2

ARTIFICIAL MARKETS AND THE THEORY OF GAMES*

W. David Montgomery
California Institute of Technology

The concept of transaction costs is a common theme in most analyses of the phenomenon of market failure. Few economists would disagree with the abstract proposition that if there exist gains to be made from exchange, then in the absence of transaction costs private bargains will take place and exhaust all potential gain from trade. This proposition serves not only as a characterization of an ideal state of affairs, but as a guide to means by which specific cases of market failure could be remedied. It suggests, in particular, that reduction in transaction costs should be examined as a potential remendy. Since the nature and extent of transaction costs depend crucially on the institutional structure in which private bargains take place, the analysis of the relationship between institutions and transaction costs becomes a primary concern of policy for dealing with market failure.

The theory of games has provided notable insights into the natue of bargaining processes. In this article I will apply co-operative game theory to a specific problem of air pollution control, as a device for designing and evaluating a set of institutions intended to eliminate certain transaction costs which appear to prevent profitable bargains from being consummated.

In analyzing the relationship between institutions and transaction costs it is helpful to break transaction costs down into three broad categories: tangible resource costs, social constraints, and strategic breakdowns. Some feeling for the coverage of each of these categories can be conveyed through a fanciful example.

*This article draws on material included in my doctoral dissertation, "Market Systems for the Control of Air Pollution," submitted to the Department of Economics at Harvard University. A lengthier version was presented at the Annual Meeting of the Public Choice Society in May 1972. I am indebted to Kenneth Arrow, Water Isard, James Krier and Charles Plott for suggestions and criticisms, and to the Environmental Quality Laboratory of the Calfornia Institute of Technology for research support.

Consider the case of electrical suppliers who would like to get together to fix prices. Certain private bargains can lead to gains for all participants—but three separate influences work against their achievement:

(1) the cost of airplane tickets, hotel accommodations, martinis, and the big mahogany table needed to bring the participants physically together;

(2) the probability that some outsiders will revoke the agreement and punish the participants, and

(3) the lack of a compelling equilibrium strategy for each participant which will take the bargaining process to a stable outcome which exhausts gains from trade.

The proposition that private bargains will take place and exhaust gains from trade when transaction costs are zero is true only if we include all three elements in our definition of transaction costs. If we exclude, for example, strategic breakdown from consideration as a transaction cost, we will find that some profitable trades do not occur because of bargaining deadlock.

Tangible resource costs can be dealt with in a straightforward manner in each specific case. If, for example, it is less expensive for a polluter to discover who is affected by his pollution and initiate negotiations than vice versa, then assigning rights to sufferers would reduce transaction costs below what they would be if rights were assigned to polluters. Indeed, the introduction of well-defined rights always tends to reduce tangible resource costs, by eliminating the need for preliminary negotiations to determine who has the right to do what. In the analysis which follows it will be assumed that well-defined rights are granted, thus eliminating one substantial source of transaction costs.

The second element in transaction costs, social constraints, can be identified with bargains, or aspects of bargains, which are prohibited by society, possibly because they violate some ethical or political norm. In the game-theoretic analysis these elements will be identified with constraints on choice sets for coalitions of various participants. Such constraints exist even in general formulations of market games. In a market exchange game, for example, each coalition is constrained to allocate commodities among its members in a manner which makes the total of each commodity in the final allocation equal to the total in the initial endowment. This amounts to the constraint that there be no "stealing" from some other coalition—a constraint which will be in force only if some specific institutions exist.

The third element in transaction cost, strategic breakdown, has been related to institutional design in several papers. The prisoner's dilemma is perhaps the most common of the cases in which rational strategies fail to exhaust gains from trade. In the prisoner's dilemma each player has a dominant strategy and, when each player chooses his dominant strategy, the outcome is inferior from the point of view of both. It has been shown that in particular examples of the prisoner's dilemma game the introduction of certain liability rules makes the solution in dominant strategies an optimal one.[1]

Another type of strategic breakdown, particularly relevant to externalities, results from "hold-out" strategies. If a polluter, for example, must obtain agreement from every person in a geographical area in order to operate, a reasonable strategy for each person would be to demand almost all the available gains from trade, since he holds a veto power.[2] If each person pursues such a strategy, no agrement is likely.

[1] C. Plott and R. Meyer, "Characterization of Public Goods, Externalities, and Exclusion," (unpublished).

[2] K. J. Arrow, "The Organization of Economic Activity: Issues Pertinent to the Choice of Market Versus Nonmarket Allocation," in U.S. Congress Joint Economic Committee, *The Analysis and Evaluation of Public Expenditures: The PPB System*, Vol. I, U.S. Government Printing Office, Washington, 1969.

There is an important difference in terms of impact on tangible resource costs and strategic breakdown between institutions which give rise to competitive markets and those which do not. There is a clear practical difference between going into a supermarket to buy a bottle of milk and hiring a lawyer to negotiate a settlement to a traffic accident. The ability of competitive markets and prices to reduce tangible resource costs of information gathering and processing is great. More importantly, in competitive markets with large numbers of participants the strategies which lead to Pareto optimality are Nash equilibrium strategies, and it can be expected that bargains will be struck.

The institution on which the paper focuses is the market in rights to pollute. This institution is constructed on an analogy to such arrangements as the market in taxi medallions in New York City. This market seems to be a conspicuously successful method of allowing a limited number of operators to exploit gains from trade. Any New Yorker will argue that the market is harmful to taxi-riders, and because of this fact it is desirable to evaluate such a market with certain social constraints on private bargains in mind. In particular, the number of licenses issued must reflect these considerations.

