UNIVERSAL MEASURED SERVICE

OF LOCAL TELECOMMUNICATIONS SERVICES

A Report Prepared by

J. Brander, J. Hartwick, and J. MacKinnon

for the

Department of Communications

Ottawa, Canada

March 31, 1981

Contract 03SU.36100-0-9507
5417-2-5(DOC)
31.26(DEA)
DGCE Doc. No. 167

UNIVERSAL MEASURED SERVICE

OF LOCAL TELECOMMUNICATIONS SERVICES

A Report Prepared by

J. Brander, J. Hartwick, and J. MacKinnon

for the

Department of Communications

Ottawa, Canada

March 31, 1981

Contract 03SU.36100-0-9507
5417-2-5(DOC)
31.26(DEA)
DGCE Doc. No. 167

TABLE OF CONTENTS

PART I

Optimal Pricing Considerations For Local Telephone Service

1. Introduction

Nearly all goods and services are priced so that payment increases as individual consumption increases. Cups of coffee, pairs of shoes, kilowatt-hours of electricity, and long distance telephone calls all have a per unit price so that a consumer pays more for every unit consumed. This kind of pricing is referred to as usage sensitive pricing (USP). Local telephone service is priced differently. In Canada, users pay a monthly fee for access to the system, but are allowed to make as many local phone calls as they wish at no extra charge.

We might expect a pricing system to be insensitive to usage if the cost of monitering use is very high compared to the cost of producing the service. In most of Europe, however, local telephone service is subject to usage sensitive pricing. It is possible that conditions have been sufficiently different in Europe and North America that, in the past, usage sensitive pricing of local telephone service might have been economically efficient for Europe but not for North America (see Mathewson and Quirin, 1972). However, conditions have changed sufficiently in the past few years that it seems unlikely that existing pricing policies in North America continue to be efficient.

In the United States there is a slow but steady movement towards usage sensitive pricing of local telephone service. There are three particularly important reasons for this change. First, local service has traditionally been subsidized by long distance service; that is, revenues generated by monthly access charges have not been covering local costs. It is, of course, difficult to determine the "costs" of local service as distinct from the costs of long distance service when both are jointly produced. Also a precise definition of subsidization is required. Subsidization and discrimination are taken up in Section 6. Despite the difficulty in being precise about costs and the degree of subsidization it is, nevertheless, commonplace that local service has been subsidized, in some relevant sense, by long distance service. (See, for example, Baude (1979)). In any case, long distance telephone service is now in competition with a growing private telecommunications industry in the U.S. If telephone companies continue to subsidize local service with high long distance rates, they may lose many of their long distance customers and much of their revenue. If local service is to generate more revenue, usage sensitive pricing promises to be a less painful (and more efficient) method than increasing access charges.

A second reason for the change in the U.S. is that local telephone lines are being used increasingly to link computers. Such lines could be active most of the time, drastically increasing the load on the local telephone networks, involving either binding capacity constraints and reduced quality of service, or expensive upgrading of local systems. In either case, usage sensitive pricing promises to be a helpful tool. Finally, the cost of monitering usage has fallen dramatically in recent years, making usage sensitive pricing much more attractive now than in the past.

The purpose of this section is to survey the important theoretical considerations that should be kept in mind in designing a pricing policy for local service. The point of view is that of the regulator rather than the shareholders of telephone companies. The regulator may face a constraint in that a minimum rate of return to shareholders is required, but it is assumed that the regulator's objective is to act in the public interest.

Perhaps the most basic insight of optimal pricing is that, in the absence of special circumstances, prices should be set equal to marginal cost. At least, marginal cost pricing is the standard against which alternative pricing methods should be compared. In any case, we start with a brief discussion of marginal cost pricing. There are, however, many special circumstances that imply that departures from marginal cost pricing are desirable. First, demand for telephone service has a strong cyclical character which makes it difficult to identify a single marginal cost. The effective marginal cost is higher in periods of peak demand than in offpeak periods. The second subject to be discussed is peak-load pricing.

The second major complication is that marginal cost pricing will not always cover costs. Telephone companies are supposed to earn a "normal profit" so departures from simple marginal cost pricing may be required from this source. Efficient pricing subject to a profit constraint leads to consideration of nonlinear pricing (especially two-part tariffs), self-selected (or optional) pricing schedules, price discrimination, and cross-subsidization. Each of these topics will be carefully discussed.

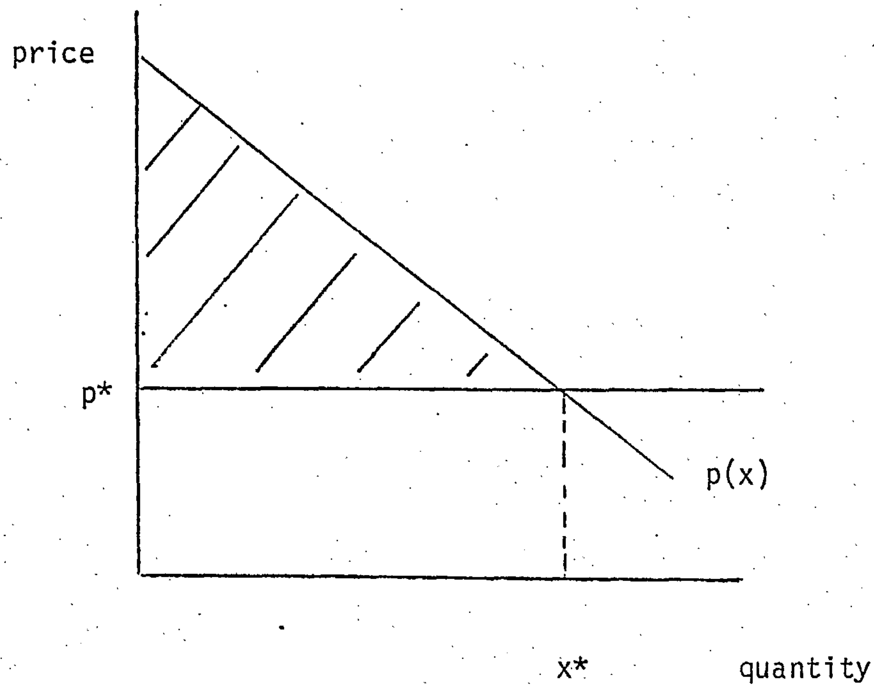
A third special feature of telephone service is that the decision to join the telephone network has an externality associated with it. Specifically, if one consumer joins, all other consumers benefit because they now have the option of phoning the new consumers. Because the new consumer would not capture all these benefits himself, he may decide not to join the system even when it would be socially desirable that he do so. This calls for adjustment of pricing schedules. Consumption externalities will be discussed. It should probably be mentioned at the outset that, although consumption externalities are important in early stages of development of a telephone network, once telephone service is nearly universal, as it is in Canada, further efficiency gains from adjustments to account for consumption externalities are likely to be trivial in comparison to the gains to be made from, for example, peak-load pricing.

There are also some additional concerns which deserve at least some attention, including ex post pricing and the Averch-Johnson effect. These two issues are taken up briefly.

2. The Measurement of Economic Welfare

Before "optimality" or "efficiency" in telephone pricing can be discussed it is necessary to have some clear specification of the objective in mind. The basic assumption is that a consumer's benefit from consuming an item can be measured by the maximum amount he would be willing to pay for it. The difference between this maximum

Figure 1: Consumer's Surplus



and the amount actually paid is a surplus to the consumer and is referred to as consumer's surplus. The consumer has a demand curve for telephone service. Let the quantity of telephone service be denoted by x. Then the maximum price the consumer is willing to pay for an additional unit of local telephone service, if he is already consuming x, is shown by the demand curve at quantity x: p(x). This is shown on Figure 1. If the price for local service were p*, then total consumer's surplus from consuming telephone service would be represented by the shaded region because, for each unit consumed, the incremental consumer's surplus is p(x) - p*.

There are difficulties with using consumer's surplus as a measure of individual utility. Some difficulties can be avoided if the marginal utility of income is not affected by the total amount of telephone service consumed. It is also convenient if the demands for other goods are not much affected by the price of local telephone service, although surplus measures do make sense even with interdependent demands. These two assumptions seem like very reasonable simplifying approximations for telephone service. They would be less reasonable for a good which consumed a large fraction of a consumer's budget, like housing services.

Given that consumer surplus is a reasonable measure of utility for a single consumer there is still the problem of aggregation.

Is one dollar of surplus to consumer A as valuable to society as a dollar of surplus to consumer B. If B is very much poorer than A the answer is likely to be no. Nevertheless, in deriving optimal pricing rules it is usually assumed that dollars of surplus count the same wherever they accrue. This approach is usually defended in one of three ways. First, it can be argued that there are government agencies explicitly concerned with redistribution of income and that they are always involved in adjustments through the tax and transfer system so as to keep the distribution equitable. Therefore, industry regulators should be concerned solely with efficient pricing for their industry and not try to make distributional judgements. (More formally, if some agent is optimizing the distribution of income, it must be the case that the social marginal value of an extra dollar is the same everywhere in society.)

A second justification is that, if some change in pricing generates a net surplus, it is always possible to carry out a set of compensating payments that would leave everyone better off. This justification is strongest, of course, when compensations are actually made. Finally, it can be argued that if surplus is maximized independent of where the benefits accrue, then society is getting better off on average. Thus, although one consumer might be made worse off by a particular change, he would expect to benefit from other policies.

For the case of telephone service, the distributional effects of changing pricing policies are likely to be small, so it seems reasonable to ignore them by focussing on surplus measures. Some economists oppose the use of surplus measures and the literature of economics is full of related analysis and discussion. Our objective here is not to survey this literature, but perhaps a few citations are appropriate. Standard references are Hicks (1940-41; 1943), Harberger (1971) and Willig (1976). A textbook treatment can be found in Varian (1978).

Our assumptions imply that individual utility functions can be approximated by the form

$$U_i = u_i(x_i) + y_i \tag{1}$$

where i is an index over consumers running from 1 to n, and y_i refers to the dollar expenditure on other goods by consumer i. Total consumer surplus from telephone service is then

$$S = \sum_{i=1}^{n} u_i(x_i) - p\Sigma x_i - T \tag{2}$$

where p is the price of local service and T represents the usage insensitive fixed monthly charges. Profits earned by the telephone company accrue to shareholders and thus also contribute to the utility of members of society. Using π to denote profits yields $\pi = p\Sigma x_i + T - c(\Sigma x_i)$ where c is the cost of producing telephone service. Adding S and π yields the net surplus, which is taken to be the regulator's objective function

$$W = \sum_{i=1}^{n} u_i(x_i) - c(\Sigma x_i) \tag{3}$$

This net benefit function is the total consumer benefit from consuming telephone service minus the cost of producing it. This can be rewritten

$$W = B(x) - c(x) \tag{4}$$

where $B = \Sigma u_i(x_i)$: the gross benefit to consumers from consumption of x. By the definition of B, we have $dB/dx = du_i/dx_i$, and, because of utility maximization by consumers, $du_i/dx_i = p$. Formally, the consumer's problem is to

maximize $\quad U_i = u_i(x_i) + y_i$

subject to $\quad px_i + t + y_i \leq I_i$

where I_i represents the consumer's total income, and t represents the monthly fixed fee paid by this consumer for local service. Writing down the Lagrangian function and associated first order conditions for a maximum:

$$L = u_i(x_i) + y_i + \lambda[I_i - px_i - t - y_i]$$

$$\partial L/\partial x_i = 0 \rightarrow du_i/dx_i - \lambda p = 0$$

$$\partial L/\partial y_i = 0 \rightarrow \lambda = 1$$

Therefore, the consumer sets $du_i/dx_i = p$ so $dB/dx = p$. Because $dB/dx = p$ it is also true that B is the integral of market demand

$$B(x^*) = \int_0^{x^*} p(x)dx$$

where p(x) is the (inverse) market demand for x. This leads to a third way of writing the regulator's problem:

$$\text{Maximize } W = \int_0^{x^*} p(x)dx - c(x^*) \qquad (5)$$

3. Marginal Cost Pricing

The desirability of marginal cost pricing is so well established in the literature on regulation that it doesn't seem worthwhile to dwell on it here. However the basic result is easily derived from what we have so far. The regulator's problem is to maximize (4) so p = mc is the solution: price should be set equal to marginal cost. The basic insight is that price reflects the marginal value to a consumer, and therefore to society, of an extra unit of telephone service. If this extra value exceeded the marginal cost to society of extra telephone service, then clearly additional consumption should take place. Similarly if the price (and marginal benefit) were less than marginal cost, consumption should be cut back. Only if price is just equal to marginal cost is the optimal amount of telephone service consumed.

Some observers treat local telephone service as two commodities: access to the system and usage. This is appropriate if costs can be identified with access and usage separately, which seems quite reasonable for telephone service. In this framework marginal cost pricing implies that both monthly access charges and usage charges would be desirable, with each charge set at the appropriate marginal cost.

This argument for marginal cost pricing makes several implicit assumptions. First of all it assumes that the marginal social cost of production is the same as the private marginal cost. If there are other distortions in the economy, such as imperfect competition and taxation so that other prices differ from marginal cost, this is not true. This problem is associated with the theory of the second best, which was first articulated in general form by Lipsey and Lancaster (1956). Whether or not this second-best problem is grounds for abandoning marginal cost pricing as a guideline has not been properly settled, and since it is a question that is essentially empirical in nature, it must be answered on a case by case basis. Later articles on the theory of the second best include McManus (1959), Green (1961), Davis and Whinston (1965), Farrell (1968), and Boadway and Harris (1977).

Strictly speaking, imposing a minimum profit constraint on the public utility turns the problem into a second-best problem. However it is conceptually useful to distinguish between the profit constraint and second-best considerations that arise because of imperfect competition and taxation elsewhere in the economy. As a practical matter, these latter second best effects are generally ignored. One hopes the effects are small; in any case, trying to calculate them is not feasible.

The basic presumption is that distortions cause prices to exceed marginal cost elsewhere in the economy, with the result that too little of these goods is produced. To the extent that telephone service is a substitute for underproduced goods, its price should exceed marginal cost. If telephone service were sufficiently complementary to underproduced goods, however, it's optimal price could be below marginal cost. Substitutes tend to dominate the marketplace; therefore in the absence of information to the contrary, one would expect that second-best adjustments would call for slightly higher prices in regulated industries. We have already assumed, for simplicity, that telephone service demand is independent of demand for other goods, which implies that telephone service is neither a substitute nor a complement for other goods and consequently, that these second-best effects are negligible. This seems like quite a reasonable assumption for local telephone service.

One objection to marginal cost pricing is based on equity considerations. If marginal cost is below average cost, as is likely for most public utilities, including telephone service, then consumers paying marginal cost prices would not pay the full cost of their consumption and would require a subsidy from society at large. This argument is developed in Coase (1970). We have assumed that dollars of surplus are worth the same everywhere which effectively assumes away this equity problem. One way of dealing with this problem, of course, is to require that the telephone company cover its full costs.

Another problem is that marginal cost might not be easily identifiable as a single number. In particular, marginal cost might vary from time to time, which leads to our next topic: peak-load pricing. As for general references on marginal cost pricing, Hotelling (1939) is often cited as a classic statement of the case for marginal cost pricing. Ruggles (1949-50) reviews much of the early literature. A rather nontechnical account of marginal cost pricing and related issues is in Kahn (1970). See also Nelson (1964) and Turvey (1970).

4. Peak-load Pricing

Demand for local telephone service varies in a sharp but regular pattern over the day and week. During periods when usage is below capacity the marginal cost of extra phone calls is very low, reflecting extra operating costs only. Conversely, in periods when the system is being used up to capacity, the marginal cost includes a capacity cost since extra consumption would require provision of extra capacity. Thus marginal costs differ from period to period and simple extension of the marginal cost pricing principle suggests that different prices should be charged.

The basic insight can be derived rather directly in the following simplified model. Consumer i gets utility from consumption in the peak period, denoted x_i^1 and from consumption in off-peak periods, denoted x_i^2. Total consumption in the peak period is denoted x^1 and total consumption in off-peak periods is x^2. Let $x = x^1 + x^2$.

Capacity is denoted by k and there is a capacity constraint: $x^1 = k$. Following Williamson (1966) we assume linear cost: $c = F + bx + \beta k$. F is some fixed cost, b is marginal operating cost and β is the marginal cost of adding to capacity. This model addresses the problem of deterministic cyclical demand. There is also a stochastic character to demand which is taken up later.

The regulator's problem is to maximize total utility minus total cost as in (3), so substituting $x^1 = k$ in the cost function yields the following problem

$$\text{Max } W = \Sigma u_i(x_i^1, x_i^2) - F - bx - \beta x^1 \tag{6}$$

$$\partial W/\partial x_i^1 = 0 \rightarrow \partial u_i/\partial x_i^1 - b - \beta = 0$$

$$\partial W/\partial x_i^2 = 0 \rightarrow \partial u_i/\partial x_i^2 - b = 0$$

As before, the consumer sets the marginal utility of extra consumption equal to the price, so if the peak period price is p^1 and the off-peak price is p^2, then the above conditions yield

$$p^1 = b + \beta$$

$$p^2 = b$$

The peak period price should equal marginal operating cost plus marginal capacity cost, and the off-peak price should equal marginal operating cost only. Thus our simple model captures the basic insight. The intuitive rationale might go as follows. Capacity is determined, in some sense, by peak demand. In the off-peak periods there is excess capacity. We can see that charging any price above marginal operating cost for the off-peak period must be inefficient. Imagine $p^2 = b + a$ for some positive a. Then there would be consumers willing to pay more than b but not as much as b + a for extra telephone service. Since the extra cost to society of their call is only b and the marginal value to the consumer (and society) exceeds b, the call is socially desirable, yet the price equal to b + a prevents the call. Thus the socially desirable price is $p^2 = b$. Note that the solution to (6) implicitly characterizes the optimal capacity that should be built. Additional capacity should be built up to the level at which the marginal willingness of consumers to pay for the service in the peak period is just equal to the marginal cost of builting and operating extra capacity.

The literature on peak-load pricing is voluminous, partly because it is a very important practical problem, but largely because there are many considerations in addition to the basic insight, all of which have been ignored by our simple formulation. Indeed, one could identify at least seven separate issues that have received special attention. We shall look at each issue in turn, but it seems reasonable to begin with a short, although perhaps cryptic, list:

1. Ambiguous definition of output and capacity
2. Nonlinear costs
3. Variable coefficients of production
4. Shifting peak
5. Choosing time and duration of periods
6. Uncertainty
7. Binding profit constraint

The first item is not difficult to deal with although it does need to be understood and did generate a certain controversy in the literature (Steiner 1957, 1958; Hirshliefer,1958; Williamson, 1966, 1974; Wilson, 1972. The papers by Steiner and Williamson are also classic early formulations of the peak-load pricing problem as is Boiteux, 1960.) If all periods in the cycle are of equal length there are natural interpretations of output, demand, and capacity: output is output over the entire period and capacity is the maximum amount of output that can be produced in any one period. Similarly demand is well-defined as the demand for total output in period j at whatever prices are prevailing. If, however, periods are of different length the symbols must be interpreted with care. For example, if one period were 4 hours and another were 8 hours, we could no longer refer to a single "capacity". Presumably, capacity in the longer period would be greater. Similarly, demands over the two periods are not directly comparable. The way we have derived the optimal pricing results, market demands do not appear. Williamson (1966) used an approach based on maximization of an expression corresponding to (5) rather than to (3) so that demand functions do enter directly. To deal with the problem of non-equal periods, he defined demand in period j as the demand that would be forthcoming over the entire cycle at price p_j, if demand were always as it was in period j. He then weighted each demand function by the proportion of the cycle of which it applied. This yields an objective function as follows.

$$W = \sum_{j=1}^{m} \int p_j(x_j) \cdot w_j \; dx_j - c(\Sigma x_j)$$

where w_j is the proportion of period j in the total. The derivation used here avoids most of the confusion by working with utility rather than demand functions.

At a more conceptual level, there is a difficulty in defining the output of a telephone company. Number of calls, number of minutes, distance, nature of call, and amount and type of access are all relevant parts of "output". Also, quality is service is important. These qualifications apply with equal force to the definition of capacity.

The second issue is nonlinear costs. Generalizing the model to nonlinear costs changes nothing. Pressman (1970) has a very useful formulation of the peak-load pricing problem including nonlinear costs. We can rewrite (6) as

$$\text{Max } W = \Sigma u_i(x_i^1, x_i^2) - c(x^1, x^2, k)$$

$$\text{Subject to } x^1 \leq k,\ x^2 \leq k.$$

Substituting $\Sigma u_i(x_i^1, x_i^2) = B(x^1, x^2)$ and assuming $x^1 = k$ and $x^2 < k$ yields the Lagrangian

$$L = B(x^1, x^2) - c(x^1, x^2, k) + \lambda[k - x^1]$$

$$\partial L/\partial x^1 = p_1 - \partial c/\partial x^1 - \lambda = 0$$

$$\partial L/\partial x^2 = p^2 - \partial c/\partial x^2 = 0$$

$$\partial L/\partial k = - \partial c/\partial k + \lambda = 0$$

Therefore $p_1 = \partial c/\partial x^1 + \partial c/\partial k$

$$p_2 = \partial c/\partial x^2$$

This is exactly the result we had before except that marginal costs can vary. $\partial c/\partial x^1$ and $\partial c/\partial x^2$ are marginal operating costs and $\partial c/\partial k$ is marginal capacity cost. With nonlinear costs, however, the determiniation of optimal capacity becomes more difficult because marginal costs become endogeneous. Estimating optimal capacity is obviously a very important problem. Also, additional capacity improves the quality of service. This is not something that has been analyzed in the theoretical literature on peak load pricing.

The third issue, variable coefficients of production, concerns a substantive issue. Most of the literature on peak-load pricing implicity assumes a fixed proportions technology. Basically, there is only one way to produce telephone calls so the cost of capacity is well-defined and similarly, operating costs are also uniquely defined. More generally, however, there may be different types of capacity with different costs. One could imagine that telephone services could be produced in a very capital intensive way, using computers and very little labour, or in a more labour intensive way. The former method would involve high capacity costs and low operating costs while the latter might involve lower capacity costs and higher operating costs. Furthermore, a firm could possibly mix different types of capacity. In electricity generation this issue is of considerable importance, in telephone pricing it is less important.

A general treatment of peak-load pricing with variable proportions is Panzar (1976). Related papers, in which a few different fixed proportions technologies are available, are Crew and Kleindorfer (1976), Wenders (1976), and Turvey (1968).

Panzar assumes what he refers to as a "neoclassical" production function: output in any period is a function of capital, which cannot be altered in the short run, and variable inputs: $x_j = x_j(L_j, k)$ where L_j is a vector of variable inputs used in period j, and k represents fixed capital. The "neoclassical" feature of this production function is that it is assumed to be continuously differentiable with respect to all inputs. The maximum capacity for any period depends on k and is reached when the marginal products of the variable factors are driven to zero.

It is fairly clear that because marginal products of variable factors approach zero continuously, it will never be optimal to run the system at capacity. It only pays to increase output up to the point where the value of the marginal product for each variable factor is just equal to the cost of that factor. This must occur before capacity is reached.

The standard formulation of peak-load pricing implicitly assumes that marginal products remain constant but drop discontinuously to zero at capacity output. This discontinuity in the marginal product functions makes corner solutions (i.e. production at capacity) desirable. Normally the implicit production function is of the fixed proportions type. There may be 100 units of capital. Each call requires a certain amount of energy and labour and uses a certain portion of the capital.

Thus the marginal product of labour and energy are constant until capacity is reached.

Panzar's second striking result is that users in all periods should contribute towards capacity costs, rather than just peak period users. This depends critically on short-run decreasing returns to scale. It is reasonable to suppose that, holding capital fixed, returns to scale in the variable factors will be decreasing over most output ranges. It is possible, however, that for periods of very low demand short run returns to scale would be constant or even increasing, which would imply that contributions of users in these periods should be equal to or below marginal operating costs.

Panzar does derive a peak-load result in the traditional spirit. Specifically, periods with higher output rates should have higher prices than periods with lower output rates, and users in higher output periods make a greater contribution above marginal costs than do users in periods of lower output.

Wenders (1976) and Crew and Kleindorfer (1976) have results in a similar vein, but which are less extreme due to the less extreme characterization of production in their models. The implication of this work which recognizes that different production techniques are possible is that loading all capacity costs on peak period users is excessive. It makes sense to charge of range of prices over the different periods according to the strength of demand in each period instead of using the stark two-price regime implied by earlier work, and which regulators have found unpalatable in any case.

Putting this variable coefficient problem aside, the next problem to consider is the possibility of a shifting peak. As suggested by the name, a shifting peak occurs when the imposition of peak load prices causes the peak period to change. One would imagine that a shifting peak is very likely if marginal capacity cost is very high while marginal operating cost is comparatively low. If, at equal prices, the peak period has only slightly higher demand than some other period, it seems quite plausible that the imposition of the full marginal capacity cost on that one period would suppress demand enough so that another period would emerge as the peak.

So far we have made the assumption that the peak period remains fixed, in order to present the basic insight as clearly as possible. However, the shifting peak does not cause any theoretical problem. The solution is that capacity costs should be borne by users in different periods. This is a second reason that the entire capacity cost should not be borne by users from just one period.

The optimal prices can be characterized quite easily in a slightly modified version of the model we have been using. In the case of the shifting peak the solution involves having consumption up to capacity in more than one period. Assuming there are just two periods, the regulator's problem is

$$\text{Maximize } B(x_1,x_2) - c(x_1,x_2, k)$$
$$\text{subject to } k = x_1, \; k = x_2$$

where B is a consumer benefit function and c is a cost function depending on capacity k and usage in each period. The associated Lagrangian function is

$$L = B(x_1,x_2) - c(x_1,x_2,k) + \lambda_1[k - x_1] + \lambda_2[k-x_2]$$

which yields the following solutions

$$p_1 = \partial c/\partial x_1 + \lambda_1$$

$$p_2 = \partial c/\partial x_2 + \lambda_2$$

$$\partial c/\partial k = \lambda_1 + \lambda_2$$

The partial derivatives $\partial c/\partial x_1$ and $\partial c/\partial x_2$ reflect marginal operating costs in each period. λ_1 and λ_2 are both positive and, when added together, equal marginal capacity cost. Thus users in both periods should contribute toward capacity cost. λ_1 and λ_2 are different, however. Efficiency requires that price equate supply and demand in each period. Since quantity consumed is the same in each period, the period with stronger demand must have a higher price. If marginal operating costs are the same in each period, then the period with stronger demand must have a higher λ and must contribute more toward capacity costs than the other periods.

This observation that different prices should be charged even when capacity is reached in each of two periods led Steiner to refer to this case, rather misleadingly, as an example of price discrimination. We tend to think of price discrimination as occuring when different consumers are charged different prices even though marginal cost is the same for both. In this case, the full marginal cost is different in the two periods because, given capacity, it includes a scarcity cost (equal to λ) that differs between the two periods. Essentially, if a user in one period is to consume more, another consumer in that period must consume less. The full marginal cost is the value to this other consumer, which is higher in the period of stronger demand.

It is clear that additional periods, each of which may or may not share the peak, can be added to this model with no extra complications. A well-explained account of the shifting peak problem is Kahn (1970).

In almost all of the literature on peak-load pricing it is implicitly assumed that the number and duration of the periods is exogenously given or obvious from the nature of the problem. In actual application, however, the structure of the periods must be determined. Should the day (or week) be broken into two, three, four or more periods for peak load pricing, and how long should each period be. In the case of telephone service, it seems that there is a natural peak corresponding to business hours, a natural intermediate period in the evening, and a natural off-peak at night. Even in the case of telephone service, however, it is not clear where the boundaries should be drawn. Furthermore, even within these "natural" periods, demand varies systematically. (There is an obvious lull, for example,during the lunch period.) In fact, there is a continuously varying pattern of demand over the day, week, and year.

Unfortunately, systematic theoretical analysis of this aspect of the peak-load problem is rather difficult. Craven (1971) examines a version of the problem in which demand at every instant depends on prices at that instant but is independent of the rest of the price profile. However, substitution across periods seems likely to be the most important feature of this problem of dividing the cycle into periods. One would expect that the costs of monitering by the telephone company and of information processing by consumers would make very complicated systems undesirable. Probably the efficient way to solve the problem for any particular case would be by simulation, using whatever information about demand and cost was available. If 2, 3 and 4 period per day regimes were compared for a variety of plausible boundaries, one could be fairly confident of selecting a reasonable periodic structure. Such a simulation exercise would not be trival, because simultaneous selection of periods and prices would be required.

Another implicit assumption made so far and made in the early literature is that demand in each period is a deterministic function of prices. In practise, of course, demand has a strongly stochastic character, so that the telephone company and regulator cannot know with certainty what demand will be forthcoming on any particular time at any particular price. This uncertainty itself affects the pricing problem. In addition, it implies that demand will sometimes exceed capacity, in which case there must be some method of rationing consumption among demanders.

Brown and Johnson (1969), Carlton (1977) and Visscher (1974) all consider the problem of setting a single price when demand varies stochastically over a cycle. Although this work does not directly concern peak-load pricing, it raises some relevant issues. First of all, it shows that optimal pricing prescriptions are sensitive to the method of rationing. If there were no transactions costs one would expect that the commodity would be rationed to consumers who valued it most highly. A consumer with low willingness to pay would always sell to a consumer with higher willingness to pay. This seems absurd for local telephone service. The best assumption is that, should demand exceed capacity, consumption is allocated randomly. Unfortunately, the one paper that treats peak-load pricing when demand is both cyclical and uncertain (Crew and Kleindorfer, 1976) assumes that, under rationing, consumption is allocated to users with the greatest willingness to pay.

Nevertheless, a few general statements about the effect of uncertainty can be made. First of all, the basic motive for peak-load pricing remains. Also, with demand uncertainty, optimal capacity tends to be greater than with certainty. Peak prices tend to be lower under uncertainty, while off-peak prices may be either lower or higher. The basic idea is that price should be set equal to expected costs. In the peak period there is some chance that capacity will not be reached so the price is a weighted average of marginal capacity cost + marginal operating cost and marginal operating cost by itself. There must be less than the peak price in the deterministic case which is just marginal capacity cost + marginal operating cost.

Consideration of demand uncertainty shows that optimal pricing is related to optimal reliability of the telephone system. Reliability is often thought of as a separate issue, however, and in the case of local telephone service (but not long distance), the objective of achieving near 100% reliability seems to be regarded as an independent objective. Very high levels of reliability would be appropriate if marginal capacity costs were low compared to the willingness to pay of consumers who were unable to complete telephone calls under rationing. Since some calls are of great value (emergency calls to ambulance services, etc.), a plausible case for very high levels of reliability might be made in the case of telephone service.

The final issue concerning our simple characterization of the peak load pricing problem is the possibility of a minimum profit constraint. Ordinary peak-load pricing, in which prices are set equal to the appropriate marginal cost, will not necessarily generate enough revenue to cover total cost. For example, if the cost function is $c = F + bx_1 + bx_2 + \beta k$ then pure peak load pricing ($p^1 = b + p$, $p^2 = b$ if period 1 is the peak period), will involve a loss equal to F. Generally, whenever there are large fixed costs, marginal cost pricing is unlikely to cover costs.

The marginal cost pricing doctrine, even extended to peak-load pricing, implies, therefore, that subsidization of telephone companies might be necessary. For various reasons this is unacceptable. One argument has already been mentioned: if consumers do not pay the full cost of the goods and services that they consume, then there is a transfer from society at large to these consumers. In addition, if a public utility has access to general revenue it has little incentive to produce efficiently. There are of course problems with rate of return regulation, but at least there is some discipline imposed on the utility's managers.

The peak-load problem with a profit constraint is taken up in Mohring (1970) and Pressman (1970). Also, the peak-load problem can be regarded as a special case of the multi-product problem examined by Baumol and Bradford (1970). In this problem the possibility that demand in each period depends on prices in all periods becomes important. Without a binding profit constraint this demand interdependence has no effect on the optimal pricing rules so it has not been discussed so far. In the profit-constrained case, however, demand interdependence does complicate the results.

We start with the case of demand independence. The problem is to maximize $W = B(x^1,x^2) - c(x^1,x^2,k)$ subject to a capacity constraint: $x^1 = k$ and a profit constraint $R(x^1,x^2) - c(x^1,x^2,k) \leq \pi_0$, where R is revenue and π_0 is the profit target. The Lagrangian function is

$$L + B(x^1,x^2) - c(x^1,x^2,k) + \lambda[k - x^1] + \mu[p^1x^1 + p^2x^2 - c(x^1,x^2,k) - \pi_0]$$

The first order conditions are

$$p^1 - mc_1 - \partial c/\partial k + \mu[p^1 + x^1dp^1/dx^1 - mc_1 - \partial c/\partial k] = 0$$

$$p^2 - mc_2 + \mu[p^2 + x^2dp^2/dx^2 - mc_2] = 0$$

where $mc_i = \partial c/\partial x^i$ for i = 1,2. Letting $\eta^i = -(p^idx_i/dp^i)/x^i$, which is the elasticity of demand for use of telephone service in period i, the following results are obtained:

$$p^1 = \gamma^1(mc_1 + \partial c/\partial k)$$

$$p^2 = \gamma^2(mc_2)$$

where $\gamma^i = (1 + \mu)/(1 + \mu + \mu/\eta_i)$

Provided the profit constraint and capacity constraint are both binding, $0 > \mu > -1$ and $\gamma^i > 1$, so both periods pay a price above the associated marginal cost, even the off-peak period. The more inelastic demand is (i.e. the smaller η^i is), the higher the mark up over marginal cost for that period. It is even possible, if demand in the peak period were very elastic while demand in the off-peak period were very inelastic, that the optimal off-peak price could be higher than the optimal peak price..

Incorporating interpendent demands makes the expressions for prices more complicated. In particular, each price depends not only on own demand elasticities but also on cross elasticities of demand and the elasticity for the other period as well. Provided that telephone calls in one period are a substitute for telephone calls in the other, both prices will be greater than the associated marginal cost.

