REPRINTED FROM

TELECOMMUNICATIONS DEMAND MODELLING

An Integrated View

Edited by:

Alain DE FONTENAY
Bell Communications Research
Livingston, New Jersey, U.S.A.

Mary H. SHUGARD
Bell Communications Research
Livingston, New Jersey, U.S.A.

David S. SIBLEY
Bell Communications Research
Morristown, New Jersey, U.S.A.

1990

NORTH-HOLLAND
AMSTERDAM • NEW YORK • OXFORD • TOKYO

TELECOMMUNICATIONS DEMAND MODELLING
An Integrated View
A. de Fontenay, M.H. Shugard, D.S. Sibley (Editors)

PERFORMANCE INDICATORS FOR REGULATED INDUSTRIES

by W.E. Diewert
Dept. of Economics
Univ. of British Columbia

1. INTRODUCTION

How should a natural monopoly (such as telephone, electricity, water, natural gas and some transportation services) be regulated in the interests of society? This question has been discussed by economists for well over 100 years, but there is still no agreement on how to answer it. Fifty years ago, economists thought that forcing the monopolist to sell each product at a price equal to marginal cost would suffice to bring about an "ideal" allocation of resources.[1] The deficits that this policy would lead to for enterprises with large fixed costs were to be supported by the general taxation powers of the state.

The price equals marginal cost solution to the control of a monopoly was successfully attacked on a number of grounds: (i) Fleming [1945; 336] and Wilson [1945; 456] pointed out that there was no natural index available to evaluate the performance of managers, or put another way: how can the regulator know what each marginal cost is? (ii) Wilson [1945; 457] also raised a problem that was later stressed by Domar [1974; 4]: how can the regulator motivate the manager of the regulated industry to take the "right" course of action? (iii) Wilson [1945; 458-459] also noted the fact that any deficits generated by the regulated industry were to be covered out of general revenue and this would create a tremendous incentive for an empire building manager to expand unduly. In the end, political bargaining would determine the allocation of resources due to the imprecision of the price equals marginal cost rule in an intertemporal context. (iv) Coase [1945; 113] [1946; 176] pointed out that the Pigou-Hotelling-Lerner solution to the monopoly problem would redistribute income to consumers of products in which fixed costs form a high proportion of total costs. (v) Finally, Hotelling [1939; 155] and Coase [1946; 179] both noted that if taxes on fixed factors could not cover the government's revenue needs, then covering the deficits of regulated enterprises will have adverse efficiency effects in the rest of the economy. Thus marginal cost pricing, by itself, does not seem to be a solution to the problem of regulating monopolies. Scott [1952] [1978; 154] noted that the level of profits serves as an adequate indicator for the performance of a firm that is subject to constant or diminishing returns to scale (or, in more modern terminology, we would say that the technology set of the firm is convex). However, profit is a totally inadequate performance indicator for a firm that is subject to increasing returns to scale (or more generally has a nonconvex technology set). Thus Scott [1978; 155] suggested using the change in profits (where inputs and outputs are valued at constant prices) as an appropriate indicator of firm performance for a firm that has the potential to engage in monopolistic behaviour. This criterion has the advantage that it depends only on observable price and quantity data (and not on difficult to observe things like "marginal cost" or "the elasticity of demand"). Similar performance indicators for regulated firms have been developed by Vogelsang and Finsinger [1979], Finsinger and Vogelsang [1981] [1982] [1985], Tam [1981] [1985], Gravelle [1985] and Sappington and Sibley [1985].

Setting price equal to marginal cost is a necessary condition for solving a somewhat vaguely specified social maximization problem. The new literature on performance indicators makes further progress by specifying alternative objective functions or bonus functions that a manager of a regulated enterprise is supposed to maximize (subject to various constraints specified by the social planner). In most cases, an ideal objective function is not empirically observable; in such cases, it is necessary to form approximations to the ideal functions that are observable. The specification of an approximation for a bonus function is called an incentive scheme.

The literature on performance indicators and incentive schemes seems very promising but there are at least two problems with it: (i) the welfare measures are either vaguely specified or are based on the Dupuit [1844] and Marshall [1920] consumer and producer surplus concepts (i.e., areas under market demand and supply curves) and (ii) the schemes are partial equilibrium in nature. Thus the primary purpose of the present paper is the development of performance indicators for firms that are subject to increasing returns to scale (or a nonconvex technology) in a rigorous general equilibrium setting.

We conclude this section by outlining the contents of subsequent sections.

In section 2, we introduce our notation and assumptions that describe a general equilibrium model of a region that has a public utilities sector and trades with the rest of the world, taking world prices as fixed.

In section 3, we show how a Pareto optimal allocation of resource may be characterized. We also define various marginal cost and willingness to pay functions that will plan a role in our subsequent analysis.

In section 4, we consider various general equilibrium bonus functions or performance indicators for the manager of the regulated sector. At first sight, it may seem that we have handed the manager the impossible task of optimizing performance over the entire economy. However, this pessimistic view neglects the fact that since we will assume that the nonutilities part of the economy behaves competitively, market prices summarize the required information about the allocation of resources in these markets in an efficient manner, a point stressed by Arrow [1964] and Arrow and Hurwicz [1977]. Thus in sections 5 and 6, we develop first and second order approximations to our theoretical bonus functions. These approximations turn out to depend only on observable price and quantity data.

Since most firms produce many outputs and utilize many inputs, our general equilibrium model contains a rather large number of variables. For readers who are unaccustomed to dealing with a sea of symbols, we present a geometric treatment of our analysis (for a special case of very low dimensionality) in section 7.

Section 8 concludes with some comments on various performance indicators that have been suggested by other authors.

2. A REGIONAL GENERAL EQUILIBRIUM MODEL WITH A REGULATED SECTOR

Consider an economy that occupies a definite geographical area. We assume that there are three classes of goods in this economy: (i) I interregionally traded goods, (ii) M utility services (e.g., telephone, water, gas, electricity, sewage, cablevision, postal, bus, rail and air services, etc.) and (iii) N other domestic goods (types of labour, capital, natural resource, transportation and retailing services).

There are four classes of economic agent in our regional economy: (i) K competitive firms, (ii) H households, (iii) one consolidated utilities sector which supplies the M utility services mentioned in the previous paragraph in a possibly monopolistic manner, and finally, (iv) a consolidated government sector which supplies certain public goods such as defense, law administration, protection services and possibly education and medical services. The government sector also taxes the other economic agents and may also provide subsidies and grants. The competitive producers take prices as given and maximize profits. We assume that these producers have convex technology sets. Each household takes interregional and local prices (and commodity taxes) as given and maximizes a utility function subject to a budget constraint.

There are three sets of feasibility constraints in the regional economy which correspond to the three classes of goods.

The first such constraint is:

$$(1) \qquad w\cdot[\Sigma_{k=0}^{K} x^k - \Sigma_{h=0}^{H} a^h + \Sigma_{h=0}^{H} \bar{a}^h] \geq 0$$

where $w \equiv [w_1,\ldots,w_I] >> 0_I$ is a strictly positive vector of "world" prices that the regional economy faces; $x^k \equiv [x_1^k,\ldots,x_I^k]$ is sector k's net output vector of interregionally traded goods for k = 0, 1,...,K where sector 0 is the public utilities sector and sectors 1 to K are the competitive firms (if $x_i^k < 0$, then sector k is utilizing good i as an input); $a^h \equiv [a_1^h,\ldots,a_I^h]$ is the net consumption vector of household h for interregionally traded goods (household supplies are indexed with a negative sign) for h = 1,...,H while a^0 is the government's net requirements vector for traded goods (government supplies are indexed with a negative sign); $\bar{a}^h \equiv [\bar{a}_1^h,\ldots,\bar{a}_I^h] \geq 0$ is household h's nonnegative endowment vector of traded goods for h = 1,...,H, and $\bar{a}^0$ is the government's endowment vector. Constraint (1) says that exports minus imports in the regional economy evaluated at world prices should be nonnegative.[2]

Our second set of constraints is

$$(2) \qquad \Sigma_{k=0}^{K} y^k - \Sigma_{h=0}^{H} b^h + \Sigma_{h=0}^{H} \bar{b}^h \geq 0_M$$

where $y^0 \equiv [y_1^0,\ldots,y_M^0]$ is a nonnegative vector of public utility services produced by the regulated sector; $y^k \equiv [y_1^k,\ldots,y_M^k] \leq 0_M$ is sector k's demand for public utility services vector for k = 1,...,K (demands are indexed by negative signs); $b^h \equiv [b_1^h,\ldots,b_M^h] \geq 0_M$ is household h's demand for public utility services for h = 1,...,H (household demands are indexed by positive signs); $b^0 \geq 0_M$ is the government sector's demand for public utility services, and $\bar{b}^h \equiv [\bar{b}_1^h,\ldots,\bar{b}_M^h] \geq 0_M$ is household h's endowment vector of public utility services for h = 1,...,H while $\bar{b}^0 \geq 0_M$ is the government's endowment vector (if utility service m cannot be stored, then $\bar{b}_m^h = 0$ for all h). The M constraints in (2) just tell us that the supplies of public utility services should be equal to or greater than the corresponding demands by firms, households and the government.