I. *The Failure of Markets in Externalities*

Externalities are commonly cited as a cause of market failure. It is instructive to ask if, in the absence of transaction costs broadly defined, markets could mediate external effects. Arrow has answered the question by extending the coverage of the general equilibrium theory of competitive markets.[3] He argues that when externalities are present we can define new commodities, each of which is indexed as an externality of type i, produced by actor j and suffered by actor h. When the commodity space is expanded in this manner, the competitive equilibrium of the expanded market economy is Pareto optimal.

Transaction costs can be cited as the reason why such expanded markets do not exist. In many cases, tangible resource costs exceed all conceivable gains from trade on such markets under all institutions, and the failures are not worth remedying through private bargains. In such cases non-market allocation mechanisms turn out to offer the only possibility of achieving some of the potential gains from trade, although even these may involve such costs as to be undesirable. The case which will be examined in this paper is that in which tangible resource costs are negligible, but in which strategic perversities or social constraints have prevented the achievement of mutually advantageous bargains. Arrow argues that strategic breakdown will be a common phenomenon, since many of the markets in externalities will have only two participants, by definition.[4] In proving the Pareto optimality of a market system it is assumed that every participant responds as a price-taker, and that some mechanism varies prices to clear all markets. If we consider each market in an externality as an isolated system, with but two participants, it becomes very unlikely that the price-taker strategy is individually rational.

The standard analysis of the two-person market as a game leads to the conclusion that the competitive equilibrium is not a Nash equilibrium with only two players, since if one participant is a price-taker the best strategy for the other player is to become a price-maker and maximize his utility, subject to the other's offer curve. This situation is illustrated in Figure 1. As the number of players increases, however, and as long as all are about the same size, the competitive equilibrium comes closer to the Nash equilibrium.[5] Therefore, we may find that

[3] *Ibid.*

[4] *Ibid.*.

[5] Lloyd Shapley and Martin Shubik, "Concepts and Theories of Pure Competition," in M. Shubik, ed., *Essays in Mathematical Economics: In Honor of Oskar Morgenstern*, Princeton, 1967.

markets fail to mediate externalities because individual strategy choices fail to lead to any agreement on the bargain to be struck. If a market in licenses is to exhaust gains from trade it must be constructed as a competitive market, having many buyers and sellers. The possibility of constructing such a market exists because of the relationship of substitutability between many externalities.

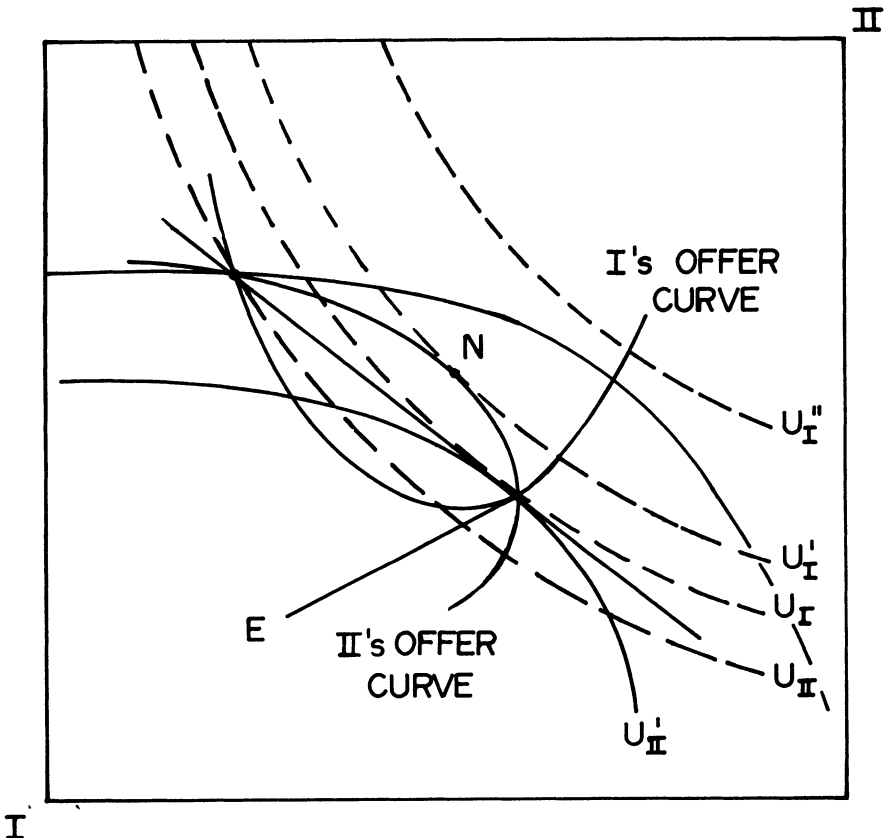

FIGURE 1

Let us consider that specific market failure which results in the phenomenon of air pollution. We could define a market by calling the sulfur dioxide produced by power plant A and suffered by Mr. Jones one commodity, and the sulfur dioxide produced by power plant B and also suffered by Mr. Jones another commodity. Each commodity gives rise to a one-to-one market, and there is no reason to expect each participant to behave as he would in a supermarket. Even if the right to the commodity is clearly defined, the question of whether a lawsuit, a guerilla action, or a transfer of money will be the form which the bargain takes is unresolved. Through the introduction of appropriately defined rights, however, a competitive market can be induced.