This concludes our survey of peak load pricing. Clearly there are many considerations involved in designing an appropriate peak load schedule. Nevertheless, the main insight remains important: prices in each period should be systematically related to marginal cost in that period. Such a pricing system would encourage efficient use of the telephone system. Actually estimating the benefits of a change to peak load pricing, so that they could be compared with the costs of implementation, is not likely to be easy. Even in a static context, demand and cost functions must be estimated. The greater problem, however, is dynamic. Because of technological development the pattern of telephone use is changing rapidly so that estimates based on historical, or even current, data are not likely to be very accurate predictors of future benefits. They are likely, however, to underestimate future benefits, so such estimates would be of some use.

5. Nonlinear Pricing

So far we have assumed that telephone service would be provided at a price that did not vary with quantity. Total outlay, R, would just be price times quantity, R = px: outlay is a linear function of quantity. However, the complication that a telephone company might face a binding minimum profit constraint gives rise to the possibility that more flexible pricing tools should be used.

In general, prices may vary with quantity so that outlay is a nonlinear function of quantity: R = p(x)x. The resulting price system is referred to as "nonlinear pricing", "nonuniform pricing", or "quantity-dependent pricing".

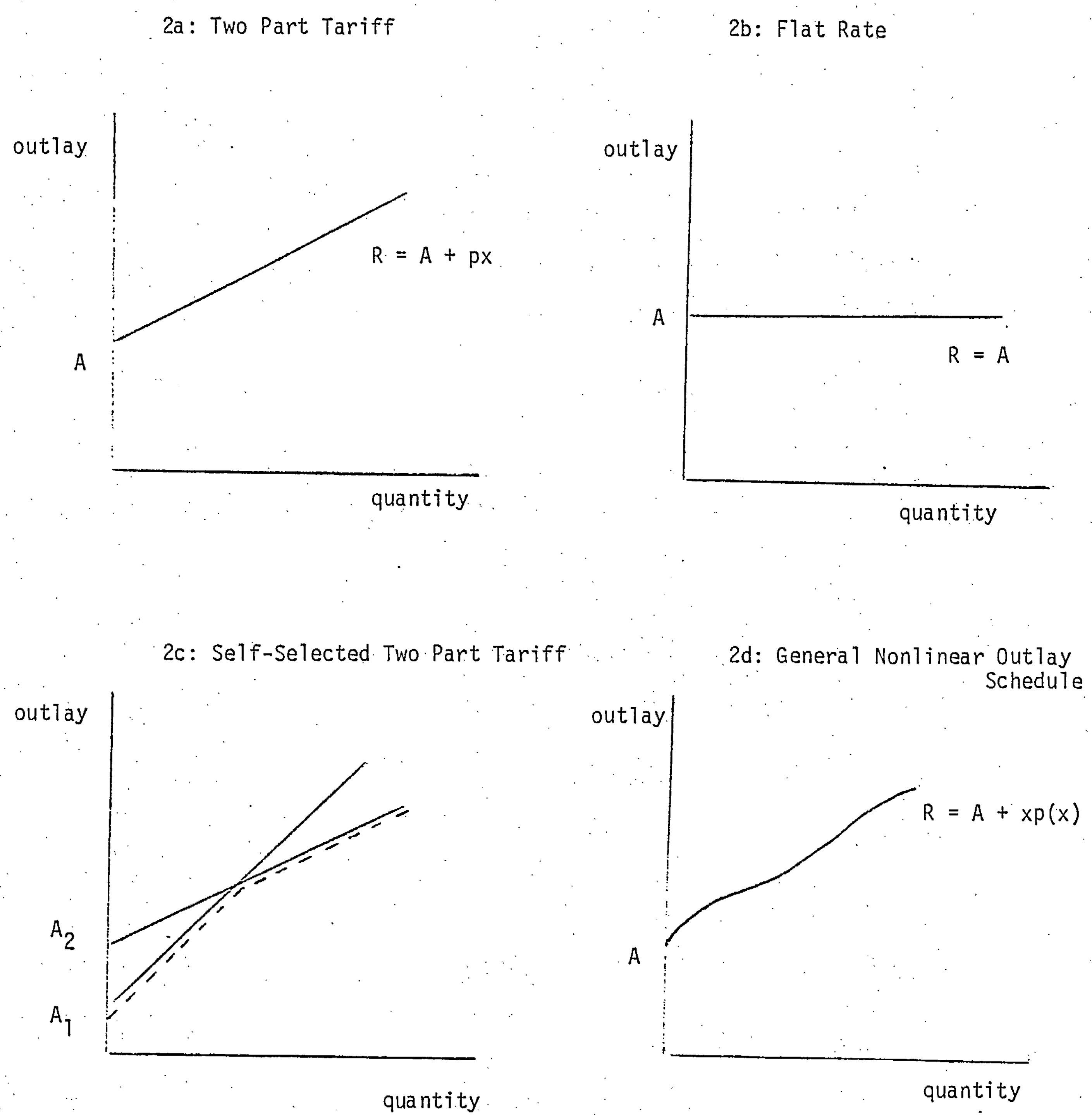
The simplest case is that of the two part tariff. A two part tariff involves charging each consumer an access or entry fee, A, and a constant usage price, p, per unit consumed. The outlay schedule is, therefore

$$R(x) = \begin{cases} 0 \text{ if } x = 0 \\ A + px \text{ if } x > 0 \end{cases}$$

See figure 2a.

A flat rate schedule is a special case of the two part tariff in which the usage price is set equal to zero; the consumer pays an entry fee for the right to consume an unlimited amount. Thus local telephone service has a flat rate schedule.

Figure 2



2a: Two Part Tariff
outlay
R = A + px
A
quantity
2b: Flat Rate
outlay
A
R = A
quantity
2c: Self-Selected Two Part Tariff
outlay
A2
A1
quantity
2d: General Nonlinear Outlay Schedule
outlay
R = A + xp(x)
A
quantity

Instead of offering each consumer a single two-part tariff a telephone company may adopt the more sophisticated strategy of offering a choice between different two part tariffs. Two tariffs, (A_1, p_1) and (A_2, p_2), might be offered with $A_1 < A_2$ and $p_1 > p_2$. (See figure 2c.) Which schedule is preferred by the consumer depends on his level of consumption. In effect the true outlay schedule is the lower envelope of the two outlay schedules in figure 2c:

$$R(x) = \min \{A_1 + p_1 x, A_2 + p_2 x\}$$

In general a nonlinear pricing structure could be quite complex; outlay schedules need contain no linear segments at all. The general form of a nonlinear outlay schedule is

$$R(x) = \begin{cases} 0 \text{ if } x = 0 \\ A + p(x) \cdot x \text{ if } x > 0 \end{cases}$$

Higher values of A combined with negative values of $p'(x)$ imply volume discounts.

Discussions of the literature of two part pricing and nonlinear pricing usually begin with Gabor (1956) and Oi (1971). Early advocates of two part pricing were Coase (1946) and Lewis (1941). Feldstein (1972) considers mixing equity and efficiency objectives and Ng and Weisser (1974) characterize optimal budget-constrained two part tariffs. Panzar (1978) and Faulhaber and Panzar (1977) consider optimal (or self-selected) two part tariffs. The general nonlinear problem is taken up by Goldman, Leland and Sibley (1977), Spence (1977, 1980), Willig (1978), Roberts (1979), and Mirman and Sibley (1980).

Before describing the character of optimal two part or general nonlinear tariffs there are two practical conditions that must be satisfied. First, the prices must be such that consumers are not induced to resell the commodity. This is not likely to be a serious problem for telephone service since relying on another consumer's telephone is rather awkward. The second condition, however, is a serious consideration for telephone use. Specifically, the total quantity used by each consumer must be measurable at relatively low cost. Measuring equipment has a large fixed cost component for every exchange which suggests that usage sensitive pricing might not be appropriate for small exchanges.

Optimal Two Part Tariffs

The simplest nonlinear structure is the two part tariff, and two part tariffs are already widely used by telephone companies in Europe and the United States. The basic problem is that pure marginal cost pricing does not cover costs. However, the consumption levels chosen under marginal cost pricing are the socially efficient levels of consumption. If there were just one consumer (or many very similar consumers) a two part tariff could achieve the "first best" solution. The usage price, p, would be set equal to marginal cost and an access charge less than the consumer surplus associated with price p would be charged to cover the resulting deficit. This is clearly superior to the uniform pricing case in which price must be set equal to average cost.

The problem with two part tariffs is that consumers differ substantially so that, at any access fee - price combination there are some consumers just at the margin who will stop consuming if the access fee (or price) is raised. Any positive access charge will exclude some consumers whose consumption is socially desirable. On the other hand, prices above marginal cost induce every consumer to consume less than the socially desirable quantity. The optimum two part tariff generally involves a mix of these two distortions.

Before setting out a formal statement of the two part pricing problem one aspect of telephone pricing should be mentioned. From the production point of view provision of access (i.e. a telephone line and phone) is essentially a different good from additional phone calls given that the phone and line are in place. Thus the monthly access charge reflects both the cost of access and the access fee part of a two part tariff on phone usage. Even a pure marginal cost solution would then involve a positive access charge equal to the (monthly) marginal cost of access and a positive usage price equal to the marginal cost of additional phone calls. In addition, any once and for all marginal connection costs would be recovered by a once and for all connection charge.

The pure two part tariff case is the case in which there is no (monthly) access cost, just a pure once and for all connection cost and a marginal usage cost. If access really is a separate good (on a month by month basis) from usage then the optimal pricing problem becomes a multi-product pricing problem. Multi-product pricing is discussed later.

A formal two part tariff model based on Faulhaber and Panzar (1977) follows. Consumers are assumed to vary in their tastes and are indexed by a continuous variable, θ which is the "taste" parameter. A consumer of type θ has utility function: $U = u(x,\theta) + y$. Faced with a two part tariff (A,p) the consumer will choose his level of consumption so as to maximize U subject to his budget constraint:

$$A + px + y \leq I$$

As before, the consumer's solution involves setting the marginal utility of telephone use equal to the price: $MU_x = p$. We define the consumers gross surplus S as $u(x,\theta) - px$. Thus, provided $S \geq A$ the consumer will choose to purchase telephone service:

$$x(A,p,\theta) = \begin{cases} 0 \text{ if } S(p,\theta) < A \\ x(0,p,\theta) \text{ if } S(p,\theta) \geq A \end{cases}$$

If $S \geq A$ then the consumer gets some net benefit from telephone use. We define the marginal consumer type $\hat{\theta}$ as the consumer type for whom $S = A$. This defines $\hat{\theta}$ by the equation $S(p,\hat{\theta}) = A$. We assume that θ ranks consumers according to their "taste" for x so that i) at each p a higher θ corresponds to a higher level of consumption and ii) the marginal consumer type, $\hat{\theta}$, is unique.

For convenience the variable θ is scaled so that θ is distributed over the interval [0,1] according to density function $m(\theta)$. Total benefit to consumers is then

$$B(A,p) = \int_{\hat{\theta}}^{1} (S(p,\theta) - A)m(\theta)d\theta$$

Let N equal the total number of consumers consuming telephone service. Then profit of the telephone company is

$$\pi = A \cdot N(A,p) + p \cdot X(A,p) - c(X(A,p))$$

and the objective of the regulator is to maximize $B + \pi$ subject to $\pi \geq \pi^*$, the minimum required profit. The Lagrangian function is

$$= B(A,p) + (1 + \lambda)\pi(A,p)$$

which gives rise to first order conditions

$$(1 + \lambda)[(p - c')X_A + AN_A] + \lambda N = 0$$

$$(1 + \lambda)[(p - c')X_p + AN_p] + \lambda X = 0$$

Unfortunately, solving these first order conditions for p and A requires knowledge of the derivatives X_A, N_A, X_p and N_p (or equivalently, the elasticities of usage and number of consumers with respect to both the access fee and usage price) at the optimum.

The following general statements can be made about the solution:

1) welfare with a two part tariff is strictly greater than with either a uniform price or a flat rate.
2) the optimal price exceeds marginal cost and the optimal access charge is positive.

Both of these statements are sensitive to the assumption that it is possible to rank consumers monotonically according to their taste for x regardless of p(i.e., that demand curve do not cross.)

To summarize, the basic idea is that the optimal two part tariff involves trading off two distortions: high access charges drive out socially desirable consumers while high prices cause all consumers to consume too little. The information required to set an optimal two part tariff is disturbingly detailed. Furthermore, one suspects that the gain in going from usage charges equal to marginal costs (with access charges to cover the deficit) to optimal two part tariffs is rather small.

Optimal (Self-Selected) Two Part Tariffs

Several U.S. telephone companies have allowed consumers to select which of two or three two part tariffs to be billed under. One of the options is generally a flat rate tariff. These optional schemes are very useful in helping to overcome consumer resistance to measured service. In addition, optional two part tariffs can generate pareto improvements over pure flat rate schemes. That is, large consumers, small consumers and telephone company can all be made better off by instituting optional two parts in place of a single flat rate system. The basic insight is rather simple (see Panzar, 1978). Consider a set of optional two part tariffs which includes the original flat rate system as an option. Consumers will choose a measured system (with usage price above zero) only if by doing so they become better off. Those who do choose a measured system (with a usage price close to marginal cost) will be induced to consume more efficiently. That is, each will consume less than before: only up to the point at which the marginal benefit equals the usage price. The improvement in efficiency also allows the telephone company to gain.

Despite this rather pleasant result, optional tariffs which include the original flat rate system as an option are not likely to have a large effect on overall efficiency. Very large users will continue to select the flat rate and place burdens on the system. If local service is to generate larger revenues, some of that revenue should come from high volume users. For this reason many economists regard pareto-improving optional two part tariffs that retain attractive flat rate options as only a minor improvement over existing flat rate schemes.

One variation on optional two part tariffs is ex post pricing. (Danbsy and Panzar, 1981). Consumers often don't know in advance which of two or more optional two part tariffs would be better for them ex post. In addition, many consumers have a sufficiently variable calling pattern that the best tariff for one month is not necessarily the best tariff for all months. Therefore consumers will be willing to pay a premium to be able to have the tariff that is revealed to be least cost at the end each month applied to their usage. Such a system can be welfare-improving and may reduce general consumer resistance to measured service. Effectively, the premium can be thought of as an insurance premium paid to reduce the variance of monthly payments.

As mentioned, two part tariffs are a special case of nonlinear pricing. For that matter, optional two part tariffs are equivalent to a single more general nonlinear pricing system. In principle the efficiency gains from using arbitrary nonlinear pricing schedules are greater than from using simple two part tariffs. Also, some industries do use multi-part or block tariffs, which are more sophisticated examples of nonlinear pricing structures. However, for local telephone service, it is questionable whether more general nonlinear pricing structures would generate much economic benefit. In any case, relatively little attention has been paid to the possibility of using pricing structures more sophisticated than two part tariffs.

6. Multi-product Pricing, Price Discrimination and Cross Subsidization

Multi-product pricing and price discrimination are conceptually different issues; however, they are formally very similar and both give rise to problems with cross-subsidization. There are two reasons why optimal multi-product pricing is relevant for local telephone service. First of all, local telephone service is only one of several services provided by telephone companies, the other important service being long distance service. Secondly, local telephone itself is probably most accurately regarded, at least from the production side, two products: access and usage. (There is also a third product: the telephone itself.)

If there is no binding profit constraint the multi-product aspect of telephone service does not affect the optimality of marginal cost pricing. However, if there is a binding profit constraint a new insight emerges: the markups over marginal cost for each product should be related to the own and cross price elasticities of demand. Specifically, high markups should be associated with low elasticities. This idea was first developed by Ramsey (1927) in the context of optimal taxation. Optimal pricing formulations of the idea are associated with Boiteux (1956) and Baumol and Bradford (1970).

The basic problem is set up below. We assume that each product is to have a uniform price, although in principle nonlinear pricing could be incorporated in the same problem (see Spence (1980) and Mirman and Sibley (1980)). As before the problem is to maximize the sum of producer and consumer surplus subject to the profit constraint:

$$\text{Max } B(x) - C(x) \text{ subject to } R(x) - C(x) \geq \pi^*$$

where $x = (x_1,\ldots,x_m)$ and R = revenue.

The Lagrangian is $L = B(x) - C(x) + \lambda[R(x) - C(x)]$ from which the first order conditions are obtained:

$$L_{x_k} = \partial B/\partial x_i - \partial C/\partial x_i + \lambda[\partial R/\partial x_i - \partial C/\partial x_i] = 0$$

As before $\partial B/\partial x_i = p_i$. $\partial C/\partial x_i$ and $\partial R/\partial x_i$ are denoted MC_i and MR_i (for marginal cost and marginal revenue) respectively. Thus the first order condition can be written:

$$p_i - MC_i = \lambda(MC_i - MR_i)$$

In the case in which cross elasticities of demand are zero $MR_i = p_i + x_i dp_i/dx_i$ so the condition can be rewritten

$$(P_i - MC_i)/p_i = \frac{\lambda}{1+\lambda}\frac{1}{\varepsilon_i} \qquad (7)$$

where $\varepsilon_i \equiv -\frac{dx_i}{dp_i}\frac{p_i}{x_i}$, the own elasticity of demand. This is the so-called "Ramsey rule". If cross elasticities are important the markup expression is somewhat more complicated.

In this case x_i is interpreted as the output of product i. However, the same analysis applies to price discrimination for a single product. Price discrimination is defined as charging different prices to different groups of consumers of the same product despite equal marginal costs. If we interpret x_i as consumption by group i and x_j as consumption of the same product by group j then the resulting optimal price discrimination formula is derived exactly as above. One difference is that cross-elasticities must be zero in the price discrimination case so that formula (7) applies generally.

The price discrimination result is due to Hartwick (1978). For telephone companies the important type of discrimination is between business users and residential consumers. Also, it is possible to consider offering different two part tariffs to different groups of users. An analysis of optimal discriminating two part tariffs for business and residential customers is Brander and Spencer (1981).

The objective of Ramsey optimal multi-product pricing and Ramsey optimal price discrimination is to maximize surplus subject to the profit constraint. This objective is agnostic to pure transfers of income between consumers. Thus large transfers can easily be implied in order to achieve small efficiency gains so that cross-subsidization is possible. Under Ramsey optimal pricing some groups of consumers, or consumers of a particular product, may subsidize other consumers.

Ramsey pricing is related to the "value-of-service" concept used by telephone companies in setting rates. Specifically if some groups have a high "value of service" they are likely to have low elasticities of demand, and Ramsey optimal pricing will involve relatively high mark-ups over marginal cost for these groups. Telephone companies have used the value of service concept to justify charging high rates to business users for local service.

Telephone companies also subsidize local service from long distance service. This appears to have little or no justification from a Ramsey optimal or value of service perspective. Instead, it is a result of trying to achieve universal service by charging low rates for access.

Many people find it hard to justify cross-subsidization on equity grounds. Why should users with low elasticities subsidize users with high elasticities. Consequently it has been suggested that pricing be constrained to be "subsidy-free" in the following sense:

Each consuming group should pay no less than the incremental cost of its service and no more than the stand-alone cost of its service.

The incremental cost of x_i is $C(x_i,\ldots,x_i,\ldots,x_n) - C(x_1,\ldots x_{i-1},0,x_{i+1},\ldots,x_n)$ and the stand-alone cost of x_i is $C(0,\ldots,0,x_i,0,\ldots,0)$.

Under this approach price discrimination and multi-product pricing would satisfy the so-called "anonymous equity" criterion. (See Faulhaber (1975, 1979) and Willig (1979) on these and related matters.)

Cross subsidization is related to sustainability problems. The situations in which cross-subsidization is large are precisely the cases in which nonregulated firms are likely to find it profitable to enter the industry and compete with the regulated firm. Such competition is not allowed in Canada, but in the U.S. this has become a serious problem. Private companies can offer long distance service more cheaply than the telephone companies who use long distance service to subsidize local servive. Two papers on sustainability are Baumol, Bailey and Willig (1977) and Panzar and Willig (1977).

Ramsey optimal pricing would involve usage sensitive prices for local service. Also, if it is decided that the large subsidy from long distance service to local service is undesirable, usage sensitive prices are the obvious tool to reduce the subsidy.

7. Other Issues

i) Consumption Externalities

The main issues in optimal pricing have been addressed. However there are some other concerns which should be mentioned. One issue that is often mentioned in connection with telephone service is consumption externalities. The basic point is that when a new subscriber joins the telephone system he confers benefits on all other users because they now have the option of phoning him. Thus any subscriber does not capture the full benefits of his joining. If his private benefits are less than the marginal cost of access but total benefits exceed the marginal cost of access, then marginal cost pricing will stop him from joining even though it is socially desirable that he do so. (See Littlechild (1975) for a theoretical treatment of this point.)

This consumption externality is often used as a justification for subsidizing access charges. This consumption externality is undoubtedly important when telephone penetration is low. The externality, if left unadjusted for, would keep penetration or total access too low by the criterion of social efficiency. When penetration is as high as it is in Canada, however, the externality problem ceases to be an important source of inefficiency. Indeed the pursuit of "universal access" as an explicit objective can be regarded as a solution to the externality problem.

In the unconstrained case, the specific solution to the externality problem would be to charge an access fee equal to the marginal cost of access minus the benefit to existing consumers from having one more subscriber in the system. Even with a profit constraint, the optimal access fee would be relatively low if consumption externalities were important. To generate the required revenue, usage prices or prices of other services would have to be raised.

ii) The Averch-Johnson Effect

So far we have examined the optimal pricing problem from the point of view of the regulator. However, actual prices are not set by regulators, but are proposed by the regulated firms and either approved by the regulator or renegotiated. Regulation may take the form of insuring that the regulated firm not earn more than an appropriate rate of return on capital. Under such circumstances profit maximizing firms will have a tendancy to use "too much" capital and consequently not produce at minimum cost. This is known as the Averch-Johnson effect and was first analysed by Averch and Johnson (1962). (See also Baumol and Klevorick (1970) and Bailey (1973).)

This point is tangential to a survey of optimal pricing. Nevertheless, one point should be made. Giving regulated firms greater flexibility to use two part, multi-part or general nonlinear pricing systems tends to increase the Averch-Johnson distortion. In any case every discussion of optimal pricing should contain a reminder that regulated firms do not have any particular incentive to pursue general efficiency or equity goals. Economists generally assume that they will maximize profits. Consequently, giving regulated firms greater flexibility is not as benign as it might seem.

This problem of incentives is difficult and there is some interesting recent work that focusses on setting up the regulatory environment in such a way that profit maximizing behaviour by firms leads them to charge optimal prices from a welfare point of view. Vogelsang and Finsinger (1979) suggest such a regulatory scheme for a multi product firm. The basic idea is that the regulator insists that the firm meet all demand forthcoming at whatever prices the firm charges and that following each period of positive profit prices be constrained so that, if such prices had been charged last period, no excess profits would have been generated. An adjustment is made if profit in the previous period was negative. Subject to these constraints the firm is allowed to charge whatever prices it likes.

Interestingly, this algorithm improves welfare every period and approaches a Ramsey optimal pricing structure. However the V-T algorithm has some weaknesses. Specifically, it is static in the sense that it assumes that that exogenous conditions such as tastes and technology do not vary from period to period. Secondly, it can involve large losses for the firm in some periods. Warskett and de Fontenay (1981) suggest a similar algorithm which does not have this second drawback and which, they argue, is likely to be capable of modification so as to perform well in a dynamic environment. Both the V-F and W-F algorithms impose weak information requirements on the regulator; unfortunately they impose rather strong information requirements on the regulated firms. Furthermore, it is not clear how such algorithms would perform with two part tariffs or other nonlinear pricing schemes. This work seems very promising but is not yet at the operational stage.

8. Concluding Remarks

The objective of this survey is to bring together those aspects of optimal pricing theory that might be relevant for pricing of local telephone service. Not every possible consideration has been addressed. (For example, Mitchell (1981) considers the problem that arises when different telephone exchanges with different demand and cost conditions are forced to have the same prices so that some kind "average" optimal price is required.) However, the main issues in optimal pricing have been described carefully.

The underlying question is: should local telephone service have usage sensitive prices. Perhaps a few directs comments on this question are appropriate. First of all, the implication of all the optimal pricing considerations is that, ignoring the cost of implementation, USP should be used. However, the actual welfare gains in moving from a suboptimal pricing configuration to a second-best optimum tend to be small in comparison to the total benefits of the service. (This is a manifestation of the so-called "iron law of deadweight loss".) Consequently if the costs of implementation are high, adopting USP might not be desirable. In any case, there is little point in trying to achieve fine tuning in setting optimal prices. Most of the welfare gains would be achieved by having prices reasonably close to marginal cost and by using access surcharges to make up any deficit. This would involve peak load pricing since cyclical variation over the day and week is a very important aspect of telephone demand.

There is considerable consumer resistance and some industry resistance to usage sensitive pricing. At least part of this resistance seems to be due to misunderstanding. Specifically, telephone service in Canada has been provided according to two important principles: value of service pricing and universal access. These principles are sometimes advanced as reasons for resisting USP. It is, therefore, worth pointing out that these two principles would actually favour adoption of usage sensitive pricing.

First of all, consider value of service pricing. This means charging businesses higher rates for local service than households. This kind of pricing can be carried out more efficiently under usage sensitive pricing than under the current flat rate scheme. Indeed, Ramsey pricing involves this very idea: low elasticities should be associated with higher markups. Generally, low elasticities are associated with high value of service so Ramsey pricing will generally be consistent with value of service pricing. Even if businesses do not really have low elasticities so that "value of service" pricing is just an excuse for charging firms higher rates than households for equity reasons, such an objective can be pursued more efficiently with usage charges and access fees than with acces fees alone.

Secondly we consider universal access. It is sometimes argued that usage prices would force poor people and old people to give up telephone service, or at least to suffer severe hardship. In fact, for a given revenue target, usage sensitive pricing would lower access charges and would make having a telephone easier to afford. At current flat rates old people who make relatively few phone calls actually subsidize high use households. (More accurately, old people come closer to covering their costs than do high use households, since both are subsidized by business users and by long distance service.) Furthermore, if it is regarded as important for equity reasons that old people or poor people have telephones they can simply be given lower rates. This kind of price discrimination is a much more effective method of achieving the equity objective than the current flat rate system. It just doesn't make sense to constrain an entire pricing system to be inefficient to meet the needs of certain users when those needs can be easily met by an efficient pricing system.

The issue of whether usage prices are "equitable" is logically separate from the question of whether local service rates should be higher overall so as to reduce the subsidy from long distance service. However, if higher revenues are required from local service, usage sensitive prices will be the most efficient and most equitable method of generating these revenues. Thus we might expect to see the introduction of usage rates being coincident with higher local rates.

The main analytic issue, then, is simply whether the economic benefits of USP would outweight the economic costs. Normally such a question would be answered by looking at current evidence. However, the telephone industry is changing so rapidly, both on the production side and the demand side, that current evidence is not likely to provide reliable estimate of costs and benefits even a few years from now. Most of the developments are in the direction of making usage sensitive pricing more attractive and less costly. Therefore current estimates should be regarded as conservative lower bounds on the potential benefits from usage sensitive pricing.

Introduction

This report deals with pricing of telecommunications services, particularly use of network facilities by businesses and households. The central issue is the desirability of usage sensitive pricing (USP) and local measured service (LMS) for households and businesses. USP requires making total payments by any consumer sensitive to his total usage. This will generally involve charging each consumer in relation to the costs his use imposes on the telephone system, although value-based discriminating usage prices are also possible. We feel that the principal efficiency gains from USP are almost entirely captured in cost based pricing. In the past it was difficult and costly to assess the costs of providing services on a call by call basis and thus rules of thumb were used to establish prices. These rules tended to emphasize value of service and ignore cost of provision. Technical breakthroughs have made the monitoring of use of the network on a subscriber by subscriber basis much less costly and thus USP seems feasible. USP has always been considered a procedure for fostering efficient use of resources in the production of services.

We survey, in Part I, theoretical aspects of pricing in settings like those characterizing the telecommunications system. The focus is on efficient pricing, which centres on relating price to marginal cost. In Part II, we survey the practices for charging for usage of various telecommunications systems. Attention is focussed on Canadian, U.S., and European practices and recent experiences. Institutional and technical issues are considered. In Part III, a survey of recent studies examining residential subscribers' response in terms of revised use patterns to changes in charges or prices. Of particular interest is the recent GTE experiment with local measured service (LMS) in Illinois. Our conclusions are presented in the last part. Since the issue involving most Canadians is the possible introduction of local measured service and related charges for their local residential use, LMS receives rather more attention. Central to our investigation was the question of eliciting more empirical information about the effects of LMS on use and welfare. Our observations appear in Part III and in the Concluding Remarks. Our remarks were influenced by our conviction that the telecommunications industry is in the midst of large shifts in demands resulting from new services which the network can provide (e.g. videotech devices, data transfer, etc.) and in supply resulting from new devices and equipment (e.g. electronic switches, satellite transmission, optical fibers, etc.).

PART II

Institutional and Technical Aspects of Measuring Telecommunications Services

Introduction

We survey recent pricing practises and related institutional matters in Canada, the U.S. and Europe in the section. It will be seen that the approach to pricing in various European countries is different from that being pursued in the U.S. Recent Canadian pricing practises are analyzed and charges compared with those in other countries. Technical aspects of monitoring usage on a call by call basis are surveyed with a view to isolating the relative costs of metering and billing local calls on a call by call basis. Two Appendices provide documentation of considerations of local measured service and usage sensitive pricing in the U.S.

Notes on the Current Telecommunications Scene in France

The seventh five-year plan in France made telecommunications development a priority. Projections are for 28 million subscriber lines by 1987, a substantial increase over the 6.2 million lines in 1974. The huge expansion in the capacity of the system provides a market for the new devices which have emerged since the transistor-electronics revolution in technology. The key elements in the revolution of the technology of telecommunication are electronic switching, digitalizing of the signals and fibre optical lines. Electronic switching is cheaper per line and has ready metering capabilities. Digital transmission of signals (relative to analogue signals) has the advantage of no loss in precision of the signal as distance is traversed. This makes digital signals virtually a necessity for data transmission. One is not usually concerned with minor distortion in voice transmission, but in other message forms minor distortion can make the transmission sufficiently unreliable to be useless. Fibre optical cables are fabricated from glass which in turn is derived from silicon, a most abundant and cheap input. Thus fibre optical cables will be substantially cheaper than copper cables and are much more compact.

France has a strategy to "wire" the country with modern equipment capable of delivering voice and non-voice messages with compatible or flexible devices. (The social, political and economic implications of this strategy were sketched in the remarkable report prepared for the President of France: S. Nora and A. Minc, The Computerization of Society, MIT Press, 1980.) The private sector is manufacturing the equipment but the planning and design specification is handled by public officials. The wave of videotech devices has been anticipated and there is a plan to have the French Antiope device reach an enthusiastic market. Each subscriber will receive free an electronic telephone director (a screen and console attached to his phone line). Telephone users will be obliged to become proficient with the device and presumably will be enthusiastic customers for the Antiope unit for phone-line compatible two-way communication. Via Antiope a user will have access to data banks, catalogues and ordering facilities, news services, etc. The demands on phone line and switching capacity will be substantially greater than in pre-videotech days, and a rationalization of usage with marginal cost pricing seems to be a necessity. One of the difficulties one encounters in analyzing changes in phone usage in reponse to price changes is that the demand schedule will be altered perhaps greatly by the new services which phone lines will be capable of providing. The effects of price changes on the old demand schedule may be small relative to the effects of shifts in the demand schedule in response to new capabilities of the phone line. We discuss rates in a cross section of European countries below but we note here that France does not yet have detailed measured rates of the kind now installed or in the planning stages in many European countries.

Notes on the Current Telecommunications Scene in the U.S.

Similar technical changes in telecommunications apparatus are being dealt with in the U.S. However the new technology makes metering calls so much less costly than with older devices and one observes not only rapid change in the hardware of telecommunications (changes such as electronic switching, fibre optical transmission cables and digitalization of signals) but changes in metering from flat rate for local subscribers to full local metering of calls. Two other forces have made local metering a natural development. The "interconnect" trend has made the long distance calling market more competitive and presumably there is downward pressure on prices and revenues. ("Interconnection" of non-Bell equipment with the Bell switched network was opened with the Carter phone court decision of 1968. In May 1970 the federal Communications Commission authorized "specialized common carriers" of inter-city traffic to connect with the local Bell (AT&T) switched network. MCI Communications Inc. was an early competitor in the newly opened inter-city traffic.)

The MCI system operates as follows: A customer in a city covered by MCI service who wants to make long distance call dials several "access" digits to gain entry to the MCI system. He then dials the number of the party he wants to reach. The signal travels on MCI's own microwave system and when it reaches the city "called", it reenters the local (frequently Bell) system and arrives at the distant phone. We reproduced some charges as of July 9, 1980 for Bell service and MCI service.

Call from New York

to	Minutes	Bell Charge	MCI Charge
Annaheim	5	$1.41	$0.74
Atlanta	10	2.45	1.33
Boston	17	3.62	1.86
Chicago	2	0.57	0.27
Dallas	33	8.26	4.56
Denver	15	3.81	2.11
Detroit	1	0.34	0.12
Los Angeles	16	4.27	2.36
Pittsburgh	14	3.36	1.61
San Francisco	2	0.63	0.30
Washington, D.C.	10	2.30	1.10

(Source: N.Y. Times, July 10, 1980)

With competition in the long distance market, traditional cross subsidization of local calling is now not as easy if at all possible. In order to get more revenues from the local calling sector or at least to bring revenues in line with costs, some form of measured service or charge per amount of usage seems natural. In fact the Federal Communications Commission, in Docket 20003 (Sept. 24, 1976) (Appendix II below) recommended that operating companies consider usage sensitive pricing (USP) as a response to revenue shortfalls brought about by increased competition. Thirdly, the U.S. has been swept by the technical changes which make the telephone line a service with many new functions, particularly for accessing computers, data banks, and in the future two-way videotech facilities. There seems little doubt that the demand schedule for the representative subscriber has or will shift substantially outwards. System wide capacity will become more scarce and must be priced accordingly. Some form of user charges resembling marginal cost pricing seems like the natural approach. As we note in Appendix I, all large companies are moving to a metered local service of some sort. The 1979 NARUC Annual Report on Utility and Carrier Regulation (Washington, D. C., 1980, pp. 593-596) contains reports of the views of various state regulatory commissions concerning local measured service, also. New York City has had local measured service since 1974 for both business and residential users. For business users in many large American cities, including New York City, Chicago, Los Angeles and Boston, business users have no flat rate option.