The third set of feasibility constraints is:

$$(3) \qquad \Sigma_{k=0}^{K} z^k - \Sigma_{h=0}^{H} c^h - \Sigma_{h=0}^{H} \bar{c}^h \geq 0_N$$

where $z^k \equiv [z_1^k,\ldots,z_N^k]$ is sector k's net supply vector for local goods (if sector k is utilizing good n as an input, then $z_n^k < 0$) for k = 0,1,...,K; $c^h \equiv [c_1^h,\ldots,c_N^h]$ is household h's (or the government's if h = 0) net demand vector for local goods and services (if household h supplies a type of labour service that corresponds to good n, then $c_n^h < 0$), and $\bar{c}_1^h \equiv [\bar{c}_1^h,\ldots,\bar{c}_N^h] \geq 0$ is household h's (or the government's if h = 0) nonnegative endowment vector of local goods. If local good n is a type of labour service, then we assume $\bar{c}_n^h = 0$ for all h; i.e., labour services cannot be stored. The N constraints in (3) say that the supply of each local good should not be less than the corresponding demand.

We assume that the preferences of household h can be represented by a quasiconcave, continuous, nondecreasing utility function f^h defined over a closed, convex set Ω^h which is a subset of R^{I+M+N} for $h = 1,\ldots,H$.

The technology sets of the competitive sectors are closed convex subsets S^k of R^{I+M+N} for $k = 1,\ldots,K$. The technology set for the public utility sector is a closed subset S^0 of R^{I+M+N}. We also assume that each technology set satisfies a free disposal property. We do not assume that S^0 is necessarily convex, since this would rule out the increasing returns to scale or decreasing cost phenomena that characterize the provision of public utility services in real life economies.

3. PARETO OPTIMALITY AND EFFICIENCY

Consider the following constrained maximization problem:

$$(4)\quad \max_{a^h,b^h,c^h,x^k,y^k,z^k} \{f^1(a^1,b^1,c^1) : (1),(2),(3);\ (x^k,y^k,z^k) \in S^k \text{ for } k = 0,1,\ldots,K;\ f^h(a^h,b^h,c^h) \geq u^*_h \text{ for } h = 2,\ldots,H;\ b^h \geq 0_M \text{ and } (a^h,b^h,c^h) \in \Omega^h \text{ for } h = 1,2,\ldots,H;\ y^0 \geq 0_M \text{ and } y^k \leq 0_M,\ k=1,\ldots,K\}.$$

In (4), the government net demand vectors a^0, b^0 and c^0 are held fixed. The above constrained maximization problem may be described as follows. We attempt to maximize the welfare of household 1 subject to the following constraints: (i) the demand equal to or less than supply constraints (1), (2) and (3) in the previous section must be satisfied; (ii) each household (except of course household 1) must attain at least a prespecified level of welfare u^*_h for $h = 2,\ldots,H$, and (iii) each sector must choose a technologically feasible set of inputs and outputs.

Under appropriate regularity conditions, a finite maximum for the objective function in (4) will exist; call this maximum u^*_1. The production

vectors (x^{k*},y^{k*},z^{k*}) for $k = 0,1,\ldots,K$ and the net consumption vectors (a^{h*},b^{h*},c^{h*}) for $h = 1,\ldots,H$ that solve (4) represent a Pareto optimal allocation of resources for the regional economy; i.e., no household's welfare can be increased without decreasing the utility or welfare level of one or more other households. In an ideal world, we would like to restrict attention to these Pareto optimal equilibria.

We shall find it convenient to characterize efficient equilibria in a somewhat different way. Consider the following constrained maximization problem:

$$(5)\quad \max_{a^h,b^h,c^h,x^k,y^k,z^k} \{\Sigma_{k=0}^K w\cdot x^k - \Sigma_{h=0}^H w\cdot a^h + \Sigma_{h=0}^H w\cdot \bar{a}^h : (2),(3);\ (x^k,y^k,z^k) \in S^k \text{ for } k = 0,1,\ldots,K;\ f^h(a^h,b^h,c^h) \geq u^*_h,\ b^h \geq 0_M \text{ and } (a^h,b^h,c^h) \in \Omega^h \text{ for } h = 1,\ldots,H;\ y^0 \geq 0_M \text{ and } y^k \leq 0_M \text{ for } k=1,\ldots,M\}.$$

The u^*_h which appear in (5) are the same as the u^*_h in (4). In both problems, we hold government net demand (a^0,b^0, c^0) fixed. In problem (5), we no longer have the foreign exchange constraint (1). However in (5), we are now attempting to maximize the amount of foreign exchange that the regional economy can produce subject to the aggregate demand less than supply constraints (2) and (3), the technology constraints for each sector and subject to household h attaining the welfare level u^*_h for $h = 1,\ldots,H$. Note that (5) has an extra utility constraint, $f^1(a^1,b^1,c^1) \geq u^*_1$, that was not present in (4).

If $f^1(a,b,c)$ is increasing in the components of the vector a, then it can be verified that the solution sets to (4) and (5) coincide.

To make further progress, we make two sets of definitions.

For $k = 0,1,\ldots,K$ and $w \gg 0_I$, define the sector k restricted profit function π^k by

(6) $\pi^k(w,y^k,z^k) \equiv \max_x \{w \cdot x : (x,y^k,z^k) \in S^k\}$.

If there is no x such that $(x,y^k,z^k) \in S^k$, define $\pi^k(w,y^k,z^k) = -\infty$. Our free disposability assumptions on the sets S^k imply that $\pi^k(w,y^k,z^k)$ is non-increasing in the components of y^k and z^k; see Diewert [1973; 293] for a proof.

For h = 1,...,H, define the household h restricted expenditure function e^h by

(7) $e^h(u_h,w,b^h,c^h) \equiv \min_a \{w \cdot a : f^h(a,b^h,c^h) \geq u_h, (a,b^h,c^h) \in \Omega^h\}$.

If there is no vector a such that the constraints in (7) are satisfied, then define $e^h(u_h,w,b^h,c^h) = +\infty$. Our assumption that the utility function f^h is nondecreasing in its arguments implies that $e^h(u_h,w,b^h,c^h)$ is nonincreasing in the components of b^h and c^h; see Diewert [1986; 172] for a proof.

Now let us fix y^k,z^k for k = 0,1,...,K and fix b^h,c^h for h = 1,...,H and maximize (5) with respect to x^k, k = 0,1,...,K, and a^h, h = 1,...,H. Using definitions (6) and (7), it can be seen that we may rewrite (5) as follows:

(8) $$B(w,u^*) = \max_{y^k,z^k,b^h,c^h} \{\Sigma_{k=0}^K \pi^k(w,y^k,z^k) - \Sigma_{h=1}^H e^h(u_h^*,w,b^h,c^h)$$

$$- w \cdot a^0 + \Sigma_{h=0}^H w \cdot \bar{a}^h : (2),(3);\ b^h \geq 0_M,\ h = 1,\ldots,H;\ y^0 \geq 0_M \text{ and}$$

$$y^k \leq 0_M \text{ for } k = 1,\ldots,K\} .$$

Problems (5) and (8) may be regarded as central planning problems for the regional economy. In both problems, the planner attempts to maximize the amount of foreign exchange that the economy can produce subject to the constraints of technology and subject to a prespecified welfare level constraint for each household.

