Journal of Economic Perspectives—Volume 8, Number 4—Fall 1994—Pages 111–128

The Allocation of Resources in the Presence of Indivisibilities

Herbert E. Scarf

The major problem presented to economic theory by the presence of indivisibilities in production is the impossibility of detecting optimality at the level of the firm, or for the economy as a whole, using the criterion of profitability based on competitive prices. I will explore this issue in a rather leisurely way, beginning with a discussion of the role played by competitive prices in verifying optimality when production takes place under constant returns to scale; then illustrating the failure of prices to perform this task when indivisibilities are significant; and, finally, suggesting the replacement of the pricing test by a specific quantity test. It is my hope that continued study of these quantity tests will increase our understanding of the division of labor in a large firm.

Using Prices to Detect Optimality Under Constant Returns to Scale

One of the most important professional activities of economists is to carry out exercises in comparative statics: to estimate the consequences and the merits of changes in economic policy or in our economic environment. We do comparative statics at the level of the firm when we calculate the effects of a change in factor endowments or in the price of a valuable input into production. We engage in comparative statics and dynamics for the economy as a whole when we examine the consequences of the dramatic increase in the price of imported oil in the latter part of 1973, or the second oil shock following the

■ *Herbert E. Scarf is Sterling Professor of Economics, Yale University, New Haven, Connecticut.*

fall of the Shah, or the dismantling of AT&T, or a massive change in income taxation within the United States, or the NAFTA. If these consequences spread throughout the entire economy, we evaluate them by assessing their effects on the well-being of the members of the community.

We have a remarkable paradigm for assessing well-being that has been passed on to us by generations of economic theorists and utilitarian philosophers. The utilitarian calculus, in its modern ordinal version, provides us with a simple test for evaluating the merits of a proposed change in economic activities: The change should be accepted if it has as an immediate consequence—or one that can be brought about by a suitable redistribution of income—an increase in the well-being or utility of all of the members of society.

The utilitarian test requires the possibility of major income redistributions that may not be politically viable—the movement of a clothing manufacturer from a Northern mill town to a lower wage region of the South may result in a potential Pareto improvement, but I know of no instances of an appropriate compensation to those employees whose jobs have been lost. And there are serious problems about maintaining effective incentives if lump sum transfers of income are made independently of effort and the supply of productive factors.

In spite of these and other reservations, I personally consider the welfare test to be an extraordinary intellectual construction—one which permits us to focus our discussions about the potential merits of a novel economic proposal. Last summer, for example, I participated in an extended discussion with a distinguished high energy physicist about the super-conducting super-collider. In that conversation, it became quite clear to me that the community of physicists in favor of the project had been unable to establish any ground rules about what constituted a compelling argument for the project. Of course, it wasn't the case that they had no arguments in favor of the collider; they had many of them. But most of these arguments could equally well have been presented for a project whose costs were orders of magnitude larger. There was no prior agreement or understanding between the proponents of the super-collider and their audiences about what constituted an acceptable argument for any particular level of expenditure.

Our profession does have such a line of discourse. It may, admittedly, be difficult to carry out the welfare test in an instance as complex as the super-collider; the collider is, after all, a public rather than a private good, and it is one whose potential benefits are extremely hard to predict. The utilitarian test is much easier to carry out when more conventional economic projects are proposed. The test actually leads to a simple exercise in the calculation of profitability, which, in my opinion, is one of the major theorems of microeconomic theory, a theorem which is not entirely obvious to the man on the street or even to professional economists.

Suppose that we are contemplating a hypothetical economic situation which is in equilibrium in the purest Walrasian sense. The production possibility set exhibits constant returns to scale so that there is a profit of zero at the equilibrium prices. Each consumer evaluates personal income (or wealth) at these prices and market demand functions are obtained by the aggregation of individual utility maximizing demands. The system is in equilibrium in the sense that demand equals supply for each of the goods and services in the economy.

Suppose that a technical advance is made resulting in the discovery of a new manufacturing activity subject to constant returns to scale—one which produces a good whose price is already known, at a new location, with different materials, with less expensive labor or with more sophisticated machinery. Shall the new activity be used at some positive level? The word "shall" in this question is the same word as in the question: "Shall the super-conducting super-collider be built?" The utilitarian test can be applied by inquiring whether the new activity can be combined with a plan of income redistributions in such a way as to make all consumers better off than they had previously been. On the face of it, this sounds as if we must solve a complex mathematical programming problem; but, in fact, the question has a remarkably simple answer: *If the activity is profitable at the old equilibrium prices, then there is a way to use the activity at a positive level so that with suitable income redistributions, the welfare of every member of society will increase.* There is no necessity to determine the new equilibrium prices arising after the activity is introduced; the current prices will do. And, conversely, *if the new activity makes a negative profit at the old equilibrium prices, then there is no way in which it can be used to improve the utility of all consumers, even allowing the most extraordinary schemes for income redistribution.* This is an astonishing mathematical theorem which I often ask about in graduate oral exams in microeconomics. The second assertion takes about two lines of proof; the first is more subtle: three lines of proof and a figure will do. I have never seen this theorem, which seems to me to be one of the important theoretical arguments in favor of private enterprise, in any textbook on economics.[1]

