*Journal of Economic Perspectives—Volume 8, Number 4—Fall 1994—Pages 161–181*

# Distinguished Fellow
# Herbert Scarf's Contributions to Economics

Kenneth J. Arrow and Timothy J. Kehoe

Herbert Scarf never formally studied economics. Both his undergraduate and his graduate training were in mathematics. For the past 35 years, however, Scarf has worked on problems at the core of economic theory. He has deepened our understanding of the nature of standard economic theory, linked together disparate lines of study, made theory usable by creating computational tools, and explored the areas of increasing returns and nonconvexity, areas in which standard theory has had the least to offer. His work has simultaneously sought the most general and abstract formulations, while simplifying the theory and bringing it closer to applications. Scarf's election as Distinguished Fellow of the American Economic Association in 1991 honors the depth and diversity of his many contributions in studies of inventories, the core, computation of equilibria, and integer programming, all of which have had a strong influence on the way economists have learned to think about these subjects, and some of which have led to flourishing lines of research.

Scarf's early interest in economic problems was formed in the rather heady and free-wheeling environment of the Rand Corporation in Santa Monica, California, in the 1950s. He had entered Princeton in 1951 after receiving his A.B. from Temple. Like other holders of doctorates in mathematics and economics from Princeton at that time—such as his friends Ralph Gomory, Lloyd Shapley, and Martin Shubik—he was eventually drawn to an interest in game theory and mathematical programming. Although these friendships would later prove influential for his economic research, he recalls that while at Princeton he remained innocent of these fields [21]. (As is conventional for this

■ *Kenneth J. Arrow is Emeritus Professor of Economics, Stanford University, Stanford, California. Timothy J. Kehoe is Professor of Economics, University of Minnesota, and Adviser, Federal Reserve Bank of Minneapolis, both in Minneapolis, Minnesota.*

journal, articles by Scarf are referenced according to the numbers in Table 1, while all other citations appear in the reference list at the end of the article.) In his first year at Princeton, he wrote a paper that contained a new proof of the fundamental theorem of algebra based on group invariant integration [1]. His thesis research dealt with diffusion processes on differentiable manifolds.

The Rand Corporation was created by the U.S. Air Force in the belief that the nature of warfare had been so fundamentally altered by technological developments that new ideas were necessary. Game theory and allied thinking in what came to be known as decision theory, a great deal of which had been developed at Princeton under John von Neumann of the Institute for Advanced Study and Oskar Morgenstern and Albert Tucker of Princeton, stood ready to meet the demand. Many distinguished scholars with mathematical backgrounds came to the Rand Corporation, some on a permanent basis, some for summers. Scarf joined the mathematics department at Rand in 1954, after receiving his Ph.D.

At Rand, Scarf showed his ability to follow new developments. He began to work on game theory with Shapley, with whom he wrote a paper on dynamic games with incomplete information [2]. His characteristic strengths emerged gradually as he evolved a style of time-intensive full engagement with a problem. They were first displayed in his work on inventory theory. Rand had held a conference on inventory management in 1950, and Kenneth Arrow, Theodore Harris, and Jacob Marschak presented there a pioneer dynamic inventory model with uncertainty. As detailed below, Samuel Karlin and Scarf continued the work, developing the intricate stochastic process problems in this framework. But Scarf characteristically went more deeply into the foundations: where previous work had assumed the form of the policy and only sought to determine its parameters, Scarf elegantly and relevantly formulated the cost assumptions that implied the optimal shape of the policy. The results were typical of his future work; when the answer came, it appeared as the only possible way of approaching the question, though no one had known previously how to formulate it.

This clarity of vision and thoroughness of knowledge are appreciated by his auditors and readers. Those who have heard Scarf present his work in a seminar or have read one of his papers or books can testify that his oral and written exposition are models of clarity of thought and courtesy to his audience. His students and colleagues testify that his contributions go even further: Scarf is a concerned and dedicated teacher, adviser, and colleague, generous with his energy and enthusiasm.

This courtesy and intellectual honesty go hand in hand. In his writings, seminars, classes, and conversations, Scarf pays attention to his audience, explaining what that particular audience needs to know to understand the problem he is concerned with: its significance, his proposed solution, and what is left undone by the solution. Since his audience is often broad, doing this without compromising or oversimplifying requires a deep understanding of the

*Table 1*

**Works by Herbert Scarf Referred to in This Essay**


1. "Group Invariant Integration and the Fundamental Theorem of Algebra," *Proceedings of the National Academy of Sciences of the U.S.A.*, May 1952, *38*, 439–40.
2. "Games With Partial Information," (with Lloyd Shapley). In Dresher, M., A. W. Tucker, and P. Wolfe, eds., *Contributions to the Theory of Games, III*. Princeton: Princeton University Press, 1957, 213–29.
3. *Studies in Mathematical Theory of Inventory and Production*, (with Kenneth J. Arrow and Samuel Karlin). Stanford: Stanford University Press, 1958.
4. "The Optimality of $(S, s)$ Policies in the Dynamic Inventory Problem." In Arrow, K. J., S. Karlin, and P. Suppes, eds., *Mathematical Methods in the Social Sciences, 1959: Proceedings of the First Stanford Symposium*. Stanford: Stanford University Press, 1960, 196–202.
5. "Some Examples of Global Instability of the Competitive Equilibrium," *International Economic Review*, September 1960, *1*, 157–72.
6. *Contributions to the Theory of Inventory and Replacement*, (with Kenneth J. Arrow and Samuel Karlin). Stanford: Stanford University Press, 1961.
7. "An Analysis of Markets with a Large Number of Participants." In *Recent Advances in Game Theory: Papers Delivered at a Meeting of the Princeton University Conference*, October 4–6, 1961. Princeton: Princeton University Press, 1962, 127–56.
8. *Multistage Inventory Models and Techniques*, (with Dorothy Gifford and Maynard Shelly, eds. Stanford: Stanford University Press, 1963.
9. "A Limit Theorem on the Core of an Economy," (with Gerard Debreu), *International Economic Review*, September 1963, *4*, 235–46.
10. "The Core of an N-Person Game," *Econometrica*, January 1967, *35*, 50–69.
11. "The Approximation of Fixed Points of a Continuous Mapping," *SIAM Journal of Applied Mathematics*, September 1967, *15*, 1328–43.
12. *The Computation of Economic Equilibria*, (with the collaboration of Terje Hansen). Cowles Foundation Monograph No. 24. New Haven: Yale University Press, 1973.
13. "The Solution of Systems of Piecewise Linear Equations," *Mathematics of Operations Research*, (with B. Curtis Eaves), February 1976, *1*, 1–27.
14. "Production Sets with Indivisibilities, Part I: Generalities," *Econometrica*, January 1981, *49*, 1–32.
15. "Production Sets with Indivisibilities, Part II: The Case of Two Activities," *Econometrica*, March 1981, *49*, 395–423.
16. "Neighborhood Systems for Production Sets with Indivisibilities," *Econometrica*, May 1986, *54*, 507–32.
17. "Notes on the Core of Production Economy." In Hildenbrand, W., and A. Mas-Colell, eds., *Contributions to Mathematical Economics: In Honor of Gerard Debreu*. Amsterdam: North-Holland, 1986, 401–29.
18. "Mathematical Programming and Economic Theory," *Operations Research*, May-June 1990, *38*, 377–85.
19. "The Shapes of Polyhedra," (with Ravi Kannan and László Lovász), *Mathematics of Operations Research*, May 1990, *15*, 364–80.
20. "An Implementation of the Generalized Basis Reduction Algorithm for Integer Programming," (with William Cole, Thomas Rutherford, and David F. Shallcross), Cowles Foundation Discussion Paper 990, 1991.
21. "The Origins of Fixed Point Methods." In Lenstra, J. K., A. H. G. Rinnooy Kan, and A. Schrijver, eds., *History of Mathematical Programming: A Collection of Personal Reminiscences*. Amsterdam: North-Holland, 1991, 126–34.
22. "The Generalized Basis Reduction Algorithm," (with László Lovász). *Mathematics of Operations Research*, August 1992, *17*, 751–64.
23. "The Frobenius Problem and Maximal Lattice Free Bodies," (with David F. Shallcross), *Mathematics of Operations Research*, August 1993, *18*, 511–15.
24. "Shortest Integer Vectors," (with David F. Shallcross), *Mathematics of Operations Research*, August 1993, *18*, 516–22.
25. "The Allocation of Resources in the Presence of Indivisibilities," *Journal of Economic Perspectives*, this issue.