In order to characterize an optimal solution for (8), we make the

somewhat restrictive assumption that the π^k and the e^h are once differentiable with respect to their arguments when evaluated at a solution to (8).[3] Let (y^{k*},z^{k*}), k = 0,1,...,K and (b^{h*},c^{h*}), h = 1,...,H solve (8). Then assuming that an appropriate constraint qualification condition is satisfied, the Kuhn-Tucker [1951] Theorem gives us the existence of multiplier vectors p* and r* such that the following Kuhn-Tucker conditions are satisfied:

(9) $\nabla_y\pi^0(w,y^{0*},z^{0*}) + p^* \leq 0_M$; $y^{0*} \geq 0_M$; $y^{0*} \cdot [\nabla_y\pi^0(w,y^{0*},z^{0*}) + p^*] = 0$;

(10) $\nabla_y\pi^k(w,y^{k*},z^{k*}) + p^* \geq 0_M$; $y^{k*} \leq 0_M$; $y^{k*} \cdot [\nabla_y\pi^k(w,y^{k*},z^{k*}) + p^*] = 0$, k = 1,...,K;

(11) $-\nabla_b e^h(u_h^*,w,b^{h*},c^{h*}) - p^* \leq 0_M$; $b^{h*} \geq 0_M$; $b^{h*} \cdot [\nabla_b e^h(u_h^*,w,b^{h*},c^{h*}) + p^*] = 0$, h = 1,...,H;

(12) $p^* \geq 0_M$; $\Sigma_{k=0}^K y^{k*} - b^0 - \Sigma_{h=1}^H b^{h*} + \Sigma_{h=0}^H \bar{b}^h \geq 0_M$; $p^* \cdot [\Sigma_{k=0}^K y^{k*} - b^0 - \Sigma_{h=1}^H b^{h*} + \Sigma_{h=0}^H \bar{b}^h] = 0$;

(13) $\nabla_z\pi^k(w,y^{k*},z^{k*}) + r^* = 0_N$, k = 0,1,...,K ;

(14) $-\nabla_c e^h(u_h^*,w,b^{h*},c^{h*}) - r^* = 0_N$, h = 1,...,H ;

(15) $r^* \geq 0_N$; $\Sigma_{k=0}^K z^{k*} - c^0 - \Sigma_{h=1}^H c^{h*} + \Sigma_{h=0}^H \bar{c}^h \geq 0_N$; $r^* \cdot [\Sigma_{k=0}^K z^{k*} - c^0 - \Sigma_{h=1}^H c^{h*} + \Sigma_{h=0}^H \bar{c}^h] = 0$;

where $\nabla_y\pi^k(w,y^{k*},z^{k*}) \equiv [\partial\pi^k/\partial y_1^k,\ldots,\partial\pi^k/\partial y_M^k]$ is the vector of first order partial derivatives of π^k with respect to the components of y, etc.
If the solution vectors $y^{0*},b^{1*},\ldots,b^{H*}$ are strictly positive and the solution vectors $y^{1*},\ldots,y^{K*}$ are strictly negative, then conditions (9), (10) and (11) reduce to the following more familiar Lagrange multiplier conditions:

(16) $p^* = -\nabla_y\pi^k(w,y^{k*},z^{k*}) = -\nabla_b e^h(u_h^*,w,b^{h*},c^{h*})$; k = 0,1,...,K; h=1,...,H.

When k = 0, $-\nabla_y\pi^0(w,y^0,z^0) \equiv p^0(w,y^0,z^0) \equiv [p_1^0(w,y^0,z^0),\ldots,p_M^0(w,y^0,z^0)] \geq 0_M$

is a vector of <u>quasi marginal cost functions</u> for producing marginal units of the various types of utility services. Thus the first set of conditions in (16), $p^* = p^0(w,y^{0*},z^{0*})$, is the counterpart to the familiar price equals marginal cost rule that occurred in the early Pigou-Hotelling-Lerner literature. We call these functions quasi marginal cost functions because they are not equal to the usual marginal cost functions that occur in the literature. The true (net) cost function for sector 0 equals $-\tilde{\pi}^0$, where $\tilde{\pi}^0(w,y^0,r) \equiv \max_z \{\pi^0(w,y^0,z) + r \cdot z\}$, and the system of true marginal cost functions for sector 0 is defined by $\tilde{p}(w,y^0,r) \equiv -\nabla_r \tilde{\pi}^0(w,y^0,r)$. This distinction between quasi and true is not a semantic one, since Arrow and Hurwicz [1977; 81-82] showed that an efficient allocation of resources will not in general be achieved by marginal cost pricing. This very important point was elaborated on by Guesnerie [1975; 13-14] and Arnott and Harris [1977]. Of course, the distinction between quasi and true vanishes if N=0 so that there are no domestic goods. For k>0, $-\nabla_y \pi^k(w,y^k,z^k) \equiv p^k(w,y^k,z^k)$ is sector k's vector of <u>producer willingness to pay functions</u>. Alternatively, $p^k(w,y^k,z^k)$ can be regarded as a system of (inverse) demand functions for public utility services for k = 1,...,K.[4] For h = 1,...,H, $-\nabla_b e^h(u_h^*,w,b^h,c^h) \equiv P^h(u_h^*,w,b^h,c^h)$ is a vector of household h <u>consumer willingness to pay functions</u> for marginal units of the various types of utility services. Under our hypotheses on the utility functions f^h, it can be shown that[5] $P^h(u_h^*,w,b^h,c^h)$ is a system of inverse conditional Hicksian [1946; 331] demand for public utility services functions (conditional on c^h) for household h.

Equations (13) and (14) may be rewritten as follows:

$$(17)\quad r^* = -\nabla_z \pi^k(w,y^{k*},z^{k*}) = -\nabla_c e^h(u_h^*,w,b^{h*},c^{h*});\ k = 0,1,\ldots,K;\ h=1,\ldots,H$$

where $r^* \equiv [r_1^*,\ldots,r_N^*]$. Equations (17) may be interpreted as follows. For k = 0,1,...,K, $-\nabla_z \pi^k(w,y^k,z^k) \equiv r^k(w,y^k,z^k)$ is a vector of <u>net quasi marginal cost functions</u> for sector k for producing extra units of the various domestic goods (if $z_n^k < 0$ so that the nth domestic good is being used as an input by sector k, then $r_n^k(w,y^k,z^k) \equiv -\partial\pi^k(w,y^k,z^k)/\partial y_n^k \geq 0$ is the sector k willingness to pay for an extra unit of this input). For h = 1,...,H, $-\nabla_c e^h(u_h^*,w,b^h,c^h) \equiv R^h(u_h^*,w,b^h,c^h)$ is a vector of household h <u>conditional net willingness to pay functions</u> for marginal units of the various domestic goods (if $c_n^h < 0$ so that household h is supplying the nth domestic good as a type of labour service, then $R_n^h(u_h^*,b^{h*},c^{h*}) \equiv -\partial e^h(u_h^*,w,b^{h*},c^{h*})/\partial c_n^h \geq 0$ is the wage rate that will just induce household h to supply an extra unit of this service without causing a change in household h's welfare level). Thus in an efficient equilibrium in our regional economy, for each good n, each sector k will equate the net quasi marginal cost of producing an extra unit of domestic good n to r_n^* and each household h will equate its net conditional willingness to pay for an extra unit of good n to r_n^* also. These conditions (17) usually do not appear in the regulation of monopolies literature, since most of the analysis has taken place in a partial equilibrium environment[6] which neglected factor markets and interregional trade.

If the reader prefers to work with a model of a closed economy, this can be accomplished as follows: set I = 1 and w = 1 and interpret the single good in category one as a domestic numeraire good that is either produced or utilized by every producer and household. Under this interpretation, the profit functions π^k become ordinary production functions and the restricted expenditure functions $a^h = e^h(u_h,b^h,c^h)$ are inverse functions for the utility functions $u_h = f^h(a^h,b^h,c^h)$. The reader should now go back and reinterpret the constrained maximization problems (5) and (8).

We conclude this section by providing interpretations for the vectors of partial derivatives of the restricted expenditure and profit functions with respect to the components of the vector of interregional prices w.

If $e^h(u_h^*,w,b^{h*},c^{h*})$ is differentiable with respect to the components of w, and if a^{h*} denotes the a^h solution to (5) or (7), then:

(18) $a^{h*} = \nabla_w e^h(u_h^*, w, b^{h*}, c^{h*})$. (Shephard's Lemma).

The derivative property (18) is a modification of a property due originally to Hotelling [1932; 594] and Shephard [1953; 11]; see Diewert [1986; 171]. In a similar manner, it can be shown[7] that if $\pi^k(w, y^{k*}, z^{k*})$ is differentiable with respect to the components of w, then the solution vector x^{k*} of interregionally traded goods to the maximization problems defined by (5) or (6) is equal to the vector of partial derivatives of π^k with respect to the components of w; i.e.,

(19) $x^{k*} = \nabla_w \pi^k(w, y^{k*}, z^{k*})$. (Hotelling's Lemma).

