58

An Approach to the Pricing of Broadband Telecommunications Services

Richard P. McLean
and
William Sharkey

Columbia Institute for Tele-Information
Graduate School of Business
809 Uris Hall
Columbia University
New York, New York 10027
(212) 854-4222

An Approach to the Pricing of Broadband Telecommunications Services

Richard P. McLean[*]

and

William W. Sharkey[†]

March 1993

Abstract

In this paper, we describe two cost allocation methodologies which may be used to determine prices for a regulated telecommunications supplier offering heterogeneous services on a broadband network. Both methodologies can be characterized by sets of plausible axioms, that one could argue should be satisfied by any pricing rule. One of the approaches leads to the Aumann-Shapley pricing rule, which is well known in the literature. The other approach leads to a pricing rule based on the Shapley value of a related cooperative game. While these approaches are similar in motivation, they differ in the technical requirements which must be imposed on the underlying cost function, and we argue that Shapley value pricing is more appropriate in a telecommunications context. We are able to explicitly determine Shapley value prices for customers who differ from one another in terms of their arrival rates, service rates, costs of lost work, and the number of simultaneous channels which they require.

---

[*]Department of Economics and RUTCOR, Rutgers University, New Brunswick, NJ

[†]Bellcore, Morristown, NJ

An Approach to the Pricing of Broadband Telecommunications Services[1]

Richard P. McLean

and

William W. Sharkey

1. Introduction

In this paper, we consider cost allocation issues arising in a model of a stochastic service system in which there are costs associated with waiting and blocking due to system congestion. We view this "service center" as a regulated firm, or public enterprise, which must recover its total costs through a regime of full cost allocating prices. A regulated telecommunications firm is responsible for providing service to its customers that satisfies certain predefined quality standards. While the quality of service is defined in terms of the blocking probabilities and waiting times that customers experience, these costs are not themselves meant to be allocated in a regulatory framework. Instead, the regulatory issue is one of allocating the long run cost of operating a telecommunications network which is capable of providing satisfactory service to various classes of users.

We believe that our framework encompasses two important problems: pricing of broadband telecommunications services, and pricing in local and regional computer networks. First consider pricing of telecommunications services. Telecommunication service providers are turning increasingly to fiber optical transmission, and to packet switching networks, that are capable of providing a vast array of services on the public network. As a result, it will soon be possible to provide new services, including file transfer, interactive graphics, image transfer, and real time video, on the same public network that is used for traditional voice communication. The transformation of an analog, copper based network to a digital fiber network, however, is expensive. It can be argued that the cost of this transformation must be allocated among the subscribers to new services, which the new technology will make possible, and among subscribers to traditional services such as voice

[1]We thank the referees for a number of helpful comments on an earlier version of this paper.

and low speed data. As long as new and old services continue to be provided by a common regulated supplier, we believe that the methodology to be developed in this paper may provide a useful framework for the determination of relative prices of different service classes.

The issues involved in pricing of broadband telecommunications services are most easily addressed in the case of a (digital) packet switching network. In such a network, every message which arrives at the gateway to the network is divided into cells, or packets, of uniform size. Each cell is assigned a header describing its destination, its priority class, and sufficient information to allow the full message to be reconstructed at the final destination. Every service type has its own set of demand characteristics, in terms of the tolerance for delay in transmission, and the costs associated with information lost in transmission. For example, voice, interactive data, and real time video cannot tolerate significant delay. On the other hand, voice, and to a lesser extent interactive data, can accept some cell loss and still provide an acceptable quality of service. File transfer can accept some delay in transmission, but is extremely sensitive to lost cells.[2]

In a traditional circuit-switched digital network similar issues arise, although to a lesser extent. Under the narrow band Integrated Services Digital Network (ISDN) standard, two data circuits with a capacity of 56 kilobytes per second, and a single voice circuit are carried on the same channel.[3] In such a network rapid data transmission is possible, but only relatively crude interactive data services can be offered. With this technology, the heterogeneity of service type in terms of channel capacity required is a crucial component of the problem. In a broadband ISDN network, as described in the preceding paragraph, heterogeneity of user requirements for channel capacity also arises, since the portion of the traffic for which delay is intolerable will be carried at a constant bit rate. This means that sufficient channel capacity must be available at pre-specified intervals in order to carry the entire capacity of the offered load, whether it is a simple voice message, or an interactive video transmission. In contrast, most data services can be transmitted at a variable bit rate, since satisfactory service only requires error free transmission within a reasonable period of time.

A second area of application of the pricing methodology that we will develop is the pricing of services offered on a computer network. The Internet is a prime example, in which thousands of worldwide academic, governmental, and industrial research facilities are linked on a common network. Currently the Internet is a relatively low capacity network which primarily supports electronic mail, remote computer login, and file transfer services. While institutions are charged a fee for access to the Internet, there are at present no usage fees associated with the volume of traffic sent over the network. Furthermore, actual users of the system typically pay no fees of any form.

[2]Lost packets must be retransmitted, and so the higher the rate of cell loss, the greater the need for duplicate transmissions.

[3]In contrast, a standard setting body has recommended two speeds for broadband ISDN traffic in a packet switching network, at 155 and 600 megabits per second [Hong and Suda (1991).

There are plans to upgrade the Internet, both in terms of available bandwidth, and in terms of the number of users who will have access to it. It is reasonable to suppose that future users will be a heterogeneous community, who will wish to use the network for a variety of different reasons. It is also reasonable to suppose that users of the future Internet will be charged usage fees, both to deal with the incentive issues raised in the following section, and to recover the costs associated with network expansion. We believe that the pricing methodology to be described in this paper resolves the cost allocation issue in a manner that will tend to promote efficient utilization of the network.

2. Methodologies for Pricing of Telecommunications Services

In this section, we briefly review a number of the methodological approaches that have been suggested for the pricing the outputs of regulated firms generally, with particular emphasis on the literature on pricing within queuing systems. We believe these approaches to be complementary to the cost allocation approach adopted in this paper, in the sense that no one approach can be applied in every conceivable pricing scenario that a telecommunications provider (or social planner) is likely to encounter in practice. At the same time, we believe that the cost allocation approach treats a number of issues that have not yet been addressed in the existing literature, and that these methods should be available to firms and their regulators who must determine real prices for a variety of heterogeneous services.

We turn first to a large literature which has examined the role of pricing in queuing systems from an incentive point of view. Generally these papers have been concerned with the role of pricing rules in eliciting truthful information from network users regarding their value of completed service and their cost of waiting. Several of the papers also focus on the use of priority queuing disciplines as a method of improving aggregate system performance.

Naor (1969) was one of the first to consider the use of prices to improve system performance in a queuing network. In Naor's model, there is a single customer type, characterized by an arrival rate λ and service rate $\mu = 1 / r$. Each customer is assumed to have a value of service of v, and a cost per unit of waiting plus service time of c, measured in the same units as v. Both v and c are assumed known by the network administrator. The network administrator is allowed to set a price p for entry into the system, and customers are assumed to balk if their expected waiting cost exceeds $v - p$. For every price p, there is therefore a critical queue size $n(p)$ such that customers will choose to balk whenever the size of the queue exceeds $n(p)$. Naor demonstrates that the revenue maximizing queue length is less than the socially optimal queue length, which is, in turn, less than the queue length which results if $p = 0$.

Sanders (1985) considers a somewhat different model, in which the network administrator is assumed to be able to impose a flow control algorithm which limits entry into the network. However, it is now assumed that customers have private information about the benefit of completing a job, and about the cost of waiting in a queue. As is well known in the incentives literature, if users are asked to directly reveal their private information, they will have an incentive to overstate both the benefits of completing service and the marginal delay cost. However, by appealing to the mechanism design literature, it can be shown that indirect mechanisms exist under which truthful reporting can be assured. Chakravorti (1993) proposes a similar approach in which truth telling behavior is induced while the overall budget balance is achieved, correcting a substantial practical difficulty with the Sanders model.

Mendelson and Whang (1990) consider a framework in which a network administrator can determine waiting time for individual users through the choice of a priority pricing rule. It is assumed that there are n customers with waiting costs c_i which satisfy $c_1 > \ldots > c_n$, and that service rates are homogeneous, with $\mu_i = 1$ for all i. Customers have private information about utility functions and waiting costs. The priority scheme that minimizes expected waiting cost per unit of time is the one which assigns highest priority to customer class 1, second highest priority to class 2, and so on, down to class n.

