A theory of interdependent demand for a communications service

Jeffrey Rohlfs

Bell Laboratories
Murray Hill, New Jersey

The utility that a subscriber derives from a communications service increases as others join the system. This is a classic case of external economies in consumption and has fundamental importance for the economic analysis of the communications industry. This paper analyzes the economic theory of this kind of interdependent demand. We begin by defining "equilibrium user set" as a set of users consistent with all individuals' (users and nonusers) maximizing their utilities. There are typically multiple equilibria at any given price, and which equilibrium is attained depends partly on the static model, partly on the initial disequilibrium conditions, and partly on the disequilibrium adjustment process. Some general properties of equilibrium user sets are derived. Then we turn our attention to some specific models based on simple characterizations of communities of interest. The implications for pricing are discussed, with special reference to the problem of starting up a new communications service (e.g., a video communications service).

1. Introduction

■ The utility that a subscriber derives from a communications service increases as others join the system. This is a classic case of external economies in consumption and has fundamental importance for the economic analysis of the communications industry. It suggests that although marginal cost pricing may be superior to allocated-cost formulae, it is still not completely appropriate.

This can be illustrated with respect to an historical policy of the industry: promoting universal service. This policy might be justified on the basis of marginal cost pricing, so long as new subscribers pay the incremental cost of expanding the system to accommodate them—even if they do not pay their "allocated"

Jeffrey Rohlfs received the A.B. in economics from Amherst College (1965) and the Ph.D. in economics from M.I.T. (1969). Currently, he is a Member of the Technical Staff at Bell Labs, where his research interests include applied microeconomics, micro theory, and econometrics.

In the course of this study, I talked to many people, and their knowledge and ideas contributed to much of the analysis in this paper. M. Wish has been a collaborator in some previous related work, and he has greatly influenced my thinking. My analysis of the general problem was greatly stimulated by a discussion I had with E. Gilbert, who developed some preliminary results about the maximum equilibrium set. I have also had many profitable discussions with W. Ballamy, J. Berrier, A. Ciesielka, D. Deutsch, A. Gersho, E. Goldstein, A. H. McKeage, D. Mitra, and R. Sanders. F. Sinden, W. Taylor, N. Valcoff, M. Wilk, and E. Zajac provided helpful comments on previous written and oral presentations of this material.

share of average costs. A still lower price, perhaps much lower, might be justified if the externalities are taken into account. The total benefits that *all* subscribers derive from the expansion of the service may be sufficient to justify the incremental costs—even if the new subscribers are unwilling to pay the entire incremental costs.

Recently Artle and Averous[1] made what appears to be the first published analysis of these externalities in communications.[2] They formulate a simple model in which the incremental utility of the service to an individual depends only on the number of telephone subscribers—not on who they are. This is the uniform calling model discussed in Section 3 of this paper. They also assume that the cost of providing telephone service depends only on the number of subscribers. This enables them to derive and interpret the necessary conditions for a social welfare optimum. Their expression has some important similarities (but also some differences) with the usual necessary conditions for a social optimum with respect to a pure public good.

The authors then use these notions to develop a dynamic demand model. They show that interdependent demand can sustain continual growth in a stationary population with stationary income. The mechanism is as follows. New subscribers join. This increases the incremental utility of the service and induces marginal nonusers to join. That in turn induces further growth, etc., etc. The authors offer this as a possible explanation for the continual growth of telephone service observed in all empirical studies of the industry.

Squire studies the problem using a somewhat different model.[3] He considers usage of the system as well as number of telephones, and assumes that the cost of providing the service is a function of these two variables. Squire specifies individual demand curves (based on a fixed number of subscribers) for incoming and outgoing calls. This enables him to develop an expression for optimal usage of the system, based on a modified consumer-surplus concept. He then derives the optimal price per call (charged only to the person making the call) consistent with this optimal usage. He finally develops an expression for the optimal size of the system and the price per telephone consistent with this optimum.

This paper makes a much more detailed analysis of the demand side of the market than attempted by Artle and Averous or by Squire. We begin by defining an "equilibrium user set" as a set of users consistent with all individuals' (users and nonusers) maximizing their utilities. A basic result is that there are typically

---

1 See [1].

2 An earlier attempt to model interdependent demand is Marris [2]. Marris develops a general theory of demand for new products but does not consider communications in particular. The interdependent aspects of demand have a much different interpretation in his analysis than in this paper or in Artle and Averous [1] or Squire [3] (discussed later in the text). Also, irreversibility plays a much larger role in Marris' analysis. Nevertheless, Marris develops some of the same concepts used in this paper; e.g., critical mass.

3 In [3].

multiple equilibria at any given price. For example, a very small equilibrium user set may be consistent with utility maximization, since the smallness of the user set in itself makes the service relatively unattractive to potential users. However, a much larger user set may also be possible for the same population at the same price. In this case the largeness of the user set would make the service attractive and allow a high level of demand to be sustained. In any planning (public or private) for the communications service, special attention must be paid to which equilibrium user set is likely to be attained.

The next section of this paper develops a general theory of demand. It derives the following results:

(1) The static model determines the attained equilibrium user set (at a given price) within a certain set of bounds.
(2) A possibly narrower set of bounds (for a given price) is derived, given the initial user set.
(3) Within the bounds defined in (2), the equilibrium attained (at a given price) depends entirely on the disequilibrium adjustment process.

The following two sections of the paper develop specialized models based on various simple characterizations of communities of interest. The simplest of all is the uniform calling pattern, which assumes that no one has any special community of interest (other than the entire population). This model is the only one in which the equilibrium theory can be developed in terms of the *number* of users, without paying attention to who they are. We can therefore define a demand curve, which turns out to have an inverted *U* shape. See Figure 1 on page 28.

Zero demand is a stable equilibrium for all positive prices. The upward-sloping part of the inverted *U* consists of unstable equilibria and constitutes the "critical mass" of the service (at any given price). If the critical mass is exceeded, demand expands to the downward sloping part of the inverted *U*. Points on the latter are stable equilibria and represent the maximum level of demand sustainable at a given price.

Unfortunately (for ease of analysis), the uniform calling pattern may not be very realistic. People typically belong to groups, each of which has a strong community of interest within itself. And they typically have a few principal contacts who alone account for a substantial part of their communication. These complications are briefly discussed in the section entitled "Nonuniform Calling Patterns."

