COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut

COWLES FOUNDATION DISCUSSION PAPER NO. 146

Welfare Economics and the Theory of Second Best

Otto A. Davis and Andrew B. Whinston

September 20, 1962

Welfare Economics and the Theory of Second Best*

by

Otto A. Davis† and Andrew B. Whinston††

July 2, 1962

†Graduate School of Industrial Administration
Carnegie Institute of Technology
Pittsburgh, Pennsylvania

††Cowles Foundation
Yale University
New Haven, Connecticut

*To be presented as a Cowles Foundation discussion paper.

Welfare Economics and the Theory of Second Best*

by

Otto A. Davis and Andrew B. Whinston

1. Introduction:

The belief seems to be prevalent that the article by R. G. Lipsey and K. Lancaster, "The General Theory of Second Best,"[1] has cast doubt

[1] Review of Economic Studies, XXIV, (December, 1956), pp. 11-32.

upon the usefulness of the propositions of welfare economics in actual policy decisions. The main proposition of the theory of second best can be stated briefly. In the words of Lipsey and Lancaster,

> "The general theorem for the second best optimum states that if there is introduced into a general equilibrium system a constraint which prevents the attainment of one of the Paretian conditions, the other Paretian conditions, although still attainable, are, in general, no longer desirable. In other words, given that one of the Paretian optimum conditions cannot be fulfilled, then an optimum situation can be achieved only by departing from all the other Paretian conditions."[2]

*This research was jointly sponsored by the Cowles Commission for Research in Economics under Task NR-047-006 with the Office of Naval Research and by the Graduate School of Industrial Administration, Carnegie Institute of Technology. The authors are indebted to numerous persons for helpful comments and criticisms. Professor Francesco Forte, University of Torino, proposed this topic to us and gave assistance on various points. Professors T. C. Koopmans, Cowles Foundation, and Kelvin Lancaster, Johns Hopkins University, made numerous suggestions which resulted in alterations in the paper. Finally, W. Brainard and T. N. Srinivasan made very helpful constructive criticisms. Only the authors are responsible for any possible errors.

[2] Ibid., p. 11.

Although Lipsey and Lancaster qualify their results by the inclusion of the term "in general" in the statement of their theorem, they give no hint as to when the Pareto conditions should or should not be satisfied.[1] Hence, even if their results had not been interpreted as

[1] Of course, not everyone has qualified the assertions of the theorem of second best. For example, Paul A. Samuelson, writing before the publication of the Lipsey-Lancaster paper, comments, "First, what is the best procedure if for some reason a number of the optimum conditions are not realized? What shall we do about the remaining ones which are in our power? Shall we argue that 'two wrongs do not make a right' and attempt to satisfy those we can? Or is it possible that failure of a number of the conditions necessitates modifying the rest? Clearly the latter alternative is the correct one. A given divergence in a subset of the optimum conditions necessitates alterations in the remaining ones. Thus, in a world where almost all industries are producing at marginal social cost less than price (either because of monopoly or external economies) it would not be desirable for the rest to produce up to the point where marginal cost equals price." Foundations of Economic Analysis, Cambridge: Harvard University Press, 1947, pp. 252-3.

meaning that all Pareto conditions should be discarded when one or more of them is violated, the theorem would still mean that the propositions of welfare economics do not provide guides by which policy proposals can be judged since it is generally held as apparent that all Pareto conditions are not and can never be expected to be satisfied. Each policy proposal has to be considered on its own merits and in the context of the entire system. "Piecemeal policy recommendations" are no good. As long as there exists at least one given and unchangeable violation of one Pareto condition, then, as their results are stated, there is no a priori or practical way to determine whether the remaining conditions should be satisfied. Again in the words of Lipsey and Lancaster.

> "Specifically, it is not true that a situation in which more, but not all, of the optimum conditions are fulfilled is necessarily, or is even likely to be, superior to a situation in which fewer are fulfilled. It follows, therefore, that in a situation in which there exist many constraints which prevent the fulfillment of the Paretian optimum conditions, the removal of any one constraint may affect welfare or efficiency either by raising it, by lowering it, or by leaving it unchanged."[1]

---

[1] *Ibid.*, p. 12.

We shall attempt to demonstrate the conditions under which a violation of one Pareto condition means that it is socially desirable that certain other of the Paretian conditions be violated. In so doing, we shall dispel what seems to be the commonly held interpretation of the Lipsey-Lancaster theorem -- that a violation of one Pareto condition means that none of the remaining Pareto conditions are desirable. Our approach will be somewhat similar to that used by Lipsey and Lancaster in that we too use the method of maximization subject to constraints. However, we have fundamental objections to parts of the Lipsey-Lancaster formulation of the second-best problem and, as we shall demonstrate later, we hold that their formulation introduces extraneous considerations, does not provide adequate "guidelines" for policy applications, and leads, at the very best, to misinterpretations of the meaning of their theorem.

2. A Behavioral Constraint and the Second Best:

Before discussing the theorem of second best, let us consider a simple example of a second best problem. We formulate this example in the "traditional manner" which uses the Lagrangean maximization method and, hence, limits the analysis since "corner solutions" are not admitted; but this formulation is adequate for our purpose here and it does provide a framework for our later discussion and critique of the Lipsey-Lancaster theorem of second best.

Suppose there exists some differentiable function of n variables $(x_1, \ldots, x_n)$.

$$(1.1) \qquad F(x_1, \ldots, x_n)$$

This function is to be maximized subject to the following constraint on the variables.

$$(1.2) \qquad G(x_1, \ldots, x_n) = 0$$

This problem represents a formalization of the typical choice situation in economic analysis. From this problem comes the familiar first-order conditions for a Pareto optimum.

$$(1.3) \qquad \frac{\partial F}{\partial x_i} - \lambda \frac{\partial G}{\partial x_i} = 0, \qquad i = 1, \ldots, n$$

By eliminating the Lagrangean multiplier λ these conditions can be reduced to the following form:

$$(1.4) \qquad \frac{\frac{\partial F}{\partial x_i}}{\frac{\partial F}{\partial x_n}} - \frac{\frac{\partial G}{\partial x_i}}{\frac{\partial G}{\partial x_n}} = 0, \qquad i = 1, \ldots, n-1$$

These are, of course, the Pareto conditions as they are usually stated. Identifying (1.1) as a utility function and (1.2) as a transformation function, we have immediately from (1.4) the usual verbal statement of the Pareto conditions. The marginal rates of substitution in consumption must equal the marginal rates of substitution in production. Thus far our presentation follows the Lipsey-Lancaster development.

We now proceed to construct a second best problem. Suppose that the producers of the good identified by the subscript "1" did not maximize profit but, instead, set output equal to some (non-optimal) constant without regard to price parameters.[1] Then, granted this assumption, we have a

---

[1] This example may seem somewhat artificial, but it is intended to serve only as a counter-example to the popular interpretation of the general theorem of second best, and it should be understood that this discussion is preliminary to our more careful analysis of second best problems in later sections.

---

behavioral constraint[2] which takes the form

---

[2] We use the term "behavioral constraint" and "behavioral rule" to refer to a constraint which reflects the decision-making model of the individual unit or units under consideration.

(1.5) $$x_1 = K$$

where K is some positive constant. The introduction of constraint (1.5) alters the problem. We now maximize (1.1) subject to both (1.2) and (1.5). Using Lagrangean methods, we obtain the following expressions:

$$(1.6) \qquad \frac{\partial F}{\partial x_1} - \lambda \frac{\partial G}{\partial x_1} - \mu = 0$$

$$(1.7) \qquad \frac{\partial F}{\partial x_i} - \lambda \frac{\partial G}{\partial x_i} = 0, \qquad i = 2, \ldots, n.$$

Of course, in general the solution to the altered problem will be different from the solution to the original problem. In addition, the λ appearing in (1.6) and (1.7) will assume a different value from the λ in (1.3). However, these points need not concern us here. From the point of view of the theory of the second best the important fact is that the other multiplier μ appears only in (1.6) since constraint (1.5) is effective only for expressions involving x_1.

In order to get the conditions for a maximum into their usual form, we eliminate the Lagrangean multiplier λ and obtain the following expressions:

$$(1.8) \qquad \frac{\frac{\partial F}{\partial x_1} - \mu}{\frac{\partial F}{\partial x_n}} - \frac{\frac{\partial G}{\partial x_1}}{\frac{\partial G}{\partial x_n}} =$$

$$(1.9) \qquad \frac{\frac{\partial F}{\partial x_i}}{\frac{\partial F}{\partial x_n}} - \frac{\frac{\partial G}{\partial x_i}}{\frac{\partial G}{\partial x_n}} = 0, \qquad i = 2, \ldots, n-1.$$

Let us now interpret these conditions. Note that, if the behavioral constraint

(1.5) is effective, then $\mu \neq 0$ and condition (1.8) represents a violation of the Pareto conditions (1.4). This condition is, of course, a direct result of the behavioral constraint (1.5). By the assumptions of the theory of the second best, this constraint is taken as given and public policy cannot concern itself with altering the behavior of the producers of the 1st good. However, and this is the important point, note that conditions (1.9) are of the same form as the Pareto conditions (1.4). This result means that even though the conventional optimum conditions are violated in one sector by the behavioral constraint (1.5), observance of the Pareto conditions is still desirable from the welfare point of view in all other sectors.

The above conclusion is opposite to the naive interpretation of the Lipsey-Lancaster theorem of second best, and the problem is an effective counter-example to that interpretation. Furthermore, the ease with which this counter-example was constructed serves notice that the qualifying term "in general" is extremely important. Evidently, it is desirable to determine the conditions under which the violation of one Pareto condition leads to second best conditions that are different in form from the original optimum conditions, and to be able to identify those second best conditions which do differ from the Pareto conditions. The Lipsey-Lancaster theorem gives no clue to the answer to either of these problems, and we propose to approach these questions by considering several cases where second-best problems are present in order to determine in each instance which, if any, of the second best conditions differ in form from the Pareto optimum conditions.

3. Special Assumptions and a General Theorem:

Before proceeding with our own development of second best problems, however, it is necessary that we point out the difficulties inherent in the Lipsey-Lancaster approach so that these may be avoided in our own analysis. Our objections to the Lipsey-Lancaster approach involve three interrelated points. First, we argue that their statement of the violation of the Pareto condition is inappropriate for the problem at hand. Second, the presence of a "numéraire" creates special difficulties.[1] Third, a functional form was not specified in their proof

[1] Both of these points were made by M. McManus, "Comments on the General Theory of Second Best," Review of Economic Studies, XXVI (June, 1959), pp. 209-224.

and yet the result depends upon a special functional form. We shall illustrate these difficulties by using one part of the Lipsey-Lancaster development.

In the proof of their "general theorem" Lipsey and Lancaster introduced a constraint which violates a Pareto condition. They state, "Such a constraint *will be of the form:*"[2]

[2] *Op. cit.*, p. 26. (Italics added).

$$(2.1) \qquad \frac{\frac{\partial F}{\partial x_1}}{\frac{\partial F}{\partial x_n}} = K \frac{\frac{\partial G}{\partial x_1}}{\frac{\partial G}{\partial x_n}} \quad ; \quad K \neq 1 \quad .$$

The approach taken in this paper is to represent the deviant behavior in terms of behavioral relations. For example, in the case of a monopolist we use as a supplementary constraint that output (and, hence, price) is chosen so as to satisfy the relationship "marginal revenue equals marginal cost." In general, the Lipsey-Lancaster formulation of representing the deviant behavior by (2.1) will not correctly reflect such a situation. Equation (2.1) expresses a relationship between two "sectors" of the economic system, the 1^{st} and the n^{th}. Certainly, there is little reason to suspect that this relationship is capable of characterizing monopolistic behavior or, for that matter, any other behavior which might lead to second best problems. The decision-making rule of the deviant is not readily apparent from (2.1).

Furthermore, if (2.1) is to be used to characterize any Pareto violation, then a value of K would have to be specified which, since a behavioral relationship is not used, gives solution values for the x's which are the same as would be obtained from a model which used a behavioral relationship as a constraint. How is this K to be determined? It would seem that one must (i) either know a priori the solution to the problem so that the K can be specified before the problem is actually solved or (ii) try to determine the K by studying the original situation. However, this latter method is suspect since the second best solution will, in general, differ from the original solution and should not give the same value of K. Furthermore, using a constant K value involves the use of an overly restrictive condition.[1]

[1]This point was made by McManus, op. cit., p. 211. Lipsey and Lancaster seem to agree by pointing out while discussing the Lerner "degree of monopoly" argument that, "We cannot, therefore, simply define a monopolistic policy by the relationship between accounting prices and costs, but must always consider the relationship between these relative to their relationship elsewhere. The difficulty is, which relationship to use? Our original choice was the relationship between price and marginal cost in some industry chosen at random, but, as McManus correctly points out, the status of this industry is higher than that of mere numéraire .... Perhaps the answer in the general case is to assume the monopolist's relationship between price and marginal cost to all other industries, the functional relationship being ultimately determined by demand conditions." "McManus on Second Best," Review of Economic Studies, XXVI, (June, 1959) p. 226.