problem. Students in his classes learn quickly that mathematical rigor does not depend on formalism or on heavy use of notation. The best test of whether a researcher fully understands his own work is whether he can explain it to someone who is not an expert in the field.

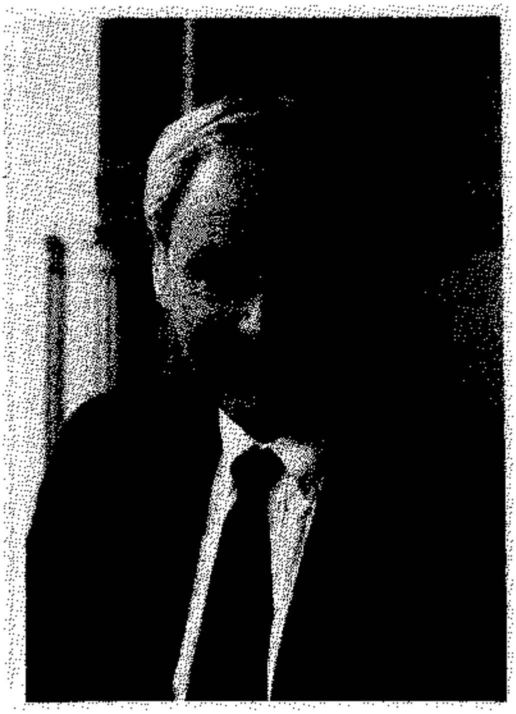
## Inventories and Introduction to Economics

While at Rand, Scarf met Arrow and Karlin, who asked him to collaborate with them on work on inventory problems at Stanford in 1956–57, a collaboration that resulted in several volumes of collected papers [3, 6, 8]. In 1957, Scarf left Rand to join the Department of Statistics at Stanford. He immediately established himself as a great teacher, offering an unusual combination of lucidity and depth. His course of mathematical methods for social scientists became a center for the new generation of graduate students in economics and psychology.

Perhaps the most widely cited work on inventory theory is Scarf's paper on the optimality of the $(S, s)$ policy [4]. It deals with the dynamic inventory problem introduced by Arrow, Harris, and Marschak (1951). Here is the situation: a retailer faces uncertain demand for its product over discrete time periods. It pays a fixed reorder cost and a unit cost when it orders this good from its producer. Over time, it pays a holding cost based on the size of its inventory and a shortage cost based on the magnitude of excess demand if it runs out of the good. Arrow, Harris, and Marschak had assumed (following actual practice) that the optimal inventory policy had the so-called $(S, s)$ or two-bin form; that is, the firm waits until its inventory falls below a lower fixed level, $s$, and then makes a reorder that restores the stock to an upper fixed level, $S$. They found the optimal policy within this restricted class, but did not show that this was indeed optimal among all possible policies. It was soon shown by Aryeh Dvoretzky, Jack Kiefer, and Jacob Wolfowitz (1953) that, with appropriate specifications on the demand process and the costs, $(S, s)$ policies need not be optimal. Karlin (1958) gave a special set of assumptions on costs and demand which sufficed for the optimality of $(S, s)$ policies.

Scarf approached the problem in a manner which became characteristic. He defined a generalization of convexity, called $K$-convexity. He showed that $K$-convexity for the total cost function is preserved by the recursion formulas of the dynamic inventory model and that it holds whenever holding and shortage costs are linear or, more generally, convex. Once the importance of $K$-convexity was grasped, it was easily shown that an $(S, s)$ policy is optimal in each period. Scarf considered a model with a finite number of periods; a student of his, Donald Iglehart (1963), extended the analysis to the infinite-horizon case.

There are obvious applications of inventory theory in monetary economics, where, for example, consumers face uncertain need for cash balances and a fixed cost of replenishing these balances. Other applications are in business



**Herbert Scarf**

cycle theory. For example, Andrew Caplin (1985), working on a thesis under Scarf's supervision, found that if retailers follow $(S, s)$ inventory policies, then variations in final demand are magnified in variations in the demand faced by producers. This finding directly contradicts the widely held notion that inventories serve as a buffer that protects manufacturers from fluctuations in sales. It is also in accord with the evidence on the relative variabilities of sales and deliveries to retailers (see, for example, Alan Blinder, 1981).

While working with Arrow and Karlin on inventories, Scarf was introduced to general equilibrium theory. At the time, Arrow was working with Leonid Hurwicz and Hirofumi Uzawa on stability of equilibrium. Scarf's recent work had given him a taste for the operations research approach to solving problems by computing solutions, and he was concerned with the applicability of the *tâtonnement* adjustment process for finding the equilibrium prices of a pure exchange economy.

To understand the issues involved in this and in much of Scarf's subsequent research, consider an economy with $n$ goods. In this economy a consumer sells his endowment vector of goods and purchases a vector of goods to consume so as to maximize utility subject to the budget constraint. The aggregate response of all of the consumers to a vector of prices $p = (p_1, \ldots, p_n)$ is summarized in an *aggregate excess demand function* $f(p) = (f_1(p), \ldots, f_n(p))$. Good by good, this function reports the sum of the consumers' demands for that good minus total supply from the endowments. Under standard assumptions on the consumers' utility functions and endowments this function is continuous and homogeneous of degree zero; in other words, demands do not change if all prices are multiplied by a positive constant. Furthermore, it obeys Walras's law, which says that the value of excess demands summed over all goods is always equal to zero.[1]

An equilibrium is a vector of prices such that excess demand is less than or equal to zero for every good, $f(\hat{p}) \leq 0$. Walras's law ensures that excess demand is actually equal to zero if the corresponding price is positive, but this concept of equilibrium allows for free goods. Léon Walras had proposed in 1874 a process that he called *tâtonnement*, or groping, to find equilibrium prices. This process, later formalized as a system of differential equations by Paul Samuelson, raises the price of a good in positive excess demand and lowers the price of a good in negative excess demand.