AT&T competitors and potential competitors persuaded the Federal Communications Commission (FCC) in 1976 to develop standards for equipment and in so doing circumvented AT&T restricting competition by dictating standards for equipment connected to its system. Also in 1976 AT&T competitors (including MCI Communications Corporation, ITT's United States Transmission Systems, Southern Pacific Communications Telenet, Graphnet, RCA American) successfully opposed a new Communications Act endorsed by AT&T which would have defined the competitive environment for non-Bell or non-AT&T companies.

In January 1981 the U.S. Government's suit filed in November of 1974 against AT&T reached the Courts. The government charged AT&T with excluding competition and stifling innovation in the telecommunications industry. The trial has not got underway because an out-of-court settlement was imminent in January and the judge deferred opening arguments. However, the Reagan administration requested another deferral in order that it could examine the Government's case. The pressure to break up AT&T, to separate its manufacturing and research organization from its operating organization, is no doubt a natural response to the vast increase in markets for new equipment brought about by the new capabilities of a linked network. New producers see a large new market to sell in. Rapid technical change has upset the equilibrium in the telecommunications industry.

Notes on the Current Telecommunications Scene in Canada

The Canadian telecommunications industry structure is unique. There is considerable public ownership (the systems in Manitoba, Saskatchewan and Alberta) and a dominant position of Bell Canada (centered in Ontario and Quebec but also with a substantial equity position in the Maritime operating companies, Newfoundland Telephone Company Ltd., New Brunswick Telephone Company Ltd., Maritime Telegraph and Telephone Company Ltd. (owns 44.45 of Island Telephone Company Ltd. serving P.E.I.)). The intercity network was essentially an AT&T monopoly in the U.S. but in Canada each separate company controls the lines in its territory and long distance service is organized by the consortium of member companies in TCTS (Trans Canada Telephone system) established in 1931. One assumes that a rate structure for long distance in part will reflect the cartel structure of the organizing institution. It appears that Canadian long distance charges have not been set to cross-subsidize local services quite as much as in the U.S. We discuss aspects of Canadian pricing below in a subsequent section.

Canada has often led in bringing new innovations on stream. The launching of the telecommunications satellite, Anik, in 1973, was a world first. Canada has seen the installation of DATAROUTE, a network for transmission of digital signals. DATAPAC is a linking of computers via the switched network. Any subscriber can dial into DATAPAC. A report in 1975 in the Financial Times indicated that charges on average to the Canadian consumer for local and long distance service rose less than 20% between 1958 and 1974 whereas the CPI rose by 70% over the same interval.

The Canadian telecommunications industry is being buffetted by the same major technical changes that we mentioned were upsetting an equilibrium in the U.S. Electronic and computer based control systems have made electromechanical systems obsolete. The new electronic systems make completely measured service a very low cost option. Rapidly changing costs, resulting from rapid technical change in the inter-city network have made revenue splitting and investment planning more difficult in TCTS (Ogle [1979; p.237]). The federal regulatory agency CRTC has received a consultant's report on these matters recently. Of great importance is the rapidly changing demand situation for telecommunications services. The new technologies have provided new services which have led to increased demand for use of both the switched and dedicated systems. Since flat rate charges can lead to significant distortions of charge per standardized use from the cost, companies are turning to some form of measured service.

For example, a subscriber might keep his line open 24 hours per day to a computer processing facility and be charged the same as a subscriber who makes a handful of three minute local calls per day under a flat rate charging system. A complete change in use patterns by households and small business can be anticipated with the mass use of videotech devices such as TELIDON and facsimile transfer units. In February 1981 Bell Canada announced the introduction of a "Display phone" which can be used as a regular telephone unit but also has a seven inch video display tube which can project data called up from a remote data storage system. Information in remote computer storage can be both called up and changed with this unit. A letter keyboard is part of the unit. The screen can display 25 rows of 40 or 80 character width. The unit also has capabilities as a computer itself and can handle electronic mail. Charges of using lines and switches should be brought in line with the costs of providing the services. The large question is whether capacity or demand is growing more rapidly under the rapid technical change. Which way will prices on average tend?

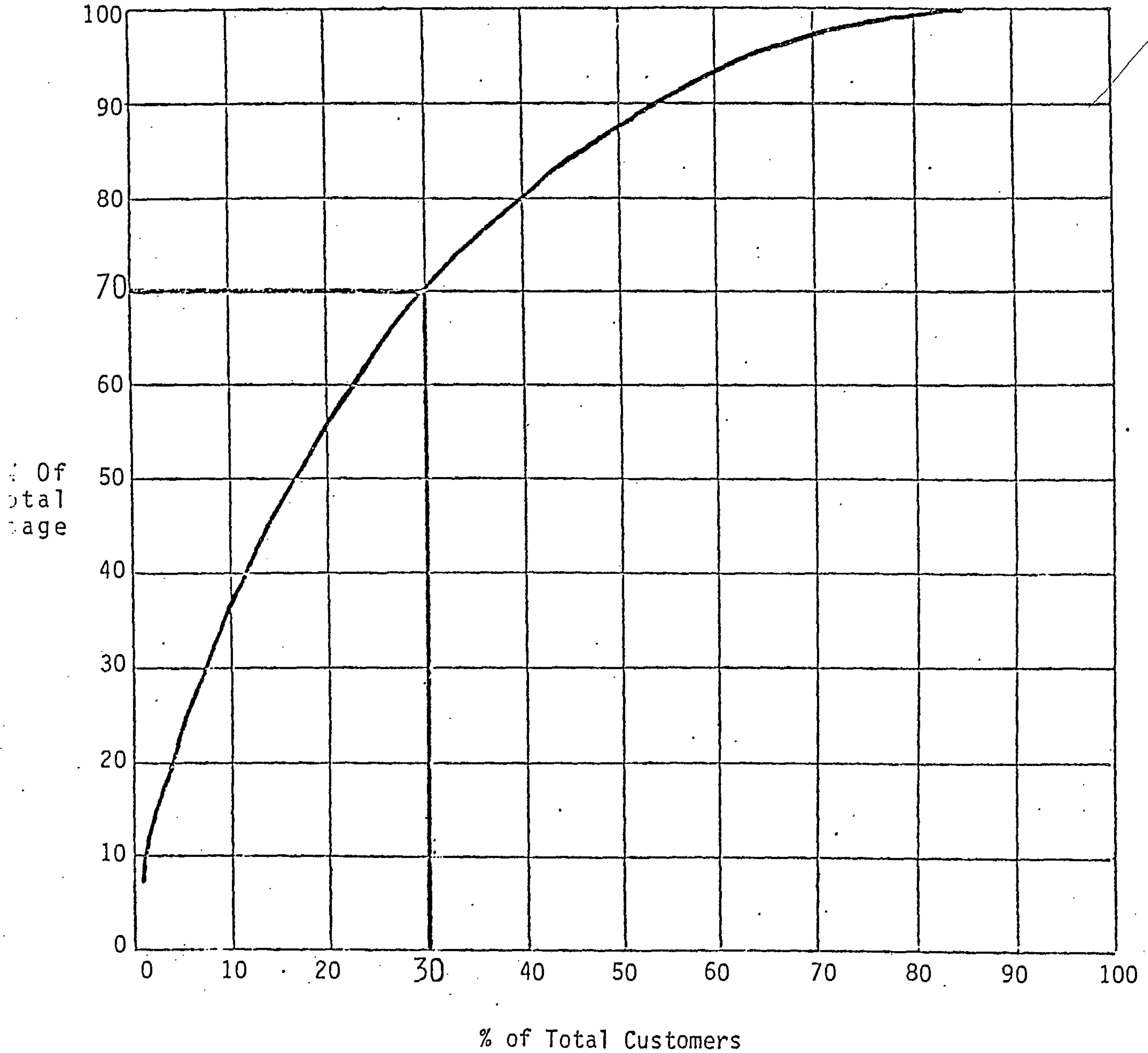
The Canadian telecommunications industry is, like the U.S. counterpart, experiencing the adjustments brought on by competitors interconnecting with parts of the existing networks. In May 1979, the CRTC granted CNCP rights to provide services which fed into the existing network. CNCP was not permitted to provide WATS (Wide Area Telephone Service) or MTS (Message Toll Service) services. The competitive pressures from CNCP on TCTS will lead to rate structure changes by TCTS and revenue changes. One would conjecture that these competitive pressures would lead toward a system in which prices or charges for each type of service were more closely aligned with costs. TCTS represents an unwieldy organization for not only deciding on rates but also on investment programs for the Canada wide transmission system. Competition as a result of "interconnection" is being experienced by the operating companies at the level of the individual subscriber and rates are being altered to reflect the new environment. "James Thackray, Bell president, said in a press release that the company was forced to seek increased rates because of an interim decision by CRTC last year that allowed customers to purchase their own equipment and attach it to Bell lines. 'When rates for these services must be constrained to remain competitive, then other rates must take up the slack', Mr. Thackray said." [Globe and Mail, February 13, 1981, p. 1].

Parameters of the Telephone Use Pattern

With regard to the relationship between the duration of calls and their frequency, the negative exponential distribution has been observed to provide a good characterization for European traffic before local measured service. For the plausible case of a mean call time of 3 minutes, the negative exponential observed would have 63% of calls completed within 3 minutes, and 28% within 1 minute. At the other extreme, only 1% of calls exceed 14 minutes.

At a less aggregative level, one negative exponential must be decomposed to allow for the fact the residential calls under a scheme of a fixed charge per call have been observed longer than business calls, and evening calls, for all subscribers, have been observed longer than daytime calls. Toll calls have been observed longer than local calls with call time increasing on average with distance. Mitchell [1979, p.9] suggests that the duration-frequency relationship is more correctly characterized as a "mixture of several exponential distributions with different mean durations."

With little documentation, Lichtenwalner [1980, p.26] reports that a minority of households make a majority of the calls. This holds for business and residential users. He presents this interesting Lorenz curve. (Figure 1)



A minority number (20-35%) make the majority of calls. This is true for business and residence. Holding times are generally shorter than we had envisioned, but it shows that flat rate pricing is discriminatory.

Figure 1

Presumably the index of usage is calls made rather than time spent using the telephone. He goes on to remark that business usage peaks during business hours (between 8:00 a.m. and 5:00 p.m.) and residential usage peaks around 9:00 p.m. and usage trails off rapidly to midnight. Between 12:01 a.m. and 6:15 a.m. less than 1% of calls per day are made.

The AT&T (Garfinkel and Linhart) paper presents some summary descriptions of telephone usage in the USA under a single party flat rate pricing scheme. These descriptions are in accord with others and we present them as Figures 2, 3, and 4. We observe that a minority of subscribers make very many calls per month leading to a noticeable skewness of the distribution in Figure 2. Figure 3 expresses this point in another way: about half the calls are placed by about one quarter of the subscribers. A similar skewness is found in conversation times under a flat rate pricing scheme. There is a long tail reflecting subscribers with long calls. The average customer holding time of 4.5 minutes is about one minute higher than that recorded for European residential calls. Finally we observe the familiar diurnal usage pattern in Figure 5, with residential use peaking in the evening and business use displaying noticeable morning and evening peaks.

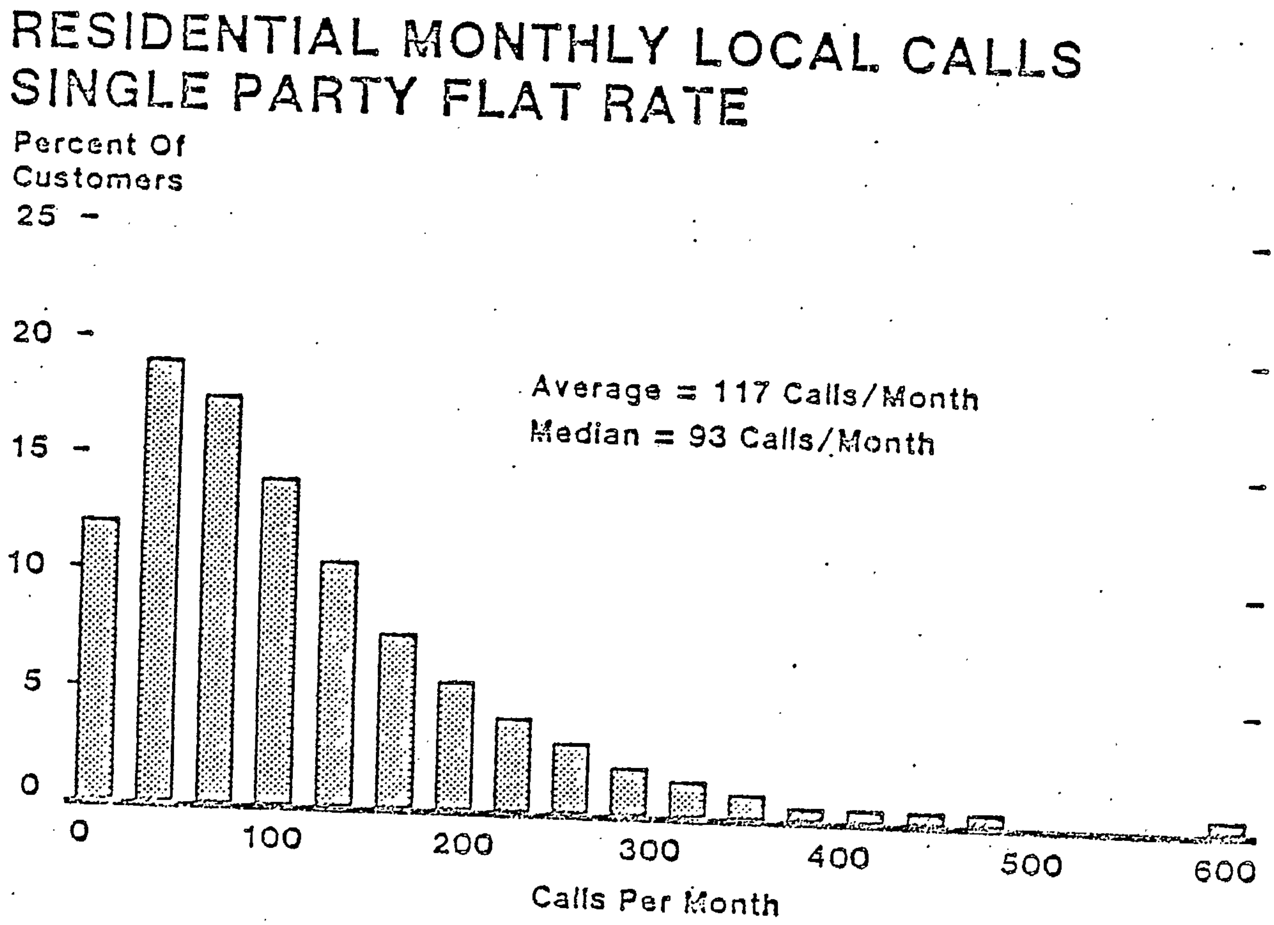

RESIDENTIAL MONTHLY LOCAL CALLS
SINGLE PARTY FLAT RATE
Percent Of Customers
25
20
15
10
5
0
Average = 117 Calls/Month
Median = 93 Calls/Month
0
100
200
300
400
500
600
Calls Per Month

Figure 2

FLAT RATE RESIDENTIAL
% CALLS VS. % CUSTOMERS

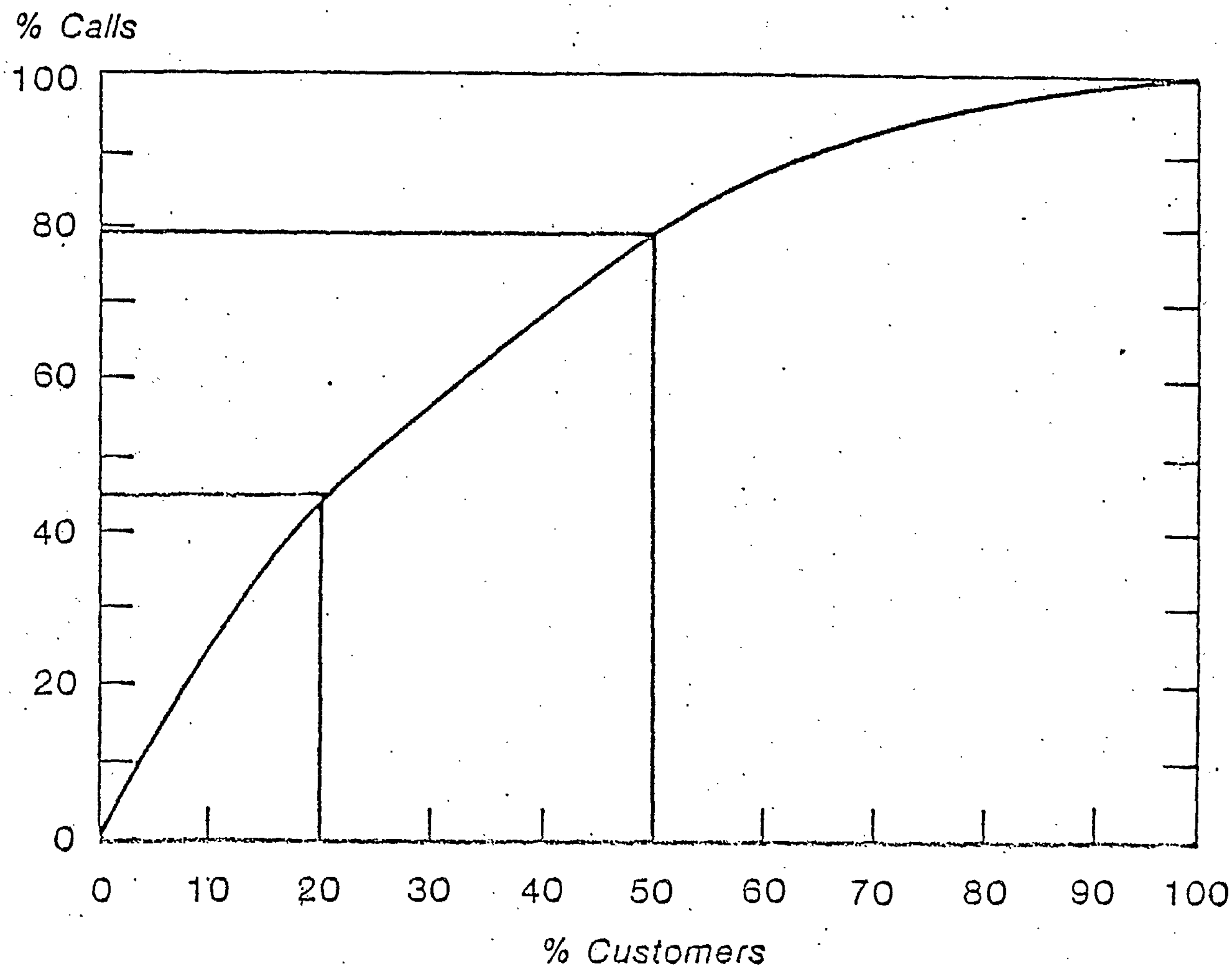


Figure 3

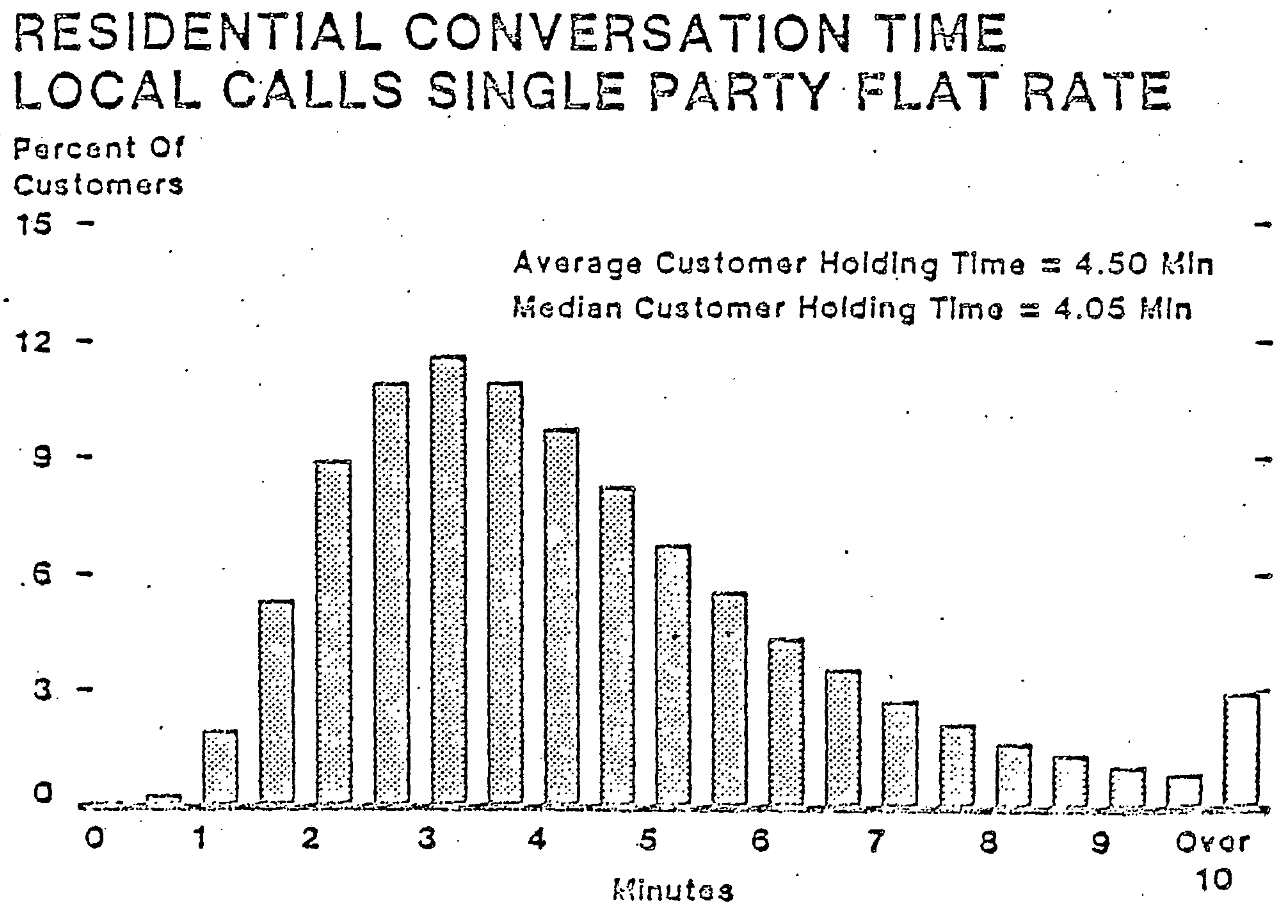

RESIDENTIAL CONVERSATION TIME
LOCAL CALLS SINGLE PARTY FLAT RATE
Percent Of
Customers
15
12
9
6
3
0
Average Customer Holding Time = 4.50 Min
Median Customer Holding Time = 4.05 Min
0
1
2
3
4
5
6
7
8
9
Over
10
Minutes

Figure 4

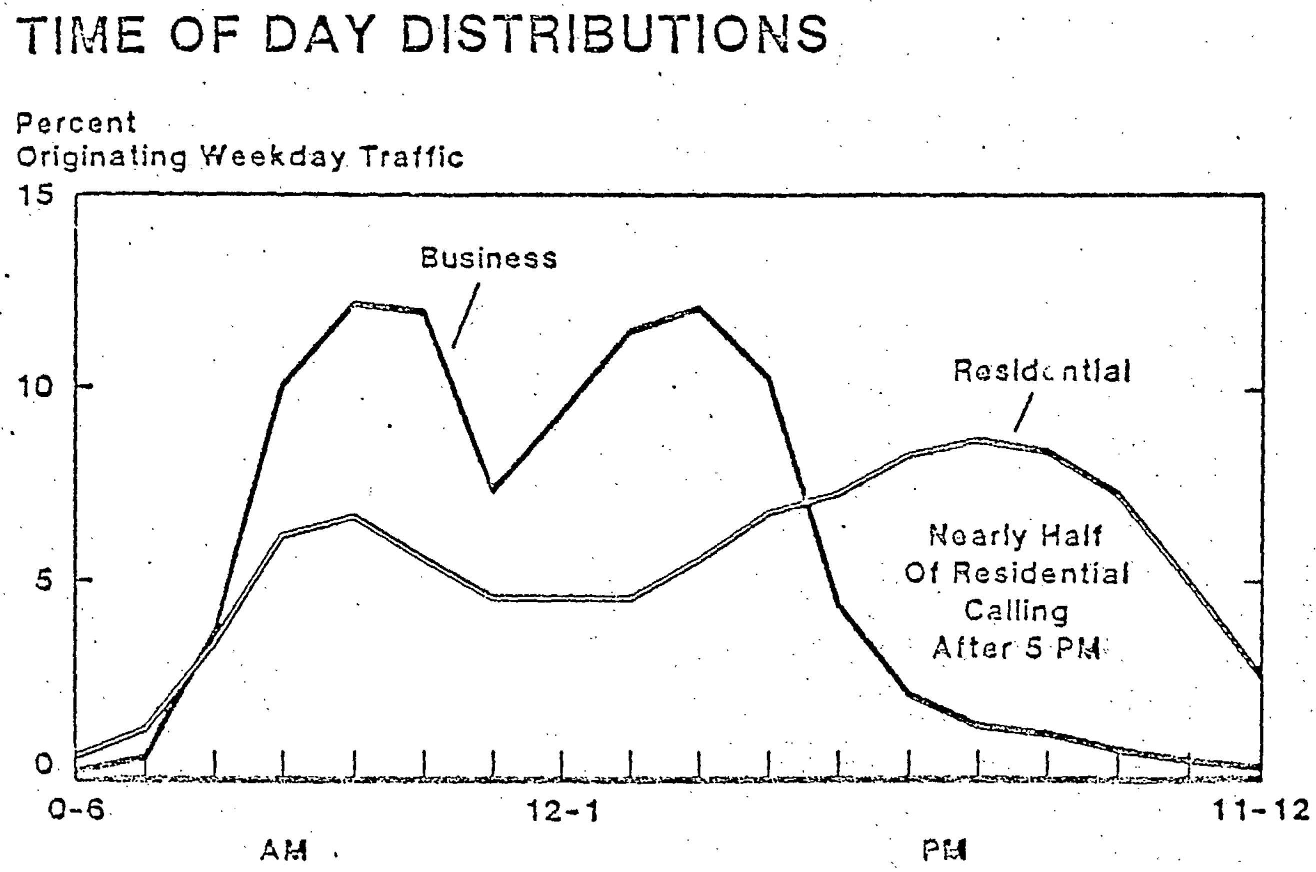

TIME OF DAY DISTRIBUTIONS
Percent
Originating Weekday Traffic
15
10
5
0
Business
Residential
Nearly Half
Of Residential
Calling
After 5 PM
0-6
12-1
11-12
AM
PM

Figure 5

A picture of the relationship between income and telephone usage is provided by these data taken from a survey of ten California exchanges using No. 1 EES switching equipment in May 1972 - July 1973.

Residential Telephone Use by Income Level under a Flat Rate

		Minutes Per Day	
Household Income	Call per Day Per User	Originating (Local Only)	Incoming (Local Plus Toll)
Under $3,000	3.48	6.18	8.24
3,000 - 5,000	2.76	4.72	5.75
5,000 - 8,000	1.45	4.91	6.30
8,000 - 10,000	1.60	5.06	5.41
10,000 - 15,000	1.42	5.28	5.64
15,000 - 20,000	1.52	4.84	4.72
20,000 - 30,000	0.97	4.30	4.85
Over 30,000	1.22	3.92	4.45

[Source: Mitchell 1978 p.521]

Observe that lower income households are relatively higher users. A Beckerite might explain such behavior in terms of the lower opportunity cost of time for poorer people. Moreover if telephone conversation has a consumption component associated with it, chatting may be a low cost leisure activity for lower income people. But lower income households are more frequently non-subscribers than higher income households. For example in British Columbia in 1978 the percentage of all households without telephone service was 3.7% while in the lowest income bracket it was 7.9%. (Statistics Canada, Household Facilities by Income and Other Characteristics, 1978, pp. 52-53.) Also in the U.S., more higher income households were subscribers than lower income ones. [See Mitchell 1978 Table 1]

B. Brandon, et al. [1981] reported on investigations of telephone usage in Chicago in the mid 1970's. Fairly comprehensive statistical tests were performed on information obtained from a representative sample of users. Chicago was covered by a partially metered system. Each call was recorded and charged for beyond some threshold level of calls. Charges were higher for calls from central Chicago to the more distant suburbs. The pricing scheme was based on the message unit per unit of usage. There was a charge of one message unit per local call. One general finding was that there was much variation in usage among subscribers with identical demographic characteristics. One might infer that some salient socio-economic dimension was absent or perhaps more usefully that conventional demographic variables are only partly successful in explaining a household's particular usage.

Specifically it was observed that "as income rises, the median number of local calls tends to rise, although no pair of income groups is significantly different. A higher income is associated with low average duration." (p.6) Income and the level of suburban calling are positively correlated. Blacks call more frequently and talk longer than whites even when other dimensions are standardized. People above 55 make fewer local calls and talk for shorter periods than younger people. Usage is positively related to the number of people in a household and the presence of teenagers is particularly significant in predicting heavy telephone usage. As AT&T papers note, subscribers when asked tend to overestimate their actual usage, both in terms of the number of calls and the aggregate time spent conversing.

Usage for local calling peaks in the evening between 7 and 9 p.m. when an average of 2 minutes per hour is taken up calling. Most local calls were placed within 5 miles of the caller's central office. Duration rises up to the 5-10 mile band and then declines. Distance was not related to other socio-economic variables. The number, average time and aggregate charges for toll calls are significantly higher for households with incomes of $20,000 or more. Finally, aggregate charges also increase with income; that is there is a positive income elasticity for "vertical services" such as "Touch-tone" and "Trimline" handsets.

The rate schedule in effect when these observations were arrived at involved the subscriber selecting his usage class. Each class had a certain number of "free" calls or message units, and beyond a cut-off point a charge was made per additional message unit. One class had a flat rate of $24.50 in 1974. The other classes had an allowance of zero, 80, 140, or 200 message units. The charge for message units above the allowed level was usually 5 3/4 cents. About 5% of subscribers chose the flat rate scheme.

Charging Formulae in Different Countries

A snapshot of some dimensions of the organization, metering and usage of telephones is provided in Table 1. Note that both Canada and the U.S. have most phones operated by private institutions. In Europe, Denmark, Finland, Italy and Spain have most phones operated by private institutions. Both Australia and New Zealand have public operation of the phone system. The remaining countries listed have the phone system publicly operated. These countries are all in Europe. The top four countries classified by phones per capita are in order: the U.S., Sweden, Switzerland, and Canada (two countries with phones predominantly privately organized and two publicly organized). The bottom four countries are Poland, Spain, France and West Germany (Spain's phones are privately organized and the rest are publicly organized). Most European countries have local measured service based on the pulse method of metering. Italy and Spain (privately run) and Poland (publicly run) were not using the pulse method for local metering. New Zealand (publicly run) and Canada and the U.S. (privately run) were also not using the pulse method to monitor local calls. We observe from other sources that the Canadian and New Zealand residential systems were on a flat rate pricing system for local calls. In the U.S. some areas had flat rates for local calling and others were charged by the call but not until recently by duration and time of day.

On the campanion table, Table 2, we have some additional detail on the charges per local call and on access and usage fees. We observe that only Canada, New Zealand, and New York State had flat rates for local calling by residential subscribers. Metering by call alone is common (7 of 17 countries) and metering by call and duration is carried out in 6 countries. We have noted elsewhere that most European countries are moving or have moved to pricing by duration of call. Thus our table fails to reflect the trend toward more complete metering of local calls. In Canada and New York State the "Service Connection Fee" was considerably lower than in other areas. This reflects a posture of cross-subsidizing

January 1976

TABLE 1

	Total Telephones	per 100 pop.	% automatic	Operation: Private	Operation: Gov't.	Conversations → in 000's: Local: Pulse Metered	Local: Other	Interurban: Pulse Metered	Interurban: Other	Internat. Outgoing: Pulse Metered	Internat. Outgoing: Other	
Canada	13,142,235	57.15	99.9	10,797,799	2,344,436	-	17,829,249	-	845,172	-	64,994	
U.S.	148,444,000	69.49	99.9	148,247,000	197,000	-	193,116,000	-	11,529,000			includes international metered pulse units
Denmark	2,316,208	44.97	99.9	2,039,822	276,386	1,630,733	-	712,342	-	-	14,479	pulse units for intern. included
Finland	1,833,993	38.89	94.8	1,244,022	589,971	2,722,445	n.a.	4,478,136	231,366	18,592 (minutes)	1,610	
France	13,833,346	26.2	96.5	0	13,833,346	29,268,100	total pulse - interurban & international included					
Germany, Fed.Rep.of	19,602,606	31.7	100.0	0	19,602,606	9,214,695	-	4,725,133	5,015	115,014 (calls)	3,329	
Italy	14,495,677	25.88	100.0	14,495,677	0	-	7,571,571	-	2,360,116	-	23,975	
Netherlands	5,047,117	36.75	100.0	0	5,047,117	1,900,989	-	1,498,132	40	46,397	1,883	
Norway	1,406,995	35.03	90.0	0	1,406,995	721,000	-	122,000	43,130	6,139	1,833	
Poland	2,577,636	7.54	90.8	0	2,577,636	-	474,092	-	177,652	748	838	
Spain	7,835,970	21.98	92.6	7,835,970	0	n.a.	-	896,341	164,427	11,298	4,459	
Sweden	5,422,795	66.07	100.0	0	5,422,795	17,600,000	-	-	4,700	25,339	1,766	
Switzerland	3,912,971	61.09	100.0	0	3,912,971	1,079,965	-	3,657,483	-	274,638	-	chargeable minutes
U. Kingdom	21,035,602	35.51	99.9	0	21,035,602	13,736,000	-	2,141,180	171,237	-	43,500	
Australia	5,266,843	39.01	96.3	0	5,266,843	3,560,000	-	258,470	86,525	-	2,225	
New Zealand	1,570,784	50.18	93.9	0	1,570,784	-	n.a.	-	79,883	-	802	

SOURCE: AT&T Long Lines, January 1, 1976.

