3-2000

Multi-period capacity expansion in globally dispersed regions

Siddhartha Syam
Marquette University, siddhartha.syam@marquette.edu

Decision Sciences, Vol. 31, No. 1 (March 2000): 173-195. DOI.

Multiperiod Capacity Expansion in Globally Dispersed Regions

Siddhartha S. Syam
College of Business Administration, P.O. Box 1881, Marquette University, Milwaukee, WI 53201-1881, email: syam@mail.busadm.mu.edu

Abstract

Multiperiod capacitated location (MCL) models specify where and when capacity expansions should be made, and how large they should be. The MCL model developed in this paper incorporates a shift from manufacturing for overseas markets to manufacturing in overseas markets. Computational results are given for problems involving up to 200 locations/destinations and 10-year planning horizons. Near-optimal solutions are provided in reasonable computing times with average convergence less than 2%. Representative variations in cost between regions are simulated in the test problems, and the managerial implications of alternative diversification strategies are also assessed.

Subject Areas

Distribution/Logistics, Lagrangian Relaxation, and Location.

INTRODUCTION

At the turn of the millennium, the rapid growth of a global economy is indisputable. International trade in consumer goods and industrial items has increased greatly in recent decades. This is brought out by the following facts: Between 1963 and 1979, real merchandise export growth averaged 11.8% per year, and between 1979 and 1991, it averaged 4.4% per year (Markusen, Melvin, Kaempfer, & Maskus, 1995). Concurrent with this trend, another megatrend in world economic history is creating new paradigms for manufacturing and trade. This trend is the rapid industrialization and growth of previously underdeveloped regions of the world known collectively as emerging markets (U.S. Department of Commerce, 1995). The tendency toward rapid global dispersion of manufacturing capacity is marked. To illustrate, multinational firms based in the United States own approximately US$65 billion in foreign production facilities (Markusen et al.). In the early 1990s, approximately 37,000 corporations owned about 170,000 foreign affiliates, and about a third of world exports is intrafirm trade between parents and affiliates (Kerr & Perdikis, 1995).

Facility location and multiperiod capacitated location (MCL) are very important topics in operations management and operations research. Much of the research in the area has focused on the advancement of computational techniques. However, there are two drawbacks: (1) Recent changes in the global environment have not been incorporated into these models and (2) There has been limited development of methods for solving MCL models, particularly in techniques for providing lower bounds (Jacobsen, 1990). In order to address these concerns, this paper introduces extensions to MCL modeling that are stated below.

A global limit and regional limits on the numbers of sites (i.e., locations where capacity expansions are permitted) are introduced, and the related geoeconomic context is extensively discussed. The MCL model is made more flexible by allowing economies of scale, diseconomies of scale, as well as constant costs in manufacturing. As a methodological extension, an efficient solution technique, which provides both upper and lower bounds, is developed. The methodology can accommodate both increasing and decreasing rates of demand for goods and services.

The rest of the paper is organized in the following manner: The next two sections contain a succinct review of relevant research and a discussion of the proposed extensions and their motivations. Subsequently, the extended MCL model is discussed, and an efficient heuristic solution methodology based on Lagrangian relaxation (Fisher, 1985) is presented. The final sections contain the details of computational testing, important managerial implications, and concluding thoughts.

RELATED RESEARCH

References in this section are limited to relevant models in the extensive facility location and MCL literature. A primary reference model in facility location is the uncapacitated (single-period) facility location (UFL) problem, which is one of determining the optimal locations for plants given certain fixed costs and variable cost rates. Capacity restrictions on the plants provide the capacitated version of the problem, and a limit on the number of facilities leads to the p-median problem. Erlenkotter (1978) has provided an extremely efficient linear programming dual-- based heuristic for the UFL problem.

The history of the MCL problem goes back to studies by Manne (1967) and others in India in the mid-1960s. MCL models extend facility location problems to issues of when to expand capacity and in what increments, in addition to the primary question of where to locate facilities. Necessarily, MCL models are harder to solve than single-period location problems. Initial research focused on finding the optimal sequence of capacity expansions at a single location, often assuming an increasing demand pattern. Manne and others showed that these expansions could be described in terms of a regeneration point theorem, which essentially says that the capacity expansions occur at points of time when the excess capacity (over demand) is zero. However, the regeneration point theorem does not hold in the case of multiple locations (Jacobsen, 1990).

Solution methodologies for the MCL problem include gradient methods (Rao & Ruthenberg, 1977) and dynamic programming (Jacobsen, 1990). Van Roy and Erlenkotter (1982) presented a multiperiod version of the UFL problem which, although highly efficient in terms of computing time, has the limitations of being uncapacitated and assuming that the variable capacity cost is zero. Heuristics involving the so-called fixed cost formulation of the MCL problem have been proposed by Hung and Rikkers (1974) and Fong and Srinivasan (1981). These heuristics attempt to achieve economies by consolidating locations, as well as consolidating capacity at a given location. Other approaches have included Luss (1979), who studied a model for two facility types, and Jacoby and Loucks (1972), who used simulation to examine expansion policies for water resource systems.

EXTENSIONS TO MCL MODELING

Although MCL models have had a long history, recent developments suggest the need for newer extensions. Some of these extensions are incorporated here and discussed below.

Limits on Sites

In 1995, Toyota Motor Corporation, Japan's leading automobile manufacturer, announced its Global Business Plan (Toyota, 1998) which, with respect to North America, called for greater localization of vehicle production, higher production capacity for existing plants, and an increased degree of local parts and material purchasing. Similarly, Honda motor company (Honda, 1998) declared that its corporate philosophy centers on the primary objective of manufacturing vehicles in the markets where they are sold. Five out of seven Hondas sold in North America are now produced here, and the proportion of North American-produced Toyota vehicles sold locally was 60% in 1996. Ford Motor Company is increasing its manufacturing capability worldwide; for instance, it has partnered with a leading engineering and automotive firm in India (Ford, 1998). Examples in other industries may be readily found; for instance, Coca-Cola produces soft drinks for local markets in Asia.

The examples cited above are part of a shift of large magnitude-a paradigm shift-from merely exporting overseas to producing overseas. The reasons for this shift are both economic and political. To illustrate, in the early 1980s American consumers greatly increased their purchases of high-quality automobiles from Japan, leading to a strong political backlash in the United States. Japanese automakers have responded with a marketing strategy of substantial local expansion, using the "Made in America" logo to blunt the criticism. In general, multinational firms worldwide have had to strike a balance between two opposing factors:

a. the desirability of producing overseas because of cost considerations, the size of emerging markets, and the necessity of cultivating political goodwill in foreign markets;
b. the significant risks of overseas investment, which include substantial loss of goodwill at home, loss of technological advantage, loss of facilities, poor product and service quality, and significant loss of value when markets and currencies collapse.