4. ALTERNATIVE OBJECTIVE FUNCTIONS FOR THE REGULATED SECTOR

Consider an observed period 0 equilibrium for our regional economy. Suppose that the firm net output vectors for the three classes of goods (interregional, public utility, and other domestic) are x^{k0}, y^{k0}, z^{k0} for $k = 0,1,\ldots,K$. Suppose that the observed household net demand vectors for the three classes of good are a^{h0}, b^{h0}, c^{h0} for $h = 1,\ldots,H$. Suppose further that the period 0 world price vector is $w^0 \gg 0_I$ and that each producer and consumer is competitively optimizing with respect to the interregionally traded goods, conditional on their allocation of domestic goods and public utility services. Thus if we denote the welfare level of household h by u_h^0, we have

(20) $u_h^0 = f^h(a^{h0}, b^{h0}, c^{h0})$; $w^0 \cdot a^{h0} = e^h(u_h^0, w^0, b^{h0}, c^{h0})$; $h = 1,\ldots,H$ and

(21) $w^0 \cdot x^{k0} = \pi^k(w^0, y^{k0}, z^{k0})$; $k = 1,\ldots,K$.

Let us define $\alpha = a^0 - \Sigma_{h=0}^H \bar{a}^h$, $\beta \equiv b^0 - \Sigma_{h=0}^H \bar{b}^h$ and $\gamma \equiv c^0 - \Sigma_{h=0}^H \bar{c}^h$. These are net government demand vectors less the corresponding endowment vectors. We shall hold α, β and γ fixed throughout the analysis which follows in subsequent sections.

The period 0 allocation of public utility services and other domestic goods is summarized by the following equations (we assume that there are no free goods):

(22) $\Sigma_{k=0}^K y^{k0} - \Sigma_{h=1}^H b^{h0} - \beta = 0_M$; $\Sigma_{k=0}^K z^{k0} - \Sigma_{h=1}^H c^{h0} - \gamma = 0_N$.

We do not assume that resources pertaining to the regulated and domestic good markets are optimally allocated in period 0. The amount of "foreign" exchange that the regional economy has produced in period 0 is (we no longer demand that interregional trade be exactly balanced):

(23) $f^0 \equiv \Sigma_{k=0}^K \pi^k(w^0, y^{k0}, z^{k0}) - \Sigma_{h=1}^H e^h(u_h^0, w^0, b^{h0}, c^{h0}) - w^0 \cdot \alpha$.

The maximal amount of foreign exchange the economy could produce and still give each household the period 0 level of utility is given by the solution to the following constrained maximization problem which is analogous to (8):

(24) $f^* \equiv \max_{y^k, z^k, b^h, c^h} \{\Sigma_{k=0}^K \pi^k(w^0, y^k, z^k) - \Sigma_{h=1}^H e^h(u_h^0, w^0, b^h, c^h)$

$- w^0 \cdot \alpha : \Sigma_{k=0}^k y^k - \Sigma_{h=1}^H b^h - \beta \geq 0_M ; \Sigma_{k=0}^K z^k - \Sigma_{h=1}^H c^h - \gamma \geq 0_N ;$

$y^0 \geq 0_M ; y^k \leq 0_M, k = 1,\ldots,K ; b^h \geq 0_M, h = 1,\ldots,H\}$.

A natural measure of the waste of resources inherent in the observed period 0 equilibrium is

(25) $W_A \equiv f^* - f^0 \geq 0$.

The inequality in (25) follows from the feasibility of the observed equilibrium for the problem (24). The waste measure defined by (25) is a variant of a measure introduced by Allais [1943] [1977].

In principle different regulatory schemes could be judged by the amount of waste that they engendered. However, we shall use the Allais measure of waste for a different purpose: namely, as a bonus or performance evaluation function that can be used to evaluate the performance of the manager of a regulated industry.

Consider the following candidate for an objective function:

$$(26)\quad G^0 \equiv \Sigma_{k=0}^{K}\pi^k(w^0,y^{k1},z^{k1}) - \Sigma_{h=1}^{H}e^h(u_h^0,w^0,b^{h1},c^{h1}) - w^0\cdot\alpha$$
$$- [\Sigma_{k=0}^{K}\pi^k(w^0,y^{k0},z^{k0}) - \Sigma_{h=1}^{H}e^h(u_h^0,w^0,b^{h0},c^{h0}) - w^0\cdot\alpha] .$$

We regard the period 0 equilibrium as being fixed. The manager of the regulated sector is supposed to choose price and quantity vectors for period 1 that maximize G^0. The period 0 equilibrium is characterized by the interregional price vector w^0, the welfare vector $u^0 \equiv [u_1^0,\ldots,u_H^0]$ and the resource allocation equations (22). We do not assume that the period 0 equilibrium is efficient, so we will sometimes refer to it as the observed distorted equilibrium. From (23), we see that the terms in square brackets in (26) add up to f^0, the amount of "foreign" exchange the regional economy produced in period 0. We assume that the world price vector w^0 and the period 0 real income vector u^0 will carry over to period 1 and we ask the manager of the regulated sector to maximize (26) subject to the following restrictions:

$$(27)\quad \Sigma_{k=0}^{K}y^{k1} - \Sigma_{h=1}^{H}b^{h1} - \beta \geq 0_M ;$$

$$(28)\quad \Sigma_{k=0}^{K}z^{k1} - \Sigma_{h=1}^{H}c^{h1} - \gamma \geq 0_N ;$$

$$(29)\quad y^{01} \geq 0_M ;\ y^{k1} \leq 0_M,\ k = 1,\ldots,K ;\ b^{h1} \geq 0_M,\ h = 1,\ldots,H.$$

It is reasonable to assume that the manager of the regulated sector can control the allocation of resources in the M regulated public utility service markets. If the manager can also control the allocation of goods in the N domestic markets, then the control variables for the manager are y^{k1}, z^{k1}, $k = 0,1,\ldots,K$ and b^{h1}, c^{h1}, $h = 1,\ldots,H$ and maximizing (26) with respect to these variables, subject to the constraints (27) - (29), is equivalent to maximizing (24) and so the optimized objective function for (26) will be

$$(30)\quad G^0 = f^* - f^0 = W_A \geq 0 ,$$

which is the amount of extra foreign exchange the economy could produce, keeping each household at its period 0 level of welfare. Thus we are giving the manager of the regulated sector the task of eliminating waste in the economy for period 1. Note that we are agreeing with Schmalensee [1979; 8] that the dominant objective of a regulatory scheme should be the improvement of economic efficiency.

Is it reasonable to suppose that the manager can control the allocation of goods in domestic markets? The answer is yes if sectors 1 to K and each household behave competitively and the government eliminates all taxes and subsidies on domestic goods in the period 1 economy. After the manager chooses his net output vector z^{01} for domestic goods, the market will optimally allocate domestic resources across the competitive firms and households, where the optimality is conditional on the choices of the manager of the regulated sector. Thus in the case where there are no distortions on domestic markets in period 1, it is perfectly reasonable to ask the manager to maximize G^0 with respect to all quantity vectors subject to the restrictions (27)-(29).

However, it may not be reasonable to suppose that the vectors u^0 and w^0 will persist in period 1. Suppose that the period 1 distribution of real income ends up being given by the vector $u^1 \equiv [u_1^1,\ldots,u_H^1]$ and the vector of world prices turns out to be $w^1 \equiv [w_1^1,\ldots,w_I^1] \gg 0_I$. Then another possible objective function for the manager of the regulated sector is

$$(31)\quad G^1 \equiv \Sigma_{k=0}^{K}\pi^k(w^1,y^{k1},z^{k1}) - \Sigma_{h=1}^{H}e^h(u_h^1,w^1,b^{h1},c^{h1}) - w^1\cdot\alpha$$
$$- [\Sigma_{k=0}^{K}\pi^k(w^1,y^{k0},z^{k0}) - \Sigma_{h=1}^{H}e^h(u_h^1,w^1,b^{h0},c^{h0}) - w^1\cdot\alpha] .$$

The first 3 sets of terms on the right hand side of (31) represent the amount of foreign exchange actually earned by the economy in period 1. The 3 sets of terms in square brackets represent a hypothetical amount of foreign exchange that would be earned in the economy if the distribution of real income remained fixed at the actual period one distribution u^1, and if the economy faced the world price vector w^1, but the allocation of regulated and domestic goods remained at their period 0 levels. The manager is supposed to maximize(31) subject to the constraints (27) - (29).

Note that $G^1 = G^0$ if $w^1 = w^0$ and $u^1 = u^0$; thus under static conditions, the two objective functions coincide.

Note also that the manager would probably prefer G^0 as an objective function, since in this case, he or she does not have to forecast what world prices w^1 and the distribution of income u^1 will be in period 1. On the other hand, the government would probably prefer G^1 as a performance indicator for the manager, since in this case, the forecast risks are imposed on the manager. Thus we could view the maximization of $(1/2)(G^0 + G^1)$ as a risk sharing arrangement between the government and the manager.