It may be worth remarking that if 17 new activities are presented simultaneously, the pricing test can be applied to the collection of activities in an arbitrary sequence, without regard to decisions made about the remaining activities. If none of the 17 activities makes a positive profit at the old equilibrium prices, then no subset of them can be used, along with income redistributions, so as to improve everyone's economic lot. If any one of the activities

[1]The second assertion: If the newly discovered activity is added to the old production possibility set, we obtain a new, and possibly larger, production possibility set. But the old equilibrium—which does not use the new activity—is still an equilibrium using the expanded production possibility set, because the new activity makes a non-positive profit at the old prices. By the first welfare theorem, the equilibrium must be Pareto optimal in the new setting with the larger set.

The first assertion: Simply draw the old production possibility set, a cone which is separated by a plane—whose normal is the equilibrium price vector—from the convex set of net trades that can be allocated among the consumers so as to improve their utility levels. The new activity ray lies above this price plane and, subject to some mild assumptions, it can be connected to the old equilibrium production plan so as to yield a feasible production plan which intersects this set of net trades and increases all consumers' utilities.

makes a positive profit, then some welfare improvement is surely possible. The activity can be introduced, a new equilibrium determined—with Pareto-improving income redistributions—and the pricing test can be applied to the remaining activities. This is an extraordinarily decentralized test; it presumably could be applied to every minor innovation on the shop floor of a large firm simply by evaluating its profitability in terms of prevailing market prices.

The market test sounds very much like a step in the simplex method for solving linear programs. An activity analysis model of the economy or a firm is given, along with a specified factor endowment and an objective function which is to be maximized subject to the constraint that the factor endowment is not exceeded. In a linear programming problem, a feasible solution to the constraints is proposed, and prices are found yielding a zero profit for the activities in use. The proposed solution is optimal if and only if the remaining activities make a profit less than or equal to zero.

The simplex method is an extremely efficient algorithm for solving linear programs: Programs involving thousands of variables can be solved routinely on a personal computer by high school students. But what is even more significant for economists is that this effective computational procedure is based on an evaluation of profitability identical to that performed by competitive markets. A visitor from another planet who was taught the simplex method for solving linear maximization problems would inevitably be led to the use of prices and profitability to detect optimality. An algorithm—a mathematical technique for solving maximization problems—suggests an institution—competitive markets—which is central to the way in which we organize our economic lives.

Is this suggestion of a major institutional structure an accident of the simplex method or can it be expected from other computational procedures as well? Is it a reasonable research strategy to address an area of economic theory which is not fully understood—at least by me—to cast it in the form of an optimization problem, and to hope that algorithms for its solution will produce a conceptual framework that is relevant to the original economic problem? I'm not sure, but it is a strategy that I have followed for a number of years in an attempt to increase my understanding of the problems posed for economic theory by indivisibilities and economies of scale.

The Failure of Prices in the Presence of Indivisibilities

Both linear programming and the Walrasian model of equilibrium make the fundamental assumption that the production possibility set displays constant or decreasing returns to scale; that there are no economies associated with production at a high scale. I find this an absurd assumption, contradicted by the most casual of observations. Taken literally, the assumption of constant returns to scale in production implies that if technical knowledge were universally available we could all trade only in factors of production, and assemble in our own backyards all of the manufactured goods whose services we would like to consume. If I want an automobile at a specified future date, I would purchase steel, glass, rubber, electrical wiring and tools, hire labor of a variety of skills on a part-time basis, and simply make the automobile myself. I would grow my own food, cut and sew my own clothing, build my own computer chips and assemble and disassemble my own international communication system whenever I need to make a telephone call, without any loss of efficiency. Notwithstanding the analysis offered by Adam Smith more than two centuries ago, I would manufacture pins as I needed them.

If production really does obey constant returns to scale, there is nothing to be gained by organizing economic activity in large, durable and complex units; in short, there is no economic justification for the existence of firms. Competitive markets would set prices for manufactured goods at every stage of production and cash would be exchanged, or accounts would be reckoned, as goods moved from one task to another. Every step in a complex manufacturing process would be tested for profitability by itself, without regard to its relationship to other potential improvements.