Using the fundamental principles of risk management, manufacturers can apply regional diversification as a strategy for limiting their exposure to the risks listed above. In this paper, regional limits on sites, based on growth trends, cost data, regional volatility, etc., are introduced as a direct way to incorporate regional diversification. It is assumed that regional limits have been predetermined; therefore, they are treated as fixed data in the MCL model. The p-median problem introduces a limit on the number of open facilities in single-period facility location problems. Similarly, in addition to regional limits on sites, a global limit on the number of sites is imposed (sites are locations at which capacity expansions are allowed).

Manufacturing Economies of Scale

MCL formulations have traditionally assumed that the variable costs of production are strictly linear in flow. This runs counter to (1) the common assumption in economics (Samuelson, 1980) of economies of scale, which is that average unit costs decrease as output is increased, at least up to some large level of output, and (2) the usual MCL assumption that capacity expansion costs are subject to economies of scale. In this paper the variable costs of labor and manufacturing are modeled as piecewise linear concave functions of output at a location. Diseconomies of scale (i.e., increasing unit costs) and constant costs are also allowed by the model.

CAPACITY EXPANSION MODEL

In common with other MCL models, the cost of capacity expansion is considered to be concave in the size of expansion. Also, as in other discrete MCL formulations, realistic plant capacity expansion options are viewed as essentially "lumpy" (Stevenson, 1996); that is, available in discrete and often large increments, rather than as continuous increases. Therefore, the model is restricted to a discrete number (three) of levels (large, medium, and small) of capacity expansion at a particular location in a particular year. The number of output levels is also discrete (three), corresponding to large, medium, and small capacity. In keeping with the concave cost assumption, unit expansion cost is a stepwise decreasing function, and total expansion cost is a piecewise linear concave increasing function of the level of expansion. The parameters and variables that are used in the MCL model are described next, followed by the model.

Model Parameters

Number of destinations = N,
Number of potential locations = M,
Number of years = I,
Number of capacity expansion levels = 3 (large, medium, and small),
Number of output level categories= 3 (large, medium, and small).

Problem Parameters

F_{ikm} = Fixed cost of kth size expansion in year i at location m,
a_{jm} = Unit manufacturing cost for output category j at location m,
b_{jm} = Unit labor cost for output category j at location m,
α_{jm} = Initial output breakpoint for output category j at location m,
β_{jm} = Terminal output breakpoint for output category j at location m,
t_{mn} = Transportation cost of shipping one unit from location m to destination n,
D_{in} = Demand in year i at destination n,
C_{0m} = Initial capacity level at location m,
C_{im} = Capacity level at location m in year i,
z_{ikm} = Magnitude of kth size expansion in year i at location m,
g_d = Regional limit on number of open facilities in region d,
$r(d)$ = Set of potential locations in region d,
p = Global limit on number of open facilities.

Variables

y_{ikm} = 0-1 variable, equal to 1 if an expansion of size k occurs at location m in year i, 0 otherwise,
w_m = 0- 1 variable, equal to 1 if at least one capacity expansion occurs at location m in the planning horizon, 0 otherwise,
v_{ijm} = 0-1 variable, equal to 1 if total output corresponds to category j in year i at location m, 0 otherwise,

x_{ijmn} = Output in output category j in year i at location m for destination n.

Model(P)

The model is discussed in detail below. To assist with interpretation, a brief heading precedes each constraint.

$$\text{Minimize } \sum_{i=1}^{I} \sum_{k=1}^{3} \sum_{m=1}^{M} F_{ikm} y_{ikm} + \sum_{i=1}^{I} \sum_{j=1}^{3} \sum_{m=1}^{M} \sum_{n=1}^{N} (a_{jm} + b_{jm} + t_{mn}) x_{ijmn},$$

subject to

$$\text{Capacity limits: } \sum_{j=1}^{3} \sum_{n=1}^{N} x_{ijmn} \leq C_{im} \quad \forall\, i, m,$$

(1)

$$\text{Redefine capacity limits: } C_{im} = C_{(i-1)m} + \sum_{k=1}^{3} z_{ikm} y_{ikm} \quad \forall\, i, m,$$

(2)

$$\text{Meet all demand: } \sum_{j=1}^{3} \sum_{m=1}^{M} x_{ijmn} = D_{in} \quad \forall\, i, n,$$

(3)

$$\text{Initial breakpoint (variable cost): } \sum_{n=1}^{N} x_{ijmn} \geq \alpha_{jm} v_{ijm} \quad \forall\, i, j, m,$$

(4)

$$\text{Terminal breakpoint (variable cost): } \sum_{n=1}^{N} x_{ijmn} \leq \beta_{jm} v_{ijm} \quad \forall\, i, j, m,$$

(5)

$$\text{One cost category for location, year: } \sum_{j=1}^{3} v_{ijm} = 1 \quad \forall\, i, m,$$

(6)

$$\text{One expansion limit for location, year: } \sum_{k=1}^{3} v_{ikm} \leq 1 \quad \forall\, i, m,$$

(7)

$$\text{Expansion only at site: } y_{ikm} \leq w_m \quad \forall\, i, k, m,$$

(8)

Global limit on sites: $\sum_{m=1}^{M} w_m \leq p,$

(9)

Regional limits on sites: $\sum_{m \in r(d)} w_m \leq g_d \quad \forall d,$

(10)

$$x_{ijmn} \geq 0 \text{ and integer } \forall\, i,j,m,n;\, y_{ikm} \in \{0,1\} \forall i,k,m.$$
$$v_{ijm} \in \{0,1\} \forall\, i,j,m;\, w_m \in \{0,1\} \forall\, m.$$

The model proposed is quite general in that every destination is also assumed to be a potential location, although it can be easily modified if the set of destinations is distinct from the set of locations. The first term of the objective function contains the costs of capacity expansions, and the second term represents the sum of manufacturing, labor, and transportation costs. The first constraint limits the outflow at any location in any year to the available capacity at that location in that year, and the second constraint defines this capacity to be the sum of the capacity in the previous year and the capacity added in the current year. Constraint (3) ensures that the demands at all destinations in all years are fully met. The fourth and fifth constraints provide, for each location, the breakpoints for each variable cost category. These breakpoints are the initial and terminal volumes of output at which the cost rate corresponding to the category is applicable. The sixth constraint ensures that only one of these categories is applicable at a location in a given year. Constraint (7) guarantees that at most one capacity expansion option is availed at a location in a given year, and constraint (8) makes sure that a capacity expansion option is exercised at a location only if the location is one of the p sites permitted. Finally, constraint (9) limits the number of sites to the parameter p, and the tenth constraint restricts the number of sites in any region to a regional limit.