Waiving for the moment the above difficulty concerning the manner in which the Pareto violation is introduced, let us illustrate our other two objections. Making what we shall find in our later analysis to be a most crucial assumption, let us assume that both (1.1) and (1.2) are separable.[2] Then, following Lipsey and Lancaster, (1.1) is to be maximized

[2]A function $f(x_1, \ldots, x_n)$ is separable if and only if

$$f(x_1, \ldots, x_n) = f_1(x_1) + f_2(x_2) + \ldots + f_n(x_n).$$

In other words, the function $f(x_1, \ldots, x_n)$ must be composed of the sum of n functions each of which involve only one of the n variables in its argument.

subject to both (1.2) and (2.1). Taking partial derivatives, we obtain as the conditions for a maximum

$$(2.2)\qquad \frac{\partial F}{\partial x_i} - \lambda \frac{\partial G}{\partial x_i} - \mu \left[\frac{\frac{\partial F}{\partial x_n} \frac{\partial^2 F}{\partial x_1 \partial x_i} - \frac{\partial F}{\partial x_1} \frac{\partial^2 F}{\partial x_n \partial x_i}}{\left(\frac{\partial F}{\partial x_n}\right)^2} - K \frac{\frac{\partial G}{\partial x_n} \frac{\partial^2 G}{\partial x_1 \partial x_i} - \frac{\partial G}{\partial x_1} \frac{\partial^2 G}{\partial x_n \partial x_i}}{\left(\frac{\partial G}{\partial x_n}\right)^2} \right] = 0, \quad i = 1, \ldots, n.$$

This is expression (7.4) in the Lipsey-Lancaster proof of the general theorem of second best. But now note that, under the crucial assumption of separability,

$$(2.3)\qquad \frac{\partial^2 F}{\partial x_i \partial x_j} = \frac{\partial^2 G}{\partial x_i \partial x_j} = 0, \qquad i \neq j$$

so that we have immediately

$$(2.4)\qquad \frac{\partial F}{\partial x_i} - \lambda \frac{\partial G}{\partial x_i} = 0, \qquad i = 2, \ldots, n-1.$$

Since the form of (2.4) is identical to that of (1.3), it appears that the Pareto condition should be satisfied for the above commodities and that for a second best solution a violation is indicated only for the 1st and nth commodities, the latter being the so-called "numéraire" and the former being the commodity for which the original violation was specified. Hence, it seems most important that one carefully specify whether the functional forms are separable or non-separable.

However, Lipsey and Lancaster desire to express the Pareto conditions in their usual ratio form or, more specifically, in terms of the so-called numéraire.[1] Eliminating the multiplier λ one may obtain

[1]We follow the Lipsey-Lancaster terminology here, and it should be noted that this diverges from the usual usage of the term "numéraire." Indeed, we shall see that this is precisely the point; the numéraire is more important than it should be.

$$(2.5)\qquad \frac{\dfrac{\partial F}{\partial x_i}}{\dfrac{\partial F}{\partial x_n} - \mu\left[- \dfrac{\dfrac{\partial F}{\partial x_1}\dfrac{\partial^2 F}{\partial x_n^2}}{\left(\dfrac{\partial F}{\partial x_n}\right)^2} + K\,\dfrac{\dfrac{\partial G}{\partial x_1}\dfrac{\partial^2 G}{\partial x_n^2}}{\left(\dfrac{\partial G}{\partial x_n}\right)^2}\right]} - \frac{\dfrac{\partial G}{\partial x_i}}{\dfrac{\partial G}{\partial x_n}} = 0\,, \qquad i = 2, \ldots, n-1\,.$$

and from this expression it would appear that the Pareto conditions should not be satisfied. The reason for this rather curious result seems to be that the "numéraire" is not neutral.[2] It appears that introducing a con-

[2] M. McManus observed this point but did not attempt to explain fully why it is true. Op. cit.

straint which violates a Pareto condition in the ratio form (2.1) necessarily implies some sort of relationship between the 1^{st} and n^{th} commodities. Hence, the choice of the n^{th} commodity as the numéraire means that it will have a non-neutral effect upon Pareto conditions stated in a ratio form (2.5) since it, as well as the 1^{st} commodity, was involved in constraint (2.1). In this regard note that if the j^{th} commodity $(1 \neq j \neq n)$ were chosen as the numéraire and utilized in the elimination of the multiplier λ_0, then, with the exception of the 1^{st} and n^{th} commodities, the Pareto conditions would take their usual form.

We have gone to some lengths to explain the difficulties inherent in the Lipsey-Lancaster approach. We have not tried to resolve these difficulties. Instead, in the pages which follow we shall attempt to avoid these difficulties in reconstructing a theory of second best.

4. Explicit and Implicit Assumptions in General Equilibrium Models:

In attempting to reconstruct a theory of second best two points are clear. First, the Pareto optimum conditions, as they are usually derived and stated (and as we have been using them), are not entirely satisfactory for the problem at hand. Not only is the ratio type of statement of these conditions rather confusing, but "corner solutions" are generally not allowed. Hence, if we wish to compare the conditions for second best optima with the conditions for a general Pareto optimum, we must begin by deriving the Pareto conditions in a more general form.

Second, the theory of second best is designed with reference to a functioning market economy which operates in a decentralized manner. Second best problems arise when one or more of the conditions necessary for the perfect operation of a decentralized price regime are violated. Recently, Gerard Debreu has studied carefully all of these conditions.[1] Therefore,

[1] Theory of Value, New York: John Wiley, 1959. Note, however, that Debreu uses the properties of convex set theory and avoids the need of assuming differentiability. We shall use the more traditional route and formulate a model which requires differentiability.

it is not necessary for us to consider all of these here, but it is necessary, at the very minimum, to point out those which we shall explicitly use and those which we shall violate subsequently in order to study second best problems.

The first of these requirements is that producers maximize profits and consumers maximize utility. This is a well-known requirement so that no further discussion is warranted, but it will be recalled that we have already considered the case where the production of one good did not satisfy this condition in our counter-example.

The second requirement is the assumption of diminishing marginal utility and decreasing marginal productivity so that production functions are convex and utility functions are concave.[1]

[1]Definition: A function $f(x)$ is convex if the domain of definition is convex, and if $(1 - \theta) f(x^*) + \theta f(x) \geq f((1 - \theta) x^* + \theta x)$ for $0 \leq \theta \leq 1$ and all x^* and x in the domain of definition of $f(x)$. Conversely, a function $f(x)$ is concave if $-f(x)$ is convex.

A third requirement is the complete absence of all externalities. In other words, the utility of each consumer depends only upon his given income and those goods which he himself both purchases and consumes; and the profits of each producer depend only upon price parameters, a given technology, and those inputs which he both purchases and utilizes.[2]

[2]See Gerard Debreu, "Coefficient of Resource Utilization," Econometrica, (July 1951), p. 277, for a rigorous definition of situations in which all technological externalities are absent. Note also footnotes 2, p. 49, and 6, p. 73, of Debreu, Theory of Value.

Finally, the economy is assumed to be fully competitive.

With the above in mind, let us state the following definitions which we shall use for the remainder of the paper.

u_i = the utility function of the i^{th} individual.

x_{ik} = the quantity of the k^{th} good consumed by the i^{th} individual.

$X_i = (x_{i1}, \ldots, x_{in})$, the "bundle" of goods consumed by the i^{th} individual.

y_{rk} = the quantity of the k^{th} good produced by the r^{th} firm.

g_{rkj} = a function indicating the quantity of the j^{th} resource utilized by the r^{th} firm in producing y_{rk} of the k^{th} good. It is assumed that $\frac{\partial g_{rkj}}{\partial y_{rk}} \geq 0$, and $g_{rkj}(0) = 0$, for all r, k and j.

L_j = the total available quantity of the j^{th} resource.

Since we are concerned here with deriving the general conditions for Pareto optimality, our criterion function -- i.e., the function to be maximized subject to certain plausible constraints -- must consist of individual utility functions. With this in mind, the above definitions allow us to consider the following vector maximization problem:

(3.1) $\max\ [u_1(X_1), \ldots, u_m(X_m)]$

subject to

(3)

(3.2) $\sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk}$, $\quad k = 1, \ldots, n$

(3.3) $\sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj}(y_{rk}) \leq L_j$ $\quad j = 1, \ldots, s$

(3.4) $x_{ik}, y_{rk} \geq 0$, $\quad i = 1, \ldots, m;\ r = 1, \ldots, z;\ k = 1, \ldots, n$.

This problem (3) is our basic model. In general, many solutions exist for problems of this type, but, as will be made clear later, this point need not bother us here. Note that (3.1) is a vector composed of individual utility functions. Constraint (3.2) requires that the sum of the quantities of the k^{th} good, $(k = 1, \ldots, n)$, which are obtained by the m consumers be less than or equal to the amount supplied by the z firms. Constraint (3.3), which states that the firms cannot use greater quantities of resources than are available in the society, carries hidden assumptions and requires a few words of explanation. We do not restrict firms to the production of one good. But by writing (3.3) in a separable form we are assuming not only that no technological externalities exist between firms but also that these externalities do not exist within the firm between the various goods which it might produce. Finally, the last constraint states plausible non-negativity requirements.

Having presented the above model of an economic system, we note that if $X_1^*, \ldots, X_m^*$ is a solution to (3), then this implies that there does not exist feasible quantities $X_1', \ldots, X_m'$ such that $u_i(X_i') \geq u_i(X_i^*)$ for each i, $(i = 1, \ldots, m)$, with at least one $u_i(X_i') > u_i(X_i^*)$. Thus any solution to (3) satisfies the Pareto welfare criterion. No individual can be made better off without making some other individual worse off. This point makes clear the reason why the possible existence of many solutions to (3) is of no concern here. We are interested only in efficient solutions i.e., Pareto optimal ones -- and have no concern for problems of distribution. Hence, any solution to (3) will do for the limited purposes of this paper.

For the purpose of deriving the general conditions for Pareto optimality, model (3) is not entirely satisfactory. It is necessary that we replace (3.1) by a scalar function. From the Kuhn-Tucker equivalence theorem we know that any vector maximization problem can be represented by a problem which has as its criterion function a positively weighted sum of the vectors.[1] In general, it is true that different weights are required

[1] H. Kuhn and A. Tucker, "Nonlinear Programming," Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, ed. J. Neymon, Berkeley: University of California Press, 1951, pp. 481-492.

for different solutions. But for each solution of (3) there exists a set of positive α_i, $(i = 1,\ldots,m)$, such that the solution to (4) below will be identical to a specified solution to (3).[2]

(4)

$$(4.1)\quad \max \sum_{i=1}^{m} \alpha_i\, u_i\,(X_i)$$

subject to

$$(4.2)\quad \sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk}, \qquad k = 1,\ldots,m$$

$$(4.3)\quad \sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj}\,(y_{rk}) \leq L_j, \qquad j = 1,\ldots,s$$

$$(4.4)\quad x_{ik},\ y_{rk} \geq 0, \qquad i = 1,\ldots,m;\ k = 1,\ldots,n;\ r = 1,\ldots,z$$

[2] Ibid.

Hence, it is entirely legitimate to use (4) as our model of an economic system since our sole interest here is in efficient solutions. Furthermore, the weight α_i can be interpreted as the reciprocal of the i^{th} consumer's marginal utility of income.[3]

[3] This interpretation is derived from the following considerations: A consumer maximization model is

$$(A)\quad \max\ u_i\ (X_i)$$

subject to

$$\text{(B)} \quad \sum_{k=1}^{n} \lambda_k x_{ik} \leq M_i .$$

where λ_k is the price of the k^{th} good and M_i is the given money income of the i^{th} consumer. Hence, the conditions for a consumer's utility maximum are:

$$\text{(C)} \quad \begin{cases} \dfrac{\partial u_i}{\partial x_{ik}} - \beta_i \lambda_k \leq 0 \\ x_{ik}\left(\dfrac{\partial u_i}{\partial x_{ik}} - \beta_i \lambda_k \right) = 0 \end{cases} \qquad k = 1,\ldots,n$$

These are interpreted as meaning that if the consumer purchases the good so that x_{ik} is positive, then strict equality must hold. Obviously, the multiplier β_i can be interpreted as the i^{th} consumer's marginal utility of income since it represents the "value" of the budget constraint M_i in (B). Defining $\alpha_i = 1/\beta_i$ we may write (C) in the form

$$\text{(D)} \quad \begin{cases} \alpha_i \dfrac{\partial u_i}{\partial x_{ik}} - \lambda_k \leq 0 \\ x_{ik}\left(\alpha_i \dfrac{\partial u_i}{\partial x_{ik}} - \lambda_k \right) = 0 \end{cases} \qquad k = 1,\ldots,n$$

since $\beta_i > 0$ is usually assumed. We shall see at a later point in our paper that these are of the same form as the Pareto optimality conditions for consumers which are derived from our general equilibrium model. Hence, it is appropriate to associate the "weights" in the criterion function (4.1) with the reciprocals of marginal utilities of individuals' incomes. This question of the relationship between "individual models" and a social welfare function is explored further in Otto A. Davis and Andrew B. Whinston, "Market Mechanisms and Social Welfare Functions," Graduate School of Industrial Administration, Carnegie Institute of Technology.

Sufficient conditions for a solution to (4) are that there exist non-negative vectors $(\lambda_1, \ldots, \lambda_n)$ and $(\mu_1, \ldots, \mu_s)$ which satisfy the conditions listed under (5) below.[1] Furthermore, the λ's

---

[1]Kuhn and Tucker derive these conditions. Op. cit.

---

will be identified with constraints (4.2) so that they may be interpreted as prices of goods; and the μ's with constraints (4.3) so that they can be interpreted as prices of resources.

We now turn our attention to the derivation of the Pareto optimum conditions. First, we consider those conditions which refer to the theory of consumer demand.

$$(5)\quad \begin{cases} (5.11) & \alpha_i \dfrac{\partial u_i}{\partial x_{ik}} - \lambda_k \leq 0 , & k = 1, \ldots, n \quad i = 1, \ldots, m \\ (5.12) & x_{ik}\left(\alpha_i \dfrac{\partial u_i}{\partial x_{ik}} - \lambda_k\right) = 0 , & k = 1, \ldots, n \quad i = 1, \ldots, m \end{cases}$$

Condition (5.11) states that at an equilibrium the i^{th} individual's marginal utility of the k^{th} good, weighed by the reciprocal α_i of his marginal utility of income, must be less than or equal to the price λ_k of that good. Condition (5.12) requires that strict equality be attained if the item is acquired -- i.e., if x_{ik} is positive. (Of course, x_{ik} can never be negative because of (4.4), but it may be identically zero.)

The equilibrium Pareto requirements in the markets for goods are:

(5)

$$(5.13) \quad \sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk} \, , \qquad k = 1, \ldots, n$$

$$(5.14) \quad \left(\sum_{i=1}^{m} x_{ik} - \sum_{r=1}^{z} y_{rk} \right) \lambda_k = 0, \qquad k = 1, \ldots, n \, .$$

The first of these (5.13) states that the quantity demanded of the k^{th} good must be less than or equal to the quantity supplied. The latter (5.14) requires that if the price λ_k is positive, then the quantity demanded must equal the quantity supplied.

