A Stochastic Model of the Dynamic Vehicle Allocation Problem

WARREN B. POWELL

Princeton University, Princeton, New Jersey

The stochastic vehicle allocation problem addresses the movement of vehicles between locations over a given planning horizon. The demand for vehicles to carry loads between locations is uncertain, and vehicles are assumed to be able to handle several loads over the course of the planning horizon. This requires tracking the movement of both loaded and empty vehicles, resulting in a network with stochastic flows. The methodology represents the flows of vehicles over the network explicitly as random variables, taking advantage of the acyclic structure of the time space network. The decision variables are formulated in terms of sending a certain fraction of the supply of vehicles at a node (which is random) over each of the outbound links. The result is a nonseparable objective function with a very simple constraint structure which lends itself readily to the Frank-Wolfe algorithm. Numerical experiments suggest very good computational efficiency.

INTRODUCTION

The stochastic vehicle allocation problem arises when a carrier must allocate vehicles over space and time in an effort to anticipate uncertain demands. Examples are railroads, which must distribute empty freight cars, and truckload motor carriers, which must supply empty trucks to different cities. The carrier may know where its vehicles are now and may even know with some certainty what the demands for those vehicles are today. Efficient allocation of vehicles, however, requires trying to anticipate future demands which must therefore be forecast, usually with considerable uncertainty.

The current approach for solving the vehicle allocation problem incorporates future demands by setting up a multitime period network with links that move forward in time as they go from one node to the next. By assuming that demands are known with certainty, the problem can be easily formulated as a capacitated transshipment problem and solved using a network simplex code. The assumption of deterministic demands, however, can potentially introduce significant errors into the model and, more seriously from a planning perspective, eliminates the possibility of analyzing strategies directed specifically at handling uncertainty in demand. For example, a railroad or motor carrier might wish to offer price discounts to shippers who place orders more than 4 days in advance. To implement such a strategy, the carrier would first have to know the cost benefits accruing from a reduction in the uncertainty in demand.

A deterministic model will in addition do a particularly poor job of estimating stockout and holding costs, and is often relatively insensitive to these parameters. It is normal, for example, to assume that the actual demand for vehicles is equal to the median demand, in which case the model will return a solution which provides too few vehicles 50% of the time. If the traffic is particularly lucrative, it is possible to estimate the demand at the 90th percentile rather than the 50th, but the marginal revenue should then be adjusted downward to reflect the probability of filling the last vehicle. If the carrier wishes to send a vehicle only if there is a demand for the vehicle, then the flows themselves become stochastic, a fact that cannot be captured even by a nonlinear deterministic model.

Previous research in the area of stochastic transshipment problems is relatively sparse. COOPER AND LEBLANC[1] considered the stochastic transportation problem where the flows from supply to demand are deterministic and with linear transportation costs. Demands are assumed to be stochastic with stockout and holding costs provided as inputs to the model. The objective is to minimize transportation costs and expected stockout and holding costs, producing a simple convex, nonlinear objective function which is easily solved. The simplicity of the model arises from the fact that flows must be sent before the demands are known and only one time period is considered.

The first, and apparently the only, effort at directly solving the multiple time period vehicle allocation problem under random demands is the recent work by

Transportation Science
Vol. 20, No. 2, May 1986

0041-1655/86/2002-0117 $01.25

JORDAN[2] and JORDAN AND TURNQUIST.[3] The problem addressed in this research is the empty freight car distribution problem, where known supplies of cars must be allocated to different classification yards over time to meet uncertain demands. The following assumptions are made in this research:

1. Supplies of and demands for empty cars are assumed random and are represented using a normal distribution. Demands are assumed to occur at the nodes of the network.
2. Travel times between terminals are random and described using a negative binomial distribution.
3. Once an empty car is sent from one yard to another, it cannot be reallocated to a third yard at a later point in time.
4. Once a car is sent to a shipper, it is lost from the system.
5. Demands that are not satisfied in one time period are carried forward to the next time period.

Assumption 4 arises because the model does not track loaded movements. Since the loaded cycle of a freight car would normally exceed a typical 7–10-day planning horizon, it is unlikely that an allocation decision would have to be made twice for the same car within the planning period. This assumption fails, however, in other applications such as trucking, where a truck may be loaded two or three times within a 7-day planning horizon. For this problem, loaded movements must be considered explicitly as well as the empties. Other limitations of this research include the use of the normal distribution, which breaks down when the mean is small relative to the variance, and the assumption that empty vehicles cannot be sent more than once within the planning period. Again, these assumptions represent reasonable approximations for empty freight car distribution but would not necessarily apply in other applications.

The purpose of this paper is to present a model that in some respects is more general than the formulation presented in [3]. The particular features of the model are motivated primarily by the truck allocation problem for truckload motor carriers. The major assumptions that are made relative to five listed earlier used by Jordan and Turnquist are as follows:

1. Supplies of and demands for trucks are assumed to be random and described by an Erlang distribution. Demands are assumed to occur on the links of the network.
2. Travel times are assumed to be deterministic.
3. An empty truck may be moved repeatedly over the planning horizon.
4. When a trailer is moved full from one city to the next, it becomes empty and must again be reallocated to handle a new demand.
5. Demands not satisfied in one time period are assumed lost from the system.

Assumptions 1, 3, and 4 represent extensions to the Jordan and Turnquist formulation. Assumptions 2 and 5, on the other hand, are more restrictive but are realistic in the context of trucking.

The organization of the paper is as follows. Section 1 presents the basic formulation of the problem and Section 2 describes how to calculate the flows of full and empty trucks on each link in the network over the planning horizon. Section 3 outlines an efficient solution algorithm, and Section 4 describes a series of numerical experiments designed to test the accuracy of certain approximations required and the efficiency of the algorithm. Also included are comparisons between deterministic and stochastic formulations of the same problem.

1. PROBLEM FORMULATION

THE STOCHASTIC vehicle allocation problem takes as given initial supplies of empty vehicles and then must route these vehicles to other nodes in the system at future points in time. The network depicting all possible movements is shown in Figure 1. Any movement to another city can be made full, if there is sufficient demand, or empty. Links representing a vehicle being held at a city until the next day are modeled as empty movements. The problem is to decide a) how many vehicles to send empty from city i to a different city j at a future point in time, b) how many vehicles should be held at node i another day, and c) how many vehicles should be allowed to handle demands from i to j.

The set of all possible movements is represented using a space-time network where each node i represents a region at a specific point in time. This network can be represented as a directed graph $G = (N, L)$ where N is the set of all possible nodes within a specified planning horizon and L is the set of all links.

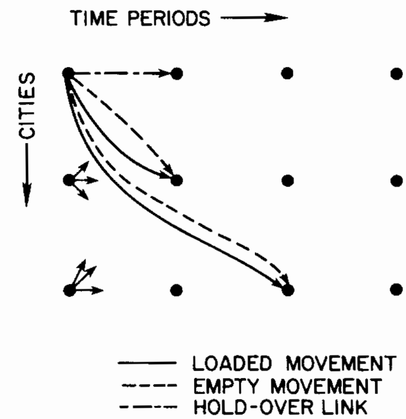

Fig. 1. Schematic of network for vehicle allocations problems.

Further define the subsets:

L^l = set of all loaded movement links
L^e = set of all empty movement links
L^h = set of links representing vehicles held in the same region from one period to the next

where $L = L^l \cup L^e \cup L^h$. To capture the temporal structure of the network define:

$\underline{\underline{i}}$ = node representing the same city as i but in the next time period (the link $(i, \underline{\underline{i}} \in L^h)$.
$\underline{i}$ = node representing the same city as i but in the previous time period (the link $(\underline{i}, i) \in L^h$).
$A_i = \{k \mid (k, i) \in L^l \cup L^h\}$
$B_i = \{k \mid (i, k) \in L^l\}$.

Note in the definitions of the sets A_i and B_i that $\underline{i} \in A_i$ but $\underline{\underline{i}} \notin B_i$.

Next define the following:

F_{ij} = flow of full vehicles from node i to j.
E_{ij} = flow of empty vehicles from i to j.
S_i = supply of vehicles at node i.
R_j = new arrivals of vehicles to node j which were sent prior to the beginning of the planning horizon. $R_j, j = 1, \ldots, N_C$, represent the initial supplies of vehicles; however, it is possible to have $R_j > 0$ for all $j \in N$.
s_{ij} = a "stockout cost" for not meeting demands.
r_{ij} = net revenue received from a full movement from i to j.
c_{ij} = cost of an empty movement from i to j.
D_{ij} = demand for vehicles from i to j.
N_C = number of cities in the system.

Throughout the paper we will use the convention that $\bar{F} = E[F]$ and $\bar{\bar{F}} = \text{Var}[F]$.

Given that the supplies of vehicles, $\{S\}$, are potentially random, it is not possible to formulate the decision variables in terms of flows of vehicles, as is done in deterministic network models. To deal with this difficulty in a manageable way, the decision variables are defined as follows:

θ_{ij} = fraction of the supply of vehicles at node i allocated for a full movement to node j.
α_{ij} = fraction of the supply at node i to be sent empty to node j.