Unfortunately, each of the two indicators defined above contains terms which cannot be evaluated using just observable price and quantity data for the two periods. Thus in the following two sections, we shall impose differentiability assumptions on our functions and attempt to find approximations to our theoretical indicators that can be evaluated using only market data.

5. FIRST ORDER APPROXIMATIONS

Recall G^0 defined by (30). We need to form approximations for the differences $\pi^k(w^0,y^{k1},z^{k1})-\pi^k(w^0,y^{k0},z^{k0})$ and $e^h(u_h^0,w^0,b^{h1},c^{h1})-e^h(u^0,w^0,b^{h0},c^{h0})$. We shall assume that the price and quantity data pertaining to periods 0 and 1 are observable and we further assume that the functions $\pi^k(w^t,y^{kt},z^{kt})$, $k=0,1,\ldots,K$, and $e^h(u_h^t,w^t,b^{ht},c^{ht})$, $h=1,\ldots,H$ are differentiable with respect to all arguments for $t = 0,1$. In particular, we assume that the following vectors of derivatives exist and are equal to corresponding nonnegative price vectors p^{kt} and r^{kt} that sector k faces in period t and P^{ht} and R^{ht} that household h faces in period t for public utility services and domestic goods respectively:

$$(32)\quad p^{kt} \equiv -\nabla_y\pi^k(w^t,y^{kt},z^{kt}),\quad r^{kt} \equiv -\nabla_z\pi^k(w^t,y^{kt},z^{kt});\ t=0,1;\ k=0,1,\ldots,K;$$

$$(33)\quad P^{ht} \equiv -\nabla_b e^h(u_h^t,w^t,b^{ht},c^{ht}),\quad R^{ht} \equiv -\nabla_c e^h(u_h^t,w^t,b^{ht},c^{ht});\ t=0,1;\ h=1,\ldots,H.$$

A first order Taylor series approximation to $\pi^k(w^0,y^{k1},z^{k1})$ is:

$$(34)\quad \pi^k(w^0,y^{k1},z^{k1}) \approx \pi^k(w^0,y^{k0},z^{k0}) + \nabla_y\pi^{k0}\cdot(y^{k1}-y^{k0}) + \nabla_z\pi^{k0}\cdot(z^{k1}-z^{k0})$$
$$= w^0\cdot x^{k0} - p^{k0}\cdot(y^{k1}-y^{k0}) - r^{k0}\cdot(z^{k1}-z^{k0})$$

where π^{k0} denotes $\pi^k(w^0,y^{k0},z^{k0})$ and we have used (21) and (32). We assume that p^{k0} and r^{k0} are price vectors (excluding any fixed charges) that sector k faces for regulated goods and domestic goods respectively in period 0 for $k = 1,\ldots,K$.

An alternative first order approximation is:

$$(35)\quad \pi^k(w^0,y^{k1},z^{k1}) \approx \pi^k(w^1,y^{k1},z^{k1}) + \nabla_w\pi^{k1}\cdot(w^0-w^1)$$
$$= w^1\cdot x^{k1} + x^{k1}\cdot(w^0-w^1)$$
$$= w^0\cdot x^{k1}$$

where we have used the counterpart to (21) for period 1 and Hotelling's Lemma (19) applied to period 1 data.

A first order approximation to $e^h(u_h^0,b^{h1},c^{h1})$ is:

$$(36)\quad e^h(u_h^0,w^0,b^{j1},c^{h1}) \approx e^h(u_h^0,w^0,b^{h0},c^{h0}) + \nabla_b e^{h0}\cdot(b^{h1}-b^{h0}) + \nabla_c e^{h0}\cdot(c^{h1}-c^{h0})$$
$$= w^0\cdot a^{h0} - P^{h0}\cdot(b^{h1}-b^{h0}) - R^{h0}\cdot(c^{h1}-c^{h0})$$

where $\nabla_b e^{h0}$ denotes $\nabla_b e^h(u_h^0,w^0,b^{h0},c^{h0})$ and we have used (20) and (33). Recall that P^{h0} and R^{h0} are price vectors (excluding any fixed charges) that household h faces for regulated goods and domestic goods respectively during period 0.

Substitute (36), (35) for k = 0 and (34) for k = 1,...,K into the definition of G^0, (30), and we obtain the following first order approximation to G^0:

$$(37)\quad G^0 \approx w^0\cdot(x^{01} - x^{00}) - \Sigma_{k=1}^{K} p^{k0}\cdot(y^{k1} - y^{k0}) + \Sigma_{h=1}^{H} P^{h0}\cdot(b^{h1} - b^{h0})$$
$$- \Sigma_{k=1}^{K} r^{k0}\cdot(z^{k1} - z^{k0}) + \Sigma_{h=1}^{H} R^{h0}\cdot(c^{h1} - c^{h0}) .$$

In order to "simplify" (37), we assume that there are no domestic tax distortions so that all producers and consumers face the same domestic goods price vector r^t in period t. Using also (22), we find that (37) reduces to the following first order approximation for G^0:

$$(38)\quad G^0 \approx w^0\cdot(x^0 - x^{00}) + r^0\cdot(z^{01} - z^{00}) - \Sigma_{k=1}^{K} p^{k0}\cdot(y^{k1} - y^{k0}) + \Sigma_{h=1}^{H} P^{h0}\cdot(b^{h1} - b^{h0}).$$

Let us analyze the terms on the right hand side of (38). $w^0\cdot x^{00} + r^0\cdot z^{00}$ is the regulated sector's period 0 net revenue from supplying world and domestic goods (other than the regulated public utility services) minus input cost. If the regulated sector does not produce any outputs other than the M regulated services, then $w^0\cdot x^{00} + r^0\cdot z^{00}$ is minus period 0 cost for sector 0. In general, these terms will represent minus the net cost of producing the regulated services vector y^{00} in period 0. The terms $w^0\cdot x^{01} + r^0\cdot z^{01}$ represent minus the net cost of producing the period 1 regulated services vector y^{01} but valued at period 0 prices. Thus $w^0\cdot(x^{01} - x^{00}) + r^0\cdot(z^{01} - z^{00})$ represents minus the change in cost for the regulated sector, where all inputs are valued at period 0 prices. The term $-p^{k0}\cdot(y^{k1} - y^{k0})$ represents the net increase in sector k's utilization of public utility services valued at period 0 prices (remember, $y^{kt} \le 0_M$ for k = 1,...,K by our sign conventions for the treatment of inputs and outputs). The term $P^{h0}\cdot(b^{h1} - b^{h0})$ represents the increase (decrease if negative) in household h's consumption of public utility services valued at period 0 prices (our sign conventions in this case imply $b^{ht} \ge 0_M$).

In an entirely analogous manner, we can derive the following first order approximation to the performance indicator G^1 defined by (31):

$$(39)\quad G^1 \approx w^1\cdot(x^0 - x^{00}) + r^1\cdot(z^{01} - z^{00}) - \Sigma_{k=1}^{K} p^{k1}\cdot(y^{k1} - y^{k0}) + \Sigma_{h=1}^{H} P^{h1}\cdot(b^{h1} - b^{h0})$$

Note that the right hand side of (39) is analogous to (38) except that period 1 prices and tax rates have replaced the corresponding period 0 prices and tax rates. The first two inner products on the right hand side of (39) represent the regulated firm's cost reduction going from period 0 to 1 evaluated at period 1 prices, the next K terms represent the net increase in the utilization of public utility services by the competitive firms in the economy evaluated at period 1 prices (a measure of producer benefit), and the last H terms represent the net increase in the consumption of public utility services by households evaluated at period 1 prices (a measure of consumer benefit).

Recall that if the manager of the regulated sector solves the problem of maximizing G^1 subject to the various resource constraints, then the period 1 economy will be efficient. If we replace the unobservable performance indicator G^1 by the approximation on the right hand side of (39), then the manager will be motivated to choose his outputs, inputs and prices for regulated goods in such a way so that the allocation of resources will be approximately efficient in period 1. The manager is allowed to set arbitrary prices for the public utility services, but he is obligated to supply whatever quantities are demanded at those prices.

If instead of using the right hand side of (39) as a performance indicator, we wish to use it as a bonus function for the manager of the public utilities sector, then we need only use a suitable monotonic transformation of the right hand side, so that the required payments to the manager will not be enormous.

The performance indicators defined by (58) and (59) can be evaluated in principle using only observable price, quantity and tax data for the two periods.

