Maximizing Profits for North American Van Lines' Truckload Division: A New Framework for Pricing and Operations

WARREN B. POWELL — *Department of Civil Engineering and Operations Research, Princeton University, Princeton, New Jersey 08544*

YOSEF SHEFFI — *Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

KENNETH S. NICKERSON — *Department of Operations Research, Stanford University, Stanford, California 94305*

KEVIN BUTTERBAUGH — *North American Van Lines — Commercial Transport Division, PO Box 988, Ft. Wayne, Indiana 46801*

SUSAN ATHERTON — *North American Van Lines — Information Services*

The Commercial Transport Division of North American Van Lines dispatches thousands of trucks from customer origin to customer destination each week under high levels of demand uncertainty. Working closely with upper management, the project team developed a new type of network model for assigning drivers to loads. The model, LOADMAP, combines real-time information about drivers and loads with an elaborate forecast of future loads and truck activities to maximize profits and service. It provided management with a new understanding of the economics of truckload operations; integrated load evaluation, pricing, marketing, and load solicitation with truck and load assignment; and increased profits by an estimated $2.5 million annually, while providing a higher level of service.

The basic operating characteristics of a truckload motor carrier are deceptively simple: shippers provide information on trailer loads of freight to be hauled from an origin city, on a given day, to a destination city. The carrier must provide a truck at the origin at the right time or potentially lose the load. While en route, the truck is dedicated to its assignment; once the load is delivered at the destination, the truck is available for another dispatch. Thus the freight is not handled at intermediate locations, as would be the case with less-than-

0091-2102/88/1801/0021$01.25

TRANSPORTATION — ROAD
NETWORKS/GRAPHS — APPLICATION

truckload carriers, and in general, the trucks do not visit multiple origins and multiple destinations on a single trip, as would be the case with household movers. In a truckload operation, the vehicle fleet typically moves in response to demand without any base terminals or fixed schedules.

This seemingly simple technology actually poses some difficult challenges to large truckload carriers who operate thousands of trucks and can dispatch hundreds of drivers per hour. Dispatchers must assign individual drivers to loads while minimizing *deadhead* miles (empty miles from a driver's location to a pickup point). Thus, in order to satisfy the customers' needs and minimize its own costs, a carrier must have its trucks placed as close as possible to the pickup locations. Truck positions, however, are the result of previous dispatching decisions, and so current decisions affect future service and profitability. Thus, when assigning individual trucks to loads, dispatchers try also to ensure that these trucks will be well positioned once their mission is complete.

Randomness in shipper demands makes this task very difficult. At the beginning of a typical day a carrier may know only 30 to 40 percent of the loads it will carry that day and as little as 10 percent of the next day's loads, and thus dispatchers are never certain of what will be available at the end of each run. In addition, decision lead times are very short. When a driver calls in, the dispatcher must either assign the truck to a known load, dispatch it empty to a *deficit* region (one in which there are typically more loads than trucks), or hold the truck in anticipation of a future load at its current location. These decisions are made continuously by dozens of dispatchers simultaneously over the entire fleet positioned across the country. In this environment the Operations Department tends to respond to immediate problems with little time for gathering information and little opportunity for planning ahead.

Carrier Issues

An important management concern in truckload operations is minimizing empty miles. Empty movements result both from structural imbalances in freight flows between producing and consuming regions and from the random nature of shipper demands. The structural imbalances create a recurring need to reposition trucks from surplus to deficit regions. Even if flows were perfectly balanced, however, randomness would cause demand requests to materialize in places and times which did not correspond exactly to current truck positions. As a result, the need to maintain a high level of equipment utilization would force additional deadheading. In fact, random effects are strong enough to cause a traditional surplus region to become a deficit region on any given day.

The high pressure environment, coupled with the demand uncertainty, provides only limited time for the carrier to plan each dispatch, creating the potential for excessive empty miles. The more serious effect of this fire-fighting mode of dispatching, however, is that the carrier passes up business opportunities without knowing they were there. The carrier's loss of this "invisible freight" results from

the fact that shipper representatives often call for service following a list of approved carriers. A carrier ranked high on this list will get as many loads as it has trucks on the day and at the place where the loads are available. It will lose any load which it cannot commit to carrying, even though such loads may be very profitable.

In addition to getting enough hauling capacity to the right locations, truckload carriers evaluate each load request before accepting it. (Different loads may carry different revenues even between the same origin-destination pair.) Freight from certain large shippers with which the carrier has long-term relationships is always accepted if trucks are available. Other loads, however, can be accepted or rejected depending on whether they add positive marginal contributions to the system.

How much a given load contributes to the carrier's profit should be determined by taking into account not only the revenue minus direct cost (direct contribution) of that load, but also the expected earnings of the trailer upon its arrival at the destination. These earnings in turn should reflect the outbound loaded opportunities at the destination and the supply of trucks there. Furthermore, the entire contribution of each load has to be balanced against the availability of other loads and the opportunity cost of using a trailer to carry the load under study. In short, load acceptance is an important issue for truckload carriers, and it should be based on systemwide considerations that account for all regions at all future time periods. The carrier cannot carry out such an analysis manually in the short time it has to respond to the shipper, and thus loads are typically accepted if the direct contribution is positive and the capacity is there.

Beyond load acceptance and truck dispatching, truckload carriers also have to determine movement prices. Since all carriers use the same equipment and provide the same service in comparable transit times, pricing is very important. Without

---

In a truckload operation, the vehicle fleet typically moves in response to demand without any base terminals or fixed schedules.

---

the infrastructure and fixed schedules of other transportation modes, the underlying structure supporting a truckload carrier is the market itself — the multidirectional patterns of loads and empties that move continuously throughout the country. In addition to dispatching trucks and screening load requests, truckload carriers are engaged in an extremely delicate balancing act of setting prices in thousands of interrelated traffic markets simultaneously. These prices must reflect complex backhaul opportunities, which are a function of the flows in the system, which in turn are a function of prices.

Interestingly, these issues are less problematic for small truckload carriers that operate a few trucks in a limited number of traffic lanes. While such carriers are more susceptible than large ones to demand fluctuations, they can easily solve issues of dispatching, load screening, and

pricing due to the virtual lack of interactions across truck assignments and traffic lanes.

North American Commercial Transport (NACT)

The truckload market in the US is currently characterized by excess capacity and therefore by strong competition. It is a market where many customers (shippers) perceive the product offered as generic in terms of the equipment and level of service offered.