The model has $3 \bullet I \bullet M \bullet N$ ordinary integer variables, and $6 \bullet / \bullet M + M\ 0-1$ integer variables. The number of constraints is $13 \bullet / \bullet M + I \bullet N + D + 1$. The expansion costs F_{ikm} are computed as $f_{ikm} z_{ikm}$, where f_{ikm} is the unit cost of the *kth* level expansion in year i at location m. Concavity of expansion cost is imposed by ensuring that $f_{i1m} > f_{i2m} > f_{i3m}$, providing piecewise linear expansion costs. Similarly, manufacturing and labor costs at a location are made piecewise linear concave in output by imposing the restrictions that $a_{1m} > a_{2m} > a_{3m}$, and $b_{1m} > b_{2m} > b_{3m}$.

SOLUTION METHODOLOGY

It is known that the single-period uncapacitated facility location model (Cornuejols, Nemhauser, & Wolsey, 1990) and, by extension, the multiperiod capacitated version of the problem, are NP-hard. It may be computationally infeasible to search for the optimal solution to the MCL problem, which is classified as one of combinatorial optimization. Heuristic methods are often the most practical for this class of problems. A heuristic methodology based upon Lagrangian relaxation (Fisher, 1985) is proposed in this paper. It involves lower and upper bounding procedures.

As a first step, constraints (1) and (2) are reformulated and consolidated into a single constraint as follows:

$$\sum_{j=1}^{3} \sum_{n=1}^{N} x_{ijmn} \leq C_{0m} + \sum_{k=1}^{3} \sum_{s=1}^{i} z_{skm} y_{skm} \quad \forall i,m.$$

(11)

Lower Bounding Technique

Lagrangian duals are obtained by dualizing (i.e., moving) complicating constraints from the constraint space into the objective function, multiplied by dual multipliers that penalize their absence from the constraint space. The objective function value obtained by solving the relaxed problem provides, for fixed multiplier values, a lower bound on the optimal objective function value of the original problem. If the multipliers are systematically updated, the lower bounds converge progressively (but not necessarily monotonically) to the optimal objective function value *of the Lagrangian dual. In the absence of a duality gap, this value is equal to the* optimal objective function value of the primal problem (Geoffrion, 1974).

The MCL model proposed here contains several complicating constraints. The first step in solving the problem is to relax the capacity constraint (11), which is equivalent to relaxing (1) and (2). Dualizing this constraint, associating multipliers μ_{im} with it, and defining the coefficient $G_{jmn} = a_{jm} + b_{jm} + t_{mn}$, the following Lagrangian dual problem (D) is obtained:

Maximize $\varphi > (\mu)$: $\mu \geq 0$, where $\varphi(\mu)$ is the following model (D1):

$$\text{Minimize} \sum_{i=1}^{I} \sum_{k=1}^{3} \sum_{m=1}^{M} F_{ikm} y_{ikm} + \sum_{i=1}^{I} \sum_{j=1}^{3} \sum_{m=1}^{M} \sum_{n=1}^{N} G_{jmn} x_{ijmn} + \sum_{i=1}^{I} \sum_{m=1}^{M} \mu_{im} \left(\sum_{j=1}^{3} \sum_{n=1}^{N} x_{ijmn} - C_{0m} - \sum_{k=1}^{3} \sum_{s=1}^{i} z_{skm} y_{skm} \right)$$

subject to constraints (3) through (10).

The objective function of (D1) can be further consolidated by combining coefficients:

$$\sum_{i=1}^{I} \sum_{k=1}^{3} \sum_{m=1}^{M} \left\{ F_{ikm} y_{ikm} = \mu_{im} \sum_{s=1}^{i} z_{skm} y_{skm} \right\} - \sum_{i=1}^{I} \sum_{m=1}^{M} \mu_{im} C_{0m} + \sum_{i=1}^{I} \sum_{j=1}^{3} \sum_{m=1}^{M} \sum_{n=1}^{N} \left(G_{jmn} + \mu_{im} \right) x_{ijmn},$$

By using coefficients $Q_{ijmn} = G_{jmn} + \mu_{im}, H_{ikm} = (F_{ikm} - \sum_{s=1}^{I} \mu_{sm} z_{ikm})$, and by further consolidating the coefficients of the y_{ikm} coefficients, it is seen that the objective function becomes equivalent to:

$$\sum_{i=1}^{I} \sum_{k=1}^{3} \sum_{m=1}^{M} H_{ikm} y_{ikm} + \sum_{i=1}^{I} \sum_{j=1}^{3} \sum_{m=1}^{M} \sum_{n=1}^{N} Q_{ijmn} x_{ijmn} - \sum_{i=1}^{I} \sum_{m=1}^{M} \mu_{im} C_{0m}.$$

The problem is still very difficult to solve because of the complicating breakpoint constraints (4) and (5). In order to make the problem more tractable, constraints (4) and (5) are relaxed. Using non-negative multipliers λ_{ijm} and γ_{ijm} for (4) and (5) respectively, the following Lagrangian dual problem is obtained:

Maximize $\varphi(\mu, \lambda, \gamma)$: $\mu, \lambda, \gamma \geq 0$, where $\varphi(\mu, \lambda, \gamma)$ is the following model (D2):

Minimize

$$\sum_{i=1}^{I}\sum_{j=1}^{3}\sum_{m=1}^{M}\sum_{n=1}^{N} Q_{ijmn}x_{ijmn} + \sum_{i=1}^{I}\sum_{j=1}^{3}\sum_{m=1}^{M} \lambda_{ijm}\left(\alpha_{jm}v_{ijm} - \sum_{n=1}^{N} x_{ijmn}\right) + \sum_{i=1}^{I}\sum_{j=1}^{3}\sum_{m=1}^{M} \gamma_{ijm}\left(\sum_{n=1}^{N} x_{ijmn} - \beta_{jm}v_{ijm}\right) + \sum_{i=1}^{I}\sum_{k=1}^{3}\sum_{m=1}^{M} H_{ikm}y_{ikm} - \sum_{i=1}^{I}\sum_{m=1}^{M} \mu_{im}C_{0m},$$

subject to constraints (3), (6), (7), (8), (9), (10), and the integer restrictions on the x_{ijmn}, y_{ikm}, and the v_{ijm} variables. Next, coefficients $E_{ijmn} = (Q_{ijmn} - \lambda_{ijm} + \gamma_{ijm})$, and $R_{ijm} = (\alpha_{jm}\lambda_{ijm} - \beta_{jm}\gamma_{ijm})$ are consolidated.