The network administrator, who knows the set of possible types but cannot identify the type of a particular customer, seeks to maximize net benefit by setting a price p_i for a customer who claims to be of type i. Given the above priority pricing mechanism, each customer is allowed to announce a priority class and choose a level of demand. An optimal pricing mechanism is one that maximizes aggregate net benefit, which is the sum of each customer's utility of completed service less the cost of waiting in the queue. A Nash equilibrium is a vector of strategies such that no customer can gain by deviating, taking as given the strategies of the other customers. An equilibrium price vector is said to be "incentive compatible" if every customer reports his or her true type. Mendelson and Whang derive a pricing mechanism that is both optimal and incentive compatible. In an earlier result, Dolan (1978) considered a model in which customers report their delay costs and the network manager determines a service discipline which serves customers in decreasing order of reported costs. Each customer is charged a price equal to the marginal delay costs that his position in the queue imposes on all customers with lower reported delay costs. It is proven that this pricing mechanism induces truthful reporting, not only as a Nash equilibrium, but also as a dominant strategy equilibrium. That is, each customer has the incentive to report true waiting costs, whatever strategies the other customers choose.

An alternative approach to pricing, distinct from the incentive literature, is the Ramsey-Boiteux model, in which a firm, or its regulator, sets prices to maximize a social welfare function subject to a profit constraint on the firm providing service.[4] Suppose that $q = (q_1, \ldots, q_n)$ represents a

[4]See Baumol and Bradford (1970) and Boiteux (1956) for a discussion to this approach.

vector of possible outputs for the firm and let $C(q)$ represent the total cost of producing the output vector q. If demands for each of the firm's products are independent, it is possible to write the inverse demand function $p_i = P_i(q_i)$ which expresses the amount that customers are willing to pay for the last unit produced when output is q_i. The Ramsey pricing rule in the case of independent demands is given by the formula

$$\frac{p_i - \frac{\partial C}{\partial q_i}}{p_i} = -\frac{k}{\eta_i},$$

where η_i is the elasticity of demand. The number k is chosen to satisfy the budget constraint, where $k = 1$ corresponds to unconstrained profit maximization, and $k < 1$ corresponds to budget constrained pricing when there are increasing returns to scale, so that prices in excess of marginal cost are required to recover total costs.

While the Ramsey pricing rule has desirable theoretical properties, there are several problems in its practical implementation. First, the rule requires accurate information about demand elasticities facing the firm, and in the case of interdependent demands, both own price and cross price elasticities must be estimated. In a competitive environment in which technological innovation leads to new service offerings, as in the telecommunications industry, these elasticities may be difficult to obtain. Second, Ramsey prices may be perceived as inherently unfair, or in a competitive environment they may invite entry when entry is not warranted on grounds of industry cost minimization. This follows because the rule requires that the markup of price above marginal cost should be the greatest in those markets in which the estimated elasticity of demand is smallest. The rule in effect institutionalizes a transfer, or subsidy, from customers in inelastic markets to those in elastic markets.

Suppose, for example, that the cost function has the form $C(q) = \sum_{i=1}^{n} C^i(q_i)$, with $C^i(q_i) = f_i + m_i\, q_i$. Then costs are separable, and the cost function for each output has a fixed cost f_i and a marginal cost m_i. Under the Ramsey rule with independent demands, aggregate fixed costs, $\sum_{i=1}^{n} f_i$, are allocated in inverse proportion to the elasticities of demand η_i. For example, if $\eta_1 = 0$ and $\eta_i < 0$ for $i = 2, \ldots, n$, then aggregate fixed costs, $\sum_{i=1}^{n} f_i$, would be allocated to service 1.

Note that in the Ramsey pricing approach, the optimal prices p_i are computed as the solution to a social welfare maximization problem that requires knowledge of consumer demand functions as well as the cost function. In contrast, cost allocation prices, to be discussed in this paper, are characterized by a set of axioms, and are based only on appropriate information about the firm's costs. In the Ramsey model, prices serve a dual function of allocating costs and rationing individual consumer's demands. In this sense they are demand compatible. If additional information regarding consumer demand functions is available, however, there are straightforward procedures in which demand compatible cost allocation prices can be determined. (We will briefly describe such methods in section 6.)

The literatures on Ramsey pricing and incentive pricing address a number of important issues that are relevant to the pricing of broadband telecommunications services. Without denying the importance of these approaches, we believe that the cost allocation approach to be developed in this paper, provides useful additional insights into the pricing of broadband services. In particular, where a group of heterogeneous services are offered (differentiated in terms of arrival rates, service times and requirements for channel capacity) the cost allocation approach offers a methodology for setting relative prices for service classes that is independent of, but complementary to, incentive questions that have been previously considered.

3. Cost Functions Which Arise from Telecommunications Design Problems

Let $N = \{1, \dots, n\}$ denote a finite set of n "service classes" (e.g. voice, video, data, etc.) Demands for service, i.e. "calls," arrive at a transmission point consisting of k channels, and for simplicity, we assume that the arrival process for calls of type i is Poisson with rate λ_i. Calls contribute to overall system congestion in two ways. First, each type i call has a duration, or size, which is assumed to be exponentially distributed with mean r_i and variance r_i^2. In a telecommunications setting r_i measures the length of a message of type i that is transmitted over the network. (In much of the literature r_i is expressed in terms of the service rate $\mu_i = \frac{1}{r_i}$.)

Second, the cost of providing a service depends on the number of processors that are simultaneously required, and the different protocols required in transmission. Thus arriving calls also contribute to system congestion through the number, d_i, of simultaneous channels which are required for the duration of a type i call.

In this paper we address the following question: how should a regulator (social planner, or system designer) allocate the costs associated with an optimally designed telecommunications system? We emphasize that this problem has two aspects. First, the system planner must determine an optimal system as a function of customer characteristics. Second, the planner must choose a mechanism for allocating the cost of this optimally designed system, based on these customer characteristics. The first issue is addressed by studying various optimization problems, and is the subject of the remainder of this section. The second issue is one of cost allocation, and will be discussed in the remainder of the paper.

To begin, let k represent the number of channels in the system, B the buffer size, and $\lambda = (\lambda_1, \dots, \lambda_n)$, $r = (r_1, \dots, r_n)$, and $d = (d_1, \dots, d_n)$ the arrival, service and channel requirement vectors described above. Let $g(k,B)$ represent the cost of building a system with k

channels and a buffer of capacity B. Typically g is an increasing function of k and B. Given k, B, and the values of λ, r and d, let $\beta_i(k,B;\lambda,r,d)$ represent the blocking probability for a call of type i. The precise form of $\beta_i(k,B;\lambda,r,d)$ depends on the queue discipline. If $B = 0$, for example, then β_i is simply the steady state probability that at least $k - d_i + 1$ channels are in use. Finally, let $w_i(k,B;\lambda,r,d)$ represent the expected delay (i.e. time in queue prior to commencement of service) for a call of type i.

A number of different models can be used to define a cost function for a general telecommunications design problem, and we will discuss two of them. In each of these models, the planner takes r, λ and d as fixed parameter vectors, and must choose values of k and B to maximize an objective function.

Model 1: Let $\overline{\beta}_1, \ldots, \overline{\beta}_n$ and $\overline{w}_1, \ldots, \overline{w}_n$ be "acceptable" blocking probabilities and expected delays. In this model, the design problem is to solve:

$$\min_{k,B} \; g(k,B)$$

(3.1) s.t.

$$\beta_i(k,B;\lambda,r,d) \le \overline{\beta}_i \text{ for all } i$$

$$w_i(k,B;\lambda,r,d) \le \overline{w}_i \text{ for all } i .$$

The solution set is parameterized by $\lambda, r, d, \overline{\beta}$, and $\overline{w}$, and while it is certainly the case that all of these parameters are important determinants of cost, we view d, $\overline{\beta}$, and $\overline{w}$ as outside the control of individual customers. As a result we will focus on λ and r as the appropriate objects to be priced in the cost allocation schemes presented in later sections. Hence, we will write the cost function associated with problem (3.1) (i.e. the minimized value of g subject to the constraints) simply as $F^1(\lambda,r)$, suppressing the dependence on $d, \overline{\beta}$, and $\overline{w}$.

Model 2: Let c_i be the value of a call to a type i customer if his call is not blocked. Alternatively, c_i represents the economic loss associated with a blocked call of type i. Let γ_i represent the economic loss associated with a unit of time spent waiting in the queue. If the system designer desires to maximize social surplus, then the problem becomes

$$\max_{k,B} \left\{ \sum_{i\in N} c_i\lambda_i\left[1-\beta_i(k,B;\lambda,r,d)\right] - \sum_{i\in N} \gamma_i\lambda_i w_i(k,B;\lambda,r,d) - g(k,B) \right\}.$$

Maximizing social surplus is equivalent to the following minimization problem:

$$(3.2) \qquad \underset{k,B}{Min} \left\{ g(k,B) + \sum_{i \in N} c_i \lambda_i \beta_i(k,B;\lambda,r,d) + \sum_{i \in N} \gamma_i \lambda_i w_i(k,B;\lambda,r,d) \right\}.$$

As in model 1, the solution set is parameterized by λ, r, d, c and γ. For precisely the same reasons as those given above, we will suppress the dependence of the cost function for problem (3.2) on d, c and γ, and write it simply as $F^2(\lambda,r)$.