A variety of assumptions and modeling approximations, in addition to those listed in the introduction, are employed to develop a formal optimization problem. These include:

1. Vehicles are treated as continuous variables.
2. Exogenous supplies of vehicles, R_j, are continuous random variables described by an Erlang density function with known parameters.
3. The demands for vehicles D_{ij} are continuous random variables described by an Erlang density function with known parameters. Since the nodes i and j represent both regions and time periods, the parameters of the distribution of D_{ij} are allowed to vary over time.
4. The supplies S_i are necessarily continuous random variables where the parameters of the distribution of S_i are determined endogenously by the model. The distribution itself will be approximated by an Erlang distribution.
5. The problem is modeled over a finite planning horizon which is implicitly defined by the nodes included in the set N. No special steps were taken to handle the end effects due to truncating the planning horizon.
6. Vehicles allocated to move loaded from i to j but which cannot do so due to insufficient demand are assumed to be held over in the same region until the next time period.
7. The random variables $\{R_j\}, j = 1, \ldots, N$ and $\{D_{ij}\}$, $i, j \in L^l$, are all assumed to be independent. Other independence assumptions are also required to simply certain derivations; these assumptions are best introduced as needed.

Assumption 6 above is central to the model both for the added realism it introduces as well as the limitation it imposes on the model. An alternative assumption would have defined a single set of decision variables $\{\theta\}$ where θ_{ij} is the fraction moved from i to j, where as many of these would be moved loaded as possible. This approach was used by POWELL et al.[6] and has the result that if the external supplies $\{R_j\}$ are deterministic then so are the flows over the network. In reality, a truck will usually only be moved from i to j if it can be moved full, implying that the model in [6] will generate considerably more empty miles than is necessary. Assumption 6 avoids this particular limitation but imposes a different, although lesser, restriction. Specifically, while it allows trucks to be moved only when full, it forces a truck allocated to move full over a particular link (i, j) to be held over from i to $\underline{\underline{i}}$ if the demand on the link is too low. The model does not allow this truck to be moved over another link which may have excess demand. The assumptions being made here do represent an extension of previous research and allows us to investigate the problem of a network with stochastic flows.

Given the vectors $\{\theta\}$ and $\{\alpha\}$, the flows of fulls and empties are given by

$$F_{ij} = \min\{\theta_{ij} S_i, D_{ij}\} \quad \forall\, i, \quad j \in B_i \tag{1}$$

$$E_{ij} = \alpha_{ij} S_i \qquad \forall\, i, \quad j \in B_i. \tag{2}$$

The flow of empties on the holdover link must include the overflow from other links, and hence is written:

$$E_{i\underline{i}} = \alpha_{i\underline{i}} S_i + \textstyle\sum_{j \in B_i} (\theta_{ij} S_i - F_{ij}). \qquad (3)$$

Note that Equations 1–3 imply that it is possible that if $\alpha_{ij} > 0$ and $\theta_{ij} > 0$, empties may be sent from i to j at the same time that the demands exceed the number of vehicles being sent full. Such an anomaly is unlikely to have any significant affect since the model will generally tend to increase θ_{ij} and decrease α_{ij} until most of the demands are being satisfied. However, this possibility does serve to highlight a drawback of the structure of the model.

Having defined the basic variables required, the objective function is simply to maximize the expected net revenue minus stockout costs over the planning horizon, as follows:

$$\max \pi(\theta, \alpha) = \textstyle\sum_{i \in N} \cdot [\sum_{j \in B_i} (r_{ij} \bar{F}_{ij} - c_{ij} \bar{E}_{ij} - s_{ij} (\bar{D}_{ij} - \bar{F}_{ij})) - c_{i\underline{i}} \bar{E}_{i\underline{i}}] \qquad (4)$$

subject to

$$\textstyle\sum_{j \in B_i} (\theta_{ij} + \alpha_{ij}) + \alpha_{i\underline{i}} = 1. \qquad (5)$$

Equation 4 represents a nonlinear, nonseparable objective function with a very simple set of constraints. Unfortunately, it has not been possible to establish concavity of $\pi(\theta, \alpha)$ and hence we must satisfy ourselves at this time that we may only find a local optimum.

The next section presents the relationships needed to calculate (4) given vectors $\{\theta\}$ and $\{\alpha\}$.

2. CALCULATING THE OBJECTIVE FUNCTION

TO CALCULATE $\pi(\theta, \alpha)$ given $\{\theta\}$ and $\{\alpha\}$ the initial supplies of vehicles at each city are assumed given (or alternatively, the means and variances of the supplies are assumed known). It is then necessary to find the means and variances of the supplies of vehicles at each node in the network for all future points in time.

Flows into node j consist of the external supply of vehicles sent prior to the beginning of the planning horizon, the flow of fulls and empties from nodes in the set A_j, empty vehicles that were held at $\underline{j}$ and the overflow of vehicles at $\underline{j}$ that were allocated to move full to the nodes in B_j but were held over as a result of insufficient demand. Stating this relationship mathematically gives

$$\begin{aligned} S_j &= \textstyle\sum_{i \in A_j} [F_{ij} + E_{ij}] + R_j \\ &= \textstyle\sum_{i \in A_j} [\alpha_{ij} S_i + F_{ij}] \\ &\quad + \textstyle\sum_{k \in B_{\underline{j}}} [\theta_{\underline{j}k} S_{\underline{j}} - F_{\underline{j}k}] + R_j. \end{aligned} \qquad (6)$$

The first term on the right hand side of (6) is the flow of fulls and empties on all links leading into node j where $\alpha_{ij} S_i$ is the flow of empties and F_{ij} is the flow of fulls as given by (1). Included in this term is the holdover link $(\underline{j}, j)$ where $F_{\underline{j}j} = 0$. The second term in (6) is the vehicles that could not be filled from $\underline{j}$ to nodes in $B_{\underline{j}}$ (which excludes j) and hence had to be held over. The moments of R_j are assumed known and are exogenous to the model.

Taking expectations of both sides of (6) gives the expected supply of empty vehicles at j:

$$\begin{aligned} \bar{S}_j &= \textstyle\sum_{i \in A_j} [\alpha_{ij} \bar{S}_i + \bar{F}_{ij}] \\ &\quad + \textstyle\sum_{k \in B_{\underline{j}}} [\theta_{\underline{j}k} \bar{S}_{\underline{j}} - \bar{F}_{\underline{j}k}] + \bar{R}_j. \end{aligned} \qquad (7)$$

To find the variance of S_j, the (assumed) independence between the supplies of S_i, $i \in A_j$, implies that flows from different points into j are independent. The flow of empties and fulls on the same link, however, will not be independent. A more difficult problem is the variance of the flow on the overflow link $(\underline{j}, j)$ as these flows are all related to the supply $S_{\underline{j}}$. In addition, the overflows, represented by the second term in (7), will be correlated with the vehicles being held over from $\underline{j}$ to j, represented by $\alpha_{\underline{j}j} S_{\underline{j}}$ in the first term in (7). In view of these relationships, the variance of S_j is found to be

$$\begin{aligned} \bar{\bar{S}}_j &= \textstyle\sum_{i \in A_j} [\alpha_{ij}^2 \bar{\bar{S}}_i + \bar{\bar{F}}_{ij} + 2\alpha_{ij} \mathrm{Cov}(S_i, F_{ij})] \\ &\quad + \textstyle\sum_{k \in B_j} \mathrm{Cov}(\alpha_{\underline{j}j} S_{\underline{j}}, \theta_{\underline{j}k} S_{\underline{j}} - F_{\underline{j}k}) \\ &\quad + \textstyle\sum_{k \in B_j} [\theta_{\underline{j}k}^2 \bar{\bar{S}}_{\underline{j}} + \bar{\bar{F}}_{\underline{j}k} - 2\theta_{\underline{j}k} \mathrm{Cov}(S_{\underline{j}}, F_{\underline{j}k})] \\ &\quad + \textstyle\sum_{k \in B_j} \sum_{\substack{l \in B_j \\ l \neq k}} \mathrm{Cov}(\theta_{\underline{j}k} S_{\underline{j}} - F_{\underline{j}k}, \theta_{\underline{j}l} S_{\underline{j}} - F_{\underline{j}l}) + \bar{\bar{R}}_j. \end{aligned} \qquad (8)$$

Equations 7 and 8, together with Equations 1 and 2, are the governing relationships for the model in the sense that, given the decision variables $\{\theta\}$ and $\{\alpha\}$, and the initial supplies of vehicles, all the flows over the network can be determined. To carry out some of the steps the assumption that demands are described by an Erlang distribution is used. If $d_{ij}(t)$ is the density function of D_{ij}, then $d_{ij}(t)$ is assumed to be given by

$$d_{ij}(t) = \mu_{ij} (\mu_{ij} t)^{l_{ij}-1} e^{-\mu_{ij} t} / (l_{ij} - 1)! \qquad (9)$$

where μ_{ij} and l_{ij} are assumed known. The distributions of the supplies of empties, S_j, are approximated using an Erlang distribution, where if $s_j(t)$ is the density function of S_j, then

$$s_j(t) = \lambda_j (\lambda_j t)^{\kappa_j - 1} e^{-\lambda_j t} / (\kappa_j - 1)!. \qquad (10)$$

For the initial time periods, λ_j and κ_j are assumed known, or the supply may be known deterministically.