For the production sector of the economy the Pareto optimum conditions are:

(5)

$$(5.15) \quad \lambda_k - \sum_{j=1}^{s} \mu_j \frac{\partial g_{rkj}}{\partial y_{rk}} \leq 0 \, , \qquad r = 1, \ldots, z \quad k = 1, \ldots, n$$

$$(5.16) \quad y_{rk} \left(\lambda_k - \sum_{j=1}^{s} \mu_j \frac{\partial g_{rkj}}{\partial y_{rk}} \right) = 0, \qquad r = 1, \ldots, z \quad k = 1, \ldots, n$$

For the r^{th} firm, condition (5.15) requires that the sum of the marginal productivities of the j^{th} resources, $(j = 1, \ldots, s)$, in the production of the k^{th} good, times the price of that resource u_j, be greater than or equal to the price λ_k of that good. In other words, marginal cost must be greater than or equal to price. Condition (5.16) requires that if the good is produced, then strict equality must be attained between the price λ_k and marginal cost. (Again, negative production is ruled out by constraint (4.4).)

The final conditions concern resource (factor) markets.

$$(5.17) \quad \sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj}(y_{rk}) \leq L_j \, , \qquad j = 1, \ldots, s$$

$$(5.18) \quad \left[\sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj} (y_{rk}) - L_j \right] \mu_j = 0, \quad j = 1, \ldots, s.$$

The first of these (5.17) states that no more than the available quantity L_j of the j^{th} resource can be used. The latter (5.18) requires that if a factor price μ_j is positive, then all of the available quantity of the resource must be used.

5. Special Problems and Second Best Solutions:

Having presented a more general version of the conditions for Pareto optimality than those usually stated, we now turn our attention to a series of second best problems. As was mentioned earlier, second best problems arise when one or more of the conditions which are necessary for a decentralized price regime to operate efficiently are not satisfied. Hence, one or more of the Pareto conditions stated in (5) above will be violated, and this violation is taken as given. The appropriate questions, then, are: (i) What constitutes a second best solution? (ii) What are the conditions for a second best solution and do these differ from the Pareto conditions? (iii) Does there exist a viable decentralized price mechanism associated with this allocation?

For the remainder of this paper we shall be concerned with the latter two of the above questions. In order to answer the first question here, however, we need to consider the relationship between problems whose solutions are Pareto optimal and those problems which give second best solutions.

In the previous section we derived the conditions for Pareto optimality by considering a vector maximization model which contained resource and market-clearing constraints and which assumed given taste and technology. Any solution to that problem was Pareto optimal. Furthermore, although no assumptions concerning the behavior of individuals and firms were explicitly introduced into the model, the optimality conditions themselves can be interpreted as normative behavioral rules which are applicable to the particular situation in which they were derived, and these rules must be followed by consumers and firms if the solution is to be achieved. Note, however, that the Pareto optimum conditions (5) are achievable by a decentralized market mechanism which operates in a competitive manner. Indeed, these are the very conditions which, granted certain assumptions, are required for a decentralized resource allocation mechanism to achieve a Pareto optimal solution.[1]

---

[1]See T. C. Koopmans, Three Essays on the State of Economic Science, New York: McGraw-Hill, 1957, pp. 4-126, for an especially clear discussion of the fact that a decentralized, competitive price regime implies the achievement of a Pareto optimal solution if certain conditions are made.

Following the usual conventions, we shall associate Pareto optimality with models which have only resource and market-clearing constraints. Of course, additional constraints are possible; and in the context of the problems under consideration here these additional constraints will usually be of such a form that one or more of the conditions for Pareto optimality will be violated. We shall term problems which are formulated in such a manner as problems of second best. Hence, second best problems involve not only the usual market-clearing and resource constraints but other constraints also.

In general we shall formulate second best problems by making the additional constraint take the form of a decision-making or behavioral rule of one or more of the entities (consumers or firms) in the economic system. Then, granted this additional constraint, we shall derive the conditions for a second best optimum, and, just as was the case with the Pareto conditions, the second best conditions can be interpreted as normative behavioral rules which must be followed by the entities in the system if the second best solution is to be achieved. In other words, the problem of second best might be posed as follows: Granted a deviation from Pareto optimal behavior by one or more of the units in the economic system, what behavioral rules should the remainder of the units follow in order to compensate best for the deviant. Hence, given the constraints, a second best solution is also an efficient solution in the sense that there does not exist an alternative allocation which will make at least one person better off without making anyone else worse off. Of course, if it were possible to alter the behavior of the deviant so that the added constraint could be removed, then in general it would be true that second best solutions are inferior to Pareto optimal solutions in that the removal of the extra constraint would make it possible to improve the position of one or more persons without worsening the others.

Having discussed verbally the meaning of a solution to a second best problem, it is appropriate to point out that we shall be able to compare always the Pareto with the second best conditions and to inquire whether a decentralized price regime exists under which the second best solution is achievable. However, the Pareto conditions shown under (5) above, and which were derived from basic model (3), are those which a perfectly functioning competitive price mechanism would achieve since the assumptions underlying that model are compatible with a competitive market regime which operates in a decentralized manner. It is possible, however, to alter these underlying assumptions and, equivalently, to alter basic model (3) and then derive Pareto conditions. In general, some of the Pareto conditions which are derived in such a manner will differ from those listed under (5). Whenever this is true, in the interest of brevity we shall derive only those conditions which differ from those discussed in (5) above.

Finally, some brief remarks concerning the comparison of Pareto and second best conditions are warranted. It cannot be over-emphasized that optimality conditions (Pareto and second best) relate to efficient solutions only. Hence, the conditions of interest are the functional forms. This fact makes it legitimate for us to compare Pareto with second best conditions without making reference to the fact that the values of both the variables and the parameters will, in general, be different at the two (Pareto and second best) solutions.[1]

---

[1]Again, we have not found it necessary to specify the distribution of income in our models, or even to assume it constant when adding constraints to a model, since we are interested only in efficient solutions in this paper. In short, the distribution of income is of no interest here.

Case I: Interdependent Utility Functions:

One of the assumptions underlying model (3) was that no externalities existed. Let us alter this assumption slightly and suppose that the utility of the 2nd consumer depends not only upon his own consumption but also upon the 1st individual's consumption x_{11} of the 1st good. In other words, we have

$$u_2(X_2, x_{11}) \qquad \text{instead of} \qquad u_2(X_2) ,$$

and the utility functions of all the individuals and all other assumptions remain as in model (3) so that we have here one external effect in consumption. Therefore, in formulating a general equilibrium model which incorporates this new feature we have as a criterion vector

$$(6) \qquad \max \quad \left(u_1(X_1) \ , \ u_2(X_2 \ , x_{11}) \ , \ \ldots , \ u_m(X_m) \right) \ .$$

We already know, however, that the Kuhn-Tucker equivalence theorem allows us to write a vector maximization problem as a problem which has as its criterion function a positively weighed sum of the vectors. Hence, we proceed directly to write:[1]

[1]We shall always follow in the remainder of the paper the convention of writing the problem directly without discussing the application of the Kuhn-Tucker theorem to a vector maximization problem.

$$(7.1) \qquad \max \ \alpha_1 u_1 (X_1) + \alpha_2 u_2 (X_2, x_{11}) + \sum_{i=3}^{m} \alpha_i u_i (X_i)$$

subject to

$$(7) \quad \begin{aligned} &(7.2) \quad \sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk} \qquad , \quad k = 1,\ldots,n \\ &(7.3) \quad \sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj}(y_{rk}) \leq L_j \quad , \quad j = 1,\ldots,s \\ &(7.4) \quad x_{ik}, y_{rk} \geq 0 \ . \qquad \begin{array}{l} i = 1,\ldots,m \\ k = 1,\ldots,n \\ r = 1,\ldots,z \end{array} \end{aligned}$$

as the model of our economic system. Note that there is nothing in our assumption of one consumption externality to alter the form of the constraints of the model so that, with the exception of the criterion function (7.1), this model (7) is identical to our basic model (4). Furthermore, we have no behavioral constraints in the model so that the optimality conditions will represent normative behavioral rules for the achievement of a Pareto optimum.

We now turn our attention toward the derivation of the Pareto conditions for model (7). Since the constraints are unaltered from those of (4), it is obvious that, with the exception of the consumption sector, the Pareto conditions will be of the same form as those discussed under (5) above. Hence, the equilibrium Pareto requirements in the market for goods are given by (5.13) and (5.14), those for the productive sector are (5.15) and (5.16), and (5.17) and (5.18) are the conditions for the resource markets. For the consumption sector we may derive

$$(8.1) \quad \alpha_i \frac{\partial u_i}{\partial x_{ik}} - \lambda_k \leq 0, \qquad \begin{array}{l} i = 2,\ldots,m \\ k = 1,\ldots,n \\ i = 1 \text{ and } k = 2,\ldots,n \end{array}$$

$$(8.2) \quad x_{ik} \left(a_i \frac{\partial u_i}{\partial x_{ik}} - \lambda_k \right) = 0. \qquad \begin{array}{l} i = 2,\ldots,m \\ k = 1,\ldots,n \\ i = 1 \text{ and } k = 2,\ldots,n \end{array}$$

so that these conditions are similar to (5.11) and (5.12) except that the 1^{st} consumer's choice x_{11} of the 1^{st} good is omitted. This divergence from the earlier conditions is caused, of course, by the assumed externality; and for that particular choice the Pareto optimum conditions are:

$$(8.3) \quad a_1 \frac{\partial u_1}{\partial x_{11}} + a_2 \frac{\partial u_2}{\partial x_{11}} - \lambda_1 \leq 0$$

$$(8.4) \quad x_{11} \left(a_1 \frac{\partial u_1}{\partial x_{11}} + a_2 \frac{\partial u_2}{\partial x_{11}} - \lambda_1 \right) = 0$$

so that optimality requires that the 1^{st} consumer take into account the effect of his choice x_{11} of the 1^{st} good on the well-being of the 2^{nd} individual.

It is the above result which causes the difficulty here. Unlike the Pareto conditions for previous model (4), which reflected the assumptions of that model and which were compatible with the utility and profit maximizing assumptions of the theory of competitive markets, the above conditions (8.3) and (8.4) indicate that a decentralized price regime would not achieve an efficient solution to this model (7). The reason is that individuals are assumed to maximize their own utility so that the 1^{st} consumer is not motivated to take into account the effect which his choice x_{11} of the 1^{st} good has upon the utility of the 2^{nd} consumer.[1] Instead the utility

[1] This statement assumes, of course, that the 2nd consumer does not bribe or otherwise influence the 1st consumer to follow the Pareto optimal rules (8.3 and 8.4).

maximization assumption requires that the 1st individual choose all goods, including the 1st good, according to the following behavioral rules:

$$(9.1) \qquad \alpha_1 \frac{\partial u_1}{\partial x_{1k}} - \lambda_k \leq 0 \quad , \quad k = 1,\ldots,n$$

$$(9.2) \qquad x_{1k} \left(\alpha_1 \frac{\partial u_1}{\partial x_{1k}} - \lambda_k \right) = 0, \quad k = 1,\ldots,n$$

so that Pareto conditions (8.3) and (8.4) will be violated for $k = 1$.

Having presented the Pareto optimum conditions for this new model (7) and having indicated that a decentralized pricing mechanism may not attain an efficient solution, we now construct a problem of second best. As will be recalled from our earlier discussion, we formulate a second best problem by attaching a behavioral constraint which violates one or more Pareto conditions to the other constraints of the model. Hence, we consider the following problem of second best:

$$(10.1) \qquad \max \ \alpha_1 u_1 (X) + \alpha_2 u_2 (X_2, x_{11}) + \sum_{i=3}^{m} \alpha_i u_i (X_i)$$

subject to

(10)

$$(10.2) \quad \sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk} \quad , \quad k = 1, \ldots, n$$

$$(10.3) \quad \sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj} \; (y_{rk}) \leq L_j \quad , \quad j = 1, \ldots, s$$

$$(10.4) \quad x_{ik}, \, y_{rk} \geq 0 \quad , \quad \begin{array}{l} i = 1, \ldots, m \\ r = 1, \ldots, z \\ k = 1, \ldots, n \end{array}$$

$$(10.5) \quad \alpha_1 \frac{\partial u_1}{\partial x_{1k}} - \lambda_k \leq 0 \quad , \quad k = 1, \ldots, n$$

$$(10.6) \quad x_{1k} \left(\alpha_1 \frac{\partial u_1}{\partial x_{1k}} - \lambda_k \right) = 0, \quad k = 1, \ldots, n \, .$$

Note that this second best problem (10) differs from our previous model (7) only in that behavioral rules (9.1) and (9.2), which specified that the 1st consumer acted in an individually rational manner and which violated Pareto conditions (8.3) and (8.4), are added as constraints (10.5) and (10.6).[1]

---

[1]Again, it is necessary that we point out that the α_i are parameters which are specified (say, by some omniscient observer) to be identical to individual marginal utilities of incomes. We have no need to actually specify here the numerical values of the α_i (aside from the fact that all α_i are positive) and, indeed, the values may be different for each model under consideration. This fact causes no difficulty since we are interested only in comparing functional forms.

Before proceeding to derive the conditions for a second best optimum, it is appropriate that we observe that the added constraints (10.5) and (10.6), which involve the presence of $\lambda_k (k=1, \ldots, n)$, make our formulation of this second best problem (10) into a somewhat unusual mathematical form. The λ_k are the imputed prices of the k^{th} goods and are associated with constraints (10.2). Yet, these λ_k appear in the "direct" problem. This aspect of our formulation, which is a mathematical sideline to the main stream of argument here, is dealt with in the appendix. While we deal with the matter more fully there, it is sufficient for an understanding of our argument to consider the λ_k which appear in constraints (10.5) and (10.6) to be parameters which are restricted to be equal to the values of the multipliers associated with constraints (10.2).