6. SECOND ORDER APPROXIMATIONS

In order to provide second order approximations to our theoretical performance indicators defined in section 4, we shall proceed in a somewhat indirect manner. We shall obtain expressions for an average of G^0 and G^1 that are exactly correct if the profit functions π^k and the expenditure functions e^h have certain functional forms. These functional forms will be flexible; i.e., they can approximate arbitrary twice continuously differentiable profit or expenditure functions to the second order.[8] The reader who is familiar with the index number literature will realize that our procedure is analogous to the construction of a superlative index number formula.[9] The main difference is that instead of approximating ratios of functions, we now approximate differences of functions by empirically observable formulae.

Consider the following functional form for a restricted profit function π where we combine our old vectors y and z into the M+N dimensional vector s; i.e., $s^T \equiv [y^T,z^T]$ where s^T indicates the transpose of the column vector s:

$$(40)\quad \pi(w,s) \equiv b\cdot w + (1/2)(w\cdot\theta)^{-1}w\cdot Aw + w\cdot Bs + (1/2)(w\cdot\theta)s\cdot Cs$$

where $\theta > 0_I$ and b are I dimensional vectors of parameters, A is an I by I symmetric positive semidefinite matrix, B is an I by M+N matrix of parameters and C is a symmetric M+N by M+N matrix. Note that we are assuming that the vector of parameters θ is nonnegative with at least one component positive. Using a result in Diewert and Wales [1985; 20], it can be shown that the positive semidefiniteness of A implies that the $\pi(w,s)$ defined by (40) is a convex function of w, as it must be in order to be a valid profit function.

Proposition 1: π defined by (40) is a flexible functional form for any preassigned choice of $\theta > 0_I$; i.e., it can provide a second order approximation to an arbitrary twice continuously differentiable restricted

profit function at an arbitrary point w^*,s^* where $w^* \gg 0_I$. Moreover the parameters of the matrix A can be further restricted to satisfy the following I linear restrictions without destroying the flexibility property:

$$(41)\quad Aw^* = 0_I\ .$$

The functional form defined by (40) is a very convenient one for econometric purposes, since if the restrictions (41) are satisfied, it has the flexibility property with a minimal number of parameters. Moreover, linear regression techniques may be used to estimate the unknown parameters. Finally, the convexity in prices property that a profit function must satisfy may be imposed globally without destroying the flexibility of the functional form by using a technique explained in Diewert and Wales [1987].

The following two propositions require us to evaluate profit functions at the deflated prices $\tilde{w}^t$ defined by:

$$(42)\quad \tilde{w}^0 \equiv w^0/w^0\cdot\theta\ ;\ \tilde{w}^1 \equiv w^1/w^1\cdot\theta\ ;\ \theta > 0_I\ .$$

The vector $\theta \equiv [\theta_1,\ldots,\theta_I]$ which appears in (42) can be an arbitrary nonnegative (but nonzero) vector. Once θ is chosen, we use it to define π by (40).

Proposition 2: Let the restricted profit function π be defined by (40) and define the deflated world price vectors $\tilde{w}^t$ by (42). Then the following identities hold for all vectors s^0 and s^1:

$$(43)\quad \pi(\tilde{w}^1,s^1) - \pi(\tilde{w}^1,s^0) + \pi(\tilde{w}^0,s^1) - \pi(\tilde{w}^0,s^0)$$

$$= [\nabla_s\pi(\tilde{w}^0,s^0) + \nabla_s\pi(\tilde{w}^1,s^1)]\cdot[s^1 - s^0] \quad \text{and}$$

$$(44)\quad = [\tilde{w}^0 + \tilde{w}^1]\cdot[\nabla_w\pi(\tilde{w}^1,s^1) - \nabla_w\pi(\tilde{w}^0,s^0)]\ .$$

The proof is by a lengthy series of straightforward calculations, remembering that $\tilde{w}^t \cdot \theta = 1$ for $t = 0,1$.

Consider the following functional form for a restricted expenditure function e where we now define the vector s to be our old household vectors b and c; i.e., $s^T \equiv [b^T, c^T]$:

$$(45)\quad e(u,w,s) \equiv b\cdot w + (1/2)(w\cdot\theta)^{-1} w\cdot Aw + w\cdot Bs + (1/2)(w\cdot\theta)s\cdot Cs$$

$$+ w\cdot au + (w\cdot\theta)s\cdot cu + (1/2)(w\cdot\theta)c_0 u^2$$

where $\theta > 0_I$, a and b are I dimensional vectors of parameters, c is an M+N vector of parameters, A is an I by I negative semidefinite symmetric matrix, B is an I by M+N matrix of parameters, C is an M+N by M+N positive semidefinite symmetric matrix and c_0 is a scalar parameter. Recall[10] that restricted expenditure functions, e(u,w,s), have to be concave in their price arguments w and convex in their quantity arguments s, under our assumption that the household's utility function is quasiconcave. The curvature conditions that we have imposed on the matrices A and C will ensure that e defined by (45) satisfies the appropriate theoretical curvature restrictions globally.

The following proposition is a counterpart to Proposition 1.

<u>Proposition 3</u>: e defined by (45) is a flexible functional form for any preassigned choice of $\theta > 0_I$. Moreover, the parameters of the matrix A can be restricted to satisfy the I restrictions (41) without destroying the flexibility result.

The following proposition is a counterpart to (43) in Proposition 2 and may be proven in the same manner.

<u>Proposition 4</u>: Pick $\theta > 0_I$, define the restricted expenditure function e by (45) and define the deflated interregional price vectors $\tilde{w}^t$ by (42). Then the following identity holds for all s^0 and s^1:

$$(46)\quad e(u^1,\tilde{w}^1,s^1) - e(u^1,\tilde{w}^1,s^0) + e(u^0,\tilde{w}^0,s^1) - e(u^0,\tilde{w}^0,s^0)$$

$$= [\nabla_s e(u^1,\tilde{w}^1,s^1) + \nabla_s e(u^0,\tilde{w}^0,s^0)]\cdot[s^1 - s^0] .$$

Recall definitions (32) and (33). Define the corresponding deflated price and distortion vectors by:

$$(47)\quad \tilde{p}^{kt} \equiv p^{kt}/w^t\cdot\theta\ ,\ \tilde{r}^{kt} \equiv r^{kt}/w^t\cdot\theta\ ,\ \tilde{r}^t \equiv r^{0t}/w^t\cdot\theta;\ t=0,1;\ k=0,1,\ldots,K;$$

$$(48)\quad \tilde{P}^{ht} \equiv P^{ht}/w^t\cdot\theta\ ,\ \tilde{R}^{ht} \equiv R^{ht}/w^t\cdot\theta\ ; \qquad t = 0,1,\ h = 1,\ldots,H.$$

<u>Proposition 5</u>: Let $\theta > 0_I$ be fixed. Let each restricted profit function π^k be in the class of functional forms defined by (40) and let each restricted expenditure function e^h be in the class of functional forms defined by (45). Assume that the economy's domestic resource constraints (3) hold with equality in each period. Define the theoretical performance indicators $\tilde{G}^0$ by (26) (except use the deflated price vector $\tilde{w}^0$ in place of w^0) and $\tilde{G}^1$ by (31) (except use $\tilde{w}^1$ in place of w^1). Then the following identity holds <u>exactly</u>:

$$(49)\quad \tilde{G}^0 + \tilde{G}^1 = (\tilde{w}^0+\tilde{w}^1)\cdot(x^{01}-x^{00}) + (\tilde{r}^0+\tilde{r}^1)\cdot(z^{01}-z^{00}) - \Sigma_{k=1}^K (\tilde{p}^{k0}+\tilde{p}^{k1})\cdot(y^{k1}-y^{k0})$$

$$+ \Sigma_{h=1}^H (\tilde{P}^{h0}+\tilde{P}^{h1})\cdot(b^{h1}-b^{h0}) \equiv 2G^* .$$

The right hand side of (49) can in principle be calculated knowing price and quantity data for the economy for the two periods. Note that G^* is the average of the first order approximation to G^0, (38), and the first order approximation to G^1, (39), except that the prices in (38) and (39) are replaced by the corresponding deflated prices. Thus there is no need to belabour the interpretation of G^*: the same old cost reduction, producer benefit and consumer benefit terms show up, expressed at average deflated prices rather than at individual period prices. If $u^0 = u^1$ (so there is no change in the distribution of real income) and w^1

is proportional to w^0, then it is easy to show that $\tilde{G}^0 = \tilde{G}^1$. Thus under these hypotheses, G* is exact for the common theoretical performance indicator $\tilde{G} = \tilde{G}^0 = \tilde{G}^1$.