As we emphasized above, the system designer's problem has two aspects: the design of the optimal system and the allocation of the associated costs. In this section we have discussed the design issue. In the next section we describe a methodology for resolving the cost allocation problem.

4. Game Theoretic Approaches to Pricing of Telecommunications Services

In this section we are interested in the problem of allocating to system users the costs $F^1(\lambda,r)$ and $F^2(\lambda,r)$ associated with problems (3.1) and (3.2). In fact, the methodologies to be discussed are applicable to any system design problem. We will treat the arrival rates and service rates as "commodities" and will allocate their costs using a suitably chosen system of prices. Several game theoretic approaches to the resolution of this problem are possible. In this section we will present two related methods and discuss their applicability to the cost allocation problems associated with (3.1) and (3.2). In section 5 we will apply one of these schemes to special cases of the models defined in section 3.

We begin with some notation and definitions. Let $N = \{1, \ldots, n\}$, and let 2^N denote the power set of N. A *cost function* is a continuously differentiable function $C: \Re_+^n \to \Re$, with $C(0) = 0$. For a generic output vector $x \in \Re_+^n$, $C(x)$ represents the cost of producing x, and we denote the marginal cost of the ith output by $\frac{\partial C}{\partial x_i}$. If C is a cost function, and $\alpha \in \Re_{++}^n$ is a fixed output vector, then the pair (C,α) defines a *cost allocation problem*. The class of all cost allocation problems is denoted by $\mathcal{C}^n$, i.e. $\mathcal{C}^n$ is equal to $\{(C,\alpha) \mid C$ is a cost function, and $\alpha \in \Re_{++}^n\}$. A *pricing mechanism* is a function $p: \mathcal{C}^n \to \Re^n$ that maps each $(C,\alpha) \in \mathcal{C}^n$ to a price vector $(p_1(C,\alpha), \ldots, p_n(C,\alpha))$ in $\Re^n$. There are many properties that one might like a price mechanism to satisfy, and in the next two sections we propose alternative systems of axioms that will enable us to uniquely characterize two such price mechanisms: Aumann-Shapley pricing and Shapley value pricing.

4.1 Aumann-Shapley Pricing in a Telecommunications System

Consider the following set of axioms. In the statements of these axioms, p is a pricing mechanism.

(A.1) *Cost Sharing:* $\sum_{i=1}^{n} \alpha_i p_i(C,\alpha) = C(\alpha)$.

(A.2) *Monotonicity:* If $\frac{\partial C}{\partial x_i} \geq 0$ then $p_i(C\ ,\ \alpha) \geq 0$.

(A.3) *Additivity:* $p(C_1+C_2\ ,\ \alpha) = p(C_1\ ,\ \alpha) + p(C_2\ ,\ \alpha)$ for each pair of cost allocation problems $(C_1\ ,\ \alpha)$ and $(C_2\ ,\ \alpha)$,

(A.4) *Consistency:* $p_i(C\ ,\ \alpha) = p_j(C\ ,\ \alpha)$ whenever α_i and α_j enter the cost function only through their sum.

(A.5) *Rescaling Invariance:* If $C_2(x) = C_1(\gamma_1 x_1\ ,\dots,\ \gamma_n x_n)$ then $p_i(C_2\ ,\ \alpha) = \gamma_i p_i[C_1\ ,\ (\gamma_1\alpha_1,\dots,\gamma_n\alpha_n)]$.

Cost sharing simply formalizes the property that revenue must exactly recover total cost $C(\alpha)$. Monotonicity means that goods or services that unambiguously contribute to total cost should be assigned non-negative prices. Additivity is a technical requirement that states that arbitrary decompositions of a total cost function into component parts ought not to affect the final cost allocation. Consistency requires that any two commodities that have the same effect on total cost (which clearly is the case if total cost is a function only of their sum) should be assigned the same price. Finally, rescaling invariance is a property which states that a cost allocation procedure ought to be insensitive to the choice of units in which output quantities are measured. For example, measuring telecommunication services in terms of minutes of use rather than seconds of use should lead to the obvious rescaling of final prices.

Having listed these conditions, it is natural to ask if there exists a price mechanism that satisfies all five. Not only does a mechanism satisfying the above properties exist, but in fact there is only one such mechanism.

Theorem [Billera-Heath (1982), Mirman-Tauman (1982)]: There exists one and only one price mechanism satisfying axioms A.1–A.5. It is defined by the formula

$$p_i^{AS}(C,\alpha)=\int_0^1 \frac{\partial C(t\alpha)}{\partial x_i}dt$$

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

The prices defined above are called Aumann-Shapley (AS) prices. Intuitively the AS price for service i is an "average" of its marginal costs. The AS pricing rule has received considerable attention in the cost allocation literature. Indeed, Billera, Heath and Raanan (1978) first suggested this rule in the context of a telephone pricing problem. In McLean and Sharkey (1992a) we have explicitly computed AS prices in a model of a stochastic service system in which the number of servers is held constant.

Given the setup described above, one can apply this methodology to the system design problems of section 3 by treating the vector (λ,r) like the commodity vector, α, and defining the cost allocation problems $(F^1,(\lambda,r))$ and $(F^2,(\lambda,r))$. If the cost functions are differentiable, the AS pricing methodology can be applied to these cost allocation problems, and AS prices for the arrival and service rates derived. To illustrate, suppose that F^1 is differentiable and let p_i and q_i represent the AS prices that result from the cost allocation problem $F^1(\lambda,r)$. Each arrival of type i is charged p_i, and each minute of usage is charged q_i, so that the long run steady state average revenue generated by this price system is equal to $\sum_{i=1}^{n} p_i\lambda_i+\sum_{i=1}^{n} q_i r_i$. Since AS pricing satisfies the cost sharing axiom, this steady state expected revenue is exactly equal to the expected steady state costs $F^1(\lambda,r)$.

In practice, the cost functions of section 3 are not generally differentiable. In fact, while numerical methods can be used to compute the cost at particular outputs in (3.1) and (3.2), it may not even be possible to define an analytic representation of the relevant cost function. This presents a serious obstacle to the application of the Aumann-Shapley pricing rule. For this reason we propose a closely related alternative pricing methodology in the following section.

4.2 Shapley Value Pricing in a Telecommunications System

We begin by defining a "stand alone game" associated with a cost allocation problem (C,α). For each $S \subseteq N$, let $\alpha^S \in \Re^n$ be the vector with $\alpha_i^S = \alpha_i$ if $i \in S$, and $\alpha_i^S = 0$ if $i \notin S$. Thus, $\alpha^N = \alpha$. Now define a set function $\hat{C} : 2^N \rightarrow \Re$ such that $\hat{C}(S) = C(\alpha^S)$ for all $S \subseteq N$. The function $\hat{C}$ is the cost function in the stand alone problem associated with (C,α), and $\hat{C}(S)$ represents the cost of satisfying the demands α_i for only those commodities $i \in S$.

The pricing scheme which we propose is based upon the Shapley value of a general cooperative game. A game (with transferable utility) is defined by a player set $N = \{1, \ldots, n\}$ and a set function $v : 2^N \rightarrow \Re$ such that $v(\emptyset) = 0$. Let G_N denote the $2^N - 1$ dimensional vector space of all games on N. Subsets $S \subseteq N$ are coalitions, and $v(S)$ is interpreted as the reward attainable by (or cost incurred by) the coalition S through some joint course of action. Typically, one is interested in dividing the number $v(N)$ among the individuals in N, and in our context, this will mean allocating cost.

Numerous fair division schemes are available. For example, one might simply distribute $v(N)/n$ to each individual. Such a scheme is generally too simplistic, since it ignores the other values $v(S)$, which themselves contain information regarding an individual's contribution to the rewards or costs associated with each subset of users. Since there are many potential ways to divide $v(N)$, it is useful to proceed by defining axioms for the class of all fair division problems, as was done above in the case of Aumann-Shapley pricing.

More formally, we consider the class of functions $\psi : G_N \rightarrow \Re^n$, where for each $v \in G_N$ and $i \in N$, $\psi_i(v)$ is the allocation that i receives in the problem v. Consider the following properties that one might argue that any ψ should satisfy.

(S.1) *Additivity:* $\psi(v + w) = \psi(v) + \psi(w)$.

(S.2) *Efficiency:* $\sum_{i \in N} \psi_i(v) = v(N)$.

(S.3) *Symmetry:* If $v(S \cup i) = v(S \cup j)$ for all $S \subseteq N \setminus \{i,j\}$, then $\psi_i(v) = \psi_j(v)$.

(S.4) *Monotonicity:* If $v(S \cup i) - v(S) \geq 0$ for all $S \subseteq N \setminus \{i\}$, then $\psi_i(v) \geq 0$.