TABLE 2

COMPARISION OF BASIC TELEPHONE CHARGES IN AUSTRALIA AND SOME OVERSEAS COUNTRIES-

(ALL CHARGES EXPRESSED IN AUSTRALIAN DOLLARS)

(1979 Exchange Rate $1.20 Cdn = $1 Aus.)

COUNTRY - CHARGING SYSTEM	ANNUAL RENTAL - EXCLUSIVE SERVICE	SERVICE CONNECTION FEE	LOCAL CALL FEES
AUSTRALIA - measured rate -	Bus: $120 Non Bus: $ 85	$120	9 cents (untimed)
AUSTRIA - measured rate -	Bus and Residential $110.71	Actual installed cost normally amounts to about $52.72	2.7 cents (untimed)
DENMARK - measured rate -	Bus and Residential $91.97	$198.83	1. Monday to Saturday from 8Am to 6Pm 4.3 cents for 3 min call. 2. All other times 2.3 cents per 3 min call.
SWEDEN - measured rate -	Bus and Residential $43.47	$61.50	3.5 cents (untimed)
NORWAY - measured rate -	Bus and Residential varies from $52.74 to $120.24 depending on the number of subscribers in the network.	$175.80 for installation plus a deposit of $351.60 which earns interest at 6.5% PA. This deposit is refunded when the service is cancelled.	11.8 cents per 3 minute call between 8Am and 5Pm on week days. All other times 11.8 cents per call - untimed.

SOURCE: Tariffs and Prices Section
Telecom Australia

COUNTRY – CHARGING SYSTEM	ANNUAL RENTAL – EXCLUSIVE SERVICE	SERVICE CONNECTION FEE	LOCAL CALL FEES
FINLAND – measured rate –	Bus and Residential $56.85	$180.48	6 cents (untimed)
FRANCE – measured rate –	Bus and Residential-Manual exchanges varies from $47.60 to $70.15 Automatic exchanges varies from $87.69 to $100.22. In Paris the rental is $117.76	$146.16 $83.52 paid with first bill and five instalments of $12.52 at 2-monthly intervals	9.8 cents (untimed)
FEDERAL REPUBLIC OF GERMANY – measured rate –	Bus and Residential $156.23	$96.44	The timing of local calls is being progressively introduced. It is expected that by January 1980, 50% of all subscribers will have their local calls timed. *See Details Below
NETHERLANDS – measured rate –	Bus and Residential $123.42	$93.91 plus $1.56 per metre of line installed or moved on subs premises	7.1 cents (untimed)
ITALY – measured rate –	Bus: $83.05 Res: $33.55	Bus: $138.32 Res: $110.82	5.5 cents (untimed)

*NOTE: Federal Republic of Germany – Local call fees.

1. Untimed calls cost 11 cents.
2. Timed calls.

(a) Monday – Friday – 6AM to 6PM
Saturday 6AM to 2PM
11 cents per 8 minute call.

(b) All other times

11 cents per 12 minute call.

COUNTRY - CHARGING SYSTEM	ANNUAL RENTAL - EXCLUSIVE SERVICE	SERVICE CONNECTION FEE	LOCAL CALL FEES
SWITZERLAND - measured rate -	Bus and Residential varies from $83.42 to $121.93 depending on the number of subs in network.	Approx charge for internal wiring is $80.22	3.2 cents per call of 3 minutes. Calls are timed
NEW ZEALAND - Flat rate -	Bus: varies from $196.27 to $307.69 Res: varies from $84.84 to $113.12. The variation is dependent on the number of subs and type of exchange to which the sub is connected ie Auto, manual or restricted service exchange	$47.13	Unlimited free local calls
REPUBLIC OF SOUTH AFRICA - measured rate -	1. Flat rate exchanges Bus. $57.41 Res. $44.65 2. Message rate exchanges (a) Automatic exchanges and large City and suburban areas. Bus and Res: $44.65 (b) Country Exchanges Bus & Res $31.69	$31.89	Unlimited free local calls for subs connected to flat rate exchanges: 4.2 cents per call (untimed) for subs connected to messages rate exchanges

COUNTRY - CHARGING SYSTEM	ANNUAL RENTAL - EXCLUSIVE SERVICE	SERVICE CONNECTION FEE	LOCAL CALL FEES
JAPAN - measured rate -	Business: $137.28 Residence: $95.04 Rentals indicated apply in the large cities of Japan with 400,000 or more subscriber lines	Connection fee $352.00 Compulsory Purchase of Bonds <u>$660.00</u> <u>$1012.00</u>	4.4 cents per 3 minutes
UK - measured rate - (Includes 8% VA.T)	Bus: $76.26 Res: $64.53	$88 if within 5km of an exchange. Beyond 5 km excess charge of $29.33 for each 200 meters or part thereof	Calls are timed 1. <u>Peak Rate</u> 5.8 cents per call of 2 minutes. 2. <u>Standard Rate</u> 5.8 cents per call of 3 minutes. 3. <u>Cheap Rate</u> 5.8 cents per call of 12 minutes. <u>NOTES</u>: 1. Peak rate - Monday to Friday 9Am to 1Pm. 2. Standard rate - Monday to Friday 8Am to 9Am 1Pm to 6Pm 3. Cheap rate - Monday to Friday 6Pm to 8Am and on Saturday and Sunday all day.

COUNTRY - CHARGING SYSTEM	ANNUAL RENTAL - EXCLUSIVE SERVICE	SERVICE CONNECTION FEE	LOCAL CALL FEES
CANADA (MONTREAL AND TORONTO) - A mix of flat and measured rate - Bell Telephone Company of Canada	Flat Rate basis Bus: $247.43 Res: $ 75.35 Measured rate basis Bus: $147.46	Bus: $31.66 Res: $15.26	Unlimited free calls to subs paying rental on flat rate basis. 125 free calls per month for subs paying rental on measured rate basis, and 4.0 cents per call for calls in excess of the free allowance
U.S.A. NEW YORK STATE - a mix of flat and measured rate - American Telephone and Telegraph Co.	Bus Subscriber 1. Flat rate basis - varies from $135.66 to $345.26 2. Message rate basis - varies from $120.90 to $134.36. These rentals include a monthly allowance of $3.56 for local calls Res Subscriber 1. Flat rate basis - varies from $78.21 to $135,66 2. Message rate basis (a) Untimed message units varies from $75.78 to $82.49 (b) Timed message units varies from $65.08 to $71.79 These rentals include a monthly allowance of $3.56 for local calls. NOTE: Flat rate service is not offered to Business and Residential subscribers in New York City.	Business $48.14 Residential $21.84	Subs on a flat rate basis of charging are entitled to unlimited and untimed free message units in his primary local call area. Message rate Basis Business Subs: Mon to Fri: 8Am to 9.00Pm: First 5 minutes 6.7cents; Each add 1 min 1 cent 9Pm to 11Pm First 5 minutes 4.4cents; Each add 1 min 0.6cents Sat, Sunday and holidays 8Am to 11Pm: First: 5minutes 4.4cents; Each add 1 min 0.6cent All days 11Pm to 8Am First 5 minutes 2.6cents; Each add 1 min 0.18cent Residential Subs-Untimed calls Mon to Fri 8Am to 9Pm - Each message unit 7.3 cents 9Pm to 11Pm " " " 5.3 " Sat, Sunday and holidays 8Am to 11Pm " " " 5.3 cents All days 11Pm to 8Am " " " 3.5 cents Timed calls Mon to Fri 8Am to 9Pm: First 5 minutes 6.7 cents; Each add min 1 cent 9Pm to 11Pm First 5 minutes 4.4 cents; Each add 1 min 0.6 cents Sat Sunday and holidays 8Am to 11Pm First 5 minutes 5.3 cents; Each add 1 min 0.6 cents All days 11Pm to 8Am First 5 minutes 2.6 cents; Each add. 1 min 0.18 cents

TABLE 3

COMPARISION OF TRUNK CALL CHARGES - EQUATED TO CHARGEABLE DISTANCES

IN AUSTRALIA

ALL COSTS EXPRESSED IN AUSTRALIAN DOLLARS

CHARGEABLE DISTANCE	COST OF 3 MINUTE DDD CALL - FROM MONDAY TO FRIDAY - AT PEAK DAY RATES ..										
	AUSTRALIA	NEW ZEALAND	CANADA	U.S.A.	U.K.	SOUTH AFRICA	JAPAN	NORWAY	SWEDEN	FRANCE	FEDERAL REP OF GERMANY
Up to 50km	0.18	0.23	0.48	0.58	0:34	0.09	0.37	0.59	0.14	0.39	0.67
51 to 85km	0.36	0.51	0.62	0.77	1.05	0.21	0.61	1.06	0.26	0.74	1.33
86 to 165km	0.81	.74	0.82	0.88	1.05	0.36	0.99	1.06	0.42	0.74	1.66
166 to 325km	1.35	1.24	0.98	0.91	1.05	0.64	1.58	1.06	0.52	1.47	1.66
326 to 485km	1.80	1.53	1.11	0.96	1.05	0.64	1.98	1.06	0.63	1.47	1.66
486 to 645km	2.025	1.87	1.16	1.01	-	0.85	2.64	-	-	-	1.66
Over 645km	2.70	-	1.20	1.01	-	1.09	3.17	-	-	-	-
Maximium			2.23	1.16		1.53					

CHARGEABLE DISTANCE	COST OF 3 MINUTE DDD CALL - FROM MONDAY TO FRIDAY - AT CHEAPEST NIGHT RATES										
	AUSTRALIA	NEW ZEALAND	CANADA	U.S.A.	U.K.	SOUTH AFRICA	JAPAN	NORWAY	SWEDEN	FRANCE	FEDERAL REP OF GERMANY
Up to 50km	0.072	0.22	0.16	0.23	0.05	0.04	0.37	0.44	0.14	0.19	0.29
51 to 85km	0.144	0.51	0.20	0.31	0.16	0.10	0.37	0.79	0.26	0.36	0.29
86 to 165km	0.324	0.59	0.27	0.35	0.16	0.18	0.53	0.79	0.42	0.36	0.29
166 to 325km	0.54	0.96	0.32	0.36	0.16	0.32	0.88	0.79	0.52	0.73	0.29
326 to 485km	0.72	1.19	0.37	0.37	0.16	0.32	1.13	0.79	0.63	0.73	0.29
486 to 645km	0.81	1.53	0.38	0.40	-	0.42	1.58	-	-	-	0.29
Over 645km	1.08	-	0.65	0.40	-	0.54	1.98	-	-	-	-
Maximium			0.74	0.46		0.76					

TABLE 4

AN INDICATION OF THE SIZE OF LOCAL CALL AREAS IN CERTAIN OVERSEAS COUNTRIES

COUNTRY	APPROXIMATE SIZE OF LOCAL CALL AREA MEASURED RADIALLY FROM MAIN EXCHANGE
Australia	Melbourne and Sydney - 40km radially from principal trunk exchange. Other capital cities 32km radially.
Austria	The local call area comprises a community with a typical radius of 5km.
France	Paris - extends 14km radially
Federal Republic of Germany	Extends approximately 20km radially in almost all local areas throughout the Federal Republic.
Canada	Montreal - extends up to 48km radially
	Quebec City - extends up to 40km radially
Singapore	Local call area extends over the whole of the Island of Singapore
Japan	Tokyo - 19km radially
Sweden	Stockholm - 36km radially
UK	London - 25km radially
USA	New York - could vary from 8km to 30km radially
	Chicago - 14km radially
	Los Angeles - 13km radially
New Zealand	Auckland and Wellington - 32km radially
Switzerland	Berne and Geneva - 7km radially
	Zurich - 6km radially
Republic of South Africa	Pretoria - 25km radially

BAR CHART PRESENTATION OF APPROXIMATE SIZE OF LOCALCALL AREAS IN AUSTRALIA/SOME OVERSEAS COUNTRIES/SOME CAPITAL CITIES - EXPRESSED IN TERMS OF RADIAL DISTANCE FROM PRINCIPAL TRUNK EXCHANGES

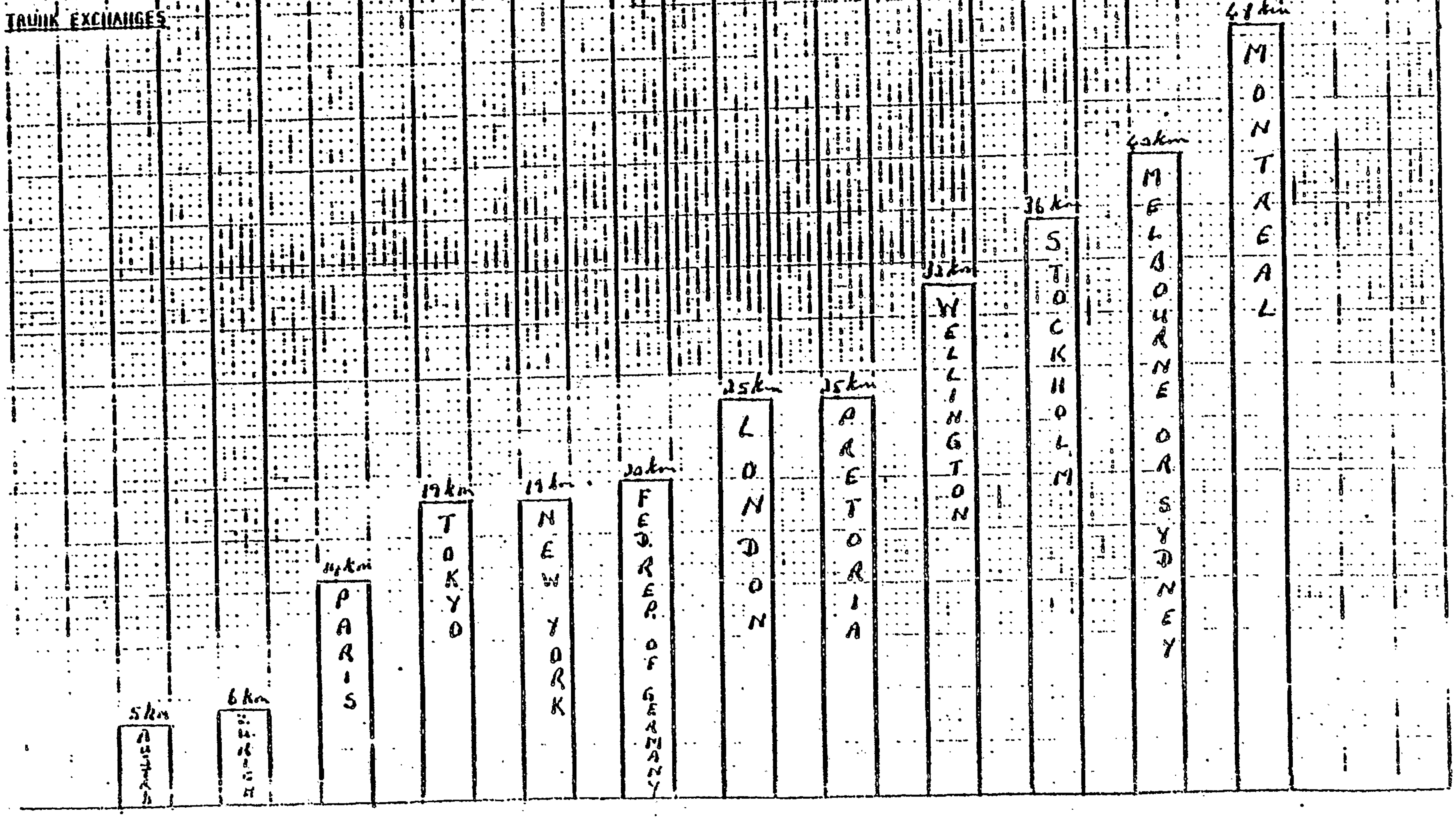


TABLE 5

HOURS OF WORK REQUIRED EACH YEAR TO PAY FOR TELEPHONE SERVICE IN AUSTRALIA AND SOME OVERSEAS COUNTRIES

	AUSTRALIA	NEW ZEALAND	JAPAN	FRANCE	FEDERAL REP OF GERMANY	DENMARK	NORWAY	SWEDEN	UK	CANADA	USA
Hours of work required each year to pay service connection fees, spread over 10 years.	2.51	1.49	*18.85	5.26	1.66	2.54	Ø8.62	1.11	2.89	0.31	0.42
Precentage of Total hours	5%	2%	32%	6%	2%	12%	16%	6%	4%	1%	1%
Hours of work required to pay annual residential rental.	17.75	35.68	17.7	36.05	26.89	11.76	19.65	7.83	21.16	15.16	15.74
Percentage of total hours	39%	60%	30%	42%	40%	54%	36%	40%	27%	47%	48%
Hours of work required to pay for 500 local calls - each of 3 minutes duration - at Peak day rates.	9.39	-	4.10	17.63	9.47	2.75	9.64	3.15	19.23	-	-
Percentage of Total hours	20%	-	7%	21%	14%	13%	17%	16%	25%	-	-
Hours of work required to pay for 100 STD calls - each of 3 minutes duration - between points 150km apart - at Peak day Rates.	16.91	23.2	18.44	26.47	28.57	4.62	17.38	7.54	34.62	16.54	16.85
Percentage of total hours	36%	38%	31%	31%	44%	21%	31%	38%	44%	52%	51%
TOTAL HOURS	46.56	60.37	59.09	85.41	66.59	21.67	55.29	19.63	77.90	32.01	33.01

*JAPAN - Includes 12.29 hours for the compulsory purchase of bonds.
ØNORWAY - Includes 5.75 hours for the compulsory purchase of bonds.

SOURCE: Tariffs and Prices Section,
Telecom Australia

TABLE 6

AVERAGE HOURLY EARNINGS OF ADULTS ENGAGED IN

MANUFACTURING INDUSTRY

COUNTRY		CATEGORY OF ADULTS	AVERAGE HOURLY EARNINGS IN AUSTRALIAN $
			$
FRANCE	(Dec 1977)	Male & Female	2.78
UK	(Oct 1977)	Male	3.05
NEW ZEALAND	(Oct 1977)	Male & Female	3.17
AUSTRALIA	(Oct 1977)	Male	4.79
CANADA	(Dec 1977)	Male & Female	4.97
U.S.A.	(Dec 1977)	Male & Female	5.24
JAPAN	(1977)	Male & Female	5.37
SWEDEN	(Dec 1977)	Male & Female	5.55
FED REPUBLIC OF GERMANY	(Oct 1977)	Male	5.81
NORWAY	(Dec 1977)	Male	6.12
DENMARK	(Dec 1977)	Male & Female	7.82

NOTE: Hourly earnings indicated are prior to payment of tax.

SOURCE: Monthly Bulletin of Statistics - December 1978, Issued by Department of International Economics and Social Affairs, Statistical Office, New York.

access in North America, I believe. This has resulted in a high rate of penetration of phone rentals in the total market. Recall that Canada and the U.S. did indeed rank in the top three countries of phones per capita. Sweden has a relatively low annual rental fee but moderate "access fees" (Sweden also charges per call) and also displays a high level of phones per capita.

The other salient and implicit dimension of local charging is the area over which a subscriber can call without incurring toll charges. We observe that the cities with the largest areas are Montreal, Melbourne or Sydney, Stockholm, and Wellington, New Zealand. Those with the smallest areas are a city in Austria, Zurich, Paris, and Tokyo. (Distance allowable under a local calling charge is one of those dimensions of implicit pricing which makes quick comparison of relative prices among cities and countries particularly difficult. Ideally, one would like to standardize across access areas. Should one use in this case geographic area or number of subscribers as the criterion of comparable access?)

Outlays By Subscribers in Different Countries

Since telephone service comprises a bundle of district subservices (local calling, toll calling, peak and off-peak calling) one has to aggregate over the subservices in order to arrive at a representative outlay. We report on two recent surveys of evaluating relative subscriber outlays across countries. Recall that outlay is a disbursement by the subscriber and may or may not be closely aligned with resource costs of providing the service. In order to avoid using international exchange rates to compare outlays, some analysts express outlays in hours worked at a representative wage rate in the country in question.

In Table 5 we have relative hours worked in different countries for a standardized basket of telephone services, for a residential subscriber. Note that the access charge has been amortized over ten years so that it looms relatively small in the calculation. The basket of subservices comprises 500 local calls of 3 minutes or less per year and 100 direct dial toll calls of 3 minutes or less between two points 150 km apart. The four countries with the lowest outlay per residential subscriber measured by hours employed required to pay for the service are Sweden, Denmark, Canada and the U.S. The four countries with the highest outlays are France, the U.K., West Germany and New Zealand. In another study by a private research organization in Europe, the outlay for a standard basket of telephone services was compared across 13 European countries. The basket was based on a representative user in the U.K. Calls were evaluated as if they were 3 minutes long and the average distance of a trunk call was 100 km. Three international calls were included in the basket. The basket comprised 649 local calls, 112 trunk calls and 3 international calls. (A composite of business and residential users was used. There was also averaging over peak and off-peak prices.)

TABLE 7

Index of Relative Outlays for a

Standard Basket of Telephone Services 1978

Sweden	42.1
Denmark	68.8
Ireland	98.7
Switzerland	100.0
Finland	100.1
Italy	102.0
France	103.6
U.K.	105.9
Norway	126.9
Belgium	135.6
Netherlands	143.2
Austria	155.5
West Germany	183.8

SOURCE: Telephony, July 10, 1978, p.76.

These results are not dissimilar to those above. Sweden and Denmark are countries with low subscriber outlays and West Germany, France and the U.K. are countries displaying higher outlays per subscriber. These index numbers are not related to wages and thus are constructed quite differently than the above measures of relative subscriber outlays.

TABLE 8

Degree of Cost Coverage by Service, in Percent

Service	Switzerland 1976	Sweden 1973/74	United Kingdom 1976	West Germany 1974
Telephone Services				113
Rentals	85-90	47		NA
business	- -	- -	85	- -
residence	- -	- -	87	- -
Calls			128	
local	ca. 60	58		NA
trunk	> 100	245		NA
Private Lines	206	NA	108	NA
Telegram	67	- -	43	52
Telex	119	NA	84	120
Total Telecommunications	108	- -	104	112

Sources: Switzerland: PTT - Geschäftsbericht 1976
Sweden: Televerket-Författningssamling Serie A:29
United Kingdom: Post Office Reports and Accounts, 1976-77
West Germany: Deutsche Bundespost, Geschäftsbericht, 1975

NA = not available.

A more detailed look at telephone rates for five European countries (Norway, Sweden, Switzerland, the U.K. and West Germany) was reported in Mitchell [1978b]. Charging formalae involved an installation fee, a monthly subscription fee, local usage charges (often metered by duration) and rates for trunk calls. In each of these countries a public agency ran the system. Generally, toll revenues were used to offset deficits arising from the provision of local service (see Table 8). Of note in the survey is the discussion of how adjustments in peak, and off-peak relative rates affected the diurnal and weekly patterns of toll telephone use. There is evidence of significant cross-elasticities of demand between periods with respect to price.

A similar resonse was reported for long distance traffic in the U.S. in the 1960's. In 1963 AT&T introduced off-peak rates starting at 9 p.m. and the number of calls between 9 and 10 p.m. doubled. In 1965 the off-peak rates were advanced to the post 8 p.m. period. A leveling off of use appeared immediately but a gradual increase in "bunching" of calls in the 8 to 9 p.m. hour emerged later. Another observation is that residential and business charges are very similar in these European countries in contrast with the situation in Canada and the U.S. where a multiplicity of rate options are available to different types of users (see our section on Canadian pricing experience).

Canadian Pricing Experience

Beigie [1973] presented a thoughtful survey of pricing of telecommunications services in Canada. (The structure of rates he examined did not change by 1981 although inflation resulted in a shift up in nominal charges.) His benchmark of desirable pricing was marginal cost pricing. He saw departures in four spheres and expressed dissatisfaction with these violations of the rule: price at marginal cost. First he was surprised at some forms of bulk toll calling discounts. For example under a particular SPPL (short period private line) arrangement, a user can make relatively short distance toll calls in moderate volume much cheaper than under the usual toll pricing schedule. These discounts appeared unrelated to a traffic volume or pattern which would suggest that the telephone company was following some variant of marginal cost pricing.

Secondly he expressed dissatisfaction at the "averaging" implicit in extended area service (EAS). That is one subscriber who makes calls through only a single central office is charged the same as another subscriber whose calls go through more than one central office and over the connecting trunk. The latter subscriber makes more "demands" on capital and is implicitly subsidized by the former since both pay the same monthly charge. Recall that EAS, a pervasive approach to pricing local service, results in a flat rate subscription applying to calls covering relatively wide geographic areas and encompassing more than one central office. EAS is usually implemented after opinion has been sampled in the area covered. There is both one-way and two-way EAS. In the former, a subscriber can call "toll-free" from a fringe area of a city into a city but city residents must pay a toll in calling from the city to the fringe. The vagueness of criteria for introducing EAS make regulatory decision difficult as was well illustrated in the CRTC Telecom Decision 81-3 concerning the British Columbia Telephone

Company. (A hybrid form of service combining elements of toll and EAS is the Residence Optional Calling Plan (ROCP) which allows for reduced rates (discounts of about 60%) for blocks of long distance calls between exchanges located up to 40 miles apart.) The third departure from marginal cost pricing he noted was the charging of different rates for apparently the same service. For example business users were charged higher installation and flat rates for the same rights to use the local network. Finally, he expressed concern over the prohibition on "interconnection" or the use of devices manufactured by "outside" firms on lines owned by a company with the right to offer service. Since 1973, there have been rulings by CRTC permitting some "interconnection", particularly that of CNCP in the sphere of non-voice transmission by business subscribers (Ogle [1979; p.235]).

Beigie is favorably disposed to detailed measured service. His model of desirable pricing is that for teletype. Charges are by 6 second units. Distances are calculated on the basis of WATS zones. There is however no variation in charge for time of day or week. (There is also the anomaly -- charges are based on one (1) minute intervals between Canada and the U.S.) Beigie relates these teletypewriter rates to the then new toll rates for Bell Canada which involved one charge for the first minute and another charge for subsequent minutes. This rate scheme was a departure from the earlier specific charge for the first three minutes. Considerable opposition emerged from users when the one minute scheme was introduced. A pricing scheme which appears clearly more efficient may not gain popular acceptance without a campaign of information dissemination. Some critics of the new set charge for the first minute contended that this pricing scheme was designed to capture a large new revenue source in non-voice communication over the switched network. Presumably data and related non-voice traffic can be transmitted in digitalized bursts often of less than one minute whereas voice calls have a mean of between three and five minutes. One wonders if huge increases in non-voice transmission over the switched network might not lead to substantial increases in the cost of a voice call if capacity does not grow as fast as traffic. Beigie suggests that one might conceive of capacity being approached on the system during the nighttime hours if business non-voice communication is directed to that time period given the traditional low prices for nighttime use. There is room for pricing which would permit fixed capital to be used at an efficient capacity for seven days a week and 24 hours per day. A voice call may become for households a rare and luxurious item. Much existing residential toll calling might take place in prepackaged non-voice modes. The interactive component of voice calls would be sacrificed but much lower relative costs to the customer would be obtained. The subscriber would presumably pick up his handset, dictate a message, dial the number of the party he wishes to send the message to and hang up. The receiving party would have a unit to reconstruct the message from the electronic signal transmitted. The charge for the use of the switched network would presumably be miniscule since the time taken to transmit the electronically prepackaged message would be miniscule.

Beigie is equivocal about reduced rates for bulk users. Such schemes as WATS (wide area telephone service) and TELPAK are not explained from the standpoint of marginal cost pricing. WATS involves essentially a leased link from an establishment to a particular geographic zone. The price is fixed for the link and not by the call. There are alternate pricing formulas under which a subscriber can acquire a WATS line for 10 hours of calling per month. For calls running in the block above 10 hours, a charge per unit time is imposed. (There is also the option of acquiring an INWATS line for only incoming calls or an OUTWATS line for only outgoing calls.) WATS relies on the established switched network. TELPAK is a partly dedicated scheme. A subscriber leases a number of channels capable of transmitting voice and non-voice traffic. These channels can be connected to the switched network at point of origin or destination. (The very low rates charged for TELPAK in the U.S. in the late 1960's were, it is alleged, intended to discourage entry into the long distance transmission business by potential rivals such as MCI.)

Dedicated systems are not complicated to price. One calculates the carrying costs of the initial outlay amortized over the life of the installation and adds on a charge for maintenance. However partly dedicated systems present the usual complexities of transmission system pricing. A charge must cover the capital and maintenance for the dedicated part, but then a charge must be made for use of the network, a charge varying with traffic (usually time of day) and distance. If rates are not set at the appropriate marginal cost, then there will exist an incentive for users to acquire their own private dedicated system. This latter appears to be a possibility being considered for communication within large corporations. Since much traffic moves by microwave signals, and the cost of antennas has been declining, the incentive for certain users to bypass the existing switched network has emerged. In other words, competitive suppliers of communication system have grown up at many stages of the telecommunication system. The reason for this burgeoning of competitors is twofold -- there are many <u>new</u> products peripheral to the system (attachments) which can be produced competitively with traditional suppliers such as Western Electric in the U.S. and Northern Electric in Canada. Many of these new products can be tailor-made and can be differentiated from a rival's offering. New niches in the market have opened up which traditional suppliers have failed to fill. Secondly, costs of entry to the sector providing network services have become lower. MCI and CNCP can provide a competitive toll network based on microwave transmission. They do not have to duplicate the hardware of the existing toll system, this latter having developed with lines linking exchanges in different cities.

It seems characteristic of network systems such as roadways, railways, telecommunications, and perhaps airways, that peripheral or low density links are subsidized by high density links within a firm. The introduction of competition appears first in high density links and this in turn forces existing firms to lower prices on these links. Prices are then raised on peripheral links and often service gets cut back or withdrawn. (This did not appear in the early stages of deregulation of airlines in the U.S., however; currently there is significant withdrawal of service in the peripheral areas (New York Times, March 8, 1981).) In conversation, James Alleman of GTE suggested that the charges for connecting relatively isolated subscribers had been set too low relative to costs in the past. In this case there was cross subsidization of capital costs rather than operating costs.

A case can be made for bulk use charges of the WATS or TELPAK sort in terms of minimizing risk on the part of the subscriber and supplier. The subscriber knows his monthly outlay in advance and the supplier knows his revenue. The uncertainty is spread over other users of a non-dedicated system who are not certain that there will be capacity for their calls at all times. Users of WATS lines also incur a cost relative to normal measured toll service in the form of queues which form within a company for access to the WATS line. Thus a WATS installation can dampen the demand for calls by means of queue rationing in place of price rationing. We discuss below that a flat rate subscription for local calling has the same risk spreading feature as does the bulk charge scheme. The supplier may end up providing more capacity than he would with measured service but both subscriber and supplier experience a disbursement and revenue stream respectively with little variation. If costs are relatively insensitive to the volume of traffic, then such a pricing scheme has merit.

The final unusual characteristic of the pricing system which attracted Beigie's attention was the variation in charges with distance. Beigie observed a noticeable <u>decline</u> in <u>charge per mile</u> as the <u>distance</u> a call traversed <u>increased</u> under the standard toll charge scheme. Charge per call is still of course higher for longer distance calls than shorter distance calls. Under TELPAK there was no decline in charge per mile. It seems plausible, in light of the technology involved that toll calls might have the charge per mile decline with the distance over which the call travels. The call travels from the sbscriber to the local central switching office, then it is routed either to the microwave network consisting of transmitting towers about 30 miles apart or to the nearest satellite transmission station. At the receiving end, the call leaves the long distance network, moves through the local exchange and reaches the receiving subscriber. Costs involved are in switching and long distance transmitting. In the microwave system the costs would be lower in moving a signal from tower to tower than in "accessing" the network. The call then uses switching resources and transmission facilities, which appear to clearly have cost per mile declining with distance. A marginal cost pricing schedule would involve the observed decline in charge per mile as distance increases. Also the charge should increase with call duration but again at a declining cost per minute. (A capacity charge should also be included by having rates vary with traffic or as a proxy, by time of day.) For the satellite transmission, the cost per mile declining with distance is even more striking. Total costs per call are presumably invariant over wide ranges of distances making the cost per call relatively constant. Thus a decline in charge per mile with distance seems obviously in order.

Why would TELPAK charges not decline in dollars per mile for increased distances? Since TELPAK is largely a dedicated system, the charge is presumably largely for installed capital per user. It is easy to see how increased distances would involve increased capital outlays and how those outlays would be roughly in proportion to distance traversed. There are then straightforward arguments for having a different pricing scheme related to the distance dimension for regular toll service and for TELPAK service.

Is there a marginal cost pricing rationale for charging residential and business subscribers different flat rates and/or different installation charges for the same right of usage of the system? Certainly if a representative business user on average makes more calls and occupies the line longer than a representative residential user, a clear case can be made for charging business subscribers more. Such is the case, but recovering the respective costs via installation and flat rate charges is a very crude form of marginal cost pricing. In fact it is probably closer to average cost pricing. The case for charging precisely by usage is in accord with marginal cost pricing but nevertheless charging proportional to usage (average cost pricing) is probably a reasonable second-best solution.

To the economist, the charging of different rates for the same service is explained on efficiency grounds in some instances (e.g. peak load pricing) and on revenue generation grounds in other instances (e.g. price discrimination). Operating companies interpret the same pricing patterns in terms of "value of service" (e.g. Skelton [1980]). Value of service pricing is a concept borrowed from transportation economics. There the concept is used to describe a procedure in which the cost of shipment for a specific commodity is decided by a rule of thumb or markup formula which involves charging for the transportation service an amount related to the value of the commodity being shipped. This is a procedure which economizes on the obviously costly process of determining the cost to the company of making the shipment. Rules of thumb do have this desirable efficiency character but the procedure of value of service pricing is nonetheless a crude form of price discrimination in the textbook sense of price discrimination. It is not a procedure in which charges are determined by costs of production of the service. We see then, Bell Canada justifying charging residential subscribers in different communities flat rates varying directly with the number of telephones in the respective areas "covered by" the flat rate and the charging of business users in an area a higher flat rate than residential users because "the value of service to a business customer generally is greater than to a residential customer" (Skelton [1980, p.8]). The monthly cost of providing the service by Bell may be related to (i) the number of phones in an area and (ii) to the type of user but invoking "value of service" as a basis for pricing in these cases is really invoking a mixture of revenue arguments (price discrimination) and efficiency arguments (average use related to average cost of provision). In any case, "value of service" pricing corresponds neither in principle nor in practice to marginal cost pricing or USP more generally. Although message-rate (charge based on number of calls alone) service for business is widely available in Canada, in 1980, only 4.2% of business lines were being charged under a message rate scheme. In Table 9 , we have flat rates for residential and business users for some Canadian and U.S. cities.

