls late at night at lower rates.

As prices continued to fall, distance dependence decreased, but time of day discounts remained in place. Some forms of discrimination increased markedly, though; the subsidization of local service by long distance fees grew explosively in the 1950s, 1960s, and especially 1970s, as is shown in [Mueller].

The breakup of the Bell System led to substantial changes. The cross-subsidy of local service by long distance revenues started declining. Together with dramatic improvements in technology and competition, this led to further declines in prices. In the new competitive environment it also led to simplification of pricing. Distance dependence was eliminated completely, and time of day variation was reduced to typically at most two tiers. Some of the most popular plans have uniform rates around the clock.

So far I have discussed the evolution of the phone system in the U.S. alone. In most countries, phone service started out as a government monopoly, and it is only recently that privatizations have begun to change that situation. According to [Mitchell], pricing of basic service has tended to be uniform, with business customers paying the same rates as residential ones. (Thus it is ironic that the competitive and unregulated capitalist marketplace in the U.S. around 1900 would lead in discriminating against business users!) There has been differentiation in pricing based on distance, but that has tended to decrease with time. Time of day variation in pricing (which is one way of discriminating in favor of residential users) came to Europe considerably later than to the U.S..

12. Cell phones

The prominence of the Internet has tended to overshadow the rise of another great high-tech success, namely that of the wireless industry. There are more users of cell phones in the world than there are Internet users. They also produce much higher revenues than Internet users. Even in the U.S., with its high Internet penetration, cellular industry revenues are much higher than those from the Internet.

Cell phones are a nice counterexample to the claim that is sometimes heard that metered services

cannot grow fast. Wireless services have been growing rapidly, but have been metered from the beginning, and rates remain high. The U.S. cell phone industry collects on average about \$0.25 per minute of wireless calls, as compared to \$0.10 per minute for the wired long distance phone service (and nothing for a large fraction of local calls).

Although some time-of-day pricing plans have been introduced in cellular telephony, the general trend has been towards simple plans. Pre-paid plans, which limit users' liability, are wildly popular, especially in Europe, where they account for most of the new subscribers. In the U.S., the most important innovation in pricing in recent years was the introduction of the AT&T Digital One-Rate™ plan in 1998. It was a form of block pricing, with initial rates of \$90 for 600 minutes per month, \$120 for 1,000 minutes, and \$150 for 1,400 minutes. One of its most distinguishing and most attractive features is that this price is all-inclusive, covering the long distance part of the call, as well as any roaming fees. The marginal per-minute revenue might appear to be only 7.5 cents per minute (as one moves from one block option to another), and roaming fees are often far higher than that. Thus this plan will definitely lose money for the carrier on some customers during some periods. However, it all averages out, since roaming is a fraction of total usage, and many customers use only a fraction of the full allotment of time they pay for. Thus the carrier is playing the role of an insurance company, absorbing some losses but gaining from customers' willingness to pay for the protection of not having to worry about an accumulation of many small charges.

The Digital One-Rate™ plan has been widely imitated, at least in the U.S. Carriers are competing to offer plans that allow various family or employee groups to call each other at no charge, and so on. This has had the effect of increasing average usage in the U.S. above the level seen in other countries that are usually regarded as being far more advanced in wireless communication. According to a press release from the U.S. cell industry association [CTIA], between the last quarter of 1998 and the last quarter of 1999, subscribers have increased their local calling from an average of 130 minutes per month in the last quarter of 1998 to 180 minutes per month in the last quarter of 1999. By contrast, in most of the rest of the world, average usage per subscriber is dropping, since pre-paid plans are spreading, and they have the effect of repressing usage. For example, in the U.K., average usage of a cell phone dropped from 4.8 minutes per subscriber per day in the second quarter of 1999 to 4.2 minutes a year later [Ofitel]. (To be more precise, the drop was from 3.5 minutes of outgoing calls and 1.3 minutes of incoming calls to 3.2 and 1.0 minutes, respectively.) At the same time, the average usage of a wired phone in the U.K. increased from 15.7 minutes of outgoing calls per day in 1999 to 17.3 minutes a year later. Although up to a third of the wired minutes are for Internet access, this does show that wired

phone usage is far higher than that of cell phones. Coupled with the effect of block pricing plans in the U.S., this strongly suggests that there is much more that can be done to stimulate voice usage, and that simple pricing plans are the best way to do it. In particular, this suggests that the main application of the increased bandwidth of 3rd generation wireless systems will be not for Internet access, but for more voice calls.