Additivity is an important "regularity" condition. If a fair division problem can be decomposed into the sum of two distinct problems, then the sum of the solutions of the component problems is the solution of the original problem. Efficiency means that exactly $v(N)$ is allocated to the individuals. The symmetry property may be paraphrased as follows: two individuals who "affect" the problem in exactly the same way (as they will if $v(S \cup i) = v(S \cup j)$ for all $S \subseteq N \setminus \{i,j\}$) should receive the same payoff. Finally, monotonicity requires that an individual's payoff (or cost allocation) should be non-negative if his contribution to the worth (or cost) of every coalition that he might join is non-negative. Note that, as a consequence of the axioms, an individual will receive a payoff of zero if his contribution to every coalition is zero.

These properties are satisfied by one and only one function $\varphi : G_N \rightarrow \Re^n$ known as the Shapley value. To define the Shapley value precisely, we need additional notation. Let Ω represent the $n!$ orderings of the members of N. For each $i \in N$ and each $R \in \Omega$, let $X_i(R)$ be the set of predecessors of i in the order R. More formally, if $R = (j_1, \ldots, j_n) \in \Omega$ with $i = j_k$, then

$X_i(R) = \{j_1, \ldots, j_{k-1}\}$. (Note that $X_i(R) = \varnothing$ if $i = j_1$.) The marginal contribution of player i in the order $R \in \Omega$ is defined as

$$\Delta_i(R,v) = v(X_i(R) \cup i) - v(X_i(R)).$$

The Shapley value $\varphi : G_N \to \Re^n$ is the operator defined by

$$\varphi_i(v) = \sum_{R \in \Omega} \frac{1}{n!} \Delta_i(R,v)$$

for each $i \in N$. The Shapley value of player i can be interpreted as an expected marginal contribution in a random ordering of the players when the $n!$ possible orderings are equally likely. The characterization theorem for the Shapley value can now be stated.

Theorem [Shapley, (1953)]: A function $\psi : G_N \to \Re^n$ satisfies axioms S.1–S.4 if and only if for each $v \in G_N$ and for each $i \in N$, $\psi_i(v) = \varphi_i(v)$.

If $\hat{C}$ is the stand alone cost function derived from the cost allocation problem (C,α), then $\hat{C}$ is a game in which the individual players are the commodities $1, \ldots, n$. $\hat{C}(S)$ represents the cost associated with the members of S on their own, and the fair division problem is the problem of allocating the total cost $\hat{C}(N)$ "fairly" among the commodities. If we use the Shapley value to accomplish this, then $\varphi_i(\hat{C})$ represents the expected marginal contribution to total cost attributable to commodity i, and each of the above axioms has a straightforward interpretation. Additivity (S.1) implies that, when costs can be decomposed into, say labor costs, raw material costs, etc., then the portion of cost allocated to commodity i is simply the sum of i's contribution to the costs of the component problems. Efficiency (S.2) requires that total cost be exactly recovered. Symmetry (S.3) requires that two commodities that affect costs in the same way should contribute equal amounts. Monotonicity (S.4) says that any commodity that does not decrease the cost of any group of commodities to which it is added should make a non-negative contribution to total cost. That is, no commodity should be subsidized.

If we apply the Shapley value to $\hat{C}$, then

$$\varphi_i(\hat{C}) = \sum_{R \in \Omega} \frac{1}{n!} \left[C(\alpha^{X_i(R) \cup i}) - C(\alpha^{X_i(R)}) \right] \tag{4.1}$$

represents the contribution to total cost $C(\alpha)$ made by commodity i. In a cost allocation framework $\varphi_i(\hat{C})$ can be interpreted as an average of incremental costs.

We are now ready to define the Shapley value (SV) price mechanism by

(4.2) $$p_i^{SV}(C,\alpha) = \frac{\varphi_i(\hat{C})}{\alpha_i}$$

where $\hat{C}$ is the stand alone problem derived from (C,α).

We first note that p^{SV} satisfies all but one of the Aumann-Shapley axioms. Since φ satisfies efficiency (S.2), it follows that $\sum_{i=1}^{m} \alpha_i p_i^{SV}(C,\alpha) = \sum_{i=1}^{m} \varphi_i(\hat{C}) = \hat{C}(N) = C(\alpha)$, and therefore p^{SV} satisfies the cost sharing axiom (A.1) of Aumann-Shapley pricing. If C is differentiable, with $\frac{\partial C}{\partial x_i} \geq 0$ for all x, then $C(\alpha^{S\cup i}) - C(\alpha^S) \geq 0$ for all $S \subseteq N\backslash i$, and therefore $p_i^{SV}(C,\alpha) = \frac{\varphi_i(\hat{C})}{\alpha_i} \geq 0$ since φ satisfies monotonicity (S.4). Therefore p^{SV} satisfies the monotonicity axiom (A.2) of Aumann-Shapley pricing. Suppose that $C = C_1 + C_2$, and let $\hat{C}_1, \hat{C}_2$, and $\hat{C}$ be (respectively) the stand alone games associated with (C_1,α), (C_2,α), and (C,α). Clearly $\hat{C} = \hat{C}_1 + \hat{C}_2$ and, since φ satisfies additivity (S.1), it follows that $\varphi(\hat{C}) = \varphi(\hat{C}_1) + \varphi(\hat{C}_2)$. Therefore $p_i^{SV}(C,\alpha) = \frac{\varphi_i(\hat{C})}{\alpha_i} = \frac{\varphi_i(\hat{C}_1)}{\alpha_i} + \frac{\varphi_i(\hat{C}_2)}{\alpha_i} = p_i^{SV}(C_1,\alpha) + p_i^{SV}(C_2,\alpha)$ for each $i \in N$, and we conclude that p^{SV} satisfies the Aumann-Shapley additivity axiom (A.3). Finally, it is possible to demonstrate that $p^{SV}(C,\alpha)$ satisfies the Aumann-Shapley rescaling invariance axiom (A.5) whenever C is differentiable, using a simple argument based on the mean value theorem.

The distinction between AS pricing and SV pricing lies in the approach to symmetry/consistency. In McLean and Sharkey (1992a) we showed that Shapley value prices can be fully characterized in terms of the cost sharing axiom (A.1) and the following new axioms.

(B.2) *Discrete Monotonicity:* If $C(\alpha^{S\cup i}) \geq C(\alpha^S)$ for all $i \in N$ and $S \subseteq N\backslash i$, then $p_i(C,\alpha) \geq 0$.

(B.3) *Discrete Additivity:* If (C,α), (C_1,α) and $(C_2,\alpha) \in \mathcal{C}^n$, and for every $S \subseteq N$, $C(\alpha^S) = C_1(\alpha^S) + C_2(\alpha^S)$, then $p(C,\alpha) = p(C_1,\alpha) + p(C_2,\alpha)$.

(B.4) *Cost Share Proportionality:* If $C(\alpha^{S\cup\{i\}}) = C(\alpha^{S\cup\{j\}})$ for all $S \subseteq N\backslash\{i,j\}$, then $\alpha_i p_i(C,\alpha) = \alpha_j p_j(C,\alpha)$.

Discrete monotonicity (B.2) and discrete additivity (B.3) are stronger than their Aumann-Shapley counterparts (A.2) and (A.3). The hypothesis of discrete monotonicity does not require that C be monotonic on $\Re^n$ (as is required with axiom A.2), nor does the hypothesis of discrete additivity require that $C(x) = C_1(x) + C_2(x)$ for all x (as in axiom A.3). We also note that cost sharing, discrete monotonicity and discrete additivity together imply that a commodity that makes no contribution to total cost in the stand alone game is assigned a price of zero. The cost share proportionality axiom is an appropriate symmetry postulate for our setup, and it plays a role in the axiomatization of SV pricing analogous to that played by consistency in the axiomatization of AS pricing.

The characterization theorem for Shapley value pricing can now be stated.

Theorem [McLean and Sharkey, (1992a)]: There exists one and only one price mechanism on $\mathcal{C}^n$ satisfying cost sharing, discrete monotonicity, discrete additivity, and cost share proportionality. For each $(C,\alpha) \in \mathcal{C}^n$ it is defined by the formula $p_i^{sv}(C,\alpha) = \frac{\varphi_i(\hat{C})}{\alpha_i}$, where $\hat{C}(S) = C(\alpha^S)$ for all $S \subseteq N$.

Shapley value pricing can be applied to the cost functions of section 3 even if these functions are not differentiable, and we view this as an advantage of SV pricing over AS pricing. In the next section we will apply the Shapley value pricing formula to cost functions derived from the general telecommunications design problems (3.1) and (3.2).