For future time periods, the mean and variance of S_j are calculated using (7) and (8), from which κ_j and λ_j are found using

$$\kappa_j = \max\{1, [\bar{S}_j^2/\bar{\bar{S}}_j]\} \tag{11}$$

$$\lambda_j = \kappa_j/\bar{S}_j \tag{12}$$

where $[x]$ in (11) denotes the nearest integer to x.

The calculation of the parameters κ and λ in this way guarantees that the need to keep κ integer does not affect the mean of the distribution of S_j. Keeping κ_j integer is not entirely necessary but allows all the calculations to be done in closed form. Note that the use of an Erlang distribution for the random variables S_j is purely a modeling approximation required to calculate the moments of the flow of full vehicles outbound from node j. No formal experiments have been conducted to test the accuracy of this fit, but simulation results presented in Section 4 suggest it provides a very good approximation for the purposes of calculating the expected flows of fulls and empties.

The only difficult calculations required to calculate Equations 1, 7 and 8 are the expressions for $\bar{F}_{ij}$, $\bar{\bar{F}}_{ij}$, $\mathrm{Cov}(S_i, F_{ij})$ and $\mathrm{Cov}(F_{jk}, F_{jl})$, where the last expression is needed to calculate the covariances in the last term of (8). After a considerable amount of algebra, the equations for $\bar{F}$, $\bar{\bar{F}}$ and $\mathrm{Cov}(S_i, F_{ij})$ are found to be (dropping the subscripts i and j):

$$\bar{F} = \frac{\theta\kappa}{\lambda}\left[1 - \frac{l}{\kappa}\sum_{n=1}^{\kappa}\frac{n}{(\kappa-n)} \cdot \binom{\kappa+l-n-1}{l} r(l,\kappa-n)\right] \tag{13}$$

$$\bar{\bar{F}} = \frac{\theta^2\kappa(\kappa+1)}{\lambda^2}\left[1 - \sum_{n=1}^{\kappa+1}\frac{l\cdot n(2\kappa+1-n)}{\kappa(\kappa+1)(\kappa-n+1)} \cdot \binom{\kappa+l-n}{l} r(l,\kappa-n+1)\right] - \bar{F}^2 \tag{14}$$

$$\mathrm{Cov}(S,F) = \frac{\theta\kappa(\kappa+1)}{\lambda^2}\left[1 - \frac{l}{\kappa+1}\sum_{n=1}^{\kappa+1}\frac{n}{\kappa-n+1} \cdot \binom{\kappa+l-n}{l} r(l,\kappa-n+1)\right] - \bar{S}\bar{F} \tag{15}$$

where

$$\binom{a}{b} = \frac{a!}{b!(a-b)!} \tag{16}$$

and

$$r(a,b) = \frac{\mu^a(\lambda/\theta)^b}{(\mu+\lambda/\theta)^{a+b}}. \tag{17}$$

The expression for $\mathrm{Cov}(F_{jk}, F_{jl})$ was derived but proved to be significantly more complex than Equations 13–15. In the interests of computational efficiency, the overflow from different links is assumed to be independent. After a few manipulations, this produces the following approximate expression for $\bar{\bar{S}}_j$:

$$\begin{aligned}\bar{\bar{S}}_j \cong & \textstyle\sum_{i\in A_j}[\alpha_{ij}^2\bar{\bar{S}}_i + \bar{\bar{F}}_{ij} + 2\alpha_{ij}\mathrm{Cov}(S_i, F_{ij})] \\ & + \alpha_{jj}(\bar{S}_j \textstyle\sum_{k\in B_j}\theta_{jk} - \sum_{k\in B_j}\mathrm{Cov}(S_j, F_{jk})) \\ & + \textstyle\sum_{k\in B_j}[\theta_{jk}^2\bar{\bar{S}}_j + \bar{\bar{F}}_{jk} - 2\theta_{jk}\mathrm{Cov}(S_j, F_{jk})] \\ & + \bar{\bar{R}}_j\end{aligned} \tag{18}$$

which further reduces to

$$\begin{aligned}\bar{\bar{S}}_j \cong & \textstyle\sum_{i\in A_j}[\alpha_{ij}^2\bar{\bar{S}}_i + \bar{\bar{F}}_{ij} + 2\alpha_{ij}\mathrm{Cov}(S_i, F_{ij})] \\ & + \textstyle\sum_{k\in B_j}[\bar{\bar{F}}_{jk} + \theta_{jk}\bar{\bar{S}}_j(\alpha_{jj} + \theta_{jk}) \\ & - (\alpha_{jj} + 2\theta_{jk})\mathrm{Cov}(S_j, F_{jk})] \\ & + \bar{\bar{R}}_j.\end{aligned} \tag{19}$$

All the terms in (19) are easily calculated using Equations 13–15.

The relationships in this section can be efficiently applied using standard network list processing techniques. To calculate all the flows over the network, the algorithm would loop over all the nodes starting with those on the first time period and moving forward in time. For each node i, $\bar{S}_i$ and $\bar{\bar{S}}_i$ would already be known, either because they were input to the model (if i falls in the initial time period) or because all the calculations needed to find them would have already been completed. Given $\bar{S}_i$ and $\bar{\bar{S}}_i$, an Erlang distribution for S_i is fitted using (11) and (12). Next, looping over all the links (i, j) emanating from i, the procedure finds $\bar{F}_{ij}$, $\bar{\bar{F}}_{ij}$ and $\mathrm{Cov}(S_i, F_{ij})$ which are then accumulated at each node $j, j \in B_i$, in such a way as would allow calculating $\bar{S}_j$ and $\bar{\bar{S}}_j$ when these nodes are encountered later. At the same time, the expected flows of fulls, $\bar{F}_{ij}$, and empties, $\bar{E}_{ij} = \alpha_{ij}\bar{S}_i$, are stored for each link, to be used later when actually calculating the objective function.

The next section addresses the problem of developing a solution algorithm to maximize the expected net revenue.

3. SOLUTION ALGORITHM

THE STRUCTURE of the objective function (4) and the constraint set (5) suggests the use of the Frank-Wolfe algorithm, which has proven useful particularly in the area of stochastic network problems.[1,3] The principal reason for its ease of application is that the linearized objective function decomposes into a set of trivial subproblems, one for each node in the network.

The linearized objective function is given simply by:

$$\min_{\{\theta,\alpha\}} \sum_i \left[\sum_{j\in B_i} \left(\frac{\partial \pi(\theta^o, \alpha^o)}{\partial \theta_{ij}} \theta_{ij} + \frac{\partial \pi(\theta^o, \alpha^o)}{\partial \alpha_{ij}} \alpha_{ij} \right) + \frac{\partial \pi(\theta^o, \alpha^o)}{\partial \alpha_{i\underline{i}}} \cdot \alpha_{i\underline{i}} \right] \quad (20)$$

subject to

$$\textstyle\sum_{j\in B_i} (\theta_{ij} + \alpha_{ij}) + \alpha_{i\underline{i}} = 1 \ \forall\ i \quad (21)$$

where $\{\theta^o, \alpha^o\}$ is the current solution. Let (θ_L, α_L) be the optimal solution of (20). This problem decomposes into separate subproblems for each node. After calculating the set of derivatives $\partial\pi/\partial\theta_{ij}$, $\forall\ j \in B_i$ and $\partial\pi/\partial\alpha_{ij}$, $\forall\, j \in B_i \cap \underline{\underline{j}}$, the solution of (20) simply finds the largest derivative and sets the corresponding θ_{ij}^L or α_{ij}^L, depending on which derivative is largest, equal to 1, with all the rest set equal to zero. This "all or nothing" solution guarantees convergence but can be relatively slow. It has been shown[4] that methods which divide the flow among competing "good" paths can significantly enhance convergence. For this reason, the following scheme is proposed. Let U_i and V_i be the sets of nodes where a small increase in θ_{ij} or α_{ij}, respectively, would show an increase in net revenue. In other words,

$$U_i = \left\{ j \mid j \in B_i, \frac{\partial \pi(\theta^o, \alpha^o)}{\partial \theta_{ij}} > 0 \right\}$$

$$V_i = \left\{ j \mid j \in B_i \cap \underline{\underline{j}}, \frac{\partial \pi(\theta^o, \alpha^o)}{\partial \alpha_{ij}} > 0 \right\}.$$

Next define weights u_{ij} and v_{ij} where

$$u_{ij} = \begin{cases} \gamma & \dfrac{\partial \pi(\theta^o, \alpha^o)}{\partial \theta_{ij}}\ j \in U_i \\ 0 & \text{otherwise} \end{cases} \quad (22)$$

$$v_{ij} = \begin{cases} \gamma & \dfrac{\partial \pi(\theta^o, \alpha^o)}{\partial \alpha_{ij}}\ j \in V_i \\ 0 & \text{otherwise} \end{cases} \quad (23)$$

where γ is a predetermined scaling factor. Finally, determine θ_{ij}^L and α_{ij}^L using the following logit function:

$$\theta_{ij} = \begin{cases} e^{u_{ij}}/(\sum_{k\in U_i} e^{u_{ik}} + \sum_{k\in V_i} e^{v_{ik}}), & j \in U_i \\ 0 & \text{otherwise} \end{cases} \quad (24)$$

$$\alpha_{ij}^L = \begin{cases} e^{v_{ij}}/(\sum_{k\in U_i} e^{U_{ik}} + \sum_{k\in V_i} e^{v_{ik}}) & j \in V_i \\ 0 & \text{otherwise.} \end{cases} \quad (25)$$

The advantage of this approach is that most of the flow is concentrated along the path showing the greatest increase in net revenue. The extent of this concentration is governed by the size of the parameter γ. Note that in applying (24) and (25), if the sets U_i and V_i are empty, then the algorithm would revert back to an all or nothing approach.