The Commercial Transport Division of North American Van Lines is one of the nation's largest truckload motor carriers, with annual revenues in excess of $260 million and a fleet of more than 5,800 trailers. In its competitive market, NACT strives to be customer-oriented rather than operations-oriented. In other words, NACT does not operate in a fixed pattern and expect shippers to use its services whenever they fit their needs. Instead, it tries to understand its customers' needs and to tailor trucking services to fit those needs. This emphasis on a high level of service, in conjunction with the competition from hundreds of small but highly efficient regional operators, challenges NACT to utilize its size to its advantage. Properly managed, a large carrier can provide shippers with the right truck at the right place at the right time more often than any small operator or a combination of small operators.

To achieve these advantages, NACT proposed the development of a computer model that would help it manage the complexities of a large operation and reap the benefits of its scale. While research on the theoretical aspects of this problem had been going on at universities for several years, no satisfactory solution was available.

The project resulted in a sophisticated new package named LOADMAP (Load Matching and Pricing) that challenges fundamental operating procedures used by truckload carriers. Standard practices of minimizing empty miles within artificial boundaries have been replaced with logic which single-mindedly maximizes profits and, interestingly, also increases service levels systemwide. Outputs from the model also assist sales and marketing personnel to identify those loads and traffic lanes with the highest marginal contribution to profits.

The development of LOADMAP required more than an application of existing management science techniques. It involved the development of a new stochastic network optimization model that handles forecasting uncertainties in a novel way, overcoming important practical difficulties in deterministic models [White 1972; Ouimet 1972] and in simpler stochastic models that have been tried previously [Jordan and Turnquist 1983; Powell, Sheffi, and Thiriez 1984]. Following Powell [1987], this new modeling framework is, in fact, directly applicable to solving a wide range of problems which have spatial, temporal, and stochastic elements (such as rail car distribution and rental vehicle management).

The Modeling Approach

The problem LOADMAP solves can be grasped by focusing on the decision to dispatch a truck to haul a load from, say, region 1 on Monday, arriving at region 5 on Wednesday (Figure 1). Once the truck

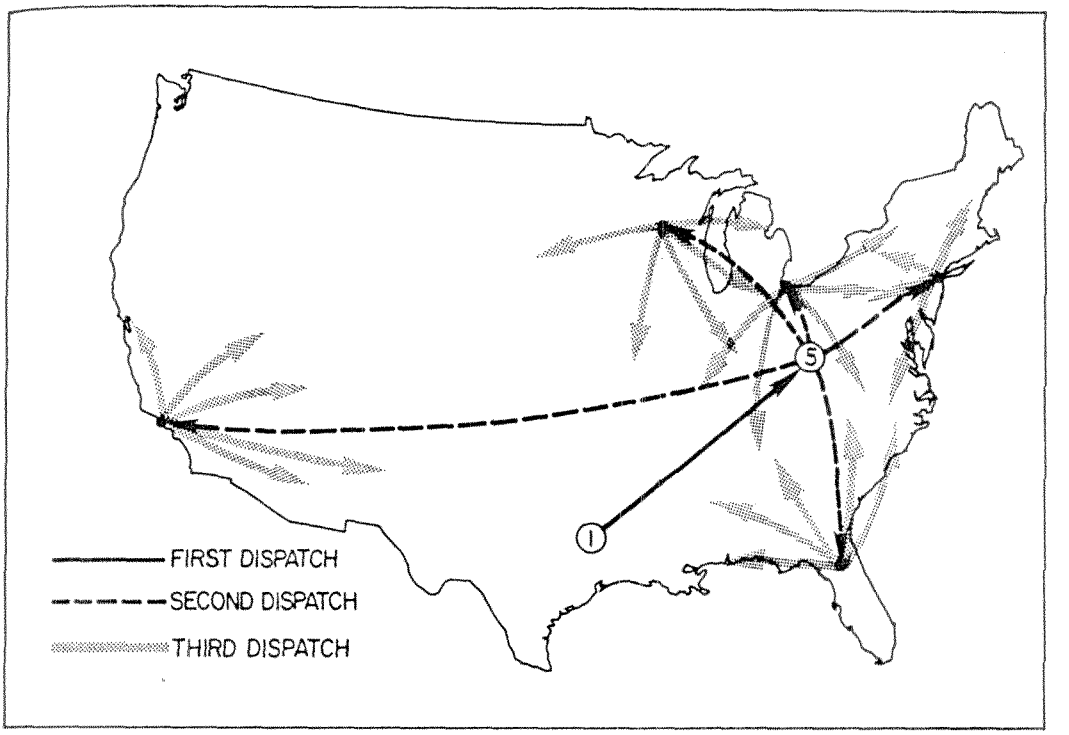

Figure 1: The dispatch of a truck from region 1 to region 5 must be evaluated in the context of the different dispatches that may be made out of region 5, which in turn must account for any subsequent downstream dispatches.

arrives at region 5, it may be dispatched in any one of five different directions, arriving at a new destination anywhere from Thursday to Monday. From these points the truck may find yet another five possible outbound dispatches, creating 25 different possible trajectories by the end of its second move (beyond the one evaluated). Using a more realistic average of 30 possible outbound dispatches from a region, after three moves there would be $30^3 = 27{,}000$ possible trajectories. In addition, since we do not know what loads will actually be available as the truck arrives in each destination, we can only probabilistically guess which trajectory should be used. This is further complicated by the presence of other trucks in those destination regions — these trucks may take these "probabilistic loads" before the truck under study. The problem then is to decide how to evaluate the contribution of a movement, given all its possible ramifications.

Note that the movement from region 1 to region 5 has to be compared with a possible move from region 1 to other regions — with all possible trajectories out of those regions. It also has to be compared with the possibility of holding the truck in region 1, out of which many potential trajectories may materialize in the future.

To solve this problem LOADMAP builds on a set of 60-80 geographically defined regions used by NACT. It then works

with a set of historical data that is updated in real time with information on loads and truck movements. The historical data is based on several months of actual loads and is updated daily using time-series models developed by NACT. This data base includes the following:

— The expected number of loads between each pair of regions over the planning horizon,
— The expected direct contribution (revenue minus direct operating costs) of each load and its expected transit time, and
— The expected cost and transit time for moving empty between any two regions.

The real-time data includes the following:

— The current location and status of each truck in the system (including the expected arrival time of trucks in transit and the estimated time for them to be ready for a new assignment),
— A list of all known (booked) loads which have been called in by shippers but not yet assigned to a driver, and
— The direct contribution of each known load, its pickup date and the time required for the move.

This information is used to develop an estimate of what will happen to a truck once it is sent into a region. The complex and uncertain set of possible trajectories is handled by tracing the future path of a truck in three stages:

(1) *The deterministic movements* are the sequence of one or more movements whose characteristics (revenue, cost, departure time and arrival time) are known at the time of the first dispatch. These movements include carrying known loads, empty moves, and holding a truck in position.

(2) *The first uncertain dispatch* represents the first movement beyond the sequence of deterministic moves. The characteristics of that move are not known at the time of the first dispatch.