5. An Application of Shapley Value Pricing

In this section we analyze two special cases of models 1 and 2 in which buffer capacity is equal to zero and server requirements d are homogeneous, with $d_i = 1$ for all i. We will offer two types of pricing schemes. First, we consider an "arrival only" model. In this model we treat the λ_i's as "demands" so that the vector λ is the telecommunications analogue of the commodity vector α. In the arrival only model we will therefore determine a price per arrival so that expected total revenue is equal to expected cost. In our second model, the "arrival/usage" model, we treat both arrivals and durations as "commodities" so that (λ,r) corresponds to α. In this model, we will compute prices for both arrivals and units of service time so that expected revenue equals expected cost.

5.1 The Cost Functions in a Single Server, Zero Buffer Model

Suppose $d_i = 1$ for all i and $B = 0$. Then the blocking probability is the same for all call types and is given by the Erlang loss function[5] which we write simply as

$$\beta(k;\lambda,r) = \frac{(\lambda \cdot r)^k/k!}{\sum_{i=0}^{k}(\lambda \cdot r)^i/i!} \tag{5.1}$$

---

[5]See Roberts (1981).

where $\lambda \cdot r = \sum_{j=1}^{n} \lambda_j r_j$.

Since B is constrained to be equal to 0, there is no delay, and the objective functions of Models 1 and 2 are functions of the single decision variable k, the number of channels. We then write the objective function as $g(k)$, and the two problems can be written as

(5.2)
$$\underset{k=0,1,2,\ldots}{Min} \quad g(k) \quad s.t. \quad \frac{(\lambda \cdot r)^k / k!}{\sum_{i=0}^{k} (\lambda \cdot r)^i / i!} \leq \overline{\beta}$$

and

(5.3)
$$\underset{k=0,1,2,\ldots}{Min} \quad g(k) + (c \cdot \lambda) \frac{(\lambda \cdot r)^k / k!}{\sum_{i=0}^{k} (\lambda \cdot r)^i / i!}.$$

Let $\hat{F}^1(\lambda,r)$ and $\hat{F}^2(\lambda,r)$ denote the cost functions corresponding to optimal solutions to (5.2) and (5.3) respectively. These integer optimization problems can be solved for any given pair (λ,r) but an optimal solution $\hat{k}(\lambda,r)$ will not depend smoothly on λ and r. We can partially deal with the problem of non-differentiability when it is known that the optimum k will be "large" so that $\sum_{j=0}^{k} \frac{(\lambda \cdot r)^j}{j!} \cong e^{\lambda \cdot r}$. Using this approximation and using the gamma function representation of $k!$, the following "relaxed" problems may be used to approximate the cost functions $\hat{F}^1(\lambda,r)$ and $\hat{F}^2(\lambda,r)$ as defined:

(5.4)
$$\underset{k \geq 0}{Min} \quad g(k) \quad s.t. \quad \frac{(\lambda \cdot r)^k e^{-\lambda \cdot r}}{\Gamma(k+1)} \leq \overline{\beta}$$

(5.5)
$$\underset{k \geq 0}{Min} \quad (c \cdot \lambda) \frac{(\lambda \cdot r)^k e^{-\lambda \cdot r}}{\Gamma(k+1)} + g(k).$$

Figure 1 illustrates $\hat{F}^1(\lambda,r)$ and $\hat{F}^2(\lambda,r)$ as functions of λ, assuming $n = 1$, parameter values $c = r = 1$, and capacity costs $g(k) = \frac{\sqrt{k}}{10}$.

The planner's optimization problem can be enriched if we allow for a finite but non-zero buffer size. A non-zero buffer can improve system performance by reducing blocking probability, and by allowing for mixes of different forms of traffic (e.g. data and voice which have different tolerances for delay) to be carried efficiently on the same network. With a finite, but non-zero, buffer capacity, however, the computation of both blocking probabilities and waiting times in (3.1) and (3.2) is extremely complex. In order to provide some insights into the analysis of the problem with waiting costs, we will consider an example in which there is infinite costless buffer capacity, so that blocking costs are not present, and system performance is characterized entirely by waiting costs.

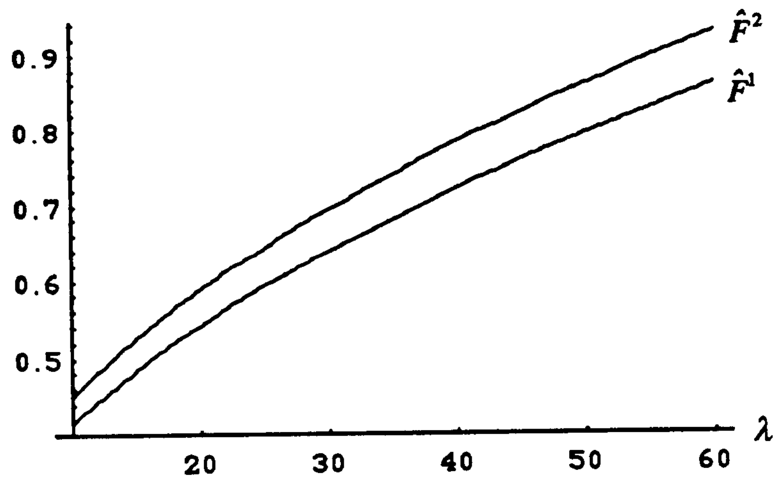

Figure 1

Suppose that service times for each customer class are independently and identically distributed as an exponential distribution with mean $\bar{r}$. Let r be the vector of service times with $r_i = \bar{r}$ for all i. Then the expected delay in a k server system is given by[6]

$$w(\lambda,r,k)=\left[1+\sum_{i=0}^{k-1}\frac{k!(k-\lambda\cdot r)}{i!k(\lambda\cdot r)^{k-i}}\right]^{-1}\frac{\bar{r}}{(k-\lambda\cdot r)}.$$

Thus (assuming buffer capacity is costless) the specializations of Models 1 and 2 generate the following cost functions:

(5.6) $$\min_{k=0,1,2,\ldots} g(k) \quad s.t. \quad w(\lambda,r,k)\le \bar{w}$$

and

(5.7) $$\min_{k=0,1,2,\ldots} g(k)+(\gamma\cdot\lambda)w(\lambda,r,k).$$

We let $\bar{F}^1(\lambda,r)$ and $\bar{F}^2(\lambda,r)$ denote the cost functions corresponding to (5.6) and (5.7) respectively.

5.2 Shapley Value Pricing: The Arrival Only Model

Let F represent one of the cost functions derived for the telecommunications design problem in section 3. In the arrival only model, we treat r as a fixed parameter, and consider the cost

[6] See, e.g. Federgruen and Groenevelt (1988, p.738).

allocation problem (G, λ), where $G: \Re^n \rightarrow \Re$ is defined by $G(\cdot) = F(\cdot, r)$. The Shapley value price for the arrival rate of a call of type i is given by $p_i^{SV}(G,\lambda) = \frac{\varphi_i(\hat{C})}{\lambda_i}$ where $\hat{C}(S) = G(\lambda^S) = F(\lambda^S, r)$. Since $\sum_{i \in N} \lambda_i p_i^{SV}(G,\lambda) = G(\lambda) = F(\lambda, r)$, these prices fully allocate cost. However, it is not entirely clear what price should be charged when some arriving calls are not "accepted" into the system, as will necessarily happen in models with a zero buffer. We propose a modified SV price defined as

$$\pi_i(G,\lambda) = \frac{p_i^{SV}(G,\lambda)}{1-\beta_i}$$

where β_i is the blocking probability for a call of type i. In models with a finite buffer, $1 - \beta_i$ is less than 1, so that $\pi_i(G, \lambda) > p_i^{SV}(G, \lambda)$.

The prices $\pi_i(G, \lambda)$ can now be used as part of the following pricing scheme. If an arriving call of type i is blocked, it pays nothing; if it is accepted, it pays $\pi_i(G, \lambda)$. The steady state number of calls of type i that are accepted is $\lambda_i (1-\beta_i)$, so the steady state expected revenue is given by $\sum_{i \in N} \lambda_i (1-\beta_i) \pi_i(G,\lambda) = \sum_{i \in N} \lambda_i p_i^{SV}(G,\lambda) = G(\lambda)$. In addition to satisfying this "modified" cost sharing property, it is easy to see that π_i also inherits the discrete monotonicity, discrete additivity, and cost share proportionality properties of the prices p_i^{SV}.

In the arrival only model, each arriving call is charged a single price, and we have argued that π_i is the appropriate price to charge accepted calls upon their arrival. These prices can also be used to compute prices for accepted units of service time. In particular, $\bar{\pi}_i = \pi_i / r_i$ is the resulting price that should be charged per unit of time for an accepted call of type i. Again the prices $\bar{\pi}_i$ recover costs since $\lambda_i(1-\beta_i)r_i$ is the steady state expected number of accepted time units associated with calls of type i. Hence steady state revenue is $\sum_{i \in N} \lambda_i (1-\beta_i) r_i \bar{\pi}_i = \sum_{i \in N} \lambda_i p_i^{SV}$.