Finding the Derivatives

The solution algorithm requires an efficient method for calculating the derivatives $\partial\pi(\theta^o, \alpha^o)/\partial\theta_{ij}$ and $\partial\pi(\theta^o, \alpha^o)/\partial\alpha_{ij}$. These derivatives must reflect the impact of a small change in θ_{ij} or α_{ij} will have both on revenues on the link (i, j) as well as on the rest of the network. The derivatives can be calculated by using the following recursions:

$$\frac{\partial \pi}{\partial \alpha_{ij}} = -c_{ij} \frac{\partial \bar{E}_{ij}}{\partial \alpha_{ij}} + \frac{\partial \bar{S}_j}{\partial \alpha_{ij}} \frac{\partial \pi}{\partial \bar{S}_j} + \frac{\partial \bar{\bar{S}}_j}{\partial \alpha_{ij}} \frac{\partial \pi}{\partial \bar{\bar{S}}_j} \quad (26)$$

$$\frac{\partial \pi}{\partial \theta_{ij}} = \left[(r_{ij} + s_{ij}) \frac{\partial \bar{F}_{ij}}{\partial \theta_{ij}} + \frac{\partial \bar{S}_j}{\partial \theta_{ij}} \frac{\partial \pi}{\partial \bar{S}_j} + \frac{\partial \bar{\bar{S}}_j}{\partial \theta_{ij}} \frac{\partial \pi}{\partial \bar{\bar{S}}_j} \right] + \left[-c_{i\underline{i}} \cdot \frac{\partial E_{i\underline{i}}}{\partial \theta_{ij}} + \frac{\partial S_{\underline{i}}}{\partial \theta_{ij}} \frac{\partial \pi}{\partial \bar{S}_{\underline{i}}} + \frac{\partial \bar{\bar{S}}_{\underline{i}}}{\partial \theta_{ij}} \frac{\partial \pi}{\partial \bar{\bar{S}}_{\underline{i}}} \right]. \quad (27)$$

The first term in (26) represents the impact of an increase in α_{ij} on profits on link (i, j). The second and third terms look at the impact of α_{ij} on $\bar{S}_j$ and $\bar{\bar{S}}_j$ and the subsequent impact on profits on all links in future time periods. Equation 27 has the same structure but is complicated by the fact that increasing θ_{ij} will increase the flow both on link (i, j) and on the overflow link $(i, \underline{i})$. The use of both (26) and (27) requires also knowing $\partial\pi/\partial\bar{S}_i$ and $\partial\pi/\partial\bar{\bar{S}}_i$ which describe the impact of an increase in $\bar{S}_i$ and $\bar{\bar{S}}_i$ on all future profits. Numerical experiments described in Section 4, however, support the use of the approximation that $\partial\pi/\partial\bar{\bar{S}}_i \cong 0$. This approximation is used in this section; interested readers are referred to Appendix A for the relationships which do not use this approximation.

To calculate (26) observe that

$$\frac{\partial \bar{S}_j}{\partial \alpha_{ij}} = \frac{\partial \bar{E}_{ij}}{\partial \alpha_{ij}} = \bar{S}_i, \quad (28)$$

which gives

$$\frac{\partial \pi}{\partial \alpha_{ij}} = \left(-c_{ij} + \frac{\partial \pi}{\partial \bar{S}_j} \right) \bar{S}_i. \quad (29)$$

To calculate (27) requires

$$\frac{\partial \bar{S}_j}{\partial \theta_{ij}} = \frac{\partial \bar{F}_{ij}}{\partial \theta_{ij}} \quad (30)$$

and

$$\frac{\partial \bar{S}_{\underline{\underline{i}}}}{\partial \theta_{ij}} = \bar{S}_i - \frac{\partial \bar{F}_{ij}}{\partial \theta_{ij}}. \tag{31}$$

Combining (27), (30) and (31) gives

$$\frac{\partial \pi}{\partial \theta_{ij}} = \left(r_{ij} + s_{ij} + c_{i\underline{\underline{i}}} + \frac{\partial \pi}{\partial \bar{S}_j} - \frac{\partial \pi}{\partial \bar{S}_{\underline{\underline{i}}}} \right) \frac{\partial \bar{F}_{ij}}{\partial \theta_{ij}} - \left(c_{i\underline{\underline{i}}} - \frac{\partial \pi}{\partial \bar{S}_{\underline{\underline{i}}}} \right) \bar{S}_i. \tag{32}$$

Finally, the use of (29) and (32) requires knowing $\partial \pi / \partial \bar{S}_i$. Applying the same logic behind (27) gives

$$\frac{\partial \pi}{\partial \bar{S}_j} = \sum_{k \in B_j} \left[(r_{jk} + s_{jk}) \frac{\partial \bar{F}_{jk}}{\partial \bar{S}_j} - c_{jk} \frac{\partial \bar{E}_{jk}}{\partial \bar{S}_j} + \frac{\partial \bar{S}_k}{\partial \bar{S}_j} \frac{\partial \pi}{\partial \bar{S}_k} \right] - c_{j\underline{\underline{j}}} \frac{\partial \bar{E}_{j\underline{\underline{j}}}}{\partial \bar{S}_j} + \frac{\partial \bar{S}_{\underline{\underline{j}}}}{\partial \bar{S}_j} \frac{\partial \pi}{\partial \bar{S}_{\underline{\underline{j}}}}. \tag{33}$$

The first term in (33) reflects the direct effect of an increase in $\bar{S}_j$ on the revenues for each link (j, k) as a result of the impact of flows and empties on that link, as well as the subsequent impact on future profits. The second term incorporates the effect of an increase in $\bar{S}_j$ on the total flow on the holdover link on $\bar{S}_{\underline{\underline{j}}}$. Equation 33 can be simplified by observing that

$$\frac{\partial \bar{S}_k}{\partial \bar{S}_j} = \frac{\partial \bar{F}_{jk}}{\partial \bar{S}_j} + \frac{\partial \bar{E}_{jk}}{\partial \bar{S}_j}, \quad k \in B_j \tag{34}$$

and

$$\frac{\partial \bar{E}_{jk}}{\partial \bar{S}_j} = \alpha_{jk}, \quad k \in B_j. \tag{35}$$

For the holdover link $(j, \underline{\underline{j}})$,

$$\frac{\partial \bar{E}_{j\underline{\underline{j}}}}{\partial \bar{S}_j} = \alpha_{j\underline{\underline{j}}} + \sum_{k \in B_j} \left(\theta_{jk} - \frac{\partial \bar{F}_{jk}}{\partial \bar{S}_j} \right) \tag{36}$$

and

$$\frac{\partial \bar{S}_{\underline{\underline{j}}}}{\partial \bar{S}_j} = \frac{\partial \bar{E}_{j\underline{\underline{j}}}}{\partial \bar{S}_j}. \tag{37}$$

Combining (33)–(37) gives

$$\frac{\partial \pi}{\partial \bar{S}_j} = \sum_{k \in B_j} \left[\left(r_{jk} + s_{jk} + c_{j\underline{\underline{j}}} + \frac{\partial \pi}{\partial \bar{S}_k} - \frac{\partial \pi}{\partial \bar{S}_{\underline{\underline{j}}}} \right) \frac{\partial \bar{F}_{jk}}{\partial \bar{S}_j} - \left(c_{jk} - \frac{\partial \pi}{\partial \bar{S}_k} \right) \alpha_{jk} - \left(c_{j\underline{\underline{j}}} - \frac{\partial \pi}{\partial \bar{S}_{\underline{\underline{j}}}} \right) \theta_{jk} \right] - \left(c_{j\underline{\underline{j}}} - \frac{\partial \pi}{\partial \bar{S}_{\underline{\underline{j}}}} \right) \alpha_{j\underline{\underline{j}}}. \tag{38}$$

Expressions for $\partial \bar{F}_{ij} / \partial \theta_{ij}$ and $\partial \bar{F}_{ij} / \partial \bar{S}_i$ are given in Appendix B.

Equations 29, 32, and 38 can be calculated quickly and easily by starting with the last node in the network, which would fall in the last time period, and working backward in time. For the nodes $\{j\}$ in the last time period, $\partial \pi / \partial \bar{S}_j = 0$. Using (38) for $\partial \pi / \partial \bar{S}_j$ (29) and (32) may be calculated in a single pass over the network.