The final section of the paper discusses some implications of the preceding demand analysis for supply and pricing of the service. An important distinction is made between viability of the service (existence of a nonnull equilibrium user set that can be served with nonnegative profits) and the start-up problem (how to attain such a user set, starting from a small or null initial user set).

Viable nonnull equilibrium user sets (if they exist) are always superior to the null set from a static point of view. We can compare such sets to determine the static social optimum or the overall market equilibrium corresponding to a static supply model. However, this kind of analysis is incomplete and may be misleading without consideration of the start-up problem. Achieving the static optimal user set may require ruinous (albeit temporary) promotional costs.

Appropriate solutions to the start-up problem depend in large part on the demand model. In the uniform calling model, the start-up problem is simply a question of getting beyond the critical mass. Community of interest groups may make the practical start-up problem much easier, but they also introduce some special problems. If an individual's demand is contingent on a few principal contacts' being users, there may exist many small self-sufficient user sets. These allow the possibility of a long-term introductory program, in which the seller gradually expands the size and number of such sets.

This paper presents only a limited discussion of costs and supply. The reason is that costs of a communications service are very complex and merit a separate study in their own right. This is a very fruitful topic for future research.

2. General theory of demand

■ Let the population consist of n individuals. As in Artle and Averous' work, we define a set of binary variables:

$$q_i = \begin{cases} 0 & \text{if individual } i \text{ does not subscribe to the communication service} \\ 1 & \text{if the individual } i \text{ does subscribe to the communication service} \end{cases} \tag{1}$$

for $i = 1, \ldots, n$.

We assume there are also m other goods in the economy. To model interdependent demand, we specify a pair of utility functions for each individual:

$$U_i^0 = U_i^0\,(r_{i1}, \ldots, r_{im}) \tag{2}$$

$$U_i^1 = U_i^1\,(q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n, r_{i1}, \ldots, r_{im}) \tag{3}$$

where

U_i^0 = Utility of individual i if he does not subscribe to the communications service,

U_i^1 = Utility of individual i if he does subscribe to the communications service, and

r_{ij} = Consumption of (noncommunications) good j by individual i.

Equations (2) and (3) implicitly assume independent utilities with respect to all goods in the economy *other* than the communications service in question. In addition, we make the usual monotonicity assumptions:

$$\frac{\partial U_i^k}{\partial r_{ij}} \geqslant 0 \text{ for all } j \text{ and } > 0 \text{ for some } j; \quad \text{and} \tag{4}$$

$$U_i^0 \leqslant U_i^1 \tag{5}$$

for all $i, k, q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n, r_{i1}, \ldots, r_{im}$.

We also make a specialized assumption applicable to a communications service:

$$\frac{\partial U_i^1}{\partial q_w} \geqslant 0 \tag{6}$$

for all $i \neq w$, $q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n, r_{i1} \ldots, r_{im}$. That is, a subscriber's utility never decreases as additional individuals subscribe (and none drop out).

This seems like a reasonable working assumption. We can, of course, imagine some exceptions; e.g., the value of the service to others would probably be lessened if a large number of life insurance salesmen subscribed to the service to solicit other subscribers. However, we assume that such occurrences are the exception rather than the rule—that, in general, the availability of a communications link is not detrimental to either party.

We assume utility maximization, which we analyze in two steps. (1) We evaluate the maxima of U_i^0 and U_i^1 (with respect to $r_{i1}, \ldots, r_{im}$) subject to individual i's budget constraint. Let us denote these maxima as $\hat{U}_i^0$ and $\hat{U}_i^1$. (2) We then compare $\hat{U}_i^0$ and $\hat{U}_i^1$ to see if the individual demands the communication service. This defines a demand variable for each individual:

$$q_i^D = \begin{Bmatrix} 0 \text{ if } \hat{U}_i^0 > \hat{U}_i^1 \\ 1 \text{ if } \hat{U}_i^0 \leqslant \hat{U}_i^1 \end{Bmatrix} \tag{7}$$

for $i = 1, \ldots, n$.

The basic methodology of this paper is to ignore interrelationships between the communications market and other markets and concentrate on relationships *within* the communications market. Thus, we make the *ceteris paribus* assumption that prices of all goods other than the communications service are fixed and that each individual has a fixed budget constraint. This allows us to express the demand variables as functions of price and the set of subscribers:

$$q_i^D = q_i^D\ (p, q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n) \tag{8}$$

for $i = 1, \ldots, n$, where $p =$ the price of the communications service.

It follows from previous assumptions that all the q_i^D are monotonically decreasing (equality allowed) with respect to p. That is, an increase in p can never change q_i^D from 0 to 1; a decrease in p can never change q_i^D from 1 to 0. However, a change in p may have no effect on q_i^D. It also follows from previous assumptions that all the q_i^D are monotonically increasing (equality allowed) with respect to all q_w $(w \neq i)$.

☐ **Equilibrium user sets.** Naturally, there is a correspondence between demanding the service and being a subscriber. We define an *equilibrium user set* as a set of users such that

$$q_i = q_i^D\ (p, q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n) \tag{9}$$

for all i.[4] Thus, in equilibrium all users demand the service; all nonusers do not demand it.

Equation (9) defines equilibrium with respect only to the demand side of the market. It describes user sets that are consistent with utility maximization at a given price. These constitute necessary but not sufficient conditions for an overall market equilibrium. The latter additionally requires that the user set and price be consistent with some specified model of supply behavior.

For fixed $p = \bar{p}$, equations (9) are a system of n equations in n binary variables. Such a system does *not* generally have a unique solution. In fact, unique solutions did not arise in any of the simple models investigated in this paper (except in the trivial case where price is so high that there can never be any demand at all).

Consequently, the equation

$$q = q^D, \tag{10}$$

where

$$q = \sum_{i=1}^{n} q_i \quad \text{and}$$

$$q^D = \sum_{i=1}^{n} q_i^D$$

may be indeterminate (for fixed $\bar{p}$). That is, it may either hold or fail to hold depending on which set of users constitutes the sum q.

For this reason, the general theory of interdependent demand cannot be developed in terms of the sum q. It is necessary to work with the individual q_i. The basic analytical concept is not the demand curve—i.e., equilibrium pairs (q, p)—but rather equilibrium user sets.