I am, I believe, not alone in thinking that the essence of economies of scale in production is the presence of large and significant indivisibilities in production. What I have in mind are assembly lines, bridges, transportation and communication networks, giant presses and complex manufacturing plants, which are available in specific discrete sizes, and whose economic usefulness manifests itself only when the scale of operation is large. If the technology giving rise to a large firm is based on indivisibilities, then this technology can be described by, say, an activity analysis model in which the activity levels referring to indivisible goods are required to assume integral values, like $0, 1, 2, \ldots,$ only. When factor levels are specified and a particular objective function is chosen, we are led directly to that class of difficult optimization problems known as integer programs.

For a theorist, the major problem presented by indivisibilities in production is the failure of the pricing test for optimality or for welfare improvements. Return to our previous discussion of the economy which is in full Walrasian equilibrium, and imagine, as before, that a new activity is discovered. But let us now assume, in contrast to our earlier example, that this new activity—perhaps a decision about the number of manufacturing plants of a particular type to be constructed— involves a discrete choice that can only be carried out at integral levels. One can argue easily that if the activity makes a negative profit at the old equilibrium prices, then there is no way to use it at a discrete or continuous level so as to improve the utility of every agent in the economy. The problem arises with the converse; it is perfectly possible that the activity make a positive profit at the old prices and still not be capable of being used at any discrete level to yield a Pareto improvement.

Even more problems arise if 17 activities are presented to us, all of which must be run at an integral level. A welfare improvement will typically require the selection of a subset of the activities, some of which are profitable at the old equilibrium prices, and some of which are not. There is no algorithm based on prices and profitability which permits us to make a sequence of welfare improvements by introducing one activity at a time, or even to detect which activities should ultimately be used. There is no pricing test in the presence of indivisibilities.

The absence of a pricing test is a truth that must be confronted. It certainly does imply that total decentralization by means of competitive prices is impossible if the technology involves serious indivisibilities. In my own view, this is a compelling reason for the existence of large firms, and it suggests that some serious insight about the large firm might be gained by considering such a firm to be essentially an algorithm for the solution of mathematical programming problems in which some of the variables are restricted to integer values. I hope that insights from this source would complement other insights about the functioning of large enterprises that are presented in a narrative rather than mathematical form, that are based on a careful analysis of particular historical cases, or that involve flows of information in hierarchical structures. The subject is sufficiently complex so that many voices should be heard.

An Example

An example may be useful. Consider a problem involving a single good that can be produced by a variety of technologies. Each technology is embodied in a particular type of manufacturing plant with a specific cost of construction, with a specific capacity, and with a specific unit cost of manufacturing. The level of demand for the product is given exogenously, and we are required to construct a series of plants and to manufacture sufficient product to satisfy this demand at minimum cost. We now have a mathematical programming problem in which some of the variables, the number of plants of each type to be constructed, are integral, and the remaining variables, the amounts manufactured at each plant, are continuous.

The example is artificial in many ways. Perhaps its most serious flaw is the obvious lack of any dynamic considerations.The construction cost is presumably paid at the time of construction, when a capacity for producing the maximum output per period is established. But demand for output manifests itself in a sequence of periods over time, possibly in a predictable though varying fashion, or possibly with a good deal of uncertainty. Moreover, it is plausible to assume that manufactured goods can be kept in inventory, at some cost, so as to satisfy future demand. These elements can certainly be introduced into our problem, but with a considerable increase in complexity. In order to make my points as simply as possible, I will assume that demand is constant over time and that no inventories are kept; the unit costs may then be thought of as the discounted sum of unit costs incurred over time as this constant demand is satisfied.

For fixed construction costs, capacities and unit costs, the optimal construction plan depends crucially on the level of demand. Some levels will call for considerable excess capacity in various plants, and other levels will not. How can we tell whether a proposed construction and manufacturing plan, which meets the demand requirement, does, in fact, minimize total cost? Competitive prices will not work for this class of problems. There is only one option: the price test must be supplemented, or replaced, by what can be called a "quantity test."

At this point, I have an expository difficulty about which I must be quite explicit. I would like to present an elementary example illustrating the particular quantity test required to demonstrate optimality without being cluttered by too much detail; this naturally leads to an example with a small number of plants, say, two. But programming problems with only two integer variables are easy to solve. In addition, when there are only two types of plants, the saving in cost achieved by a truly optimal solution is typically small compared to the cost of approximately optimal solutions, which are themselves quite easy to find. This is not true for larger problems, and I ask your indulgence on this issue.