The Solution Algorithm

Once the derivatives are determined, new "trial" values for $\{\theta\}$ and $\{\alpha\}$, denoted $\{\theta_L\}$ and $\{\alpha_L\}$, are determined either by applying the Frank-Wolfe algorithm, resulting in an "all-or-nothing" choice of $\{\theta_L\}$ and $\{\alpha_L\}$, or by using the heuristic splitting formulas (24) and (25). Finally, a one dimensional search is applied to find the best convex combination of the current set $\{\theta\}$ and $\{\alpha\}$ and the trial values $\{\theta_L\}$ and $\{\alpha_L\}$.

The complete algorithm is summarized as follows:

Step 1. Initialization
Determine initial values $\{\theta^o\}$ and $\{\alpha^o\}$ by solving the problem deterministically. Set $N = 0$.

Step 2. Forward pass
Calculate the moments of the flows on each link, $\bar{F}_{ij}$, $\bar{\bar{F}}_{ij}$, and $\bar{E}_{ij}$ and $\bar{\bar{E}}_{ij}$, and the moments of the supplies at each node, $\bar{S}_j$ and $\bar{\bar{S}}_j$, beginning with the first node and moving forward in time.

Step 3. Backward pass
Calculate the derivatives $\partial \pi / \partial \theta_{ij}$ and $\partial \pi / \partial \alpha_{ij}$ using the recursion for $\partial \pi / \partial \bar{S}_j$, beginning with the last node and working backward in time.

Step 4. Direction finding
Determining a trial solution $\{\theta_L{}^N\}$ and $\{\alpha_L{}^N\}$ by applying the logit splitting functions (Equations 24 and 25) using the concentration parameter $\gamma = \gamma(N)$.

Step 5. Step size
Find a stepsize β which solves:

$$\max_{0 \le \beta \le 1} \pi(\theta^N + \beta(\theta_L{}^N - \theta^N), \alpha^N + \beta(\alpha_L{}^N - \alpha^N))$$

using a Golden section search (see, for example, [7, p. 90]). Set

$$\theta^{N+1} = \theta^N + \beta(\theta_L{}^N - \theta^N)$$

$$\alpha^{N+1} = \alpha^N + \beta(\alpha_L{}^N - \alpha^N).$$

Step 6. Convergence
If $\pi(\theta^{N+1}, \alpha^{N+1}) - \pi(\theta^N, \alpha^N) < \varepsilon$, stop. Otherwise, set $N = N + 1$ and go to Step 2.

The initialization step makes use of a network simplex code to solve the problem deterministically. All supplies and demands are set equal to the means of their distributions and then assumed to be known exactly. All the nodes in the last planning period are joined to a supersink with links that generate no revenue. The other links in the network are divided between full movement links, with revenue r_{ij} and upper bound equal to the expected demand, and empty movement links with "revenue"—c_{ij} and no upper bound.

In Step 4, the algorithm calls for using Equations 24 and 25 for determining the trial values $\{\theta_L\}$ and $\{\alpha_L\}$, with a concentration parameter $\gamma(N)$ that is allowed to vary with the iteration number. If $\gamma(N)$ is set equal to a very large constant for all N, the effect is equivalent to using the Frank-Wolfe algorithm, where the solution of the linearized subproblem (20) results in an all-or-nothing choice for $\{\theta_L\}$ and $\{\alpha_L\}$. Let $\gamma_o = \gamma(0)$ be an appropriately scaled constant, where $\gamma_o \geq 0$. A sequence that is tested in the next section is $\gamma(N) = 2^N \cdot \gamma_o$, which is designed so that the choice of $\{\theta_L\}$ and $\{\alpha_L\}$ converges to the solution of (20). The use of $\gamma_o = 0$ is equivalent to dividing the flow among all the paths that show an increase in profit. The next section presents the results of numerical experiments that test the performance of the algorithm.

4. NUMERICAL EXPERIMENTS

A SERIES of numerical experiments were conducted to a) test the validity of certain independence assumptions required to calculate flows, b) evaluate the accuracy of the approximation that $\partial\pi/\partial\bar{\bar{S}}_j = 0$, c) compare the efficiency of alternative search algorithms, and d) demonstrate the importance of incorporating uncertainty in the model.

Model Validation

Two independence assumptions were made in Section 2 in order to efficiently calculate the moments of the flows over the network. The first is that flows coming from two separate nodes into the same node are independent, and the second is that the overflow from two links emanating from the same node i onto the overflow link $(i, \underline{i})$ are independent. A Monte Carlo simulation program was written which randomly generated initial supplies of trucks and then, using a set of decision variables $\{\theta_{ij}\}$ and $\{\alpha_{ij}\}$ as input as well as randomly generated demands on all the links, moved these supplies forward over the network. A network with 10 cities and 7 time periods was used. Initial supplies of vehicles were assumed to be described by an Erlang distribution with a standard deviation set at 20 and a mean chosen from a uniform distribution between 15 and 25. Demands were assumed to be described by an Erlang distribution with a mean chosen from a uniform distribution between 0 and 8 with a standard deviation equal to the mean. In both cases, adjustments were made to the standard deviation in the process of fitting the Erlang distribution to ensure the integrality of the shape parameter (see Equations 11 and 12).

The simulation was run with 100 repetitions, from which average flow values (for fulls and empties) were computed, along with confidence intervals for the true population means. A plot of observed versus predicted is shown in Figure 2, demonstrating excellent agreement between the two models. Ninety percent confidence intervals were constructed around each sample average, and it was found that 88% of the predicted flows fell within the appropriate confidence interval. The standard error between the observed and predicted means was 1.26. Based on these results, it seems safe to conclude that the analytical model is accurately estimating the mean link flows when the network is operated under the assumptions described in Section 1.

Evaluating the Approximation that $\partial\pi/\partial\bar{\bar{S}} = 0$

It was shown in Section 3 that the approximation $\partial\pi/\partial\bar{\bar{S}} = 0$ considerably simplifies the calculation of the derivatives $\partial\pi/\partial\theta_{ij}$ and $\partial\pi/\partial\alpha_{ij}$. To test the validity of the approximation, the same network described above was used, although now it is necessary to specify costs and revenues. Link distances were generated

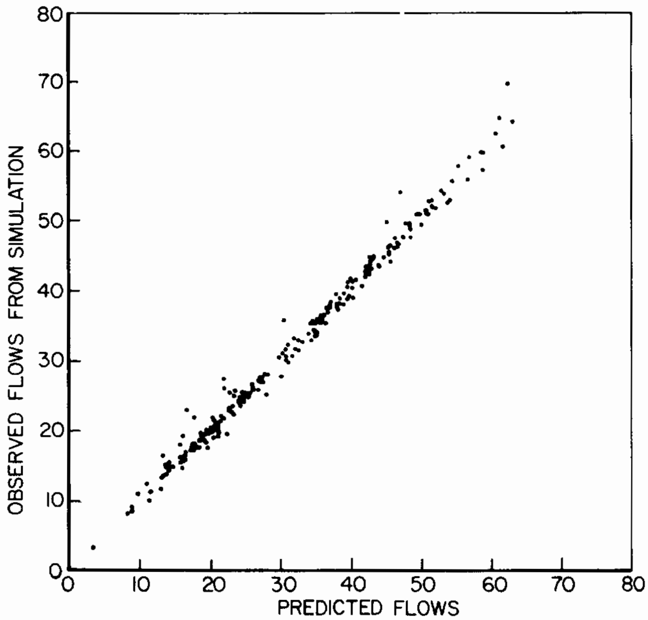

Fig. 2. Validation of estimates of predicted link flows.

randomly from a uniform distribution from 100 to 1000 miles. The cost of moving an empty truck was set at \$0.80 per mile, and the net revenue from moving a loaded truck was set at \$0.15 per mile. The cost of holding a truck at a location for 1 day was set at \$100.

Applying the relationships described in Appendix A, it was found that $\partial\pi/\partial\bar{\bar{S}}_j$ was smaller than $\partial\pi/\partial\bar{S}_j$ typically by a factor of 10^{-4}. After performing 20 iterations of a straight Frank-Wolfe algorithm, the difference in the optimum solution found was negligible. Using the exact derivatives, the optimum objective function was found to be 69167 after 30 iterations and required 20.2 CPU seconds. The approximate derivatives, on the other hand, produced an objective function of 69184 after the same number of iterations and required only 9.1 seconds. In view of these results, the assumption that $\partial\pi/\partial\bar{\bar{S}}_j = 0$ will be made in all remaining experiments.

Comparing Alternative Search Algorithms

There are two mechanisms by which the speed of the search process can be controlled. The first is in the solution of the linearized subproblem (20), where either a straight all-or-nothing choice of the trial solution $\{\theta_L\}$ and $\{\alpha_L\}$ may be used, or the logit splitting formulas (24) and (25) may be used. The second is the choice of step size, where a one dimensional search may be used or an alternative procedure such as one which simply sets the step size using a predetermined sequence. Beginning with the solution of the linearized subproblem, the primary question is the choice of the scaling parameter γ at each iteration. The decision was made to begin with an initial value γ^o, and to double it after every iteration so that as the algorithm progresses, the values of $\{\theta_L\}$ and $\{\alpha_L\}$ found using the logit splitting function converges to that produced by actually solving (20), which yields an all or nothing solution. Increasing γ at every iteration, then, helps to ensure convergence.