With the above remarks in mind, we turn our attention toward the derivation of the conditions for a second best optimum. We are interested in seeing whether these conditions differ in form from the Pareto optimum conditions. Consider first those which refer to the theory of consumer demand.

$$(11.11) \qquad a_i \frac{\partial u_i}{\partial x_{ik}} - \lambda_k \leq 0 , \qquad \begin{matrix} i = 2, \ldots, m \\ k = 1, \ldots, n \end{matrix}$$

$$(11.12) \qquad x_{ik}\left(a_i \frac{\partial u_i}{\partial x_{ik}} - \lambda_k \right) = 0, \qquad \begin{matrix} i = 2, \ldots, m \\ k = 1, \ldots, n \end{matrix}$$

Note that these second best conditions (11.11) and (11.12) are identical in form to Pareto optimum conditions (8.1) and (8.2) except for the fact that the 1^{st} consumer ($i = 1$) is omitted from the above conditions. Hence, with the noted exception, the attachment of additional constraints to model (7) in order to form the second best problem (10) did not cause a change in the form of the optimality conditions in consumption. This result stands in marked contrast to the popular interpretation of the Lipsey - Lancaster theorem of second best.

The 1^{st} consumer's insistence upon individual rationality created, in this particular instance, our second best problem. The conditions for a second best optimum in regard to the 1^{st} individual are stated below where the multipliers η_k and ρ_{k1} ($k = 1,\ldots,n$), refer to constraints (10.5) and (10.6) respectively.

$$(11.13)\quad \alpha_1 \frac{\partial u_1}{\partial x_{11}} + \alpha_2 \frac{\partial u_2}{\partial x_{11}} - \lambda_1 - \sum_{k=1}^{n} \eta_k \alpha_1 \frac{\partial^2 u_1}{\partial x_{1k} \partial x_{11}} - \rho_1 \left(\alpha_1 \frac{\partial u_1}{\partial x_{11}} - \lambda_1 \right) - \sum_{k=1}^{n} \rho_k x_{1k} \left(\alpha_1 \frac{\partial^2 u_1}{\partial x_{1k} \partial x_{11}} \right) \leq 0$$

$$(11.14)\quad x_{11} \left[\alpha_1 \frac{\partial u_1}{\partial x_{11}} + \alpha_2 \frac{\partial u_2}{\partial x_{11}} - \lambda_1 - \sum_{k=1}^{n} \eta_k \alpha_1 \frac{\partial^2 u_1}{\partial x_{1k} \partial x_{11}} - \rho_1 \left(\alpha_1 \frac{\partial u_1}{\partial x_{11}} - \lambda_1 \right) - \sum_{k=1}^{n} \rho_k x_{1k} \left(\alpha_1 \frac{\partial^2 u_1}{\partial x_{1k} \partial x_{11}} \right) \right] = 0$$

$$(11.15)\quad \alpha_1 \frac{\partial u_1}{\partial x_{1k}} - \lambda_k - \sum_{e=1}^{n} \eta_e \alpha_1 \frac{\partial^2 u_1}{\partial x_{1e} \partial x_{1k}} - \rho_k \left(\alpha_1 \frac{\partial u_1}{\partial x_{1k}} - \lambda_k \right) - \sum_{e=1}^{n} \rho_e x_{1e} \left(\alpha_1 \frac{\partial^2 u_1}{\partial x_{1e} \partial x_{1k}} \right) \leq 0 \qquad k = 2,\ldots,n$$

$$(11.16) \qquad x_{1k}\left[\alpha_1 \frac{\partial u_1}{\partial x_{1k}} - \lambda_k - \sum_{e=1}^{n} \eta_e \alpha_1 \frac{\partial^2 u_1}{\partial x_{1e} \partial x_{1k}} - \rho_k \left(\alpha_1 \frac{\partial u_1}{\partial x_{1k}} - \lambda_k\right) - \sum_{e=1}^{n} \rho_e x_{1e} \left(\alpha_1 \frac{\partial^2 u_1}{\partial x_{1e} x_{1k}}\right)\right] = 0, \qquad k = 2,\ldots,n$$

$$(11.17) \qquad \alpha_1 \frac{\partial u_1}{\partial x_{1k}} - \lambda_k \leq 0. \qquad k = 1,\ldots,n$$

$$(11.18) \qquad x_{1k}\left(\alpha_1 \frac{\partial u_1}{\partial x_{1k}} - \lambda_k\right) = 0. \qquad k = 1,\ldots,n$$

These conditions are, of course, different from the Pareto optimum conditions (8.3) and (8.4) which, by the formulation of the second best problem, had to be violated. In addition to the above conditions (11.13 - 11.18) we know that for consistency the multipliers η_k and ρ_k, ($k = 1,\ldots,n$), must satisfy certain conditions. These are considered in (11.25) and (11.26) below.

The second best conditions for the market for goods are

$$(11.19) \qquad \sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk} \;, \qquad k = 1,\ldots,n$$

$$(11.20) \qquad \left(\sum_{i=1}^{m} x_{ik} - \sum_{r=1}^{z} y_{rk}\right) \lambda_k = 0, \qquad k = 1,\ldots,n.$$

Note that these are the same as the Pareto optimum conditions (5.13) and (5.14).

The second best conditions for the productive sector are as follows:

$$(11.21) \qquad \lambda_k - \sum_{j=1}^{s} \mu_j \frac{\partial g_{rkj}}{\partial y_{rk}} \leq 0, \qquad \begin{matrix} k = 1, \ldots, n \\ r = 1, \ldots, z \end{matrix}$$

$$(11.22) \qquad y_{rk} \left(\lambda_k - \sum_{j=1}^{s} \mu_j \frac{\partial g_{rkj}}{\partial y_{rk}} \right) = 0, \qquad \begin{matrix} k = 1, \ldots, n \\ r = 1, \ldots, z \end{matrix}$$

and these are of identical form as Pareto optimum conditions (5.15) and (5.16).

The second best conditions for the factor market are:

$$(11.23) \qquad \sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj} (y_{rk}) \leq L_j \, , \qquad j = 1, \ldots, s$$

$$(11.24) \qquad \mu_j \left(\sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj} (y_{rk}) - L_j \right) = 0, \qquad j = 1, \ldots, s$$

and these are the same as Pareto optimum conditions (5.17) and (5.18).

Assuming that the model possesses properties sufficient for a solution to exist,[1] we observe that the multipliers η_k and ρ_k,

[1] See the discussion in the appendix for the definition of a solution to this type of problem.

$(k = 1, \ldots, n)$, must satisfy the following conditions since (11.17) and (11.18) must hold by the assumptions of the problem:

$$(11.25)\quad x_{11}\left[\sum_{k=1}^{n}\eta_k\frac{\partial^2 u_1}{\partial x_{1k}\partial x_{11}}+\rho_1\left(\alpha_1\frac{\partial u_1}{\partial x_{11}}-\lambda_1\right)+\sum_{k=1}^{n}\rho_k x_{1k}\left(\alpha_1\frac{\partial^2 u_1}{\partial x_{1k}\partial x_{11}}\right)\right]=\left[\alpha_2\frac{\partial u_2}{\partial x_{11}}\right]x_{11}$$

$$(11.26)\quad x_{1k}\left[-\sum_{e=1}^{n}\eta_e\frac{\partial^2 u_1}{\partial x_{1e}\partial x_{1k}}-\rho_k\left(\alpha_1\frac{\partial u_1}{\partial x_{1k}}-\lambda_k\right)-\sum_{e=1}^{n}\rho_e x_{1e}\left(\alpha_1\frac{\partial^2 u_1}{\partial x_{1e}\partial x_{1k}}\right)\right]=0,\quad k=2,\ldots,n$$

Having derived the conditions for the achievement of a second best optimum, we make the following observations: First, these second best conditions are compatible with the underlying assumptions of utility and profit maximization. Hence, if the other assumptions underlying market allocation models are satisfied here, and if a solution exists and is stable, then a competitive pricing mechanism exists which is associated with the solution to the above problem. Second, these conditions for a second best optimum are of the same form as those for a Pareto optimum with the exception of the second best conditions which refer to the 1st consumer who, by the assumptions of the problem, departed from the Pareto conditions. Finally, if at least some of the additional constraints are not redundant, then a second best solution will be "inferior" to a Pareto optimum solution in the sense that a removal of those additional constraints would allow some consumers to be made better off without any consumers being made worse off.

Case II: A Production Externality:

Let us now return to our original assumption that there are no externalities in consumption so that all utility functions are again of the form

$$u_i \; (X_i) \;, \qquad i = 1, \ldots, m.$$

Recalling that $g_{rkj}(y_{rk})$ was defined as a function which indicated the quantity of the j^{th} resource used by the r^{th} firm in producing y_{rk} of the k^{th} good, let us suppose that only the first firm is affected by an externality. Specifically, assume that the amount of the 1^{st} resource used by the 1^{st} firm in producing the 1^{st} good depends not only on the amount of that good y_{11} which it produces, but also upon the amount of that good y_{21} which the 2^{nd} firm produces. Hence, we have the following model of an economic system:

$$(12.1) \qquad \max \sum_{i=1}^{m} \alpha_i \, u_i \, (X_i)$$

subject to

$$(12.2) \qquad \sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk} \;, \qquad k = 1, \ldots, n$$

$$(12.3) \quad g_{111}(y_{11}, y_{21}) + \sum_{k=2}^{n} g_{1k1}(y_{1k}) + \sum_{k=1}^{n} \sum_{r=2}^{z} g_{rk1}(y_{rk}) \leq L_1$$

(12) $$(12.4) \quad \sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj}(y_{rk}) \leq L_j, \qquad j = 2, \ldots, s$$

$$(12.5) \quad x_{ik}, y_{rk} \geq 0, \qquad i = 1, \ldots, m; \; r = 1, \ldots, z; \; k = 1, \ldots, n .$$

Obviously, constraint (12.3) is the one which reflects the externality. It contains the "externality term" $g_{111}(y_{11}, y_{21})$ as explained above, and the remainder of that constraint concerns the 1^{st} resource in all other possible uses.[1] Constraint (12.4) refers to the use of all other resources and, as postulated, no externalities are present there.

---

[1]Note that there is nothing in the model to restrict a firm to the production of only one good. Hence the term $\sum_{k=2}^{n} g_{1k1}(y_{1k})$ is needed to indicate possible uses (if any) of the 1^{st} resource by the 1^{st} firm in the possible (but not required) production of other goods $(k = 2, \ldots, n)$.

Let us now turn our attention to the derivation of the Pareto optimum conditions for model (12). It is obvious from the manner in which the externality enters the constraints that the Pareto conditions for consumption and the market for goods are not affected. In other words, the Pareto conditions in consumption are given by (5.11) and (5.12), and those for the goods market are (5.13) and (5.14). The Pareto conditions for the factor market will reflect the form of constraints (12.3) and (12.4) so that we have

$$(13.1)\quad g_{111}(y_{11}, y_{21}) + \sum_{k=2}^{n} g_{1k1}(y_{1k}) + \sum_{k=1}^{n} \sum_{r=2}^{z} g_{rk1}(y_{rk}) \leq L_1$$

$$(13.2)\quad \left[g_{111}(y_{11}, y_{21}) + \sum_{k=2}^{n} g_{1k1}(y_{1k}) + \sum_{k=1}^{n} \sum_{r=2}^{z} g_{rk1}(y_{rk}) - L_1 \right] \mu_1 = 0.$$

for the 1^{st} resource, and for all other resources ($j = 2, \ldots, s$) the conditions are given by (5.17) and (5.18). Finally, the Pareto conditions in production also reflect the externality. For the 2^{nd} firm's choice of output of the 1^{st} good we have

$$(13.3)\quad \lambda_1 - \mu_1 \frac{\partial g_{111}}{\partial y_{21}} - \sum_{j=1}^{s} \mu_j \frac{\partial g_{21j}}{\partial y_{21}} \leq 0$$

$$(13.4)\quad y_{21}\left(\lambda_1 - \mu_1 \frac{\partial g_{111}}{\partial y_{21}} - \sum_{j=1}^{s} \mu_j \frac{\partial g_{21j}}{\partial y_{21}} \right) = 0$$

and the remaining conditions are

$$(13.5)\quad \lambda_k - \sum_{j=1}^{s} \mu_j \frac{\partial g_{rkj}}{\partial y_{rk}} \leq 0, \qquad \begin{array}{l} r = 1, 3, \ldots, z \\ k = 1, \ldots, n \\ \text{and for} \\ r = 2; \ k = 2, \ldots, n \end{array}$$

$$(13.6)\quad y_{rk} \left(\lambda_k - \sum_{j=1}^{s} \mu_j \frac{\partial g_{rkj}}{\partial y_{rk}} \right) = 0, \qquad \begin{array}{l} r = 1, 3, \ldots, z \\ k = 1, \ldots, n \\ \text{and for} \\ r = 2; \ k = 2, \ldots, n. \end{array}$$

Since these Pareto conditions are normative behavioral rules which must be followed if an efficient solution is to be achieved, it is interesting to note that while conditions (13.5) and (13.6) require individually rational (profit maximizing) action in a competitive setting -- i.e., firms choose quantities y_{rk} such that marginal cost equals price if the good is produced -- conditions (13.3) and (13.4) require that the 2nd firm produce a quantity y_{21} of the 1st good which equates social rather than private cost with price. In other words, the 2nd producer must take into account the effect which his production of the 1st good has on the 1st firm if a Pareto efficient solution is to be achieved. Hence, constraints (13.3) and (13.4) are in conflict with the assumption of profit maximization.[1] In a setting of competitive markets the 2nd producer would have no motivation to use Pareto conditions (13.3) and (13.4) as behavioral rules so that a decentralized price regime would not achieve the Pareto optimal solution.

[1] Of course, this excludes the possibility that the 1st firm might bribe the 2nd to act in the Pareto optimal manner.

The above result makes natural our formulation of a problem of second best. Suppose that the 2nd producer makes his choice of output y_{21} of the 1st good without regard for the 1st producer. In other words, assume he follows the behavioral rules:[2]

[2] Note that while it was necessary in the previous case to constrain our deviant 1st consumer in all his choices because of the fact that individual utility functions were not assumed to be separable, it is only necessary to constrain the 2nd producer here in the choice of the 1st good since separability is assumed in production.

$$(14.1) \qquad \lambda_1 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{21j}}{\partial y_{21}} \leq 0$$

$$(14.2) \qquad y_{21}\left(\lambda_1 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{21j}}{\partial y_{21}}\right) = 0$$

The second best problem is formulated, of course, by attaching behavioral constraints (14.1) and (14.2) to the other constraints of model (12). However, having displayed the second best conditions for the previous case in their explicit (and long and complicated) form, it is desirable in the interest of brevity to change the form in which constraints (14.1) and (14.2) are written. For given prices λ_1 and μ_j, $(j = 1,\ldots,s)$, the above constraints represent a restriction on the choice of y_{21}. Hence, we may write (14.1) and (14.2) in the form

$$(15) \qquad H(\lambda_1, \mu_1, \ldots, \mu_s, y_{21}) \leq 0$$

and attach (15) to the other constraints of model (12) in order to formulate a problem of second best.