☐ **Disequilibrium analysis.** Given that several equilibrium user sets exist for a given price (*ceteris paribus*), it is important to know which ones (if any) are most likely to occur. This requires analyzing what happens if the market is initially in disequilibrium. Our procedure is as follows. We specify a very general disequilibrium adjustment process. We then investigate the extent to which the user sets resulting from this process depend on the static model, the extent to which they depend on the initial disequilibrium conditions, and the extent to which they are indeterminate, depending on a more detailed specification of the adjustment process.

In this section, we restrict our attention to the demand side of the market and assume a given price for the communications service. We further assume that adjustments of consumption in other markets can be made rapidly and costlessly. This seems like a reasonable simplifying assumption, allowing us to analyze dis-

[4] We can also define equilibrium user sets with respect to any given set of discriminatory prices; i.e., a set of users such that

$$q_i = q_i^D(p_i, q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n) \tag{9a}$$

for all i, where p_i is the price charged to individual i.

equilibria in the communications market without considering possible disequilibria in the rest of the economy.

Now suppose there is an arbitrary initial user set. It may be based on utility maximization for current or previous states of the world, past selling efforts of the supplier of the service, or anything else. We assume that adjustments to this user set occur according to the following adjustment process. (1) An individual in equilibrium ($q_i^D = q_i$) never changes his status from user to nonuser or vice versa. This is reasonable, since such a change would always reduce his utility (except in the knife-edge case where $\hat{U}_i^0 = \hat{U}_i^1$, in which case the change in status has no effect on utility). (2) The length of time an individual can remain continually in disequilibrium ($q_i^D \neq q_i$) is bounded. He eventually must change his status. This is also reasonable, since the change always increases his utility (except in the knife-edge case where $\hat{U}_i^0 = \hat{U}_i^1$, in which case the change in status has no effect on utility).

The adjustment process is essentially a model of utility maximization with inertia. It is very general in that it makes no assumption about the speed of adjustment. This speed may vary from individual to individual. It may depend on the user set or actions of the seller. Or it may change over time.

A limitation of this process is that it does not allow individuals to collude and subscribe together. This is relatively unimportant if an individual's demand is contingent on a large user set, since such collusion would be difficult with very large groups. However (as will be seen later) we do have to consider relaxing the assumption in models where an individual's demand is contingent on a few of his principal contacts' being users.

It is important to note that the adjustment process does *not* necessarily converge to an equilibrium user set. Consider the following example. A demands the service if and only if B is a user; B demands the service if and only if C is a user; C demands the service if and only if A is a user. Suppose the initial user set is A. One possible version of the adjustment process is as follows. C joins because A is a user. Then A disconnects because B is not a user. Then B joins because C is a user. Then C disconnects because A is not a user. Then A joins because B is a user. Then B disconnects because C is not a user. We are now back to the original user set, and the process can be repeated indefinitely.

Nevertheless, the user sets resulting from the adjustment process can be bounded, as shown in the following theorems.

Theorem 1: If the initial user set is the entire population, the adjustment process can only *remove* individuals from the user set; no individual can ever be added who has previously dropped out.

Proof: If the entire population is an equilibrium user set, no one is added or removed, and the theorem is satisfied. If the entire population is *not* an equilibrium set, let $r_1, r_2 \ldots$ represent the sequence of individuals who change status. (If individuals change status simultaneously, we list them in arbitrary order.) Now r_1 must be a removal (not an addition), since the entire population consists of users, and there is no one left to be added. Given that $r_1, \ldots, r_k$ are removals, r_{k+1} must also be a removal for the following reason. The only possible additions would be the individuals $r_1, \ldots, r_k$. But all of these dropped out (and therefore did not demand the service) when the user set contained the current user set. Thus, they cannot demand the service according to the monotonicity assumption.

It follows that all of the r_i must be removals. Q.E.D.

Theorem 2: If the initial user set is the entire population, the adjustment process converges to an equilibrium user set in finite time.

Proof: The process must converge in finite time for the following reason. All changes of status are removals. Since only n (the size of the population) individuals can be removed, there are at most n changes in status. These must all occur in finite time.

After all the changes in status occur, no user can fail to demand the service, for otherwise the process would continue. Moreover, no nonuser can demand the service because of the monotonicity condition. Thus, the final user set is an equilibrium user set. Q.E.D.

Theorem 3: If the initial user set is the entire population, the adjustment process converges to the union of all equilibrium user sets regardless of the order of removals.

Proof: Let X be an arbitrary equilibrium user set; let R be the equilibrium result after individuals $r_1, \ldots, r_k$ have been removed according to the adjustment process. X cannot contain r_1, since r_1 was removed when the user set was the entire population (and hence contained X). Since X does not contain r_1, it cannot contain r_2 either. (The entire population minus r_1 contains X-r_1.) Nor can X contain $r_3, \ldots, r_k$. Thus, $X \subseteq R$.

But X is an arbitrary equilibrium user set. Thus, R contains all equilibrium user sets. Since R is itself an equilibrium user set, it is the union of all equilibrium user sets. Q.E.D.

This set will hereafter be referred to as the "maximum equilibrium user set."

By entirely symmetrical reasoning we can establish the following. If the initial user set is null,

(1) the adjustment process can only add individuals, and no one is ever removed who previously joined;
(2) the adjustment process converges to an equilibrium user set in finite time;
(3) the adjustment process converges to the intersection of all equilibrium user sets, regardless of the order in which individuals are added.

This set will hereafter be referred to as the "minimum equilibrium user set."

It is important to note that the minimum and maximum equilibrium user sets need not be the same. For example, consider the following simple model. Each individual demands the service if three of his five principal contacts are users. The minimum equilibrium user set is null. There are no users; so no one has three of his five principal contacts as users; so no one demands the service. The maximum equilibrium set is the entire population. Everyone has all five of his principal contacts as users; so everyone demands the service. In addition, there may be any number of equilibrium user sets between these two extremes, depending on the distribution of principal contacts.

In this example, the static model tells us practically nothing about what equilibrium will actually be attained. We only know that it will be zero or 100 percent or somewhere in between. This is an extreme case, but in general the static model determines the actual equilibrium only within certain bounds—the minimum and maximum equilibrium sets. Moreover, in all of the models investigated in this paper, there exists the possibility that these bounds may be far apart. In a practical situation, this difference may mean the difference between marketing success and failure.

The above theorems show that if the initial user set is sufficiently large, convergence to the maximum equilibrium user set is assured (according to the assumed adjustment process). If the initial user set is sufficiently small, convergence to the minimum equilibrium set is assured. For intermediate initial disequilibrium user sets, the actual equilibrium attained may also depend on a more detailed specification of the adjustment process than given above.[5] It may be critical whether or not the disequilibrium nonusers subscribe before the disequilibrium users drop out.