Model (D2) is separable into subproblems (SI), (S2), and (S3), which are shown next. (S1), (S2), and (S3) have naturally integer-valued solutions. Accord- ing to Geoffrion (1974), this means that they possess the integrality property, which implies that the bounds from them cannot improve upon those obtained from linear programming relaxations of model (P). However, the subproblems are solvable by extremely fast greedy methods as discussed below, resulting in the rel- atively quick solution times reported in the computational results.

Model (S1)

$$\text{Minimize} \sum_{i=1}^{I}\sum_{k=1}^{3}\sum_{m=1}^{M} H_{ikm}y_{ikm},$$

subject to

$$\sum_{k=1}^{3} y_{ikm} \leq 1 \quad \forall\, i, m,$$

(12)

$$y_{ikm} \leq w_m \quad \forall\, i, k, m,$$

(13)

$$\sum_{m=1}^{M} w_m \leq p,$$

(14)

$$\sum_{m \in r(d)} w_m \leq g_d \quad \forall\, d,$$

(15)

$$y_{ikm} \in \{0,1\} \forall i, k, m; w_m \in (0,1)\ \forall\, m.$$

Model (S2)

$$\text{Minimize} \sum_{i=1}^{I}\sum_{j=1}^{3}\sum_{m=1}^{M}\sum_{n=1}^{N} E_{ijmn}x_{ijmn},$$

subject to

$$\sum_{j=1}^{3} \sum_{m=1}^{M} x_{ijmn} = D_{in} \;\; \forall\, i, n,$$

(16)

$$x_{ijmn} \geq 0 \text{ and integer } \;\; \forall\, i, j, m, n.$$

Model (S3)

$$\text{Minimize} \sum_{i=1}^{I} \sum_{j=1}^{3} \sum_{m=1}^{M} R_{ijm} v_{ijm},$$

subject to

$$\sum_{j=1}^{3} v_{ijm} = 1 \quad \forall\, i, m,$$

(17)

$$v_{ijm} \in \{0,1\} \;\; \forall\, i, j, m.$$

Solution procedures for subproblems

(S1), (S2), and (S3) are solved using greedy algorithms (Nemhauser & Wolsey, 1988). Subproblem (S2) involves only the x_{ijmn} variables, while (S3) involves only the x_{ijm} variables. Both are linear integer knapsack problems and, as such, can be optimally solved by standard greedy algorithms (Nemhauser & Wolsey). The computational complexity of all three subproblems are provided in the Appendix. Algorithm (L1) for (S1) is based on its optimal properties and the constraints of the model. The algorithm and the optimal properties of (S1) are contained in the Appendix.

To improve efficiency, the following constraint, which is implied by constraints (7) and (11) was imposed and found to be quite effective in reducing the gap between lower and upper bounds.

$$\sum_{j=1}^{3} \sum_{n=1}^{N} x_{ijmn} \leq C_{0m} + \sum_{s=1}^{i} z_{s3m} \;\; \forall\, i, m.$$

(18)

The constraint is used to restrict outflows in (S2). It is even more effective if the global and regional limits on the numbers of sites are imposed. Denoting the optimal objective function value of (S2) as LB2, and the optimal objective function value of (S3) as LB3, a lower bound for the primal problem is provided by the sum of LBI, LB2, and LB3, less $\sum_{i=1}^{I} \sum_{m=1}^{M} \mu_{im} C_{0m}$.

Upper Bounding Heuristic

The Lagrangian solution methodology for (P) involves both lower and upper bounds on the optimal objective value of the problem. In order to obtain a good upper bound, a heuristic solution to the problem is developed. The basis for the heuristic is a computation, for each location, of the minimum cost that is achievable at the location, assuming that it is the sole source of supply. Based upon an ordering of these minimum costs, capacity expansions at locations are heuristically assigned. Subsequently, minimum costs flows are determined by a greedy procedure for the resulting networks in each year. Finally, the solution is modified to ensure that all prevailing cost categories are not violated, and a grand total of costs encompassing all assigned sites and all years is computed. The procedure, labeled (UB), is invoked during problem initialization. It is described in the Appendix.

Algorithm to Solve Problem (P)

The MCL problem is solved by a systematic application of the upper and lower bounding techniques discussed above. The lower bounding method is iterative, with subproblems (S1), (S2), and (S3) solved at each iteration, using updated values of the multipliers. The dual function $\phi(\mu, \lambda, \gamma)$ is piecewise linear concave in the multipliers, and the method of subgradient optimization (Held, Wolfe, & Crowder, 1974) is used to maximize it. In the absence of a duality gap, the subgradient method converges, in the limit, to the optimal value of the primal problem (Polyak, 1967). The algorithm for problem (P) is shown next, while the updating of multipliers is discussed below.

Algorithm CAPEXP

Step 1

Initialization: Check for problem feasibility. Initialize the multipliers to zero, the incumbent lower bound (ILB) to $-\infty$, and set a tolerance ε for convergence between lower and upper bounds. Fix an iteration limit, and conduct procedure UB to obtain the upper bound.

Step 2

Solve the lower bound problem by solving subproblems (S1), (S2), and (S3). The lower bound is then given by $\text{LB} = \left(\text{LBl} + \text{LB2} + \text{LB3} - \sum_{i=1}^{I} \sum_{m=1}^{M} \mu_{im} C_{0m}\right)$. Update the incumbent lower bound if necessary: If LB > ILB, then ILB = LB.

Step3

Check for convergence. If (UB - ILB)/UB ≤ ε, terminate, since the solution is acceptable. Otherwise, use subgradient optimization to update the multipliers. If the number of iterations is less than the iteration limit, return to Step 2, otherwise stop.