The Japanese i-mode system is frequently cited as proving that wireless Internet access can be a lucrative service. It has several interesting aspects. One is that apparently the spending on data by i-mode users is about equal to those users' increased spending on voice. Another is that most of the i-mode Internet services, such as downloading jingles, are paid for through flat monthly fees.

13. Residential access to the Internet

For the mass public, the Internet became visible and accessible only in the mid-1990s. Even before then, though, there was a growing industry in online services, with CompuServe, Prodigy, and AOL the most prominent companies in it. Growth was rapid, but not at the rate we have witnessed on the Internet in the last half a dozen years. Each network had its proprietary user interface, and a limited selection of content providers. Communication among users was not emphasized. As recently as 1996, pricing was based on a fixed monthly rate that covered a small number of hours of connect time, and fees for each additional hour. In many cases, some areas (for example of Prodigy) were designated as "premium" and charged at higher rates, and many of the databases on CompuServe had fees for items retrieved. In the early days, there were also charges for all email messages. (For the CompuServe price schedule of early 1996 with all its complexity, see p. 60 of [OECD].) All these fees were heartily disliked by the users. Email charges were dropped first. Then the industry converted to flat rate billing as a result of intense customer pressure (exemplified by the quote from [Swisher] in Section 8) and competition from new ISPs. (AT&T WorldNet was the first major ISP to offer flat rate pricing.) The dominant pricing model today in the U.S. is the flat rate, unlimited access plan for anywhere between \$14.95 and \$21.95 per month. Since most people have their local calls covered by a fixed monthly rate, their entire Internet access cost is flat rate. There are even some "free" access plans, supported by advertising.

Residential broadband access to the Internet in the U.S., through cable modems and DSL links, is universally priced on a flat rate basis. There is quality and price differentiation (as in limits on bandwidth of ADSL links, or between the Home and Work offerings that represent a primitive form of the Paris Metro Pricing of [Odlyzko2]), but those play a role only in the rare cases where customers select a service.

14. Non-Internet data networks

Most of the corporate spending for data transmission is for private lines, in which a dedicated connection of a fixed bandwidth is provided by telecommunications carriers between two points. (See [CoffmanO1] for estimates of the size of private line and public data networks.) Increasingly such links carry IP traffic, and are thus part of the Internet. Still, they are not visible to general users, since they are hidden behind firewalls. The price of such a link depends on the bandwidth, distance, and various regulatory and political constraints (especially for international lines), but not on the traffic carried. Most of the ISPs that do not own fiber networks carry their traffic over private lines leased from larger carriers. In general, as is shown in [CoffmanO1], for longer private lines, distance dependence in pricing has been decreasing over the last two decades. In other respects, though, pricing of the long distance connections is becoming more complicated. In the days of strict government regulation, before the 1983 breakup of AT&T, it used to be simple, based just on the air distance between the endpoints of the link. This pricing was far divorced from reasonable cost allocation. It led large customers to place facilities in various out-of-the-way places, in order to minimize transmission costs for connecting their branches. This in turn led to intensive mathematical investigations of the Steiner tree problem and related questions. Today, in a competitive market, two conflicting tendencies can be discerned. On one hand, we are moving away from simple tariffs, and towards negotiations and spot market. This is especially noticeable for high capacity links. This trend is in accord with the general thesis of this paper, since prices of private lines, especially high bandwidth ones, are high. On the other hand, there is also an opposite tendency, in which corporations are outsourcing their entire data networks, and avoiding the complexity of dealing with a variety of carriers, price plans, as well as network administration itself.