Arrow, Hurwicz, Uzawa, and others had found that the *tâtonnement* process always leads to an equilibrium if excess functions satisfy either the weak axiom of revealed preference or gross substitutability. They were hopeful that the method would work under far more general conditions and were busy searching for such conditions. Scarf [5] dashed their hopes by producing a simple example with three goods and three consumers in which there is a unique equilibrium but, unless the *tâtonnement* process is started precisely at those prices, the process will literally go around in circles forever.

This result comes as no surprise now, since a series of papers by Hugo Sonnenschein (1973), Rolf Mantel (1974) (a student of Scarf), and Gerard Debreu (1974) has shown that, with a sufficient number of heterogeneous consumers, the excess aggregate demand function is essentially arbitrary (except for continuity, homogeneity, and Walras's law) and hence that the behavior of the *tâtonnement* process is also essentially arbitrary. At the time, however, Scarf's paper had considerable influence on discouraging enthusiasm for the *tâtonnement* method. It also had the effect of focusing his own attention on the lack of an algorithm for calculating general equilibria.

## The Core

After giving a talk on his example of instability at Columbia University in 1960, Scarf discussed with Martin Shubik the relationship of the equilibria of an exchange economy to its core, a topic that Shubik had been studying. The *core* is an allocation of the aggregate endowment such that no coalition of consumers can find an alternative allocation for its members that both is feasible

[1]Homogeneity of degree zero says that $f(\theta p) = f(p)$ for any price vector $p$ and any number $\theta > 0$. Walras's law says that $p_1 f_1(p) + \cdots + p_n f_n(p) = 0$.

given their endowments and makes all of them better off. That evening, a conversation between Scarf and Lloyd Shapley, who was visiting Shubik at the time, established that any competitive allocation was in the core. Shubik proposed the conjecture that, as the number of consumers increased to infinity, the core would converge to the set of competitive allocations.

To specify an allocation in an economy with $n$ goods and $m$ consumers, we need a list of $n \times m$ numbers to describe the assignment of each good to each consumer. As the number of consumers increases, this list becomes longer and the core becomes a more complicated object. To avoid the awkwardness in constantly changing the dimensionality of the core, Scarf used an approach employed by Shubik (1959), who in turn had followed Francis Edgeworth: he restricted his attention to economies with a fixed number of types of consumers, specified in terms of utility functions and endowments, and increased the total number of consumers by replicating all of these consumers the same number of times. Assuming that the core assigned the same consumption bundle to all consumers of the same type, Scarf [7] was able to prove that, if any allocation was in the core for any number of replications, then there existed a vector of prices such that the prices along with that allocation constituted an equilibrium.

Subsequently, Debreu pointed out to Scarf that the assumption of identical core allocations for consumers of the same type, far from being restrictive, was a simple consequence of strict concavity of the utility function. He also suggested a substantial simplification of the central proof, which was incorporated into a joint paper [9]. This paper became the starting point for a large literature on the relationship between the core and the set of equilibria in economies with large numbers of consumers. One of the most notable contributions to this literature was that of Robert Aumann. Having heard Scarf discuss his original paper at a conference at Princeton in 1962, Aumann (1964) produced a model with a continuum of consumers in which the core and the set of equilibrium allocations were the same.

In 1963, Scarf moved to the Cowles Foundation at Yale, which he had visited in 1959–60. He has remained at Cowles, except for visiting appointments at Rome, Cambridge, and Stanford, since then. The atmosphere at Cowles is ideally suited to Scarf. As he described it in the introduction to his book on computation of equilibria [12, p. x]: "The standards of mathematical rigor and clarity of thought which prevail at Cowles are well known to the economics profession. But perhaps more important is the persistent though subtle suggestion that the highest aim of even the most theoretical work in economics is an ultimate practical applicability."

## Computation of Economic Equilibria

After arriving at Yale, Scarf returned to the problem of finding a method for computing economic equilibria. His work on the core had suggested a possible approach: if he could find a method for calculating core allocations, then this method would serve to approximate equilibrium allocations, at least when the number of consumers was large. Scarf had been able to find an algorithm for finding allocations in the core of exchange economies with three consumers, and he was working on extending this procedure to more general situations.

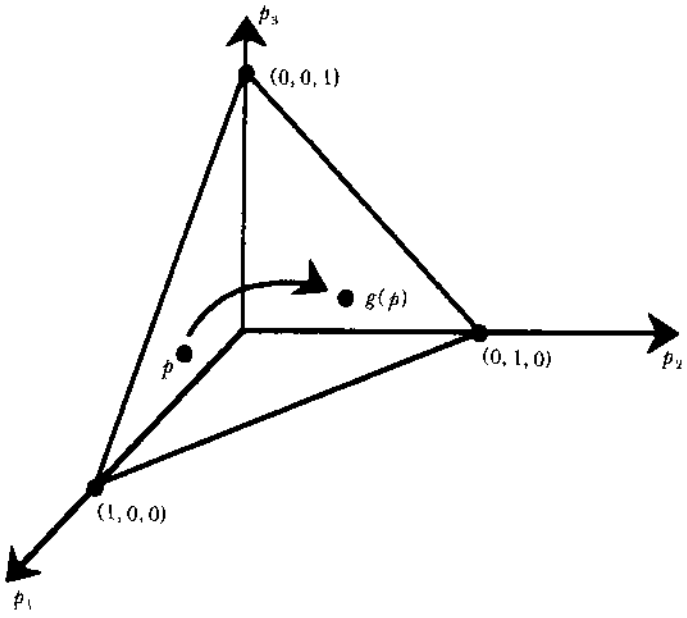
An exchange economy can be thought of as a cooperative game. The general specification of such a game considers all possible subsets, or *coalitions*, of agents and the sets of outcomes available to them. In an exchange economy, for example, a coalition is a subset of the set of consumers and the available outcomes are the utilities that can be generated by distributing the total endowment of the coalition among its members. An assignment of outcomes is in the core of the game if there is no coalition that can attain an outcome that is preferred by all of its members.

Scarf had produced a simple example of a game with three agents with nonconcave utility functions that had no core, so he knew that additional assumptions had to be imposed to guarantee existence of an outcome. The condition that he used was that the game be "balanced." This condition is complicated to enunciate; it is attractive, however, because it is satisfied by any exchange economy with concave utility functions considered as a cooperative game. Scarf was able to prove the existence of an outcome in the core of a balanced cooperative game. His approach relied on Brouwer's fixed point theorem, however, which, at that time, he considered an approach too abstract ever to lead to a constructive algorithm.

The breakthrough in finding an algorithm for calculating outcomes in the core occurred in 1965 when Aumann, who was visiting Cowles at the time, showed Scarf a recent paper by Carlton Lemke and Joseph Howson (1964). This paper provided an algorithm for calculating Nash equilibria of two-agent noncooperative games. Although the application was very different from the one that Scarf was studying, this paper excited Aumann and Scarf because, up until that time, the only argument for the existence of Nash equilibria in this type of game had relied on the sort of fixed point theorem that Scarf wanted to avoid. Using the fundamental insight of Lemke and Howson's paper, Scarf was able to develop an algorithm for approximating an outcome in the core of a balanced cooperative game [10].