Noting that constraint (15) involves only the variable y_{21}, we observe that the form of the second best conditions will be the same as the form of the Pareto conditions for consumption (5.11) and (5.12), for the goods market (5.13) and (5.14), and for the resource market (13.1) and (13.2) supplemented by (5.17) and (5.18) for $j = 2,\ldots,s$. Even in the production section the variable y_{21} affects only the 2nd firm's choice of output for the 1st good. With this exception it follows that the second best condition will be of the same form as Pareto optimum conditions (13.5) and (13.6). For the 2nd producer's choice of the 1st good the second best conditions are:

$$(16.1) \quad \lambda_1 - \mu_1 \frac{\partial g_{111}}{\partial y_{21}} - \sum_{j=1}^{s} \mu_j \frac{\partial g_{21j}}{\partial y_{21}} - \eta \frac{\partial H}{\partial y_{21}} \leq 0$$

$$(16.2) \quad y_{21} \left(\lambda_1 - \mu_1 \frac{\partial g_{111}}{\partial y_{21}} - \sum_{j=1}^{s} \mu_j \frac{\partial g_{21j}}{\partial y_{21}} - \eta \frac{\partial H}{\partial y_{21}} \right) = 0$$

$$(16.3) \quad \lambda_1 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{21j}}{\partial y_{21}} \leq 0$$

$$(16.4) \quad y_{21} \left(\lambda_1 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{21j}}{\partial y_{21}} \right) = 0.$$

Hence, with the exception of the conditions for the 2nd producer's choice of the 1st good, the second best optimum conditions are of the same form as the Pareto optimum conditions. This result again stands in marked contrast to the popular interpretation of the Lipsey - Lancaster theorem of second best.

Finally, note that all second best conditions are consistent with the assumptions of models of a decentralized, competitive price regime. Hence, there exists such a market mechanism associated with the solution to this problem of second best.

Case III. Mutual Externalities in Production

In both of the cases previously considered a violation of one Pareto optimum condition resulted in a problem of second best where, with the exception of the specified violation, the conditions for a second best optimum were of the same form as the original Pareto conditions. These results are exactly opposite to the popular interpretation of the general theorem of second best. But this opposite result—that a violation of one Pareto condition leaves all the others still desirable—is not always true. We now examine a case which is designed to demonstrate the element of validity in the popular interpretation of the Lipsey - Lancaster position.

Suppose, as in Case II, that the quantity of the 1st resource used by the 1st firm in producing the 1st good depends upon both the quantity y_{11} which it produces and upon the amount y_{21} which the 2nd firm produces. Furthermore, assume that the interaction is mutual so that the amount of the 1st resource used by the 2nd firm in producing the 1st good depends upon the quantity y_{21} which it produces and upon the amount of that good y_{11} which the 1st firm produces. Also, we assume initially that the mutual externality terms $g_{111}(y_{11}, y_{21})$ and $g_{211}(y_{11}, y_{21})$ are non-separable. All other relationships and assumptions are the same as those specified for the basic model (3).

Given the above assumptions, we have the following model of an economic system:

(17.7) $$\max \sum_{i=1}^{m} \alpha_i u_i(X_i)$$

subject to

$$(17.2)\qquad \sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk}, \qquad k = 1,\ldots,n$$

$$(17.3)\qquad g_{111}(y_{11}, y_{21}) + \sum_{k=2}^{n} g_{1k1}(y_{1k}) + g_{211}(y_{11}, y_{21}) + \sum_{k=2}^{n} g_{2k1}(y_{2k}) + \sum_{k=1}^{n} \sum_{r=3}^{z} g_{rk1}(y_{rk}) \leq L_1$$

$$(17.4)\qquad \sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj}(y_{rk}) \leq L_j \qquad j = 2,\ldots,s$$

$$(17.5)\qquad x_{ik}, y_{rk} \quad 0 \qquad \begin{array}{l} i = 1,\ldots,m \\ r = 1,\ldots,z \\ k = 1,\ldots,n \end{array}$$

where constraint (17.3), which concerns the 1st resource and contains our externality terms, differs in the specified manner from constraint (12.3) of the previous model.

Let us now consider the derivation of the Pareto optimum conditions. Once again, it is obvious from the manner in which the externalities enter the constraints that the Pareto conditions for this model (17) are of the same form as those of the basic model (4) for both consumption and the goods market. Hence (5.11) and (5.12), and (5.13) and (5.14) are applicable here. Clearly, the form of the Pareto conditions for the resources market must reflect the altered form of the constraints although the basic idea and the interpretation are unaltered. These conditions which must be satisfied at an optimum solution are:

(18.11)
$$g_{111}(y_{11}, y_{21}) + \sum_{k=2}^{n} g_{1k1}(y_{1k}) + g_{211}(y_{11}, y_{21}) + \sum_{k=2}^{n} g_{2k1}(y_{2k}) + \sum_{k=1}^{n} \sum_{r=3}^{z} g_{rk1}(y_{rk}) \leq L_1$$

(18.12)
$$\mu_1 \left[g_{111}(y_{11}, y_{21}) + \sum_{k=2}^{n} g_{1k1}(y_{1k}) + g_{211}(y_{11}, y_{21}) + \sum_{k=2}^{n} g_{2k1}(y_{2k}) + \sum_{k=1}^{n} \sum_{r=3}^{z} g_{rk1}(y_{rk}) - L_1 \right] = 0$$

(18.13)
$$\sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj}(y_{rk}) \leq L_j , \qquad j = 2,\ldots,s$$

(18.14)
$$\mu_j \left(\sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj}(y_{rk}) - L_j \right) = 0, \qquad j = 2,\ldots,s .$$

These conditions require no further interpretation.

However, the Pareto conditions for the production sector are of greatest interest here. These conditions are as follows:

(18.15)
$$\lambda_k' - \sum_{j=1}^{s} \mu_j \frac{\partial g_{rkj}(y_{rk})}{\partial y_{rk}} \leq 0 \qquad \begin{array}{l} r = 3,\ldots,z \\ k = 1,\ldots,n \\ \text{and for} \\ r = 1,2;\ k = 2,\ldots,n \end{array}$$

$$y_{rk}\left(\lambda_k - \sum_{j=1}^{s} \mu_j \frac{\partial g_{rkj}(y_{rk})}{\partial y_{rk}}\right) = 0, \tag{18.16}$$

$r = 3, \ldots, z$; $k = 1, \ldots, n$ and for $r = 1, 2$; $k = 2, \ldots, n$.

In the context of competitive markets these conditions require that firms act in an individually rational (profit maximizing) manner. Hence, these normative, behavioral rules require that output be chosen so that marginal cost equals price. However, the 1st and 2nd firms' choices of the 1st good are omitted from the above conditions, and for these choices the Pareto optimum conditions are listed below:

$$\lambda_1 - \mu_1 \frac{\partial g_{211}(y_{11}, y_{21})}{\partial y_{11}} - \mu_1 \frac{\partial g_{111}(y_{11}, y_{21})}{\partial y_{11}} - \sum_{j=2}^{s} \mu_j \frac{\partial g_{11j}(y_{11})}{\partial y_{11}} \leq 0 \tag{18.17}$$

$$y_{11}\left[\lambda_1 - \mu_1 \frac{\partial g_{211}(y_{11}, y_{21})}{\partial y_{11}} - \mu_1 \frac{\partial g_{111}(y_{11}, y_{21})}{\partial y_{11}} - \sum_{j=2}^{s} \mu_j \frac{\partial g_{11j}(y_{11})}{\partial y_{11}}\right] = 0 \tag{18.18}$$

$$\lambda_1 - \mu_1 \frac{\partial g_{111}(y_{11}, y_{21})}{\partial y_{21}} - \mu_1 \frac{\partial g_{211}(y_{11}, y_{21})}{\partial y_{21}} - \sum_{j=2}^{s} \mu_j \frac{\partial g_{21j}(y_{21})}{\partial y_{21}} \leq 0 \tag{18.19}$$

$$(18.20) \qquad y_{21}\left(\lambda_1 - \mu_1 \frac{\partial g_{111}(y_{11}, y_{21})}{\partial y_{21}} - \mu_1 \frac{\partial g_{211}(y_{11}, y_{21})}{\partial y_{21}} - \sum_{j=2}^{s} \mu_j \frac{\partial g_{21j}(y_{21})}{\partial y_{21}}\right) = 0$$

Identifying, as usual, the Pareto conditions with normative behavioral rules, it is apparent from (18.17) and (18.18) that the 1st firm should consider the effect which its output has on the 2nd firm when the production level y_{11} of the 1st good is chosen. Correspondingly, (18.19) and (18.20) require the 2nd firm also account for its effect upon the 1st firm when choosing an output y_{21} of the 1st good. In other words, the optimality conditions require that both firms choose outputs which equate marginal social cost with price, but marginal social cost is not equal to marginal private cost in either instance because of the presence of the terms for marginal interaction cost,

$$\mu_1 \frac{\partial g_{211}(y_{11}, y_{21})}{\partial y_{11}} \quad \text{and} \quad \mu_1 \frac{\partial g_{111}(y_{11}, y_{21})}{\partial y_{21}}$$

in the Pareto conditions for the 1st and 2nd firms respectively.

Having developed the Pareto optimum conditions for this model, we turn to second best problems. Several such problems are possible here, but we initially pose the following one: Suppose that the 2nd firm is privately owned and attempts to make all its decisions in an individually rational -- i.e., profit maximizing -- manner.

It is obvious that in the formulation of a problem of second best this firm must be constrained to set marginal cost equal to price in its decisions concerning the 1st good. Hence, we pose the behavioral rule

$$(19.1) \qquad \lambda_1 - \mu_1 \frac{\partial g_{211}(y_{11}, y_{21})}{\partial y_{21}} + \sum_{j=2}^{s} \mu_j \frac{\partial g_{21j}(y_{21})}{\partial y_{21}} \leq 0$$

$$(19.2) \qquad y_{21} \left(\lambda_1 - \mu_1 \frac{\partial g_{211}(y_{11}, y_{21})}{\partial y_{21}} + \sum_{j=2}^{s} \mu_j \frac{\partial g_{21j}(y_{21})}{\partial y_{21}} \right) = 0$$

which will violate Pareto conditions (18.19) and (18.20).

Before formulating and considering a problem of second best, we observe that it is not at all clear that behavioral rules (19.1) and (19.2) can be accomplished in the absence of information concerning the decisions or decision rules of the 1st firm in its choice of a production level y_{11} of the 1st good. Since we are assuming non-separability for the externalities, the term $\frac{\partial g_{211}(y_{11}, y_{21})}{\partial y_{21}}$ is a function of both y_{21} and y_{11}, and y_{11} is not under the 2nd firm's control.[1] However, this fact need not cause our analysis to run into

---

[1] For an analysis of both separable and non-separable technological externalities in production see, Otto A. Davis and Andrew B. Whinston, "Externalities, Welfare, and the Theory of Games," Journal of Political Economy, LXX (June, 1962).

trouble. Our purpose is not to analyze the difficulties which the 2nd firm might experience in trying to make its output decisions but, rather, to derive the conditions for a second best optimum. Hence, we may utilize (19.1) and (19.2) in the construction of a second best problem, although the resulting conditions may not be easily attainable.

Let us suppose now that the 1st firm is owned by the government and that it desires to choose an output which will maximize welfare rather than attempting to maximize profits without regard to the effect of the output choice upon welfare. Thus, the 1st firm desires to follow, as a decision rule, the second best conditions which apply to it.

Our interest, of course, is in deriving the second best condition not only for the 1st firm but for all the sectors of the economic system, and in determining whether these conditions differ in form from the Pareto optimal conditions. We formulate the second best problem by the usual method of attaching constraints (19.1) and (19.2) to the other constraints of model (17).[1]

---

[1]Unlike the previous case, we shall not rewrite and simplify the form of the additional constraints (19.1) and (19.2) here. The reason for using the complicated form is that, at a later point, we desire to assume separability in order to show exactly how this assumption affects the second best conditinns.

Noting that, for given prices, constraints (19.1) and (19.2) involve only the variables y_{11} and y_{21}, we immediately observe the following facts: The second best optimum conditions for consumption, the goods market, and the resources market are of the same form as the Pareto optimum conditions for consumption (5.11) and (5.12), for the goods market (5.13) and (5.14), and for the resources market (18.11 - 18.14). Even in the production sector the second best conditions for all firms, with the exception of the 1st and 2nd firms in their decisions concerning the 1st good, are of the same form as the Pareto conditions (18.15) and (18.16).