While private line networks have been growing, their growth rates in recent years have been moderate, about 20% per year in bandwidth and 10% in revenues in the U.S., [CoffmanO1]. (Remarkably enough, these growth rates are similar to the 28% growth rate of data traffic estimated for the late 1970s and early 1980s in [deSolaPITH]. It might seem paradoxical that the growth rate would slow down in recent years, but that is probably accounted for by growth shifting to other data networks.) Much more rapid growth has been taking place in the Frame Relay and ATM networks, which have seen growth rates approaching and sometimes exceeding 100% per year. Frame Relay is expected to overtake private lines in revenues, at least in the U.S., within a couple of years. Both Frame Relay and ATM networks multiplex traffic from many customers. Most of the traffic is on PVCs (permanent virtual circuits), which are analogs of private lines in providing point-to-point connections. Pricing is

predominantly on the basis of the size of the inlet to the public network, and independent of the traffic actually carried. (For the small fraction of traffic carried on SVCs, switched virtual circuits, there are usually usage sensitive charges, but for typical usage patterns, those are small compared to the fixed monthly charges.) Prices are also independent of distance, except that international virtual circuits are priced higher than domestic U.S. ones. (Generally there is no difference in pricing of international circuits among countries.) Thus the trend in this growth area is towards maximal simplicity.

15. Dedicated business connections to the Internet

The initial design decision about the Internet called for flat rate pricing. It was taken by ARPA. It is often claimed that all dedicated connections to the Internet in the U.S. are paid through flat rates. However, a substantial, although unknown, fraction are provided through so-called “burstable” rates that are illustrated in Table 15.1. Traffic is measured over 5-minute intervals, and the higher of the two traffic figures (for the two directions) is accumulated. At the end of the month, the top 5% of these 5-minute samples are discarded, and the highest remaining sample (the 95-th percentile) is used to determine the charges. Thus if in a 30-day month the lowest 8208 of the 8640 samples are below 384 Kbps, with at least one of those low 8208 samples above 256 Kbps, the charge for that month will be \$2500. These rates often appear attractive to users who have light traffic yet value the availability of broadband links to achieve low transaction latency, but have low average utilization. However, the pricing structure is such that in many cases users pay more with burstable than with flat rates. Unfortunately there are no studies on whether burstable rates are getting more popular or not, nor on what fraction of users who do pay burstable rates save with them over flat rate prices.

Table 15.1. UUNet burstable rates for Internet access:

flat rate T1 (1.54 Mbps): \$2500/month

burstable T1:

95-th percentile of bandwidth usage	monthly price
< 128 Kbps	\$1300
128-256	1900
256-384	2500
384-512	2750
> 512	3000

Straightforward charging by the byte by service providers in the U.S. appears to be most common in Web-hosting, where a company contracts out the maintenance of its external Web server. Typically the customer pays a fixed monthly fee that depends on the bandwidth of the connection to the Internet and storage space. This fee covers some allotment of bytes that are sent out, and additional ones beyond that limit are charged at rates that are typically in the range of \$0.03-0.05 per MB.

Charging according to usage is also increasing as an internal corporate cost allocation mechanism in the U.S. With growth in data traffic and resulting expenses, central corporate networking groups are moving towards charging different business units for their services according to the volume of traffic those units generate.

Perhaps the most important conclusion one can draw from the frequent use of usage charging in the U.S. is that measuring traffic volumes is not too hard. It is often claimed that any kind of byte-counting would be extremely expensive and not worth undertaking, and would lead to endless billing disputes. However, neither of these problems seems to be a major obstacle, at least not for the simple byte-counting that is involved in either burstable rates or internal corporate cost allocations. (Attempting to use pricing for real-time congestion control would be much more complicated, though.) Concerns have been raised about the fairness of charging users more just when service is worst (which is when packets dropped due to congestion cause retransmission), or for packets sent in error or by malicious outsiders. In practice, though, these problems appear not to be serious.

Charging by volume appears to be more common in other countries, especially those that have to pay for expensive links to the U.S., where much of the most popular Internet content resides. The book [McKnightB] has several papers and references to other papers that describe early experiments with usage-sensitive charging in places such as Chile and New Zealand. There are also some more recent cases, such as those of JANET in Britain and Telstra in Australia. See [Odlyzko5] for more details.