In this case, the two commodities would normally be perfect substitutes for Mr. Jones. This offers the hope that a market in a more broadly defined commodity, sulfur oxide, will have a competive equilibrium. Such a market would not only reduce tangible resource costs, but also replace indeterminate private bargains, of various forms, with predictable market behavior.

II. *Rights to Pollute*

Suppose that some quantity of rights to use the waste disposal capacity of the atmosphere is defined. It could be, for example, that some regulatory agency sets emission standards. Even subsequent to this attempt to deal with a market failure, gains from trade may still exist in two dimensions. Polluters and sufferers might achieve mutual benefit from changing the quantity of rights held by the polluters, and polluters as a separate group might find mutual benefit in rearranging rights among themselves. In this article, a game-theoretic model will be constructed and used to design a system for exploiting gains from trade among polluters as a separate group. Because a reduction in air pollution is a public good, it is not yet clear that a market system can be used to exploit gains from trade between polluters and sufferers.

It is assumed that emissions of air pollutants can be represented as a vector $E = (e_1, \ldots, e_n)$ where e_i is the average emission rate of firm i. Air quality is represented as a vector of average atmospheric concentrations measured at m points in a region. That is, air quality is a vector $Q = (q_1, \ldots, q_m)$ where q_j is air quality at location j. Rights are granted in order to achieve a desired air quality Q^*.

Each firm is assumed to maximize profits. To avoid unnecessary detail, it is assumed that prices for all inputs and outputs of firms in the system are independent of their activity levels. Under these conditions, we can define cost functions $F_i(e_i)$, which represent the difference for firm i between profits with emissions at some profit-maximizing level $\bar{e}_i$ and profits with emission rate e_i.[6]

It is further assumed that emissions are mapped into air quality by a meteorological diffusion matrix H, so that $E \cdot H = Q$. The goal of the game is to bring about adoption of an efficient emission vector E^{**} which minimizes $\sum_i F_i(e_i)$ subject to $EH \leqslant Q^*$. It is assumed throughout that $F_i(e_i)$ is twice differentiable and bounded from below.

It is assumed that some initial allocation $A = (a_1, \ldots, a_n)$ of rights to use the atmosphere is made. The operational significance to the firm of these rights comes down to permission to emit contaminants into the atmosphere at a certain rate. That is, they come down to constraints of the form $e_i \leqslant a_i$ on the permissible emission levels. Therefore, to any initial allocation of rights, there corresponds an initial emission vector which will be adopted if no voluntary rearrangement of rights is permitted. Designating the initial emission vector $E^o = (e^o_1, \ldots, e^o_n)$, firms operating under constraints $e_i \leqslant a_i$ will choose emission levels $e^o_i = \min(\bar{e}_i, a_i)$. Conversely, to any emission vector there corresponds some initial allocation of rights (to individual firms) which will lead to its adoption.

If the initial allocation of rights gives rise to an emission vector $E^o \neq E^{**}$, then some other allocation will increase joint profits and still result in adequate air quality. It will be shown that it is in fact possible to define rights to use the atmosphere and rules regarding their transfer such that voluntary private bargains among firms can lead to the adoption of the efficient emission vector E^{**}.

A simple example is useful to help clarify the issues involved. Consider a region containing two firms and two air quality monitoring points, and in which cost functions are convex and independent. It is possible to show the optimum graphically for this case. In Figure 2, the shaded area corresponds to the set of emission vectors which produce at least desired air quality. The curves c-c and c'-c' are isocost curves. They are contours of function $\sum_i F_i(e_i)$, representing those

[6]Proved in W.D. Montgomery, "Markets in Licenses and Efficient Pollution Control Programs," *Journal of Economic Theory*, 5, 1972.

combinations of e_1 and e_2 which give rise to the same value of $\sum_i F(e_i)$. The further one goes from the origin, the less cost will be. The shape of the curves follows from convexity. The point (e_1^{**}, e_2^{**}) at which an isocost curve is tangent to the boundary of the constraint set is the efficient emission vector. Note that only one constraint is binding at this point: air pollution is less than the level required by the other constraint.