Having found an algorithm for the core, Scarf realized that he could approximate equilibrium prices directly, without relying on the relationship of the core and equilibria. His approach was to approximate fixed points, and to understand how it works, it is necessary to understand a bit about Brouwer's fixed point theorem. Brouwer's theorem guarantees that certain systems with the same number of equations and unknowns have solutions. Specifically, this theorem deals with functions defined on the *unit simplex*, the set of points whose components are nonnegative and sum to one. Suppose that the value of such a function $g(p) = (g_1(p), \ldots, g_n(p))$ also lies in the simplex. Then we say that $g$

*Figure 1*
**Mapping the Unit Simplex into Itself**




maps the simplex into itself because $g$ associates any point $p$ in the simplex with another point $g(p)$ also in the simplex as depicted in Figure 1. Brouwer's fixed point theorem says that if $g$ maps the simplex into itself and is continuous, then it has a fixed point $g(\hat{p}) = \hat{p}$. In other words, there is a solution to the system of $n$ equations and $n$ unknowns, $g_j(p_1, \ldots, p_n) = p_j$.

It requires some work to move from knowing that a fixed point exists for an arbitrary function $g$ to knowing that an equilibrium exists for an economy with an aggregate excess demand function $f$. We start by using the homogeneity of degree zero of excess demand to impose the normalization that prices should sum to one, to restrict price vectors to the unit simplex. We then modify the function formed by adding prices to the corresponding excess demand, $g(p) = p + f(p)$, to obtain a function with the following properties: First, $g(p)$ is in the unit simplex whenever $p$ is. Second, a price vector $\hat{p}$ is an equilibrium for the economy, whenever it is a fixed point of $g$.[2]

[2]Consider the function

$$g_j(p) = \frac{\max[p_j + f_j(p), 0]}{\max[p_1 + f_1(p), 0] + \cdots + \max[p_n + f_n(p), 0]}, \qquad j = 1, \ldots, n.$$

It is easy to use Walras's law to show that $g$ is continuous whenever $f$ is and that $\hat{p} = g(\hat{p})$ if and only if $f(\hat{p}) \leq 0$.

Although Brouwer's theorem can be used to prove the existence of equilibrium for an economy, Scarf had originally tried to avoid using it in his work on computation because the traditional proofs of this theorem, of which there were several, said that a fixed point existed but said nothing about how to find it. In 1966 Scarf developed an algorithm for computing equilibria based on a procedure for calculating fixed points [11].

The idea behind this algorithm is fairly simple. The researcher starts by specifying a finite, but large, grid of points on the simplex. The algorithm considers a set of $n$ such points that are close together and asks if these points come close enough to satisfying the conditions for a fixed point in that the values of the function evaluated at these points are, in a specific sense, close to the points themselves. If the set of points does not satisfy this property, then the algorithm moves to an adjacent set of points formed by all but one of the previous points and one new point and asks the question again. It terminates when it encounters an approximate fixed point. The algorithm has two key features: The first is that at any step, except at the start and the end, the algorithm can always carry out the replacement operation of moving from one set of points to another by dropping a point in the first set and adding an additional point. The second key feature is that at any step, except at the start and the end, the set of points under consideration has exactly two adjacent sets of points to which it could move, one of which it has just visited and the other that it will visit at the next step.

The algorithm must terminate with an approximate fixed point in a finite number of steps. There are only a finite number of points in the grid and, consequently, only a finite number of sets of points that the algorithm could consider. The only way the algorithm could fail to converge would be to cycle; it cannot cycle because then some set of points would be the first such set to be considered twice; and this cannot be because the algorithm would have already visited the adjacent sets of points. This argument that the algorithm must terminate is the insight that Scarf derived from Lemke and Howson's paper, which is otherwise different in its mechanics and interpretation. Scarf likes to illustrate Lemke and Howson's insight with a story [12, p. 48]:

> Let us imagine a house consisting of a finite number of rooms, each of which has precisely two doors. Assume that one of the rooms has a door leading to the outside of the house. Then there must be at least one other door leading to the outside! And that other door may be found by this simple rule: Begin with the known outside door and proceed from room to room, never departing from a room by the door used in entering it. One can never return to a room previously entered.

For the algorithm to be efficient, it must be able to march through the simplex considering only a tiny fraction of the total number of points in the grid. For the algorithm to be easy to use, it must be easy to specify the grid of points and easy to carry out the replacement operation needed to move from one set of points to another. Scarf originally specified the sets of points to be considered using a concept called *primitive sets*, that he had used in his algorithm for computing an outcome in the core. In 1967, a student in his mathematical economics class at Yale, Terje Hansen, found that, if the grid of points is specified as all points of the form $(m_1/D, \ldots, m_n/D)$, where $m_1, \ldots, m_n$ are nonnegative integers that sum to $D$, then the concept of primitive set can be specified so that the replacement operation can be performed with a few simple additions and subtractions.

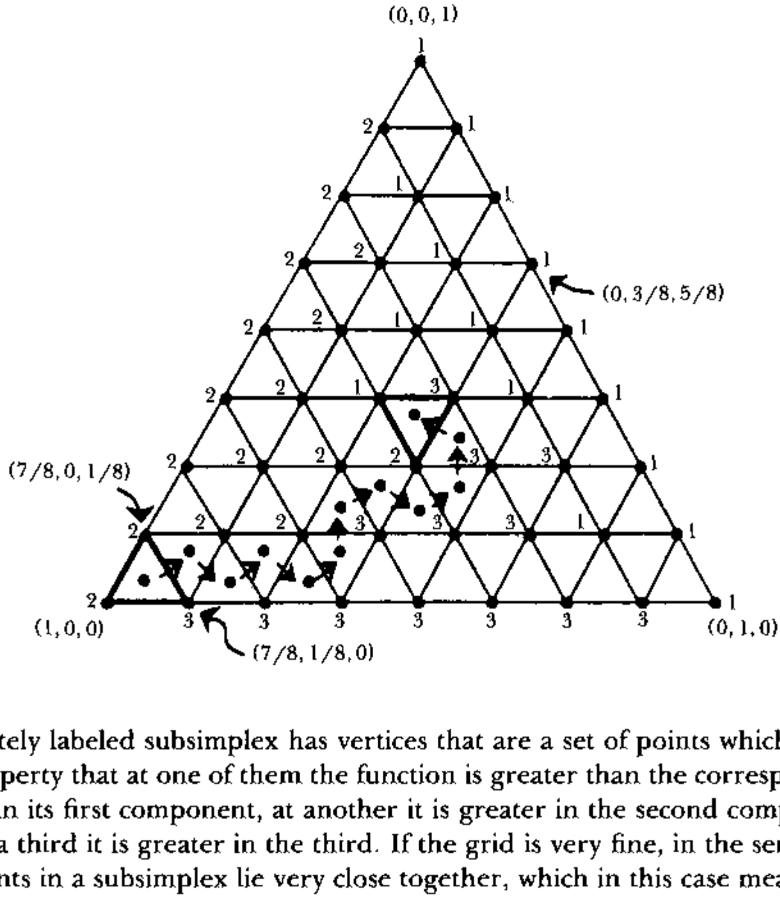
In 1968, Scarf received an unpublished paper from Harold Kuhn (1968) that proposed an alternative approach to approximating fixed points based on a traditional concept from combinatorial topology, the simplical subdivision of the simplex, rather than on primitive sets. Scarf and Kuhn were intrigued by the similarity of the replacement operation discovered by Hansen and the analogous step in Kuhn's algorithm. Discussing this similarity and drawing a few pictures, they came to realize that the two algorithms were virtually identical, although their initial geometric interpretations had differed.