(3) *All further uncertain movements* are those beyond the first uncertain dispatch until the end of the planning horizon.

These three stages reflect the amount of information available at the time of a dispatch decision. In stage 1, we know the exact contribution of each movement (negative if it is an empty move) and both the destination and the time of arrival at each destination. In stage 2, we can only estimate stochastically the outbound opportunities, but we know at least the location of a truck when it begins the second stage (since we are able to track the vehicle deterministically through the first stage). By the time we get to stage 3, not only are we forced to forecast the available opportunities, but we do not even know the location of the truck.

Our solution approach models each stage with a level of detail that matches the quality of the information available.

Stage 3: End Effects

Stage 3 deals with those truck movements that are the furthest into the future. It is modeled by developing a single number, $p(j,s)$, termed an *end effect*, which gives the value of a truck in region j on day s. This number is calculated by using dynamic-programming-style recursions to track the forward trajectory of a truck based on historical loaded and empty flows (see appendix). An end effect is thus an estimate of the expected contribu-

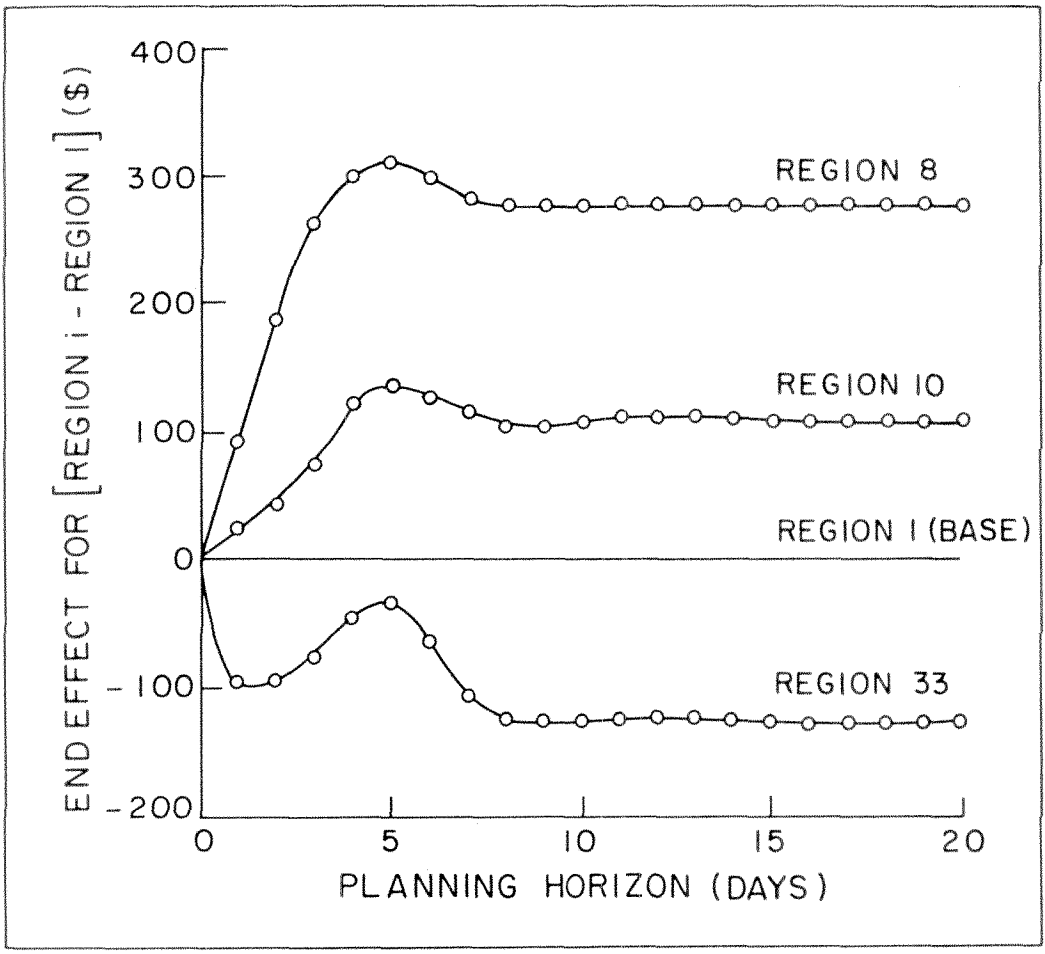

Figure 2: The average contribution of a truck in different regions, relative to region 1, becomes constant as the planning horizon is increased, suggesting that planning past 10 or 15 days will have no impact on decisions today.

tion of a truck, until the end of the planning horizon, independent of the forecasted number of trucks in a region. It indirectly reflects, however, the past supply of trucks and the opportunities available to them in that region on that day of the week.

Figure 2 shows the relative end effect values of three regions calculated for each day until the end of a 20-day planning period. The relative end effects are calculated by subtracting some base region (in this case the end effect of region 1) from the end effects of each of the other regions. The importance of relative end effects is that they describe the short term advantage of using a truck in one region over another. As the curves clearly show, after approximately 10 days the end effects for all regions become constant.

Figure 2 is based on actual data and it demonstrates that once the planning horizon is longer than 10 days, the relative end effects would not change by

increasing the planning period. (LOADMAP sets this period to 20 days as a safety factor.) Figure 2 also highlights the fact that the value of a truck in different regions can be very different, reflecting the dispatch opportunities out of that region.

Stage 2: The Marginal Truck

The goal in modeling stage 2 of a truck trajectory is to capture the marginal value of an additional truck in a region at some future time. It is modeled by looking at what a dispatcher would do, say, on day 3 in Boston. We do not know with certainty what options will be available to the dispatcher, but we can consider what might happen by evaluating historical trends and assuming that the dispatcher will use trucks in the most profitable manner. Table 1 illustrates the situation by showing the first nine possible ways (out of many more possibilities) in which a truck in Boston on day 3 might be used: first, loaded to Pittsburgh, second, loaded to Chicago, and so on. If, when we actually reach day 3, there is a load to Pittsburgh, the historical average contribution of that load (revenue minus direct operating costs) is $75 (see the "expected contribution" column). We then need to factor in how valuable a truck will be in Pittsburgh, where it might be leaving on day 5. This is given by the end effect, which in this (realistic) example is $1,880. The $1,880 is the expected total contribution a truck will earn starting in Pittsburgh on day 5 until the end of the planning horizon (which was set at 20 days for this example). The $1,880 by itself is not meaningful; what is important is its value relative to the end effects of other regions.