16. Software agents and other technological palliatives

The histories of communication services presented in the preceding sections largely conform to the thesis of the Introduction; simple pricing along with higher quality, lower prices, and increased total spending is the natural evolutionary path for communication services. Deviations from this trend tend to be associated with expensive and infrequent transactions. However, the Internet is in many ways unique, and so historical analogies might not apply. In this section we consider some factors that might make historical analogies invalid. (Additional ones are discussed in [Odlyzko5].)

As was mentioned several times earlier in the paper, yield management techniques are spreading to

a large extent because modern computing and communication technologies are making them possible. Manufacturers as well as service providers can get real-time feedback on sales and orders and tailor their production and pricing according to changing conditions. But we can turn this argument around. Similar computing and communication technologies are also available to consumers. Couldn't software agents be used to automate the price and quality negotiation tasks that have historically driven users to prefer simple pricing?

My prediction is that agents will be used, but only to a limited extent, and will not affect the drive for simplicity in pricing. In many high-tech areas, invocations of the words "software agent" or "genetic algorithms," or "fuzzy logic" are sometimes used to suggest that some technological magic will take care of all the hard problems without requiring serious thought. In practice, this has not happened. All these approaches have proved useful, but not one has been a panacea. There are two related factors that operate, and are likely to continue operating, in pricing as well as other fields. One is that these systems that are supposed to simplify life are not all that easy to master, and so are not as widespread as their proponents had hoped. The other is that even when these systems are used, they serve to encourage the growth of complexity that eats up any gain that had been achieved.

Both factors mentioned above are treated in more detail in [Odlyzko5]. Here I will just mention that the general conclusion about limited application of the supposedly revolutionary pricing approaches is supported by the fate of companies such as Priceline.com and eBay. Although they attracted great attention from the public and especially from investors, both have seen their stock valuations plummet by the end of 2000.

Another way to deal with the complexity that optimization requires and at the same time provide the end users with the simplicity they desire is to have intermediaries. These enterprises would negotiate with communication providers for basic services (using all the modern tools of auctions, futures, and derivatives) and offer them in packages to users at flat rates. This proposal is feasible, but does not refute any of the arguments for simplicity presented here. It basically consists of renaming various agents in the communications marketplace. The evidence of this paper is that simplicity is of paramount importance for individuals and small organizations. The company that is responsible for their communications needs is the effective carrier, whether that company has its own physical facilities, or leases them from other carriers.

17. Conclusions

The history of communication suggests strongly that as services become less expensive and are more widely used, the balance shifts away from the need to segment the market, and thereby to extract maximal revenues and to maximize utilization efficiency of the infrastructure. Instead, customer desire for simplicity becomes dominant.

Simplicity is likely to be much more important on the Internet than in other communication services. Customers do not care about the network, they care about their applications. Those applications are growing rapidly in number, variety, and importance, as the Internet becomes what Bill Gates has called the “digital nervous system” of more and more organizations. We will not want to worry how much to pay for a packet from site X to site Y that was generated by our request for something from site A, which then contacted site B, etc. We will be happy to pay extra for simple schemes that make our lives easy.

Flat rate is by far the simplest pricing plan, and, as predicted by Anania and Solomon [AnaniaS], it continues to dominate the data transmission market. The historical information of this paper only strengthens the arguments of Anania and Solomon and of the papers [Odlyzko3, Odlyzko4] in favor of continuing with flat rates for data transmission over core fiber optic networks. However, there are and will continue to be settings where such pricing may not be feasible. One such area is in the U.S. long distance voice telephony, where access charges are by far the largest cost component. Another such area is likely to be in wireless communication. Although the bandwidth there is growing, it is orders of magnitude lower than on fiber, and will remain orders of magnitude lower. Hence wireless bandwidth will continue to be relatively scarce (at least relative to that on fiber backbones) and technical and economic methods to ration it may continue to be required.

When usage sensitive pricing is required, customer preferences argue for only the simplest possible schemes, such as the Paris Metro Pricing proposal of [Odlyzko2]. However, it is best to avoid even schemes such as Paris Metro Pricing. There are alternatives that have a usage sensitive component, yet approximate flat rate pricing from the customer point of view. One such alternative is block pricing, which provides a user with a large allotment of time (in cases of phone calls) or bytes (for data). There are various choices that can be made with block pricing. The key point is to satisfy users’ desires for simplicity, predictability, and risk avoidance.