To make the algorithm more concrete, consider the simplex pictured in Figure 2. This simplex is divided into subsimplices, which in this case are small triangles whose vertices are the grid of points of the form $(m_1/8, m_2/8, m_3/8)$, where $m_1$, $m_2$, and $m_3$ are nonnegative integers such that $m_1 + m_2 + m_3 = 8$. Each one of these points has a label that is the number 1, 2, or 3. A point in the grid on the boundary is labelled 1 if $m_1 = 0$, 2 if $m_2 = 0$, and 3 if $m_3 = 0$; thus, for example, the point $(0, 3/8, 5/8)$ receives the label 1 as pictured. On the corners, where more than one label is possible according to this rule, we follow some fixed rule, like always choosing the first possible label. If the grid points inside the simplex are each given an arbitrary label 1, 2, or 3, then there is a classic combinatorial puzzle: is there necessarily a subsimplex, a little triangle, whose vertices have all of the labels 1, 2, and 3? The answer is yes, and Scarf's algorithm provides a method for finding such a completely labeled subsimplex, one whose vertices have the three labels.

To see the connection between a completely labeled subsimplex and fixed points, let us suppose that the labels of grid points interior to the simplex are not arbitrary at all, but instead depend on the function $g$: we give a point $v$ a label $i$ for which $g_i(v) \geq v_i$. For example, suppose that we are considering the grid point $v = (6/8, 1/8, 1/8)$ and that when we evaluate the function $g$ at $v$ we obtain $g(6/8, 1/8, 1/8) = (0.23, 0.41, 0.36)$. (Remember that if we are calculating the equilibrium of a model economy, $g(v)$ depends on the aggregate excess demands evaluated at the price vector $v$.) Since $g_2(v) = 0.41 \geq 1/8 = v_2$, we can give $v$ the label 2, as in Figure 2.[3] Given this labeling rule, we see that a

[3]Notice that, since $g_3(v) \geq v_3$, we could have given $v$ the label 3; but we apply the convention of using the first applicable label. Notice too that this way of assigning labels is consistent with the labels on the boundary of the simplex since $g_j(v)$ is always nonnegative, and $v_j = 0$, and $g_j(v) \geq v_j = 0$.

*Figure 2*
**Scarf's Algorithm for Computing Fixed Points**




completely labeled subsimplex has vertices that are a set of points which satisfy the property that at one of them the function is greater than the corresponding vertex in its first component, at another it is greater in the second component, and at a third it is greater in the third. If the grid is very fine, in the sense that the points in a subsimplex lie very close together, which in this case means that $D$ is very large, the continuity of $g$ would imply that at any point $p$ in the completely labeled subsimplex the inequalities $g_j(p) \geq p_j$ are all almost satisfied. Since the components of $g(p)$ and $p$ both sum to one, however, this says that $g(p)$ is almost equal to $p$. Hence $p$ can serve as an approximate fixed point, which means that it can serve as an approximate equilibrium of the underlying model economy.

The crucial steps to finding an approximate fixed point are to choose a fine enough grid of points and to find the desired subsimplex. Scarf's algorithm provides a method for finding the desired subsimplex. As in Figure 2, it starts in the corner of the simplex, where the labels on the boundary guarantee that there is a subsimplex with vertices that have the labels 2 and 3. It then moves into the simplex, into the house from the door to the outside in Scarf's tale, by considering the unique other subsimplex that shares with the first its interior side, which has vertices with the labels 2 and 3. It examines the new vertex to see if the labeling rule assigns it the label 1. If it does, the algorithm stops with a completely labeled subsimplex. If it does not, the algorithm finds the unique new side of the subsimplex that has vertices with the labels 2 and 3, the other door out of the room, and moves through it to the adjacent subsimplex. By always moving through a side of a subsimplex whose vertices have the labels 2 and 3, and never moving through a side with the labels 2 and 2 or 3 and 3, the algorithm respects the rule that every room has precisely two doors and it guarantees that if it encounters a vertex with label 1, then it has found a completely labeled subsimplex. It is easy to see that the only way that the algorithm would try to move outside the simplex would be if we were to return to the starting point, where there is a subsimplex with a side on the boundary of the simplex with vertices with the labels 2 and 3. This cannot happen, however, since the algorithm can never cycle. Consequently, we must encounter a completely labeled subsimplex.[4]

This whole procedure can be readily extended to higher dimensions, to economies with many goods. In fact, Scarf's book [12] contains a Fortran program for the algorithm that is only 27 lines long. What is needed is that the user supply a subroutine that provides a label for any new vertex encountered by the algorithm. This label, of course, requires the evaluation of the excess demand function of the specific model economy that the user is studying.

An obvious drawback of Scarf's algorithm is that how close its solution comes to satisfying the equilibrium conditions depends on how close the points in the completely labeled subsimplex are to each other, which in turn depends on how large $D$ is. Unfortunately, the closer the points in any subsimplex are, the more work that the algorithm has to do. Furthermore, if we compute an answer where the grid is too coarse—in the sense that the points are too far apart to provide a good approximation to an equilibrium—then refining the grid requires us to discard the information obtained in the previous try and to start all over again from the corner of the simplex.

Scarf's work inspired a line of papers in operations research devoted to finding more efficient ways of overcoming this problem. Early users of Scarf's

[4]The crucial replacement step is trivial to carry out. The second triangle encountered by the algorithm in Figure 2, for example, has as vertices the vectors

$$\begin{bmatrix}7\\1\\0\end{bmatrix}, \quad \begin{bmatrix}7\\0\\1\end{bmatrix}, \quad \begin{bmatrix}6\\1\\1\end{bmatrix}$$

divided by 8. Since the third vertex has the label 2, we move to a new simplex by keeping the first and third vertices and dropping the second vertex, which also has the label 2. Elementary geometry tells us that we can do this by completing a parallelogram, adding the first and third vertices and subtracting the second,

$$\begin{bmatrix}7\\1\\0\end{bmatrix} + \begin{bmatrix}6\\1\\1\end{bmatrix} - \begin{bmatrix}7\\0\\1\end{bmatrix} = \begin{bmatrix}6\\2\\0\end{bmatrix}.$$

algorithm attempted to overcome the problem by using the approximate fixed point as a starting point for some other numerical technique (like Newton's method, which is not guaranteed to work but usually does well if started near a solution). Later, Curtis Eaves (1972), Orin Merrill (1971), and Gerard van der Laan and Adolphus Talman (1980) developed modifications of Scarf's algorithm that allowed the grid to be continually refined during the course of the procedure and for solutions of arbitrarily high accuracy to be found without having to restart at a corner. Two of Scarf's students at Yale—Michael Todd and Ludo van der Heyden—also contributed to this operations research literature.[5]