Adding the direct contribution to the end effect, we can rank the different options that may be available to the dispatcher in terms of expected profitability as shown in Table 1 (see the "total" column). We can now assume that if the dispatcher were to have one truck in Boston on day 3, he would use it on the highest ranked available option. The second truck would then be used on the next highest option and so on. As a result, we can develop the probability of dispatching the first, second, . . . k-th truck from Boston on day 3 to any one of the possible assignments it might have. (The idea here is that the first truck in a region has all the

Assignment	Expected Contribution	End Effect	Total	Dispatch Probabilities Truck 1	Truck 2	Truck 3	Truck 4
Loaded to Pittsburgh	75	1880	1955	.295	.049	.006	0
Loaded to Chicago	190	1741	1931	.167	.080	.020	.004
Loaded to Miami	250	1567	1817	.021	.013	.004	0
Loaded to Baltimore	40	1500	1540	.015	.010	.003	0
Loaded to New York	55	1452	1507	.427	.580	.449	.259
Loaded to Dallas	330	1175	1505	.027	.077	.109	.102
Empty to New York	-35	1509	1474	.001	.004	.007	.007
Loaded to Denver	450	1014	1464	.007	.023	.036	.038
Empty to Harrisburg	-40	1478	1438	.030	.117	.227	.298
Expected Contribution				1713	1550	1477	1433

Table 1: Possibilities open to a truck located in Boston on day 3.

options open to it out of the loads available then and there, the second truck has all the options except the best one taken by the first truck, and so on.) The dispatch probabilities for the first four trucks for each of the nine assignment options are shown in the last four columns of Table 1.

Since they are based on historical data, the dispatch probabilities incorporate the forecasted number of loads from Boston on day 3 to each of the potential destinations. Thus, the first truck has a 29.5 percent probability of being dispatched on the first option (loaded to Pittsburgh) but its most likely dispatch is loaded to New York City, an assignment that ranks only fifth on the list. The second truck has a lower probability of being dispatched to Pittsburgh but an even higher probability of moving loaded to New York.

Consider now the column labeled "Truck 1" in Table 1, depicting the dispatch probabilities of the first truck. Multiplying these probabilities by the total expected contribution of each assignment (the "total" column in Table 1), the column sum total will be the expected contribution of the first truck in Boston on day 3. These calculations can be performed for all possible trucks to give the value of each additional truck in the region and day under study. As should be expected, the marginal truck values decline monotonically from one truck to the next.

The calculation of the marginal truck values for each potential truck in each region at each point in time is the heart of the model. These calculations mean that LOADMAP handles forecasting uncertainties in a natural and rigorous way, capturing relationships between total profits and the number of trucks sent into a region. The smoothness of the trade-off between the number of trucks in a region and the value of each additional truck gives the model considerable stability in both dispatching and the calculation of the value of an additional load into a region.

Stage 1: Known Contributions

Having developed a probabilistic model of truck movements through the second and third stages of its trajectories (in decreasing detail), modeling the first stage is more straightforward. It simply requires adding up the direct contributions of the deterministic loaded and empty moves that make up this stage. Thus if LOADMAP evaluates a movement from the current location of a truck to, say, Washington, DC, arriving on day 2, and from there to Boston, arriving on day 3, the contribution of that part of the movement is simply the sum of these two known direct contributions (to Washington, DC, and then to Boston). The challenge here is to choose among the many thousands of possible sequences of loaded and empty moves.

How much a given load contributes to the carrier's profit should be determined by taking into account not only the revenue minus direct cost of that load, but also the expected earnings of the trailer upon its arrival at the destination.

The Network Structure

Now that we know how to evaluate a given move, the problem is to determine how to optimally dispatch the trucks to maximize both customer service and total expected profits over the planning horizon. This problem can be represented by a time/space network model in which each node represents a particular region on a given day.

To handle correctly intra-regional deadhead movements, each region/day is actually represented by a set of three nodes: an inbound node, an outbound node, and a node for the stochastic cluster emanating from the region/day under consideration. The link from an inbound node to an outbound node of the same region is associated with intra-regional deadheading cost. Loaded move links lead from the outbound node of one region to the inbound node of another, while empty move links lead from the inbound node of one region to the outbound node of another. Such a representation models the fact that the empty movements go directly to a pickup point within a region, while loaded movements have to drop off the load and then deadhead to a pickup point. Notwithstanding this three node representation of each point, we assume in the remainder of this paper that each region/day is represented by a single node, for clarity of exposition.

The LOADMAP network incorporates two types of links: (1) *"deterministic" links*, which are used to represent the movement of specific known loads, possible empty moves, or holding a truck in a region (capturing the stage 1 movements); and (2) *"stochastic" links*, which are used to capture the value of an additional truck in a region (representing the movements in stages 2 and 3).

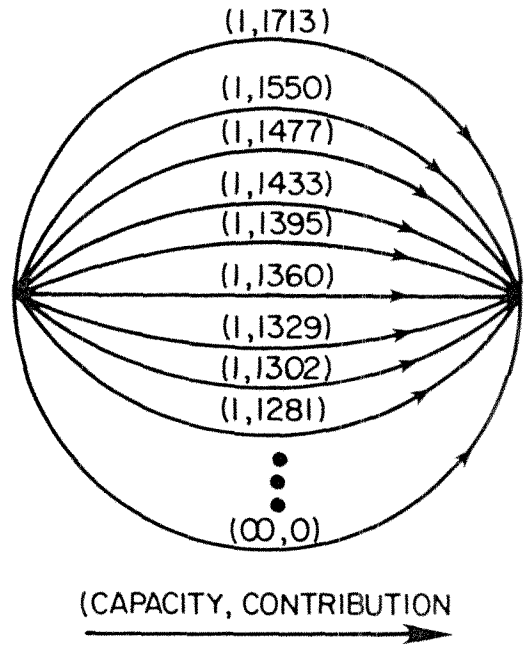

Figure 3: A cluster of stochastic links gives the value of each additional truck in a region on a given day until the end of the planning horizon. Each additional truck is worth less due to the reduced opportunities available.