Updating of Multipliers

A subgradient of the function $\phi(\mu, \lambda, \gamma)$, at a particular iteration, is a vector consisting of the following elements:

$$sg_{im}^{11} = \left(\sum_{j=1}^{3} \sum_{n=1}^{N} x_{ijmn} - C_{0m} - \sum_{k=1}^{3} \sum_{s=1}^{i} z_{skm} y_{skm}\right) \forall\, i, m.$$

$$sg_{ijm}^{17} = \left(\alpha_{jm} v_{ijm} - \sum_{n=1}^{N} x_{ijmn}\right) \forall\, i, j, m.$$

$$sg_{ijm}^{18} = \left(\sum_{n=1}^{N} x_{ijmn} - \beta_{jm} v_{ijm}\right) \forall\, i, j, m.$$

The values of x_{ijmn}, v_{ijm}, and y_{skm} are the values obtained from the lower bound procedure at the iteration. Let the multipliers at iteration t be μ_{im}^t, λ_{ijm}^t and γ_{ijm}^t, respectively. The step size at iteration t is given by $\rho_t = \delta^t \bullet \ (\text{UB} - \text{ILB})/\|\eta^t\|^2$, where δ^t is a scalar between 1 and 2, and η^t is the vector of subgradients at iteration t. Convergence results for this step size are provided in Polyak (1967). At the end of iteration t, using a step size of ρ^t, the multipliers are updated according to the following formulae: $\mu_{im}^{t+1} = \mu_{im}^t = \rho^t \bullet sg_{im}^{11}$, $\lambda_{ijm}^{t+1} = \lambda_{ijm}^t = \rho^t \bullet sg_{im}^{17}$ and $\gamma_{ijm}^{t+1} = \gamma_{ijm}^t + \rho^t \bullet sg_{im}^{18}$, subject to the non-negativity requirement on the multipliers. The updated multipliers are used to generate the lower bound problem for the next iteration.

COMPUTATIONAL TESTING

This section provides the details of two computational experiments that were conducted on a Pentium-II personal computer running at 300 megahertz with 128 megabytes of memory. The objectives were to verify that the solution methodology is effective and to assess the effects of varying certain parameters of the model.

Data Generation

A representative spread in manufacturing, capacity expansion, and labor costs between five global regions is estimated using data available in the Statistical Abstract of the World (Gale Research, 1994), which also contains secondary references on global data. Using the 1990 data provided, the interregional spread in manufacturing and capacity expansion costs is estimated to be about 15:1, and about 75:1 in labor costs (the cheapest and most expensive regions included are Asia and North America, respectively). Further evidence of a wide disparity in labor costs between regions may be found in Roth (1994); for instance, the average hourly manufacturing wage rate in 1993 in Germany was $24.87, in the United States, $16.40, and $0.54 in Russia.

At each location, the largest capacity expansion in a given year is uniformly distributed between 2 and 2.5 times the medium expansion, while the small expansion is uniformly distributed between 0.6 and 0.8 times the medium expansion. Economies of scale (i.e., concavity of costs) are introduced by making the unit costs of large-scale output (and expansion) 0.90 times the unit costs of small-scale levels, and medium-scale unit costs equal to 0.95 times those at small-scale levels. Regional cost differences are captured by multiplying certain base costs by regional multipliers, which are uniformly distributed within the relevant cost spread.

Demand at a location is uniformly distributed between 1,000 and 5,000 units in the first year. In subsequent years, demand at a location depends on whether it is classified as belonging to the increasing demand category or the decreasing demand category. If demand is increasing, annual demand in a given year (after the first) is 1.5 times the previous year's demand, whereas if demand is decreasing, annual demand is 0.9 times last year's demand. Although these growth rates are not typical of all situations, certain goods and services are known to have very high growth rates in demand, particularly in industrializing countries. As an instance, the computer software market grew at an annual rate of 42% in Brazil and 55% in Portugal during 1993-1995 (U.S. Department of Commerce, 1997).

The base costs are as follows: Unit manufacturing cost is uniformly distributed between $10 and $5, unit labor cost is uniformly distributed between $7 and $2, unit expansion cost (small expansion) is uniformly distributed between $2 and $1, and unit transportation cost is $10. At a given location, the initial output breakpoint for the medium cost category is equal to average demand, and the initial breakpoint for the low cost category is equal to the average demand multiplied by the ratio of the large capacity expansion to the medium expansion.

Computational Experiments and Results

The parameters of the computational experiment are shown in Table 1. The first column identifies the problem set, and column 2 gives the number of locations, which is equal to the number of destinations. The third column shows the number of years in the capacity planning horizon. The fourth column provides the total number of locations in which capacity expansions are permitted, while column 5 shows the similar limit for each of the five regions.

Ten problems were solved for each of the configurations shown in Table 1. The computational results for experiment A are shown in the next table. The lower bounding procedure is set to terminate if either (a) problem convergence is less than 1%, (b) it has conducted 150 iterations, or (c) it has conducted 40 iterations without improvement in the lower bound. Column 1 in Table 2 identifies the problem set, and the second column provides the average convergence between upper and lower bounds, where convergence is $C = 100 \bullet$

(upper bound - lower bound)/ upper bound. Column 3 shows the minimum convergence and column 4 provides the maximum convergence. Columns 5, 6, and 7 show the average, minimum, and maximum computational times, respectively (in seconds).

The results provided in Table 2 show that the solution methodology achieves near-optimal results for the problems tested, with convergence gags in the neighborhood of 1 % for many problem sets. The average convergence gap for all problems combined is 1.63%. The smaller problems are solved relatively quickly, averaging 42 seconds for problems with 50 location/destinations and three-year planning horizons. Very large problems involving 200 location/destinations and 10-year planning horizons have an average solution time of about 3,000 seconds.

MANAGERIAL IMPLICATIONS

It has been noted that regional diversification is a primary strategy of leading firms such as Honda, which have already declared their intent to manufacture products in the markets where they are sold. A diversification strategy is an effort to either spread the risk or seek political goodwill by not concentrating all sites in the lowest cost regions of the world. While diversification can be quite effective from a marketing viewpoint, managers need cost information in order to make an informed decision. The effects of alternative strategies that companies may pursue are examined here.

First, an experiment in which the base scenario involves 100 locations, a three-year planning horizon, and no limits on sites, is conducted. A global limit of 20 sites is imposed in the second scenario, and additional regional limits of four sites per region are imposed in the third scenario. It is found that the average cost increase, relative to the base case is, with only a global limit, about 4.11% and, with additional regional limits, about 14.20%. The primary effect of the regional limits is to ensure that solutions do not exclude high-cost regions such as North America, particularly since they are integral to very successful marketing strategies ("Made in America"). It is also found that a risk-averse or conservative approach that excludes volatile regions altogether, leads to cost premiums averaging 300%.

The risk-versus-cost trade-off can be made specific if numerical risk indexes are constructed for particular regions. Such risk indexes should reflect managerial and/or expert risk assessments regarding regional attributes such as instability, inflation risk, currency-exposure risk, etc. To illustrate the use of risk indexes, consider two scenarios, both involving five regions, 40 locations, and a global limit of 20 sites. The regions are shown in decreasing order of unit costs and increasing order of risk. The first scenario (X) assumes a narrower degree of spread in cost as well as risk relative to the second scenario (Y). A number of models with varying regional limits are applied to the two scenarios.