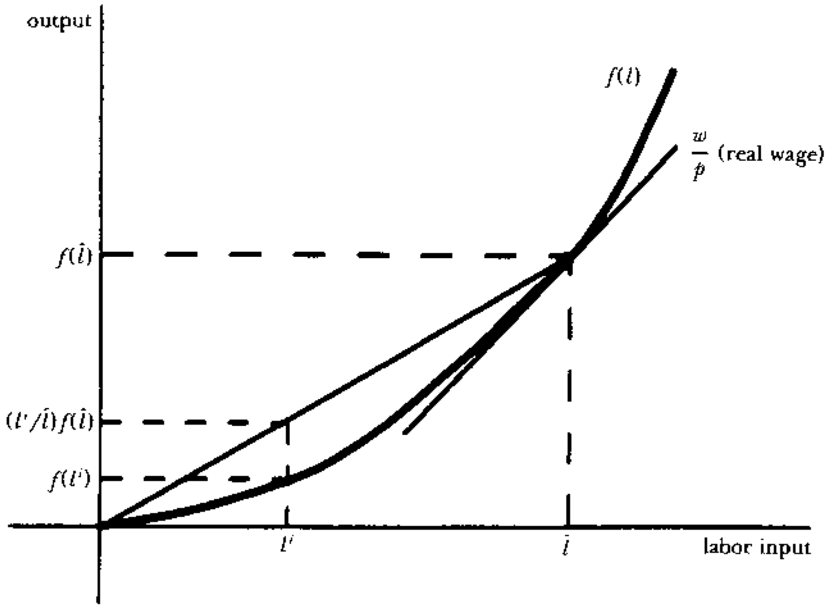
Most of Scarf's students at Yale over the period 1968–80 did research not on theoretical topics dealing with the algorithm, however, but on its applications. Indeed, there is what many would characterize as a Yale school of economists who use applied general equilibrium models to do economic policy analysis. Students of Scarf in this group include Andrew Feltenstein, Timothy Kehoe, Ana Martirena-Mantel, Marcus Miller, Donald Richter, Jaime Serra-Puche, John Shoven, John Spencer, and John Whalley. Shoven and Whalley, two of the leaders in this field, are responsible for a pair of definitive surveys of applied general equilibrium methods (Shoven and Whalley, 1984, 1992). Serra-Puche has combined an academic career with public service in his home country of Mexico. He is currently Secretary of Trade and Industrial Development there and has been directly responsible for much of the internal deregulation and open trade policies of the Mexican government, including its recent negotiations with the governments of Canada and the United States over the North American Free Trade Agreement.

## Increasing Returns and Integer Programming

In the course of his work on the core with Debreu, Scarf became interested in economies with production. Debreu and he realized that their proof of convergence of the core to the set of competitive allocations would easily extend to economies where all coalitions had access to the same constant-returns-to-scale production technology. Scarf also realized that, if the production technology exhibits increasing returns, then the economy may not have a competitive equilibrium. He was able to construct some examples, however, in which the economy has a nonempty core even though it has no competitive equilibrium.

[5]When Bruce Kellogg, Tien-Yien Li, and James Yorke (1976) and Steven Smale (1976) provided an alternative approach to finding fixed points based on solving a system of differential equations, Scarf and Eaves [13] wrote a paper that established the intimate relation between this approach and that based on simplical subdivisions. This paper also established the relation between Scarf's work and the then emerging literature on the differentiable approach to general equilibrium of such researchers as Debreu (1970), Egbert Dierker (1972), and Andreu Mas-Colell (1975).

*Figure 3*
**A Two-good Economy with Increasing Returns**




Consider the example depicted in Figure 3. There is an economy with two goods, a consumption good $y$ and inelastically supplied labor $l$. The technology for producing the consumption good using labor is specified by a production function $f(l)$ that exhibits increasing returns in the sense that average productivity $f(l)/l$, the slope of the line between the origin and $(l, f(l))$, is increasing in labor input. There are $m$ consumers who value the consumption good alone and have endowments $l^i$ of the labor. Consider now the allocation where all of the labor $\bar{l} = l^1 + l^2 + \cdots + l^m$ is used to produce the consumption good, which is then divided among the consumers in proportion to their labor input, $y^i = (l^i/\bar{l})f(\bar{l})$. To be in the core this allocation must have the property that no coalition can attain higher levels of consumption, and therefore of utility, for its members using its members' endowments. It is easy to see that this allocation is in the core, and the intuition is simple and seemingly general: the coalition of all the consumers can achieve total output higher than that available to any other coalition precisely because there are increasing returns.

This sort of example led Scarf to conjecture that increasing returns could only make existence of an allocation in the core more likely than it was in the constant-returns case, and with constant returns he was able to prove existence using a simple extension of his argument for exchange economies. He speculated that cooperative game theory might provide the tools for studying economies with increasing returns. It was apparent that standard general equilibrium theory floundered in such economies: The economy depicted in

Figure 3 has no competitive equilibrium because at the aggregate production plan $(\bar{l}, f(\bar{l}))$ there does not exist a price for consumption and a wage such that this production plan is profit maximizing. Indeed, given the real wage ratio depicted in the figure, the production plan is profit *minimizing* among efficient points, since all efficient points except $f(\bar{l})$ lie above the line tangent to $f(l)$ at $f(\bar{l})$.

Scarf explored this approach to analyzing economies with increasing returns in a set of notes [17] written in 1963 but not published until 1986. In these notes, he established conditions sufficient for existence of an allocation in the core, but he also developed a procedure for constructing examples with empty cores. Specifically, he proved that, for any production set that exhibits increasing returns, one can construct a group of consumers specified by utility functions and endowments so that the corresponding cooperative game has an empty core. In fact, the procedure applies to the example just given; we merely need to allow some endowments of the consumption goods and some utility for leisure. Scarf's notes were widely circulated in unpublished form. His positive results on sufficient conditions for existence led to a literature on coalition production economies, with notable contributions by Claude Oddou (1976), William Sharkey (1979), and Tatsuro Ichiishi and Martine Quinzii (1983). Scarf himself turned his attention to the theory of optimization in production problems with increasing returns and indivisibilities.

In focusing on the mathematics of optimization, Scarf was inspired by the development of activity analysis and linear programming theory in the 1940s and 1950s. Researchers such as George Dantzig and Tjalling Koopmans had developed a theory for analyzing production problems with constant returns that changed the way that economists thought about the role of prices in determining economic efficiency. To get some flavor of this theory, consider an economy that has two alternative techniques for producing output using labor. The first produces four units of output using four of labor; the second produces one unit of output using two of labor. Suppose that there are six units of labor available and that the two techniques, or activities, can be run at any (nonnegative) levels. Because the second activity produces more output per unit of labor, for every amount of labor input, it makes sense to use only this activity. The output maximizing production plan is to run the second activity at level $x_2 = 1.5$ and not to run the first activity at all, $x_1 = 0$.[6]

The duality theorem of linear programming offers a simple check for efficiency: A proposed production plan is efficient if and only if there exist prices for the outputs and inputs such that activities employed at positive levels in the plan make zero profit and all other activities earn profits less than or equal to zero. In this case, the price of output, taken as numeraire, is equal to one and the price of labor is one. At these prices the first activity makes

[6]More particularly, the linear programming problem whose solution determines the efficient use of resources is max $x_1 + 4x_2$, subject to $2x_1 + 4x_2 \leq 6$, and $x_1, x_2 \geq 0$.

negative profits and the second activity makes zero profits. A corollary of the duality theorem gives a simple test for new candidates for efficient activities: If a new activity makes positive profits at the old prices, it should be incorporated into the efficient production plan; if it does not make positive profits, it can be ignored. If someone in our model economy invents a new process that produces three units of output using two of labor, for example, it should necessarily be incorporated into the efficient production plan.

