

Internet pricing in light of the history of communications

Andrew Odlyzko

AT&T Labs - Research, Florham Park, NJ 07932, USA

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.

Flat rates are the simplest form of pricing. Although they have generally been regarded as irrational, and economically and socially undesirable, they have serious advantages. Consumers like them, and are willing to pay extra for them. Further, flat rates are extremely effective in stimulating usage, which is of advantage in a rapidly growing service like the Internet.

Keywords: flat rates, Internet pricing, usage sensitive charging, communications pricing

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, and thus going against the trend of other communication services. (Extensive references are available.⁷) 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 often 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.¹²

The technology and economics of early telephony made the reasoning behind that 1905 conclusion even more compelling than the arguments supporting Varaiya's call for abandoning flat rate for Internet access. (The primitive switching technology, involving human operators, created high marginal costs and diseconomies of scale, unlike the Internet, where almost all the costs are fixed.¹⁰) 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 (based on data from⁶) 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. 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 Japan and the U.K. are re-introducing limited forms of flat rate pricing for local calls. The pressure for such unmetered plans in other countries is also growing. Is this some temporary aberration? The thesis of this paper is that it is not. The historical trend has been almost uniformly towards simplifying pricing structures. Furthermore, simple prices, especially flat rates, have many virtues that are not widely appreciated.

2. THE VIRTUES OF FLAT RATES AND SIMPLICITY

Although the literature is full of denigrations of flat rates, they can be justified in conventional economic terms. They represent a form of bundling.⁴ Bundling frequently offers sellers higher revenues than they could obtain by selling a-la-carte, by taking advantage of uneven preferences for different items among consumers. However, the bundling argument by itself is not decisive. On the other hand, there are additional arguments that, together with the bundling one, do make a very strong case for flat rates. They are based on human preferences, and so come from psychology (although they are beginning to make inroads into economics through the subfield of behavioral economics).

Usage-sensitive pricing is effective. The problem is that we may not like its effects. In particular, such pricing lowers demand, often by substantial factors. Figure 1 (drawn from AOL online press releases) 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, that same figure shows statistics for French Internet users. They pay by the minute for their local connections, even when their ISP charges are in the form of flat monthly rates. (These are Internet users, not Minitel ones, who even in peak years for that service spent under 3 minutes online per day.) The time spent online by these French users is less than a third the time of AOL's U.S. subscribers, and has been growing considerably more slowly. That the difference in usage is caused by pricing, and not by culture, is shown by the graph of Telecom New Zealand usage, which has been moving from the current French level to that of AOL after the introduction of flat rates in May 1999.

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 rate local calling. 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.⁵

It is easy to understand why consumers or even governments might favor increased usage. For service providers, though, profitability is usually perceived as the main requirement (especially now that the dot-com bubble has burst). Increasing usage is often perceived as running counter to that goal. Carriers have often attempted to use techniques such as market segmentation and price discrimination to grow. For example, a century ago, that was the basic policy of the Bell System:

AT&T also favored [local measured service pricing] because they believed that it would increase network membership. The president of AT&T, Frederick Fish, believed that customers valued access and that charging a low fee for network membership would maximize the number of subscribers. According to Fish, the number of users was an important determinant of the value of telephony to individual subscribers. His desire to maximize network connections led the firm to adopt a pricing structure in which prices to residential customers were actually set below the marginal cost of service in order to encourage subscriptions. These losses were made up through increased charges to business customers.⁸

However, it is questionable whether that was the best policy. The rapid growth of the telephone industry in the U.S. in the early years of the 20th century apparently owes much to the independent phone companies, which forced the preservation and even extension of flat rate local service.

Fish's belief that value was created largely through the number of customers connected to the network, rather than the number of calls, made it difficult for him to recognize the utility of flat-rate service as a pricing policy to preempt entry. But there were other possible ways to conceptualize the process of network formation. AT&T officials could have conceived the value of network participation as based largely on the number of connections, rather than the number of members connected to the network. Maximum use of the network for a given number of subscribers would have been achieved under flat-rate service.⁸

On the Internet, increasing usage is likely to be imperative for carriers. Technology will be increasing available bandwidth. Service providers whose customers do not move on to more bandwidth-intensive applications will wither. Profitability is necessary in the long run, but bare survival will require persuading customers to move up the technology curve. Flat or at least simple rates are among the best methods for doing that.

The logic of quality and price differentiation is impeccable. In principle such practices can improve the efficiency of the economy. Unfortunately they run up against 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.'¹³

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:

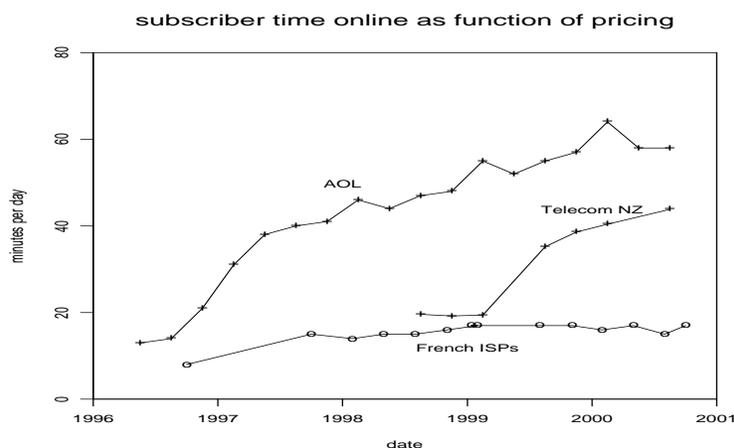


Figure 1. Time spent online as function of charging method. AOL and New Zealand Telecom XTRA ISP service introduced flat rate plans in October 1996 and May 1999, respectively, leading to surges in usage. French ISP subscribers pay for each minute online.