With this caveat in mind, let us consider an example involving only two types of plants. The first type of plant—the Smokestack plant— is of ancient design, huge, made of red brick with steam pouring from its chimneys; it has a large capacity, is moderately inexpensive to construct per unit of capacity and has a fairly high marginal cost of production. The second plant—the High Tech plant—is a gleaming marvel of computerized technology; it has a capacity of medium size, is expensive to set up per unit of capacity, but has a lower marginal cost of production. Specific numerical values are provided in Table 1.

If capacity could be built continuously rather than in discrete units, the cost per unit of capacity in the Smokestack plant would be $53/16$ and the cost of supplying a unit of demand would be $53/16 + 3 = 6.3125$. The average

Table 1
Production Costs: Smokestack versus High Tech

	Smokestack	*High Tech*
Capacity	16	7
Construction Cost	53	30
Marginal Cost	3	2
Average Cost	6.3125	6.2857

Table 2
Cost Minimizing Choices of Plants and Output Levels

Demand	*#Smokestack*	*#High Tech*	*Output 1*	*Output 2*	*Total Cost*
55	3	1	48	7	347
56	0	8	0	56	352
57	1	6	15	42	362
58	1	6	16	42	365
59	2	4	31	28	375
60	2	4	32	28	378
61	3	2	47	14	388
62	3	2	48	14	391
63	0	9	0	63	396
64	4	0	64	0	404
65	1	7	16	49	409
66	2	5	31	35	419
67	2	5	32	35	422
68	3	3	47	21	432
69	3	3	48	21	435
70	0	10	0	70	440

construction and manufacturing cost from a High Tech plant is $30/7 + 2 = 6.2857$. What is, of course, uncomfortable about the example is the closeness of these two average costs.

If plants could be constructed at an arbitrary size, the market test—using either average or marginal cost as a criterion—would require that all demand be satisfied from High Tech plants alone. But the optimal solution is considerably different if plants must be built in discrete sizes, and the pricing test for optimality fails dramatically. Table 2 illustrates the cost minimizing choices of plants and the aggregate output levels of each type of plant for an interval of demand values.

If the capacities at both plants were larger, the number of plants of each type would be considerably less sensitive to the level of demand; a given configuration of underutilized plants would be optimal for a large interval of demands. It can easily be shown that the number of Smokestack plants becomes a periodic function of demand after some point ($d = 91$ in this example). But it is clear from this table that the optimal integer solution cannot be obtained simply by rounding the fractional solution in which only High Tech plants are used.

Quantity Tests for Optimality

Let us focus on a particular value of demand, say, 60, for which the optimal solution is to build two Smokestack plants, four High Tech plants, and manufacture 32 and 28 units respectively, for a cost of $378. Suppose that an alternative solution had been proposed: that we build three Smokestack plants (at a cost of $159 and providing a capacity of 48), two High Tech plants (at a cost of $60 and capacity of 14), and that we manufacture 46 units at the Smokestack plant and 14 units at the High Tech plant, for a total cost of $385. Is there a quantity test revealing that this proposal, which also satisfies the demand of 60, is not optimal?

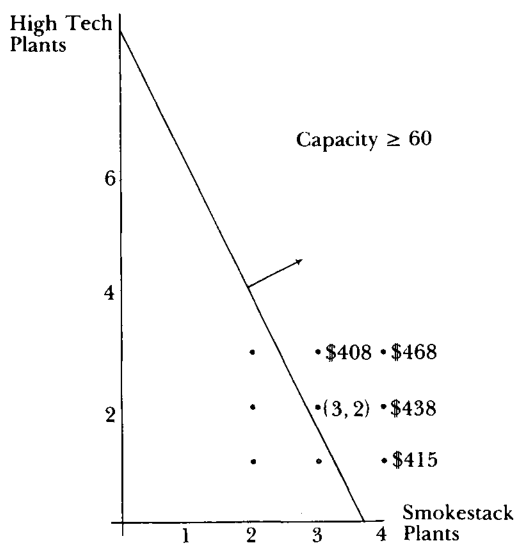
The most elementary quantity test is to plot the point (3, 2) in the plane, and examine its 8 neighbors, which are obtained by increasing or decreasing the number of plants of each type by unity. In other words, for any particular feasible construction plan given by a pair (#Smokestack plants, #High Tech plants), we examine those alternative construction plans obtained by adding to this pair of integers each of the 8 vectors:

$$(1,0),(1,1),(0,1),(-1,1),(-1,0),(-1,-1),(0,-1),(1,-1)$$

and testing each one of them to see whether it produces another feasible plan at lower cost.