Results related to the duality theorem—the separating hyperplane theorem, the Kuhn-Tucker theorem, and the welfare theorems—are central to modern economic theory. Scarf was concerned that there did not seem to be analogous results for economies with increasing returns and indivisibilities in production. Specifically, he was concerned with production models, like the model discussed above, where activity levels are constrained to be integers. In this case, which is called an *integer programming problem,* there is no simple test, like the pricing test discussed above, to verify whether a production plan is optimal. The earlier situation is so simple that the solution to the integer programming problem can be found by examining each of the feasible points, those points with integer coordinates. It is $x_1 = 1$, $x_2 = 1$. It is easy to verify that there are no prices such that both activities earn zero profit.

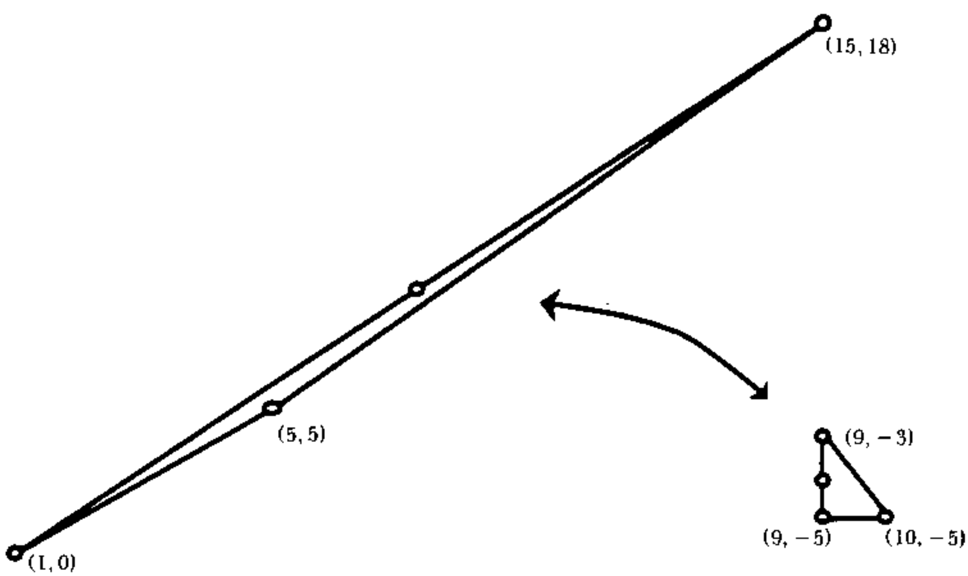
Scarf's old friend Ralph Gomory had developed the first algorithms for solving integer programming in 1958. A number of alternative algorithms have been proposed over time, but most are unpredictable and unreliable: a small change in one of the parameters can change an easy-to-solve problem into an intractable one. In the language of the theory of complexity developed by computer scientists in the 1970s, the integer programming problem is an *NP complete problem*, which means that this sort of problem is a member of a large class of mathematical problems for which there is no known algorithm that is guaranteed to solve the problem in a number of arithmetical steps that is a polynomial function of the amount of data necessary to specify the problem. Furthermore, these problems are all equivalent in the sense that a polynomial algorithm for one problem could be translated into a polynomial algorithm for any other NP complete problem.

The development of an efficient algorithm for solving integer programming problems would be a major breakthrough in at least three areas: in applied mathematics, where it would provide an approach for solving a wide variety of related problems; in operations research, where it would provide a method for solving many important problems faced by large firms; and in economics, where it could provide insights into the organization of firms and into decision making within these firms. In an article in this issue, Scarf [25] presents an extensive discussion of the importance of integer programming in economics.

Scarf has been interested in the theory of integer programming since the early 1960s and has worked on it almost exclusively since the late 1970s. One of his goals has been to develop a simple test for optimality of a production plan, analogous to the pricing test in the linear programming case. Using the same concept of primitive sets that had proved useful in his earlier work on the core and on computation of equilibria, Scarf was able to develop such a test. The grid of points on which these primitive sets are defined are *lattice points*, those points with integer coordinates, the only points from which the solution to an integer programming problem can be drawn. As in Scarf's work on fixed points, a solution to an integer programming problem is associated with a primitive set whose vectors are completely labeled; here the labeling rule depends on which constraints of the problem are satisfied at the lattice point. Unlike the case with his work on fixed points, however, Scarf found that the primitive sets appropriate for an integer programming problem were not regularly shaped analogues of triangles. Instead, they could have more vertices and be very long in some dimensions and very thin in others. In general, a primitive set is a collection of lattice points that are vertices of a convex polyhedron that does not contain another lattice point. The number of these vertices and the precise shape of the polyhedron depend on the specific integer programming problem. The interested reader should consult the series of articles [14, 15, 16] that Scarf has written on this subject.

Scarf has been working since the mid-1980s on developing algorithms for solving integer programming problems. His approach has led him to the study of the geometry of numbers, a branch of mathematics developed by the Polish mathematician Hermann Minkowski in the 1890s [19, 23, 24]. In the geometry of numbers there is a classic problem that is relevant to the determination of

*Figure 4*
**A Unimodular Transformation**



whether or not a collection of lattice points forms a primitive set in Scarf's algorithm: When does a given convex body contain a lattice point? Consider, for example, the question of whether the triangle in two dimensional space with vertices $(1, 0), (5, 5), (15, 18)$ contains a lattice point. One approach to answering this problem would be to conduct an exhaustive search of all the integer vectors $(h_1, h_2)$ whose values satisfy $1 \leq h_1 \leq 15$ and $0 \leq h_2 \leq 18$. (Actually, this problem involving a triangle in two dimensional space is easy to analyze, but it serves as an illustration of problems involving general convex polyhedra in $n$ dimensional space.)

Another approach is to first find a linear change of variable that moves the vertices of the triangle closer together and to then conduct the exhaustive search. Such a change of variables should be a *unimodular transformation*, that is, the change of variable and its inverse should transform lattice points into lattice points. A simple unimodular transformation, shown in Figure 4, transforms the triangle with vertices $(1, 0), (5, 5), (15, 18)$ into one with vertices $(9, -5)$, $(10, -5), (9, -3)$.[7] We now have to search over the transformed integer vectors $(h'_1, h'_2)$ whose values satisfy $9 \leq h'_1 \leq 10$ and $-5 \leq h'_2 \leq -3$. Doing so yields the vector $(9, -4)$ that lies in the triangle. Since the change of variables is linear, we know this point corresponds to one in the original triangle; since the change of variable is unimodular, we know that this point corresponds to a lattice point. The lattice point contained in the original triangle is $(8, 9)$.