The stochastic links are arranged in clusters where a cluster emanates from each node (region/day combination) in the network, in addition to all other links that may emanate from it. Figure 3 depicts one such cluster, in which each of the stochastic links, except the last one, has an upper bound of one (truck) and a link contribution which equals the value of that marginal truck. While Figure 3 depicts only 10 links (the first four of which correspond to the example in Table 1), an actual cluster might contain up to 50 stochastic links. The last link, with zero contribution and no upper bound, is added to the cluster in case the model sends a

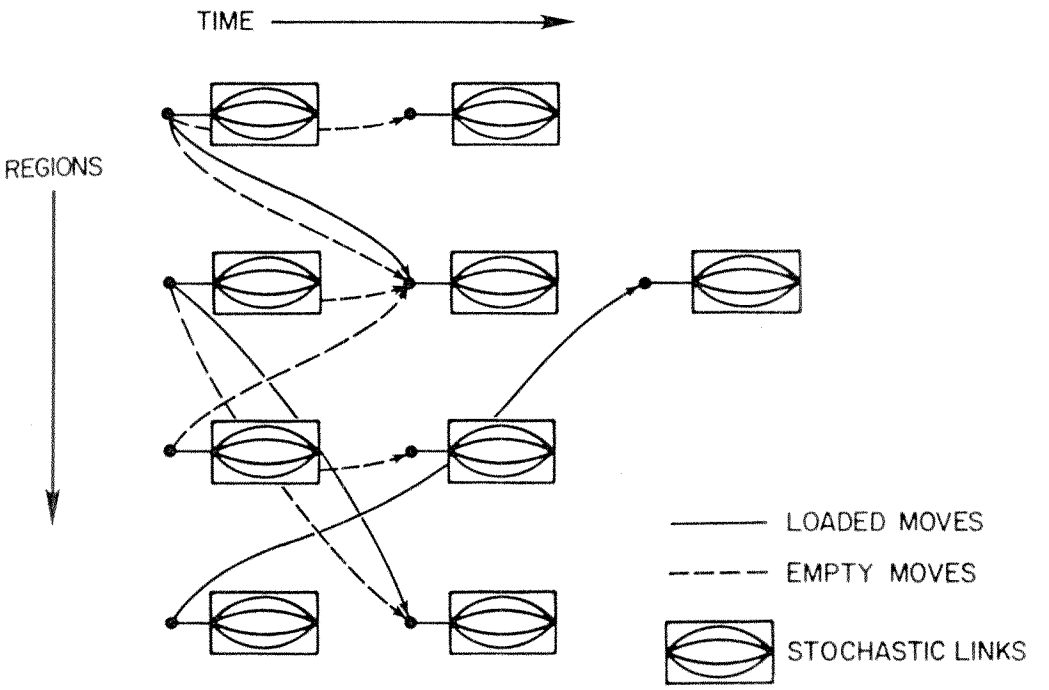

Figure 4: Loaded and empty moves out of the first time period can be represented using a time/space network, with each path ending in a cluster of stochastic links.

larger-than-expected number of trucks to that region/time.

A simplified view of the resulting time/space network is shown in Figure 4. In this network the links represent possible truck movements. Solid links represent known loads each with a known contribution, dashed links represent empty moves or holding actions (also with known contributions) and the stochastic links are drawn in boxes. Every truck enters the network on the region/day where it first becomes available. The trucks then "flow" over the network, "picking up" positive and negative contributions as they move across each link.

The network structure used by LOADMAP means that at the beginning of any given day the following options are available for each truck: (1) holding in place for loads that may be called in later in the day (the value of this option is quantified by stochastic links emanating from the current region on the first day); (2) moving loaded if a load is available; (3) moving empty to a nearby region arriving later the same day ("same day empties"); (4) moving empty over a longer distance, arriving the next day ("overnight empties"); or (5) holding in the region until the next day ("moving into" either the deterministic or stochastic links emanating from the next day's node).

Regardless of the type of movement chosen, all truck trajectories ultimately end in a movement over one of the stochastic links. This link then summarizes all future expected costs and revenues from the time it is entered through the end of the planning horizon. Note that

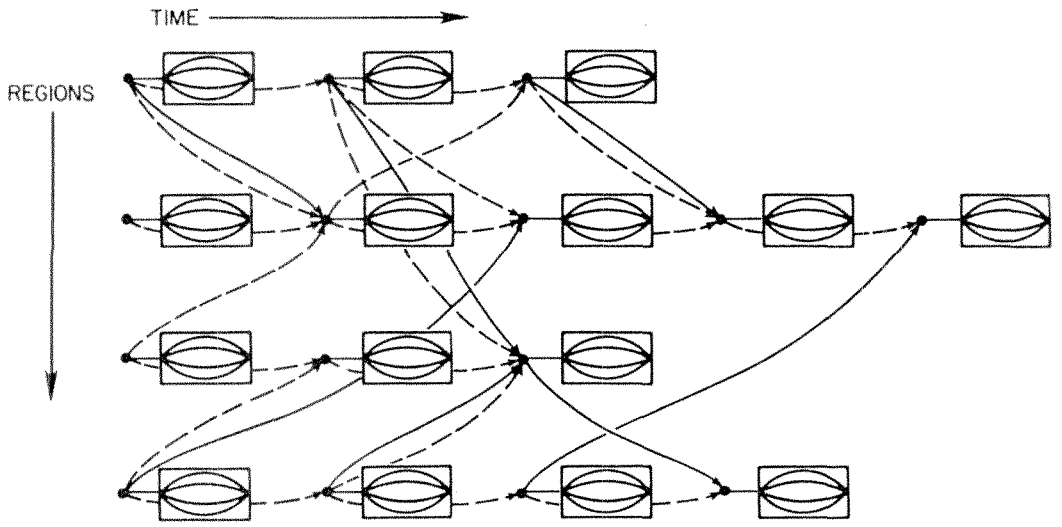

Figure 5: Multiple dispatches over known loads and empty repositioning can be represented in an expanded network. The network extends until all known loads are represented, with each truck eventually ending in a cluster of stochastic links.

the deterministic portion of the network extends as far into the future as there are known loads. When the model runs out of known loads, it simply generates one last set of stochastic links and stops. This is shown in Figure 5, which depicts known loads emanating from future time points and the corresponding extension of the network into the future. Note that the actual network structure connects the end of each stochastic link cluster to a single "supersink" node; this is not drawn in Figures 4 and 5 for clarity of presentation.

This structure gives LOADMAP the ability to build automatically on the amount of information available to the carrier. When more loads are called in earlier, the number of deterministic links grows, and a larger portion of each truck trajectory within the planning period is known with certainty. Dispatching decisions will then automatically become more accurate. This property also means that the model can be used to quantify the value of advanced booking.

Optimization of Truck Dispatching

At North American, the model is run four times a day to provide updated instructions based on current conditions. Using a tailored start procedure and an efficient adaptation of the network simplex code, LOADMAP optimizes a 10,000 link network in 2-4 CPU seconds. The total run time of the model is less than 15 CPU seconds on an AMDAHL 580 computer.

One of the crucial decisions that the model has to make when no known loads are available in a given region is whether to send the truck empty elsewhere or to hold it for loads that may come in later that day. To do this correctly, the model has to forecast what will happen in the remainder of each day, as the day is unfolding and loads are called in. The modeling logic here is based on a detailed statistical analysis of historical data, where we assumed that loads are booked

in accordance with a Poisson process at a rate that varies over the course of the day. The expected number of loads that have yet to be called in is then updated each time the model is run, thus providing an accurate estimate of the value of holding a truck versus repositioning it at another region.