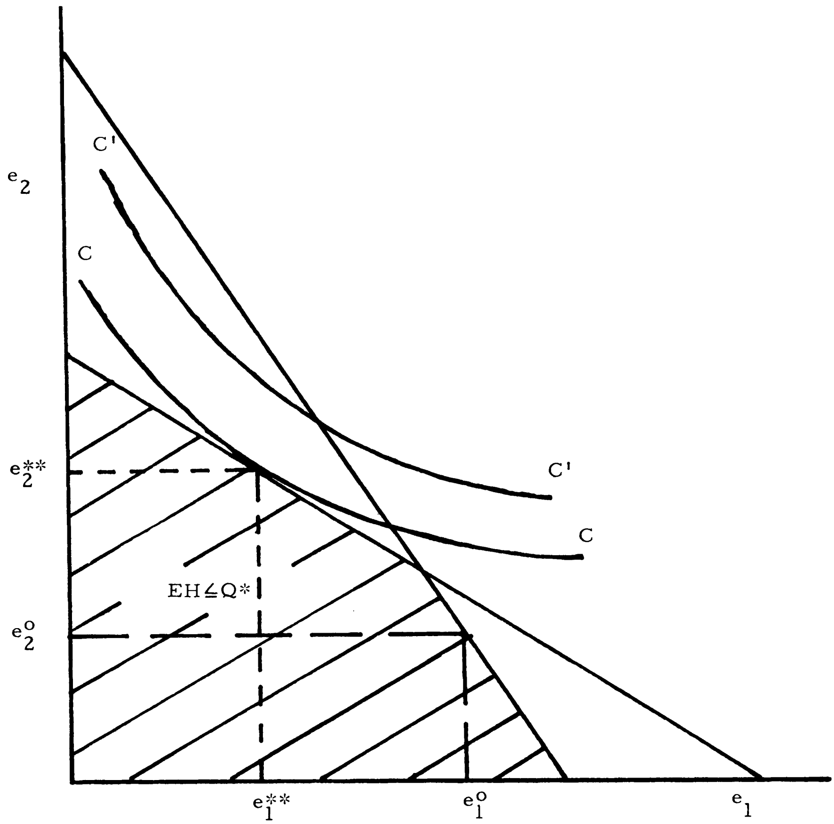

FIGURE 2

The corner of the constraint set is the only point at which both constraints are satisfied. Because of the shape of the isocost curves, the air quality vector associated with this point entails a higher joint total cost than does the air quality vector associated with E**, which gives better air quality at one point.

With the help of this example, some of the problems of allocation and exchange of rights can be isolated. Without knowledge of cost functions, it is impossible to solve for E**. Therefore, it is not likely that the initial allocation of rights will lead to E**. Suppose that the initial allocation is such that E^o (in Figure 2) is adopted. It is clear that the total cost of E^o is higher than the total cost of E**, and also that an allocation of rights which supports E** would be preferred by both parties if firm 1 (for which $e_1^{**} < e_1^o$) were offered adequate compensation by firm 2 (for which $e_2^{**} > e_1^o$). Since $F_1(e_1^{**}) + F_2(e_2^{**}) < F_1(e_1^o) + F_2(e_1^o)$

by hypothesis, it follows that firm 2 can offer such a payment (making firm 1 better off than under the initial allocation) and remain better off itself.

The problem is to see how a voluntary rearrangement of rights can bring about a change from the allocation supporting E^o to that supporting E^{**}

Rights can be made transferable in numerous ways, not all of which have the desirable properties normally associated with the competitive equilibrium of a market system. A number of such ways of defining and transferring rights will be mentioned, to provide a contrast to the system which makes possible market-like behavior.

The simple rule that any reduction in emission by firm 2 leaves firm 1 free to increase its emission by exactly the same amount is not usable. Suppose one unit of emission by firm 1 contributes twice as much to pollution at point j as does one unit of emission by firm 2. Then, let rights to emit be allocated to the two firms in such a way that if each emits at the maximum level allowed by his rights, air quality at point j is just adequate. If firm 2 sells an emission right to firm 1, and if the rights are undifferentiated, air pollution at point j will increase to an unsatisfactory level.

It is only through some such rearrangement of rights that an efficient emission vector can be reached; but it is equally clear that in order to achieve desired air quality, some restrictions must be put on permissible rearrangements. Let the restriction be that only those contracts are permitted which reallocate rights in such a way that any emission vector chosen will satisfy the air quality constraints as a whole. In this example, this leads to imposing on the joint choice of emission levels (established in the contract) the constraints

$$h_{11}e_1 + h_{12}e_2 \leq q_1^*$$

$$h_{12}e_1 + h_{22}e_2 \leq q_2^*$$

Then, adequate air quality will be preserved and the firms will be able to arrive at a mutually beneficial arrangement which results in the adoption of the emission vector E^{**}. Moreover, their situation has a very natural representation as a non-constant sum game.

In general we assume that there is a set of n firms in the region, denoted I = (1, . . . ,n). An "air quality management system" is a set of rules for an n-person game to be played by the polluting firms in the region. The management system first imposes an emissions constraint on each firm. In the absence of any binding agreement with others, this emission rate must not be exceeded. Second, the management system defines acceptable behavior for each and every group of firms. This may involve specifying exact emissions for each member, but in general it will involve making a requirement that certain inequalities must be satisfied by the emission vector adopted through the group decision process. It will be shown that the process of reaching such a decision can be interpreted as rearranging the rights granted to individual firms.

To formalize the concept of a management system consider the set of integers I = (1, . . . ,n) which represents the set of firms. The power set of this set is the set of all its subsets. The power set is denoted by $\bar{I}$, where $\bar{I}$ =((1), (2) , . . . ,(n), (1,2), (1,3) , . . . , (1,2, . . . ,n)). This set has 2n - 1 elements. A management system associates with each coalition a set of inequalities which must be satisfied by any emission vector adopted by that group for its members. We call this set of inequalities the "choice constraints for coalition S." Thus, in the example where 1 = (1,2), a management system is represented by the mapping

$$\{1\} \longrightarrow e_1 \leq e_1^o$$

$$\{2\} \longrightarrow e_2 \leq e_2^o$$

$$\{1,2\} \longrightarrow h_{11}e_1 + h_{21}e_2 \leq q_1^*$$
$$h_{12}e_1 + h_{22}e_2 \leq q_2^*.$$

It is convenient to be able to refer to the "coalition choice set," which is the set of emission vectors satisfying the coalition choice constraints.