A series of experiments were conducted to help determine the best value of γ^o. Observing that the derivatives $\partial\pi/\partial\theta_{ij}$ tended to fall in the range of 100 to 2,000 for links in the first time period, experiments were run using $\gamma^o = 5 \times 10^{-3}$, 5×10^{-4} and 5×10^{-5}. The differences in the rate of convergence were not large. The best value, namely $\gamma^o = 5 \times 10^{-4}$, produced an objective function value of 66722 after 10 iterations, as compared to 66489 after 10 iterations of a pure Frank-Wolfe algorithm (which performed the worst).

Considerably greater savings were realized by improving the choice of stepsize. Calculating the optimum step size to an accuracy of $\varepsilon = 0.001$ over the reduced interval from 0 to 0.1 (instead of from 0 to 1) was found to make up over 90% of the total execution time. One method of reducing this requirement is to reduce the accuracy of the one dimensional search. Two alternatives are the following. If $\beta^{(N)}$ is the step size on the Nth iteration, then it is possible to use a predetermined sequence[5] where

$$\beta^{(N)} = a/N, \tag{39}$$

where a is a given constant. A value of $a = 0.1$ was chosen for the experiments reported here. A second alternative is to begin with a fixed step size, and to reduce the step size only when a reduction is needed to produce an increase in the objective function. This method is equivalent to:

$$\beta_{(N+1)} = \begin{cases} \beta^{(N)} & \text{if } \pi[\theta^{(N)} + \beta^{(N)}(\theta_L - \theta^{(N)}), \alpha^{(N)} + \beta^{(N)}(\alpha_L - \alpha^{(N)})] > \pi[\theta^{(N)}, \alpha^{(N)}] \\ \eta^K \cdot \beta^{(N)} & \text{otherwise} \end{cases} \tag{40}$$

where η is a predetermined constant less than 1 and K is the smallest positive integer such that

$$\pi[\theta^{(N)} + \eta^K \cdot \beta^{(N)}(\theta_L - \theta^{(N)}), \alpha^{(N)} + \eta^K \cdot \beta^{(N)}(\alpha_L - \alpha^{(N)})] > \pi[\theta^{(N)}, \alpha^{(N)}]. \tag{41}$$

Throughout this research, $\eta = 0.75$ was used and an initial step size of $\beta^{(1)} = 0.1$ was used.

Figure 3 compares the effectiveness of these different methods for finding the step size by showing the objective function versus the elapsed CPU time, which includes the time required to calculate the derivatives,

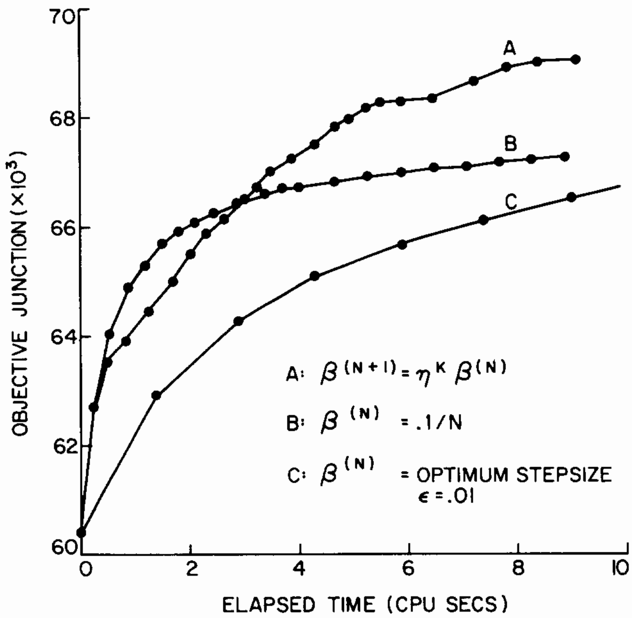

Fig. 3. Rate of convergence for different choices of step size.

but does not include data input or initialization. As is evident from the figure, the method using Equation 40 is substantially faster than the arithmetic sequence (39) or the optimum step size.

The Importance of Uncertainty

There are two issues associated with uncertainty. The first is the methodological issue of whether a deterministic model produces a sufficiently accurate solution, while the second addresses the more substantive question of the value of perfect (or better) information. The answer to both questions is, of course, highly problem dependent, but a few experiments do help to shed some light on the question.

The accuracy of the deterministic model can be seen by comparing three numbers, the value of the optimal solution of the deterministic objective function, the value of the stochastic objective function using the solution provided by the deterministic model and finally the value of the optimal solution of the stochastic objective function. These numbers were calculated for the same network described earlier, while varying the coefficient of variation of the demand distribution from 0.10 to 1.0. (Note that because of the integrality of the shape parameter, the actual coefficient of variation can only take on values from the sequence $\sqrt{1/k}$, for $k = 1, 2, 3, \cdots$.) The results of this experiment, shown in Table I, demonstrate the large discrepancy that can exist between the deterministic objective function (which is not affected by the variance of the demand distribution) and the corresponding value of the stochastic objective function. After optimizing the stochastic objective function, net revenue increased 14.4% for the most highly variable demands down to an increase of just 2.7% for the least variable demands. It is important to realize that some of this improvement can be attributed to the use of fractional vehicles, an important problem when flows are small.

The figures in Table I also provide some insight into the question of the cost of uncertainty. Profits increase over 57% when the coefficient of variation decreases from 1.0 to 0.6, and increases another 31% when the coefficient of variation is decreased to 0.4. Because of the assumptions underlying the model, these numbers are probably overstating the true cost of uncertainty. The methodology described in this paper does, however, provide a framework for developing improved models for investigating the cost of uncertainty.

5. DIRECTIONS FOR FURTHER RESEARCH

A STOCHASTIC formulation of the vehicle allocation problem raises a number of questions that do not arise in a deterministic model. The presence of stochastic supplies raises the question of the choice of appropriate decision variables, since the flow on a link, as used in a deterministic model, is no longer a well defined concept. The use of the decision variables $\{\theta\}$ and $\{\alpha\}$, which specify the fraction of the available supply that should move along each link, is only one of several possible approaches which should be compared in a systematic way.

A second but related issue is the development of a better model of the dispatching process at each location. It was assumed in this research that a certain fraction of the supply of vehicles is allocated for demands on a given link, and that if this supply exceeds the actual demand, then these vehicles must be held over at that location for another day. Such an assumption probably overstates the number of vehicles that could not move due to insufficient demand. A more accurate model would require developing a better understanding of how the dispatching process actually responds to uncertainty.

APPENDIX A

IN THIS appendix, the derivatives $\partial\pi/\partial\theta$ and $\partial\pi/\partial\alpha$ are found without using the approximation that $\partial\pi/\partial\bar{\bar{S}} = 0$. The presentation will work from Equations 26 and 27 in the text. Beginning with Equation 26, observe that

$$\frac{\partial \bar{S}_j}{\partial \alpha_{ij}} = \frac{\partial \bar{E}_{ij}}{\partial \alpha_{ij}} = \bar{S}_i. \tag{A.1}$$

TABLE I

Comparison of Solutions from Deterministic and Stochastic Models

	Coefficient of Variation of Demand Distribution					
	0.1	0.2	0.4	0.6	0.8	1.0
Optimal deterministic objective function[a]	163.4	163.4	163.4	163.4	163.4	163.4
Stochastic objective function with deterministic solution[a]	142.9	138.1	118.2	100.8	87.7	60.4
Optimal stochastic objective function[a]	146.7	142.7	125.1	108.9	96.3	69.1
Percent improvement over deterministic solution	2.7	3.3	5.8	8.0	9.8	14.4

[a] All numbers in thousands of dollars over planning horizon.

Next, differentiating (19) gives

$$\frac{\partial \bar{\bar{S}}_j}{\partial \alpha_{ij}} = 2\alpha_{ij}\bar{\bar{S}}_i + 2\,\mathrm{Cov}(S_i, F_{ij}) + \delta_{ij} \textstyle\sum_{k\in B_j}[\bar{S}_{\underline{j}}\theta_{\underline{j}k} - \mathrm{Cov}(S_{\underline{j}}, F_{\underline{j}k}] \quad \text{(A2)}$$

where

$$\delta_{ij} = \begin{cases} 1 & j = \underline{i} \\ 0 & \text{otherwise.} \end{cases}$$

Combining (26), (A.1), and (A.2) gives

$$\frac{\partial \pi}{\partial \alpha_{ij}} = \bar{S}_i\left(-c_{ij} + \frac{\partial \pi}{\partial \bar{S}_j}\right) + \left\{2\alpha_{ij}\bar{\bar{S}}_i + 2\,\mathrm{Cov}(S_i, F_{ij}) + \delta_{ij}\sum_{k\in B_{\underline{j}}}[\bar{S}_{\underline{j}}\theta_{\underline{j}k} - \mathrm{Cov}(S_{\underline{j}}, F_{\underline{j}k})]\right\}\frac{\partial \pi}{\partial \bar{\bar{S}}_j}. \quad \text{(A.3)}$$

The derivatives $\partial\pi/\partial\bar{S}_j$ and $\partial\pi/\partial\bar{\bar{S}}_j$ are calculated later.