TABLE 9

	Flat Rate Individual Business Line per Month		
	1969	1979	Increase
Ottawa, Ontario	$13.70 (Can.)	$23.15 (Can.)	$ 9.45 (Can.)
Birmingham, Alabama	$19.50 (U.S.)	$37.95 (U.S.)	$18.45 (U.S.)
Albany, N.Y.	$17.00 (U.S.)	Flat Rate not available	
Montreal, Quebec	$16.25 (Can.)	$27.35 (Can.)	$11.10 (Can.)
Toronto, Ontario	$16.25 (Can.)	$27.35 (Can.)	$11.10 (Can.)
Atlanta, Georgia	$19.00 (U.S.)	$33.10 (U.S.)	$14.10 (U.S.)
Miami, Florida	$15.75 (U.S.)	$29.25 (U.S.)	$13.50 (U.S.)

	Flat Rate Individual Residence Line per Month		
	1969	1979	Increase
Ottawa, Ontario	$5.30 (Can.)	$ 7.50 (Can.)	$2.20 (Can.)
Birmingham, Alabama	$6.10 (U.S.)	$12.85 (U.S.)	$6.75 (U.S.)
Albany, N.Y.	$6.00 (U.S.)	$12.53 (U.S.)	$6.53 (U.S.)
Montreal/Toronto	$5.85 (Can.)	$ 8.55 (Can.)	$2.70 (Can.)
Atlanta, Georgia	$6.50 (U.S.)	$11.85 (U.S.)	$5.35 (U.S.)
Miami, Florida	$5.95 (U.S.)	$12.10 (U.S.)	$6.15 (U.S.)

(Source: P. E. Skelton, Testimony to Restrictive Trade Practices Commission of Canada, May 1980)

Metering for Local Measured Service[1]

Step-by-step offices still represent more than 60% of switching entities in the U.S., 90 years after their introduction. It is not easy to add features to the switch. However, two approaches are practicable. The company can move to LMS by adding on peripheral devices or integral devices. The different approaches have different costs and advantages. We will note these below.

The peripheral devices are attached to each subscriber's tip-and-ring as in Figure 6. The integral devices are attached between the line finder and first selector, at a point of concentration. The Figure shows an input ratio of ten to one for the peripheral and integral devices and this is a minimum value. Either type can be active--that is not only record call elements but also participate in the switching function in varying degrees--or passive, i.e. record only the call elements.

Peripheral devices can be added on without requiring any office rearrangements to reassign subscribers. They make sense in a transition phase as for example when new electronic switches are expected to be brought on stream in the near future. They also are desirable for situations in which only some subscribers using a switching center are being

TYPICAL ACCESS POINTS

(PERIPHERAL & INTEGRAL)

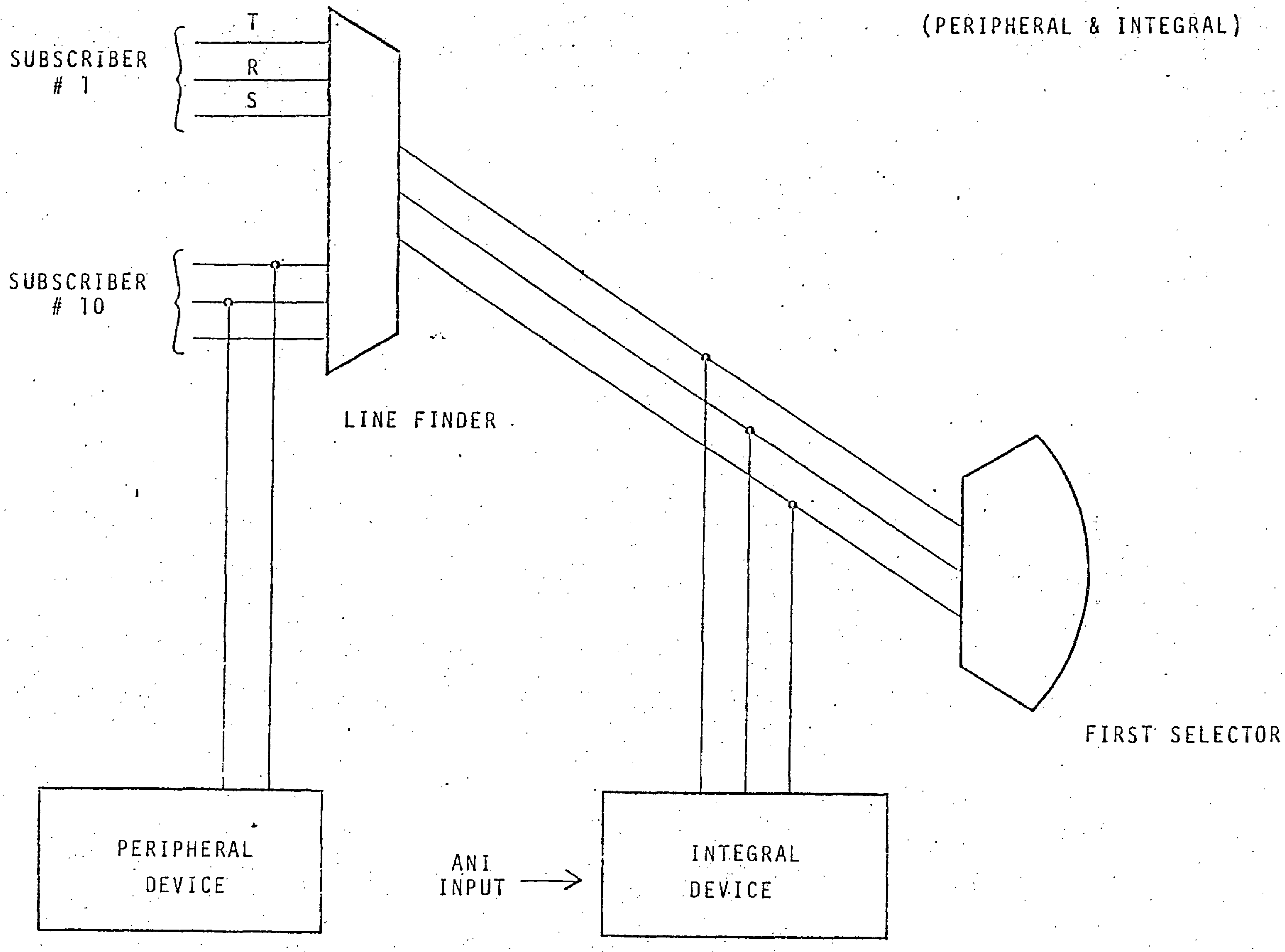
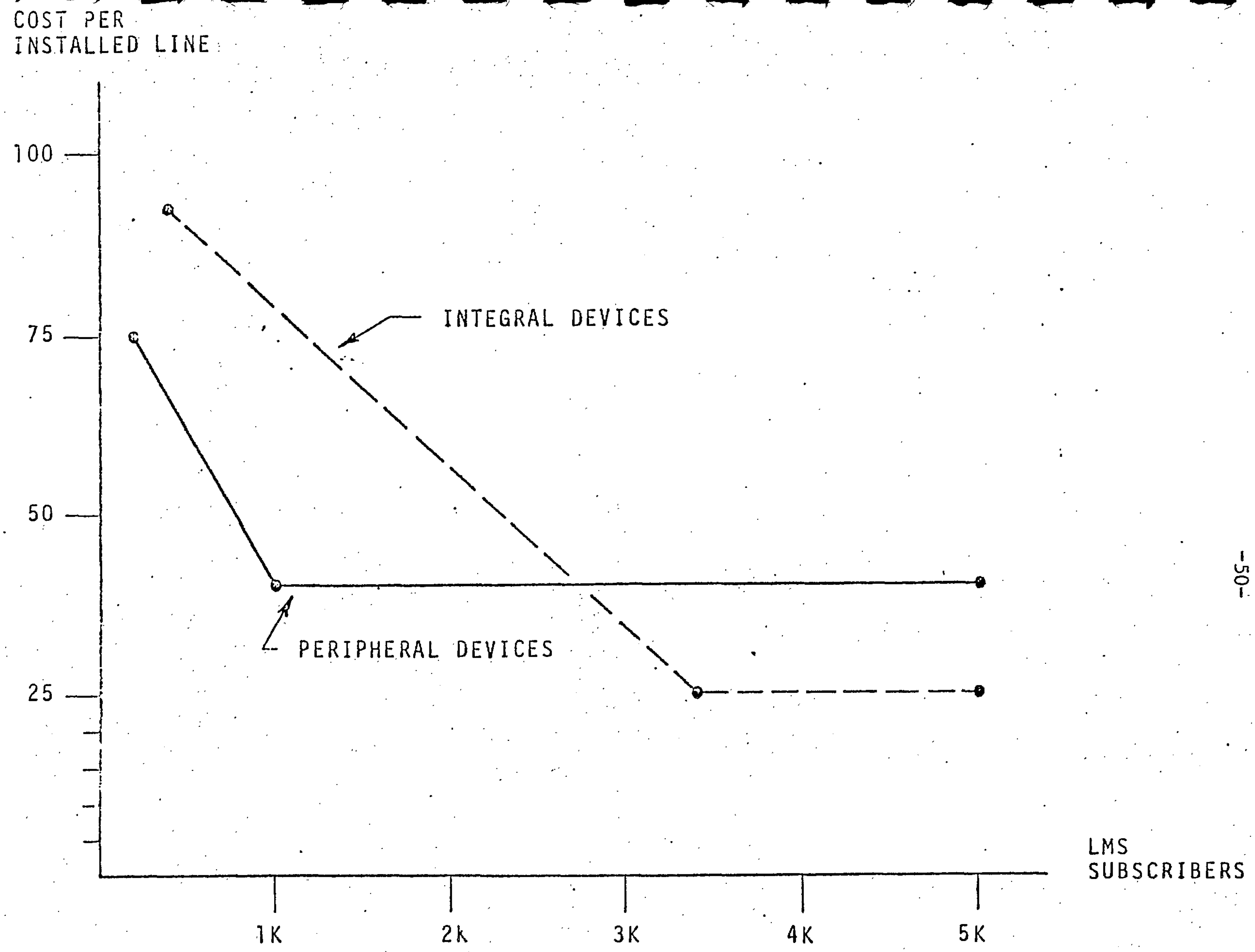


Figure 6

charged under LMS. For example as LMS is phased in, only business subscribers will be moved away from a flat rate in a transitional period. The accessibility of the Tip and Ring leads and the modularity of the available peripheral devices make gathering data under LMS in Step-by-step offices a practical approach. Since connections are made on a subscriber by subscriber basis, peripheral devices seem best suited for exchanges with under 3,000 subscribers.

The integral device can allow for vertical service offerings. It can record such call details as abbreviated or touch dialing, automatic redial, and other originating custom calling features. The installation cost is almost 40% less than the peripheral devices since the connection is at a point of concentration. Typically, the connection is made at tip, ring, and sleeve leads between the line-finders and first selector stage. The sleeve lead of each subscriber must also be connected in conjunction with ANI (automatic number identification) equipment.

Integral devices are available which provide bulk billing of local calls and details of toll calls. It turns out that currently many companies in the U.S. are converting existing paper tape toll recording systems to magnetic tape and the cost of this change is comparable in a LAMA (local automatic message accounting) office to that of integral devices to do the job of LMS, toll billing and providing the above mentioned vertical service enhancement. However the cost of retrofitting the tributary offices in CAMA (central automatic message accounting) environments tends to dilute the advantage of this step. There are substantial



PERIPHERAL VS. INTEGRAL PRICE COMPARISON FOR LMS EQUIPMENT

Figure 7

COMPUTERIZED LMS & AMA IN A X-BAR OFFICE

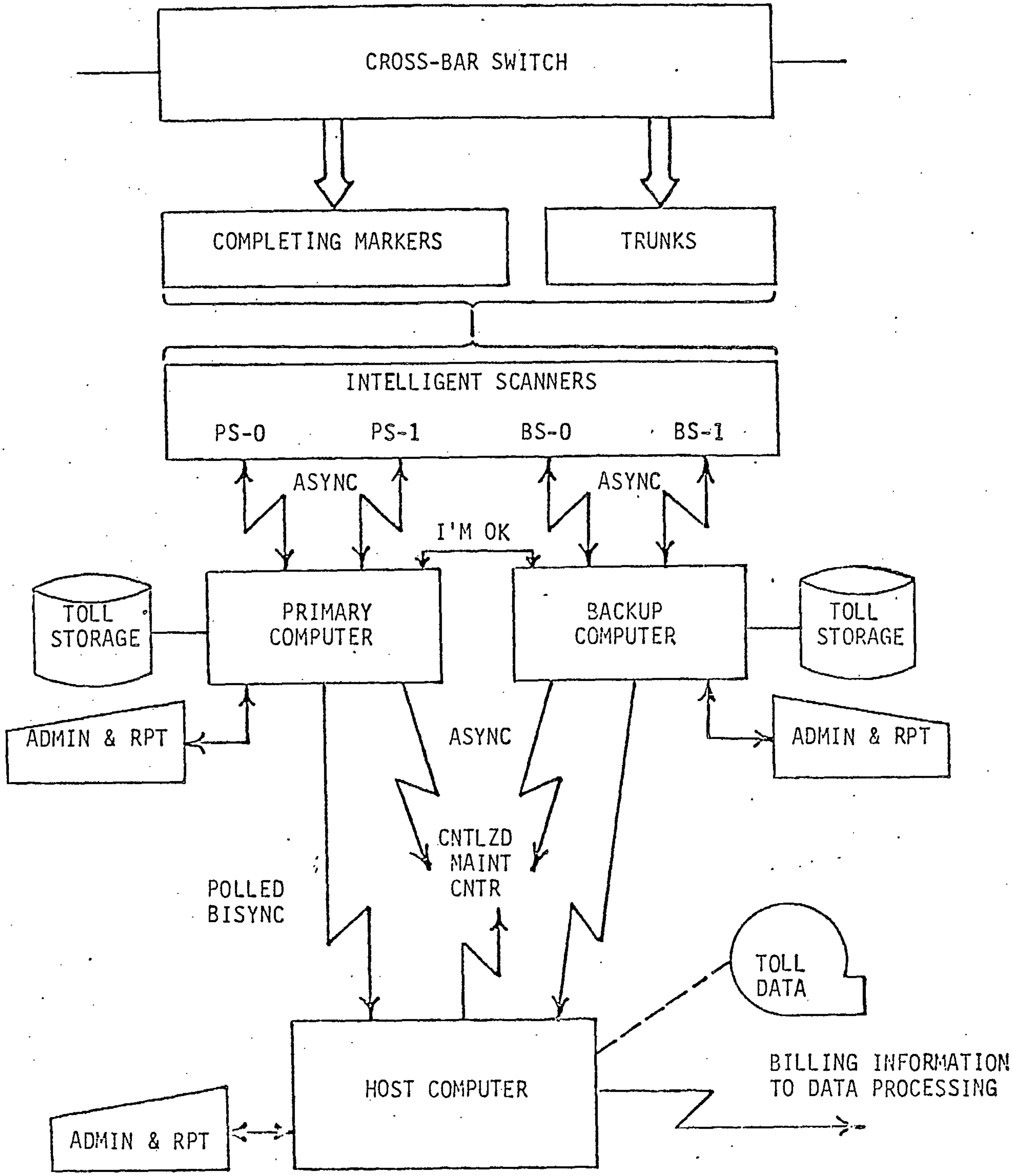


Figure 8

fixed costs associated with integral devices and these should be spread over at least 3,000 subscribers. With peripheral devices, the fixed costs are less. See the trade-off illustrated in Figure 7. However if toll recording and vertical services are taken into the calculation, then active devices which include these features in addition to LMS, whether peripheral or integral, might be cost effective in almost any sized office.

Although details of a local call will be recorded, only bulk information will be fed into the billing module. The peripheral and integral devices reduce the data to a bulk format, typically, and transmit the data to the billing complex. In Figure 8, there is an illustration of a configuration which utilizes both peripheral and integral devices with information being collected by a minicomputer at a central point. An alternative approach would have the information teleprocessed directly to that facility and call details for LMS would be retained on a storage device.

We discussed above, peripheral and integral devices for step-by-step switching entities. Much of what was reported concerning the use of peripheral devices for LMS and AMA is directly relevant for #5 Crossbar switching devices. Peripheral or outboard devices are economical solutions for implementing LMS in cases for which relatively few subscribers are to be monitored. However for large-scale LMS and AMA, integral or inboard solutions are the least cost alternative. Inboard solutions for a Crossbar system are classified as equipment that fits within either the switch accessing interface (i.e. the line finder or first selector which we observed for the case above of a step-by-step switching entity) or the switch itself (i.e. the completing marker and trunk). The cost of installation of inboard devices are less than those for peripheral devices covering all subscribers in a switching entity because there are fewer connections to make.

There are again substantial secondary benefits of inboard solutions. On a #5 Cross Bar, a completing marker and trunk interface offers significant benefits in terms of maintenance information, equipment retirement, telephone usage details (e.g. completes and incompletes), more accurate call billing and timing, plus toll data collection.

Inboard or integral devices can accommodate call service growth opportunities without requiring call switch replacement. This is known as accommodating vertical growth.

The IBM inboard solution to LMS and AMA for a Crossbar System is as follows. The Intelligent Scanner is connected to the Completing Markers and Trunks to pick up the four element call information on all calls handled by the Crossbar Switch. In Figure 2 there is an illustration of the computer configuration recommended to support the application. Presumably this configuration would also serve as a support for the integral devices attached to a step-by-step switching device discussed above. Except for the nature of the actual link to the telephone system, the system in Figure 8 is a general support system for providing LMS. In Figure 8, the computers are paired: one is primary and the other is a backup unit. Each subsystem monitors the operation of the other. In the event of a malfunction, the backup system assumes the billing function from the primary system. An alarm sounds as a the switchover occurs. The system will register a cause of malfunction and the system status at the time of switchover where practical. The backup system runs warm in parallel with the primary, gathering and assembling billing data but not actually assuming the billing function until the primary malfunctions. Periodically, each primary and backup computer is polled by a Host computer. Binary Synchronous Communications is recommended using a voice grade dial-up network. The estimated throughput on a 4800 bps line is approximately 40,000 assembled calls per hour. Periodic maintenance reports indicating switch malfunctions will be printed at the Host and a central maintenance center if desired.

The system monitors and processes a call in the following way. A marker seizure alerts the scanner to collect initial entry data. A scan of all marker leads is initiated. Called number, line equipment number, trunk frame, trunk block, trunk select, party identification and various marker progress indicators are all collected during seizure from the marker interface. It is then determined whether the system will generate billing or whether the call will be billed elsewhere or not billed at all.

After the call information is collected from the marker, the specific trunk seized is identified from the information received. The trunk I.D. along with the called line Equipment Number and called number is sent to a temporary call queue area in disk storage. After the call is completed, the line equipment location number obtained from the marker is translated into a calling or billing number by use of the line translation table which resides in the disk storage. The line translation table also provides the billing class of the calling line. The billing class and the called area of office code are combined with other information to determine the type of call and the Message Billing Index. This information is utilized in message unit reduction and is stored in the completed call site.

Trunk relays are scanned periodically for call supervision. Digital filtering is applied to reduce noise hits. The customer is considered to have answered if answer supervision exists for a specified interval. When it has been determined that a call has been answered the connect time is stored in the same call queue area with the rest of the initial entry information. Later the disconnect timing entry is stored in the call queue area thus completing the assumbly of call details. Billing can be made on the basis of "first party disconnect", "last party down", or "calling party disconnect". The assumbled call is then moved to another buffer area in disk pending transmission via binary synchronous communication to a Host System.

The Host concept is a computer system which provides for the collection of billing data from remote ticketing sites via a dial-up voice grade network. Thus toll data can be collected also. In fact the data from multiple switch sites can be presented to the data processing center from a Host system via magnetic tape or over a high speed data link.

Cost information was not provided for this system by the IBM representative but we see clearly why; 1) much of the cost is fixed or independent of the number of subscribers and to a lesser extent independent of the traffic, and 2) the cost of expanding the system both horizontally (more subscribers and/or more traffic) and vertically (more services to the subscriber) is sensitive to the type of LMS and AMA system selected. The decision to proceed with LMS requires careful consideration of the time horizon over which the augmented switching entity will be kept in service. There is no evidence of a planned rush to scrap existing electromechanical switching entities such as Cross Bar and step-by-step in favor of electronic switching systems with LMS and AMA built in. Thus retrofitting is the issue facing many companies and not a minor matter in the investment plans of those companies.

Portable monitoring equipment is available for LMS recording and was used in some surveys by the Continental Telephone Co. of St. Louis, MO.

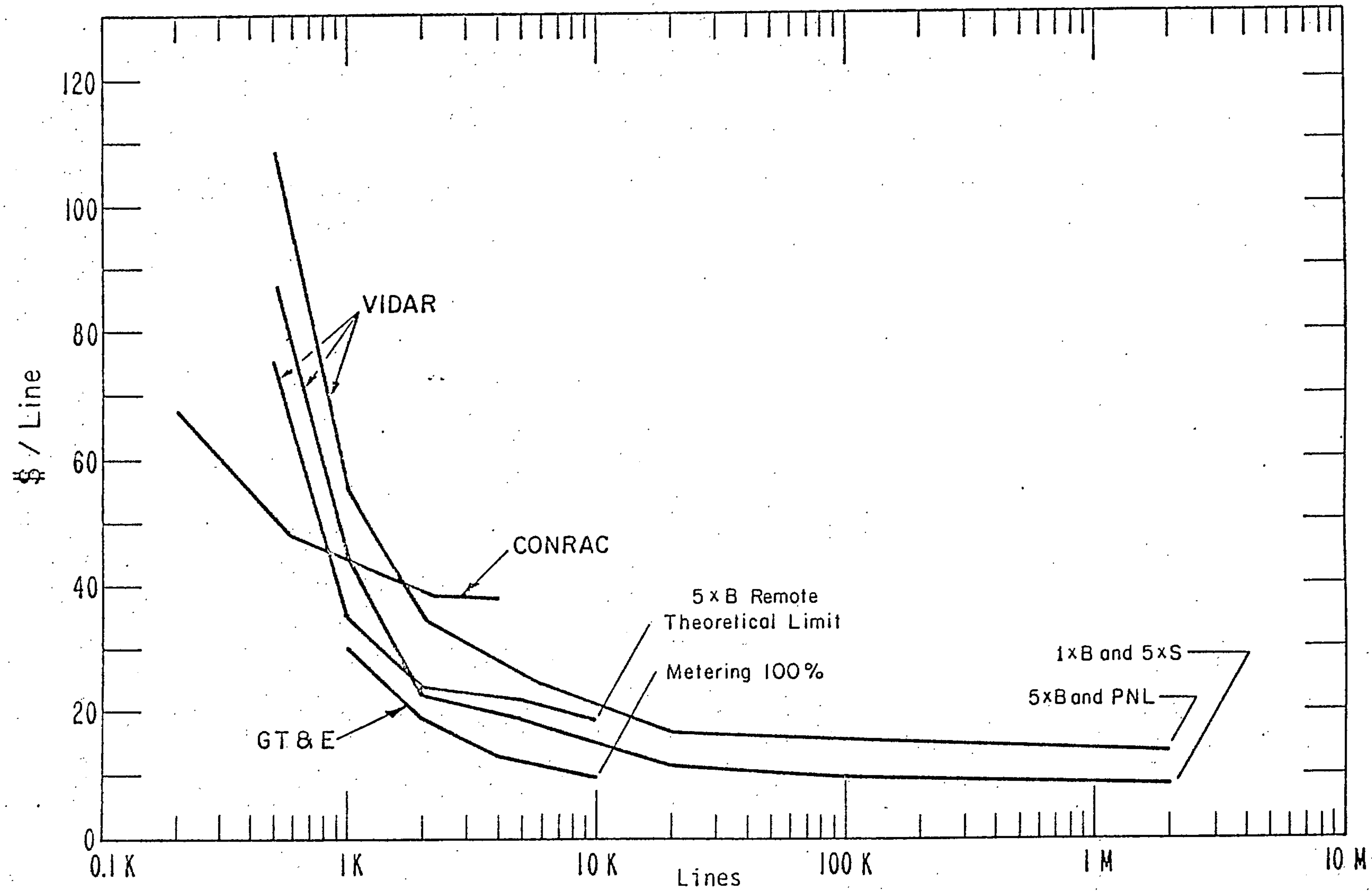
Each remote unit monitors 240 lines and is capable of storing 15000 call records. Each unit weighs 135 lbs. and uses 48 volts of power. The unit records calling number, time of day, type of supervision, call start time, call duration and called number. Both touch tone and rotary dial pulses can be processed. The unit can be used for any type of switching entity (step-by-step, electronic, or digital) since it gets its pulse from the tip and ring at the frame. A printer is required at the remote units. The system can be "quick connected" with Amphenol plugs if an office is prewired with such devices.

It uses 9 track, 1600 bp. tape on an 8½ inch reel at 25 inches per second. Each reel can store up to 125,000 calls. the units are affected by electrical noise from exchange cables and cannot measure loops over 1400 ohms. Party line customers cannot be monitored. These remote units are most suited for monitoring PBX (personal and business exhange) switching entities, but Continental used them throughout its systems.

The costs of peripheral versus integral devices for LMS and AMA reported in Figure 7 are in line with those reported in Alleman [1977]. We have reproduced Alleman's summary chart in Figure 9. These figures on a per line basis include the capital costs of data processing and billing.

We note that a remote unit from Vidar (the company which supplied equipment for LMS metering of New York City's 3.6 million subscribers) is lower cost per line than an integral unit for a switching entity with less than about 11,000 lines. The Conrac schedule relates to equipment provided by Conrac for converting small C.O. (central offices) or switching entities to USP. Like many reported cost figures, there is no breakdown by type of switch being monitored or by quality of monitoring being carried out. The actual recording and data processing equipment added on can vary widely in terms of cost and ancillary services, vertical and horizontal, provided.

Lichtenwalner [1980] indicates that the capital cost per switching entity rises from about $30,000 to $36,000 as lines rise from 0 to 6000.



The Average Metering Investment Costs

Figure 9

Thus the capital cost of LMS and AMA for a 1000 line central office is about $31 per line and for a 3000 line office is about $11 and for a 6000 line office falls to about $6. These figures are similar to Alleman's in the neighborhood of 1000 line central offices but are below Alleman's on average for central offices above 1000 lines. (Alleman's GT&E figures are similar to Lichtenwalner's.) One cannot overlook the possibility that Lichtenwalner and Alleman may have drawn on the same sources for their estimates.

The GT&E schedule relates to observations which have become available as a result of that company's experiment with USP in Illinois. This was a small scale experimental situation and the data for costs of monitoring are possibly an underestimate of actual costs which must be met in a long term commercial application. (Metering costs for step-by-step switching entities were netted out of the figures reported in Figure 9.)

Mountain Bell reported costs on a flow (not capital basis) for its system centered on Denver. Their figures were 14 cents per line for EES, 19 cents per line for 5XB and 92 cents for SXS. The administrative costs were reported to be 24 cents per line. This leaves a gross figure for step-by-step of about three times that for the electronic switching entity. (The nominal magnitudes were in U.S. dollars for a publication dated 1973.) Step-by-step equipment was expected to be scrapped in the near future and so the metering equipment was amortized over a relatively short time horizon.

This brings us to the critical timing decision facing telephone companies. If USP is brought in quickly, much "dated" equipment will have to be retrofitted with new metering divices which in turn will likely be scrapped in a matter of years (not decades). The sunk costs of new metering equipment on older equipment will induce companies to delay scrapping older equipment longer than they would have in the absence of retrofitting. If USP is phased in slowly, inboard metering devices will be installed with new switching entities and on average one expects a more rapid scrapping of older switching equipment. The payoff from having USP is traded off against the payoff from not scrapping equipment already paid for. The scrap value of integral devices for step-by-step switching entities must be close to zero whereas for peripheral devices, the scrap value could be substantial since such equipment could be used in other applications.

A related aspect of the timing of the implementation of USP is the matter of upgrading existing service. If the introduction of USP involves the phasing out of party lines before a date set in the absence of USP, then part of the cost of such upgrading represents a cost of USP. Party lines are still common in rural areas in the U.S. This seems to be one reason why USP has been implemented in a dense urban area such as New York City rather in less densely settled areas. Mitchell [1978] reported that New York Telephone estimated that the cost of installing metering equipment in its large No. 1 cross bar exchanges in New York City was about $15 per line. Record keeping and billing cost estimates were from $.001 to $.003 per call.

The Pulse System of Metering: The European Approach

The institutional setting in which telephone or telecommunications services generally are provided is quite different in Europe and much of the world from that in the U. S. and Canada. Public ownership is the rule for the most part outside the U.S. and much of Canada. Strangely enough, the form which USP is assuming in Europe is quite different from the form being developed in the U.S. In Europe a billing unit called a message unit is set and the charge for a call is expressed in so many message units. These message units correspond in a rigid one-to-one relationship to a technical datum called the pulse. The entire telephone network throbs with these regularly spaced pulses at all times and as a subscriber makes a call, the number of pulses he "confronts" is registered in his account. The pulse frequency varies with the distance over which a call is made and in many cases with the time of day. To bring charges in line with costs the calibration of the pulse frequency and charge per message unit has to be carefully worked out. The system is obviously technically much simpler than a measured service system which records for every call, number calling, number called, time of day, duration, and distance. It also has the advantage of preserving the privacy of a subscriber's personal communication network. It has been deemed unsatisfactory by some U.S. observers because a subscriber cannot verify that his bill is accurate since no detail per call is recorded (see for example Lichtenwalner [1980] p.39). Actually for the pulse system, a subscriber can purchase a device for his handset which records the pulses "incurred" immediately before the eyes of the user. It would seem that such a device could serve quite adequately as a check on the monthly bill. The subscriber could compare totals but also could, while a call is being made, verify that the pulses incurred conformed to his understanding of the approximate charge for the call. The segments of the monthly total could be observed as they were incurred. Thus it does appear that one can overcome the problem of the presumed absence of verifiability of the subscriber's bill under the pulse system.

The recent experience with the pulse system for USP has been surveyed very well by Mitchell [1979]. Mitchell's survey deals with six European countries -- Denmark, Norway, Sweden, Switzerland, the United Kingdom and West Germany. Some form of local call timing based on the pulse method or earlier more mechanically oriented versions of the current technology have been in effect in Denmark since 1950, in the U.K. since 1958, in Finland since 1960, in Spain since 1971, in Norway since 1975, in Switzerland since 1978, and in West Germany since 1980. Sweden will convert to the pulse system in 1982. First we should emphasize that the estimated capital cost per line for such LMS and message unit billing is between $4 and $10 in 1978 U.S. dollars. Mitchell [1978 AER] estimates that full LMS and AMA along the U.S. model (so-called 4 element billing) costs between $5 and $50 per line for electromechanical offices (step-by-step and crossbar) and between $2 and $5 for electronic switching offices. If detail is not required, then the pulse method is much cheaper since from a technical point of view the monitoring mechanism for calls is much simpler and the billing procedures involve only summing message units and multiplying them by the single chosen pricing factor. Since we noted above that there are straightforward ways for an individual to monitor his personal calling costs under the pulse system, we must conclude that the implicit "cost" or shortcoming of the pulse system of USP relative to full 4-element monitoring system is the degree of verifiability, the degree of flexibility in price setting per call (e.g. one long distance call compared with another of a different distance), and the ease of monitoring aggregate traffic by the operating company. The qualifications of the desirability of the pulse system are all matters of degree and not basic issues of the relative "capabilities" of one system relative to the other. The pulse system does preserve the privacy of the billed party and seems more in line with western traditions of civil rights.

The two major issues facing designers of a pulse system are the timing of the pulse for various calls, and the pricing of the first interval of a call. Subsidiary issues are the pricing of the message unit and the organization of calling areas or districts. It turns out that some pulse systems have been calibrated in practice with a number of minutes between consecutive pulses for local calls. For calls averaging three minutes, it becomes a complicated calculation to arrive at a charge which reflects actual time elapsed for the call. The matter is made difficult not only because pulses come in discrete and sometimes relatively widely spaced intervals but because as a caller joins the network, he may make contact as a pulse just "beats" or on any faction of an interval between beats. Various averaging or smoothing algorithms have been developed to bring charges on average in line with call duration. These are described in some detail in Mitchell [1979]. With new electronic devices available for creating pulses, it should be possible to keep the interval between pulses very small, and the problems associated with discreteness and billing for actual call time elapsed will become of trivial magnitude. Mitchell reports, however, that the basic daytime pulse interval in these European systems ranges from 2 to 8 minutes. Daytime pulse intervals in the five distance bands of the Swiss trunk network are currently at 60, 36, 22.5, 16.3, and 12.8 seconds, also not trivially small intervals. Devices for generating pulses have been electromechanical. The new electronic pulse generators in West Germany will provide for intervals from .3 to 1600 seconds. In one country (Denmark) dialing and ringing time will be included in the length of a call. A single pulse was often associated with one local call in the past and this accounts in part for the seeming long interval between pulses currently in use and for the charges per pulse. Figures cited by Mitchell for 1978 in U.S. dollars have the charge per pulse ranging from a low of 3.5¢ for Denmark to a high of 12.5¢ in Norway.

The demarcation of areas for local calls clearly influences the quantity of service demanded since it is an indirect form of pricing by the operating company. The desiderata involved in arranging these areas include treating each custom similarly. People who live near boundaries between districts face different outlays compared with those well away from a boundary. These difficulties can be mitigated by delimiting overlapping local calling areas and assigning a particular subscriber to the area with the central office which he is closest to. This approach has been set out in West Germany and the pattern of overlapping areas is referred to as a fish scale configuration.