For the governmentally owned 1st firm the second best optimum conditions for the 1st good are:

$$(20.1) \qquad \lambda_1 - \mu_1 \frac{\partial g_{211}(y_{11}, y_{21})}{\partial y_{11}} - \mu_1 \frac{\partial g_{111}(y_{11}, y_{21})}{\partial y_{11}} - \sum_{j=2}^{s} \mu_j \frac{\partial g_{11j}(y_{11})}{\partial y_{11}} + \eta\mu_1 \frac{\partial^2 g_{211}(y_{11}, y_{21})}{\partial y_{21}\,\partial y_{11}} + \rho\, y_{21}\, \mu_1 \frac{\partial^2 g_{211}(y_{11}, y_{21})}{\partial y_{21}\,\partial y_{11}} \leq 0$$

$$(20.2) \qquad y_{11} \left[\lambda_1 - \mu_1 \frac{\partial g_{211}(y_{11}, y_{21})}{\partial y_{11}} - \mu_1 \frac{\partial g_{11}(y_{11}, y_{21})}{\partial y_{11}} - \sum_{j=2}^{s} \mu_j \frac{\partial g_{11j}(y_{11})}{\partial y_{11}} + \eta\mu_1 \frac{\partial^2 g_{211}(y_{11}, y_{21})}{\partial y_{21}\,\partial y_{11}} + \rho\, y_{21}\, \mu_1 \frac{\partial^2 g_{211}(y_{11}, y_{21})}{\partial y_{21}\,\partial y_{11}} \right] = 0 .$$

Although we have seen that [illegible] the 1st [illegible] the governmental firm (j=1) [illegible] in an individually rational (profit maximizing) manner since the set of best conditions were of the same form as the Pareto conditions (18.15) and (18.16), the above second best conditions (20.1) and (20.2) differ in form from the analogous Pareto conditions (18.17) and (18.18). Hence, if the governmentally owned 1st firm desires to maximize welfare, it must __not__ produce that quantity of the 1st good which will equate price and marginal social cost. It must ignore Pareto conditions (18.17) and (18.18) and try to choose, instead, some other quantity which satisfies the more complicated expressions (20.1) and (20.2)[1]. This result means, in effect, that the

[1]This result may still be contrasted with the popular interpretation of the Lipsey - Lancaster position, which is caused, perhaps, by some rather unfortunate wording in parts of their paper. For example, they state, "...given that one of the Paretian optimum conditions cannot be fulfilled, then an optimum can be achieved only by departing from all the other Paretian conditions." __Op. cit.__, p. 11.

governmentally owned 1st firm must try to "influence" the decision of the 2nd firm via the functional relationship of the externality. As was noted earlier, this influence is theoretically possible because

$$\frac{\partial g_{211}(y_{11}, y_{21})}{\partial y_{21}}$$ is a function of both y_{21} and y_{11}, and this term

is a part of the postulated behavioral rule of the 2nd producer. However, as we noted earlier in regard to the behavioral rule of the 2nd producer, there is little reason to suspect that second best conditions (20.1) and (20.2) could be achieved in the absence of communication and cooperation between the 1st and 2nd firms. The reason for this is that under non-separability

$$\frac{\partial g_{111}(y_{11}, y_{21})}{\partial y_{11}} \text{ is a function of both } y_{11} \text{ and } y_{21} \text{ also.}$$

Hence, this situation involves all the difficulties associated with mutual interaction so that it seems unlikely that a decentralized price regime could achieve this second best solution.[1]

---

[1]See Davis and Whinston, op. cit.

Finally, we note that the second best optimum conditions for the 2nd producer's choice of the 1st good, and the conditions which must be imposed upon the multipliers (η, ρ), are easily derived but are both lengthy and uninteresting. Hence, we omit them here and simply observe that, because of the formulation of the problem, the second best conditions for the 2nd firm's choice of output for the 1st good obviously differ from the Pareto optimum conditions (18.19) and (18.20).

Our main result thus far in Case III has been to show that for firms mutually affected by non-separable technological externalities, the violation of the Pareto conditions which refer to the good under consideration by one of these firms will lead to second best optimum conditions for the other firm affected by the externality which are of a different form from the Pareto optimum conditions. For all firms not affected by the externalities, and for the remaining sectors of the economy, the second best conditions are of the same form as the Pareto optimum conditions.

The above result naturally raises the following question: Suppose that the managers of the governmentally owned 1st firm failed to perceive that they might reduce social loss by communicating with the 2nd firm and specifying the decision rules for that good under consideration -- (20.1) and (20.2) for the 1st firm, and (19.1) and (19.2) for the 2nd firm -- which are being used.[1] Instead, assume that, whether or not communication takes place, the governmentally owned 1st firm uses Pareto conditions (18.17) and (18.18) as its decision rules. In other words, it attempts to choose outputs at which price equals marginal social cost. It is clear that this is a problem of second best which differs from the above one. This problem requires that in addition to constraints (19.1) and (19.2), constraints (18.17) and (18.18) must be attached to model (17). The 2nd firm attempts to act in an individually rational manner and the 1st firm attempts to satisfy the Pareto conditions. Since the addition of constraints to a maximization model never increases the value of the functional, we obtain the following result for this instance: Under the assumption of non-separable technological externalities which are mutual between two firms, if one firm acts in an

---

[1] Of course, it is obvious that in this situation welfare only achieves its absolute maximum when the two firms jointly maximize profits so that the Pareto optimum conditions are achieved. See Davis and Whinston, *op. cit.*

individually rational (profit maximizing) manner, then social welfare is lowered if the other firm is required to satisfy a Pareto condition rather then the more complicated second best conditions as represented in (20.1) and (20.2).

Both of our above results depended upon the very crucial assumption of non-separability. In order to demonstrate the importance of this assumption let us suppose that the externality terms $g_{111}(y_{11}, y_{21})$ and $g_{211}(y_{11}, y_{21})$ are separable. Then, under this assumption,

$\frac{\partial g_{111}(y_{11}, y_{21})}{\partial y_{11}}$ is a function of y_{11} only, and correspondingly,

$\frac{\partial g_{211}(y_{11}, y_{21})}{\partial y_{21}}$ is a function of y_{21} only. Hence, as a contrast to our previous results, neither of the two firms should experience unusual difficulty in following postulated behavioral rules. Furthermore, the cross derivative

$$(21) \qquad \frac{\partial^2 g_{211}(y_{11}, y_{21})}{\partial y_{21}\, \partial y_{11}} = 0$$

so that the terms in the second best conditions (20.1) and (20.2) which involve (21) must vanish. This means that (20.1) and (20.2) assume, in this instance, the same form as Pareto optimum conditions (18.17) and (18.18). Hence, if the mutual externalities are separable, all of the second best conditions are of the same form as the Pareto conditions except for the case of the single firm (the 2nd producer) which was initially assumed to violate a Pareto condition. Again, this result is opposed to the popular interpretation of the Lipsey - Lancaster theorem of second best.

For a final second best problem in this case, let us assume that the 1st firm also is privately owned. Thus suppose that both the 1st and the 2nd firms behave in an individually rational manner. For the 1st firm's decisions concerning the 1st good, this assumption means that it desires to use as decision rules

$$(22.1) \qquad \lambda_1 - \mu_1 \frac{\partial g_{111}(y_{11}, y_{21})}{\partial y_{11}} - \sum_{j=2}^{s} \mu_j \frac{\partial g_{11j}(y_{11})}{\partial y_{11}} \leq 0$$

$$(22.2) \qquad y_{11}\left(\lambda_1 - \mu_1 \frac{\partial g_{111}(y_{11}, y_{21})}{\partial y_{11}} - \sum_{j=2}^{s} \mu_j \frac{\partial g_{11j}(y_{11})}{\partial y_{11}}\right) = 0$$

This second best problem is formed by attaching as constraints (22.1) and (22.2) as well as (19.1) and (19.2) to the other constraints of model (17).

Note, however, that constraints (22.1) and (22.2), and (19.1) and (19.2) involve only the variables y_{11} and y_{21} so that it follows that the second best conditions in consumption are of the same form as Pareto optimum conditions (5.11) and (5.12), those for the goods market retain the form (5.13) and (5.14), those for the resources market remain of the form (18.11 - 18.14), and with the exception of the 1st and 2nd firms' production of the 1st good the second best conditions in production are of the same form as Pareto conditions (18.15) and (18.16). Hence, the second best optimum conditions diverge from the Pareto optimum conditions only for the 1st and 2nd firms in the production of the 1st good, and for these firms' production of this particular good a divergence was a part of the second best problem.

It seems unnecessary for us to present here the second best conditions for the last of the above situations since these conditions are long and complicated. However, it is worthy of note that, if communication and coordination between these two firms does not take place, and if the externalities are non-separable, these two firms are unlikely to achieve the second best solution. If coordination and communication do take place, then it is almost contradictory to even pose a second best problem since the logical thing for the two firms to do is to maximize joint profits and this will result in the achievement of the Pareto optimum in the context of the particular problem under consideration.[1]

---

[1] See Davis and Whinston, op. cit.

Case IV: Monopoly

So far our argument has been concerned with second best problems which are associated with externalities. Before generalizing our results, it seems desirable to treat the more traditional problem of monopoly. Actually, the question of obtaining efficiency in a decentralized system which is characterized by the existence of a monopoly is rather complex and raises issues which were not present in our above analysis.

Some of these issues can be avoided by recourse to additional assumptions. Designating the 1^{st} firm in the production of the 1^{st} good as the monopoly, we assume that

$$(26.1) \qquad \lambda_1 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{r1j}}{\partial y_{r1}} < 0 , \qquad r = 2, \ldots, z$$

for all possible values of y_{r1} so that $y_{r1} = 0$, $(r = 2, \ldots, z)$ always. In this situation there is no question concerning whether other firms "should" produce the 1st good. Correspondingly, we also assume that

$$(26.2) \qquad \lambda_1 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{11j}}{\partial y_{11}} = 0$$

is possible for some positive values of y_{11}. This avoids the issue of whether a good should be produced when it is only possible to produce it under conditions of monopoly.

However, the above considerations are not the fundamental ones here. Even the meaning of monopoly is not without ambiguity in the context of practical problems of second best. First, while the partial equilibrium meaning of monopoly is fairly clear -- or at least it is agreed that a monopolist faces a negatively sloped demand curve which is known to him and which allows him to manipulate his price by altering only the quantity of output of his own good -- the characterization of a monopoly in a general equilibrium system is not an entirely settled question. Since the demand curve has to be specified, one wonders whether to follow Hicks, Slutsky, or Friedman.[1] Or suppose one decided to derive individual demand curves

---

[1] See, e.g. Milton Friedman, "The Marshallian Demand Curve," Journal of Political Economy, LVII (December, 1949), pp. 413-95.

by an agreed upon method under which consumers maximized utility subject to a given income and price parameters which were varied, and then one summed individual demands to obtain market demands. Suppose that supply curves were also derived, and the demand-supply functions used to solve for the monopolist's price in terms of the other variables in the system. One might suspect that such a demand curve would be appropriate. However, one's suspicions would not be true here since a demand curve derived by such a method would be based upon the behavioral rules that consumers maximized individual utility and producers maximized individual profits; and the very purpose of second best problems is to determine normative behavioral rules for producers and consumers granted the presence of a deviant in the system. Hence, if the derived normative behavioral rules were different from those originally assumed in the derivation of the monopolist's demand curve, that curve would no longer be completely valid for the general equilibrium system.

In addition to the above question, there remains the difficult problem of constructing a model which will be useful to the governmental policy makers in the system. The policy makers cannot, of course, deductively derive the monopolist's demand curve since individual utility functions are unknown. Instead, the demand curve would have to be inductively estimated by econometric methods, and these methods never seem to find demand as a function of more than a few variables. Hence, if the policy makers could not determine demand as a function of more than a few variables, they could not possibly estimate the complicated interactions which, as we shall shortly see, the inclusion of additional variables causes in both Pareto and Second Best problems. But if the interactions cannot be estimated, then the officials would not be able to advise or make policy adjustments.[1]

---

[1]Of course, the monopolist himself might perceive his demand curve in the partial equilibrium sense of a simple relation between price and his own output.

Granted the above difficulties, our approach is to consider the monopoly problem as two cases. First, we shall assume that the government's econometricians could only derive demand as a function of one variable so that we have

$$(27) \qquad \lambda_1 = f(y_{11})$$

for this case, and in the next section we shall consider a more complicated case which will indicate how answers are changed as more interactions can be included.

With the above remarks in mind, we make the following observations: First, since it is assumed here that possible individual interactions on the monopolist's demand curve are too small to be estimated, the appropriate model for deriving the conditions for Pareto optimality from this practical standpoint is our basic model (4) so that the monopolist too will be considering price as a given parameter. Hence, we have (5.11) and (5.12) for the Pareto conditions in the consumption sector, (5.13) and (5.14) for the goods market, (5.15) and (5.16) for the production sector, and (5.17) and (5.18) for the factor markets.

Of course, an unconstrained monopolist must be assumed to take into account the fact that alterations in his output y_{11} affect his price λ_1 and hence his total revenue $f(y_{11})y_{11}$. Under the usual assumption that a monopolist equates marginal revenue and marginal cost, we have the behavioral rules:

$$(28.1) \qquad \frac{df}{dy_{11}} y_{11} + f(y_{11}) \leq \sum_{j=1}^{s} \mu_j \frac{\partial g_{11j}}{\partial y_{11}}$$

$$(28.2) \qquad y_{11} \left(\frac{df}{dy_{11}} y_{11} + f(y_{11}) - \sum_{j=1}^{s} \mu_j \frac{\partial g_{11j}}{\partial y_{11}} \right) = 0 .$$

We shall shortly formulate our second best problem, but since we have no interesting complications here we may write (28.1) and (28.2) in the more succinct form

$$(29) \qquad \phi \ (\mu_1, \ldots, \mu_s, y_{11}) \leq 0$$

in order to simplify the expression of the second best conditions and yet not obscure relevant parts of the argument. We now present the following second best problem:

$$(30.1) \qquad \max \sum_{i=1}^{m} \alpha_i u_i(X_i)$$

subject to

$$(30.2) \qquad \sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk} \quad , k = 1, \ldots, n$$

(30.3) $\sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj}(y_{rk}) \leq L_j \;, \quad j=1, \ldots, s$

(30)

(30.4) $\phi(\mu_1, \ldots, \mu_s, y_{11}) \leq 0$

(30.5) $\lambda_1 = f(y_{11})$

(30.6) $x_{ik}, y_{rk} \geq 0 \;, \quad i=1, \ldots, m; \; r=1, \ldots, z; \; k=1, \ldots, n \;.$

Observing that the added constraints contain only the variable y_{11}, we have immediately the following results: the second best optimum conditions in consumption are of the same form as the original Pareto conditions (5.11) and (5.12); those for the goods and resources markets are of the identical form as Pareto conditions (5.13) and (5.14), and (5.17) and (5.18) respectively; and with the exception of the 1st firm's production of the 1st good, (r=1, k=1), the second best conditions in production are of the same form as Pareto conditions (5.15) and (5.16). For the 1st firm in the production of the 1st good the second best conditions are

(31.1) $\lambda_1 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{11j}}{\partial y_{11}} - \eta \frac{\partial \phi}{\partial y_{11}} \leq 0$

(31.2) $y_{11} \left(\lambda_1 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{11j}}{\partial y_{11}} - \eta \frac{\partial \phi}{\partial y_{11}} \right) = 0 \;.$

Our obvious conclusion is that, if one Pareto condition is violated by a monopolist when the demand function can only be estimated in the indicated manner, all other Pareto conditions are still desirable and must be satisfied at this kind of second best solution.