Table 1: Parameters for Experiment A.

Problem Set	Number of Locations/ Destinations	Number of Years in Expansion Horizon	Global Limit on Locations with Capacity Expansions	Regional Limit on Locations with Capacity Expansions
1	50	3	16	8
2	50	5	16	8
3	50	10	16	8
4	75	3	25	13
5	75	5	25	13
6	75	10	25	13
7	100	3	33	17
8	100	5	33	17

9	100	10	33	17
10	200	3	66	33
11	200	5	66	33
12	200	10	66	33

Table 2: Computational results, Experiment A.

Problem Set	Average Percent Convergence	Minimum Percent Convergence	Maximum Percent Convergence	Average Seconds	Minimum Seconds	Maximum Seconds
1	2.14	1.72	3.56	42	31	53
2	1.04	0.90	1.41	26	12	83
3	1.35	0.81	2.14	138	25	177
4	1.57	1.32	1.87	56	53	58
5	0.96	0.84	1.03	29	13	99
6	0.88	0.56	1.51	87	18	209
7	1.63	1.50	1.68	114	106	124
8	2.16	0.82	6.91	131	17	220
9	2.77	0.68	7.50	289	119	479
10	1.63	1.38	1.77	704	617	847
11	1.82	1.62	1.92	1220	787	1982
12	1.73	1.56	1.99	3004	2156	3989

Table 3: Risk indexes: Scenarios X and Y.

	Region				
Scenario	I	II	III	IV	V
X	2	5	10	15	20
y	2	8	15	30	50

Note: A lower risk index corresponds to smaller risk.

Table 4: Regional limits: Models A, B, C, D.

	Region				
Model	I	II	III	IV	V
A	20	20	20	20	20
B	7	5	3	3	2
C	10	5	2	2	1
D	12	8	0	0	0

Model A is the base model which represents the riskiest strategy, permitting the concentration of all sites in any region. B, C, and D represent alternative diversification strategies with an increasing tendency to concentrate sites in low-risk regions. Further, consider two additional scenarios, U and W. Scenario U assumes the same risk indexes as Y, but incorporates a strong "learning" effect (from previous or similar endeavors), which leads to sharply reduced marginal rates of increase in cost (about half of the rates previously assumed) as production is increased. Scenario W has the same parameters as U, but involves a change in corporate policy that permits a marked increase in the maximum capacity permitted at any single site (about three times that previously assumed). The risk indexes in Scenarios X and Y are shown in Table 3, and the Models A, B, C, and D are shown in Table 4.

The results of solving the models are shown in Figures 1, 2, and 3 as risk and cost combinations. The vertical axes show the percentage increase in cost over the base model (A). The horizontal axes provide the sum of risk indexes corresponding to the sites selected in a particular solution. The connected combinations provide the approximate equivalent of an "efficient frontier" similar to those found in financial analysis. All points above the efficient frontier should be excluded on the basis of unnecessary cost or risk. It should be noted that combinations that are excluded under one scenario may become viable under another (as in the area between the plots in Figure 1).

Figure 2 reveals that corporate learning may lead to beneficial effects in terms of smaller increases in cost when relocating sites to lower risk (higher cost) regions, particularly at sites in the regions with smallest risk. This is because differences in cost are marginal in the most volatile (but cheapest) regions but more substantial in stable (but expensive) regions. Combining corporate learning with large capacity as in scenario W provides very interesting results. It can be seen in Figure 3 that the scenario W plot occasionally dips into negative territory, implying that sites with lower risk may correspond to lower total cost.

The result appears to be counterintuitive, since low-risk regions are also associated with higher costs, but it is actually logical in certain circumstances. These circumstances are the simultaneous presence of (i) significant corporate learning, which leads to "flattened" marginal cost rates at high volumes of output; and (ii) higher capacities (whether by design or by accident) in lower risk sites, compared to higher risk sites. In this situation, a lower risk site may produce high enough volume to incur smaller total cost relative to a higher risk site. This effect is more probable when the cost rate and risk differentials between the particular regions involved are not very large.

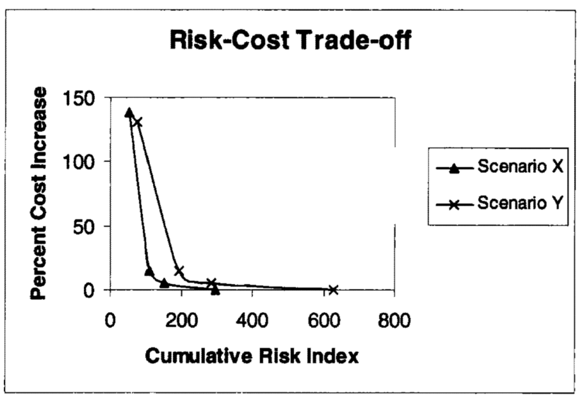

Figure 1: Risk versus cost: X and Y.

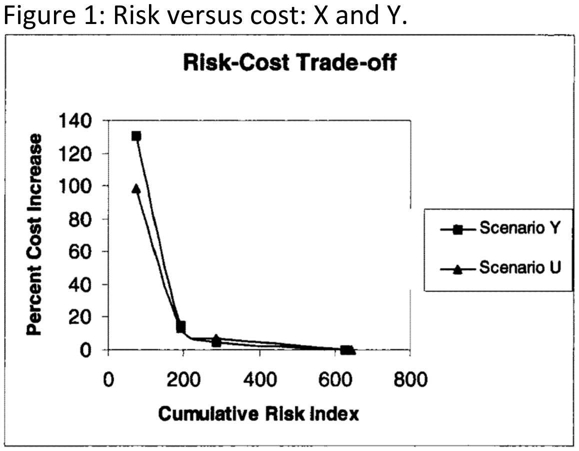

Figure 2: Risk versus cost: Y and U.

Finally, a possibility that is not explored here is that it may be feasible to introduce revenue functions which, in conjunction with the efficient frontier, may lead to an optimal location/expansion strategy similar to an optimal portfolio in financial analysis. Considerations of space and scope preclude the full investigation of this potential managerial application of the modeling approach developed in this paper.