☐ **Particular initial user set.** The minimum and maximum equilibrium user sets provide bounds on user sets that are possible for *any* given initial user for *any* version of the adjustment process described above.[6] This subsection provides bounds on user sets that can be attained from a *particular* initial user set for *any* version of the adjustment process.

Let S be an arbitrary initial user set. We now define the following two adjustment sequences.

Optimistic sequence

(1) First all nonusers who demand the service subscribe in arbitrary order, but no users drop out. This converges to the same user set $\overline{\overline{S}}$, irrespective of the order in which individuals subscribe. (Proof is analogous to that of Theorem 3.)

(2) Then all users who do not demand the service drop out in arbitrary order. This converges to the same user set S, irrespective of the order in which individuals drop out. Moreover, $\overline{S}$ is an equilibrium user set. (Proof analogous to Theorems 2 and 3.)

Pessimistic sequence

(1) First all users who do not demand the service drop out in arbitrary order, but no nonusers subscribe. This converges to the same user set $\underline{\underline{S}}$, irrespective of the order in

[5] P. 22.
[6] P. 22.

which individuals drop out. (Proof analogous to Theorem 3.)

(2) Then all nonusers who demand the service subscribe (in arbitrary order). This converges to the same user set $\underline{S}$, irrespective of the order in which individuals subscribe. Moreover, $\underline{S}$ is an equilibrium user set. (Proof analogous to Theorems 2 and 3.)

Now let $R_1, R_2, \ldots$ be a sequence of user sets resulting from applying an arbitrary version of the adjustment process to S. As previously discussed, this sequence need not converge to an equilibrium user set. However, we can place the following bounds on the sequence:

(1) $\underline{\underline{S}} \subseteq R_i \subseteq \overline{\overline{S}}$ for all i. This follows directly from the monotonicity assumption.

(2) After some finite period of time, $\underline{S} \subseteq R_i \subseteq \overline{S}$ for all i.

Proof: Let $x_1, \ldots, x_k$ be a sequence of individuals who drop out in part (2) of the optimistic sequence. Since x_1 does not demand the service given user set $\overline{\overline{S}}$, he cannot demand it given any user set R_i. Thus, if x_1 is a user, he is continually in disequilibrium. He must drop out in finite time and can never thereafter rejoin. Once x_1 drops out, we can apply the same reasoning sequentially to $x_2, \ldots, x_m$. Thus, after some finite period of time, all $R_i \subseteq \overline{S}$. The proof that $\underline{S} \subseteq R_i$ is exactly symmetrical. Q.E.D.

Thus, the optimistic and pessimistic sequences define bounds on user sets attainable from a particular initial user set. These bounds may (or may not) be considerably narrower than the bounds provided by the minimum and maximum equilibrium sets. In any event, within these bounds the equilibrium user set attained depends entirely on a detailed specification of the adjustment process.

☐ **Additive utilities.** In order to proceed further we must make more assumptions. To simplify the problem we propose a model of additive utilities. That is, we define a vector $\underset{\sim}{f}$ and a matrix $\underset{\sim}{V}$ such that

$$U_i^0 = f_i\,(r_{i1}, \ldots, r_{im}) \tag{11}$$

$$U_i^1 = f_i\,(r_{i1}, \ldots, r_{im}) + \sum_{j \neq i} v_{ij}\, q_i, \tag{12}$$

where v_{ij} $(i \neq j)$ is the incremental utility to individual i of a communications link with individual j, $(v_{ij} \geqslant 0.)$

The additive model assumes that these incremental utilities do not depend on consumption of other goods or on other communications links available to the individual. These do seem to be reasonable simplifying assumptions, but there are some problems with them. The growth of telephone service has had fundamental effects on social and business customs, and these would not be captured in an additive model. It has also resulted in substantial changes in communities of interest, which are assumed to be fixed in equation (12). However, the additive model would be commensurately better for analyzing smaller differences in market penetration or for analyzing a service that does not provide so revolutionary an improvement in communications as did the invention of the telephone.

Equation (12) also assumes that the service has no value except to communicate with others who have the service. The service is worthless if no one else subscribes. This assures that the null set is an equilibrium user set at any positive price.

This assumption sounds reasonable enough, but there are some possible exceptions. An individual may have noncommunications applications for the hardware. If the service is new, he may find it prestigious or derive self-satisfaction from being an innovator. However, these kinds of considerations go beyond the scope of this paper.

The additivity assumption is quite useful and allows us to derive a convenient expression for q_i^D as shown below.

The maxima $\hat{U}_i^0$ are defined by the *ceteris paribus* conditions and do not depend on anything in the communications industry. Maximizing equation (12) with respect to $r_{i1}, \ldots, r_{im}$, subject to individual i's budget constraint, we obtain

$$\hat{U}_i^1 = \hat{U}_i^0 - h_i(p) + \sum_{j \neq i} v_{ij} q_j \tag{13}$$

for some function h_i where $h_i(0) = 0$, $h_i(p) > 0$ for all i.

It follows that

$$q_i^D = \begin{Bmatrix} 0 \text{ if } \sum_{j \neq i} v_{ij} q_j < h_i(p) \\ 1 \text{ if } \sum_{j \neq i} v_{ij} q_j \geqslant h_i(p) \end{Bmatrix} \tag{14}$$

where $v_{ij} \geqslant 0$ for all i, j.

We also assume constant marginal utility of money for a given individual. This means that $h_i(p)$ is a linear function:

$$h_i(p) = b_i p. \tag{15}$$

We can therefore write equation (14) as follows:

$$q_i^D = \begin{Bmatrix} 0 \text{ if } \sum_{j \neq i} w_{ij} q_j < p \\ 1 \text{ if } \sum_{j \neq i} w_{ij} q_j \geqslant p \end{Bmatrix} \tag{16}$$

where $w_{ij} = \dfrac{v_{ij}}{b_i}$ for all $i \neq j$.

☐ **Further simplification.** Both the monotonicity and the additivity assumptions greatly simplify the problem (at some cost in realism). However, we still must deal with the matrix $\underset{\sim}{V}$ which is the size of the population squared. Thus for a city with a population of one million, $\underset{\sim}{V}$ would have one trillion entries. Clearly, further simplification is required. The following two sections consider some possibilities for breaking the problem down to manageable size.