In Figure 1, those combinations of Smokestack plants and High Tech plants which together provide a capacity of 60 units or more lie on or above the frontier—an unbounded region forming the constraint set for the two integer variables. (I have not drawn the iso-cost lines in this figure since cost depends not only on the number of plants of each type but also on the variables which do not appear in the figure: the levels of output from each type of plant.) The reader will notice that the four neighbors of (3, 2) given by (2, 2), (2, 1), (3, 1)

Figure 1
Looking in the Immediate Neighborhood



and (2, 3) all lie below the frontier and therefore do not provide sufficient capacity to satisfy the demand of 60. The remaining four neighbors (4, 2), (4, 3), (3, 3) and (4, 1) do provide adequate capacity for the demand of 60, but their total costs—assuming that High Tech plants are used to full capacity—are each larger than the cost of $385 associated with (3, 2). The plan (3, 2) is therefore a *local minimum* for this very natural quantity test, but it is not the *global minimum* when the demand is 60. Some alternative to this particular quantity test is required, if we are looking for a test with the property that a local minimum is global for any demand specification.

The Unique Minimal Quantity Test

For a quantity test to detect optimality it must be based on an examination of a set of neighbors that are related in some intrinsic fashion to the underlying problem, rather than being merely adjacent in an elementary geometric sense. For our problem, there is a *unique, minimal* set of neighbors all of which must be examined to be certain about detecting optimality if we wish to use translates of the same set for all feasible points and all levels of demand. They are obtained by subtracting each of the following neighbors from the proposed plan:

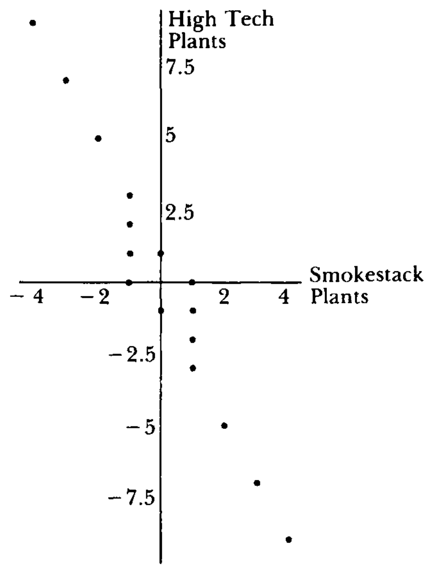
$$\begin{array}{l}(0,1),\\(1,0),(1,-1),(1,-2),\\(-1,3),(-2,5),(-3,7),\\(4,-9),(7,-16).\end{array}$$

This set of neighbors, displayed along with its negatives in Figure 2, has the important property that if one of its translates is excluded from the test set, then there will be some level of demand and some feasible solution which is not optimal but which cannot be improved by moving to any neighbor in the smaller test set.[2] If all of them are examined, the local quantity test based on this set of neighbors will yield the global optimal solution for any level of demand.

To apply the neighborhood test of Figure 2, imagine that we lift up the set of those neighbors that decrease cost and translate them from the origin to the proposed solution (3, 2), as in Figure 3. The non-optimality of the plan (3, 2), for the demand of 60, is easily seen in the figure by noticing that if we subtract

[2]Figure 2 displays eight of the nine neighbors listed above (I've left out (7, −16) in order to keep the scale reasonable) and their negatives as well. These nine neighbors decrease cost when subtracted from a trial solution, and their negatives increase cost. The negatives are useful if an infeasible solution is given and we are searching for a feasible neighbor, or for a variant of the problem in which we are looking for a configuration of plants that maximizes output subject to a cost constraint.

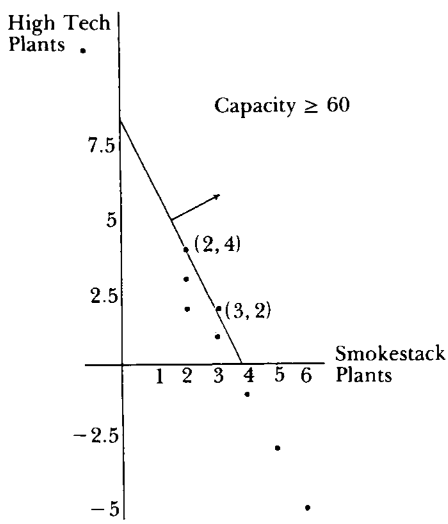
Figure 2
A Global Quantity Test



the neighbor (1, − 2) from (3, 2), we reach the new plan (2, 4) which is feasible, lowers cost, and is, in this instance, optimal.

There is a clear algorithm suggested by these considerations: 1) Propose some construction plan which produces a capacity sufficient to meet demand; 2) If one of its neighbors in the unique minimal test set is also feasible and leads to a lower cost, move to that alternative plan; 3) If there are no such neighbors,

Figure 3
Applying the Global Quantity Test



the original proposal is optimal. This algorithm depends crucially on our requirement that the local test set be the same for every feasible point; the set of neighbors would otherwise have to be determined anew at each iteration.