... 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.¹¹

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-Rate™ 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. is due largely to 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, and the full paper presents 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. This is illustrated in Tables 2 and 3, which show the evolution of prices of postal and telephone services in the U.S. (Data drawn primarily from.^{15,16})

Table 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

What Tables 2 and 3 show is primarily the tendency for prices to become distance-insensitive as a service evolves. This is a particularly interesting example of the trend towards simplification in prices. A Martian who saw just the evolution of price schedules might be tempted to argue that the costs of providing uniformly high quality to all transmissions cannot be very high. After all, if the extra cost of network facilities to send messages to distant locations is not worth charging for, then the cost of overprovisioning is likely to be small as well. An even more convincing argument can be developed if one studies the reasons for distance dependence in pricing more carefully, as is done in the full paper. Such distance dependence is often used primarily as a means of charging according to perceived value. Thus the decrease in distance dependence indicates that the extra hassle that varying prices impose on users is not worth the additional profit they bring.

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. Here we just sketch the two main arguments for flat rates. One is based on conventional economic arguments, viewing flat rate pricing as a form of bundling. This enables the service providers to take advantage of users’ uneven preferences for components of the bundle and increase revenues. The second argument for simple pricing is based on customer willingness to pay more for simplicity. This was noted above in connection with Internet access and local phone calls, and many more examples are cited in the full paper. In particular, there is evidence from the recent INDEX experiment that confirms this in a quantitative form.

Table 3. Domestic U.S. telephone calling rates. Price of station-to-station, daytime, 3-minute phone call from New York City.

year	Philadelphia	Chicago	San Francisco
1917	\$0.75	\$5.00	\$18.50
1926	0.60	3.40	11.30
1936	0.50	2.50	7.50
1946	0.45	1.55	2.50
1959	0.50	1.45	2.25
1970	0.50	1.05	1.35

Why are users willing to pay more for flat rate plans, or more generally for simple ones? One reason is to avoid what Nick Szabo¹⁴ has called “mental transaction costs;” “Yes, I can save by optimizing my usage, but do I want to, if the savings amount to pennies, and require my attention dozens of times a day?” The choices available to us are growing explosively, but our time isn’t. Cutting down on the mass of things we have to worry about is valuable.

There are also other reasons for the willingness to pay extra for flat rates.⁴ A very important factor is the insurance effect. (“How do I know how big a bill my teenagers will run up?”) The popularity of prepaid calling cards for wired and wireless telephony is a highly visible example of the attractiveness of limiting risks, even for affluent customers.

3. CONCLUSIONS

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.

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 most 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. The historical trends documented in this paper, together with projections of technological advances, argue in favor of continuing with this scheme for transmission over core fiber optic networks. (This was already predicted before the age of the Internet by Anania and Solomon,¹ and this paper adds yet more arguments to theirs.) However, there are and will continue to be settings where flat rate pricing may not be feasible. One such area is currently in 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 Paris Metro Pricing.⁹ However, it is best to avoid even that scheme. 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).

Further along the spectrum towards true flat rate is the “expected usage pricing” proposal. 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.

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: More detailed arguments, data, and references can be found in the long manuscript,¹⁰ which will eventually be published in a revised form under a different title as a book.

REFERENCES

1. 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/>).
2. 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.
3. 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/>).
4. 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>).
5. A. Hansell, Now AOL everywhere: Internet’s giant prepares to leap off the desktop, *New York Times*, July 4, 1999.
6. World Telecommunications Indicators Database (5th ed.), ITU, 1998. Available at (<http://www.itu.int/publications/bookstore.html>).
7. 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/>).
8. J. Nix and D. Gabel, AT&T’s strategic response to competition: Why not preempt entry?, *J. Economic History* 53, no. 2 (1993), 377-387.
9. 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>).
10. A. M. Odlyzko, The history of communications and its implications for the Internet. Available at (<http://www.research.att.com/~amo>).
11. 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.
12. J. W. Stehman, *The Financial History of the American Telephone and Telegraph Company*, Houghton Mifflin, 1925.
13. 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.
14. N. Szabo, Micropayments and mental transaction costs, presented at the 2nd Berlin Internet Economics Workshop. Available at (<http://www.best.com/~szabo/micropay.ps>).
15. U.S. Department of Commerce, Bureau of the Census, *Historical Statistics of the United States: Colonial Times to 1970*, 1975.
16. 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>).