Let E_s be a vector having elements whose subscripts run over the set $S \in \bar{I}$. E_s is the projection of some n-dimensional vector E on the subspace of dimension $\bar{S}$. Now define $\Omega(S)$ as the set of vectors E_s which satisfy the choice constraints for coalition S. An alternative way of defining an air quality management system is to define it as the set-valued function $\Omega(S)$. Each coalition S will be allowed to minimize $\sum_{i \in S} F_i(e_i)$ subject to the constraint $E_s \in \Omega(S)$.

The establishment of an initial allocation of rights and the prohibition of rearrangements of these rights is a trivial management system. The prohibition can be justified by the fact that some rearrangements of rights can result in a degradation of air quality. By setting proper coalition choice constaints, however, it is possible to ensure that desired air quality is achieved. Indeed, we make this a requirement on coalition choice constraints: they must be such that no matter what the coalition structure may be, any permitted choices of emission rates will produce air quality at least as good as Q*. This is not the only requirement, since it is fulfilled even by the system which prohibits the formation of coalitions. The second requirement is that some coalition or coalitions must be in a position to make choices resulting in the adoption of E**, the efficient emissions vector.

A very simple management system having these two properties is one in which each firm is assigned as initial allocation of rights $A = (a_1, \ldots, a_n) > 0$ giving rise to an initial emission vector $E^o = (e^o_1, \ldots, e^o_n) > 0$. Then, let the management strategy be as follows: if the number of firms in a coalition is less than n, each firm must adopt the same emission which is permitted by the initial allocation of rights. For the coalition of the whole, coalition choice constraints are the air quality constraints. Formally,

$$\text{For all } S \in \{\bar{I} - I\}$$
$$e_i \leq a_i$$

and for $S = I$, $E \cdot H \leqslant Q^*$

This case could also be generated by saying that formation of any coalition except the coalition of the whole is prohibited. We allow reallocation of rights only by *unanimous consent*, and then allow any reallocation producing adequate air quality. This system always has a nonempty core. It is, however, a system very prone to deadlock in bargaining. Each participant is in a position to demand almost all the gains from trade, because of the requirement of unanimous consent.

It is also possible to define coalition choice constraints which give the management system the form of a market game. The essential feature of these choice constraints is that they make the emissions allowed to any coalition a linear combination, identical for all coalitions, of the rights held initially by its members. The coalition choice constraints have a natural interpretation as requiring that the initial allocation of rights be reallocated within a coalition in a way which preserves their totals. There are two modal types of grant: a right to emit and a right to pollute. The right to emit will be a scalar, stating the highest average rate of emission which is allowed to a firm. The right to pollute is not so simple, since the firm directly controls its emissions, not pollution. The firm must adjust its emissions so that it does not cause greater pollution at any point than it has the right to cause. The right to pollute will have to be specified for each location at which air quality is measured. Let q_{ij} be the pollution caused at point j by firm i, so that $h_{ij}e_i = q_{ij}$. Then if q_{ij}^* is the amount of pollution at point j which firm i has a right to cause, the initial allocation of rights will have the form of a set of inequalities $h_{ij}e_i \leqslant q_{ij}^*$, where j = 1, . . . ,m, or of the single inequality

$$e_i \leq \min_j \frac{q_{ij}^*}{h_{ij}} .$$

This allocation has the same abstract form as the allocation of emission rights where $e_i \leqslant a_i$ is the legal restriction.

We will develop a rule, applying to coalitions of any size, for forming choice constraints as functions of the initial allocation. The suggested rule is that the pollution resulting from any emission policy adopted by a coalition must not exceed, at any point, the pollution which would be caused by firms in the coalition if each exercised fully the rights granted in the initial allocation. If emission rights are granted, the coalition choice constraints are $\sum_{i \epsilon S} h_{ij}e_i \leqslant \sum_{i \epsilon S} h_{ij}a_i$. If pollution rights are granted, the coalition choice constraints are $\sum_{i \epsilon S} h_{ij}e_i \leqslant \sum_{i \epsilon S} q_{ij}^*$. We require in addition that the initial allocation satisfy $\sum_i q_{ij}^* = q_j^*$ for pollution rights and $\sum_i h_{ij}a_i < q_j^*$ for emission rights. We can verify the first property for either mode of allocating rights by adopting the neutral symbol $c_{ij} = q_{ij}^* = h_{ij}a_i$ and the convention $\sum_i c_{ij} < q_j^*$.

Denote a partition of the set I as Π. For any partition Π we know from the coalition choice constraints that

$$\sum_{i \epsilon S} h_{ij}e_i \leq \sum_{i \epsilon S} c_{ij} \tag{1.1}$$

for all $S \epsilon \Pi$ But since

$$\sum_{S \epsilon \Pi} \sum_{i \epsilon S} h_{ij}e_i = \sum_{i=1}^{n} h_{ij}e_i$$

and

$$\sum_{S \epsilon \Pi} \sum_{i \epsilon S} c_{ij} = \sum_{i=1}^{n} c_{ij} \leq q_j^* ,$$

we conclude that for any partition, which corresponds to a coalition structure,

$$\sum_i h_{ij}e_i \leq q_j^* . \tag{1.2}$$

This establishes that no coalition structure can give rise to the choice of an emission vector which does not satisfy the air quality constraints. In order to find out whether the coalition choice constraints will allow the achievement of efficiency, we must examine how a coalition will choose a joint strategy. Any coalition can choose the emission vector in its choice set which minimizes the sum of its members' costs. Let us consider the coalition of the whole. Its choice constraints are $\sum_i h_{ij}e_i \leq \sum_i c_{ij}$. Now we must distinguish again between emission and pollution rights. If pollution rights are granted, then $\sum_i c_{ij} = \sum_i q^*_{ij} = q^*_j$ for all j. Therefore the constraints for the coalition of the whole coincide with the air quality constraints. In this case it is possible that a private bargain should be struck among the n firms such that the efficient emission vector is adopted.