In his investigation of European pricing experience, Mitchell observed some changes in calling in response to changes in the system used for charging customers. In January 1975, Norway introduced a periodic pulse pricing system for local calls made between 8:00 a.m. and 5:00 p.m., Monday through Saturday. After 5:00 p.m. and on Sunday, each call is only charged for one pulse. A 500 person survey was conducted before and after the introduction of the periodic pulse pricing system. Prior to the changeover in pricing, the average local call was 3 minutes long. The data show the customary pattern of longer calls by residential subscribers (averaging over 4 minutes) compared with calls of business and private-exchange (PABX) users (about 2-5 minutes). Business subscribers make more than five times as many calls during the day as residential users.

The new pricing scheme (with a 3 minute pulse and a charge of 11.7¢ per pulse) resulted in reductions in both the number and duration of daytime calls by every type of sbuscriber. Residential subscribers reduced the number of their calls by 17% and conversation time by 41%. Business and PABX users reduced the number of their calls by about 11% and the duration by 7%. The data indicated that residential and business customers (and not PABX users) shifted calls to the off peak period, increasing the number of calls by 8% and the length of calls by 7%. One can extend the examination of the impact of a new price system by developing elasticity measures (see Mitchell's resourceful inquiry). However one has to keep in mind the possibility of quite different responses by subscribers in the short run and long run. One really wants to see in these analyses careful development of short and long run price elasticities.

GTE and AT&T Experience with Local Measured Service

Senior research officers of two large U.S. telephone companies (GTE and AT&S) have discussed the transition from a flat-rate tariff to a USP system recently. Cohen and Beauvais [undated] focus attention on results from GTE's experiment in Illinois with USP and Garfinkel and Linhart [undated] address the issue of assessing the impact of a new USP tariff system for local service. Although these papers contain interesting information and calculations, each is written under the assumption that USP is being put in place by the respective company and the posture is one of discussing the merits and functioning of a USP system (different systems seem to be envisaged for the two companies) and not directly the pros and cons of their form of USP relative to other forms or relative to a flat-rate system. Cohen and Beauvais (hereafter GTE) report that "the objective is to have 90% of the company's customers on measured service over a ten year horizon". Garfinkel and Linhart (hereafter AT&T) state: "A majority of business users already have measured service and the transformation of residential local service pricing is proceeding." GTE go so far as to indicate that the move to USP is not only for reasons of efficient pricing and costing of components of their telephone system but also to develop charges which will bring revenues from "exchange telephone service" under to the portion of total system costs attributable to exchange service. It has been conventional wisdom that local telephone service was subsidized in the U.S. by revenues from other parts of the telephone system and GTE are explicit on this point.

We can use those reports from GTE and AT&T to shed only a tiny bit of light on the implications of implementing local measured service (LMS) since what is reported is part of a program of advocacy and not either an impartial analysis nor a careful prediction of other consequences of a shift from a flat-rate pricing scheme to one of LMS. What is perhaps reasonable to predict, even without these documents from GTE and AT&T is that charges under LMS will be set so that in the medium-run, revenues will be brought closer to costs in the relevant sector of the company. It is not clear that some new forms of cross-subsidization will not develop within the telephone companies, forms of cross-subsidization involving local service. However, what AT&T and GTE plan in the short-run is to not have post-USP revenues significantly different from pre-USP revenues on average, the average being calculated over subscribers.

The GTE paper reports on a particular experiment with the introduction of LMS, while the AT&T paper describes current local calling characteristics under a flat-rate pricing system and a "black-box" computer program for estimating the effect of introducing LMS. The particular GTE experiment (2.5¢ of fixed charged per call plus 1¢ per minute with a 20% discount for 5-11 p.m. and a 50% discount for 11 p.m. - 8 a.m. $3.45 access charge and $19.00 maximum bill) resulted in a 19% reduction in telephone usage. There was a significant decline in revenues to the company and no appreciable improvement in the load factor. One must keep in mind that regulators placed a ceiling on bill size in the post LMS situation so that very heavy users were still not induced to curtail their telephone use. It was found that 60% of subscribers were able to reduce their monthly bill under the LMS pricing scheme but it was felt that this benefit for these customers was in part a result of an excessive discount for use during evening hours. Also, the conclusion was added: "from a policy perspective, neither age, nor income, nor occupation, nor education has been found to have any significant effect on usage.... It reinforces the economists' position that such redistribution questions [as those involved in providing low rates for say the elderly ("lifeline rates")] are not appropriate topics to be handled by the rate-making process."

In the long run, full or non-optional LMS, is considered to result in lower usage levels than under a flat-rate system and a substantial reduction in switching investment could be envisaged under an LMS pricing scheme. In the transition to LMS, new expenses are involved in 1) data processing, 2) business office, 3) operator services and 4) measurement equipment. The following table is instructive since it indicates the costs which have been found relevant in planning by GTE for LMS.

	Range of Potential Cost	Base Case Cost
Business Office Start-Up and Training (Per Account)	$ .05-$ 2.00	$ .06
Business Office Ongoing (Per Account Per Year)	$ .50-$ 2.00	$ 1.02
Measuring Cost (Per Call)	$ .001-$ .006	$ .001
Metering Cost (Per Line Installed)	$ 5.00-$60.00	$ 5.00
Data System Development (Per Line)	$ 0.75-$ 1.50	$ 1.25
Switching Facilities (Per Busy Hour CCS Installed)	$25.00-$75.00	$35.00

These costs cannot be allocated by subscriber directly since some are for the switching facility time. Except for switching facility time, they are in line with other estimates. The $5.00 per line for metering is on the higher end for EES equipment and on the lower end for electromechanical equipment. In some system wide simulations, GTE contends that the LMS and automated message accounting (AMA) is the efficient pricing approach when viewed from a medium-term time horizon (over which the costs of LMS and AMA can be amortized). It should be noted that much of this GTE paper is describing in a few sentences large and complex exercises such as system wide simulations. The reader cannot follow the trail back to determine which assumptions and parameter choices were crucial for the particular results.

GTE investigated subscribers' subjective reaction to a switch to LMS from flat-rate charging. Initial reaction was negative by a majority of subscribers but when, well into the experiment, it was seen that the bills for a majority were lower under LMS, about fifty per cent pronounced a preference for the LMS pricing scheme. Given historical calling patterns, one would expect it quite difficult to get a <u>majority</u> in favor of LMS since the minority of <u>very heavy</u> users would vote against the scheme but also since lower income people have been observed to use the telephone more in minutes per day (Mitchell[1978]), they would likely vote against LMS. Since there are more lower income families than higher income, on a straight vote, one would expect <u>a priori</u> that a majority would be opposed. When one then adds the fact that local service has been subsidized for the most part by toll service, subscribers might be wary of voting for

LMS in the short run, unsure of the charges in the longer run. The problem with asking people to state their reaction "favorable" or "unfavorable" is that intensity of preference is not taken into account. Thus one has to ask whether certain surveys accurately reflect the appropriate aspects of the issues being investigated. To overcome some of the deficiencies of questionnaires, an estimate of the change in producer and consumer surplus was reported. It was positive, reflecting a welfare improvement resulting from the introduction of LMS. (Detailed calculations were not reported.)

The AT&T paper contains a sketchy description of a computer program which is designed to indicate to a local operating company what the impact of the introduction of LMS will be, primarily on revenue. There are two aspects of particular interest, largely ancillary to the computer program, in the paper. First AT&T appears to prefer an LMS scheme of the following sort: there will be three charging formulas from which a subscriber can select. One formula involves a flat rate for local calling. The second formula involves a fixed charge plus a certain dollar value of "free" calls and then a charge per call (weighted by duration). The third formula is similar to the second but with different parameters. Customers can select the formula they wish to have applied to their usage. Apparently not all consumers select the least cost offering. This might be more subtle than at first blush it appears. If a customer is uncertain about his usage, he may choose a formula which permits him to avoid unanticipated large bills. Thus the choice of formula involves an insurance component as well as a direct use component -- the insurance premium being a charge by the company for absorbing the risk of a potentially costly (high aggregate usage) period. A flat rate transfers the cost to a subscriber of bearing variability in his monthly outlays from the subscriber to the operating company. The operating company then incurs the cost of the variability in its aggregate costs of providing service. If each household's use is not highly correlated with another's use, the operating company can spread the individual variability or act as an insurance company. Flat rates make good sense. However, there are significant differences between average use over households, and flat rate schemes involve cross-subsidization of one use group by another. The insurance element of flat rates becomes a small item relative to the dollar amounts of cross-subsidization that result from a single flat rate for classes of users with widely varying use patterns. Moreover, peaking of aggregate usage is a case of an absence of statistical independence among individual households' usage. The operating companies can incur substantial costs which get passed on to customers of providing capacity for peak use under a flat rate pricing scheme. This lack of independence among subscribers' usage is another point against the seemingly attractive idea of the operating company as insurance organization.

It is difficult to view the AT&T LMS pricing scheme as one of marginal cost pricing. The GTE scheme with a small fixed charge per call plus a charge per minute thereafter, both charges possibly varying with time of day, seems much closer to a marginal cost pricing approach. The fixed charge per call is associated with the "demand" for switching capacity and the charge per minute is associated with variable costs per call. However marginal cost pricing has never seemed to capture the loyalty of managers of public utilities in North America in general, say relative to those in Europe. Elements of cross-subsidization, price-discrimination, and "income" redistribution have persisted in the pricing of public utilities in North America, at least until recent years. An actual example of the AT&T pricing scheme is reported for Juniper, Florida.

Class of Service	Fixed Charge	Allowance
Flat Rate	$10.95	-
Standard Measured	$ 9.05	$5.10
Low-Use	$ 6.90	$2.00

Usage Charges:

Distance	Initial Minute	Each Additional Minute
Tier 1	$ .05	1¢
Tier 2	$ .11	3¢
Tier 3	$ .20	6¢

Time of Day Discounts

12 p.m. - 2 p.m.	weekdays	}	-25%
5 p.m. - 11 p.m.	weekdays and Sunday	}	
11 p.m. - 8 a.m.	weekdays	}	
8 a.m. - 5 p.m.	Sunday	}	-50%
All day	Saturday	}	

Curiously enough, actual selections were 88% for the Flat Rate, 2% for the Standard Measured and 10% for the Low-Use. Clearly different charge parameters in the formulas would lead to quite different patterns of formula selection.

The second element of interest in the paper, is the description of use patterns prevailing under a flat-rate system. We have noted these data in our section: Parameters of the Telephone Usage. With regard to the actual computerized model, we note two difficulties confronting the reader of the description. The change of a pricing scheme naturally creates changes in usage. The nature of this change is the focal point of our investigation in this study. "Repression" is the name given by telephone company officials to the reduction in usage resulting from measured service. The AT&T paper "repression" seems to be introduced in the model in order to bring outcomes from the model into line with observed results rather than to be based on a priori estimates of household elasticities of demand for telephone usage with respect to price. Secondly, the interesting question of how different pricing schemes affect non-local calling patterns is completely obscured by the procedure of attributing to a subscriber usage based on the average of all subscribers. Again insights with regard to possible elasticities cannot be drawn from the report.

In another AT&T study (Infosino [1980]) of exchanges with no EAS and only flat rates, it was found that usage was independent of income. In fact other demographic characteristics also had little relationship to usage (i.e. sex of household head, age of household head (people over 65 used the telephone less), duration of residence in the area, education of household head). A model was derived for use in predicting local residential telephone usage of the form

Call Rate = -1.34 + 1.10 x (Number in household) + 2.30 x (Race dummy variable) + .000204 x (density of phones in the area).

This model was derived from data for 10 exchanges in each of Cincinnati and California. The basic 'area' in the study was a local exchange area or "wire center". Calling rate was defined as total local calls in an exchange divided by the number of households, per unit time.

PART III

Econometric Analysis of Local Telephone Pricing

Introduction

If it were costless to charge for local telephone calls by duration, distance and time of day, and if the sole consideration in designing price systems for local telephone calls were economic efficiency, then it is unlikely that we would ever observe flat-rate pricing. Usage-sensitive pricing would be the norm for local telephone service, as it is for most services provided through the market system. In the real world, however, there are at least three reasons why flat-rate pricing may be preferable to usage-sensitive pricing in certain circumstances. Whether these circumstances in fact prevail is a question which can possibly be answered by econometric analysis.

The first reason why flat-rate pricing may be socially optimal is that measuring local telephone usage, and billing for it, is costly. By charging a price close to marginal cost, rather than a price of zero, people are induced to forego making phone calls the value of which, to them, is less than their marginal cost. This should increase the sum of producers' and consumers' surplus, by an amount equal to half the reduction in the cost of providing the telephone service, in the case where the price charged is actually the marginal cost and the demand curve is linear (see Mitchell [1978] and the chapter by Mitchell in Baude *et al.* [1979]). But whether this gain in efficiency is great enough to offset the costs of measuring and billing for local telephone usage is clearly an empirical question. If charging for local phone calls has no effect on consumers' telephone usage, or if that effect is relatively small, then the costs of usage-sensitive pricing will exceed the benefits. Thus the first question that one would want to answer by econometric methods is: how responsive is the demand for local phone calls to the price (or system of prices) charged?

A second reason for retaining flat-rate pricing is that it may serve, to some extent, as an insurance policy. The number and duration of local phone calls that people make may be, in some cases, largely determined by random events which vary from month to month. Under a system of usage-sensitive pricing, the local phone bill will also vary from month to month. If people dislike such variation, they would presumably be willing to pay more (on average) under a flat-rate system than under a system of measured local service. The additional amount they are willing to pay is a measure of the insurance value of flat-rate pricing, and this amount should be added to the costs of switching from flat-rate to measured service. Whether this argument has any empirical validity, and just how great the insurance value of flat-rate pricing is, can in principle be ascertained by empirical econometric work.

Finally, even if there is a non-trivial efficiency gain from usage-sensitive rather than flat-rate pricing, politicians might still prefer the latter because of its distributional consequences. Economists generally argue that distributional considerations should not be allowed to stand in the way of economic efficiency, on the grounds that it is far more effective to transfer income directly than it is to subsidize the prices of goods which are disproportionately consumed by groups considered worthy of subsidization. Nevertheless, it is possible that if certain groups such as the poor, the old and racial minorities make substantially more use of the local telephone system than average, and if it would be politically difficult or impossible to offset the distributional effects of a change in the pricing system by direct transfers, then there might be an argument for retaining flat-rate pricing. Once again, whether this sort of argument has any validity is first of all an empirical question.

Existing empirical work has not answered all of the above questions satisfactorily. No work at all has been done in Canada, and not very much has been done in the United States. Nevertheless, existing work does provide some useful information. The GTE measured service experiment, analysed in a number of papers including Park and Wetzel (1980) and Park et al. (1980), has certainly established that local telephone usage is sensitive to price. When people have to pay for local phone calls, they make fewer of them, and when they have to pay more for longer calls, the calls they make are shorter. That is exactly what any economist would expect. Unfortunately, the GTE results do not allow one to say with any confidence by how much local telephone usage would drop if one switched from flat-rate pricing to a usage-sensitive scheme that did not closely resemble the one used in the experiment. That point will be taken up again below.

In addition, both the GTE results and independent work using A.T.&T. data (see Brandon (1981) and Infosino (1980)) make it quite clear that local telephone usage does depend on demographic characteristics, especially on the size and age-composition of households and on race. At the same time, any systematic variation in telephone usage is swamped by random variation across households, so that subsidization of local phone calls is a very inefficient method of redistributing income.

Estimates of the Effect of Price on the Demand for Local Telephone Services

There are at least three different types of data which can be used to estimate the effect of price on the demand for local telephone calls. The most elaborate and expensive approach is to use data for individual firms and/or households. Such data can be gathered by combining survey data on household characteristics with data on the number and duration of local phone calls obtained as a byproduct of charging for local phone calls. In the case of the GTE experiment, household data were obtained both before and after local measured service went into effect, so that changes in the demand for local phone calls between the two dates can be attributed to the effect of prices. In the future, as more and more telephone companies in the United States adopt various forms of usage-sensitive pricing, it should be possible to obtain data for households which face a wide variety of pricing systems, by administering surveys and gathering related usage data in a number of different localities. However, the cost of obtaining, managing and analyzing this type of data can be very large.

A much less expensive approach would be to use cross-section data where the units of observation are municipalities or, perhaps, telephone exchanges. So far, to my knowledge, no studies using data of this type have been performed. However, within a few years it should be possible to find a wide variety of telephone pricing schemes within North America. Gathering data on total telephone usage (number of calls or number of minutes) should be easy enough for the telephone companies, and demographic data can be taken from the census. While this sort of study clearly could not answer any very precise questions about local telephone usage, it might well be able to provide reasonable estimates of the extent to which aggregate usage responds to price in the long run. That of course is precisely what we need to know to decide whether the efficiency gains from usage-sensitive pricing exceed the costs.

Another relatively inexpensive approach is to use aggregate time-series data. Suppose that one can observe telephone usage data for an entire community on a weekly, daily or even hourly basis, over a period of several years. Suppose further that the system of local telephone pricing changed substantially during that period. Then it should be straightforward to estimate a time-series model which explains the pattern of day-to-day or week-to-week fluctuations in telephone usage, first for the period prior to the price system change, and then for the period after the change (perhaps dropping a month or two in the middle to allow consumers to adjust to the new system). Any differences between the estimated levels of telephone usage between the two models (adjusting for trend effects, if any) could then be ascribed to the effect of the price change, and confidence bands could be constructed in a straightforward way.

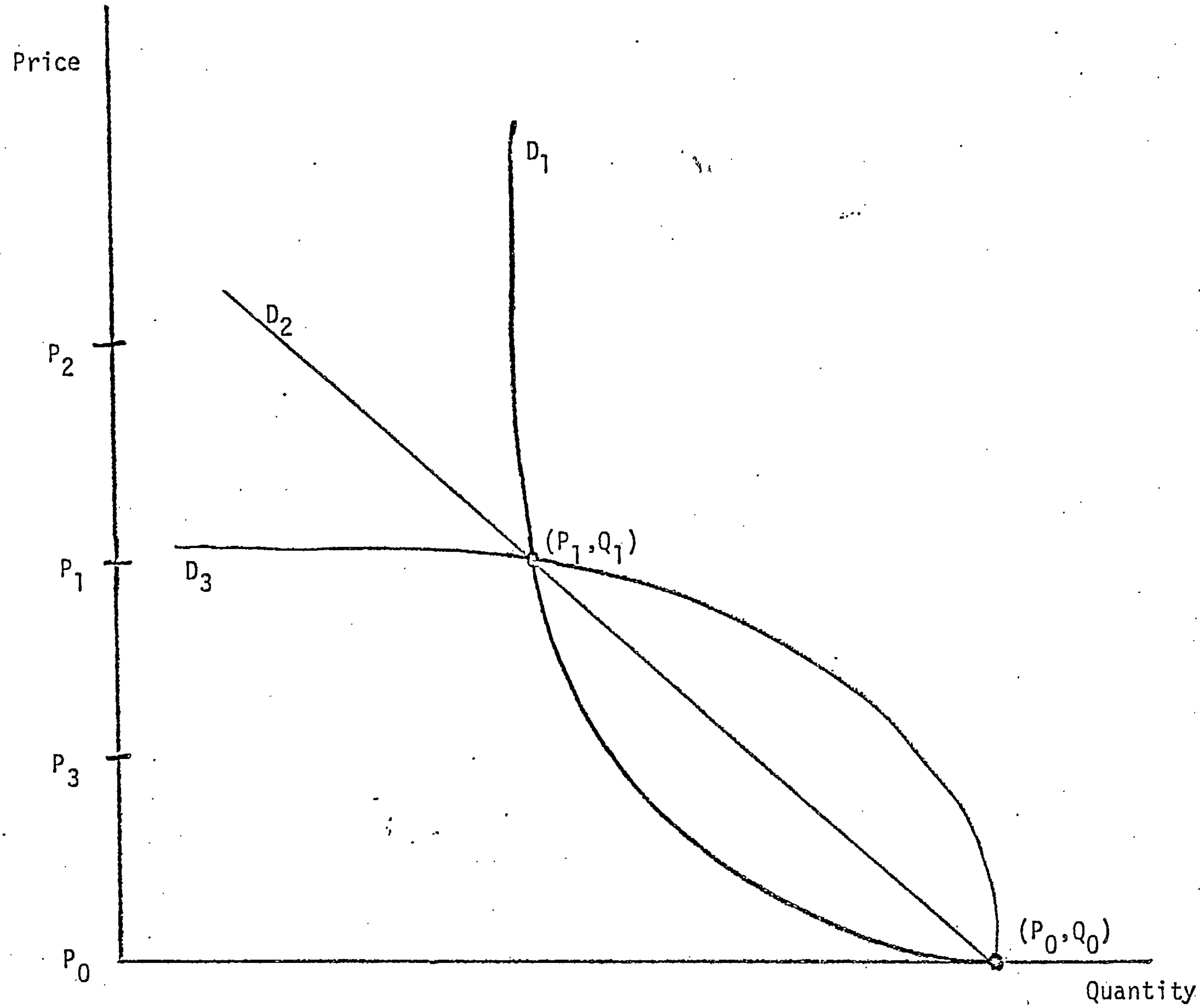
Exactly this approach has been used by Jensik (in Baude et al. (1979)), using monthly data from the GTE experiment. He concludes that measured service reduced local telephone usage by about 19 per cent. A slightly different approach has been used more recently by Wilkinson (1980). He constructed time-series models using monthly data, but estimated them over the whole period of the sample, putting in a dummy variable to account for the change in pricing system. His conclusion was that measured service reduced local telephone usage by about 23 per cent. The use of dummy variables in this context is rather dubious, since a change in the price system could be expected to influence more than just the mean of the dependent variable. But it is unlikely that the results would change much if a more general specification were used.

The basic problem with the Jensik and Wilkinson studies is not that their assumptions are dubious, but that the data they use are fundamentally limited. Jensik's estimates seem more consistent with those of other studies than Wilkinson's, so we shall, for the moment, accept that they are true. Like any estimated quantities, they are actually somewhat imprecise, but that is not the source of the difficulty. Thus let us accept as a fact Jensik's estimate that charging 1.5¢ per minute instead of zero will cause a reduction of nineteen per cent in the number of minutes of local phone calls, for households with single-party service, in certain small Illinois communities in mid-1977. Let us make a further leap of faith and assume that these communities, so far as their use of the local telephone system is concerned, are typical of North America as a whole, so that Jensik's estimate of the "repression" associated with a charge of 1.5¢ per minute may be taken as gospel.

Unfortunately, estimates of this type tell us very little about the demand curve for local telephone services. Suppose, for example, that we wanted to know the effect of charging 3.0¢ per minute. Then all we can say for certain, based on Jensik's "fact", is that demand will fall by not less than 19%, and by not more than 100%. Similarly, if we wanted to know how much demand would fall if we charged 0.75¢ per minute, all we could say for certain is that it would not fall by more than 19%, and surely would not rise.

To see why this is so, consider Figure 1. Jensik's "fact" allows us to observe two points on the demand curve for minutes of local phone calls: (P_0,Q_0) and (P_1,Q_1). Here P_0 is zero, P_1 is 1.5¢ per minute, and Q_1 is 19% less than Q_0. But there is an infinite number of demand curves which pass through the points (P_0,Q_0) and (P_1,Q_1); three of these curves, D_1, D_2 and D_3, are shown on the figure. Jensik's "fact" does not allow us to say which of these curves (if any of them) is the true demand curve. Now that does not matter if we are simply interested in what happens when we charge prices P_0 or P_1. But if we want to know what will happen when we charge, say, P_2 or P_3, it obviously makes an enormous difference which demand curve

Figure 1

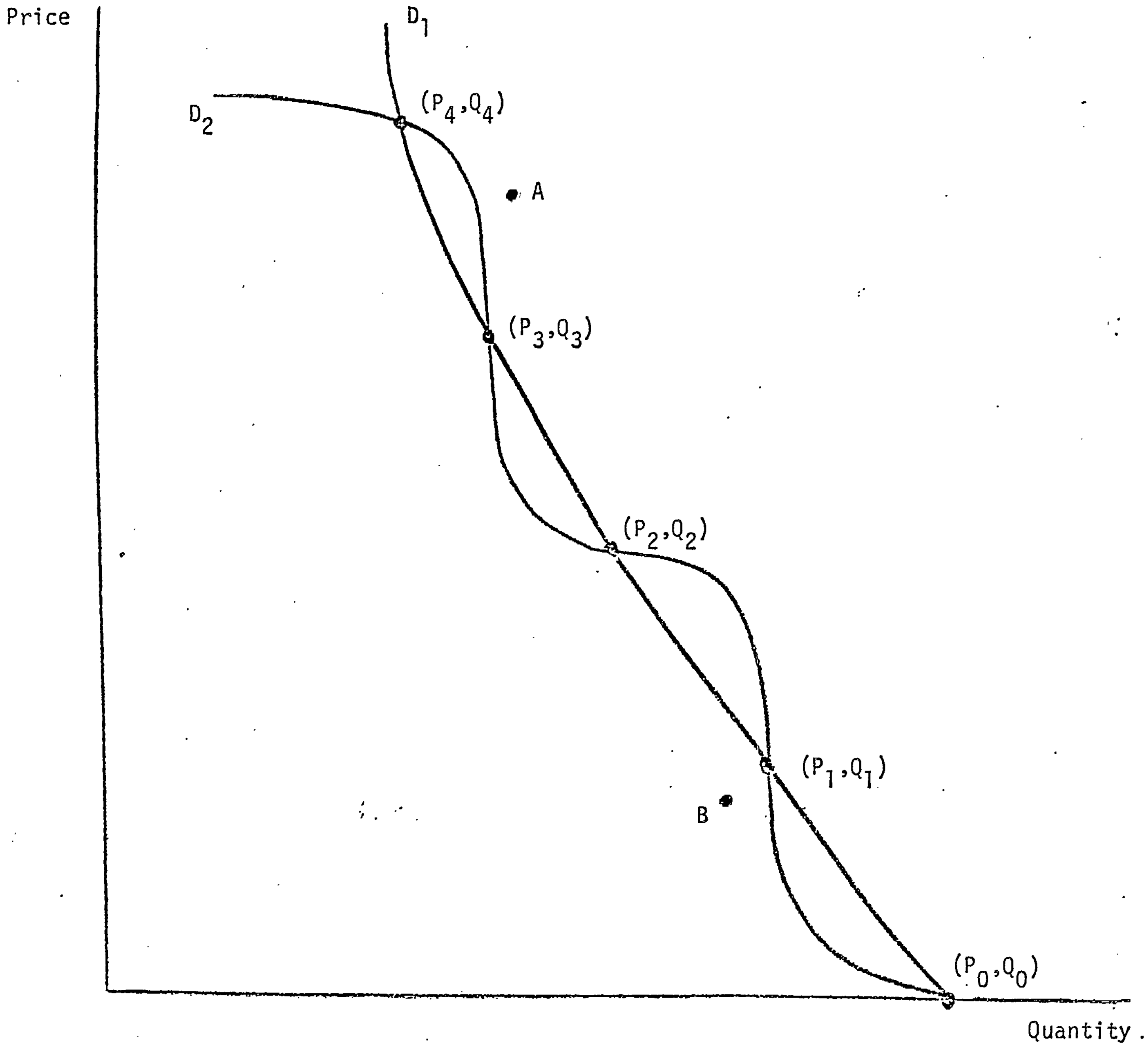


is the true one.

This point applies not just to Jensik's study, but to any study, based on either aggregate time-series or disaggregate (household or firm) cross-section data, where the only variation in price comes about because of a once-and-for-all change in the price system. Such studies may allow us to estimate with great precision the effect of that particular change, but they do not allow us to say anything at all about the demand curve.

In contrast, consider Figure 2. Here we have managed to obtain five points on the demand curve, (P_0,Q_0) through (P_4,Q_4). Now there is no guarantee that the demand curve is actually D_1, which seems to fit these points rather well. It might, conceivably, be as strange-looking as D_2, although common sense suggests that D_2 is a rather unlikely candidate for a demand curve. But whatever the true demand curve is, it certainly does not pass through points such as A and B, where both price and quantity are greater than, or less than, price and quantity at one of the points which is on the curve; the fact that demand curves slope downward sees to that. Thus if we are to be able to estimate demand curves with any precision (as opposed to simply points on those curves), we must have access to data in which price takes on several different values in the range that is of interest to us. Such data may be hard to obtain in a measured service experiment conducted in just one community, because consumers and regulatory commissions are unlikely to be in favour of either charging different prices to different consumers, or making frequent changes to the prices charged everyone. This suggests that cross-section data from a number of communities, either at

Figure 2



the aggregate level or, if sufficient resources are available, at the firm or household level, offer the best hope for estimating demand curves for local telephone services. Unfortunately, such data are not likely to be available for a few years, since it will take time to implement the differing usage-sensitive pricing schemes which are currently being proposed in the United States.

What is probably the best currently available study of the effect of price on the demand for local telephone services -- Park and Wetzel (1980) -- does not have particularly good data to work with. The GTE experiment involved three separate communities, and total monthly usage by households with single-party and multi-party service are available for each of them. During the initial period of the experiment, there was no charge for local phone calls. From mid-1977 until early 1979, two different single-party tariffs were in effect: in one community there was a charge of 2¢ per call plus 1¢ per minute, while in the other two communities there was no per-call charge and a charge of 1.5¢ per minute. After mid-1979, all communities had a charge of 2.5¢ per call and 1¢ per minute. Thus three different per call charges (zero, 2¢ and 2.5¢) and three different per minute charges (zero, 1¢ and 1.5¢) were observed, in four different combinations. While this amount of price variability is far from ideal, it does offer some hope of being able to estimate at least part of a demand curve, and of being able to separate the effects of per-call and per-minute charges.

Like Jensik (1979) and Wilkinson (1981), Park and Wetzel (1980) use monthly time-series data. Unlike the former authors, however, Park and Wetzel estimate a multivariate model, disaggregating the data by telephone exchange (Jacksonville, Clinton and Tuscola, the three different towns in which the experiment took place) and by type of service (single-party and multi-party residential). The presence of multi-party service complicates the model, because this type of service remained on a flat-rate pricing system throughout the experiment, and Park and Wetzel suspect (correctly, as it turns out) that there was some substitution from single- to multi-party service. This could happen either because people actually switched from one service class to the other, or because calls between single-party and multi-party subscribers would almost invariably be initiated by the latter to avoid usage charges. One of the objectives of the Park-Wetzel study is to estimate the actual repression effects of usage-sensitive pricing separately from its effects on substitution across service classes.

The model used by Park and Wetzel to explain single-party usage can be written as follows:

$$U_{it} = \alpha_i * \beta_t * \exp(-\eta_C P_{C,it} - \eta_M P_{M,it}) * \exp(-\delta_C P_{C,it} - \delta_M P_{M,it}) + \varepsilon_{it}$$

where U_{it} is usage in exchange i during month t, usage being defined either as number of calls or as number of minutes; α_i and β_t are constants which vary across exchanges and across months respectively, and have to be estimated; and $P_{C,it}$ and $P_{M,it}$ are the prices of calls and minutes respectively in exchange i at time t. It will be observed that in this model η_C and η_M cannot be identified separately from δ_C and δ_M. That is made possible by the equation to explain multi-party usage. Without going into details, we can write the latter as

$$U_{it} = \gamma_i * \beta_t * g_{it}(\alpha_i, \gamma_i, \eta_{it}, \delta_{it}) + \nu_{it}.$$

The function g_{it} in the multi-party equation is such that the increase in multi-party usage as a result of price changes is exactly equal to the fall in single-party usage through the second price term in the single-party equation. Thus if δ_C and δ_M were zero, there would be no increase in multi-party usage.

Using both calls and minutes data as the dependent variables, Park and Wetzel estimate the six-equation system described above by a version of multivariate generalized least squares. They allow for contemporaneous correlation of the error terms across exchanges and service classes, for seasonal heteroskedasticity, and for a form of first-order serial correlation. Assuming that their model is correct, the estimation procedures they use are certainly satisfactory.

The results of Park and Wetzel are quite consistent with those of Jensik and Wilkinson. They find that the tariff currently in effect in the three test exchanges (P_C=2.5, P_M=1.0) reduces the number of calls by 14.5% and the number of minutes by 18.8%, compared to flat-rate pricing. Of these reductions, 2.1% and 4.4% respectively are accounted for by substitution between single-party and multi-party usage. Thus the actual repression effect of usage-sensitive pricing is apparently somewhat less than earlier studies had suggested. Although no standard errors are attached to these estimates, the t-statistics on the parameter estimates from which they are derived suggest that they are really quite precise.

Although the Park-Wetzel estimates are probably the best currently available, their study does have a number of deficiencies. They apparently use nominal prices throughout, despite the fact that their data cover 52 months of the late nineteen-seventies, when the general price level was rising rapidly. Economic theory suggests that prices should have been deflated by some general price index. Doing so would give the price series a little more variation (which would certainly be desirable in this case), and would probably increase the estimates of how much usage responds to price.

Another serious deficiency of the Park-Wetzel study is its treatment of time-of-day discounts. The GTE experiment offered customers a 20% discount during the evening and on Sundays, and a 50% discount at night. This fact was ignored by Park and Wetzel, who simply used daytime rates in their equations. Repression would presumably have been greater if no time-of-day discounts had been available, so Park and Wetzel's estimates are presumably too small. It would be very difficult to quantify this assertion, however, because of the way the experiment was designed; since discount rates were a set percentage of daytime rates, there is no independent variation in the two series. Some light on this issue is shed by Jensik (1979), who notes that the reduction in telephone usage during discount periods was almost identical to the reduction during non-discount periods. The most plausible explanation for this appears to be that usage during discount periods is more price-sensitive than during the daytime, and that there may be some substitution between high and low cost calling periods. Since any usage-sensitive pricing scheme that is intended to set prices near marginal costs must involve time-of-day pricing, it would clearly be of interest to estimate demand equations by time of day, allowing for substitution across calling periods. Unfortunately, this would not seem to be possible with the GTE data.