This logic implies that the value of holding a truck in place at the beginning of a day is generally much higher than, say, at 3:00 PM, when 95 percent of the expected loads may have already been called in. As a result, the model is more likely to recommend an empty move at the end of the day than in the morning. When such a move is made, it is intended to position the truck for loads that are expected to be called in the next day.

The structure of the network underlying LOADMAP derives from state-of-the-art analytical and numerical considerations. Aside from mathematical elegance, these considerations offer critical practical features in two areas: the optimization that LOADMAP performs and the way it handles the uncertainty inherent in the problem.

In the first area, LOADMAP explicitly maximizes expected profits over the entire system. Every dispatch decision (loaded or empty) is automatically balanced against all other opportunities for the truck under consideration, including the possibility of holding it for a load that may be called in later. On a practical level this means that LOADMAP's recommendations may include the following:

— Refusal to take certain loads due to their overall negative impact on the system (even though the *direct* contribution of a refused load may be positive).
— Recommending long or unusual empty moves by correctly estimating the marginal contribution of such a movement.

In the second area, LOADMAP handles the future uncertainty explicitly, creating a unique stochastic network structure that distinguishes between known and forecasted information. By contrast, other models treat demand forecasts deterministically, a process that creates practical problems ranging from heuristic truncation of the planning horizon to unreasonable sensitivity of immediate decisions to forecasted data. These problems create difficulties in implementing model recommendations. (For example, a deterministic model may recommend moving loaded trucks out of a region that has no loads to be shipped.)

Such issues are never a problem with LOADMAP, which handles those and other problems in a way that is both rigorous and intuitive. Particularly useful features of the model include the following:

— The farther a forecast is in the future, the less effect it has on decisions made today.
— Loads that have already been called in have a greater impact than forecasts for future loads.
— Increasing the planning horizon (beyond the usual 15 to 20 days) has no effect on the size of the network and has only a small effect on execution times. Furthermore, for theoretical reasons, increasing the planning horizon beyond approximately 20 days has no effect on dispatch decisions made today.

The blending of actual optimization with explicit treatment of uncertainty allows LOADMAP to make unique tradeoffs in optimizing the dispatching decisions. For example, it is willing to take high risks in repositioning trucks into a region for high margin freight, whereas it will insist on a high probability of finding a load out of regions with low margins. This ability to balance probabilities with profit margins allows North American to leverage its size and mitigate the effects of randomness in demand. LOADMAP's recommendations to position trucks where customers are likely to have loads to be shipped allow the carrier to provide a higher level of service while maximizing its profits.

Estimation of Shipment Profitability

As the previous sections show, LOADMAP is a unique tool that allows NACT dispatchers to maximize profits on a real-time basis. It also is being used by Sales and Marketing to help develop sales priorities, identifying traffic lanes where additional freight will have the highest marginal contribution to total system profits. Thus, what would normally be considered a purely operational model is in fact an important tool for tactical planning and marketing. Indeed, LOADMAP's most profound impact is the framework it provides for understanding the economics of trucking.

We can illustrate the power of this framework by considering one of the most important and difficult problems in truckload planning — quantifying the profitability of a load to the system. Load profitability analysis is needed both for screening loads in real-time and for evaluating traffic lanes and shippers.

A widely used approach to calculating load profitability is to subtract from the revenue generated by the load the direct operating cost along with some portion of the empty costs incurred before or after the loaded move. This allocation of empty miles is not only arbitrary but also ignores the opportunity cost of using a truck on a given load, as well as the larger impact of accepting that load on the rest of the system.

Our alternative approach works as follows. Let

$r(i,j)$ = the revenue earned on a load going from i to j (where i and j represent nodes in the LOADMAP network, that is, each represents a given region on a given day).

$c(i,j)$ = direct operating cost of hauling a load from i to j.

$VP(j)$ = the marginal contribution of an additional truck at j.

$VM(i)$ = the marginal contribution of one less truck at i.

Network optimization experts will quickly recognize the relationship between the dual variables at nodes i and j (*node duals*) and the quantities $VP(j)$ and $VM(i)$. Unfortunately, the node duals generated by the network simplex algorithm only approximate $VP(j)$ and $VM(i)$. The reason is that due to the network flow constraints the value of the dual variable associated with adding a truck at a (region/day) node is not the same as the value of the variable associated with subtracting a truck there. The simplex node duals, however, can give either one of these values or even some intermediate

number. Thus these variables can be highly unstable, creating problems in their practical application.

To solve these problems we developed a special post optimality logic to calculate $VP(j)$ and $VM(i)$ directly in a rigorous way. For example, to consider the value of adding a truck in a region/time node, we consider opportunities to either increase flow into that node or decrease flow out of it. This requires finding a path from that node to the "supersink" node at the end of the planning horizon, along which we can increase flow (or decrease flow when going against a link's direction). Using this logic we can calculate the total system contribution (TSC) of a load going from region/time i to region/time j as follows:

$$TSC = r(i,j) - c(i,j) + VP(j) - VM(i).$$

The term $VP(j)$ incorporates the downstream effect of sending another truck into region j while $VM(i)$ balances the opportunity cost of not using the truck on some other activity out of i. If the TSC is positive, the load is considered attractive, whereas loads with a negative TSC should be avoided. Note that apparently profitable loads may have a negative TSC if there are even more profitable opportunities being passed up. At the same time a seemingly poor load can appear attractive if the best alternative is to hold the truck in a poor region or move it empty.

The TSC statistic is not only an intuitively reasonable measure of the value of a load to the system, it also rests on a solid mathematical foundation. The expression is drawn from optimization theory where it is known as a shadow price or reduced cost. The practical application of this approach is in its ability to give the planner the value (to the system) of each load over the entire planning horizon. This statistic, in conjunction with longer-term considerations of customer relations, is used by NACT to make clear accept-or-reject recommendations.

Marketing and Pricing

Aside from their use for evaluating each incoming shipment, TSC statistics can be used in marketing and market performance analysis, where they can identify markets with high profit potential, as well as markets which are not performing up to standard. After each run of LOADMAP, the TSC statistic of each load is calculated and stored in a special data set. Then, from time to time, the average TSC of all the loads moving in each lane is computed and reported to the marketing department. Similarly, the average TSC statistic for each customer and customer location is calculated and reported to marketing.

If a lane has a low average TSC relative to other lanes outbound from a region, it indicates that the prices in that lane do not adequately reflect systemwide balance conditions and other opportunities out of that region. Since the TSC reflects global balance conditions, a lane may have a high TSC even though prices are relatively low, if it is generally a backhaul lane. Conversely, a heavy head haul lane may exhibit a low TSC if the prices, while relatively high, are not high enough to offset backhaul costs and other opportunities.