It can be seen that these neighbors are closely related to the discrete analogue of marginal products. As the demand level increases, the optimal construction plan will either be unchanged or move to a new plan which is obtained from the previous plan by adding one of these neighbors.

The reader may feel, at this moment, that I've pulled this collection of neighbors out of a hat. I don't mean to be unduly mysterious, so let me be more specific about the role that neighbors play in detecting optimality. Consider the neighbor (1, −2). If, as in our example, we *subtract* this neighbor from a proposed construction plan, we obtain a new plan with 1 less Smokestack plant and 2 more High Tech plants. There will be a net loss in capacity of two units and a net increase in construction costs of $7. But the 2 additional High Tech plants are capable of manufacturing 14 units at a cost of $28; these 14 units were previously manufactured at the Smokestack plant for an additional $1 per unit. It follows that there is a cost saving of $7 associated with this decrease in capacity of two units.

This "marginal" change would result in a decrease in cost if the original plan had at least 2 units of excess capacity and used at least one Smokestack plant; under these circumstances the change should certainly be adopted and we should move to a new solution with lower cost (as we did in our example in moving from the configuration (3, 2) to (2, 4)). If the demand requirement had been 61 units rather than 60, the configuration (3, 2) would equally well have been feasible. But the excess capacity would now be 1 unit rather than 2; we no longer can subtract the neighbor (1, −2) and remain with sufficient capacity. On the other hand, it is easy to see that subtracting the neighbor (−3, 7) reduces capacity by only one unit, but we can only move to this neighbor if we are currently contemplating a plan with seven or more High Tech plants. Subtracting (4, −9) also reduces capacity by a single unit, but for this change to lead to a new feasible configuration, we must currently be constructing at least four Smokestack plants. And, finally, if we build seven fewer Smokestack plants and 16 more High Tech plants, there will be no change in capacity but a reduction in cost of $3, an option we should certainly take if the proposed plan involves building seven or more Smokestack plants.

If we examine the members of the minimal test set, the decrease in capacity and cost obtained by *subtracting* each of them from a proposed solution is given in Table 3, assuming that all High Tech plants are used to full capacity. The set of neighbors provides all of the discrete tradeoffs necessary to verify optimality. Of course, it is not obvious—nor can I make it obvious without an argument that is quite straightforward, but which I would rather not include in this paper —precisely why no other discrete tradeoffs are necessary.[3]

[3]A simple algorithm for constructing the set of neighbors when there are two integral variables is given in Scarf (1981b). There is sufficient regularity in Table 3 so that the reader might be able to guess what the algorithm is without looking at this paper.

Table 3
Decreases in Capacity and Cost for Each Neighbor

Neighbor	*Capacity*	*Cost*
(0, 1)	7	23
(1, 0)	16	53
(1, −1)	9	30
(1, −2)	2	7
(−1, 3)	5	16
(−2, 5)	3	9
(−3, 7)	1	2
(4, −9)	1	5
(7, −16)	0	3

These observations are quite general. Subject to very mild conditions, an arbitrary activity analysis model with integral activity levels has associated with it a unique, minimal neighborhood system, which depends solely on the technology matrix and not on the specific factor endowment, and such that a local maximum with respect to this neighborhood system is a global maximum for any particular right-hand side (Scarf, 1981a). The minimal neighborhood system is an intrinsic feature of the discrete production possibility set and is fully independent of the particular factor endowment. In suggesting a quantity test set which depends only on the specification of technical possibilities, we are maintaining the distinction between technical knowledge and factor availability which has been so fruitful in economic analysis.

In our example, the technology is simply given by the cost structure, and aside from non-negativity of the variables there is a single constraint requiring that output be greater than or equal to demand. We would have more constraints if there were demands for output at several locations, or if we were explicit about a variety of factors of production. The neighborhood system would still be independent of the demand specification and factor endowments.

In the next section, I shall describe a few structural properties of neighborhood systems suggesting that they can be examined rapidly and systematically when there are only two integral variables. Other properties are known for higher dimensions, and many more remain to be discovered. To the extent that a firm can be viewed partially as an algorithm for the solution of discrete programming problems, the firm must either know this minimal test set explicitly—or in some implicit fashion—to test for optimality as the factor endowment varies. The optimal allocation of indivisible resources is essentially a combinatorial problem, addressed in this paper by means of neighborhood systems. In a large firm, whose size arises from a technology involving indivisibilities, the resolution of these combinatorial problems must be reflected in the firm's organized decision-making procedures for selecting appropriate responses to changes in economic circumstances.