The major deficiency of the Park-Wetzel study is that the model used is never adequately tested. Conditional on the functional form employed, estimates of how much usage responds to price are remarkably precise, suspiciously so, perhaps, in view of how little price variation there is in the data. Before using those estimates to forecast the effects of prices different from those observed during the experiment, one should do everything possible to verify the validity of the functional form. For example, one could estimate several alternative models with different functional forms for the effect of prices. If any of these alternative models fit as well as or better than the original, and had substantially different implications as to the effect of prices outside the observed range, then one would have to view the original model with great suspicion. My belief is that this would almost certainly occur, since the fundamental problems of demand curve estimation, which were discussed above and illustrated by Figures 1 and 2, apply with almost as much force to the work of Park and Wetzel as they do to that of Jensik and Wilkinson. In the absence of observations on prices over the whole range of interest, demand curves can never be estimated with any degree of confidence. This fact should always be borne in mind by the readers of studies like that of Park and Wetzel.

Some evidence on what happens to telephone usage at much higher prices than those charged in the GTE experiment is provided by Wong (1981). During 1969 and 1970, Mountain Bell of Colorado offered a service called METROPAC, which allowed customers to pay a fixed charge and make an unlimited number of calls within an extended calling area. In 1971 this service was dropped, and replaced by a revised service with a lower fixed charge, a sixty minute per month free calling allowance, and a charge of 8¢ per minute beyond the free allowance. For customers who chose to subscribe to both of these services, Wong compared the empirical distribution of calls and minutes under the two different tariffs. He found that the mean number of calls dropped by 63.5% and the mean number of minutes dropped by 76.7% after the introduction of measured service.

Wong's data are far from ideal for estimating the impact of measured service on local telephone calling. They exclude customers who did not find it worthwhile to subscribe to either plan, presumably because they made few phone calls outside their local area and within the extended area. They also exclude customers who found it attractive to subscribe to one plan, but not to both. Moreover, calls within an extended calling area are somewhat different in character than many local phone calls. Nevertheless, Wong's study is the only one which considers such a large change in price, and for that reason the results are of interest.

The U.S. CPI rose by 53% between January, 1971 and July, 1977. Thus a charge of 8¢ per minute in early 1971 is roughly equivalent to a charge of 12¢ per minute at the time the GTE experiment was undertaken. According to the Park-Wetzel estimates, a charge of 12¢ per minute (with no per-call charge) would reduce the number of calls by 47.1% and the number of minutes by 68.4%. These may seem like rather large reductions, but they are actually less than those recorded by Wong. Thus, although the estimates of Park and Wetzel cannot confidently be used to forecast usage changes for prices outside the range they observed, Wong's results do at least suggest that they have not over-estimated the repression effect of large price changes.

Demand Estimates Based on Household Data

If one is simply interested in the effects of price on local telephone usage, there is no need to utilize data for individual households. Time-series data such as those used by Park and Wetzel and, in the future, aggregate cross-section data, should be more than adequate for most purposes, and far cheaper to obtain and analyse than individual household data. However, there are some questions which are hard to analyse without access to the latter.

The largest existing study using household data is described in Brandon (1981). Data were obtained from several surveys of A.T.& T. customers in Chicago, randomly selected from certain exchanges with modern switching equipment, so that it was relatively inexpensive to measure telephone usage. Chicago already has a form of local measured service, with per-call but not per-minute charges, but these charges did not vary across the sample (except insofar as different customers have chosen different service plans, with different fixed charges and numbers of "free" message units). Thus Brandon and her co-authors are not concerned with the effect of price on local telephone usage, but simply with the effect of different household characteristics.

The Brandon study does not contain many results of much economic interest. There is apparently enormous variation in telephone usage across households, a rather small part of which (no more than twenty per cent) can be explained by household characteristics. Blacks seem to make more local phone calls than whites, but fewer calls to the suburbs. Larger households make more calls than smaller households, and households with teenage children make more calls than households of the same size without them. There is little relationship between local calling and income, although wealthier households do make more calls to the suburbs.

The major contribution of the Brandon book, in my view, is its discussion, in chapters 2 to 4 and 11, of the practical issues of designing a survey to study local telephone usage, and managing the data collected in an effective manner. Anyone who imagines that this type of study can be completed quickly, cheaply and without a great deal of effort on the part of a team of several people, should certainly read this book.

The study by Infosino (1980) is rather similar in spirit to the Brandon book. It uses survey data from California and Cincinnati to estimate models of local call demand. Once again, there is no price variation in the sample. Results are by no means identical to those in Chapter 6 of Brandon (1981), but are in many respects qualitatively similar.

The only available study using household data which is designed to study the effects of usage-sensitive pricing is the paper by Park et al. (1980), which uses data from the GTE experiment. Observations are available on 641 households for six separate months, three under the old flat-rate tariff and three, one year later, under the measured rate tariff. The model that Park and his co-authors estimate is

$$(C_{it})^{.27} = Z_i\beta + T_t(Z_i\alpha) + \mu_i + \nu_i T_t + \varepsilon_{it}.$$

Here C_{it} is the number of calls by household i in month t, Z_i is a vector of demographic variables associated with household i (including a constant term), T_t is a dummy variable which takes on the value zero in months when the flat-rate tariff was in effect and the value one in months when the measured rate tariff was in effect, and μ_i, ν_i and ε_{it} are error terms.

This is a rather sophisticated model, and several features are worthy of note. First of all, the dependent variable is raised to the power .27 in order to make the error terms roughly symmetric. Household data on telephone usage, even conditional on exogenous variables, always exhibits marked skewness. Since this is not satisfactory in a regression model, it must be eliminated somehow, and taking a power transformation of the dependent variable is a popular approach. In principle, this transformation should be estimated along with the other parameters of the model, but that requires use of maximum likelihood estimation, which for the type of model dealt with here would be extremely expensive. Thus Park et al. appear to have picked the number .27 on the basis of earlier work, and do not explicitly test it (although they provide some evidence to suggest that it is satisfactory). In contrast, Infosino (1980), used a value of .5, and Brandon (Chapter 6 in Brandon (1981)) used a somewhat approach which we will not discuss here.

The second interesting feature of the Park et al. model is that different households are allowed to respond differently to the imposition of local measured service; even the structure of the error terms is allowed to change. This sort of specification is quite unusual in regression analysis, and is made possible here only by the large sample size. Park et.al. find that different households do in fact respond differently (restrictions that they respond in the same way being sternly rejected), with large users reducing usage by more, even in percentage terms, than small users. As a result, most of the overall reduction in usage is due to reduction by the larger users. This implies that optional measured service plans may be relatively ineffective in cutting aggregate telephone usage, unless the optional fixed rate is so high that very few users choose to opt for it.

The study by Park et. al is a fine piece of work, but it does leave many questions unanswered, and the methodology could not necessarily be transferred to other data sets. Like Jensik and Wilkinson, Park et.al. simply model the effect of a once-and-for-all change in pricing regime; their results tell us almost nothing about demand curves for local telephone calls. More fundamentally, they implicitly assume that telephone customers have no choice as to the pricing regime they face. In the case of the GTE experiment, that is probably not a bad assumption, although it is not strictly true. When the pricing regime changed, customers could choose to disconnect their phones, or to change from single-party to multi-party service. In addition, since the regulatory authorities imposed a ceiling on the monthly bill, extremely large customers still faced a marginal cost of zero. Customers in the first two categories would not be included in the sample used by Park et al., implying that the sample could not be entirely random. Extremely large customers would be included in the sample, but since marginal phone calls cost them nothing, the model used should not have applied to them.

In principle, a model of telephone usage by households should estimate choice of pricing regime and usage simultaneously. At the very least, households always have two choices: to have a telephone, or not. In North America, almost everyone chooses to have a telephone, and it may be empirically valid to ignore those who do not. In other parts of the world, on the other hand, many households choose not to have telephones, and a change in the pricing system might significantly affect this choice, and hence affect the characteristics of those with telephones. Under many existing and proposed measured service schemes in the United States, customers can choose from two or more different tariffs. For example, one tariff might involve a high fixed monthly charge and no charge for local calls, another might involve a lower fixed charge, a certain free allowance of "message units", and a charge for calls beyond the free allowance, and a third might involve a very low fixed charge coupled with charges for all calls. Obviously customers will not allocate themselves randomly across these different tariffs. Those who expect to make the most calls, and/or those who place the highest value on a phone bill that does not vary from month to month, would be most likely to choose the first option; while those who expect to make the fewest calls would be most likely to choose the last. Any analysis of usage data which does not explicitly take into account this endogenous choice by customers is likely to yield biased and inconsistent estimates of the parameters of interest.

To our knowledge, no study has yet been done to examine choice of tariffs by customers, either by itself or jointly with equations to explain usage. Designing such a study would be a major exercise, even if adequate data were readily available. It would be attractive to derive both choice-of-tariff and usage equations directly from a utility-maximizing framework (with uncertainty explicitly modelled, of course), but that would be a non-trivial exercise. On the econometric side, the invariably observed skewness of usage data would have to be dealt with, since existing models of discrete choice always assume normality or some other distribution which is quite different from what we observe. It is conceivable that an explicit utility-maximizing model might actually explain skewness (if, for example, a symmetric random variable entering the utility function in a certain way resulted in a skewed distribution of usage), but that is just a conjecture at this time. Thus applied work on the determination of tariff regime and usage by households would appear to be on the frontiers of existing economic and econometric theory.

Will Usage-Sensitive Pricing Pay for Itself?

One of the principal reasons for interest in the price-sensitivity of local telephone calls, is that charging for them is not costless. If consumers are not very sensitive to price, the increase in the sum of consumers' and producers' surplus that results from usage-sensitive pricing will not be great enough to offset the costs of that pricing scheme. In this section, we use the estimates of Park and Wetzel to shed light on this issue.

In its simplest form, the demand curve employed by Park and Wetzel may be written as

$$Q = Ae^{-\eta P}$$

where Q is total conversation time (in minutes per month, say), A is a constant which may vary from month to month, P is the price in cents per minute (the additional complication of per-call as well as per-minute pricing will be ignored here), and η is a parameter. The estimated value of η, according to Table 2 of Park and Wetzel, is .096; however, since their study used data from 1976 through 1979, that estimate should be adjusted downward to allow for inflation. On the other hand, there are some reasons to believe that the Park-Wetzel estimates may be too low. We shall therefore use several values of η, ranging from .06 to .10 .

The area under the above demand curve is

$$(A/\eta)e^{-\eta P} + PAe^{-\eta P};$$

the first term is the integral with respect to price (triangle (1) in Figure

3), and the second is price times quantity (rectangle (2) in Figure 3). The total cost of supplying Q minutes of local phone calls is assumed to be CQ, where C is the long-run marginal cost in cents per minute. Thus the sum of producers' and consumers' surplus is

$$(A/\eta)e^{-\eta P} + PAe^{-\eta P} - CAe^{-\eta P}.$$

Economic theory tells us that this will be maximized when P = C, something which is easily verified by differentiating the above expression with respect to P, setting the derivative equal to zero, and solving for P. What interests us here is to find out how much this expression will diminish when P is set equal to zero instead of to C; the difference puts an upper bound on the billing and measuring costs which would be acceptable under local measured service.

In Table 1, we tabulate the sum of consumers' and producers' surplus for P = C and P = 0, and the difference between them, for C = 1.0, 1.5, 2.0, 2.5, and 3.0, and for η = .06, .08 and .10. Since all quantities are proportional to A, we arbitrarily set A equal to unity.

The results in Table 1 are quite striking. The surplus gained from usage-sensitive pricing varies from .0294 cents if the cost per minute is only 1¢ and η = .06, to .4082 cents if the cost per minute is 3¢ and η = .10. If households make local phone calls totalling about six hours per month (so that A should be 360 instead of 1; see Brandon (1981), page 105), then the potential efficiency gain per household per month varies from 10.6¢ to $1.47. These figures should be compared with the costs of usage-sensitive pricing, accurate estimates of which do not seem to be available.

Figure 3

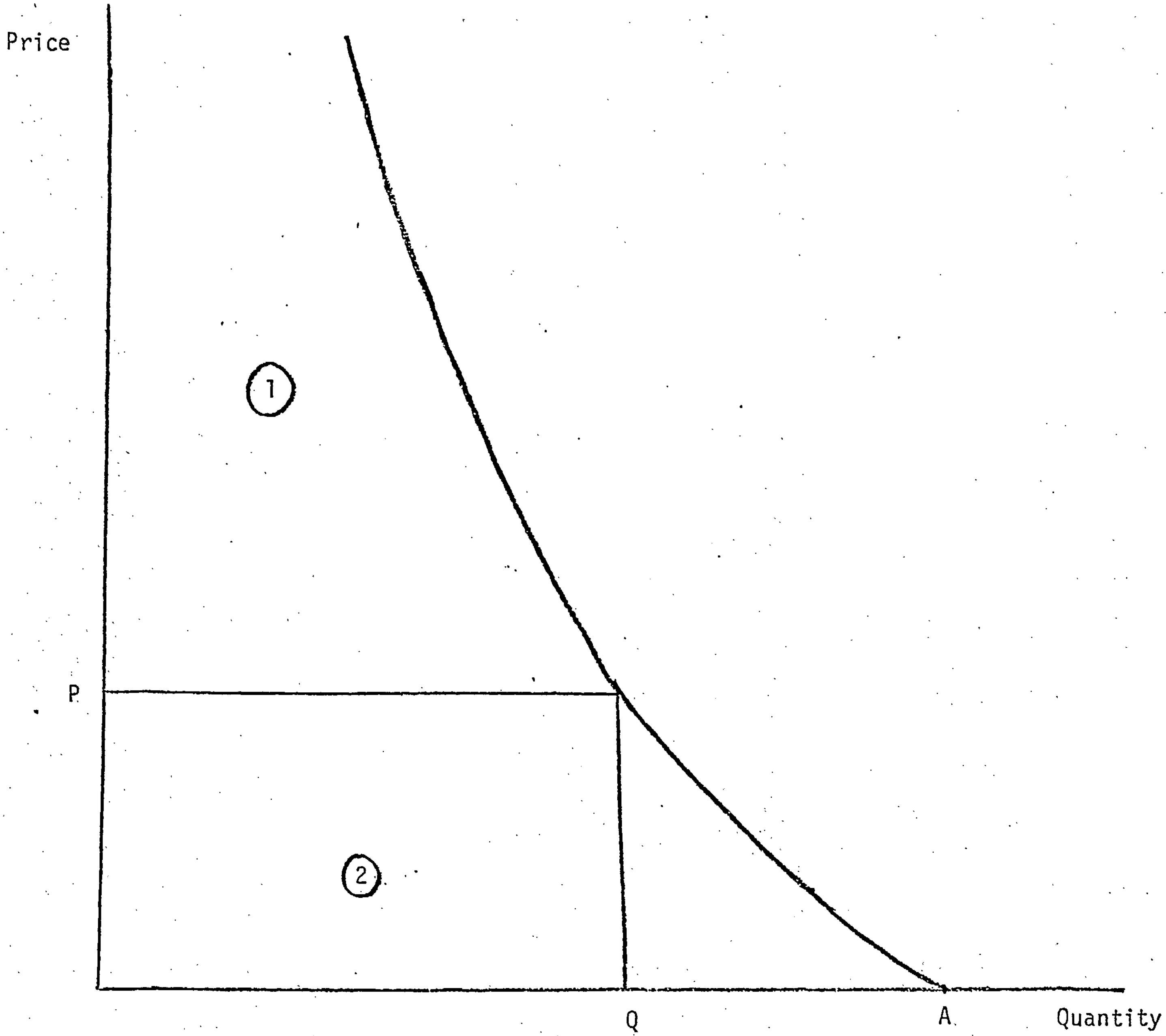


Table 1

Efficiency Gains from Usage-Sensitive Pricing

C	η	Surplus if P = 0	Surplus if P = C	Potential Gain
1.0	.06	15.6667	15.6961	.0294
	.08	11.5000	11.5390	.0390
	.10	9.0000	9.0484	.0484
1.5	.06	15.1667	15.2322	.0655
	.08	11.0000	11.0865	.0865
	.10	8.5000	8.6071	.1071
2.0	.06	14.6667	14.7820	.1153
	.08	10.5000	10.6518	.1518
	.10	8.0000	8.1873	.1873
2.5	.06	14.1667	14.3451	.1785
	.08	10.0000	10.2341	.2341
	.10	7.5000	7.7880	.2880
3.0	.06	13.6667	13.9212	.2545
	.08	9.5000	9.8328	.3328
	.10	7.0000	7.4082	.4082

Note: all figures are in 1981 cents, with demand curves normalized so that usage at a price of zero is one minute per month.

The preceding exercise is by no means definitive, since there is no reason to believe that the form of demand curve used by Park and Wetzel is correct, or that what was true of small Illinois communities in the late 1970's will be true of all Canada in the 1980's. One major problem is technological change. It seems extremely likely that innovative products (home computers, videotext, electronic banking, etc.) will greatly increase the demand for local phone calls. It also seems likely that such demands will be relatively price-elastic. If so, all the evidence gathered so far will soon be obsolete, and the potential gains from usage-sensitive pricing will be greater than those shown in Table 1.

Conclusions

Not a great deal is yet known about the demand for local telephone services. As time passes, more data from the United States should become available, and more studies should be completed using those data. That fact, coupled with the potential for technical change to make any study obsolete in short order, suggests that there would not be much point performing a measured service experiment in Canada. If such an experiment were to be performed, it should be carefully designed. In particular, it is important to confront telephone customers with several prices over the range of interest, if demand curves are to be estimated with any degree of confidence.

Concluding Remarks

We have reviewed theoretical, practical and econometric aspects of telecommunications services pricing. The particular aspects of telecommunications services pricing which make the general matter complex are jointness in production of services, heterogeneity of joint input, increasing returns to scale, network externalities from increased network coverage and peaking of uses. Also of particular importance is that the telecommunications industry has a very mixed market structure with the principal suppliers being regulated privately-owned firms. However, government firms and unregulated private firms also play an important role. Designing formulae for charging or price for services is from a practical point of view so difficult because the industry is experiencing such significant technical change on the supply side and such significant demands for new services on the demand side of the market. There are large challenges for the pricing analyst and for company decision makers in arriving at rates which maximize public benefits from the telecommunications system in the current period as well as in the longer term.

Is usage sensitive pricing and particularly local measured service a practical and useful alternative approach for Canadian telecommunications services producers? We are partial to USP and LMS *a priori* because they represent an application of the abstract notion of pricing at marginal cost. (Marginal cost pricing is a doctrine of long-standing in economics which has won acceptance because it has been shown to be efficient in the sense of providing a fixed output with the least resource cost on the input side.) Of great persuasiveness, also, in our judgment, is the fact that only one or two countries in the world are not practising or moving to practising many forms of USP including local measured service. Most European countries are moving from a message toll system of charges (price per call for telephone service) to the pulse system of local measured service. In the United States, all major companies surveyed (including AT&T or the Bell System and GTE) are moving toward measured service for network usage including of course LMS (Appendix I). The Federal Communications Commission in the U.S. suggested that USP might be the best approach for telecommunication companies in the future (Appendix II). One is struck by the fact that with new electronic switching devices, the cost of monitoring usage on a call by call basis is much lower than was the case with electromechanical switching devices. Rules of thumb (such as "value of service" pricing in the telecommunication and transportation industries) for pricing are used by producers because the cost of arriving at the appropriate cost-based charge is relatively high. Now the new technologies make the costs of precise monitoring of usage very low and appropriate usage-based charging schemes can be implemented. We note finally that New York City and most business users in large U.S. cities have been operating exclusively with some form of LMS since at least 1974. LMS is a practical alternative in the long run.

Should we switch to LMS tomorrow, to a scheme of measured service tied to costs of provision, monitoring and billing? There would be an altered pattern of usage by subscribers. Heavy users would curtail usage, presumably. It is most difficult to predict the new patterns of usage accurately and in consequence to predict the effect on individual and social welfare. Existing empirical work, including experiments by GTE in Illinois, provides a paucity of information on elasticities of demand for telephone service by household. Per call by a residential subscriber, existing information on demand elasticities is compatible with welfare gains from LMS of between .5¢ and 10.5¢, we noted in Part III above. The chief deficiency of these results is that they are based on estimates which were derived in situations in which price variation was relatively small. As we see it, the relevant context for developing usage based prices is the medium term future in which we foresee massive changes in the system or in capacity and equally massive shifts in system demands as new services come on line and are adopted on a mass scale. Pricing is after all a procedure for rationing supply or capacity and transferring subunits of capacity to users willing to pay. The rate at which new capacity and new demands will come on line seems most difficult to predict.

With regard to LMS, different suppliers of services have devised different charging schemes. The pulse system, common in Europe, aggregates calls by pulses incurred as a subscriber uses his telephone line. Monitoring is by usage but recording at the end of a billing period does not distinguish one call from another. One simply observes than one has incurred so many pulses per month. The U.S. approach involves a full monitoring and recording of each call by duration, origin and destination, time of day, and date. However, charge formulae differ across telecommunication service suppliers. GTE appears to favor a charge based on the four elements of the call regardless of the type of user. AT&T seems to be committed to options for users. One can elect to be charged a flat rate at a relatively high charge if one wishes or one can be charged on a call by call basis, the charge per call being based on the four of only small long run value, and useful information will emerge from various markets in any case. Consequently, it does not seem useful to mount a large experiment at present. Once the rate of technological change becomes less rapid in the industry, however, a well-designed experiment might prove very useful.

Appendix I: Responses of four operating companies which sponsored a major conference on USP to a request for information about their plans for implementing USP. Measured service for local rates appears to be a certainty for all companies surveyed. Note particularly that the Bell System (AT&T) is committed to local measured services options.

GTE Service Corporation

One Stamford Forum
Stamford, Connecticut 06904
203 357-2000

February 24, 1981

Professor John M. Hartwick
Department of Economics
Queen's University
Kingston, Canada
K7L 3N6

Dear Professor Hartwick,

Thank you for your letter inquiring about GTE's LMS plans. I understand that you were at a briefing by Jim Alleman on LMS and GTE's plans in this area. The situation has not changed significantly since that meeting. GTE is actively considering and working toward implementing LMS in its exchanges; however, more information is required, e.g. information on the cost sensitivity of exchange plants and the cost measurement system -- before a final committment to implement is made.

Research indicates that, if LMS "proves" in, it is most beneficial when it is applied to all classes of services within an exchange.

If I can be of further assistance, please contact me (203-357-3521) or Jim Alleman (203-357-2391).

Very truly yours,

G. Cohen

Gerald Cohen

GC:mar

UNITED TELEPHONE SYSTEM, INC.
BOX 11315 / KANSAS CITY, MISSOURI 64112

WRITER'S DIRECT DIAL NUMBER
(913) 676-3231

January 20, 1981

Prof. John M. Hartwick
Department of Economics
Queen's University
Kingston, Canada K7L 3N6

Dear Mr. Hartwick:

Re your January 6, 1981 letter inquiry: The fundamental planning process for the introduction of local measured service is currently underway in all United Telephone System Operating Companies. The exact introduction dates will depend on regulatory and consumer acceptance, measuring and billing systems being in place, and employee readiness. At present, it appears our first offering of local measured service to our customers will be by mid-1982.

Some of our companies may offer local measured service optionally for residence customers, and non-optionally for business customers. However, it is envisioned that our early offerings will be on an optional basis to both residence and business customers. But, charges for local usage will be levied to all users on an equitable basis. That is, the same rate will apply to calls of equal duration, distance and time-of-day, regardless of the customer classification.

We are attempting to evaluate digital switching technology capacities and implementation plans to meet these new measured offerings. However, where older switching systems may require local call recording outboard systems, investigation of such systems capacities and configurations are currently being pursued.

Sincerely,

George N. Fuciu

George N. Fuciu
Staff Director
Revenue Planning

GNF:crg

Continental Telephone
Service Corporation

56 Perimeter Center East
Atlanta, Georgia 30346
(404) 391-8057

Marvin W. Krehmeyer
Vice President - State Regulatory Matters

February 4, 1981

Professor John M. Hartwick
Department of Economics
Queen's University
Kingston, CANADA

Dear Professor Hartwick:

This is in response to your letter dated January 6, 1981. Continental Telephone has adopted a deliberate approach to the implementation of local measured service within its service areas due to the fact there are many small exchanges, and there is some uncertainty as to whether LMS will be cost effective in these situations. At the same time, Continental believes that there are presently good and serious reasons for LMS in certain selected exchanges, such as those where EAS is a problem or where the likelihood of neighboring large operating companies converting to measured service is very likely.

In those cases where LMS may be feasible within the Continental System, it would apply for both business and residential customers. Continental's initial offerings of LMS will be restricted to digital offices, and in this regard, current plans call for installation of digital equipment in 40% of our offices serving over 50% of our customers by the mid-eighties. This technology is only one factor which would influence the rate of implementation of LMS within the Continental System.

I trust that the above information will provide you with the data requested in your inquiry.

Very truly yours,

MWK/ph

H. E. Harvey, Jr.
Director—Tariffs and Costs

AT&T

American Telephone and
Telegraph Company
295 North Maple Avenue
Rm 17-6356H1
Basking Ridge, N. J. 07920
Phone (201) 221-8312

February 2, 1981

Mr. John M. Hartwick
Professor of Economics
Queen's University
Kingston, Canada
K7L 3N6

Dear Mr. Hartwick:

Larry Garfinkel has asked me to answer your letter of January 6, 1981. The Bell System's plan is to move ahead expeditiously with the implementation of Local Measured Service (LMS). We consider it essential to the preservation of universal telephone service and to the future economic viability of local telephone companies in this country. I have enclosed two documents - the Bell System's plan for Measured Service and a current status report - that I think will answer most of the questions posed in your letter.

The speed with which Measured Service can be made available in particular localities naturally depends to some extent on the local costs of its implementation. It is most cost effective in large metropolitan areas, which have been assigned the highest priority in our current implementation efforts. While it is true that the cost of measuring equipment is higher in offices with step-by-step switching technology than in offices with Electronic Switching Systems, the cost of measuring equipment is only one part of the total costs of implementing Measured Service, the rest of which do not significantly vary as a function of switching technology. Moreover, the cost of measuring equipment for step-by-step offices diminishes every year. So, while in certain localities current switching technology does retard the rate of implementation, the number of customers affected is relatively small.

Should you have any further questions, feel free to call Fred Mitchell, who coordinates our implementation program, on (201) 221-7326.

Sincerely,

H E Harvey Jr

Attachments

THE PLAN

The following text details "THE PLAN" for attainment of the System goal for Measured Service by 1985.

Jurisdictions are in varying stages of Measured Service development and obviously will reach the System goal using differing strategies and time sequences. Although some objective guidelines for implementation are recommended, no single strategy can be regarded as a panacea for inception of "THE PLAN." Section (B.5) is devoted to the discussion of strategies.

THE PLAN (1979-1985)

Create a new basic exchange pricing structure:

- At least three residence options should be available to customers; Premium Flat Rate, Standard Measured with an allowance and Low Use Measured without an allowance (or very small allowance).

- Flat rate business service offerings should be withdrawn and replaced with a non-optional measured offering, that will include all Measured Service pricing elements in its design. All business service offerings will be priced to cover their direct costs.

Exchange pricing will continue to sustain "Universal Service."

- Basic service is the capability for information transfer including access to and from the local and toll network, and includes appropriate maintenance.

- Local usage will be "unbundled" from all measured access lines and priced at levels to provide a contribution to Measured Residence Access. The premium flat rate residence option should be priced so that in the aggregate all flat rate customers will generate revenues sufficient to match the costs of providing this type of service.

- The rate structure should include the four elements of Measured Service pricing; frequency, duration, time-of-day and distance.

Rationale for the frequency element will relate to economic set-up costs. Duration may be supported by data displaying extremes and variations in holding time. Time-of-day pricing may be represented as critical to curtailment of investment growth, in conjunction with the desire to shift peak messages to off-peak time periods, using pricing for traffic balancing with an alternative to offset rate increases to some degree. Distance is vital in fragmenting large flat rate calling areas and demonstrates consistency with costs and customer perceptions of value.

- Large flat rate calling areas should be reduced in size to include no more than a customer's home exchange, zone or wire center and adjacent exchanges, zones or wire centers.

- Optional Calling Plans (OCP) designed with elements of Measured Service pricing may be used as a bridge for calling to areas that previously were included in the flat rate local calling area.

The OCP design should be integrated with the overall Measured Service Plan to permit a smooth flow from local schedules at the local/MTS interface. Existing OCP's should be redesigned to meet this criteria as appropriate. See Optional Calling Plans (orange book) published in 1978.

- Introduction of Measured Service should be initially directed to metropolitan areas and major exchanges of each jurisdiction.

Facility planning should contemplate availability of Measured Service that includes all pricing elements (frequency, timing, distance, and time-of-day) even when implementation plans may contemplate the introduction of each on a more gradual basis.

- To the degree that Statewide introduction of Measured Service is not required by commissions, economics will probably delay introduction in many small outstate exchanges, where the unit cost of measurement may be highest and the benefits lowest.

- It is contemplated that the differentials between Residence Standard Measured Service and existing flat rate offerings will be increased over time, encouraging more customers to select Standard Measured Service as the most attractive offering.

- Low Use Service should be provided as an entry level offering to enhance the System goal of universal service. This offering will be maintained as an economic alternative to Standard Measured Service, and the usage crossover with standard Measured Service should occur at about 30-40 messages per month.

- Message Minute Mile (MMM) design will be utilized in local rate design. This design is to be priced on a dollar and cent basis rather than in the traditional Message Unit manner. Through this pricing technique, existing rate disparities at the Local-MTS interface may be reduced or eliminated.

- The Plan contemplates a unified interdepartmental planning process to include all elements of tracking and analysis discussed in subsequent sections of this guide.

Measured Service -Current Status

C&P Virgina	1/5/81 - The Commission approved the filing for an experimental rate design - Extended Area Calling - for five exchanges. The plans include the four usage elements and are MMM priced. For more details see "Trends", Issue No. 18.
NWB Minnesota	12/29/80 - Filed for optional measured service where facilities and equipment are available. The proposed offering includes the four usage elements and is MMM priced. The aggregated monthly access line charge is $7.80 for residential customers and $22.50 for business customers.
Mountain Idaho	12/1/80 - The Commission approved the filing for optional measured service. The proposed plan includes the four usage elements and is MMM priced. The aggregated monthly access line charge is $3.55 for residential customers and $13.00 for business customers.
South Central Kentucky	10/29/80 - The Commission approved the filing for optional residential measured service in the Frankfort exchange. The plan includes all four usage elements and is MMM priced. The aggregated monthly access charge is $7.17 and includes a $3.00 usage allowance.
Bell of Pa. Delaware	10/19/80 - Commission approved the filing for optional business measured service. The plan includes all four usage elements and is MMM priced. The Monthly aggregated access charges will range from $9.62 - $16.42 depending on the rate group.
South Central Mississippi	10/13/80 - The Commission approved the filing for optional residential measured service. The offering consists of the four usage elements and is MMM priced. The aggregated access line charge will be 80% of the monthly flat rate and will include a $3.00 usage allowance.
Southern Georgia	10/3/80 - Filed for optional measured service in the Atlanta exchange. The plan consists of a flat to home area (flat rate service in primary local calling area) and measured service beyond. The measured service area includes all four usage elements and is MMM priced.
PNB Idaho	9/22/80 - The Commission approved the filing for optional business measured service in the Lewiston ESS office. The offering includes all four usage elements and is MMM priced. The aggregated monthly access line charge is $9.75.

January, 1981

South Central Louisiana	9/22/80 - Commission approved the filing for optional residential measured service. The plan includes the four usage elements and is MMM priced. The monthly aggregated access charges range from \$8.25 - \$9.70 which includes a \$3.00 usage allowance.
Southern North Carolina	9/4/80 - Filed for optional low use residential measured service where facilities and equipment are available. The plan includes the four usage elements and is MMM priced. The monthly aggregated access line is between \$5.05 - \$5.95 which includes a \$2.00 usage allowance.
Southern South Carolina	9/4/80 - Filed for optional low use residential measured service where facilities and equipment are available. The tariff includes the four usage elements and is MMM priced. The monthly aggregate access line is between \$5.40 - \$7.20 which includes a \$2.00 usage allowance.
Northwestern South Dakota	8/28/80 - The commission approved the filing for optional low use residential measured service in Sioux Falls and Rapid City. The monthly aggregated access line charge is \$6.00 which includes a 40 call, frequency-only allowance. After the 40th call, the plan includes three usage elements (frequency, duration, and distance).
Bell of Pa. Delaware	8/19/80 - Commission approved the filing for optional residential measured service. The plan includes the four usage elements and is MMM priced. The monthly access line is priced at \$2.00 below the flat rate and will include a \$1.00 usage allowance.
Cincinnati	7/31/80 - Filed for optional measured service and includes the four usage elements and MMM pricing. The plan consists of two residential offerings - Standard Measured and Low Use and a Standard Measured business offering.
Southwestern Arkansas	7/25/80 - The Commission approved the filing for optional measured service with the four usage elements and MMM pricing. The disaggregated access line charges range from \$2.30 - \$3.30 (low use offering) for residence customers and from \$8.55 - \$12.25 (with a \$5.50 usage allowance) for business customers. The effective date for these offerings is planned for September 29, 1980.
South Central Tennessee	7/15/80 - Commission approved the filing for optional residence measured service. The plan includes two offerings: Measured Lifeline Service (frequency only) and Standard Measured Service (frequency, duration, and time of day). The services will be effective October 1, 1980.

Southwestern Oklahoma

7/11/80 - Filed for optional measured service. The plan includes all four usage elements and is MMM priced. The disaggregated access line charges range from $1.60 - $4.60 for residence customers, (low use offering) and $3.80 - $14.75 for business customers.

Southwestern Texas

7/7/80 - Filed for optional measured service. The plan includes the four usage elements and is MMM priced. The disaggregated access line charges range from $4.00 - $6.10 for residence customers, (low use offering) and between $10.35 - $20.30 ($3.80 usage allowance) for business customers.

Southwestern Kansas

6/30/80 - Filed for optional residence measured service. The plan includes the four usage elements and is MMM priced. The disaggregated access line charge ranges from $3.25 - $5.35 for the low use offering.

Southwestern Missouri

6/2/80 - Implemented optional measured service. The plan includes the four usage elements and is MMM priced. There are four rate groups, and the access charges range from $3.05 - $6.75 for residence and $6.75 - 15.00 for business customers.