If, on the other hand, emission rights are granted, so that $c_{ij} = h_{ij}a_i$, the choice constraints will be $\sum_i h_{ij}e_i \leq \sum_i h_{ij} a_i$. If we require only that $\sum_i h_{ij}a_i \leq q^*_j$, the choice constraints for the coalition of the whole will not necessarily coincide with the air quality constraints. For some allocations we can find that $\sum_i h_{ij}a_i < q^*_j$ for some j. In the two firm example (e^o_1, e^o_2) is such an allocation. In this example the choice constraints for the coalition of the whole are

$$h_{11}e_1 + h_{21}e_2 \leq h_{11}e^o_1 + h_{21}e^o_2 \tag{2.1}$$
$$h_{12}e_1 + h_{22}e_2 \leq h_{12}e^o_1 + h_{22}e^o_2.$$

We know that $h_{11}e^{**}_1 + h_{21}e^{**}_2 > h_{11}e^o_1 + h_{21}e^o_2$ since in Figure 2 the point (e^{**}_1, e^{**}_2) lies above a line with slope equal to $-h_{11}/h_{21}$ passing through (e^o_1, e^o_2). Therefore, if constraint 2.1 is imposed on the coalition of the whole it will be unable to adopt the efficient emission vector. We conclude that choice constraints of the form $\sum_{i \in S} h_{ij}e_i \leq \sum_{i \in S} h_{ij}a_i$ allow the adoption of an efficient emission vector only if a_i, the initial allocation, is chosen to satisfy the *equations* $\sum_i h_{ij}a_i = q^*_j$. This condition is, however, very strong, and suggests that transferable emissions rights are not a practical system for encouraging gains from trade.

The artificial market created by a management system which defines transferable rights will only be capable of exhausting gains from trade for all initial allocations of rights if the rights granted are rights to pollute. Therefore, we will consider exclusively those management systems $\Omega(S)$ which define rights to pollute.

For any number of firms and any initial emissions vector, each member of a coalition can be made better off until the agreed emission vector minimizes the sum of costs for members of the coalition subject to the choice constraints. If E^o is an initial emissions vector, E^o is the vector of initial emission rates for firms in the coalition S. Let E^*_s be the vector in $\Omega(S)$ which minimizes joint total cost for firms in S. Then

$$\sum_{i \in S} F_i(e^o_1) \geq \sum_{i \in S} F_i(e^*_i),$$

and

$$\sum_{i \in S} [F_i(e^*_i) - F_i(e^o_i)] \leq 0.$$

Since it is always possible to find a set of payments which makes every member of a coalition better off when the formation of a coalition allows a reduction in joint total cost, we can concentrate our attention on the total cost reduction which can be achieved by a coalition. We do this by defining the characteristic function of the game associated with an air quality management system.

A characteristic function is a mapping from the power set of I into the real numbers which has two properties:

(1) $V(\phi) = 0$
(2) If R and S are any two disjoint subsets of I, $V(R \cup S) \geqslant V(R) + V(S)$.

The characteristic function of a game whose rules are given by a management system is defined as

$$V(S) = \sum_{i \in S} F_i(e_i^o) - \min_{E_s \in \Omega(S)} \sum_{i \in S} F_i(e_i).$$

For a coalition S, the characteristic function has the value which equals the difference between the sum of the costs incurred by its members if each acted independently and the least cost which can be incurred by the group when it chooses an emission vector in the coalition choice set.

The first property of a characteristic function is a matter of definition: its value is zero for the empty set. The second must be verified by proving that

$$\min_{E_s \in \Omega(S)} \sum_{i \in S} F_i(e_i) + \min_{E_R \in \Omega(R)} \sum_{i \in R} F_i(e_i) \geq \min_{E_{R+S} \in \Omega(R \cup S)} \sum_{i \in R \cup S} F_i(e_i).$$

We will abbreviate the constrained minimum operator $\min_{E_s \in \Omega(S)}$ to $\min_{\Omega(S)}$.

Note first that since there are no externalities between firms, the value of $\min_{\Omega(S)} \sum_{i \in S} F_i(e_i)$ is independent of the choices made by firms not included in S. Therefore, the characteristic function is well-defined. Now define e_i^* by

$$\sum_{i \in S} F_i(e_i^*) = \min_{\Omega(S)} \sum_{i \in S} F_i(e_i)$$

and

$$\sum_{i \in R} F_i(e_i^*) = \min_{\Omega(R)} \sum_{i \in R} F_i(e_i).$$

Since $E_S^* \in \Omega(S)$ and $E_R^* \in \Omega(R)$,

$$\sum_{i \in S} (h_{ij} e_i^* - c_{ij}) \leq 0$$

and

$$\sum_{i \in R} (h_{ij} e_i^* - c_{ij}) \leq 0.$$

Adding these two sets of inequalities, we have

$$\sum_{i \in R \cup S} (h_{ij} e_i^* - c_{ij}) \leq 0.$$

Defining $E_{R \cup S}^*$ as the vector obtained by combing E_S^* and E_R^*, we conclude that $E_{R \cup S}^* \in \Omega(R \cup S)$.