Further along the spectrum towards true flat rate is the “expected usage pricing” of [Odlyzko4]. It would be similar to the most popular Lexis/Nexis plans, with service providers offering users unlimited

access for some period such as a year. The pricing would be determined by the capacity of the link and that customer's record of prior usage. Service providers would assume some growth rate in traffic, and could put into the contracts provisions for reopening them in case of unusual behavior. This type of scheme would leave scope for negotiations and for actions that improve the efficiency of the network. ("We will lower your fee by 10% if you agree to send your backups over our network at 3 in the morning, and not at 10 in the evening.") Such an approach would have several advantages for service providers. It would stimulate usage. Further, it should also reduce turnover, as a competitor attempting to attract somebody else's customer would not have the detailed knowledge of that customer, and so would face the problem of adverse selection, in which only the least desirable customers might switch.

The general conclusion is that we should strive for simplicity, even at the cost of efficiency. That is the world of communications has been evolving for the past two centuries, and that is how it is likely to evolve in the future. There will be opportunities for optimization of the network, but they will have to be pursued in ways that do not burden the end users.

Acknowledgements: I owe much to the people who have provided me with information. I am especially grateful to Jörn Altmann and his colleagues in the INDEX project for the usage data quoted in Section 9, Sheldon Hochheiser for AT&T historical data and pointers to other useful sources, and Jerry Mansfield for the history of the U.S. Postal Service. I also thank David Applegate, Lars Aronson, Vijay Bhagavath, Greg Blonder, Frances Cairncross, Karyen Chu, Ken Church, David Cracknell, Whit Diffie, Amy Friedlander, David Gabel, Ehud Gelblum, Sam Glazer, Derek Gregory, Paul Henry, David Hochfelder, Paul Israel, Laura Jereski, Richard John, Andreas Jonason, Frank Kelly, Richard Kielbowicz, Don King, Alan Kotok, Jeff Lagarias, Henry Landau, Kenneth Lipartito, Paul Odlyzko, Hilarie Orman, Alison Oswald, Sam Paltridge, Philonoë, Jim Reeds, Jim Roberts, Larry Roberts, Angela Sasse, Ulf Stahrenberg, Hal Varian, Dave Walden, Roger Watt, and Barry Wellman, for comments and useful information.

Price indices, conversion rates, and data sources: All prices listed in this paper are in current dollars (i.e., not adjusted for inflation). Where non-U.S. prices are quoted, they have been converted to U.S. dollars at approximations to the exchange rate valid at the time. (For example, five dollars to one pound sterling was used for all of the 19th century.)

More detailed explanations of the tables and figures (as well as additional data) are contained in [Odlyzko5].

Figure 1.1 is based on quarterly financial reports from AOL, available at www-db.aol.com/corp/news/press/), as well as statistics compiled by AFA, the association of French ISPs, which are available at <http://www.afa-france.com/html/chiffres/index.htm>), and Telecom New Zealand quarterly reports, available at <http://202.27.156.72/inves>

Table 1.1 is based on data from [ITU]. For Japan, the data is for 1996. The revenue figures include wireless and data services.

Expenditure figures in Table 2.1 come from [USDOC1] and from recent annual reports of the USPS. Volume statistics for the years through 1840 are derived from Table 1.2 on p. 4 of [John], supplemented by some data in [Miles].

Table 2.2 is derived from statistics in [USDOC1, USDOC2]. The revenue figures include cell phones as well as data services.

Tables 9.2, 10.1, and 10.2 are based on [USDOC1].

Pre-1999 prices in Table 11.1 were obtained from AT&T Archives.

Table 15.1 comes from the end of 1999 online information for UUNet, available at <http://www.uu.net>).