3. Uniform calling pattern

■ The preceding section began by considering the problem in its full generality and considered some reasonable kinds of simplifying assumptions. We now take the opposite approach, beginning with a very simple model and then relaxing assumptions to make the model more complicated and realistic.

In this section we assume that all the (off-diagonal) elements in any single row of $\underset{\sim}{V}$ are equal. This implies that no one has any special community of interest other than the entire population. The *number* of subscribers affects an individual's demand, but he does not care who these subscribers are.

This may in fact be a reasonable approximation (for some purposes). We might reason that the individual communicates with a large number of people during the course of a year, many of whom he does not know in advance. The number of users may be as good a proxy as any for the incremental utility he derives from the service.

However, it is also true that most people belong to groups, each of which has a community of interest within itself. They also typically have a few principal contacts with whom they communicate more than with others. Thus, their demand for a communications service would depend on how many members of their community of interest group and which of their principal contacts subscribe to the service.

In any event, the uniform calling model seems like a good place to begin developing the theory. (This model is also adopted by Artle and Averous and Squire.)[7] It allows some strong results to be derived and provides some useful insights about interdependent demand. Results from the uniform calling model also provide convenient reference points for analyzing more complex models, which are briefly discussed in the next section of this paper.

The uniform calling model allows us to write equation (15) as follows for a large population:

$$q_i^D \begin{Bmatrix} 0 \text{ if } fw_i < p \\ 1 \text{ if } fw_i \geqslant p \end{Bmatrix}, \tag{17}$$

where $f =$ the user fraction (q/n), and

$$w_i = \sum_{j \neq i} w_{ij}.$$

This in turn allows individuals to be ordered in terms of their demand for the service. That is, if $w_i \geqslant w_j$ (i's demand exceeds j's), individual i is in every equilibrium user set that contains j.

□ **Demand curve.** Since individuals can be ordered as above, every equilibrium user set consists of *all* individuals (i) for whom $w_i \geqslant$ some K. Similarly, for any q, there is at most one equilibrium user set with q members; i.e., the q people with the highest values of w_i. (If more than one person has the minimum w_i in the user set, all persons with that w_i must be in the user set for it to be an equilibrium.)

---

[7] In [1] and [3], respectively.

Thus, every equilibrium user set can be uniquely characterized by q, the number of members in it. We can develop the equilibrium theory for this model in terms of the sum q, without specifying the individual elements q_i. In particular, we can define a demand curve; i.e., the locus of all the pairs (q, p) for which there exists an equilibrium user set (which would have to be the q people with highest w_i).

This gives us a convenient way of looking at the relationship between price and the equilibrium user sets. However, it is important to note that equation (10) is still indeterminate, and we must be careful in applying the demand curve in disequilibrium situations.

☐ **An example.** Before discussing the general properties of such a demand curve, let us consider a specific example. Suppose the population is large, and w_i is distributed uniformly between 0 and 100 over the population. For the marginal individual

$$w_i = 100(1 - f). \tag{18}$$

For an equilibrium at $0 < f < 1$, fw_i for the marginal individual must equal p. [See equation (17).] Thus, the demand curve is the locus of points where

$$100\,f(1 - f) = p. \tag{19}$$

As previously mentioned[8] the null set ($f = 0$) is an equilibrium user set for all $p > 0$. For an equilibrium at $f = 1$, p must be less than or equal to fw_i for all individuals. But the minimum of w_i and hence fw_i is 0. Thus, the only equilibrium is $p = 0$.

FIGURE 1

DEMAND CURVE FOR UNIFORM CALLING PATTERN

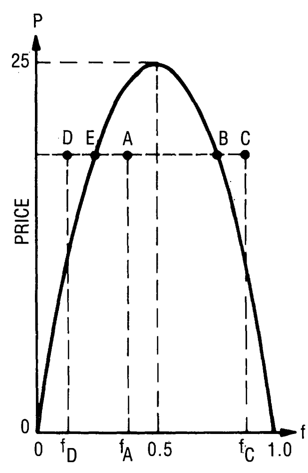


Figure 1 shows the demand curve. It consists of the entire positive p-axis plus an invested parabola going through (0,0) and (1,0) and having a maximum at (0.50,25).

The maximum equilibrium set is the right-hand side of the parabola for $0 < p \leqslant 25$; it is null for $p > 25$.

For small p, there is an enormous difference between the minimum and maximum equilibrium user sets. Thus, the actual equilibrium attained (for small p) depends critically on the initial disequilibrium conditions and the disequilibrium adjustment process.

☐ **Disequilibrium analysis.** The following analysis of disequilibrium is based on the adjustment process proposed above.[9] As before, we assume a fixed price and restrict our attention to the demand side of the market. We first consider the special case in which the initial disequilibrium users are those with the highest w_i. We then consider the general case by examining arbitrary perturbations from an initial equilibrium.

☐ **Initial users have highest w_i.** This subsection assumes that the initial user set consists of all individuals for whom $w_i \geqslant$ some K. This is necessarily true if the users form an equilibrium user set for any $\bar{p}$. Thus, the results apply to any disequilibrium brought about by a price change from an initial equilibrium.

[8] P. 26.

[9] P. 22.

Suppose we are originally in disequilibrium at A (in Figure 1), underneath the parabola. Given the user fraction f_A, the equilibrium price is higher than the actual price. All users are satisfied, but some nonusers would prefer to become users. If p remains constant, the user fraction will ultimately increase to B.

Suppose we are originally in disequilibrium at C. Given f_C, actual price exceeds equilibrium price, and f declines to B.

Suppose we are originally in disequilibrium at D. Given f_D, actual price exceeds the equilibrium price. So, f declines. As f declines, the discrepancy between actual and equilibrium price increases until the market achieves equilibrium at $f = 0$.

In all these cases, the order in which individuals join or drop out is immaterial. The optimistic and pessimistic sequences are equivalent, and all versions of the adjustment process converge to the same equilibrium user set.

In general, the positive p-axis and the downward sloping part of the parabola in Figure 1 consist of stable equilibria. The upward sloping part of the parabola consists of unstable equilibria.

The upward sloping part of the parabola can be regarded as a "critical mass" for the service. That is, for any positive price below the maximum of the parabola, the market must be forced to some initial disequilibrium beyond the critical mass before the service can grow by itself. The higher the price, the higher is the critical mass.