We can separate $\sum_{i \in R \cup S} F_i(e_i^*)$ as follows:

$$\sum_{i \in R \cup S} F_i(e_i^*) = \sum_{i \in S} F_i(e_i^*) + \sum_{i \in R} F_i(e_i^*) = \min_{\Omega(S)} \sum_{i \in S} F_i(e_i) + \min_{\Omega(R)} \sum_{i \in R} F_i(e_i).$$

Since $E^*_{R \cup S} \epsilon \Omega(R \cup S)$,

$$\min_{\Omega(R \cup S)} \sum_{i \in R \cup S} F_i(e_i) \leq \sum_{i \in R \cup S} F_i(e_i^*).$$

Therefore,

$$\min_{\Omega(R \cup S)} \sum_{i \in R \cup S} F_i(e_i) \leq \min_{\Omega(R)} \sum_{i \in R} F_i(e_i) + \min_{\Omega(S)} \sum_{i \in S} F_i(e_i).$$

This establishes that V(RwS) V V(R) + V(S).

Whether or not a positive advantage accrues from the formation of some coalition depends on whether or not $E^o = E^{**}$. Let us proceed to form partitions of the set I by repeated splitting of sets into pairs of disjoint subsets, and write

$$V(I) \geq V(I - S) + V(S) \geq V(I - S_1 - S_2) + V(S_1) + V(S_2) \geq \dots \geq \sum_i V(\{i\}).$$

If $E^o = E^{**}$, then $\sum_i F_i(e_i^{**}) = \sum_i F_i(e_i^o)$, and $V(I) = \sum_i V(\{i\})$.

In this case strict equality holds everywhere above, and no coalition can gain an advantage by forming. In this case the characteristic function is additive and the game is called "inessential." If $E^o \neq E^{**}$, we will not have strict equality everywhere: then the characteristic function is superadditive and the game is called "essential."

The proposition that when the initial emission vector is not efficient the game is essential implies the proposition that when some initial configuration is not optimal, there exist private bargains which can, in the absence of transaction costs, make everyone better off. The procedure used to establish the proposition is, however, informative. It was found first that whether or not private bargains could take place depended on whether or not there was a right to contract. It was shown that social constraints led to specific restrictions on this right, and that in order to construct a system in which a private bargain could achieve an efficient outcome we had to define specific rights and rules for their transfer.

The game derived from the management system will always have a nonempty core if the cost functions $F_i(e_i)$ are convex. It is possible to prove this directly by proving that the game is balanced because of the theorem that a balanced game always has a non-empty core.[7] The demonstration that the game is balanced is a straightforward modification of the method used by Scarf,[8] and is placed in an Appendix.

We can prove a deeper result, that the game is a market game. Theorem II in the Appendix verifies that the game is totally balanced.

Since a game is a market game if and only if it is totally balanced, the management game is a market game, as defined by Shapley and Shubik.[9] The validity of interpreting the coalition choice constraints as the definition of allowable rearrangements of rights is confirmed by the proof that the management system gives rise to a market game.

[7]Proved in Lloyd Shapley and Martin Shubik, "On Market Games," *Journal of Economic Theory*, 1 (1969), 9–25.

[8]Herbert Scarf, "On the Core of an N-Person Game," *Econometrica, 35* (2967), 50–69.

[9]Shapley and Shubik, *op.cit.*

We can indeed go further, for this co-operative market game can be decomposed into a competitive market. We can break up the coalition choice constraints into "budget" constraints on each firm. Define ℓ_{ij} as a quantity of rights to cause pollution at point j held by firm i, and ℓ_{ij}^{o} as the firm's initial allocation of these rights. Some enforcement mechanism is needed to ensure that each firm observes the following budget constraints relating its emissions and license holdings:

$$h_{ij}e_i \leq \ell_{ij}. \qquad (j = 1, \cdots, m)$$

Each firm can buy and sell rights at prices p_j to be established by the market. Then the firm will minimize

$$F_i(e_i) + \sum_j p_j(\ell_{ij} - \ell_{ij}^{o})$$

subject to the "budget" constraint. If $\sum_i \ell_{ij}^{o} = q_j^{*}$, the equilibrium prices, license holdings and emission rates in the market are such that $\sum_i F_i(e_i)$ is minimized subject to $EH \leqslant Q^{*}$.[10]

This is a competitive market with many buyers and sellers of rights. We have a predictive theory which asserts that its equilibrium will be achieved. If conditions under which price-taking is a rational strategy were not created by the specification of rights, there would be no such assurance. Among the many ways to define the right to contract, one leads to a predictable, efficient outcome.

[10]Proved in Montgomery, *op. cit.*

APPENDIX

Definition. A balanced collection of sets is a collection T for which there exists a measure δ_s such that $\delta_s \geqslant 0$ and

$$\sum_{\substack{S \in T \\ S \supset \{i\}}} \delta_s = 1.$$

Definition. A game is balanced if for every balanced collection of coalitions T a vector $(\pi_1, \ldots \pi_n)$ satisfies $\sum_i \pi_i \leqslant V(I)$ whenever $\sum_{i \in S} \pi_i \leqslant V(S)$ for all $S \in T$.

Theorem I. The game associated with an air quality management system is balanced.