Case V: Interdependent Demand:

In order to extend the above result, let us consider the following situation: Assume that the 1st firm is still the sole producer of the 1st good so that assumptions (26.1) and (26.2) are also applicable here. But assume that the 2nd good, which is produced under perfectly competitive conditions, is so "closely related" to the 1st good that total output of the second good directly affects the demand curve of the 1st good.[1]

[1] This might be the case, for example, if there were only one producer of light bulbs but electric lamps (and all other forms of lighting) were produced under competitive conditions. Hence, a market-clearing price of light bulbs would depend not only upon the quantity of light bulbs offered for sale, but also upon the number of electric lamps sold during the specified period of time.

Furthermore, we assume that this interaction is sufficiently great that the estimated demand curve will be of the form:

$$(32) \qquad \lambda_1 = h(y_{11}, \sum_{r=1}^{z} y_{r2})$$

and we desire to utilize this curve in our analysis of the situation.

It is obvious that in this situation Pareto optimality of this variety can be achieved only if these estimatable interactions are included in the model of the economic system. Furthermore, the 1st firm must behave in a perfectly competitive manner. Hence, we must impose the constraint

$$(33.1) \qquad h(y_{11}, \sum_{r=1}^{z} y_{r2}) \leq \sum_{j=1}^{s} \mu_j \frac{\partial g_{11j}(y_{11})}{\partial y_{11}}$$

$$(33.2) \qquad y_{11} \left[h(y_{11}, \sum_{r=1}^{z} y_{r2}) - \sum_{j=1}^{s} \mu_j \frac{\partial g_{11j}(y_{11})}{\partial y_{11}} \right] = 0$$

upon the 1st firm, and we assume that the situation will allow $y_{11} > 0$ to be possible so that side issues are avoided. In the interest of brevity and without loss of generality we may rewrite these constraints in the following form

$$(34) \qquad \Omega(\mu_1, \ldots, \mu_s, \lambda_1 y_{11}, \sum_{r=1}^{z} y_{r2}) \leq 0$$

and then state the following model of an economic system.

$$(35.1) \qquad \max \sum_{i=1}^{m} a_i u_i (X_i)$$

subject to

$$(35.2) \qquad \sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk}, \qquad k=1, \ldots, n$$

$$(35.3) \quad \sum_{k=1}^{n} \sum_{r=1}^{z} g_{rkj}(y_{rk}) \leq I_j , \qquad j=1, \ldots, s$$

(35)

$$(35.4) \quad \Omega(\mu_1, \ldots, \mu_s, y_{11}, \sum_{r=1}^{z} y_{r2}) \leq 0$$

$$(35.5) \quad \lambda_1 = h(y_{11}, \sum_{r=1}^{z} y_{r2})$$

$$(35.6) \quad x_{ik}, y_{rk} \geq 0 , \qquad \begin{matrix} i=1, \ldots, m \\ r=1, \ldots, z \\ k=1, \ldots, n \end{matrix}$$

Observing that our basic model (4) differs from the above model (35) only in the appearance of constraint (35.4) and (35.5) in the latter, we are able to state immediately the following results: Since for given prices the added constraints contain only variables y_{11} and y_{r2}, ($r=1, \ldots, z$), the Pareto optimum conditions for this model (35) are of the same form as the Pareto conditions for the former model (4) in the areas of the consumption sector, the goods markets, and the resources markets; and these are given by (5.11) and (5.12), (5.13) and (5.14), and (5.17) and (5.18) respectively. Even in the production sector if the 1st firm's production of the 1st good ($r=1$, $k=1$) and all firms' production of the 2nd good ($k=2$; $r=1, \ldots, z$) are omitted, then the Pareto optimum conditions here are of the same form as those previously stated as (5.15) and (5.16). For the 1st producer in the production of the 1st good the Pareto conditions are

$$(36.1) \quad \lambda_1 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{11j}(y_{11})}{\partial y_{11}} - \gamma \frac{\partial \Omega}{\partial y_{11}} \leq 0$$

$$(36.2) \quad y_{11} \left(\lambda_1 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{11j}}{\partial y_{11}} - \gamma \frac{\partial \Omega}{\partial y_{11}} \right) = 0$$

$$(36.3) \quad \Omega(\mu_1, \ldots, \mu_s, y_{11}, \sum_{r=1}^{z} y_{r2}) \leq 0$$

$$(36.4) \quad y_{11} \left[\Omega(\mu_1, \ldots, \mu_s, y_{11}, \sum_{r=1}^{z} y_{r2}) \right] = 0 .$$

Hence, the 1st firm must act in the manner prescribed by our formulation of the problem. The Pareto conditions for the production of the 2nd good are

$$(36.5) \quad \lambda_2 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{r2j}}{\partial y_{r2}} - \gamma \frac{\partial \Omega}{\partial y_{r2}} \leq 0 , \qquad r=1, \ldots, z$$

$$(36.6) \quad y_{r2} \left(\lambda_2 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{r2j}}{\partial y_{r2}} - \frac{\partial \Omega}{\partial y_{r2}} \right) = 0 \quad r=1, \ldots, z .$$

In other words, these Pareto conditions require that all firms producing the 2nd good take into account the effect which their production decisions have upon the 1st firm. Marginal social cost involves not only marginal private cost but also a marginal cost associated with the assumed and estimated interdependence.

It is, of course, obvious that a decentralized market mechanism may not achieve this Pareto optimal solution. First, if we rule out the possibility of the 1st firm bribing or otherwise influencing the producers of the 2nd good, then there is no reason why they should consider their effect upon the 1st producer when making their output decisions.[1]

[1] Of course, as we have pointed out elsewhere in a discussion of "natural units," there is always an incentive for one affected by an externality to try to influence the behavior of the other party. See Davis and Whinston, op. cit. This point is also made by Ronald Coose, "The Problem of Social Cost," Journal of Law and Economics, III (1960), pp. 1-44, and by James M. Buchanan and W. Craig Stubblebine, "Externality," Department of Economics, University of Virginia, 1962, (mimeographed).

More fundamentally, in the absence of some kind of externally imposed constraint, the 1st firm will not desire to equate price and marginal cost of the 1st good as we have stated in (33.1) and (33.2) and summarized in (34). Instead, it will desire to exploit its monopoly power and maximize profits by equating marginal revenue and marginal cost.

It is in light of the latter of the above points that we formulate our initial second best problem. Suppose that the 1st firm acts in an individually rational manner. Then, while it follows the usual decision rules for all other goods which it might produce (equate price and marginal cost), it must attempt to utilize the following behavioral rule in regard to the 1st good.

$$(37.1) \quad \frac{\partial h\left(y_{11}, \sum_{r=1}^{z} y_{r2}\right)}{\partial y_{11}} y_{11} + h\left(y_{11}, \sum_{r=1}^{z} y_{r2}\right) \leq \sum_{j=1}^{s} \mu_j \frac{\partial g_{11j}(y_{11})}{\partial y_{11}}$$

$$(37.2) \qquad y_{11} \left[\frac{\partial h(y_{11}, \sum_{r=1}^{z} y_{r2})}{\partial y_{11}} y_{11} + h(y_{11}, \sum_{r=1}^{z} y_{r2}) - \sum_{j=1}^{s} \mu_j \frac{\partial g_{11j}(y_{11})}{\partial y_{11}} \right] = 0$$

We now ask the following question: Granted that the 1st firm follows behavioral rules (37.1) and (37.2), what are the forms of the second best conditions?[1] Once again, without loss of generality we may rewrite

---

[1]Interestingly enough, it does not matter for our purposes whether or not we assume that the externality term $\sum_{r=1}^{z} y_{r2}$ enters the demand function (32) in a separable or non-separable manner since, because of the fact that price times quantity is total revenue, the expression for total revenue $h(y_{11}, \sum_{r=1}^{z} y_{r2})y_{11}$ must always be non-separable.

(37.1) and (37.2) in the succinct form

$$(38) \qquad \theta\,(\mu_1, \ldots, \mu_s, y_{11}, \sum_{r=1}^{z} y_{r2}) \leq 0$$

and formulate our second best problem by replacing constraint (35.4) in model (35) with the new constraint (38).

Observing that, for given prices, the new constraint (38) contains only the variables y_{11} and y_{r2} ($r=1, \ldots, z$), we immediately have the following result: With the exception of the conditions for the 1st and 2nd goods, the second best optimum conditions are of the same form as the original Pareto optimum conditions. For the 1st good the second best conditions are

$$(39.1) \qquad \lambda_1 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{11j}}{\partial y_{11}} - \eta \frac{\partial \theta}{\partial y_{11}} \leq 0$$

$$(39.2) \qquad y_{11} \left(\lambda_1 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{11j}}{\partial y_{11}} - \eta \frac{\partial \theta}{\partial y_{11}} \right) = 0$$

$$(39.3) \qquad \theta \left(\mu_1, \ldots, \mu_s, y_{11}, \sum_{r=1}^{z} y_{r2} \right) \leq 0$$

$$(39.4) \qquad y_{11} \left[\theta \left(\mu_1, \ldots, \mu_s, y_{11}, \sum_{r=1}^{s} y_{r2} \right) \right] = 0$$

These are not, for our purposes, very interesting. However, the second best conditions for the 2nd good are interesting. These are as follows:

$$(39.5) \qquad \lambda_2 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{r2j}}{\partial y_{r2}} - \eta \frac{\partial \theta}{\partial y_{r2}} \leq 0, \qquad r=1, \ldots, z$$

$$(39.6) \qquad y_{r2} \left(\lambda_2 - \sum_{j=1}^{s} \mu_j \frac{\partial g_{r2j}}{\partial y_{r2}} - \eta \frac{\partial \theta}{\partial y_{r2}} \right) = 0, \qquad r=1, \ldots, z.$$

In other words, even though the producer of the 1st good is a monopolist, the producers of the 2nd good must take into account the effects which their output decisions have upon him if the second best solution is to be attained.

Unless the monopolist somehow influences these producers, a decentralized price regime will not attain this second best solution.[1]

[1] Of course, in this particular instance it is theoretically possible for a governmental authority to impose a tax or subsidy (depending upon whether the externality is an external economy or diseconomy) so that the solution could be achieved.

Note that we could now formulate another second best problem by constraining the producers of the 2nd good to behave in an individually rational manner also. We shall not formally analyze this problem here since the following results are already clear: (i) The second best optimum conditions, with the exception of the 1st and 2nd goods, would be of the same form as the Pareto optimum conditions. (ii) In the absence of any successful attempt by the 1st producer to influence the decision of the producers of the 2nd good, this second best solution would be the one associated with a decentralized pricing mechanism which operated on the basis of the prescribed model.

Finally, we observe without further analysis that as more interaction variables are included in the estimated demand curve, then the corresponding entities must take these interactions into account for a Pareto solution. Furthermore, for a class of second best problems — for example, the one analyzed above — the second best conditions for those units whose decision variables enter into the monopolist's demand curve will differ, in general, from the Pareto conditions, but interaction effects will still be present in the derived behavioral rules.

6. Generalizations and Conclusions:

The fundamental proposition of modern welfare economics is that given independent preference orderings for consumers, independent technologies for producers, and certain conditions on the shapes of these functions, then, if consumers maximize utility subject to given income and price parameters, and if producers maximize profits subject to these price parameters, there exists a set of prices such that a social maximum is achieved where no individual can be made better off without making some other individual worse off.[1] Furthermore, granted these conditions and certain assumptions, this Pareto welfare maximum can be achieved via a pricing mechanism and decentralized decisions.

In this paper we have examined situations where at least one individual actor (producer or consumer) has a preference ordering, a criterion function, or a technology other than the one specified for the above theorem. In those cases examined we attempted to show that although a Pareto optimal point might exist, it would not be achieved by a decentralized pricing regime under the specified conditions. Then, granted

[1] See T. C. Koopmans, Three Essays on the State of Economic Science, New York: McGraw-Hill, 1957, pp. 41-104, for a statement and discussion of this proposition.

the behavior of the deviant member, we tried to study within the context of an economic system the strategies which other actors should play in order to compensate for the behavior of the deviant. Specifically, we attempted to determine whether their strategies should be altered if a deviant was introduced into the system.

In order to justify our method of analysis, to claarify further the meaning of a second best problem, and to generalize our results, let us consider the following series of models. First, however, we need to state the following definitions which supplement those presented earlier.

X_i^k = $(x_{i1}, \ldots, x_{ik-1}, x_{ik+1}, \ldots, x_{in})$, the "bundle" of goods consumed by the i^{th} individual with the exception of the k^{th} good.

$\hat{X}_i$ = $(X_i, x_{1l}, \ldots)$, a vector composed of X_i and all goods consumed by other consumers which influence directly the utility of the i^{th} consumer. Hence, the consumption of other consumers may or may not enter $\hat{X}_i$, according to whether consumption externalities are specified for a particular problem.

Y_r = $(y_{r1}, \ldots, y_{rn})$, a vector of the quantities of all goods produced by the r^{th} firm.

Y_r^k = $(y_{r1}, \ldots, y_{rk-1}, y_{rk+1}, \ldots, y_{rn})$, a vector of the quantities of all goods except the k^{th} which are produced by the r^{th} firm.

G_j = a function indicating the quantity of the j^{th} resource utilized by all firms in producing all goods.

λ = $(\lambda_1, \ldots, \lambda_n)$, a vector of prices (multipliers) associ-ated with goods.

μ = $(\mu_1, \ldots, \mu_s)$, a vector of prices (multipliers) associ-ated with resources.

B_{ik} = a function which represents the behavioral rule of the i^{th} consumer in choosing the k^{th} good.

D_{rk} = a function which represents the behavioral (decision) rule of the r^{th} firm in choosing the output level of the k^{th} good.