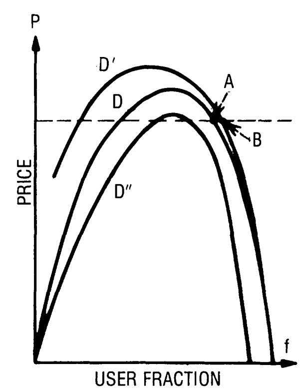
☐ **Arbitrary perturbations from equilibrium.** The preceding analysis does *not* necessarily apply for arbitrary initial conditions. An initial user set of y people may converge to very different equilibria depending on who those y people are and on a more detailed specification of the adjustment process (than given above[10]).

For example, suppose the initial user set of y people contains none of the y people with highest w_i. Suppose it contains the people with the next y highest values of w_i. Even though the initial market penetration is y, the market can achieve a critical mass of $2y$ if the y nonusers with highest w_i all subscribe before any of the initial users drop out (optimistic sequence). However, if all the initial users who do not demand the service drop out before any nonusers subscribe (pessimistic sequence), the market may fail to achieve a critical mass much lower than y.

The results of the preceding subsection do apply for small but arbitrary perturbations (in the user set) from an original stable equilibrium. For example, suppose the market is originally in the stable equilibrium at A in Figure 2, and a set (R) of nonusers subscribes. We can analyze this by constructing a demand curve conditional on all members of R being users (D' in Figure 2). The perturbation of R becoming users may cause additional users to subscribe. However, no matter what the adjustment process is (subject to the rules laid down above[11]), demand can never expand beyond B. And at B (or any point between B and A), only people in the *original* equilibrium set demand the service. Thus,

FIGURE 2

ARBITRARY PERTURBATIONS FROM EQUILIBRIUM



[10] P. 22.

[11] P. 22.

the market eventually goes back to A (so long as the price remains fixed).

In the general case the perturbation may involve users dropping out as well as nonusers joining. We analyze this by constructing D' as before and D'', demand conditional on those who drop out being nonusers. The market might converge to $f = 0$, if the perturbation is large and brings demand below critical mass. However, if the perturbation is sufficiently small, D'' will be sufficiently close to D that demand cannot go below critical mass. Thus, the market must return to the initial equilibrium.

This reasoning also applies to the stable equilibria at $f = 0$. If the perturbation is sufficiently small, D' will be sufficiently close to D that critical mass cannot be achieved, and the market must return to the initial equilibrium.

□ **General properties of the model.** Some properties of the above example apply generally to all uniform calling models. The entire positive p-axis always consists of stable equilibria. The demand curve always has an upward-sloping part, which constitutes the critical mass for the service for initial users sets with maximal w_i. It always has a downward sloping part (perhaps vertical), which consists of stable equilibria. However, both parts need not be unique, and the demand curve may be jagged. This allows the possibility of many stable equilibria for a given price.

4. Nonuniform calling patterns

■ This section considers some models that are more complex than the uniform calling model. Some specific results are presented, but they are naturally not so strong as those of the previous section. Our primary objective is to point out the analytical complexities in such models and suggest some ways of dealing with them.

□ **Community of interest groups.** Suppose the population consists of k groups ($k \leqslant n$). We assume that an individual has the same community of interest with everyone in the same group. However, this community of interest may be different for different groups. Mathematically, we assume that if individuals j and m are in the same group, $v_{ij} = v_{im}$ for all i.

□ **Disjoint groups.** The simplest case is disjoint groups. That is, $v_{ij} = 0$ unless i and j are in the same group. In such a model, we can consider each group as a separate population, and all the analysis of the previous section carries over. A critical mass can be defined for each group in terms of market penetration within that group. For given $\bar{p}$ the maximum possible number of stable equilibria is 2^k (unless the demand curve for some groups is jagged). The equilibria are characterized by which of the k groups achieve their critical mass.

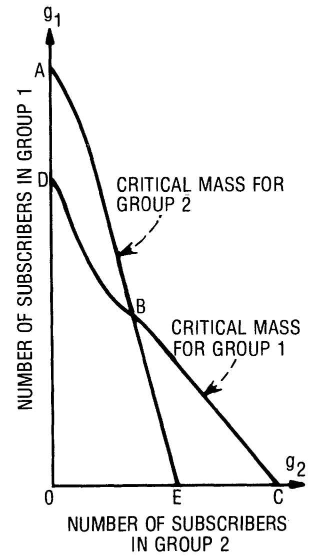
□ **Joint groups.** We now consider the case of joint groups, where v_{ij} does not necessarily equal zero for i and j in different groups.

In this model, the incremental utility of the service to an individual is a function of the g_i (number of subscribers in the ith group). Thus, we can define the critical mass for each group as a function of all the g_i (and p). This is illustrated in Figure 3 for the case of two groups.

If the initial user set in Figure 3 consists of individuals with the highest w_i in each group (a weaker condition than requiring initial users to have the highest w_i in the population), we have the following result. If all groups achieve their critical mass (initial market penetration outside *ABCO* in Figure 3), the service will expand to the maximum equilibrium set. If no group achieves its critical mass (initial market penetration within *DBEO* in Figure 3), the service will collapse to the minimum equilibrium set. If some groups achieve their critical mass but others do not (initial market penetration within *ABD* or *BCE* in Figure 3), the service may expand to the maximum equilibrium set, collapse to the minimum equilibrium set, or achieve equilibrium somewhere in between. Which of these occurs depends on the parameters of the static model, the initial market penetration in each group, and the disequilibrium adjustment process. An upper bound on the number of stable equilibria is 2^k (unless the demand surface is jagged). Each equilibrium is characterized by which of the k groups achieve their critical mass and which do not.

FIGURE 3

CRITICAL MASSES FOR MODEL WITH COMMUNITIES OF INTEREST



☐ **Further refinements.** Introducing further refinements into the model is straightforward. As before, we somehow divide the population into k groups such that within each group individuals can be ordered in terms of their demand for the service. For greater realism, we could have a large value for k, but that has the drawback of requiring us to deal with a k-dimensional quantity vector and to contend with the possibility of many different equilibria (for each price).

☐ **Few principal contacts.** An individual's demand may depend primarily on which of his few principal contacts are users. A basic analytical tool for studying such demand is the "self-sufficient" user set; i.e., a set of individuals, each of whom demands the service conditional on all others in the set being users. An equilibrium user set must, of course, be self sufficient, but the converse is not necessarily true. Someone outside the self-sufficient set may demand the service if everyone in the set has it.