In general, note that G* is defined in terms of observable price and quantity data. Moreover, G* is exactly equal to the theoretical performance indicator $(1/2)\ (\tilde{G}^0 + \tilde{G}^1)$ if the industry profit functions have normalized quadratic forms (39) and the consumer restricted expenditure functions have normalized quadratic forms (45). Since these quadratic functional forms can provide second order approximations to arbitrary profit and expenditure functions? G* has a second order approximation property and thus may be called a superlative performance indicator in analogy with the index number literature.

In order to help the reader interpret the rather formidable formula (49), we present the geometry of a very special case of our analysis in the following section.

7. THE GEOMETRY OF A SPECIAL CASE

Let I be arbitrary, M = 1, N = 0, K = 0 and H = 1. Hence there is an arbitrary number of interregionally traded goods, one public utility good, no domestic goods, no competitive firms and only one household. Furthermore, we assume that the welfare of the single household remains unchanged and that world prices remain unchanged, so that $u \equiv u^0 = u^1$ and $w \equiv w^0 = w^1$. Under these conditions, our theoretical performance indicators defined in section 4 collapse to the same thing, G say. There are only two functions in the regional economy, $\pi^0(w,y^0)$, the manager's restricted profit function, and $e^1(u,w,b)$, the single household's restricted expenditure function. The adding up constraint for public utility services implies that $y^0 = b$, where b is household one's demand for public utility services. Since b is the only managerial choice variable, let us simplify the notation and define $\pi(b) \equiv \pi^0(w,b)$ and $e(b) \equiv e^1(u,w,b)$. In period 0, the allocation of the single public utility good is b^0 and in period one, the allocation is b^1. The manager's problem is to maximize

(50) $G \equiv \pi(b^1) - e(b^1) - [\pi(b^0) - e(b^0)]$

with respect to $b^1 \geq 0$. If the observed $b^1 > 0$ solves this maximization problem and the functions π and e are differentiable at the solution, then the following first order condition will be satisfied:

(51) $-\pi'(b^1) = -e'(b^1) \equiv P^1$.

The left hand side of (51) represents the regulated firm's marginal cost of producing an extra unit of the public utility service while the right hand side represents the household's willingness to pay for an extra unit and represents a point on the consumer's Hicksian [1946; 331] (real income constant) demand curve. P^1 denotes the price the household faces in period 1 while $P^0 \equiv -e'(b^0)$ denotes the price the household faced in period 0.

Since we are assuming that $w^1 = w^0$, it is harmless to assume that $w^0 \cdot \theta = w^1 \cdot \theta = 1$. Under these conditions, the approximation to G defined by one half of (49) is:

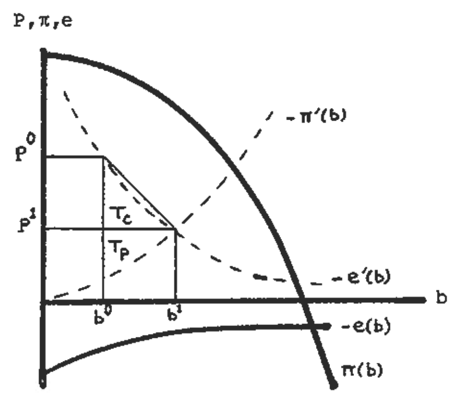
(52) $(1/2)(w^0+w^1)\cdot(x^{01}-x^{00}) + (1/2)(P^0+P^1)\cdot(b^1-b^0)$

$= \pi(b^1) - \pi(b^0) + (1/2)(P^0+P^1)\cdot(b^1-b^0)$

where the equality in (52) follows from $w^0 \cdot x^{00} = \pi(b^0) = w^1 \cdot x^{00}$ and $w^1 \cdot x^{01} = \pi(b^1) = w^0 \cdot x^{01}$, using also $w^0 = w^1$.

The algebra above is illustrated in Figure 1.

Figure 1.



The top solid curve in Figure 1 is $\pi(b)$ while the bottom solid curve is $-e(b)$. The verticle sum of these two curves is maximized at b^1 where the marginal cost curve, $-\pi'(b)$, intersects the Hicksian demand curve, $-e'(b)$.

Note that the area under $-\pi'(b)$ from b^0 to b^1 equals $-[\pi(b^1) - \pi(b^0)]$ while the area under $-e'(b)$ from b^0 to b^1 equals $-[e(b^1) - e(b^0)]$. Using these relations, it can be seen that the theoretical performance indicator (or measure of efficiency gain) G is equal to the triangular regions T_c plus T_p where the upper boundary of T_c is the dashed line. On the other hand, it can be seen that the approximate measure of gain defined by (52) is equal to T_c^* plus T_p but now the upper boundary of the triangular region T_c^* is the straight line instead of the dashed line. Of course, if the expenditure function e is in the class of functional forms defined by (45), then the triangular regions would coincide. The first order approximation to G^0 (see (38)) can be shown to be $T_p + (P^0 - P^1)(b^1 - b^0)$ while the first order approximation to G^1 (see (39)) is T_p.

We leave to the reader the task of adapting the geometry to situations where the manager chooses b^1 to the right of the intersection of the marginal cost and Hicksian demand curves and to situations where the technology set is not convex and thus the marginal cost curve can be

downward sloping. The main impression we want to leave with the reader is that the quadratic approximations defined in section 6 are very much better than the linear approximations defined in section 5.

8. ALTERNATIVE PERFORMANCE INDICATORS

Early contributors to the incentive scheme and performance indicator approach to the control of monopoly include Scott [1952], Domar [1974], Vogelsang and Finsinger [1979] and Tam [1981]. In our notation, the three most interesting performance indicators were defined by:

$$(53) \quad I_S \equiv w^1\cdot(x^1 - x^0) + p^1\cdot(y^1 - y^0) + r^1\cdot(z^1 - z^0) \;;$$

$$(54) \quad I_{FV} \equiv p^1\cdot(y^1 - y^0) - [C(y^1,w^1,r^1) - C(y^0,w^0,r^0)] \;;$$

$$(55) \quad I_T \equiv (1/2)(p^1 + p^0)\cdot(y^1 - y^0) - [C(y^1,w^1,r^1) - C(y^0,w^0,r^0)]$$

where x^t,y^t,z^t denote the net output vectors for internationally traded goods, public utility services and domestic goods respectively for the monopoly sector in period t; w^t,p^t,r^t denote the corresponding period t price vectors that the public utility sector faces for the three classes of goods, and $C(y^t,w^t,r^t)$ denotes the net cost of producing the vector of public utility services y^t in period t. I_S defined by (53) is our mathematical interpretation of Scott's [1978; 155] verbal description of his performance indicator. Note that our performance indicator G^1 defined by (39) reduced to it if H + K = 1 or if all users of the public utility services are charged the same price for the same service. I_{FV} defined by (54) is a performance indicator due to Finsinger and Vogelsang [1981; 400] [1982; 285] [1985; 165]. If $w^0=w^1$ and $r^0=r^1$ and $C(y^t,w^t,r^t)=[w^t\cdot x^t + r^t\cdot z^t]$ for t = 0,1 so that the Arrow-Hurwicz problem does not apply, then it can be shown that $I_S = I_{FV}$. The performance indicator I_T defined by (55) is due to Tam [1985; 284]. Sappington and Sibley [1985] have defined a closely related indicator. Note that I_T is related to our performance indicator G* defined in (49); in fact the indicators are identical if: (i) we use normalized prices in (55),

(ii) H + K = 1 or all users of the same public utility service are charged the same price and (iii) $w^0 = w^1$ and $r^0 = r^1$.

The above performance indicators may be subjected to a number of criticisms. (i) Scott's indicator has no formal economic or mathematical justification and the other two indicators were justified using a suspect Marshallian [1920; 487] consumer surplus concept. (ii) Scott's indicator has only a first order approximation property while the other indicators have formal approximation properties only under additional assumptions. In contrast, our performance indicator defined by the right hand side of (49) had a second order approximation property. (iii) I_{FV} and I_T are not inflation proof; i.e., they are not invariant to scale changes in the level of prices in any period. (iv) The last two indicators are based on a partial equilibrium analysis and moreover, they are subject to the Arrow-Hurwicz objection if the monopolist's production possibilities set has certain types of nonconvexities.

In spite of the above criticisms, it is clear that our analysis owes a great deal to this pioneering incentive scheme literature.[11]

9. CONCLUSION

Much of the literature on evaluating the efficiency of a regulated public utility monopoly and the literature which proposed optimal regulatory schemes made use of the consumer surplus concept. It is well known that the consumer surplus concept suffers from a number of theoretical deficiencies, particularly when it is applied in a many consumer context. Thus there appears to be a significant theoretical deficiency in the literature.