In addition to its use in evaluating lane performance and in tactical marketing,

the average TSC statistic is also used to provide an input to pricing decisions. While pricing must take into account many factors, including current market conditions and competitive pressures, the TSC statistic gives the lowest price that should be charged in a-given lane on a short-term basis. Since we require that the TSC should be greater than zero, the average price in a lane, $r(i,j)$, should satisfy

$$\underline{r}(i,j) > \underline{c}(i,j) + \underline{VP}(j) - \underline{VM}(i)$$

where the underlined quantities represent moving averages of the corresponding values. This expression helps avoid underpricing in highly competitive lanes and adds a quantitative perspective that mitigates the tendency that may exist in marketing organizations to drop prices as low as required to attract business.

In using the average lane TSC values, North American has been able to develop better sales priorities, which simultaneously reflect market and operating conditions. The implications are significant. Truckload motor carriers typically exhibit the traditional tension between Operations, which is encouraged to reduce empty miles, and Marketing, which strives to maximize revenues. With a network model that maximizes profits and a method for identifying markets with the highest profit contribution, North American has a system that gives both Operations and Marketing the same objective — to maximize profits.

Implementation

We developed LOADMAP in 1985 and installed it at North American in February 1986. After several months of testing, a number of refinements to the forecasting logic, and the addition of several reports, the model went on-line early in the summer of 1986.

Impact on Management Philosophy

Some of the most significant impacts of LOADMAP are the changes in management philosophy it produces. NACT management has compiled the following list of observations regarding the model's long-term impact on the way the truckline has been managed:

Planning ahead in Operations: Prior to LOADMAP, Operations reacted to loads as they became available. LOADMAP, however, requires load forecasts and anticipation of truck movements to meet current and future demand. Operations now forecasts loads for each day and the next and receives feedback on its forecasting performance.

Operating on a national scope: Planners, who each manage a group of regions, used to give priorities to loads in their own regions. With LOADMAP, planners often reposition trucks across regional boundaries since each load and move is evaluated on the basis of its contribution to system-wide profit.

Unifying Sales and Operations objectives: Sales and Operations often have conflicting goals on which their performance is evaluated; thus Sales gauges load volume and revenue while Operations measures empty miles. LOADMAP provides the total system contribution for every move as a means to measure the performance of both departments.

Short-term pricing: The use of LOADMAP's minimum prices (based on the requirement of a positive TSC) demonstrated

how rates on incremental loads priced on a daily basis can improve margins and help balance vehicle flows.

Customer priority by operating lane: Based on historical TSC statistics, North American can now develop customer priority lists by traffic lane based on the contribution level of each shipment. Sales then reviews the list to modify priorities based on long-term customer commitments and national contracts. This provides both Sales and Operations a common goal in servicing customers and understanding the worth of the customer's entire volume to the system.

Real-time load evaluation: Prior to LOADMAP, Operations and Sales would almost always accept loads from customers, then determine how to provide a truck to service the load. Now management believes that through the use of LOADMAP's results, loads can be screened at order entry for impact on the current system. Loads can then be accepted or rejected on the basis of the customer's priority and the load's contribution at the time of order registration. This alleviates the problem of accepting a load and then not being able to perform the service.

Recognition of each region's "booking profile": The forecasting model requires the development of profiles of how loads are called in over the course of the day. This has increased the dispatchers' sensitivity to the timing of shipment bookings in each region and to the time when vehicles could be released or held to meet anticipated demand.

Bottom-Line Impacts

Two approaches were taken to develop hard numbers on the impact LOADMAP has had on North American's bottom line. The first included analysis undertaken internally by North American, comparing LOADMAP's dispatching decisions over a period of three weeks to what actually happened in the field. This analysis was limited by the fact that it ignored LOADMAP's ability to better position the fleet, thereby increasing total revenue. (During the test period trucks were not actually dispatched by LOADMAP, so there was no reliable way of estimating the effect.) Instead, attention was focused purely on LOADMAP's ability to minimize total empty miles by optimizing across the entire fleet. The results showed that if 100 percent of LOADMAP's recommendations had been followed, the loaded movement ratio (that ratio of loaded to total miles) would have improved by 3.8 percent as a result of the reduction in empty miles. This translates to potential annual savings of $4,980,000. In practice, however, not all of the recommendations can be followed (due to restrictions in the management of the owner-operator fleet) and the consensus was that only half of these savings were truly achievable, giving a conservative estimate of $2,490,000 in annual savings.

Simulating the Optimization

In reality, there will be an impact on the revenue side as LOADMAP produces better positioning for the right loads. Unfortunately, it is impossible to measure directly the amount or quality of this invisible freight in the system. Instead, we designed an experiment that tests the dispatching capabilities of the model with a degree of precision and detail that we have not seen documented elsewhere.

First, we wrote a large-scale Monte-Carlo model that encompasses LOADMAP to simulate the entire operation of a truckload carrier. The simulation imitates the process of loads being called in and the movement of drivers with a very high level of detail. This includes simulation of such activities as drivers exceeding their duty time limits and going to sleep, matching different truck types with load types, dispatching single and double teams, and accounting for actual driver compensation schedules (based on a driver's experience and the length of haul of a load). Five times during each simulated day the simulation calls LOADMAP to determine how drivers should be dispatched and, following the execution of LOADMAP's instructions, the clock is advanced to the next dispatch period.

This simulation approximates the operation of a truckline that uses LOADMAP, and thus some of the effects of the model can be learned by running it with actual data and comparing the results to actual performance. Such an experiment, however, cannot entirely capture the invisible freight since the carrier usually does not know about freight not carried due to insufficient capacity. To measure this effect we developed a truck dispatching game built around the simulation model. In this game, teams of six dispatchers compete against each other by making dispatch decisions with the objective of maximizing total contribution while servicing the customer demands. Each team member operates a dispatch computer terminal and can execute any dispatch from his or her region. As in actual operations, team members cooperate with each other by sending empties across regional boundaries and alerting teammates to developing problems. Each dispatch is recorded by the computer, and at the end of each day all loads that were not carried are recorded as "refused."

In a game in July 1986, starting with the same truck locations and using an identical set of loads, a team of Princeton students competed against a team of MIT students. In August 1986, three teams of logistics and carrier executives competed against each other during an MIT summer course; at the end of 1986, the game was run with teams of Princeton students; and in February 1987, with teams of NACT managers. All experiments used data from NACT, factored down by 60 percent (to speed the game up) and randomized enough to mask actual loads. We followed each of the above experiments (games) by a simulation/LOADMAP run using the same data. The results all showed that LOADMAP consistently produces an 8-10 percent profit increase over the teams, all of whom performed similarly. (The NACT teams performed somewhat better than the others but still within the same range.) This profit percentage amounts to about two million dollars annually for NACT.