All self-sufficient sets necessarily belong to the maximum equilibrium user set. Moreover, if the entire self-sufficient set is contained in the initial disequilibrium user set, then the entire self-sufficient set is necessarily part of the final equilibrium user set.

In any practical problem, we could never hope to have a complete empirical list of principal contacts. The way to proceed in such cases is to specify a probability distribution which indicates (approximately) how likely various configurations of principal contacts are. This leads to some interesting combinatorial analysis but goes beyond the scope of this paper.

5. Some implications for supply and pricing

■ Costs of providing a communications service depend on the constituency of the user set as well as its size. These costs are in themselves fully as complex as interdependent demand. Such complexities go beyond the scope of this paper, which is a study of interdependent *demand*. Nevertheless, it is useful to look at some implications of the preceding sections for supply and pricing of the service. We assume that the service is provided by a (regulated) monopoly and specify a very general cost function:

$$C = C(q_1, \ldots, q_n), \tag{20}$$

where C is strictly monotonic (equality *not* allowed) in all its arguments.

We investigate various kinds of pricing strategies, some involving short-run losses. However, we assume that in (long-run) equilibrium, the monopoly must earn nonnegative profits. Some of the pricing strategies considered involve discriminatory pricing, but we do not assume that perfect discrimination is necessarily possible. Indeed, perfect discrimination would surely not be possible in any realistic situation. Consequently, the nonnegative profit restriction may be inconsistent with Pareto optimality. It is nevertheless consistent with existing real-world institutions.

We can now make a crucial distinction for planning supply of a communications service; i.e., the difference between "viability of the service" and "the start-up problem." Viability is determined solely by the static model. It means that there exists a nonnull equilibrium user set that can be served with nonnegative profits. (We also refer to such a user set as "viable.") The start-up problem is a dynamic consideration. It refers to the costs and practical difficulties of attaining a viable user set, starting from a small or null initial user set.

From a static point of view, any viable user set is superior to the null user set (in the sense that the supplier of the service and all users are at least as well off as before and possibly better off. Nonusers' utilities are unchanged).[12] If there are several viable user sets, they can be compared to determine the social optimum (subject to the nonnegative profit restriction). We can also determine the overall market equilibria consistent with various static supply models.

However, this kind of static analysis is incomplete and may in fact be misleading. We must also consider the dynamic aspects; i.e., the start-up problem. If the initial user set is small or null, the static social optimum may require ruinous (albeit temporary) promotional costs. Thus, all things considered, a smaller user set, or perhaps even the null set, may be superior.

The remainder of this section discusses various possible solutions to the start-up problem. We consider the case of a new service (e.g., a video communications service) and assume that the initial user set is null. This is necessarily an equilibrium user set at all positive prices. Thus, the service, even though viable,

[12] Of course, this issue becomes much more complicated if we consider interrelationships with other markets and possible income redistribution.

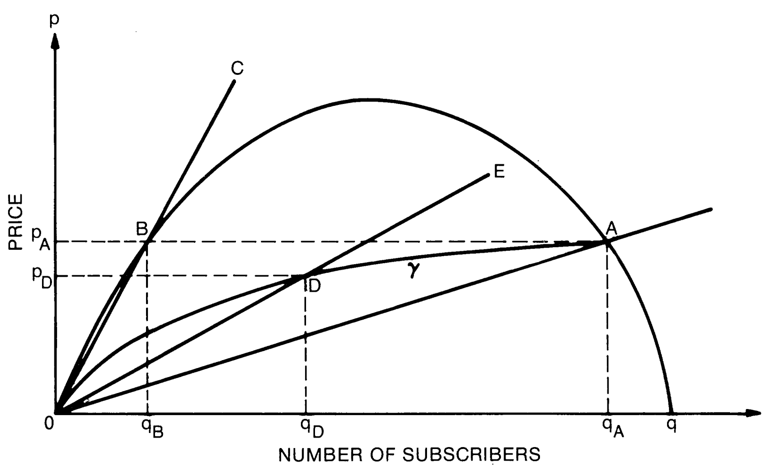
cannot get started by itself. It requires some positive action by the seller, probably involving temporary losses. The next three subsections consider some possible start-up strategies in the context of the various demand models we have studied in the previous sections.

Although we assume in this analysis that the product is viable, it is worth noting that in real life the seller would have no such guarantee. There is always a risk involved in introducing a new product or service. The seller generally faces certain losses when the product or service is first introduced and the *prospect* of future profits. In a regulatory environment, the supplier must consider how the regulators will respond if the service succeeds and how they will respond if it fails.

☐ **Uniform calling pattern.** Let us assume that the demand curve is a (nonjagged) inverted U, and suppose the desired nonnull equilibrium set is A in Figure 4. The long-run optimal price is P_A, but some method must be contrived to get beyond the critical mass.

FIGURE 4

SOLVING THE START-UP PROBLEM



☐ **Direct approach.** The most direct approach is to give the service free to a selected group of people for a limited time. For this method to succeed, the initial user set must, of course, be sufficiently large to achieve critical mass. Half measures are worse than useless. If critical mass is not achieved, the whole effort will be a complete failure, and demand will eventually contract to zero.

The success of this approach may also depend on how the initial user set is selected. The discussion above[13] shows that the optimal initial user does not necessarily consist of those with maximal w_i. (In the example, an alternative choice of y people allowed a critical mass of $2y$ to be achieved.) However, if any other set is chosen, there should be some assurance that the initial users will not discontinue the service before those with higher w_i subscribe.

[13] P. 29.

If they do discontinue, the start-up effort will have failed to achieve anything.

☐ **Low introductory price.** Another way to start up the service is to have a low introductory price. This price could then be raised as the number of subscribers increased. There are any number of ways this could be accomplished. We can represent the introductory program as an expansion path in (q,p) space; e.g., γ in Figure 4. The expansion path shows how p increases as q increases.

The expansion path must pass through the origin to get the service started in this simple model. If, in addition, the expansion path is always concave downward, the introductory program has a very desirable property. Regardless of the order in which individuals enter the market, no individual ever subscribes and then later discontinues service after price and quantity rise along the expansion path. The proof is as follows. Suppose an individual enters at D. That means he is willing to pay p_D to have the service available for q_D of his communication. It follows from the uniform calling assumption that he would be an equilibrium user at *any* point along the straight line OE through D. But the concavity assumption assures that the expansion path γ is always below OE for $p > p_D$, $q > q_D$. Thus, if the individual joins at D, he cannot drop out at any point on γ to the right of D. Q.E.D.