SUMMARY

A model for multiperiod capacitated location (MCL) that incorporates recent trends in facility location has been presented in this paper. The model imposes a global limit and regional limits on the numbers of open sites. As another extension to MCL modeling, the model permits economies of scale, diseconomies of scale, and constant costs in manufacturing and labor expenditure. In common with other multiperiod models, setup costs are modeled as concave and piecewise linear in the size of the expansion. The model is discrete and permits three levels of capacity increases-large, medium, and small. Unlike other MCL studies, demand at a particular location is allowed to follow either an increasing trend or a decreasing trend.

The solution methodology developed is based on Lagrangian relaxation. A computational experiment shows that the methodology finds good or near-optimal solutions in reasonable solution times. Unlike previous MCL research, the methodology generates both lower bounds and upper bounds, providing a measure of the quality of the solution in the form of the convergence gap between upper and lower bounds. Convergence gaps for the experiment average less than 2%.

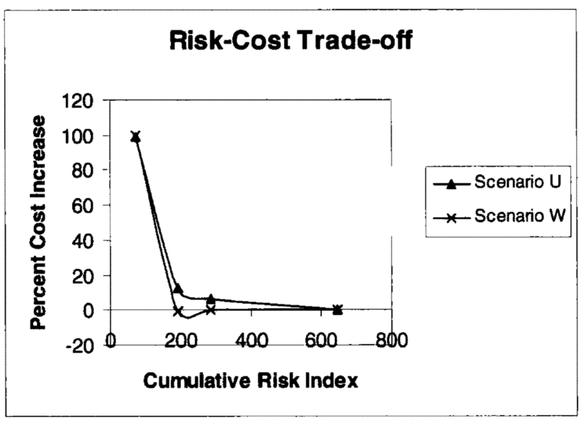

Figure 3: Risk versus cost: U and W.

Alternative diversification strategies for managers are investigated, and the idea of a cost-risk efficient frontier is explored. Tests confirm that risk reduction strategies, which limit the numbers of sites permitted in particular regions, and conservative diversification strategies that avoid certain regions altogether are often associated with significant cost premiums. While the cost premiums may be substantial, managers need to weigh them against potential political benefits and risk factors when making location or expansion decisions. [Received: April 4, 1997. Accepted: July 9, 1999.]

APPENDIX

Optimal Properties of Model (S1)

Property 1: $H_{ikm} \geq 0 \Rightarrow y_{ikm} = 0$

Proof

Let Z^* be the optimal objective value, with $y_{ikm} = 1$ in the optimal solution, S^*. Then, since no constraint prohibits this, let $y_{ikm} = 0$, with Z^{**} the corresponding objective value. Then $Z^{**} < Z^*$, contradicting the optimality of Z^*.

Property 2: Define an eligible variable as a y_{ikm} variable, which (i) has a negative coefficient and (ii) can equal 1 without violating any constraint. Given a set of eligible variables, the objective function is minimized by selecting the variable with the most negative coefficient to be 1.

Proof

(a) By similar reasoning to Property 1, at least one of the eligible variables should be 1. (b) Let y' and $y"$ be two eligible variables, and H' and $H"$ be the corresponding coefficients, with $H' < H"$. Let Z' be the objective function value, with $y' = 1$, $y" = 0$, and vice-versa for $Z"$, all other variables being equal. By the assumption that $H' < H"$, $Z' < Z"$. The property then immediately follows from similar pair- wise comparisons between all eligible variables.

Algorithm (Ll) and Procedure (UB)

Algorithm LI

[The steps of the algorithm follow directly from properties 1 and 2 and the constraints of model (S1).]

Step 1

Set $\mathrm{LB1} = 0$. Sort the coefficients H_{ikm} in ascending order. Set a counter $f = 0$. Initialize all the Y_{ikm} variables to 0. Set location counter $e = 0$.

Step2

Do Until

$f > 3 \bullet I \bullet M$ (end of sorted list) or maximum number of capacity expansions:

a. { Set $f = f + 1$. Let H_{ikm} be the f th coefficient in the sorted list. If $H_{ikm} \geq 0$, stop: all subsequent variables have $H_{ikm} \geq 0$}

Otherwise:

b. {Let $m"$ be the location corresponding to the current H_{ikm} coefficient.

Case 1: *location counter* $e = p$

If location $m"$ is not one of the p locations with $w_m = 1$, go to 2a.

Case 2: *location counter* $e < p$

or

$m"$ *is one of* p *selected locations*

i. Let $i"$ be the year for the current H_{ikm} coefficient. If for year $i"$ and location $m"$, any Y_{ikm} variable is already set to 1, go to 2a.

Otherwise:

ii. Let $d"$ be the region to which $m"$ belongs. If the number of locations with $w_m = 1$ in $d"$ is equal to the regional limit g_d, and $m"$ is not one of these, go to 2a.

Otherwise:

iii. Set $Y_{ikm} = 1$. If w_m is not equal to 1, set it to 1. Set $e = e + 1$, and increase the regional location counter for $d"$ by 1. $\mathrm{LB1} = \mathrm{LB1} + H_{ikm}$. Go to 2a.

Procedure UB

Step 1

[This step ensures that a counter ($Z(m)$) developed for each location, contains the lowest possible cost if it is assumed that only the particular location for which the cost is calculated is available.]

For each location m,

a. Sort the primal coefficients G_{3mn} in ascending order, and set $Z(m) = 0$, and $w(m) = 0$.

b. Assume that, for every combination of location m and year i, the capacity that is added is the maximum possible, that is, z_{i3m}.This means that the capacity available at location m in year i is $C_{0m} + \sum_{s=1}^{i} Z_{s3m}$, which is the maximum possible.

c. For each year do the following:

i. Initialize a counter $c = 0$, and repeat until $c > N$: Set $c = c + 1$. Let n" be the destination corresponding to the current G_{3mn} coefficient. Let Δ be the smaller of the remaining capacity at m and the remaining demand at n". Then $Z(m) = Z(m) + G_{3mn} \bullet \Delta$, and reduce the capacity at m by Δ.

ii. Add the cost of the necessary capacity expansion in the current year at the current location to $Z(m)$.

Step 2

[This step heuristically assigns sites in the order of ascending costs (as determined in Step l).]

Sort the $Z(m)$s in ascending order, set counter c to 0, and repeat until $c > M$**:** If the number of locations selected for capacity expansions [i.e., those with $w(m) = 1$] is equal top, then go to Step 3. Otherwise, let m" be the *cth* location in the sorted list. If the number of selected locations in the region corresponding to m" is equal to the regional limit, then do nothing. Otherwise, set $w(m'') = 1$.

Step 3

[This step heuristically finds the minimum cost flows in each year assuming that the site assignments in Step 2 are optimal.]