Finding $\partial\pi/\partial\theta_{ij}$ uses logic similar to that used to find (A.3) but is complicated by the fact that changing θ_{ij} affects the flow both on the link (i, j) and the holdover link $(i, \underline{i})$. Keeping this in mind gives

$$\frac{\partial \pi}{\partial \theta_{ij}} = \left[(r_{ij} + s_{ij})\frac{\partial \bar{F}_{ij}}{\partial \theta_{ij}} + \frac{\partial \bar{S}_j}{\partial \theta_{ij}}\frac{\partial \pi}{\partial \bar{S}_j} + \frac{\partial \bar{\bar{S}}_j}{\partial \theta_{ij}}\frac{\partial \pi}{\partial \bar{\bar{S}}_j}\right] + \left[-c_{i\underline{i}}\left(\bar{S}_i - \frac{\partial \bar{F}_{ij}}{\partial \theta_{ij}}\right) + \frac{\partial \bar{S}_{\underline{i}}}{\partial \theta_{ij}}\frac{\partial \pi}{\partial \bar{S}_{\underline{i}}} + \frac{\partial \bar{\bar{S}}_{\underline{i}}}{\partial \theta_{ij}}\frac{\partial \pi}{\partial \bar{\bar{S}}_{\underline{i}}}\right]. \quad \text{(A.4)}$$

The terms in the first set of brackets in (A.4) reflects the direct impact of an increase in θ_{ij} on link (i, j) and the subsequent impact on profits from a change in $\bar{S}_j$ and $\bar{\bar{S}}_j$. The second set of brackets reflects the impact of θ_{ij} on the overflow of vehicles that could not be filled to the holdover link, where again the impact on the moments of $S_{\underline{i}}$ on future profits must be accounted for.

To calculate the terms of (A.4) note that $\partial\bar{S}_j/\partial\theta_{ij} = \partial\bar{F}_{ij}/\partial\theta_{ij}$ and $\partial\bar{S}_{\underline{i}}/\partial\theta_{ij} = \bar{S}_i - \partial\bar{F}_{ij}/\partial\theta_{ij}$. Differentiating (19) gives:

$$\frac{\partial \bar{\bar{S}}_j}{\partial \theta_{ij}} = \frac{\partial \bar{\bar{F}}_{ij}}{\partial \theta_{ij}} + 2\alpha_{ij}\frac{\partial}{\partial \theta_{ij}}\mathrm{Cov}(S_i, F_{ij}) \qquad j \neq \underline{i} \quad \text{(A.5)}$$

$$\frac{\partial \bar{\bar{S}}_{\underline{i}}}{\partial \theta_{ij}} = \frac{\partial \bar{\bar{F}}_{ij}}{\partial \theta_{ij}} + \alpha_{ij}\bar{S}_{\underline{j}} + 2\theta_{ij}\bar{S}_{\underline{j}} - (\alpha_{\underline{j}j} + 2\theta_{ij})\frac{\partial}{\partial \theta_{ij}}\mathrm{Cov}(S_i, F_{ij}) - 2\,\mathrm{Cov}(S_i, F_{ij}). \quad \text{(A.6)}$$

Combining (A.4), (A.5), and (A.6) gives

$$\frac{\partial \pi}{\partial \theta_{ij}} = \left(r_{ij} + s_{ij} + \frac{\partial \pi}{\partial \bar{S}_j}\right)\frac{\partial \bar{F}_{ij}}{\partial \theta_{ij}} + \left(-c_{i\underline{i}} + \frac{\partial \pi}{\partial \bar{S}_{\underline{i}}}\right)\left(\bar{S}_i - \frac{\partial \bar{F}_{ii}}{\partial \theta_{ij}}\right) + \left[\frac{\partial \bar{\bar{F}}_{ij}}{\partial \theta_{ij}} + 2\alpha_{ij}\frac{\partial}{\partial \theta_{ij}}\mathrm{Cov}(S_i, F_{ij})\right]\frac{\partial \pi}{\partial \bar{\bar{S}}_j} + \left[\frac{\partial \bar{\bar{F}}_{ij}}{\partial \theta_{ij}} + (\alpha_{i\underline{i}} + 2\theta_{ij})\left(\bar{S}_i - \frac{\partial}{\partial \theta_{ij}}\mathrm{Cov}(S_i, F_{ij})\right) - 2\,\mathrm{Cov}(S_i, F_{ik})\right]\frac{\partial \pi}{\partial \bar{\bar{S}}_{\underline{i}}}. \quad \text{(A.7)}$$

Calculating $\partial\bar{F}_{ij}/\partial\theta_{ij}$, $\partial\bar{\bar{F}}_{ij}/\partial\theta_{ij}$ and $\partial\,\mathrm{Cov}(S_i, F_{ij})/\partial\theta_{ij}$ simply involves differentiating (13), (14), and (15), with respect to θ_{ij}. These expressions are given in Appendix B.

The next problem is calculating the derivatives $\partial\pi/\partial\bar{S}_j$ and $\partial\pi/\partial\bar{\bar{S}}_j$. These calculations can be handled efficiently by setting up a recursion where, for example, $\partial\pi/\partial\bar{S}_j$ would be expressed in terms of $\partial\pi/\partial\bar{S}_k$ and $\partial\pi/\partial\bar{\bar{S}}_k$, $k \in B_j$. Beginning with $\partial\pi/\partial\bar{S}_j$, the recursion is given by:

$$\frac{\partial \pi}{\partial \bar{S}_j} = \sum_{k\in B_j}\left[(r_{jk} + s_{jk})\frac{\partial \bar{F}_{jk}}{\partial \bar{S}_j} - c_{jk}\frac{\partial \bar{E}_{jk}}{\partial \bar{S}_j} + \frac{\partial \bar{S}_k}{\partial \bar{S}_j}\frac{\partial \pi}{\partial \bar{S}_k} + \frac{\partial \bar{\bar{S}}_k}{\partial \bar{S}_j}\frac{\partial \pi}{\partial \bar{\bar{S}}_k}\right] - c_{j\underline{j}}\frac{\partial \bar{E}_{j\underline{j}}}{\partial \bar{S}_j} + \frac{\partial \bar{S}_{\underline{j}}}{\partial \bar{S}_j}\frac{\partial \pi}{\partial \bar{S}_{\underline{j}}} + \frac{\partial \bar{\bar{S}}_{\underline{j}}}{\partial \bar{S}_j}\frac{\partial \pi}{\partial \bar{\bar{S}}_{\underline{j}}}. \quad \text{(A.8)}$$

To calculate $\partial\bar{\bar{S}}_k/\partial\bar{S}_j$, Equation 19 is again used to find:

$$\frac{\partial \bar{\bar{S}}_k}{\partial \bar{S}_j} = \frac{\partial \bar{\bar{F}}_{jk}}{\partial \bar{S}_j} + 2\alpha_{jk}\frac{\partial}{\partial \bar{S}_j}\mathrm{Cov}(S_j, F_{jk}) \quad k \neq \underline{j} \quad \text{(A.9)}$$

$$\frac{\partial \bar{\bar{S}}_{\underline{j}}}{\partial \bar{S}_j} = \sum_{k\in B_j}\left[\frac{\partial \bar{\bar{F}}_{jk}}{\partial \bar{S}_j} - (\alpha_{jj} + 2\theta_{jk})\frac{\partial}{\partial \bar{S}_j}\mathrm{Cov}(S_j, F_{jk})\right]. \quad \text{(A.10)}$$

The expressions for $\partial\bar{F}_{jk}/\partial\bar{S}_j$, $\partial\bar{\bar{F}}_{jk}/\partial\bar{S}_j$ and $\partial\,\mathrm{Cov}(S_j, F_{jk})/\partial\bar{S}_j$ are presented in Appendix B. Combining

(A.8)–(A.10), together with (34)–(37) in the text, gives

$$\frac{\partial\pi}{\partial\bar{S}_j}=\sum_{k\in B_j}\Bigg[(r_{jk}+s_{jk})\frac{\partial\bar{\bar{F}}_{ik}}{\partial\bar{S}_j}-c_{jk}\alpha_{jk}+\left(\frac{\partial\bar{F}_{ij}}{\partial\bar{S}_j}+\alpha_{jk}\right)\frac{\partial\pi}{\partial\bar{S}_k}$$

$$+\left\{\frac{\partial\bar{\bar{F}}_{jk}}{\partial\bar{S}_j}+2\alpha_{jk}\frac{\partial}{\partial\bar{S}_j}\mathrm{Cov}(S_j,F_{jk})\right\}\frac{\partial\pi}{\partial\bar{\bar{S}}_k}$$

$$+\left\{\frac{\partial\bar{\bar{F}}_{jk}}{\partial\bar{S}_j}-(\alpha_{jj}+2\theta_{jk})\frac{\partial}{\partial\bar{S}_j}\mathrm{Cov}(S_j,F_{jk})\right\}\frac{\partial\pi}{\partial\bar{\bar{S}}_j}\Bigg]$$

$$+\left(\frac{\partial\pi}{\partial\bar{S}_j}-c_{jj}\right)\left(\alpha_{jj}+\sum_{k\in B_j}\left(\theta_{jk}-\frac{\partial\bar{F}_{jk}}{\partial\bar{S}_j}\right)\right). \quad \text{(A.11)}$$