How, in general, are these neighbors to be determined for a given technology matrix? I find it astonishing that there is a canned computer program that can be found either in *Mathematica* or *Maple* which automatically calculates the set of neighbors if presented with the underlying activity analysis matrix.[4] The program is not designed with this particular question in mind; its purpose is to compute a very sophisticated object in a field of mathematics known as algebraic geometry, a topic which is far removed from issues of economic theory. But here we see one of the remarkable, though rare, virtues of the translation into mathematical form of an everyday problem: words, phrases, and concepts which bear no apparent relationship to each other in ordinary discourse may become synonymous in the language of mathematics.

Comparative Statics in the Presence of Indivisibilities

Can the minimal neighborhood system be used to analyze changes in optimal behavior resulting from a modification in our economic environment? One type of modification is an exogenous change in factor endowments or, in our example of plant selection, a change in demand for output. The minimal neighborhood system permits us to analyze this type of change quite readily in the sense that, for a general activity matrix, changes in the optimal solution associated with increases in the factor endowment or demand are given by precisely these neighbors.

A more complex change results from a modification in the technology rather than the factor endowment. In our earlier discussion in which production exhibited constant returns to scale, this feature was illustrated by the introduction of a new activity whose profitability could be tested at the old equilibrium prices. If indivisibilities are present, a new activity analysis matrix results in a new minimal neighborhood system.

Our numerical example is so elementary that the only changes in technology are modifications in the costs and capacities of the two competing types of plants. To see the consequences for the minimal neighborhood system, let me first remark that the set of neighbors is unaltered if both capacities change by a common factor and if the average costs—the cost per unit of capacity plus marginal cost—are changed by a possibly different common factor for both types of plant.

Let us make such a rescaling, followed by slight increases in capacity at the Smokestack plant and marginal cost at the High Tech plant so that the parameters are now given by Table 4. At this point, the average cost at

[4]The program finds a mathematical object known as a Groebner Basis for a polynomial ideal whose generators are defined by the columns of the activity analysis matrix.

Table 4
Changes Which Bring Average Costs Closer

	Smokestack	*High Tech*
Capacity	1609	700
Construction Cost	5300000	3000000
Marginal Cost	3000	2010
Average Cost	6293.97	6295.71

the Smokestack plant is slightly lower than that of the High Tech plant and the previous minimal test set is increased by four new neighbors; it is now given by:

$$
\begin{aligned}
&(0,1),\\
&(1,0),(1,-1),(1,-2),\\
&(-1,3),(-2,5),(-3,7)\\
&(4,-9),(7,-16)\\
&(-10,23),\\
&(17,-39),(27,-62),(37,-85).
\end{aligned}
$$

The changes in capacities and costs obtained by subtracting each of these neighbors from a proposed feasible solution are given in Table 5. As we see, if

Table 5
Decreases in Capacity and Cost for each Neighbor

Neighbor	*Capacity*	*Cost*
(0, 1)	700	2307000
(1, 0)	1609	5300000
(1, −1)	909	2993000
(1, −2)	209	686000
(−1, 3)	491	1621000
(−2, 5)	282	935000
(−3, 7)	73	249000
(4, −9)	136	437000
(7, −16)	63	188000
(−10, 23)	10	61000
(17, −39)	53	127000
(27, −62)	43	66000
(37, −85)	33	5000

the change in costs increases the competitiveness of the plants—causes their average costs to converge—the number of neighbors will expand. If the plants are closer in efficiency, a higher level of scrutiny is required to detect optimality.[5]

For integer programming problems with two variables, the set of neighbors can be linearly ordered, as in our examples. Small changes in the specification of the problem will always result in modifying our degree of resolution by adding or deleting an interval of neighbors at the end of the list. When the parameters change continuously, the unique procedure for detecting optimality changes in the most elementary fashion possible for a discrete, ordered set of points: the set grows or shrinks at one end. One of the major themes of my current research is to describe the ways in which the set of neighbors changes when the number of discrete choices is larger than two and the neighbors are no longer organized linearly. All of the present evidence suggests that, for the general integer programming problem, the minimal test set gains or loses members at a small number of locations on its boundary.

These examples also illustrate some unexpected structural elements of the set of neighbors: the set seems to be composed of a small number of linear segments. This is a very desirable feature, since the question of whether a member of a linear set of neighbors can be added to a proposed feasible solution so as to retain feasibility and decrease cost is easy—rounding will do. It is not difficult to argue that this structure is valid for an arbitrary problem with two integer variables; the set of neighbors always consists of a small number of intervals (Scarf, 1981b). This observation permits us to construct what computer scientists call a "polynomial"—a really fast—algorithm for integer programs with two variables.