For each year, determine the cheapest flows using a greedy procedure and assuming that each location selected in Step 2 has the maximum capacity possible as defined in Step l, and other locations have zero capacity.

Step 4

[This step restores feasibility and computes costs for the flows found in Step 3.]

Restore feasibility

Step 3 determines flows that do not violate capacity constraints (1) and (2) (since maximum capacities are used), but which may violate the cost break- point constraints (4) and (5). This step finds the cost categories corresponding to the flows found in Step 3 by comparing them to the breakpoints, α_{jm} and β_{jm} in constraints (4) and (5). At this point, the flows x_{ijmn} are identified with the correct indices $i, j, m,$ and n for each flow, and no constraint of the model is violated.

Compute costs

For each location that is selected in Step 2, (a) initialize a location cost counter and (b) for each year, do the following:

Initialize a year cost counter, and find the total flow out of the location. This total corresponds to one of the three cost categories at the location. Determine the applicable manufacturing and labor cost category. The cost of the flow is then computed as (flow • cost coefficient of applicable manufacturing/labor cost category). Add this cost to the year cost counter. Then add the cost of the capacity expansion necessary to achieve this flow to the year cost counter. Next, add the transportation cost of the outflow. Finally, add the content of the year counter to the location's cost counter.

Step 5

[This step sums the costs (associated with the individual sites) calculated in Step 4 to provide a heuristic upper bound.]

Add the costs of the selected locations computed in Step 4 to provide a grand total. This grand total cost provides the necessary upper bound on the optimal objective value of the primal problem.

Computational Complexity of Algorithms for (S1), (S2), and (S3)

1. Algorithm (Ll)for (SJ)

Let $\Omega = / \bullet M \bullet 3$. Then, the heapsort algorithm used for Step 1 requires $O(\Omega\log\Omega)$ operations. Step 2 requires one pass through the sorted list [$O(\Omega)$ operations]. For each element in the list there is a potential pass through three Y_{ikm} variables and a potential pass through g_d locations. Therefore, complexity of Step 2 is [$O(3 + g_d) \bullet \Omega$]**.** The computational complexity of (L1) is $O[\Omega\log\Omega + (3 + g_d) \bullet \Omega]$.

2. Greedy algorithm for (S2) and (S3)

Step 1 involves sorting the E_{ijmn} coefficients. Let $\theta = I \bullet M \bullet N \bullet 3$. The complexity of the heapsort is $O(\theta\log\theta)$ operations. Step 2 involves a pass through the sorted list, with complexity $O(\theta)$. The complexity of the greedy algorithm is $O(\theta + \theta\log\theta)$. With similar reasoning, the complexity of (S3) is $O(\Omega + \Omega\log\Omega)$.

Siddhartha Syam is an assistant professor of management at Marquette University in Milwaukee, Wisconsin. He received his MBA from the University of Missouri and a PhD in management science from Texas A&M University. His research interests include the application of quantitative methods to issues in operations management, particularly in the areas of logistics, facility location, and scheduling. He has published in *Decision Sciences, Naval Research Logistics Quarterly,* and the *European Journal of Operational Research,* among others.

References

Cornuejols, G., Nemhauser, G. L., & Wolsey, L. A. (1990). The uncapacitated facility location problem. In P B. Mirchandani & R. L. Francis (Eds.), Discrete location theory. New York: John Wiley, 119-171.

Erlenkotter, D. (1978). A dual-based procedure for uncapacitated facility location. Operations Research, 26, 992-1009.

Fisher, M. L. (1985). An applications oriented guide to Lagrangian relaxation. Interfaces, 15, 10-21.

Fong, C. O., & Srinivasan, V (1981). The multi-region dynamic capacity expansion problem. Parts I and II. Operations Research, 29, 787-816.

Ford Motor Company. (1998). Company information on the Internet. Available: www.ford.com.

Gale Research. (1994). Statistical abstract of the world. Detroit, MI: Gale Research Inc, 194-200, 417-423, 974-980.

Geoffrion, A. M. (1974). Lagrangean relaxation for integer programming. Mathematical Programming Study, 2, 82-114.

Held, M., Wolfe, P, & Crowder, H. (1974). Validation of subgradient optimization. Mathematical Programming Study, 6, 66-68.

Honda Motor Company. (1998). [Company information on the Internet.] Available: www honda.com.

Hung, H. K., & Rikkers, R. F. (1974). A heuristic algorithm for the multi-period facility location problem. Joint ORSA/TIMS National Meeting, Boston, MA. Jacobsen, S. K. (1990). Multiperiod capacitated location models. In P B. Mirchan

dani & R. L. Francis (Eds.), Discrete location theory. New York: John Wiley, 173-208.

Jacoby, H. D., & Loucks, D. P (1972). Combined use of optimization and simulation models in river basin planning. Water Resources Research, 8, 14011414.

Kerr, W A., & Perdikis, N. (1995). The economics of international business. London, U.K: Chapman and Hall, 8-13.

Luss, H. (1979). A capacity expansion model for two facility types. Naval Research Logistics Quarterly, 26, 291-303.

Manne, A. S. (Ed.). (1967). Investments for capacity expansion: Size, location, and time phasing. Cambridge, MA: MIT Press, 19-141.

Markusen, J. R., Melvin, J. R., Kaempfer, W. H., & Maskus, K. E. (1995). International trade: Theory and evidence. New York: McGraw-Hill, 5-15. Nemhauser, G. L., & Wolsey, L. A. (1988). Integer and combinatorial optimization. New York: John Wiley, 60, 666-667.

Polyak, B. (1967). A general method for solving extremum problems. Soviet Mathematics Doklady, l , 593-597.

Rao, R. C., & Ruthenberg, D. P (1977). Multilocation plant sizing and timing. Management Science, 23, 1187-1198.

Roth, T. (Sept. 30, 1994). Gordian knot. Wall Street Journal, p. R4. Samuelson, P A. (1980). Economics. New York: McGraw-Hill, 439-455. Stevenson, W. J. (1996). Production/operations management. Chicago: Irwin, 197-203.

Toyota Motor Corporation. (1998). [Company information in online newspaper, The Times@Toyota.] Available: www toyota.com.

U.S. Department of Commerce. (1995). The big emerging markets: 1996 outlook and source book. Lanham, MA: Beman Press, 52-260.

U.S. Department of Commerce. (1997). Best market reports: Computer software. Economic Bulletin Board. Washington D.C.: Office of Business Analysis. Van Roy, T J., & Erlenkotter, D. (1982). A dual-based procedure for dynamic facility location. Management Science, 28, 1091-1105.