5.3 Shapley Value Pricing: The Arrival/Usage Model

In the arrival/usage model, the λ_i and r_i are both priced, and again, modification of the SV prices is required along the lines of the arrival only prices, described above. Let $p_{\lambda_i}^{SV}(F, \lambda, r)$ and $p_{r_i}^{SV}(F, \lambda, r)$ denote, respectively, the SV prices for the arrival rate and duration of a call of type i, corresponding to (4.1) and (4.2) when α represents the output vector $(\lambda, r) = (\lambda_1, \ldots, \lambda_n; r_1, \ldots r_n)$. Note that the stand alone game $\hat{C}$ associated with $(F,(\lambda,r))$ now has $2n$ "players," whereas the stand alone game associated with (G,λ) in the arrival only model of section 5.2 has n players. Intuitively, $p_{\lambda_i}^{SV}$ is associated with "access" to the system, while $p_{r_i}^{SV}$ is related to the "use" of the system, once access has been gained. If only accepted calls are to be charged a fee then the appropriately modified prices are

$$\pi_{\lambda_i}(F,\lambda,r)=\frac{p^{SV}_{\lambda_i}(F,\lambda,r)}{(1-\beta_i)}$$

and

$$\pi_{r_i}(F,\lambda,r)=\frac{p^{SV}_{r_i}(F,\lambda,r)}{\lambda_i(1-\beta_i)},$$

where $(1-\beta_i)$ is the probability that a call of type i is accepted.

We interpret $\pi_{\lambda_i}(F\ ,\ \lambda\ ,\ r)$ as a "setup charge" that each accepted call must pay. Once an accepted call pays the setup fee and enters the system, it then pays $\pi_{r_i}(F\ ,\ \lambda\ ,\ r)$ per unit of time for its duration. Note that π_{λ_i} is measured in dollars per type i call, while π_{r_i} is measured in dollars per unit of time per call. (This explains the appearance of λ_i in the denominator of the expression defining π_{r_i}.) The portion of steady state total revenue attributed to setup fees is equal to $\sum_{i\in N}\lambda_i(1-\beta_i)\pi_{\lambda_i}(F,\lambda,r)=\sum_{i\in N}\lambda_i p^{SV}_{\lambda_i}(F,\lambda,r)$, while that attributed to usage fees is $\sum_{i\in N}\lambda_i(1-\beta_i)r_i\pi_{r_i}(F,\lambda,r)=\sum_{i\in N}r_i p^{SV}_{r_i}(F,\lambda,r)$. Hence, $\sum_{i\in N}\lambda_i(1-\beta_i)\pi_{\lambda_i}+\sum_{i\in N}\lambda_i(1-\beta_i)r_i\pi_{r_i}=F(\lambda\ ,\ r)$ and the prices in the arrival/usage model also satisfy a modified cost sharing property.

As we have argued above, the prices π_{λ_i} and π_{r_i} have interpretations as setup and usage fees. They can be combined into an average price per unit of service time by defining

$$\pi_i^*=\frac{\pi_{\lambda_i}}{r_i}+\pi_{r_i}.$$

Both $\overline{\pi}_i$ and π_i^* represent prices per unit of duration time for an accepted call of type i. The computations above demonstrate that $\sum_{i\in N}\lambda_i(1-\beta_i)r_i\overline{\pi}_i=\sum_{i\in N}\lambda_i(1-\beta_i)r_i\pi_i^*=F(\lambda,r)$, so both $\overline{\pi}$ and π^* recover costs. We note that one could also define a price per time unit of an accepted call by pricing r_i only. In such a "usage only" model the Shapley value pricing formula could be applied to the cost allocation problem (H,r), where $H(\cdot)=F(\lambda,\cdot)$, treating λ as a parameter. We will not, however, pursue this approach in this paper.

5.4 Some Comparative Statics Results

It is now possible to conduct a crude comparative statics analysis by varying the parameters of interest in both models 1 and 2 and observing the effects on the computed Shapley value prices. We will consider the arrival only model for the case in which buffer capacity B is equal to 0, and in which server requirements for each customer type are homogeneous with $d_i = 1$ for each i. Hence blocking probabilities are the same for each service type and are equal to β. We use the approximate cost functions $\hat{F}^1(\lambda,r)$ and $\hat{F}^2(\lambda,r)$ defined by (5.4) and (5.5). In (5.4), $\beta=\overline{\beta}$ is a parameter which is given as part of the design problem. In traditional voice telecommunications networks, a blocking probability of one percent during busy periods is often specified. In (5.5), β is endogenously determined.

As a benchmark, we define a pricing rule based on the system wide average cost as follows.[7]

$$p_i^{AC}=\frac{1}{(1-\beta)}\frac{r_i}{\lambda\cdot r}\hat{C}(N)=\frac{r_i}{(1-\beta)}\frac{F(\lambda,r)}{\lambda\cdot r}. \tag{5.9}$$

Since λ_i measures the arrival rate for calls of type i and r_i measures their duration, $\lambda\cdot r$ is the expected total duration of calls arriving in the system. Therefore $\frac{F(\lambda,r)}{\lambda\cdot r}$ is the average cost per unit of duration when no distinction is made between different types of calls, and p_i^{AC} is the modified price which each accepted arrival of type i must pay based on this average cost.

In Tables 1-4 we report on results for the case of three demand types and various values of the underlying parameters. In all cases we assume a linear cost of capacity given by $g(k)=k/10$, and in Model 1 the pre-specified blocking probability is equal to $.01$. Since in Model 1, the cost function $\hat{F}^1(\lambda,r)$ does not depend on the cost of lost calls c, it follows that π_i^{SV} is independent of c in this case. Furthermore, (5.4) reveals that in Model 1, $\hat{F}^1(\lambda,r)$ depends on λ and r only through the inner product $\lambda\cdot r$.

λ	r	π	p^{AC}	$\hat{k}$
(10,10,10)	(1,1,1)	(.138,.138,.138)	(.138,.138,.138)	41
(20,10,10)	(1,1,1)	(.128,.136,.136)	(.132,.132,.132)	52
(10,10,10)	(2,1,1)	(.255,.136,.136)	(.264,.132,.132)	52
(30,20,10)	(1,1,1)	(.122,.125,.133)	(.125,.125,.125)	74
(10,10,10)	(3,2,1)	(.365,.250,.133)	(.374,.249,.125)	74

Table 1: Arrival Only SV and AC Prices for Model 1

It is instructive to compare π and p^{AC} in Tables 1 and 2. Roughly speaking the following result holds. π_i tends to be less than p_i^{AC} when $\lambda_i r_i$ is greater than $(\lambda\cdot r)/3$. Since $\lambda\cdot r$ is a measure of the total work transmitted over the system, $(\lambda\cdot r)/n$ measures the average work per customer type when there are n customers. Thus, as a general rule, "large" customers, i.e. type i calls with relatively large $\lambda_i r_i$, pay less than average cost under the Shapley value pricing rule.

[7] Such a pricing rule has been proposed in the context of pricing of broadband telecommunications services by Egan (1987, p. 489).

c	λ	r	π	p^{AC}	$\hat{k}$
(10,10,10)	(10,10,10)	(1,1,1)	(.159,.159,.159)	(.159,.159,.159)	48
(20,10,10)	(10,10,10)	(1,1,1)	(.167,.158,.158)	(.161,.161,.161)	48
(10,10,10)	(20,10,10)	(1,1,1)	(.146,.156,.156)	(.151,.151,.151)	60
(10,10,10)	(10,10,10)	(2,1,1)	(.281,.157,.157)	(.298,.149,.149)	60
(30,20,10)	(10,10,10)	(1,1,1)	(.170,.165,.158)	(.164,.164,.164)	49
(10,10,10)	(30,20,10)	(1,1,1)	(.138,.142,.151)	(.141,.141,.141)	85
(10,10,10)	(10,10,10)	(3,2,1)	(.394,.276,.156)	(.413,.275,.138)	82

Table 2: Arrival Only SV and AC Prices for Model 2

From Tables 1 and 2 we can see that in both models π_i is monotonically decreasing in λ_i and λ_j for $j \neq i$. Tables 3 and 4 illustrate more precisely the behavior of these prices with respect to changes in r_i. Recall that $\overline{\pi}_i = \pi_i / r_i$ is the price per accepted unit of duration time. Then holding the parameters c, λ, r_2 and r_3 constant, we observe that $\overline{\pi}_i$ is monotonically decreasing in r_i, in both Models 1 and 2. We also can observe that in Model 1, though not in Model 2, π_i is monotonically decreasing in r_j for $j \neq i$. Observe that $\overline{\pi}_2 = \pi_2$ and $\overline{\pi}_3 = \pi_3$ since we are analyzing the case $r_2 = r_3 = 1$.