A remarkable accomplishment of mathematical programming is the generalization of this result to problems with an arbitrary number of integer variables. For any fixed number of integer variables, there is an integer programming algorithm which uses themes similar to, though not identical with minimal test sets, and which executes in "polynomial" time—very rapidly —as the other parameters of the problem vary. These algorithms have more than theoretical interest: they have been coded by experts, and seem to be

[5]Consider the new neighbor $(-10, 23)$ pointing in the direction of a cost reduction if excess capacity is 10 or more and if at least 23 High Tech plants are under consideration. Why is this not a neighbor under the previous cost structure? Under the cost structure of Table 1, if we had decided to build 10 more Smokestack plants and 23 fewer High Tech plants, there would have been a decrease in capacity of 1 unit. Construction costs would have decreased by $160, but marginal costs would have increased by $161, for a net increase in cost of $1. This choice would not have lead to a decrease in cost even if there had been excess capacity in the original proposal.

Of course, this argument might suggest that 10 fewer Smokestack and 23 more High Tech plants should have been built under the older cost regime, since this would result in a reduction in cost. But this would be feasible only if excess capacity was at least 1 unit and the number of contemplated Smokestack plants was 10 or more. If this had been so, the lack of optimality would already have been detected by the earlier neighbor leading to 7 fewer Smokestack and 16 more High Tech plants.

among the best general purpose mixed integer programming algorithms currently available.[6]

For Studying Large Firms, Pack Up Your Derivatives in Mothballs

But let us leave this example with only two discrete choices concerning types of plants, and remember that in a large manufacturing enterprise there will be many discrete choices involving a large menu of tasks and machinery, each of which has its own capacity, set-up cost and marginal cost. The equipment may be placed in a number of different locations on the shop floor; the work may be passed from one piece of machinery to another with complex requirements of scheduling and precedence, and the tasks may alter from one job lot to another as the product specification varies. Demands may be revised capriciously and unexpectedly over time; output may be shipped to many different regions. The enterprise may have a host of competitors or none at all. In the absence of internal market prices, combinatorial arguments and quantity tests are necessary to regulate the flow of activity inside the enterprise in an optimal fashion.

My message boils down to a simple straightforward piece of advice; if economists are to study economies of scale, and the division of labor in the large firm, the first step is to take our trusty derivatives, pack them up carefully in mothballs and put them away respectfully; they have served us well for many a year. But derivatives are prices, and in the presence of indivisibilities in production, prices simply don't do the jobs that they were meant to do. They do not detect optimality; they aren't useful in comparative statics; and they tell us very little about the organized complexity of the large firm. Neighborhood systems are the discrete approximations to the marginal rates of substitution revealed by prices. They are relatively easy to compute, seem to behave pretty well under continuous changes in the technology, and will ultimately lead to even better algorithms than we have now.

We know much more about the structure of neighborhood systems than I have been able to describe here—not enough, perhaps, to derive a really satisfactory theory of the internal organization of the large firm at the present time. But my own intuition is that this is an important way to proceed. I am confident that serious, ultimately useful insights about the large firm will eventually be obtained by thinking very hard and long about indivisibilities in production.

[6]The seminal paper, exhibiting the first polynomial algorithm when the number of variables is fixed at an arbitrary level, is Lenstra (1983). Lenstra's argument requires the approximation by ellipsoids of many convex bodies in high dimensions. This approximation is avoided in Lovász and Scarf (August 1992). An elementary exposition of the Lovász and Scarf variant of Lenstra's algorithm may be found in Scarf (1990). Computational experience is discussed in Cook, Rutherford, Scarf, and Shallcross (1993).

■ *I would like to thank my colleagues Truman Bewley, William Brainard, Alvin Klevorick, William Nordhaus, T.N. Srinivasan and James Tobin for their thoughtful comments.*

References

Cook, William, and Thomas Rutherford, Herbert E. Scarf, and David Shallcross, "An Implementation of the Generalized Basis Reduction Algorithm for Integer Programming," *ORSA Journal on Computing*, Spring 1993, *5*, 206–12.

Lenstra, Hendrik W. Jr., "Integer Programming with a Fixed Number of Variables," *Mathematics of Operations Research*, 1983, *8*, 538–48.

Lovász, László, and Herbert E. Scarf, "The Generalized Basis Reduction Algorithm," *Mathematics of Operations Research*, August 1992, *17*, 751–64.

Scarf, Herbert E., "Production Sets with Indivisibilities Part I: Generalities," *Econometrica*, January 1981a, *49*, 1–32.

Scarf, Herbert E., "Production Sets with Indivisibilities Part II: The Case of Two Activities," *Econometrica*, March 1981b, *49*, 395–423.

Scarf, Herbert E., "Mathematical Programming and Economic Theory," *Operations Research*, May–June 1990, *38*, 377–85.