References

- [AltmannRV] J. Altmann, B. Rupp, and P. Varaiya, Internet demand under different pricing schemes, in *Proc. ACM Conference on Electronic Commerce (EC-99)*, ACM, 1999, pp. 9-14.
- [AnaniaS] L. Anania and R. J. Solomon, Flat—the minimalist price, pp. 91-118 in *Internet Economics*, L. W. McKnight and J. P. Bailey, eds., MIT Press, 1997. Preliminary version in *J. Electronic Publishing*, special issue on Internet economics, <http://www.press.umich.edu/jep/>).
- [BouchS] A. Bouch and M. A. Sasse, It ain't what you charge, it's the way that you do it: A user perspective of network QoS and pricing, *Proc. IFIP/IEEE International Symp. Integrated Network Management (IM'99)*, IEEE, 1999, pp. 639-654.
- [BouchSDM] A. Bouch, M. A. Sasse, and H. DeMeer, Of packets and people: A user-centered approach to Quality of Service, in *Proc. IWQoS 2001*. Available at <http://www.cs.ucl.ac.uk/staff/A.Sasse/pub.html>).
- [Chu] K. Chu, User reactions to flat-rate options under time charges with differentiated quality of access: Preliminary results from INDEX, available at <http://www.marengoresearch.com/isqe/>).
- [CoffmanO1] K. G. Coffman and A. M. Odlyzko, The size and growth rate of the Internet. *First Monday*, Oct. 1998, <http://firstmonday.org/>). Also available at <http://www.research.att.com/~amo>).

- [CoffmanO2] K. G. Coffman and A. M. Odlyzko, Internet growth: Is there a “Moore’s Law” for data traffic?, *Handbook of Massive Data Sets*, J. Abello, P. M. Pardalos, and M. G. C. Resende, eds., Kluwer, 2001. To appear. Available at <http://www.research.att.com/~amo>).
- [CTIA] CTIA (Cellular Telecommunications Industry Association), April 11, 2000 press release, “CTIA Reports 1999 Survey Results,” available at http://www.wow-com.com/news/ctiapress/body.cfm?record_id=857).
- [deSolaPITH] I. de Sola Pool, H. Inose, N. Takasaki, and R. Hurwitz, *Communications Flows: A Census in the United States and Japan*, North-Holland, 1984.
- [EdellV] R. Edell and P. Varaiya, Providing Internet access: What we learn from INDEX, *IEEE Network*, vol. 13, no. 5 (1999), pp. 18-25. Also available at http://www.path.berkeley.edu/~varaiya/papers_ps.dir/networkpaper.pdf).
- [FishburnOS] P. C. Fishburn, A. M. Odlyzko, and R. C. Siders, Fixed fee versus unit pricing for information goods: competition, equilibria, and price wars, *First Monday*, vol. 2, no. 7 (July 1997), <http://www.firstmonday.org/>. Also to appear in *Internet Publishing and Beyond: The Economics of Digital Information and Intellectual Property*, B. Kahin and H. Varian, eds., MIT Press, 2000. Available at <http://www.research.att.com/amo>).
- [Gregory] D. Gregory, The friction of distance? Information circulation and the mails in early nineteenth-century England, *J. Historical Geography*, 13 (1987), 130-154.
- [Hansell] A. Hansell, Now AOL everywhere: Internet’s giant prepares to leap off the desktop, *New York Times*, July 4, 1999.
- [ITU] World Telecommunications Indicators Database (5th ed.), ITU, 1998. Available at <http://www.itu.int/publications/bookstore.html>).
- [JacksonCW] D. C. Jackson, W. H. Crumb, and G. W. Wilder, *Report on the Telephone Situation in the City of Chicago; In Respect to Service, Rates, Regulation of Rates, etc.; submitted to The Committee on Gas, Oil and Electric Light of the City Council of the City of Chicago*, Gunthorp-Warren Printing Co., 1907.
- [John] R. R. John, *Spreading the News: The American Postal System from Franklin to Morse*, Harvard Univ. Press, 1995.