r_1	π_1	π_2	π_3	$\overline{\pi}_1$	$\overline{\pi}_2$	$\overline{\pi}_3$
1	.138	.1383	.1383	.138	.1383	.1383
2	.255	.1361	.1361	.128	.1361	.1361
3	.369	.1349	.1349	.123	.1349	.1349
4	.480	.1341	.1341	.120	.1341	.1341
5	.590	.1335	.1335	.118	.1335	.1335
6	.698	.1331	.1331	.116	.1331	.1331
7	.806	.1327	.1327	.115	.1327	.1327
8	.913	.1324	.1324	.114	.1324	.1324
9	1.019	.1321	.1321	.113	.1321	.1321
10	1.125	.1320	.1320	.113	.1320	.1320

Table 3: Arrival Only SV Prices for Model 1:
r_1 variable, $r_2=r_3=1$, $c=\{10,10,10\}$, $\lambda=\{10,10,10\}$

r_1	π_1	π_2	π_3	$\bar{\pi}_1$	$\bar{\pi}_2$	$\bar{\pi}_3$
1	.159	.1590	.1590	.159	.1590	.1590
2	.281	.1572	.1572	.141	.1572	.1572
3	.398	.1566	.1566	.133	.1566	.1566
4	.512	.1566	.1566	.128	.1566	.1566
5	.624	.1568	.1568	.125	.1568	.1568
6	.734	.1572	.1572	.122	.1572	.1572
7	.842	.1577	.1577	.120	.1577	.1577
8	.950	.1583	.1583	.119	.1583	.1583
9	1.057	.1590	.1590	.117	.1590	.1590
10	1.163	.1597	.1597	.116	.1597	.1597

Table 4: Arrival Only SV Prices for Model 2:
r_1 variable, $r_2=r_3=1$, $c=\{10,10,10\}$, $\lambda=\{10,10,10\}$

6. Additional Issues

6.1 Asymmetric Treatment of Customer Classes

There are numerous game theoretic situations in which it is desirable to treat two players differently even though they "affect" the game in the same way. In the context of cost allocation, this means that two commodities that "affect" costs in the same way may be assessed different cost shares. For example, a regulator of a provider of telecommunications services might wish to treat residential and commercial service differently (in terms of cost shares) even if they affect costs the same way (e.g. when they enter the cost function only through their sum). We will now discuss a generalization of the Shapley value pricing approach, called weighted Shapley value pricing, that allows for differential pricing of commodities in situations like the one described above. First we introduce the weighted Shapley value.

Let $w = (w_1, \ldots, w_n) \in \Re_{++}^n$ and define a probability distribution $p^w(R)$ as follows: If $R = (j_1, \ldots, j_n)$ let

$$p^w(R) = \left[\frac{w_{j_1}}{w_{j_1}}\right]\left[\frac{w_{j_2}}{w_{j_1}+w_{j_2}}\right]\cdots\left[\frac{w_{j_{n-1}}}{\sum_{t=1}^{n-1} w_{j_t}}\right]\left[\frac{w_{j_n}}{\sum_{t=1}^{n} w_{j_t}}\right].$$

The weighted Shapley value is then defined as

$$\phi_i^w(v) = \sum_{R\in\Omega} p^w(R)\,\Delta_i(R,v).$$

If $w_1 = w_2 = \ldots = w_n$, then $p^w(R)\equiv 1/n!$ and we recover the (symmetric) Shapley value.

It is natural to generalize Shapley value pricing to weighted Shapley value pricing by defining

$$p_i^{wSV}(C,\alpha) = \frac{\phi_i^w(\hat{C})}{\alpha_i}$$

where $\hat{C}$ is the stand alone game derived from $(C\ ,\ \alpha)$. As we have discussed above, the weighted Shapley value price will allow substitute commodities to be assessed different cost shares. To further explore the nature of the non-symmetry of weighted Shapley value prices consider the following generalization of the cost share proportionality axiom.

(B.4w) *Weighted Cost Share Proportionality:* If $C(\alpha^{S\cup i}) = C(\alpha^{S\cup j})$ for all $S \subseteq N \setminus \{i,j\}$, then $\dfrac{\alpha_i p_i(C,\alpha)}{w_i} = \dfrac{\alpha_j p_j(C,\alpha)}{w_j}$.

In McLean and Sharkey (1992b) we proved:

Theorem: If $p : \mathcal{C}^n \rightarrow \Re^n$ satisfies A.1, B.2, B.3, and B.4w, if and only if, $p_i(C,\alpha) = \dfrac{\phi_i^w(\hat{C})}{\alpha_i}$ for all $i \in N$.

Other pricing schemes satisfy weighted cost share proportionality, such as $\overline{p}_i^w = \left(\dfrac{w_i}{\sum_{j\in N} w_j}\right)\dfrac{C(\alpha)}{\alpha_i}$.

Indeed $\dfrac{\alpha_i \overline{p}_i^w}{w_i} = \dfrac{\alpha_j \overline{p}_j^w}{w_j}$ for every commodity, even if the hypothesis of the weighted cost share proportionality axiom is not satisfied. In general, however, $\overline{p}^w$ and p^{wSV} are different because $\overline{p}^w$ does not satisfy discrete monotonicity.

Weighted SV prices provide additional flexibility for a system administrator who must find cost allocating prices for a cost function associated with an optimally designed system. For example, if the system administrator wishes to collect twice as much revenue from type i calls as from type j calls, even though calls of type i and j affect system costs identically, in the sense that $C(\alpha^{S\cup i}) = C(\alpha^{S\cup j})$ for all $S \subseteq N \setminus \{i,j\}$, then the weighted cost share proportionality axiom implies that a weight vector in which $w_i = 2w_j$ should be chosen, and this approach is justified by the full set of axioms as outlined in the above theorem. In particular, suppose that calls of types i and j have identical characteristics, in the sense that $\lambda_i = \lambda_j$, $r_i = r_j$, and $d_i = d_j$. Then weighted cost share proportionality guarantees that $p_i^{wSV} = 2p_j^{wSV}$, so that prices as well as cost shares are proportional to weights.

6.2 Subadditivity and Shapley Value Pricing

Subadditivity is a property of a cost function that has been used to characterize technologies that give rise to natural monopoly. Formally, a cost function C is subadditive if $C(x + y) \leq C(x) + C(y)$ for all non-negative output vectors x and y. Cost allocation problems in which the cost function is subadditive have the desirable property that SV pricing is "individually rational." That is, in a cost allocation problem (C, α), $\alpha_i p_i^{SV}(C,\alpha) \leq \hat{C}(\{i\}) = C(\alpha^{\{i\}})$. This follows immediately from the definition of the Shapley value, since subadditivity implies $\hat{C}(S \cup \{i\}) - \hat{C}(S) \leq \hat{C}(\{i\})$ for all $S \subset N\backslash\{i\}$.

It can be demonstrated, using a result of Smith and Whitt (1981), that the cost functions $\hat{F}^i(\lambda,r)$ and $\bar{F}^i(\lambda,r)$, $i = 1,2$, of section 5 are subadditive in λ whenever service times are equal for all customer types and $g(k)$ is itself subadditive. To see this, define $\beta^*(k,x) = \dfrac{x^k / k!}{\sum_{i=0}^{k} x^i / i!}$. For the case $n = 1$ it follows that $\beta(k;\lambda,r)$ equals $\beta^*(k,\lambda r)$. Smith and Whitt demonstrate that the function $L(k, \lambda, r) = \lambda\beta^*(k,\lambda r)$ is subadditive in k and λ. That is,

$$L(k'+k'', r, \lambda'+\lambda'') \leq L(k', r, \lambda') + L(k'', r, \lambda'') \tag{6.1}$$

where k', k'', λ', and λ'' are non-negative real numbers. To demonstrate subadditivity for the general n customer case, let λ' and λ'' be vectors of arrival rates for multiple customer types having a common service time $r_i = \bar{r}$. From (5.2) we have $\hat{F}^1(\lambda',r) = g(k')$ and $\hat{F}^1(\lambda'',r) = g(k'')$ for some values of k' and k'' such that $\beta^*(k',\bar{r}\sum_{j=1}^{n}\lambda_j') \leq \bar{\beta}$ and $\beta^*(k'',\bar{r}\sum_{j=1}^{n}\lambda_j'') \leq \bar{\beta}$. From (6.1) it follows that

$$\beta^*\left[k'+k'',\bar{r}\left(\sum_{j=1}^{n}\lambda_j' + \sum_{j=1}^{n}\lambda_j''\right)\right] \leq \frac{\sum_{j=1}^{n}\lambda_j'}{\sum_{j=1}^{n}\lambda_j' + \sum_{j=1}^{n}\lambda_j''}\beta^*\left[k',\bar{r}\sum_{j=1}^{n}\lambda_j'\right] + \frac{\sum_{j=1}^{n}\lambda_j''}{\sum_{j=1}^{n}\lambda_j' + \sum_{j=1}^{n}\lambda_j''}\beta^*\left[k'',\bar{r}\sum_{j=1}^{n}\lambda_j''\right] \leq \bar{\beta}$$

Therefore $k' + k''$ is feasible for $(\lambda' + \lambda'',r)$, and given subadditivity of g, $\hat{F}^1(\lambda'+\lambda'',r) \leq g(k'+k'') \leq g(k') + g(k'') = \hat{F}^1(\lambda',r) + \hat{F}^1(\lambda'',r)$. Similar arguments demonstrate subadditivity of $\hat{F}^2$. Since Smith and Whitt demonstrate that delay has a similar subadditivity property, the functions $\bar{F}^1$ and $\bar{F}^2$ are subadditive as well.