Proof. Consider a vector $(\pi_1, \ldots \pi_n)$ such that

$$\sum_{i \in S} \pi_i \leq \sum_{i \in S} F_i(e_i^o) - \min_{\Omega(S)} \sum_{i \in S} F_i(e_i)$$

for all S in some arbitrary balanced collection T. For each coalition S there exists an emission vector E^s such that $\sum_{i \in S} F_i(e_i^s) = \min_{\Omega(S)} \sum_{i \in S} F_i(e_i)$, and

$$\sum_{i \in S} h_{ij} e_i^s \leq \sum_{i \in S} c_{ij}$$

and

$$\sum_{i \in S} \pi_i \leq \sum_{i \in S} [F_i(e_i^o) - F_i(e_i^s)].$$

We must prove that there exists a vector E such that

$$\sum_i h_{ij} e_i \leq \sum_i c_{ij}$$

and

$$\sum_i \pi_i \leq \sum_i [F_i(e_i^o) - F_i(e_i)].$$

(1) Define $$e_i = \sum_{\substack{S \in T \\ S \supset \{i\}}} \delta_s e_i^s.$$

Therefore
$$\sum_i h_{ij} e_i = \sum_i [h_{ij} \sum_{\substack{S \in T \\ S \supset \{i\}}} \delta_s e_i^s]$$

$$= \sum_{i=1}^{n} \sum_{\substack{S \in T \\ S \supset \{i\}}} \delta_s h_{ij} e_i^s$$

$$= \sum_{i \in S} \sum_{S \in T} \delta_s h_{ij} e_i^s$$

$$= \sum_{S \in T} \delta_s \sum_{i \in S} h_{ij} e_i^s.$$

Since $\sum_{i\in S} h_{ij}e_i^s \leq \sum_{i\in S} c_{ij}$,

$$\sum_{S\in T} \delta_s \sum_{i\in S} h_{ij}e_i^s \leq \sum_{S\in T} \delta_s \sum_{i\in S} c_{ij}.$$

But $\sum_{S\in T} \delta_s \sum_{i\in S} c_{ij} = \sum_{i=1}^{n} c_{ij} \sum_{\substack{S\in T\\ S\supset\{i\}}} \delta_s = \sum_{i=1}^{n} c_{ij}.$

Therefore, e_i defined as in this proof satisfies the choice constraints for the coalition I.

(2) By convexity and the definition of δ_s,

$$\sum_{\substack{S\in T\\ S\supset\{i\}}} \delta_s F_i(e_i^s) \geq F_i\Big(\sum_{\substack{S\in T\\ S\supset\{i\}}} \delta_s e_i^s\Big),$$

and by definition $e_i = \sum_{\substack{S\in T\\ S\supset\{i\}}} \delta_s e_i^s$.

Therefore $\sum_{\substack{S\in T\\ S\supset\{i\}}} \delta_s F_i(e_i^s) \geq F_i(e_i)$,

and $\sum_{i=1}^{n} [F_i(e_i^o) - \sum_{\substack{S\in T\\ S\supset\{i\}}} \delta_s F_i(e_i^s)] \leq \sum_{i=1}^{n} [F_i(e_i^o) - F_i(e_i)].$

By assumption $\sum_{i\in S} \pi_i \leq \sum_{i\in S} [F_i(e_i^o) - F_i(e_i^s)]$.

For each S, we can multiply both sides of the inequality by δ_s and preserve the sense of the inequality. Therefore

$$\sum_{S\in T} \delta_s \sum_{i\in S} \pi_i \leq \sum_{S\in T} \delta_s \sum_{i\in S} [F_i(e_i^o) - F_i(e_i^s)].$$

But $$\sum_{S\epsilon T} \delta_s \sum_{i\epsilon S} \pi_i = \sum_{S\epsilon T} \sum_{i\epsilon S} \delta_s \pi_i$$

$$= \sum_{i=1}^{n} \sum_{\substack{S\epsilon T \\ S\supset\{i\}}} \delta_s \pi_i$$

$$= \sum_{i=1}^{n} \pi_i \sum_{\substack{S\epsilon T \\ S\supset\{i\}}} \delta_s = \sum_{i=1}^{n} \pi_i,$$

and $$\sum_{S\epsilon T} \delta_s \sum_{i\epsilon S} [F_i(e_i^o) - F_i(e_i^s)] = \sum_{i=1}^{n} \sum_{\substack{S\epsilon T \\ S\supset\{i\}}} \delta_s [F_i(e_i^o) - F_i(e_i^s)]$$

$$= \sum_{i=1}^{n} \sum_{\substack{S\epsilon T \\ S\supset\{i\}}} \delta_s F_i(e_i^o) - \sum_{i=1}^{n} \sum_{\substack{S\epsilon T \\ S\supset\{i\}}} \delta_s F_i(e_i^s)$$

$$= \sum_{i=1}^{n} F_i(e_i^o) - \sum_{i=1}^{n} \sum_{\substack{S\epsilon T \\ S\supset\{i\}}} \delta_s F_i(e_i^s).$$

Therefore,

$$\sum_{i=1}^{n} \pi_i \leq \sum_{i=1}^{n} [F_i(e_i^o) - \sum_{\substack{S\epsilon T \\ S\supset\{i\}}} \delta_s F_i(e_i^s)]$$

$$\leq \sum_{i=1}^{n} [F_i(e_i^o) - F_i(e_i)].$$

Definition. A game is totally balanced if the game formed by replacing the set of players I with subset R $\subseteq$ I is also balanced.

Theorem II. The game derived from the management system is totally balanced.

Proof. Immediate by replacing I by R in Theorem I, and letting T be a balanced collection of subsets of R.