Finally, a recursion for $\partial\pi/\partial\bar{\bar{S}}_j$ is needed. The steps required to set up this recursion closely follow those used to develop that used to calculate $\partial\pi/\partial\bar{S}_j$. Repeating these steps produces the following relationship:

$$\frac{\partial\pi}{\partial\bar{\bar{S}}_j}=\sum_{k\in B_j}\Bigg[\left(r_{jk}+s_{jk}+\frac{\partial\pi}{\partial\bar{S}_k}-\frac{\partial\pi}{\partial\bar{S}_j}\right)\frac{\partial\bar{F}_{jk}}{\partial\bar{\bar{S}}_j}$$

$$+\left(\frac{\partial\bar{\bar{F}}_{jk}}{\partial\bar{\bar{S}}_j}+\alpha_{jk}^2+2\alpha_{jk}\frac{\partial}{\partial\bar{\bar{S}}_j}\mathrm{Cov}(S_j,F_{jk})\right)\frac{\partial\pi}{\partial\bar{\bar{S}}_k}$$

$$+\left(\frac{\partial\bar{\bar{F}}_{jk}}{\partial\bar{\bar{S}}_j}+\alpha_{jj}\theta_{jk}-(\alpha_{jj}+2\theta_{jk})\right. \quad \text{(A.12)}$$

$$\left.\cdot\frac{\partial}{\partial\bar{\bar{S}}_j}\mathrm{Cov}(S_j,F_{jk})\right)\cdot\frac{\partial\pi}{\partial\bar{\bar{S}}_j}\Bigg]+\alpha_{jj}^2\frac{\partial\pi}{\partial\bar{\bar{S}}_j}.$$

Expressions for $\partial\bar{F}_{ij}/\partial\bar{\bar{S}}_{ij}$, $\partial\bar{\bar{F}}_{ij}/\partial\bar{\bar{S}}_i$ and $\partial\ \mathrm{Cov}(S_i, F_{ij})/\partial\bar{\bar{S}}_i$ are given in Appendix B.

APPENDIX B—EXPRESSIONS FOR THE DERIVATIVES

LET $r_\theta(a, b) = \partial r(a, b)/\partial\theta$ and $r_\lambda(a, b) = \partial r(a, b)/\partial\lambda$, where $r(a, b)$ is given by Equation 17 in the text. It is easily verified that

$$r_\theta(a, b) = r(a, b)\,\frac{a\lambda - b\mu\theta}{\mu\theta^2 + \theta\lambda} \quad \text{(B.1)}$$

and

$$r_\lambda(a, b) = r(a, b)\,\frac{b\mu\theta - a\lambda}{\mu\lambda\theta + \lambda^2}. \quad \text{(B.2)}$$

The derivatives $\partial\bar{F}_{ij}/\partial\theta_{ij}$, $\partial\bar{\bar{F}}_{ij}/\partial\theta_{ij}$ and $\partial\ \mathrm{Cov}(S_i, F_{ij})/\partial\theta_{ij}$ are given by (dropping the subscripts i and j):

$$\frac{\partial\bar{F}}{\partial\theta}=\frac{\bar{F}}{\theta}-\frac{\theta l}{\lambda}$$

$$\cdot\sum_{n=1}^{\kappa}\left[\frac{n}{(\kappa-n)}\binom{\kappa+l-n-1}{l}r_\theta(l,\kappa-n)\right] \quad \text{(B.3)}$$

$$\frac{\partial\bar{\bar{F}}}{\partial\theta}=\frac{2(\bar{\bar{F}}+\bar{F}^2)}{\theta}-\frac{\theta^2 l}{\lambda^2}$$

$$\cdot\sum_{n=1}^{\kappa+1}\left[\frac{n(2\kappa+1-n)}{(\kappa-n+1)}\binom{\kappa+l-n}{l}\right. \quad \text{(B.4)}$$

$$\left.\cdot r_\theta(l,\kappa-n+1)\right]-2\bar{F}\frac{\partial\bar{F}}{\partial\theta}$$

$$\frac{\partial\ \mathrm{Cov}(S,F)}{\partial\theta}$$

$$=\frac{(\mathrm{Cov}(S,F)+\bar{S}\bar{F})}{\theta}-\frac{\theta\kappa l}{\lambda^2}$$

$$\cdot\sum_{n=1}^{\kappa+1}\left[\frac{n}{\kappa-n+1}\binom{\kappa+l-n}{l}r_\theta(l,\kappa-n+1)\right] \quad \text{(B.5)}$$

$$-\bar{S}\frac{\partial\bar{F}}{\partial\theta}.$$

The derivatives of $\bar{F}$, $\bar{\bar{F}}$ and Cov(S, F) with respect to $\bar{S}$ and $\bar{\bar{S}}$ were calculated using the chain rule as follows:

$$\frac{\partial\bar{F}}{\partial\bar{S}}=\frac{\partial\bar{F}}{\partial\lambda}\frac{\partial\lambda}{\partial\bar{S}}+\frac{\partial\bar{F}}{\partial\kappa}\frac{\partial\kappa}{\partial\bar{S}} \quad \text{(B.6)}$$

$$\frac{\partial\bar{F}}{\partial\bar{\bar{S}}}=\frac{\partial\bar{F}}{\partial\lambda}\frac{\partial\lambda}{\partial\bar{\bar{S}}}+\frac{\partial\bar{F}}{\partial\kappa}\frac{\partial\kappa}{\partial\bar{\bar{S}}}. \quad \text{(B.7)}$$

Similar expressions are used for the derivatives of $\bar{\bar{F}}$ and Cov(S, F) with respect to $\bar{S}$ and $\bar{\bar{S}}$. The derivatives $\partial\lambda/\partial\bar{S}$, $\partial\kappa/\partial\bar{S}$, $\partial\lambda/\partial\bar{\bar{S}}$ and $\partial\kappa/\partial\bar{\bar{S}}$ were calculated using

$$\lambda = \bar{S}/\bar{\bar{S}} \quad \text{(B.8)}$$

$$\kappa = \bar{S}^2/\bar{\bar{S}}. \quad \text{(B.9)}$$

Differentiating (B.8) and (B.9) gives

$$\frac{\partial\lambda}{\partial\bar{S}}=\frac{1}{\bar{\bar{S}}} \quad \text{(B.10)}$$

$$\frac{\partial\lambda}{\partial\bar{\bar{S}}}=-\frac{\bar{S}}{\bar{\bar{S}}^2} \quad \text{(B.11)}$$

$$\frac{\partial\kappa}{\partial\bar{S}}=\frac{2\bar{S}}{\bar{\bar{S}}} \quad \text{(B.12)}$$

$$\frac{\partial\kappa}{\partial\bar{\bar{S}}}=-\frac{\bar{S}^2}{\bar{\bar{S}}^2}. \quad \text{(B.13)}$$

The derivatives of $\bar{F}$, $\bar{\bar{F}}$ and Cov(S, F) with respect to κ were found numerically. The derivatives with respect to λ are very similar to those with respect to θ, using $r_\lambda(a, b)$ in place of $r_\theta(a, b)$ in addition to other minor changes. As the steps are very straightforward, the expressions are not included.

ACKNOWLEDGMENTS

THIS MATERIAL is based upon work supported by the National Science Foundation under grant CEE-8203476.

REFERENCES

1. L. COOPER AND L. J. LEBLANC, "Stochastic Transportation Problems and Other Network Related Convex Problems," *Naval Res. Log. Quart.* **24,** 327–337 (1977).
2. W. C. JORDAN, "The Impact of Uncertain Demand and Supply on Empty Rail Car Distribution," Ph.D. dissertation, Cornell University, Ithaca, N.Y., January 1982.
3. W. C. JORDAN AND M. A. TURNQUIST, "A Stochastic, Dynamic Model for Railroad Car Distribution," *Trans. Sci.* **17,** 123–145 (1983).
4. L. J. LEBLANC, R. V. HELGASON AND D. E. BOYCE, "Improved Efficiency of the Frank-Wolfe Algorithm," Working paper No. 81-131, Owen Graduate School of Management, Vanderbilt University, Nashville, Tenn., 1982.
5. W. B. POWELL AND Y. SHEFFI, "The Convergence of Equilibrium Algorithms with Predetermined Step Sizes," *Trans. Sci.* **16,** 45–55 (1982).
6. W. B. POWELL, Y. SHEFFI AND S. THIRIEZ, "The Dynamic Vehicle Allocation Problem with Uncertain Demands," In *Ninth International Symposium on Transportation and Traffic Theory,* (J. Volmuller and R. Hamerslag (eds.), VNU Science Press, The Netherlands, 1984.
7. P. E. GILL, W. MURRAY AND M. H. WRIGHT, *Practical Optimization,* Academic Press, New York, 1981.