While these results support the (independently obtained) NACT estimate, it is interesting that in all cases both costs and revenues in the simulation/LOADMAP runs were higher than those of the teams. LOADMAP tended to run more empty miles but also to take more loads. In fact, LOADMAP typically left behind significantly fewer loads than did the teams. More importantly, out of the

"must take" loads (loads tendered by important clients) LOADMAP refused only one or two versus 18 to 23 such loads refused by the teams (these numbers are for a standard 12-day game or simulation run). These results demonstrate that, consistent with NACT philosophy, LOADMAP increases customer level of service while maximizing profit.

The merit of better service cannot be measured strictly in the light of short-term profit. The main value of these results is that they demonstrate how LOADMAP can increase the carrier's ability to provide the right truck at the right place at the right time all the time. In that sense, LOADMAP is not only just a management science model that saves money; it is a management science model that makes money.

APPENDIX

Two elements of the algorithm need to be explained in greater detail: the end effects, $p(j,s)$, and the dispatch probabilities shown in Table 1.

The end effects are calculated using a simple backwards recursion. For notational simplicity, assume that out of a given region i on day s we have $n = 1, 2, \ldots, N$ options available to a truck, where an option might be to move loaded or empty to another region, or to hold in the same region until tomorrow. Now define

P = number of periods in the planning horizon;

$u_n(i,s)$ = the average number of trucks historically used for the n-th dispatch option;

$q_n(i,s)$ = the fraction of trucks out of region i, day s used for the n-th option,
$= u_n(i,s) / \sum_k u_k(i,s)$;

$t(i,j)$ = the number of time periods ($= 1, 2, \ldots$) required to move loaded or empty from region i to region j (the model uses two different travel times for these movements, but for simplicity we will assume here that they are the same);

$r_n(i,s)$ = direct contribution of the n-th option, where it is usually positive if the movement is loaded and is minus the empty movement costs if the option is an empty movement; and

$w_n(i,s)$ = the expected value of the n-th option.

The last quantity, $w_n(i,s)$, is calculated as follows:

If $s + t(i,j) < P$, then

$$w_n(i,s) = r_n(i,s) + p(j,s+t(i,j));$$

If $s + t(i,j) \geq P$, then

$$w_n(i,s) = r_n(i,s)\ [(P-s)/t(i,j)],$$

where the last equation accounts for contributions within the planning horizon if the movement ends beyond the planning horizon. The first equation puts the expected value of the n-th option as the direct contribution from that option plus the expected value of terminating at the destination, given by the end effect.

The end effects can now be calculated by starting at $s = P$ and then working backwards through time. Initially set

$$p(j,P) = 0 \text{ for all } j.$$

Now start with $s = P - 1$, then $P - 2$,

and so on, calculating at each step

$$p(j,s) = \sum_n q_n(j,s)\, w_n(j,s) \text{ for all } j.$$

An important property that can be shown easily is that as P becomes very large (in practical terms, large means 15 or 20 days), the end effects can be expressed in the following form:

$$p(j,s) = g_j + b\,(P - s)$$

where g_j is a region specific adjustment factor (which captures the differences between regions) and b is a growth factor that does not depend on the region a truck starts in. This means that if the planning horizon is sufficiently large, the end effects simply grow linearly at the same rate for all regions. As a result, for large P we can write the relative value of a truck in region i on day t versus region j on day s using

$$p(i,t) - p(j,s) = g_i - g_j + b\,(s - t),$$

which of course is independent of the length of the planning horizon. It is this important property that makes our model independent of the length of the planning horizon.

Having calculated the end effects, the next step is to find the truck dispatch probabilities, shown as the last four columns in Table 1. We will assume below that the options have been ordered so that $w_1(i,s) \geq w_2(i,s) \geq \ldots \geq w_N(i,s)$. Assume we are working out of region i on day s, and let

$d(k,n)$ = *the probability the k-th truck is* dispatched on the n-th option,

f_n = the forecasted number of trucks that will be used on the n-th option (if the n-th option represents an empty move, this is typically taken to be the historical number of empties used for this purpose),

X_n = a random variable, with mean f_n, denoting the actual number of trucks used for the n-th option,

Y_n = a random variable denoting the cumulative number of trucks used on the top n options,

$$= \sum_{i=1}^{n} X_i.$$

The probability that the k-th truck is dispatched on the n-th option is equivalent to the joint probability that Y_{n-1} is less than k (if it were greater than or equal to k, then we would have dispatched the k-th truck on one of the first n-1 options) and that Y_n is greater than or equal to k (if this were not true, we would be dispatching the k-th truck on option $n+1$ or greater). Thus

$$d(k,n) = \text{Prob}\,[\, Y_{n-1} < k \text{ and } Y_n \geq k\,]$$

which after some manipulations becomes

$$d(k,n) = \text{Prob}\,[\, Y_{n-1} < k\,] - \text{Prob}\,[\, Y_n < k\,].$$

We assume that the random variables X_n are distributed according to a Poisson distribution with mean f_n, and hence the variables Y_n also have Poisson distributions. Thus the dispatch probabilities are simply differences between two Poisson distributions.

References

Jordan, W. C. and Turnquist, M. A. 1983, "A

stochastic dynamic model for railroad car distribution," *Transportation Science*, Vol. 17, No. 2, pp. 123-145.

Ouimet, G. P. 1972, "Empty freight car distribution," *MS thesis*, Queens University, Kingston, Ontario.

Powell, W. B. 1987, "An operational planning model for the dynamic vehicle allocation problem with uncertain demands," *Transportation Research*, Vol. 213, No. 3, pp. 217-232.

Powell, W. B.; Sheffi, Y.; and Thiriez, S. 1984, "The dynamic vehicle allocation problem with uncertain demands," *Proceedings of the Ninth International Symposium on Transportation and Traffic Theory*, Dormuller and Hamersalg, eds., VNU Science Press, The Netherlands, pp. 357-374.

White, W. W. 1972, "Dynamic transshipment network: An algorithm and its application to the distribution of empty containers," *Networks*, Vol. 2, No. 3, pp. 211-236.

Keith J. Margelowsky, Vice-President, Administration, North American Commercial Transport Division, A Division of North American Van Lines, Inc., PO Box 988, Fort Wayne, Indiana 46801-0988, writes "The impact of this system on our operation is divided into these areas:

Changes in Management Philosophy

More proactive management in daily truckload operations.

Operation on a national scope rather than regional.

Common goals between Sales and Operations Departments.

Pricing decisions which consider daily incremental impacts.

Major Uses of LOADMAP

Recommended repositioning of vehicles.

Load prioritization.

Evaluation of potential loads.

The impact of LOADMAP can be measured by the improvement to divisional contribution margin. We have performed a number of evaluations, and our conservative estimate is a $2.9 million improvement."