- [Kielbowicz] R. B. Kielbowicz, *News in the Mail: The press, Post Office, and Public Information, 1700-1860s*, Greenwood Press, 1989.
- [Kraepelien] H. Y. Kraepelien, Local telephone pricing: The configuration theory, U.S. Dept. Commerce, National Tech. Info. Service, report PB-280 194, March 1978.
- [MacKieMV] J. K. MacKie-Mason and H. R. Varian, Pricing congestible network resources, *IEEE J. Selected Areas Comm.*, 13 (1995), 1141-1149. Available at <http://www.sims.berkeley.edu/~hal/people/hal/papers.html>.
- [McKnightB] L. W. McKnight and J. P. Bailey, eds., *Internet Economics*, MIT Press, 1997. Preliminary version of many papers available in J. Electronic Publishing, special issue on Internet economics, <http://www.press.umich.edu/jep/>.
- [Miles] P. Miles, History of the Post Office, *The Bankers' Magazine, and Statistical Register*, vol. 7 (new series, sometimes listed as vol. 12 in the original series), no. 6 (Nov. 1857), pp. 337-365 and no. 7 (Dec. 1857), pp. 433-448.
- [Mitchell] B. M. Mitchell, Pricing policies in selected European telephone systems, pp. 437-475 in *Proc. Sixth Annual Telecommunications Policy Research Conference*, D. C. Heath, 1979.
- [Mueller] M. L. Mueller, Jr., *Universal Service: Competition, Interconnection, and Monopoly in the Making of the American Telephone System*, MIT Press and AEI Press, 1997.
- [Odlyzko1] A. M. Odlyzko, The bumpy road of electronic commerce, in *WebNet 96 - World Conf. Web Soc. Proc.*, H. Maurer, ed., AACE, 1996, pp. 378-389. Available at <http://www.research.att.com/~amo>.
- [Odlyzko2] A. M. Odlyzko, Paris Metro Pricing for the Internet, in *Proc. ACM Conference on Electronic Commerce (EC-99)*, ACM, 1999, pp. 140-147. Based on a 1997 unpublished manuscript, A modest proposal for preventing Internet congestion. Both available at <http://www.research.att.com/~amo>.
- [Odlyzko3] A. M. Odlyzko, The economics of the Internet: Utility, utilization, pricing, and Quality of Service. Available at <http://www.research.att.com/~amo>.

- [Odlyzko4] A. M. Odlyzko, The Internet and other networks: Utilization rates and their implications. Presented at the 1998 Telecommunications Policy Research Conference. *Information Economics & Policy*, vol. 12 (2000), pp. 341-365. Available at <http://www.research.att.com/~amo>.
- [Odlyzko5] A. M. Odlyzko, The history of communications and its implications for the Internet. Available at <http://www.research.att.com/~amo>.
- [OECD] OECD, Information infrastructure convergence and pricing: The Internet, report OCDE/GD(96)73, available at <http://www.oecd.org/dsti/sti/it/cm/prod/ONLINE.HTM>.
- [OfTel] U. K. Office of Telecommunications, Nov. 2000 Market Information Update, available at <http://www.oftel.gov.uk/market/miu1100.pdf>.
- [Panzar] J. C. Panzar, The Pareto domination of usage-insensitive pricing, pp. 425-436 in *Proc. Sixth Annual Telecommunications Policy Research Conference*, D. C. Heath, 1979.
- [Stehman] J. W. Stehman, *The Financial History of the American Telephone and Telegraph Company*, Houghton Mifflin, 1925.
- [Swisher] K. Swisher, *Aol.Com: How Steve Case Beat Bill Gates, Nailed the Netheads, and Made Millions in the War for the Web*, Times Books, 1998.
- [Szabo] N. Szabo, Micropayments and mental transaction costs, presented at the 2nd Berlin Internet Economics Workshop. Available at <http://www.best.com/~szabo/micropay.ps>.
- [USDOC1] U.S. Department of Commerce, Bureau of the Census, *Historical Statistics of the United States: Colonial Times to 1970*, 1975.
- [USDOC2] U.S. Department of Commerce, Bureau of the Census, *Statistical Abstract of the United States 1999*, 1999. Available online at <http://www.census.gov/prod/www/statistical-abstract-us.html>.
- [Varian1] H. R. Varian, Price discrimination, pp. 597-654 in *Handbook of Industrial Organization*, vol. I, R. Schmalensee and R. D. Willing, eds., Elsevier, 1989.
- [Varian2] H. R. Varian, Differential pricing and efficiency, *First Monday*, vol. 1, no. 2 (Aug. 1996), <http://firstmonday.org/>.
- [Zajac] E. E. Zajac, *Political Economy of Fairness*, MIT Press, 1995.