The above arguments demonstrate that in the arrival only model with $r_i = \bar{r}$ for every customer type, SV pricing satisfies the individual rationality property, so that $\lambda_i p_i^{SV}(G,\lambda) \leq G(\lambda^{\{i\}})$ for every customer i. Hence, in the arrival only model with equal service times, no customer is assigned a cost share which exceeds the "stand alone" cost of serving that customer by himself. Finally, we note that whenever the cost function C is subadditive, these same arguments demonstrate that weighted SV prices are also "individually rational," i.e. $\alpha_i p_i^{wSV}(C,\alpha) \leq C(\alpha^{\{i\}})$.

The individual rationality property of Shapley value pricing is weaker than the property of "subsidy freeness" or "sustainability" of a pricing rule. Specifically, a pricing rule p in a cost allocation problem (C,α) is subsidy free if, for every subset S of services, $\sum_{i\in S} \alpha_i p_i(C,\alpha) \leq C(\alpha^S)$. It follows that the vector of revenue shares $(\alpha_1 p_1 , \ldots , \alpha_n p_n)$ is contained in the core of the stand alone game $\hat{C}$. Subadditivity is not sufficient for subsidy freeness, and a stronger property of the cost function is required. A cost function is "submodular" if $\frac{\partial^2 C}{\partial x_i \partial x_j} \leq 0$. If C is submodular then C is subadditive, and both the Aumann-Shapley and Shapley value pricing rules can be shown to be subsidy free. It is possible that cost functions which arise from telecommunications design problems are approximately submodular in many cases. It is, however, beyond the scope of the present paper to pursue these matters further.

6.3 Demand Compatible Shapley Value Pricing

Shapley value prices are not necessarily consistent with consumer demand functions. That is, when applied to arbitrary output vectors α it is necessary to assume that customer demands at α are completely inelastic with respect to price. In the context of Aumann-Shapley pricing, Mirman and Tauman (1982) have demonstrated that it is possible to embed the AS price mechanism in an equilibrium setting such that, given appropriate continuity assumptions, there exists at least one output vector α for which Aumann-Shapley prices are also equilibrium prices. That is, given $p^{AS}(C,\alpha)$, utility maximization by consumers is consistent with the aggregate output α. In practice, one could compute a "demand compatible" Aumann-Shapley price vector by solving an appropriate fixed point problem.

The approach in Mirman and Tauman (1982) can be adapted to Shapley value pricing, and we will now briefly outline their "model B" in our context. Suppose that customer $i \in N$ is endowed with y_i dollars of exogenous income. A customer's utility u_i is a function of the rate λ_i at which he makes calls, as well as the average duration r_i of each call that he makes. Given prices for arrival rates and usage, the customer maximizes utility subject to the budget constraint determined by these prices, and his income y_i. A SV pricing equilibrium consists of demands $\bar{\lambda}$ and $\bar{r}$ such that for each $i \in N$ $(\bar{\lambda}_i , \bar{r}_i)$ solves

$$\max \; u_i(\lambda_i, r_i)$$
$$\text{s.t.}$$
$$p^{SV}_{\lambda_i}\left(F,(\bar{\lambda},\bar{r})\right)\lambda_i + p^{SV}_{r_i}\left(F,(\bar{\lambda},\bar{r})\right)r_i \le y_i$$
$$\lambda_i, r_i \ge 0.$$

Thus, a SV pricing equilibrium formalizes the notion of demand compatible SV prices: when consumers react to the SV price vector $p^{SV}\left(F,(\bar{\lambda},\bar{r})\right)$, the resulting "utility maximizing demands" will be precisely $(\bar{\lambda}, \bar{r})$. Mirman and Tauman prove the existence of an AS pricing equilibrium under the standard economic assumptions of continuity and quasiconcavity of utility functions, etc. Although we will not pursue the issue further, one could use their techniques, and prove the existence of a SV pricing equilibrium under similar assumptions.

7. Conclusion

In this paper, we have presented two cost allocation methodologies which may be used to determine prices for a regulated telecommunications supplier offering heterogeneous services on a broadband network. In our approach, the system planner must first determine an optimal telecommunications system as a function of customer characteristics, and then must choose a mechanism for allocating the cost of this optimally designed system. We have addressed the first issue by studying various optimization problems. To analyze the second issue of cost allocation, we have proposed pricing rules that are characterized by sets of plausible axioms that one could argue should be satisfied by any pricing rule. One of these approaches leads to Aumann-Shapley pricing, which has been extensively studied in the literature. The other approach leads to a pricing rule based on the Shapley value of a related cooperative game. While these approaches are similar in motivation, they differ in the technical requirements which must be imposed on the underlying cost function, and we have argued that Shapley value pricing is more appropriate in a telecommunications context. These methodologies can be used to determine Shapley value prices for customers who differ from one another in terms of their arrival rates, service rates, costs of lost work, and the number of simultaneous channels which they require.

To illustrate our methodology, we have presented numerical results for two kinds of cost allocation problems associated with a queuing system in which buffer capacity is equal to zero and server requirements are homogeneous. In the "arrival only" model, we treat the arrival rates as "demands" and determine a price per arrival so that expected total revenue is equal to expected cost. In the "arrival/usage" model, we treat both the arrivals and durations as "commodities" and compute prices for both arrivals and units of service time so that expected revenue equals expected cost.

We believe that the cost allocation approach, developed in this paper, provides useful insights into the pricing of broadband services. In particular, where a group of heterogeneous services are offered, which are differentiated in terms of arrival rates, service times and requirements for channel capacity, our cost allocation approach offers a methodology for setting relative prices for service classes that complements incentive approaches to pricing that have been previously considered in the literature.

References

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

Billera, L.J. and D.C. Heath (1982), "Allocation of Shared Costs: A Set of Axioms Yielding a Unique Procedure," *Mathematics of Operations Research*, 7, 32-9.

Billera, L.J., D.C. Heath, and J. Raanan (1978), "Internal Telephone Billing Rates–A Novel Application of Non-Atomic Game Theory," *Operations Research, 26*, 956-65.

Boiteux, M. (1956), "Sur la gestion des Monopoles Publics astreints a l'equilibre budgetaire," *Econometrica, 24*, 22-40.

Chakravorti, B. (1993), "Optimal Flow control of an M/M/1 Queue with a Balanced Budget," submitted to *IEEE Transactions on Automatic Control.*

Dolan, R.J. (1978), "Incentive Mechanisms for Priority Queuing Problems," *Bell Journal of Economics, 9*, 421-36.

Egan, B. (1987), "Costing and Pricing the Network of the Future," *International Switching Symposium, 1987 Proceedings, Vol. 2*, 483-90.

Federgruen, A. and H. Groenevelt (1988), "Characterization and Optimization of Achievable Performance in General Queuing Systems," *Operations Research, 36*, 733-41.

Hong, D. and T. Suda (1991), "Congestion Control and Prevention in ATM Networks," *IEEE Network Magazine*, July, 1991, 10-16

McLean, R.P. and W.W. Sharkey (1992a), "Alternative Methods of Cost Allocation for Stochastic Service Systems," unpublished manuscript.

McLean, R.P. and W.W. Sharkey (1992b), "Weighted Cost Allocation Mechanisms," unpublished manuscript.

Mendelson, H. and S. Whang (1990), "Optimal Incentive-Compatible Priority Pricing for the M/M/1 Queue," *Operations Research, 38*, 870-83.

Mirman, L. and Y. Tauman (1982), "Demand Compatible, Equitable, Cost Sharing Prices," *Mathematics of Operations Research, 7* , 40-56.

Naor, P. (1969), "On the Regulation of Queue Size by Levying Tolls," *Econometrica, 37*, 15-24.

Roberts, J.W. (1981), "A Service System with Heterogeneous User Requirements–Application to Multi-Services Telecommunications Systems," in *Performance of Data Communications Systems and their Applications*, G. Pujolle, ed. Amsterdam: North Holland Publishing Co.

Sanders, B.A. (1985), "A Private Good/Public Good Decomposition for Optimal Flow Control of an M/M/1 Queue," *IEEE Transactions on Automatic Control*, AC-30, 1143-5.

Smith, D.R. and W. Whitt (1981), "Resource Sharing for Efficiency in Traffic Systems," *Bell System Technical Journal, 60*, 39-55.