The above condition can be very important if the cost of connecting an individual to the network is large.

Perhaps the most interesting program with a low introductory price is usage-proportional pricing. The remainder of this subsection investigates that plan under various assumptions.

We first consider the possibility that all subscribers have equal usage, proportional to the number of subscribers. It follows that price is also proportional to number of subscribers [instead of proportional to a quadratic function of number of subscribers as in the equilibrium model, equation (19)], and the expansion path of the service is a straight line.

One possibility would be to let $p = p_A \dfrac{q}{q_A}$. In that case demand would expand along OA. Price would increase automatically as q increased. When q reached the optimum, q_A, price would just equal its optimum, p_A.

An alternative is to let $p = p_A \dfrac{q}{q_B}$. Demand would expand along OB until the critical mass was reached. Then, for further expansion price should be fixed at $p_A(-\epsilon)$, and demand would expand along BA.

Unfortunately, the equal-usage assumption may not be very realistic. If not, usage-proportional pricing would exclude individuals with a lot of low value usage, and would admit (for a low price) individuals with a small amount of high value usage.

This might not be a bad idea, even in the long run. In general, usage-proportional pricing would be appropriate if the network were being used at capacity, and costs were closely related to total usage, not necessarily to the number of subscribers. A fixed price per subscriber would be appropriate if the network were not being used at capacity, and much equipment had to be committed for each subscriber.

The latter is perhaps more typical of a new communications service. Thus, usage-proportional pricing would probably result in some inefficiencies and misallocation of resources. Whether these inefficiencies are substantial or trivial is an empirical question. In any event, they should be compared to the inefficiencies of other start-up programs; e.g., the direct approach discussed in the previous subsection.

For a mature service the price should probably have a higher fixed component and a lower usage component. However, the externalities in consumption still have to be taken into account.[14]

A program with a low introductory price relies more on market processes than does the direct approach discussed in the previous subsection and does not depend so critically on the managers' judgment. In particular, the managers are not required to determine which individuals have the highest w_i. The individuals select themselves by choosing to subscribe to the service at the offered price.

☐ **Community of interest groups.** Community of interest groups may greatly reduce the practical difficulty of starting up the service. Maximum equilibrium demand may be achieved even if the initial user set is small—so long as that set exceeds the critical masses for some community of interest groups.

At the same time, community of interest groups place a greater burden on whatever procedure is used to select the initial users. If the initial user set is selected by managers, they must know what the community of interest groups are and decide how many individuals to select from each. In some circumstances, it would be optimal to select everyone from the same group; in other cases a more even spread would be optimal. In any event, the managers must make this choice, and the success of the program may be greatly influenced by how well they choose.

Community of interest groups also place a greater burden on the market process for programs involving a low introductory price. The expansion path involves price and $g_1, \ldots g_k$ (instead of q as in Figure 4). Efficiency requires that k different linear combinations of the g_i be concave downward. (The equivalent condition to γ's being concave downward in Figure 4.) But we have only one control variable, p. Clearly, it may be impossible to satisfy all k concavity conditions.

Thus, *any* program based on a (single) low introductory price may be inefficient in the sense that some individuals may join at the low introductory price but later drop out as the price rises. If this problem is serious, the seller may find it advantageous (and perhaps necessary) to use discriminatory pricing to assure that only "permanent" users join.

☐ **Few substantial contacts.** If an individual's demand depends on his principal contacts' being users, the start-up problem may be

[14] See Squire [3].

fundamentally different from that discussed above. It may be unnecessary for there to be hundreds or thousands of users before an individual demands the service. A user population of two or three may be self-sufficient—if they are the right two or three people.

These small self-sufficient sets do not necessarily promote further growth, but they do allow a different kind of approach to starting up the service. The seller can begin by establishing small self-sufficient user sets. He can then gradually expand the size and number of these sets until the desired equilibrium user set is attained or until the service starts to grow by itself. The practicality of this method depends primarily on the size of the (minimum) self-sufficient user sets.

The smallest possible self-sufficient user sets consists of two mutual contacts for whom both v_{ij} and $v_{ji} \geqslant p$. If there are many such pairs i, j in the population, the service will start up, expanding beyond the minimum equilibrium set, with little or no help. Selling the service to i and j requires only getting the two together. And they may organize themselves and agree to subscribe to the service (contrary to the disequilibrium adjustment process assumed above[15]). After i and j both subscribe to the service, they may attract other individuals, and the service can grow further.

In fact, such growth from self-sufficient sets of two probably accounts in large part for the success in starting up telephone service. Indeed, telephone made a substantial penetration of the market while it was entirely a private-line service.

However, even if the service is viable, its incremental utility may be insufficient for very many people to demand the service to communicate with a single principal contact. If this is the case, the start-up problem is more difficult, and we must deal with larger self-sufficient user sets.

A self-sufficient set of three mutual contacts might also be able to organize itself and have all three members agree to subscribe together. However, this becomes progressively more difficult and unlikely as the size of the self-sufficient sets increases. It would be especially difficult if all the members of the set were not mutual contacts and no one knew all the other members of the set.

The seller might be able to gather data and determine self-sufficient user sets. He could then try to sell the service to everyone in such a set simultaneously. Naturally this is more difficult, the larger is the self-sufficient set. In fact, the difficulties of organizing even six to eight people and getting them all to agree to a joint purchasing decision may be far from trivial.

Nevertheless, the seller might do well to gather data on communications patterns and try to determine self-sufficient user sets. But it may be necessary to combine this with some other kind of start-up program. This could take the form of the direct approach or the low introductory price previously discussed. However, it is also possible to have a continual program in which each new subscriber is offered a low rate until the seller can connect a self-sufficient set of users that contains him. This kind of program

[15] P. 22.

would be effective if new users had high values of w_i but little community of interest with the current user set.

References

1. ARTLE, R. AND AVEROUS, C. "The Telephone System as a Public Good: Static and Dynamic Aspects." *The Bell Journal of Economics and Management Science,* Vol. 4, No. 1 (Spring 1973), pp. 89–100.
2. MARRIS, R. *The Economic Theory of "Managerial" Capitalism*. New York: Free Press, 1964.
3. SQUIRE, L. "Some Aspects of Optimal Pricing for Telecommunications." *The Bell Journal of Economics and Management Science,* Vol. 4, No. 2 (Autumn 1973), pp. 515–525.