Granted these supplementing definitions, let us present the following model of an economic system:

(40.1) $\max \quad [u_1(\hat{X}_1), \ldots, u_m(\hat{X}_m)]$

subject to

(40.2) $\sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk}$, $\quad k=1, \ldots, n$

(40.3) $G_j (Y_1, \ldots, Y_z) \leq L_j$, $\quad j=1, \ldots, s$

(40)

(40.4) $x_{ik} = B_{ik}(X_1, \ldots, X_i^k, \ldots, X_m, Y_1, \ldots, Y_z, \lambda, \mu)$, $\quad i=1, \ldots, m$; $k=1, \ldots, n$

(40.5) $y_{rk} = D_{rk} (X_1, \ldots, X_m, Y_1, \ldots, Y_r^k, \ldots, Y_z, \lambda, \mu)$, $\quad r=1, \ldots, z$; $k=1, \ldots, n$

$$(40.6) \qquad x_{ik},\ y_{rk} \geq 0 , \qquad \begin{matrix} i=1, \ldots, m \\ r=1, \ldots, z \\ k=1, \ldots, n \end{matrix}$$

This model (40) is more general than those which we have considered. Specification of other consumers' consumptions which might enter $\hat{X}_i$ will allow externalities in consumption. Similarly, specification of the functional form of the G_j can allow or not allow externality in production. But the main difference in this and our previous models is the presence of constraints (40.4) and (40.5), the behavioral rules of consumers in their choice of goods and producers in their choice of outputs. Let us suppose for the moment that the forms of all B_{ik} (constraint 40.4) and D_{rk} (constraint 40.5) are to be chosen. We are searching, then, through whatever methods are available, for behavioral rules for consumers and producers which allow efficient solutions -- i.e., for which markets are cleared, resources are not over-demanded, and the resulting allocation is such that no one can be improved without worsening others.

In order to determine the functional forms for constraints (40.4) and (40.5) which will allow the desired class of solutions to be achieved, let us consider the following model:

$$(41.1) \qquad \max \quad [u_1(\hat{X}_1), \ldots, u_m(\hat{X}_m)]$$

subject to

$$(41.2) \qquad \sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk} , \qquad k=1, \ldots, n$$

(41)

$$(41.3) \qquad G_j(Y_1, \ldots, Y_z) \leq L_j , \qquad j=1, \ldots, s$$

$$(41.4) \qquad x_{ik}, y_{rk} \geq 0, \qquad \begin{matrix} i=1, \ldots, m \\ r=1, \ldots, z \\ k=1, \ldots, n \end{matrix}$$

Assuming that vectors $\hat{X}_i$, $(i=1, \ldots, m)$, have the same components, and that the functional forms of the G_j, $(j=1, \ldots, s)$, are identical, we observe that model (41) differs from the previous one (40) only in that constraints (40.4) and (40.5) are omitted in (41). Hence, the values of the solutions to problem (41) must form an upper bound to the values of the solutions to model (40) because of the fact that the addition of constraints to a maximization problem cannot increase the value of the solution. However, our present concern is to determine the form of constraints (40.4) and (40.5). It is obvious that the upper bound may be achieved if (40.4) and (40.5) are chosen to be redundant constraints since the problems are otherwise identical. In fact, if (40.4) and (40.5) are chosen to be the necessary conditions for a solution to model (41), then they must be redundant for model (40). Since solutions to (41) are Pareto optimal, the desired form of the behavioral rules B_{ik} and D_{rk} (constraint 40.4 and 40.5) are the conditions for Pareto optimality. Furthermore, if the $\hat{X}_i$ and G_j are specified such that no externalities exist, and if certain other conditions are satisfied,[1] then the necessary conditions to

---

[1]See Koopmans, op. cit., for a discussion of these conditions.

(41) may be associated with a perfectly competitive pricing mechanism which operates in a decentralized manner, and the multipliers λ and μ may be interpreted as market prices.

Having discussed models where the solutions are Pareto optimal, we consider now a second best problem. Suppose that we have as a model of our economic system:

(42.1) $\max \quad [\, u_1(\hat{X}_1), \ldots, u_m(\hat{X}_m) \,]$

subject to

(42.2) $\sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk}$, $\quad k=1, \ldots, n$

(42.3) $G_j(Y_1, \ldots, Y_z) \leq L_j$, $\quad j=1, \ldots, s$

(42.4) $x_{pq} = \bar{B}_{pq}(X_1, \ldots, X_p^q, \ldots, X_m, Y_1, \ldots, Y_z, \lambda, \mu)$, $\quad 0 \leq p \leq m$, $0 \leq q \leq n$

(42)

(42.5) $y_{e\ell} = \bar{D}_{e\ell}(X_1, \ldots, X_m, Y_1, \ldots, Y_e^{\ell}, Y_z, \lambda, \mu)$, $\quad 0 \leq e \leq z$, $0 \leq \ell \leq n$

(42.6) $x_{ik} = B_{ik}(X_1, \ldots, X_i^k, \ldots, X_m, Y_1, \ldots, Y_z, \lambda, \mu)$, $\quad i=1, \ldots, m$, $k=1, \ldots, n$

(42.7) $y_{rk} = D_{rk}(X_1, \ldots, X_m, Y_1, \ldots, Y_r^k, \ldots, Y_z, \lambda, \mu)$, $\quad r=1, \ldots, z$, $k=1, \ldots, n$

(42.8) $x_{ik}, y_{rk} \geq 0$, $\quad i=1, \ldots, m$, $r=1, \ldots, z$, $k=1, \ldots, n$

where constraints (42.4) and (42.5) are specified behavioral rules $\bar{B}_{pq}$ of the p^{th} consumer in choosing the q^{th} good, and $\bar{D}_{e\ell}$ of the e^{th} producer in choosing an output quantity of the ℓ^{th} good.

As was the case for problem (40), constraints (42.6) and (42.7) -- behavioral rules B_{ik} and D_{rk} -- are to be chosen. Also, we desire to choose those rules such that efficient solutions may be achieved. Since the addition of constraints can never increase the value of a solution, it is obvious that if (42.6) and (42.7) are chosen to be redundant, then an optimal and efficient solution will be achieved. Hence, let us consider the following problem:

(43.1) $\max \quad [\, u_1(\hat{X}_1), \ldots, u_m(\hat{X}_m) \,]$

subject to

(43.2) $\sum_{i=1}^{m} x_{ik} \leq \sum_{r=1}^{z} y_{rk}$, $\quad k=1, \ldots, n$

(43.3) $G_j\,(Y_1, \ldots, Y_z) \leq L_j$, $\quad j=1, \ldots, s$

(43)

(43.4) $x_{pq} = \bar{B}_{pq}\,(X_1, \ldots, X_p^q, \ldots, X_m, Y_1, \ldots, Y_z, \lambda, \mu)$, $\quad 0 \leq p \leq m,\; 0 \leq q \leq n$

(43.5) $y_{e\ell} = \bar{D}_{e\ell}\,(X_1, \ldots, X_m, Y_1, \ldots, Y_e^{\ell}, \ldots, Y_z \;\; \lambda, \mu)$, $\quad 0 \leq e \leq z,\; 0 \leq \ell \leq n$

(43.6) $x_{ik}, y_{rk} \geq 0$, $\quad i=1, \ldots, m;\; r=1, \ldots, z;\; k=1, \ldots, n$

Since problem (43) differs from (42) only in the absence of constraints (42.5) and (42.6), it is obvious that these constraints will be redundant to (42) when they are the necessary conditions to (43).

The above analysis provides justification for our formulation of second best problems. In a sense, we were trying to determine what the behavioral rules of the decision making units in the system "should be" if an efficient solution is to be achieved. Hence, we concentrated upon the derivation of optimality conditions for models similar to (41) and (43), and these conditions were interpreted as the desired, normative behavioral rules.

Furthermore, our formulation allows us to make precise the corollary to the theory of second best. Lipsey and Lancaster state this corollary as follows: "From this theorem [the Lipsey-Lancaster theorem of second best] there follows the important negative corollary that there is no *a priori* way to judge between various situations in which some of the Paretian optimum conditions are fulfilled while others are not."[1]

---

[1] Lipsey and Lancaster, *op. cit.*, pp. 11-12.

However, if one utilizes models of the type shown in (43) or at earlier points in this paper where second best problems were analyzed, then, since additional constraints cannot increase the value of the solution, the removal of a specified behavioral constraint will never worsen welfare, but can leave it unchanged or improve it.[2]

[2]In this sense one of the Lipsey-Lancaster statements of the negative corollary is unfortunately misleading. They state, "It follows, therefore, that in a situation in which there exist many constraints which prevent the fulfillment of the Paretian optimum conditions, the removal of any one constraint may affect welfare or efficiency either by raising it, by lowering it, or by leaving it unchanged.", op. cit. p. 12.

Finally, the approach to problems of second best which has been utilized in this paper allows one to compare always Pareto with second best optimum conditions. Furthermore, it is possible to state some general rules for determining which if any second best conditions will differ from the Pareto optimum conditions. Since the desired Pareto normative behavioral rules for (40) are given by the necessary conditions for (41), and since the normative second best behavioral rules for (42) are given by the necessary conditions for (43), then if we observe that (43) differs from (41) only in that certain behavioral rules -- constraints (43.4) and (43.5) -- have been specified, it is apparent that the problem really amounts to determining whether the additional constraints in (43) change the form of the necessary conditions to (41). Suppose, for example, that in (43.4) $p=1$, $q=1$ only, and in (43.5) $e=0$, $l=0$ always so that model (43) differs from (41) only in the addition of one behavioral rule which has been specified as a constraint. Obviously, then, the forms of the necessary conditions to (43) will differ from those to (41) only for those conditions for which the corresponding variable appears in the additional constraint. Of course, this result can be generalized immediately so that no matter how many additional constraints have been specified, the form of the necessary conditions to (43) will differ from the form of the conditions to (41) only for those variables which appear in the additional constraints

As we have formulated second best problems, the additional constraints involve variables whose values are determined by the choices of decision makers in the system. When only one additional constraint is added, and when the added constraint contains only variables subject to the choice of the particular decision-making unit under consideration (the deviant from the Pareto behavioral rules), then only one Pareto condition is violated -- the one initially assumed to be violated -- at a second best optimum. On the other hand, if the specified behavioral constraint of the deviant unit contains variables whose values depend directly upon the choice of other units or upon the deviant in other areas of operation, then, as we have seen for the cases of mutual externalities and interdependent demand, the situation is different. Those units whose choice affect directly the values of the variables included in the specified behavioral rule of the deviant can no longer follow a Pareto behavioral rule if the second best solution is to be attained.

It is the latter of the above situations which constitute the difficulty in "piecemeal policy recommendations." It is only when units are inter-connected functionally -- either by preference orderings, technology, or demand relations -- is it true that requiring a subgroup of this inter-connected entity to satisfy Pareto conditions, while the remainder violates these conditions, may result in a decrease in social welfare. As far as a policy is concerned, inter-connected units must be considered as an entity or, as we have put it elsewhere, a natural unit[1] if

[1]Davis and Whinston, op. cit.

a welfare maximum is to be achieved. Hence, if a policy maker is considering whether one situation is socially preferable to another, the question can be more easily answered if one of the situations includes all of the "constraint variables" of the other one. In such an instance, at least one may make certain inequality statements about social welfare even in the absence of particular knowledge of utility or production functions.

APPENDIX

In the course of the discussion in this paper we encountered the following type of maximization problem:

I. Max $F(X)$

$$\text{s.t.} \quad G_j(X\lambda^\rho) \leq 0 \qquad j = 1 \ldots m$$

$$x_i \geq 0$$

$$X = (x_1 \ldots, x_n) \qquad \lambda^\rho = (\lambda^\rho_\ell \ldots \lambda^\rho_k)$$

$$0 < \ell < k \leq m$$

λ^ρ may be considered as a vector of parameters. Assuming that F is concaved and G_j are convex functions of X for each λ^ρ we may by the Kuhn Tucker theorem[1] consider the saddle value problem

[1] Theorem 3, Non Linear Programming, Second Berkeley Symposium, Berkeley, 1951.

II. $$\phi(X\lambda) = F(X) - \sum_{j=1}^{m} \lambda_j \, G_j \, (X\lambda^\rho)$$

A solution $(X^o \lambda^o)$ to the saddle value problem and thus X^o a solution to the maximization problem I will be a function of the values taken by λ^ρ. We may express this by writing $X^o = X^o(\lambda^\rho)$ $\lambda^o = \lambda^o(\lambda^\rho)$ which is a short form for

$$\text{III(a)} \qquad x_i^o = x_i^o(\lambda_\ell^\rho \dots \lambda_k^\rho) \qquad i=1, \dots, n$$

$$\text{III(b)} \qquad \lambda_j^o = \lambda_j^o(\lambda_\ell^\rho \dots \lambda_k^\rho) \qquad j=1, \dots, m$$

In order to obtain particular values $(X^o \lambda^o)$ as a solution we must in some manner specify the values of the parameters $(\lambda^\rho \dots \lambda_k^\rho)$. For the models in this paper we have added the following conditions: If for λ_i there is λ_j^ρ such that j=i then we add the condition for each such i

$$\lambda_j^\rho = \lambda_i$$

In order to help clarify the above discussion we present the following example:

Minimize $x^2 + y^2 + z^2$

Subject to (1) $x + y + z = 1$

(2) $3x + 3y - \lambda^\rho = 0$

Letting λ be the Lagrangean multiplier for (1) and μ for (2) we obtain

$$\begin{aligned} 2x - \lambda - 3\mu &= 0 \\ 2y - \lambda - 3\mu &= 0 \\ 2z - \lambda &= 0 \\ x + y + z &= 1 \\ 3x + 3y - \lambda^\rho &= 0 \end{aligned}$$

Adding the additional condition $\lambda = \lambda^\rho$ and solving the entire system we obtain

$$(\bar{x}\ \bar{y}\ \bar{z}\ \bar{\mu}\ \bar{\lambda}\ \bar{\lambda}^\rho) = (\tfrac{1}{5}, \tfrac{1}{5}, \tfrac{3}{5}, -\tfrac{4}{15}, \tfrac{6}{5}, \tfrac{6}{5})$$

and a minimum of $\frac{11}{25}$. This may be contrasted with a solution to the same example excluding constraint (2) of $(\bar{x}\ \bar{y}\ \bar{z}\ \bar{\lambda}) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{2}{3})$ and a minimum of $\frac{1}{3}$ which is less than $\frac{11}{25}$.