In our paper, we proposed a new method for regulating a public utility monopoly, one that is solidly based on microeconomic theory. Our proposed method involved giving the public utility manager the "right" objective function to maximize. This social objective function is a very broad one; it is the amount of extra money or foreign exchange that the regional economy that is served by the public utility could earn while keeping each

household at a specified level of welfare. This theoretical objective function is a variant of one originally proposed by Allais [1943].

A difficulty with our proposed regulatory scheme is that the theoretically desirable social objective function is unobservable. However, we provide first and second order approximations to the theoretical objective function which depend only on observable price and quantity data. Thus we end up with a regulatory scheme that is "practical" in the sense that no elasticity information is required in order to implement it and it should achieve an efficient allocation of resources (to the second order). Thus the deadweight loss generated by currently used regulatory methods should be reduced by our proposed method.

In the course of our derivations, we defined some new classes of flexible functional forms for restricted profit and expenditure functions (see Propositions 1 and 3 above) that should prove to be useful in applications.

Finally, it seems appropriate to conclude by quoting Zajac [1978; 47] on the significance of deadweight loss due to inadequate regulatory schemes:

> "A deadweight loss to an economist is like an unexploited energy source to an engineer. In both cases lack of tools or the presence of constraints may thwart the realization of potential benefits. But the larger the potential benefits, the more justified is an effort to obtain more flexible or additional policy instruments to convert potential benefits into realized benefits. Of course, the obstacles may be so great that even herculean efforts are insufficient to bring about the conversion -- a situation of great frustration to both the economist and the engineer."

FOOTNOTES

1. See Walras [1980; 83], Pigou [1920; 278], Hotelling [1938; 256], Lerner [1944; 182], Meade [1945] and Fleming [1945].
2. If the region were indebted to the rest of the world, then the 0 in (1) could be replaced by a parameter b, the amount of world currency required to service the debt. Notation: $w \cdot x \equiv \Sigma_{i=1}^{I} w_i x_i$.
3. This means that each producer and consumer either produces or utilizes an interregionally traded good at the optimal solution.
4. See Diewert [1986; 25].
5. See Diewert [1986; 175].

6. Notable exceptions include Guesnerie [1975], Brown and Heal [1979] [1980] and Drèze [1980].
7. See Hotelling [1932; 594], Gorman [1968] and Diewert [1986; 142].
8. See Diewert [1974; 113] for the concept of a flexible functional form.
9. See Diewert [1976] and Denny and Fuss [1983a] [1983b].
10. See Diewert [1986; 170-176].
11. The material in this chapter is treated at greater length in Diewert [1985].

References

Allais, M. (1943), A la recherch d'une discipline économique, Tome I, Paris: Imprimerie Nationale.

Allais, M. (1977), "Theories of General Economic Equilibrium and Maximum Efficiency," pp. 129-201 in Equilibrium and Disequilibrium in Economic Theory, E. Schwoediauer (ed.), Dordrecht: D. Reidel.

Arnott, R. and R. Harris (1977), "Increasing Returns and the Inefficiency of Cost Minimization," Discussion Paper #278, Department of Economics, Queen's University, Kingston, Canada.

Arrow, K.J. (1964), "Control in Large Organizations," Management Science 10, 397-408.

Arrow, K.J. and L. Hurwicz (eds.), (1977), Studies in Resource Allocation Processes, New York: Cambridge University Press.

Brown, D.J. and G. Heal (1979), "Equity, Efficiency and Increasing Returns," The Review of Economic Studies 46, 571-585.

Brown, D.J. and G. Heal (1980), "Two-Part Tariffs, Marginal Cost Pricing and Increasing Returns in a General Equilibrium Model," Journal of Public Economics 13, 25-49.

Coase, R.H. (1945), "Price and Output Policy of State Enterprise: A Comment," Economic Journal 55, 112-113.

Coase, R.H. (1946), "The Marginal Cost Controversy," Economica 13, 169-182.

Denny, M. and M. Fuss (1983a), "The Use of Discrete Variables in Superlative Index Number Comparisons," International Economic Review 24, 419-421.

Denny, M. and M. Fuss (1983b), "A General Approach to Intertemporal and Inter-spatial Productivity Comparisons," Journal of Econometrics 23, 315-330.

Diewert, W.E. (1974), "Applications of Duality Theory," pp. 106-171 in Frontiers of Quantitative Economics, Vol. II, M.D. Intriligator and D.A. Kenrick (eds.), Amsterdam: North-Holland.

Diewert, W.E. (1976), "Exact and Superlative Index Numbers," Journal of Econometrics 4, 115-145.

Diewert, W.E. (1985), "Efficiency Improving Incentive Schemes for Regulated Industries", Discussion Paper 85-34, Department of Economics, University of British Columbia, Vancouver, Canada, November.

Diewert, W.E. (1986), The Measurement of the Economic Benefits of Infrastructure Services, Berlin: Springer-Verlag.

Diewert, W.E. and T.J. Wales (1987), "Flexible Functional Forms and Global Curvature Conditions," Econometrica 55, 43-68.

Domar, E.D. (1974), "On the Optimal Compensation of a Socialist Manager," The Quarterly Journal of Economics 88, 1-18.

Drèze, J.H. (1980), "Public Goods with Exclusion," Journal of Public Economics 13, 5-24.

Dupuit, J. (1944), "De la mesure de l'utilité des travaux publics," Annales des Ponts et Chaussées 8, translated as "On the Measurement of the Utility of Public Works," International Economic Papers 2 (1952), 83-110.

Finsinger, J. and I. Vogelsang (1981), "Alternative Institutional Frameworks for Price Incentive Mechanisms," Kyklos 34, 388-404.

Finsinger, J. and I. Vogelsang (1982), "Performance Indices for Public Enterprises," pp. 281-296 in Public Enterprise in Less-Developed Countries, Leroy P. Jones (ed.), New York: Cambridge University Press.

Finsinger, J. and I. Vogelsang (1985), "Strategic Management Behavior Under Reward Structures in a Planned Economy," Quarterly Journal of Economics 100, 263-269.

Fleming, J.M. (1944), "Price and Output Policy of State Enterprise: Comment," Economic Journal 54, 328-337.

Gorman, W.M. (1968), "Measuring the Quantities of Fixed Factors," pp. 141-172 in Value, Capital and Growth: Papers in Honour of Sir John Hicks, J.N. Wolfe (ed.), Chicago: Aldine Publishing Co.

Gravelle, H.S.E. (1985), "Reward Structures in a Planned Economy: Some Difficulties," Quarterly Journal of Economics 100, 271-278.

Guesnerie, R. (1975), "Pareto Optimality in Non-Convex Economies," Econometrica 43, 1-29.

Hicks, J.R. (1946), Value and Capital, Oxford: Clarendon Press.

Hotelling, H. (1932), "Edgeworth's Taxation Paradox and the Nature of Demand and Supply Functions," Journal of Political Economy 40, 577-616.

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

Kuhn, H.W. and A.W. Tucker (1951), "Nonlinear Programming," pp. 481-492 in The Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, J. Neyman (ed.), Berkeley: University of California Press.

Lerner, A.P. (1944), The Economics of Control, New York: Macmillan.

Marshall, A. (1920), Principles of Economics, London: Macmillan.

Meade, J.E. (1944), "Price and Output Policy of State Enterprise," Economic Journal 54, 321-328.

Pigou, A.C. (1920), The Economics of Welfare, London: Macmillan.

Sappington, D.E.M. and D.S. Sibley (1985), "Regulatory Incentive Schemes Using Historic Cost Data," unpublished paper.

Schmalensee, R. (1979), The Control of Natural Monopolies, Toronto: D.C. Heath.

Scott, M.F.G. (1952), "Criteria of Efficiency in Nationalized Industries," B. Litt. Thesis, Oxford.

Scott, M.F.G. (1978), "A Profits Test for Public Monopolies," pp. 152-168 in Policy and Politics, D. Butler and A.H. Halsey (eds.), London: Macmillan.

Shephard, R.W. (1953), Cost and Production Functions, Princeton, N.J.: Princeton University Press.

Tam, M.S. (1981), "Reward Structures in a Planned Economy: The Problem of Incentives and Efficient Allocation of Resources," Quarterly Journal of Economics 96, 111-128.

Tam, M.S. (1985), "Reward Structures in a Planned Economy: Some Further Thoughts," Quarterly Journal of Economics 100, 279-289.

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

Walras, L. (1980), "The State and the Railways," Journal of Public Economics 13, 81-100.

Wilson, T. (1945), "Price and Outlay Policy of State Enterprise," Economic Journal 55, 454-461.

Zajac, E.E. (1978), Fairness or Efficiency, Cambridge, Massachusetts: Ballinger Publishing.