Mountain Wyoming

5/31/80 - Filed for optional measured service in all ESS offices. The plan includes three usage elements (frequency, duration, and distance) and is MMM priced. The disaggregated monthly access line charges are $4.50 for residence customers and $14.50 for business customers.

Michigan

5/20/80 - Filed for low use measured service. The plan includes the four usage elements (metro areas) and is MMM priced. The proposed monthly access line charge is $5.00.

Southern Georgia

5/15/80 - Commission approved the filing for optional measured service in four pilot locations: Gainesville, Clermont, Flowery Branch, and Lula exchanges. The offering includes all four usage elements.

Mountain Arizona

5/14/80 - Filed for optional measured service in Phoenix and Tucson. The plan includes the four usage elements and is MMM priced. The access line charges for residence customers is $2.50 for Low Use and $6.00 ($5.50 usage allowance) for standard measured service. Business customers would pay $14.00 ($8.00 usage allowance).

Southern South Carolina	4/1/80 - Implemented an optional measured service pilot in the Florence exchange. The offering includes all four usage elements. The pilot will remain in effect for one year.
New Jersey	2/28/80 - Filed to extend the availability of its existing individual standard residence Measured Service to all customers as facilities become available. The monthly access line will be priced between the proposed 1FR and proposed Low Use offerings and would include a 75 message unit allowance.
New York	2/21/80 - Filed for the implementation of MMM pricing and to price flat rate service to reflect more closely its usage costs and to adopt a single statewide charge for the business Measured Service access lines.
PNB Idaho	12/31/79 - Implemented optional residence measured service in the Lewiston ESS office. The offering includes incidence, duration, and time-of-day discounts for $1.50 less than the flat rate service. A $3.00 allowance is included.
Southern North Carolina	12/1/79 - Implemented an optional Measured Service pilot in Davidson. The offering consists of two mileage tiers and includes all elements. A low use service is offered to residence customers along with the basic Measured Service.
New England New Hampshire	12/1/79 - Implemented optional residence measured service in four exchanges. The offering provides residence customers with a 30 message unit allowance (timed) at a monthly rate of $4.95. Measured business timing will be instituted on an exchange basis as facilities permit.
Wisconsin	11/29/79 - Commission ordered conversion of all flat rate business service to frequency-only measured business rate. Also included in the order was an increase in the additional local message charge from 6.7¢ to 7¢ per additional call.
C&P Maryland	11/13/79 - Filed on April 16, 1979, to implement timing for business customers. The proposal was denied on November 13. The Commission requested that any future timing proposal by the Company be cost supported

C&P
West Virginia

9/17/79 - Filed for optional Measured Service. Plan includes four-elements (MMM priced) as central offices are equipped on a statewide basis for business and residence customers. A two-element plan (frequency and duration) will be available in all other exchanges until the exchange is capable of providing the four-element offering.

PT&T
California

7/13/79 - The CPUC ordered a Measured Service offering with all elements and MMM pricing to replace the message unit service in the San Francisco-East Bay and Los Angeles areas. The access line charge for the standard measured service will remain at $3.75 but will now include a $3.00 allowance. The Zone Usage Measurement Schedule (ZUM) replaces the old message schedule. The schedule consists of 3 mileage bands and a 1 minute initial and overtime period.

Southwestern
Arkansas

7/17/79 - Implemented optional residence measured service in the Franklin central office of the Little Rock metropolitan exchange. The offering includes the four elements (frequency, duration, distance, and time of day) with a monthly access rate of $2.75 for one party service and $2.10 for two party service.

7/27/79 - Implemented optional residence measured service with frequency only in El Dorado and Jonesboro. The offering includes a 30 call allowance in El Dorado with a monthly access rate of $4.25. Jonesboro has no call allowance and a monthly access rate of $2.45.

PNB
Washington

7/1/79 - Implemented optional residence measured service in five exchanges with incidence, duration, and time-of-day pricing.

7/15/79 - Implemented optional residence measured service in three more exchanges. Offering includes incidence, duration and time-of-day for $1.50 less than the flat rate service. An allowance of $3.00 is included.

Southern
Georgia

7/1/79 - Implemented optional residence message service on a trial basis in the Atlanta exchange. This service is available to subscribers served by ESS offices at a monthly rate of $6.50 with a 25 call allowance. Each additional call is 10 cents per call.

Mountain
Idaho

7/1/79 - Implementated optional residence Measured Service on a one year trial basis in all areas served by ESS(7) The rate for the new offering is 60% of the applicable individual flat rate service. The plan includes frequency, duration, and time-of-day elements.

Michigan	3/2/79 - The Michigan Public Service Commission approved time-of-day discounts on interzone message rates for the state's three metropolitan area district exchanges (Detroit, Pontiac, and Grand Rapids).
C&P Virginia	Filed in 6/77 for withdrawal of flat rate business service including timing. Hearings ended in January, 1979, and the filing was withdrawn in February, 1979. Bill passed 3/79 to restrict MS, which included timing of calls unless a flat rate option was available.
Mountain New Mexico	2/1/79 - Implemented Optional Residence Measured Service in Albequerque San Mateo, Albequerque Academy, Los Alamos, and Santa Fe Southwest. This includes the four elements (frequency, duration, distance, and time-of-day) and MMM pricing.
Southern Florida	Implemented Optional Measured Service in five exchanges including all elements of MS.

Pembroke Pines/Hollywood	9/15/79
Orange Park	6/16/79
Metro - Miami	4/19/79
Delray Beach	4/ 1/79
Jupiter	12/31/78

Illinois	11/1/78 - Commission approved the offering of Optional Measured Residence and Business Service within 112 exchanges outside the Chicago metro area. Time-of-day pricing will be introduced.
New York	9/1/78 - Filed a two-phase plan for the implementation of a full Measured Service structure (Frequency, Duration, Time-of-Day, Distance). Also includes MMM pricing. Company proposes to introduce phase one early in 1980 and phase two in 1983.
PNB Oregon	8/1/78 - Optional residence measured service was implemented on the basis of availability of ESS equipment. Includes all elements except distance.
Ohio	7/1/78 - Completed withdrawal of flat rate business service statewide. Statewide optional residence completed 1977.
South Central Mississippi	7/1/78 - Announced pilot location for MS implementation. Commission ordered introduction of equipment necessary to provide MS capability.

Appendix II: Federal Communications Commission of the U.S. Docket 20003, Sept. 24, 1976, recommending Usage Sensitive Pricing as an approach for dealing with new market conditions for telecommunications services in the U.S.

FIRST REPORT & ORDER
DOCKET 20003
RELEASED SEPT. 24, 1976

229. Usage Sensitive Pricing and Interconnection. As found by T+E, another available alternative which could operate to neutralize indirect contribution losses, if any, due to interconnect competition is the initiation of usage sensitive pricing structures for intrastate exchange services. 161/ We set forth below examples of usage sensitive pricing measures which we believe merit study and briefly elaborate on some of our other findings with respect to usage sensitive pricing.

230. Most of the telephone companies currently utilize flat monthly rates rather than usage sensitive pricing for local exchange service. While there are many reasons for these circumstances, and they should be analyzed on a case-by-case basis, we note generally that flat rate pricing appears to provide little incentive for customers to make efficient use of service. Essentially, the cost to the customer of making a local telephone call, after he has subscribed to the service, is zero. Yet increased tarffic causes increased cost to the utility and requires more investment in traffic sensitive plant. Traffic sensitive plant may be as much as half of all exchange plant in the largest cities, and may be at least 20% of exchange plant in the smallest. 162/ Usage insensitive pricing also appears to increase the allocation of exchange plant to intrastate, thereby raising intrastate revenue requirements and rates. On the other hand, usage sensitive charges for exchange service (which could vary by time of day, distance, and duration) appear to provide an incentive for customers to conserve their use of the telephone system (and also to make such calls at off-peak hours when they do not add to the need for plant). Since there is no such suppressive incentive in the case of current flat rate local calls, the percentage of the time that exchange plant is in use that is being used for intrastate purposes is raised. Since exchange plant and expenses are allocated to interstate on the basis of relative minutes of use, the effect of usage insensitive pricing then is generally to raise intrastate revenue requirements and rates.

161/ T+E finds that several reasons exist for initiating usage sensitive pricing irrespective of the existence of interconnect competition and its alleged adverse economic impact, e.g., usage sensitive pricing encourages more efficient use of existing plant investment and a reduction in new plant investment. However, if it is viewed only as a response to interconnect competition such factors as market growth stimulation caused by interconnect competition also become important. As noted preivously, T+E finds it is likely that very modest market growth stimulation caused by interconnect competition, coupled with usage sensitive pricing, could operate fully to neutralize a revenue requirements shift, i.e., indirect contribution loss, resulting from interconnection. We agree with T+E. OTP claims usage sensitive pricing in conjunction with separations changes could fully neutralize the jurisdictional impact of interconnection.

162/ GTE lecture on usage sensitive pricing given to the Commission's Common Carrier Bureau staff February 27, 1976.

Conversely, the effect of usage sensitive pricing would generally be to lower intrastate revenue requirements and rates. A lower intrastate revenue requirement would make intrastate services, such as basic local exchange services, less dependent on any subsidy obtained from interstate via the separations mechanism. Thus, to the extent there may be an adverse effect from interconnect competition which results in a reduced interstate subsidy (i.e., reduced indirect contribution) to intrastate services including local basic exchange services, usage sensitive pricing could operate to neutralize this subsidy loss.

231. We cannot go further and conclude that basic local exchange rates are, in fact, higher today because of usage insensitive pricing since most states (California is the main exception) do not allocate the intrastate toll and local exchange revenue requirements on the basis of the jurisdictional separations formula. However, if the intrastate revenue requirement were allocated between toll and exchange services in a manner parallel to that provided for in the jurisdictional separations manual, the effect of usage insensitive pricing would be to raise local exchange rates, if a state allocated the intrastate revenue requirement on the basis of the proportion of costs directly attributable to local exchange and toll.

232. At the outset, we note that the potential cost savings due to more economically rational tariffs has not been widely explored in telephony. We received little information in this inquiry, but we have attempted to explore the matter through other sources. We find that there are potential cost savings through efficiency gains in the use of existing plant and potentially reduced investment for new plant. We recognize that the cost of metering equipment is an important factor and that the potential for cost savings may be limited in some exchanges or that some telephone companies may find it more advantageous than others. In any event, as noted above, usage sensitive pricing certainly merits thorough exploration by the telephone industry irrespective of its use as an ameliorating response to interconnect competition.

233. In designing an economically rational tariff structure, costs must be examined to determine their variability with usage. Some costs are not traffic sensitive and might be reflected in a customer charge or a "network access" charge. 163/ Some costs derive from the use of common equipment and vary with the number of calls. Other costs vary with the duration of calls. Some costs vary with distance and duration. A properly designed usage sensitive tariff should reflect these cost differences to the greatest extent possible. We have carefully evaluated testimony from the NYPSC 164/ and as a result we suggest that it would be useful to consider tariffs having some of the following elements:

1. A station equipment charge for all items of station equipment provided by the telephone company, i.e., "unbundled" terminal equipment charges.

2. A "network access" or "customer" charge to cover local loops, station wiring up to the connecting block, and dedicated central office plant (such as the main frame and the appearance of the line finder). Business office expenses could also be recovered herein.

3. A charge per call. On exchanges with little common equipment this could be waived. This charge could vary by time of day.

4. A charge per minute. This charge could also vary by time of day.

5. Charges for special services, such as local directory assistance.

163/ As noted previously, the "network access" charge could be the focal point for any direct or indirect subsidy to basic local telephone service that society finds is warranted. This assumes, of course, that terminal equipment services, such as PBX and KTS, are priced to cover their total costs.

164/ NYPSC Case No. 26775. *Supra*, n. 136.

6. On calls beyond the customer's local central office, the charges per call and per minute could vary with distance.

7. Non-recurring charges could be based on local practices, but we believe the following should be considered:

   a. a charge for the paperwork involved in a service order which could vary with what is ordered;

   b. a charge for the central office work involved in a service order which could vary with what must be done;

   c. a charge for connecting the customer's premises to the network, including a charge for basic inside wiring (up to the first connection on the premises) which could vary with the type of customer;

   d. a charge for additional inside wiring on a per foot or per jack (connecting block) basis. This amount could vary with the type of customer equipment to be connected;

   e. a charge for installing station apparatus, which could vary with the type of equipment to be connected.

We also believe that moves and rearrangements could be fully charged for at the time performed, and that non-recurring costs could be segmented from recurring charges. There is some evidence that business customers in many areas are being undercharged for moves and rearrangements because of averaging with residential customers. 165/ We also recommend a thorough examination of extended area service rates, particularly the relationship of such rates to associated costs. 166/

165/ Dittberner Associates, "Interconnection, An Economic Impact Analysis", Appendix D, OTP-Contract OTP-SE-72-113. Detailed analysis of this report by OTP indicates that the results are dependent on several assumptions and OTP found that most these assumptions were unsupported. We cite this report because it is the only study of the problem of differential cost of moves and rearrangements that has come to our attention. The issue of cross-subsidization in moves and changes is hard to study, since the cost that is not recaptured immediately must be recovered in the undifferentiated pool of recurring charges. Thus due to the capitalization of station equipment and connections into an undifferentiated account, it is not possible to determine whether the allocations presently made are justified on any of the bases we consider -- direct cost, incremental cost, fully distributed cost, etc.

166/ Extended area service is basically telephone service which includes service beyond a subscriber's exchange or zone at exchange rates rather than toll rates.

234. When it is realized that even during the business day there may be slack usage periods, e.g., before 10 AM, between 12 PM and 2 PM, and so forth, it appears that some cost savings could be affected by transferring calls to those periods. We have seen the results of a toll tariff in South Dakota that transferred a large amount of traffic by giving 20% discounts at 12 noon and after 4 PM. 167/ There may also be seasonal peaks in certain areas, and not just where tourism and recreation are important industries. Seasonal tariffs are well established and accepted in the electric utility industry, and we believe they should be explored in the field of telephony as well.

235. Our experience with respect to time of day charges has been successful. We have found demand, particularly residential demand, to be quite sensitive to price differences. We have not attempted to measure the cost savings accrued when we ordered the initial "after nine" rates in April 1963, but we do note that there was a substantial interstate net revenue increase of 10%, despite the transfer of $46 million in revenue requirements from intrastate to interstate in 1962. 168/

236. If usage sensitive pricing is initiated for basic local exchange service, it is likely that some subscribers, depending on the extent of their usage, could pay more for basic service while others could pay less. It has also been argued that usage sensitive pricing would raise the cost of service to the poor who are allegedly heavy telephone users. No evidence on this matter was submitted in the present inquiry. What evidence we have seen is for three exchanges in Illinois. While we are reluctant to generalize from so limited a sample, this evidence indicates

---

167/ South Dakota Study, "Interstate Long Distance Telecommunication Service, AT&T", 1969.

168/ Richard Gabel, Development of Separations Principles in the Telephone Industry, East Lansing, Michigan State University, Institute of Public Utilities, 1967, p. 103.

that telephone usage generally increases with income, and that the poor and elderly are relatively light users. 169/

237. The foregoing discussion is not intended to be exhaustive of all the ramifications relating to the initiation of usage sensitive pricing and possible separations changes. Usage sensitive pricing certainly deserves active study by the telephone industry, and any further consideration of competitive effects from interconnect competition must include also a look at usage sensitive pricing and separations as possible neutralizing mechanisms of any adverse economic impact that might arise. Therefore, we expect that the parties will address this subject comprehensively in the further stages of this inquiry.

169/ Illinois Commerce Commission, Docket No. 76-0069, Ex. No. 3, in Bell Ex. No. 21 there is presented an analysis of demand for basic residential exchange service and a associated demographic analysis undertaken for Bell by National Economic Research Associates, Inc. The purpose of the study was to investigate quantitatively the economic and demographic determinants of telephone availability including age, urbanism, education, race, region of residence and family type and especially income level. The study was based on reported availability, not use, of the telephone. Chief findings include: (a) availability tends to vary directly with income, age and urbanism, is higher among whites and lower in the South; and (b) availability does seem to be price sensitive, but the degree of sensitivity varies inversely with the income level; estimates of (arc) price elasticities vary from virtually zero (for households with annual income greater than $12,000) to .29 (for households with annual income less than $3,000); also varies with age and family type. T+E finds the study to be of questionable value for this inquiry. T+E indicates that at best, one can say price sensitivity seems to vary somewhat with income level, but the estimated variations at all income levels are very low. T+E notes that not only is telephone availability inelastic but almost completely so and that this would suggest (income effects aside) that modest changes in the charge for exchange service would have very little effect on the universality of service. T+E also identifi several technical problems with the study. For example, it finds there is an obvious deficiency in using reported telephone availability as a proxy for telephones in use. See T+E Deliverable A, p. 41.

Footnote

Section II

1. The first part of this section draws on material in Shelley [1980] and the second part on material in Anderson [1980].

References

Section I

Marginal Cost Pricing and the Second Best

Allingham, M.C., and Archibald, G.C., Second Best and Decentralization, *Journal of Economic Theory*, Vol. 10 (1975) 157-173.

Boadway, R. and Harris, R., "A Characterization of Piecemeal Second Best Policy," *Journal of Public Economics*, Vol. 8 (1977) 169-190.

Coase, R.H., "The Theory of Public Utility Pricing and its Application," *Bell Journal of Economics*, Vol. 1 (1970) 113-128.

Davis, O.A. and Whinston, A.B., "Welfare Economics and the Theory of Second Best," *Review of Economic Studies*, Vol. 32 (1965) 1-14.

Dusansky, R. and J. Walsh, "Separability, Welfare Economics the Theory of Second Best," *Review of Economic Studies*, Vol. 43 (1976) 49-52.

Farrell, J.J., "In Defence of Public-Utility Price Theory," (1958) in *Public Enterprise: Selected Readings*, ed. by R. Turvey (Penguin) 1968, 44-63.

Hotelling, H., "The General Welfare in Relation to Problems of Taxation and Railway and Utility Rates," *Econometrica*, Vol. 6 (1938) 242-269.

Lipsey, R.G. and Lancaster, K., "The General Theory of the Second Best," *Review of Economic Studies*, Vol. 24 (1956) 11-32.

McManus, M., "Comments on the General Theory of the Second Best," *Review of Economic Studies*, Vol. 26 (1958) 209-224.

Nelson, J.R. ed., *Marginal Cost Pricing in Practice*, Englewood Cliffs, N.J.: Prentice Hall, 1964.

Ruggles, N., "Recent Developments in the Theory of Marginal Cost Pricing," *Review of Economic Studies*, Vol. 17 (1949-50) 107-126.

Turvey, R. ed., *Public Enterprise: Selected Readings*, Penguin, 1968.

Peak Load Pricing

Bailey, E.E. and White, L.J., "Reversals in Peak and Offpeak Prices," *Bell Journal of Economics*, Vol. 4, (Spring 1973), 278-292.

Bailey, E.E., "Peak Load Pricing Under Regulatory Constraint," *Journal of Political Economy*, Vol. 80 (1972) 662-679.

Boiteux, M., "Peak Load Pricing," *The Journal of Business*, (1960).

Brown, G. and Johnson, M.B., "Public Utility Pricing and Output under Risk," *American Economic Review*, Vol. 59 (March 1969), 119-129.

Carlton, D. "Peak-Load Pricing and Stochastic Demand," *American Economic Review*, Vol. 67 (December 1977), 1006-1110.

Craven, J., "On the Choice of Optimal Time Periods for a Surplus-Maximizing Utility Subject to Fluctuating Demand," *Bell Journal of Economics*, Vol. 2 (Autumn 1971) 495-502.

Crew, M. and Kleindorfer, P., "Marshall and Turvey on Peak Load or Joint Product Pricing," *Journal of Political Economy*, Vol. 79 (November-December 1971).

__________ and __________ "Recent Contributions to the Theory of Marginal Cost Pricing: The Problem of Peak Loads," *Economic Journal*, Vol. 81, (June 1971).

__________ and __________ "Peak Load Pricing with a Diverse Technology," *Bell Journal of Economics*, Vol. 7 (Spring 1976), 207-231.

Drèze, J., "Some Postwar Contributions of French Economists to Theory and Public Policy," *American Economic Review*, Vol. 54 (June 1964).

Hirshleifer, J., "Peak-Loads and Efficient Pricing: Comment," *Quarterly Journal of Economics*, Vol. 12 (1958), 451-468.

Littlechild, S.C., "Peak-Load Pricing of Telephone Calls," *Bell Journal of Economics*, Vol. 1 (Autumn 1970) 191-210.

Mohring, H., "The Peak Load Problem with Increasing Returns and Pricing Constraints," *American Economic Review*, Vol. 60 (1970), 693-705.

Panzar, J.C., "A Neoclassical Approach to Peak Load Pricing," *Bell Journal of Economics*, Vol. 7 (Autumn 1976), 521-530.

Pressman, I., "A Mathematical Formulation of the Peak-Load Pricing Problem," *Bell Journal of Economics*, Vol. 1 (Autumn 1970) 304-326.

Steiner, P.O., "Peak Loads and Efficient Pricing: Reply," *Quarterly Journal of Economics*, Vol. 72 (August 1957) 465-468.

Turvey, R., "Peak Load Pricing," *Journal of Political Economy*, Vol. 76 (Februach 1968) 101-113.

Visscher, M.L., "Welfare Maximizing Price and Output with Stochastic Demand: Comment," *American Economic Review*, Vol. 63 (March 1973) 224-229.

Wenders, J.T., "Peak-Load Pricing in the Electric Utility Industry," *Bell Journal of Economics*, Vol. 7 (Spring 1976).

Williamson, O.E., "Peak-Load Pricing and Optimal Capacity under Indivisibility Constraints," *American Economic Review*, Vol. 56 (September 1966) 810-827.

__________, "Peak Load Pricing: Some Further Comments," *Bell Journal of Economics*, Vol. 5 (Spring 1974) 223-228.

Wilson, G.W., "The Theory of Peak-Load Pricing: A Final Note," *Bell Journal of Economics*, Vol. 3 (Spring 1972) 307-310.

Ramsey Pricing, Nonlinear Pricing and Cross Subsidization

Baumol, W., Bailey, E., and Willig, R., "Weak Invisible Hand Theorems on the Sustainability of Prices in a Multiproduct Monopoly," *American Economic Review*, Vol. 67 (1977), 350-365.

Baumol, W.J. and Bradford, D.R., "Optimal Departures from Marginal Cost Pricing," *American Economic Review*, Vol. 60 (1970) 265-83.

Boiteux, M., "Sur la Gestion des Monopoles Publics Asteints à l'Equilibre Bugetaire," *Econometrica*, Vol. 24 (1956) 22-40.

__________, "On the Management of Public Monopolies Subject to Budgetary Constraints," *Journal of Economic Theory*, Vol. 3 (September 1971) 219-240.

Brander, J.A. and Spencer, B.J., "Local Telephone Pricing: Two Part Tariffs and Price Discrimination," presented at a conference on Telecommunications in Canada: Economic Analysis of the Industry, Montreal, March, 1981.

Coase, R.H., "The Marginal Cost Controversy," *Economica*, Vol. 13 (1946) 169-82.

Dansby, R. and Panzar, J., "An Economic Analysis of Measured Service Options," presented at Telecommunications in Canada: An Economic Analysis of the Industry: Montreal, March, 1981.

Diamond, P.A., "A Many Person Ramsey Tax Rule," *Journal of Public Economics*, Vol. 4, (1975) 335-42.

Faulhaber, G., "Cross Subsidization: Pricing in Public Enterprises," *American Economic Review*, Vol. 65 (December, 1975) 966-977.

Faulhaber, G., "Cross Subsidization in Public Enterprise Pricing," in *Pricing in Regulated Industries: Theory and Application*, ed. by J.T. Wenders 1979.

Faulhaber, G.R. and Panzar, J.C., "Optimal Two-Part Tariffs with Self-Selection," Bell Laboratories Economic Discussion Paper No. 74, (1977).

Feldstein, M.S., "Equity and Efficiency in Public Sector Pricing: The Optimal Two-Part Tariff," *Quarterly Journal of Economics*, Vol. 86, (1972) 175-87.

Gabor, A., "A Note on Block Tariffs," *Review of Economic Studies*, Vol. 23, (1956) pp. 32-41.

Goldman, M.B., Leland, H.E., and Sibley, D.S., "Optimal Nonuniform Prices," Bell Laboratories Economic Discussion Paper No. 100 (1977).

Hartwick, J.M., "Optimal Price Discrimination," *Journal of Public Economics*, Vol. 9 (1978) 83-89.

Leland, H.E. and Meyer, R.A., "Monopoly Pricing Structures with Imperfect Discrimination," *Bell Journal of Economics*, Vol. 7 (1976) 449-62.

Lewis, W.A., *Overhead Costs*, New York: A.M. Kelly (1970).

Lewis, W.A., "The Two Part Tariff," *Economica*, Vol. 8 (1941) 249-70.

Littlechild, S.C., "Two-Part Tariffs with Consumption Externalities," *Bell Journal of Economics*, Vol. 6 (1975) 661-70.

Mayston, D.J., "Optimal Licensing in Public Sector Tariff Structures," in M. Parkin and A.R. Nobay (eds.), *Contemporary Issues in Economics*, Manchester University Press (1974).

Mirman, L.J. and Sibley, D., "Optimal Nonlinear Prices for Multiproduct Monopolies," *Bell Journal of Economics*, Vol. 11 (Autumn 1980) 659-620.

Mitchell, B.M., "Optimal Pricing of Local Telephone Service," *American Economic Review*, Vol. 89 (1977) 66-83.

Murphy, M.M., "Price Discrimination,Market Separation and the Multi-part Tariff, *Economic Inquiry*, Vol. 17 (1977) 587-99.

Ng, Y.K. and Weisser, M., "Optimal Pricing with a Budget Constraint - The Case of the Two-Part Tariff," *Review of Economic Studies*, Vol. 41 (1974) 337-45.

Oi, W.Y., "A Disneyland Dilemma: Two-Part Tariffs for a Mickey Mouse Monopoly," *Quarterly Journal of Economics*, Vol. 85, (1971) 77-96.

Ordover, J. and Panzar, J., "On the Nonexistence of Pareto Superior Outlay Schedules," *Bell Journal of Economics*, Vol. 11, (1980) 351-54.

Panzar, J.C., "The Pareto Domination of Usage-Insensitive Pricing," in Proceedings of Sixth Annual Telecommunications Policy Research Conference, edited by H.S. Dordiek, (1978).

Panzar, J. and Willig, R., "Free Entry and the Sustainability of Natural Monopoly," *Bell Journal of Economics*, Vol. 8, (1977) 1-22.

Ramsey, F., "A Contribution to the Theory of Taxation," *Economic Journal* Vol. 37 (March 1927) 47-61.

Roberts, K.W.S., "Welfare Considerations of Nonlinear Pricing," *The Economic Journal,* Vol. 89 (1979) 66-83.

Spence, A.M., "Nonlinear Prices and Welfare," *Journal of Public Economics,* Vol. 8 (1977) 1-18.

__________, "Multi-product Quantity-Dependent Prices and Profitability Constraints," *Review of Economic Studies,* Vol. 47 (1980) 821-41.

Stiglitz, J.E., "Monopoly, Non-linear Pricing and Imperfect Information: The Insurance Market," *Review of Economic Studies,* Vol. 44 (1977) 407-30.

Willig, R.D., "Pareto-superior Nonlinear Outlay Schedules," *Bell Journal of Economics,* Vol. 9, (1978) 56-69.

__________, "Consumer Equity and Local Measured Service," in *Perspectives on Local Measured Service,* edited by Baude et. al., (1979) 71-78.

General and Miscellaneous

Artle, R. and Averous, C., "The Telephone System as a Public Good - Static and Dynamic Aspects," *Bell Journal of Economics*, Vol. 4 (1973) 89-100.

Averch, H. and L.L. Johnson, "Behaviour of the Firm Under Regulatory Constraint," *American Economic Review*, Vol. 52 (1962) 1052-1069.

Baude, J., Cohen, G., Garfinkel, L., Krehmeyer, M., and Ogg, J., *Perspectives on Local Measured Service:* Proceedings of the Telecommunications Industry Workshop, Kansas City, Missouri, March, 1979.

Bailey, E.E., *Economic Theory of Regulatory Constraint*, Lexington, Mass.: D.C. Heath, 1973.

Baumol, W. and Klevorick, A., "Input Choices and Rate-of-Return Regulation: an Overview of the Discussion," *Bell Journal of Economics*, Vol. 1 (Autumn, 1970) 162-190.

Crew, M.A., ed. *Problems in Public Utility Economics and Regulation*, Lexington, Mass. D.C. Heath 1979.

Harberger, A., "Three Basic Postulates of Applied Welfare Economics," *Journal of Economic Literature*, Vol. 9 (1971) 785-797.

Hicks, J.R., "The Rehabilitation of Consumer's Surplus," *Review of Economic Studies*, Vol. 8 (1940-41) 108-116.

__________, "The Four Consumer's Surpluses," *Review of Economic Studies*, Vol. 11 (1943) 31-41.

Kahn, A.E., *The Economics of Regulation*, New York: Wiley, 1970.

Mathewson, F. and Quirin, G., "Metering Costs and Marginal Cost Pricing in Public Utilities," *Bell Journal of Economics*, Vol. 3 (1972) 335-339.

Mitchell, B.M., "Economic Issues in Local Measured Service," Rand Paper P-6322, 1979.

__________, B.M., "Local Telephone Costs and the Design of Rate Structures," presented at Telecommunications in Canada: Economic Analysis of the Industry, Montreal, March, 1981.

Sherman, R. and Visscher, M.L., "Rate-of-Return Regulation and Price Structure, in *Problems in Public Utility Economics and Regulation*, ed. by M.A. Crew, 1979.

Varian, H., *Microeconomic Analysis,* Norton, 1978.

Vogelsang, I. and Finsinger, J., "A Regulatory Adjustment Process for Optimal Pricing by Multiproduct Monopoly Firms," *Bell Journal of Economics,* Vol. 10 (1979) 157-171.

Wenders, J.J., *Pricing in Regulated Industries: Theory and Application,* Proceedings of an Economic Seminar at Keystone, Colorado, August 1978, published by Mountain States Telephone and Telegraph Company 1979.

Warskett, G. and de Fontenay, A., "The Regulatory Process Under Partial Information," presented at Telecommunications in Canada: Economic Analysis of the Industry, Montreal, March 1981.

Willig, R., "Consumers' Surplus Without Apology," *American Economic Review,* Vol. 66 (1976).

References

Section II

Alleman, J. H. [1977], The Pricing of Local Telephone Service, OT Special Publication 77-14, U.S. Department of Commerce.

Anderson, John [1980], "Retrofit of Other Electromechanical Offices", in What Telephone Companies Should Know About Local Measured Service (Chicago: Telephony Publishing Corp.)

Beigie, C. [1973], "Pricing of Telephone Services" in H. E. English, ed. Telecommunications of Canada, (Toronto: Methuen) pp. 100-132.

Cohen, G. and E. C. Beauvais, (undated), "The Demand for Local Exchange Service: Some Implications for Planning", (mimeo)

Garfinkel, L. and P. B. Linhart, (undated), "Revenue Analysis of Local Measured Telephone Service", Public Utilities Fortnightly, (forthcoming)

Infosino, W. J. [1980] "Relationships Between the Demand for Local Telephone Calls and Household Characteristics", The Bell System Technical Journal, 59, #6

Lichtenwalner, O. [1980], "The Dilemma-Requirements and System Selection, in What Telephone Companies Should Know about Local Measured Service (Chicago: Telephony Publishing Corp.)

Mitchell, B. M. [1978], "Optimal Pricing of Local Telephone Service", American Economic Review, Vol. 68, No. 4, September pp.517-537.

__________ [1979], "Telephone Pricing in Europe: Localizing the Pulse", Rand Corporation, P-6215.

__________ [1978b], "Pricing Policies in Selected European Telephone Systems", Rand Report, P-6169.

Ogle, E. G. [1979], Long Distance Please, (Toronto: Collins).

Shelley, James [1980], "Responsive Retrofit-Step by Step Offices", in What Telephone Companies Should Know about Local Measured Service (Chicago: Telephony Publishing Corp.)

Skelton, P.E. [1980], "Testimony" to Restrictive Trade Practices Commission of Canada, May.

References
Section III

Baude, Joseph A., *et. al.* (ed.), *Perspectives on Local Measured Service; Proceedings of the Telecommunications Industry Workshop*, Telecommunications Industry Workshop, Kansas City, 1979.

Brandon, Belinda B. (ed.), *The Effect of the Demographics of Individual Households on Their Telephone Usage*, Ballinger, Cambridge, Mass., 1981.

Infosino, William J., "Relationships Between the Demand for Local Telephone Calls and Household Characteristics", *Bell System Technical Journal*, 59 (1980), 931-953.

Jensik, John M., "Dynamics of Consumer Usage", in Baude *et. al.* (*op. cit.*), pp. 141-159.

Mitchell, Bridger M., "Optimal Pricing of Local Telephone Service", *American Economic Review*, 68 (1978), 517-537.

Opinion Research Corporation, "Attitude Research Among Residence and Business Telephone Customers 1975-1979", ORC Study Number 33604, 1980.

Park, Rolla Edward, Bridger M. Mitchell, Bruce M. Wetzel and James H. Alleman, "Charging for Local Telephone Calls: How Household Characteristics Affect the Distribution of Calls in the GTE Illinois Experiment", Rand Working Paper #WD-180-2-NSF, 1980.

Park, Rolla Edward and Bruce M. Wetzel, "Charging for Local Telephone Calls: Price Elasticity Estimates from the GTE Illinois Experiment", Rand Working Paper #WD-727-1-NSF, 1980.

Wilkinson, G.F., "The Estimation of Usage Repression Under Local Measured Service: Empirical Evidence from the GTE Experiment", paper presented at a conference on Telecommunications in Canada: Economic Analysis of the Industry, Montreal, March, 1981.

Wong, T.F., "Identifying Tariff-induced Shifts in the Subscriber Distribution of Local Telephone Usage", paper presented at a conference on Telecommunications in Canada: Economic Analysis of the Industry, Montreal, March, 1981.