In the above example we find a unimodular transformation such that the first coordinates of all of the points are close to each other. To search for a point in the triangle, we first choose $h'_1 = 9$ and search over values of $h'_2$ and then choose $h'_1 = 10$ and repeat the search. The general approach in $n$ dimensions is analogous: We find a unimodular transformation that allows us to reduce a problem in $n$ dimensional space to a small number of $n - 1$ dimensional problems. Doing this successively, reducing each of the $n - 1$ dimensional problems into a small number of $n - 2$ dimensional problems, and so on, gives a decision tree structure to the original problem. If the original problem has a large dimension, then there are a large number of branches in the decision tree. The essential ingredient to this approach is to find a fast way to come up with the unimodular transformations that allow us to do the branching. Since computations along separate branches are independent, there are substantial opportunities for parallel processing.

[7]The change of variables is

$$h'_1 = 9h_1 - 7h_2$$

$$h'_2 = -5h_1 + 4h_2.$$

Since the coefficients of this transformation are integers, it transforms lattice points into lattice points. Since the matrix of coefficients has determinant 1, its inverse also has coefficients that are integers. A unimodular transformation is defined by a matrix of integer coefficients with determinant 1 or $-1$.

Independently of Scarf, a number of applied mathematicians have been taking similar approaches to solving integer programming problems. One of them, Hendrick Lenstra (1983), has been able to find an algorithm that can solve integer programming problems with a fixed number of variables in a polynomial number of steps. Although fixing the number of variables in advance means that this is not a polynomial algorithm for the NP complete problem, it suggests that there may be fairly general and efficient algorithms for integer programming problems.

Recently, Scarf has been working with a group of mathematicians and computer scientists to implement an integer programming algorithm that can efficiently solve problems with up to 100 integer variables [20, 22]. The results so far are promising. Scarf [18, 25] speculates that this research could lead us to a new understanding of the theory of the firm and of bounded rationality. This speculation might seem somewhat far-fetched if we did not have such solid evidence of the power of Scarf's mathematical intuition and his ability to focus it on fundamental economic problems.

■ *The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.*

## References


**Arrow, Kenneth J., Theodore Harris, and Jacob Marschak,** "Optimal Inventory Policy," *Econometrica*, July 1951, *19*, 250–72.

**Aumann, Robert J.,** "Markets With a Continuum of Traders," *Econometrica*, January-April 1964, *32*, 39–50.

**Blinder, Alan, S.,** "Retail Inventory Investment and Business Fluctuations," *Brookings Papers on Economic Activity*, 1981, 2, 443–505.

**Caplin, Andrew S.,** "The Variability of Aggregate Demand With $(S, s)$ Inventory Policies," *Econometrica*, November 1985, *53*, 1395–1409.

**Debreu, Gerard,** "Economies With a Finite Set of Equilibria," *Econometrica*, May 1970, *38*, 387–92.

**Debreu, Gerard,** "Excess Demand Functions," *Journal of Mathematical Economics*, March 1974, *1*, 15–23.

**Dierker, Egbert,** "Two Remarks on the Number of Equilibria of an Economy," *Econometrica*, September 1972, *40*, 951–53.

**Dvoretzky, Aryeh, Jack Kiefer, and Jacob Wolfowitz,** "On the Optimal Character of the $(S, s)$ Policy in Inventory Theory," *Econometrica*, October 1953, *21*, 586–96.

**Eaves, B. Curtis,** "Homotopies for Computation of Fixed Points," *Mathematical Programming*, August 1972, *3*, 1–22.

**Ichiishi, Tatsuro, and Martine Quinzii,** "Decentralization for the Core of a Production Economy With Increasing Returns," *International Economic Review*, June 1983, *24*, 397–412.

**Iglehart, Donald,** "Optimality of $(S, s)$ Policies in the Infinite Horizon Dynamic Inventory Problem," *Management Science*, January 1963, *9*, 259–67.

**Karlin, Samuel,** "Optimal Inventory Policy for the Arrow-Harris-Marschak Dynamic Model." In Arrow, Kenneth J., Samuel Karlin, and Herbert Scarf, eds., *Studies in the Mathematical Theory of Inventories and Production*. Stanford: Stanford University Press, 1958, 135–54.

**Kellogg, R. Bruce, Tien-Yien Li, and James Yorke,** "A Constructive Proof of the Brouwer Fixed-Point Theorem and Computational Results," *SIAM Journal on Numerical Analysis*, September 1976, *13*, 473–83.

**Kuhn, Harold W.,** "Simplicial Approximation of Fixed Points," *Proceedings of the National Academy of Sciences of the U.S.A.*, December 15, 1968, *61*, 1238–42.

**Laan, Gerard van der, and Adolphus J. J. Talman,** "An Improvement of Fixed Point Algorithms by Using a Good Triangulation," *Mathematical Programming*, May 1980, *18*, 274–85.

**Lemke, Carlton E., and Joseph T. Howson,** "Equilibrium Points of Bi-Matrix Games," *SIAM Journal of Applied Mathematics*, June 1964, *12*, 413–23.

**Lenstra, Hendrick W.,** "Integer Programming With a Fixed Number of Variables," *Mathematics of Operations Research*, November 1983, *8*, 538–48.

**Mantel, Rolf R.,** "On the Characterization of Aggregate Excess Demand," *Journal of Economic Theory*, March 1974, *7*, 348–53.

**Mas-Colell, Andreu,** "On the Continuity of Equilibrium Prices in Constant-Returns Production Economics," *Journal of Mathematical Economics*, March 1975, *2*, 21–33.

**Merrill, Orin H.,** "Applications and Extensions of an Algorithm that Computes Fixed Points of Certain Non-Empty Convex Upper Semi-Continuous Point to Set Mappings," Technical Report No. 71-7, Department of Industrial Engineering, University of Michigan, 1971.

**Oddou, Claude,** "Theoremes D'Existence et D'Equivalence Pour des Economies Avec Production," *Econometrica*, March 1976, *44*, 265–81.

**Sharkey, William W.,** "Existence of a Core When There Are Increasing Returns," *Econometrica*, July 1979, *47*, 869–76.

**Shoven, John B., and John Whalley,** "Applied General Equilibrium Models of Taxation and International Trade: An Introduction and Survey," *Journal of Economic Literature*, September 1984, *22*, 1007–51.

**Shoven, John B., and John Whalley,** *Applying General Equilibrium*. Cambridge: Cambridge University Press, 1992.

**Shubik, Martin,** "Edgeworth Market Games." In Tucker, A. W., and R. D. Luce, eds., *Contributions to the Theory of Games, IV*. Princeton: Princeton University Press, 1959, 267–78.

**Smale, Steven,** "A Convergent Process of Price Adjustment and Global Newton Methods," *Journal of Mathematical Economics*, July 1976, *3*, 107–20.

**Sonnenschein, Hugo,** "Do Walras' Identity and Continuity Characterize the Class of Community Excess Demand Functions?" *Journal of Economic Theory*, August 1973, *6*, 345–54.