ORIGIN TO DESTINATION UNRELIABILITY IN RAIL FREIGHT TRANSPORTATION

by

CARL DOUGLAS MARTLAND
S.B., M.I.T.
(1968)

Submitted in partial fulfillment
of the requirements for the degrees of
Civil Engineer and Master of Science

at the
Massachusetts Institute of Technology

Signature of Author... Signature redacted ....
Department of Civil Engineering, (May 12, 1972)

Certified by... Signature redacted ....
Thesis Supervisor

Accepted by... Signature redacted ....
Chairman, Departmental Committee on Graduate Students
of the Department of Civil Engineering

Archives
MASS. INST. TECH.
JUL 12 1972
LIBRARIES

ABSTRACT

ORIGIN TO DESTINATION UNRELIABILITY IN RAIL FREIGHT TRANSPORTATION

by

CARL DOUGLAS MARTLAND

Submitted to the Department of Civil Engineering on May 12, 1972 in partial fulfillment of the requirements for the degree of Master of Science.

Improving freight service is one way for railroads to attain a better competitive position and thereby increase their profitability. By investigating logistics systems, this thesis identifies those characteristics of trip time distributions which are most important to shippers. Performance measures are then selected for each of these.

Data from three railroads reveals a wide range in performance provided various city pairs. Regression analyses show that the variability in mean trip times, but not that in reliability, can be explained by general trip characteristics such as mileage and the number of intermediate yards. Reliability is caused by the specific yard performance characteristics which vary greatly from yard to yard.

There is no essential tradeoff between mean trip times and reliability. Improved reliability will reduce mean trip times and could ultimately lead to benefits to shippers and railroads of well over $500 million per year.

Thesis Supervisor: Joseph Sussman

Title: Associate Professor of Civil Engineering

ACKNOWLEDGMENTS

I would like to thank the many people who helped in the creation of this thesis:

- Ken, Joe, Andy, Brian, John, Bob, and Mike for their continued interest in discussing the ideas expressed in the thesis as well as a myriad of other subjects.

- Professor A.S. Lang for finding me a "home" at M.I.T. and for continued guidance as my research.

- Professor J.M. Sussman for refusing to accept everything I wrote and for attempting to make this document readable.

- Irene Steiner for an incredibly fast and accurate typing job.

- Mr. Henry Roy Fitzgerald for supporting me for the last four years.

- My wife Nancy for keeping me going when skies were grey.

PROFESSIONAL ACKNOWLEDGEMENTS

This research was supported in part by Federal Railroad Administration Contract #DOT-FR-10006. Special thanks are due project monitors Mr. John Williams and Mr. John West of the FRA for providing valuable assistance with respect to the overall goals and strategy of this research.

A study of railroad reliability is impossible without interaction with railroad officials and access to railroad data. The following railroads have been instrumental in keeping this research relevant:

Penn Central

Boston and Maine

Southern

Southern Pacific

Frisco

C&O/B&O

Burlington Northern

Rock Island

CONTENTS

CONTENTS (Cont'd.)

TABLES

FIGURES

FIGURES (Cont'd.)

CHAPTER 1

INTRODUCTION

A. Background

In 1970 the rate-of-return for Class I railroads dropped below 2% for the first time since the depression[1]. Although this rate increased to 2.5% in 1971[2], the long-term decline in railroad profitability has by no means been reversed. Since 1940, total intercity freight ton-miles have more than tripled while rail ton-miles have barely doubled. During the same period, the rail market share in terms of ton-miles has declined from 62% to 40%, the trucking market share has increased from 10% to 21%, and the oil pipeline share has increased from 10% to 22%[3]. Between 1959 and 1969, railroads' share of intercity freight revenues fell from 33% to 25% while that of trucks rose to 70%[4]. These figures highlight the extensive impact

---

[1]AAR, Yearbook of Railroad Facts 1971, p. 24. Rate of return represents the ratio of net railway operating income (net income prior to paying fixed charges such as interest on debt and rents for leased lines or equipment) to total net investment in transportation property, as recorded under the accounting regulations of the ICC.

[2]Railway Age, 31 January 1972, p. 68.

[3]Transportation Association of America, Transportation Facts and Trends, April 1971, p. 8.

[4]Idem

of the trucking and oil pipeline industries on railroad operations over the past 30 years.

This impact has been most severe for those railroads which operate in the densely populated northeastern United States (Table I). Between 1951 and 1969, gross revenues and revenue ton-miles actually declined for Eastern District railroads; the percentage of revenues carried through to net operating income before federal income taxes declined precipitously. By the end of 1970, the Penn Central, the nation's largest railroad with gross revenues over two billion dollars, joined the Boston and Maine and the Reading Rail-roads in the growing group of bankrupt eastern railroads.

Railroads are generally conceded to be the most efficient mode for transporting heavy shipments long distances overland. Although terminal costs are high, marginal over-the-road costs are very low due to low labor and power requirements. However, trucks, because their terminal costs are lower, can move shipments short distances more economically than railroads even with much higher over-the-road costs. Hence neither their inroad into the short haul rail market nor the subsequent increase in the average rail haul from 350 to 500 miles can be regarded as a misallocation of transportation resources. In particular, the fact that the railroads in the northeast are the least profitable results in part simply because a great number of large cities are too close to one another to generate a great deal of the long haul shipments which are most profitable to railroads.

TABLE 1.1

PERCENTAGE CHANGE IN VARIOUS RAILROAD STATISTICS BETWEEN 1951 AND 1969, FOR ALL CLASS I RAILROADS AND BY DISTRICT

	US	East	South	West
Revenue ton-miles[1] (billions)	19%	-6%	53%	32%
Revenues[2]	11%	-6%	30%	24%
Revenues carried through go net operating income before federal taxes[2]	-55%	-74%	-33%	-48%
Rate of return[3]	-37%	-67%	-12%	-26%

---

[1]Railway Facts, p. 35.

[2]Moody's Transportation Manual, p. 43.

[3]Railway Facts, p. 24.

In a sense, the plight of the northeast railroads highlights the basic issue facing the industry. Since railroads are no longer the only or the best means of intercity transportation, they must re-organize their structure and their operations so that they can profitably exercise their competitive advantage in heavy, long-haul shipments. It is definitely in the nation's interest to insure that railroads can provide this service. In the past, numerous alternatives for aiding this re-organization process have been suggested or implemented. Although it is beyond the scope of this thesis to evaluate each of these, demonstrating the range of available possibilities provides an important conceptual framework for this study:

1. Improve Rail Service to Attract New and/or Retain Old Customers -- Unit trains, piggy-back trains, and high priority trains that by-pass intermediate yards or even interchanges are all responses to customers demands for a faster, more reliable service. Specialized cars have been developed to meet the unique requirements for shipping commodities such as auto parts, automobiles, perishables, and grain.

2. Reduce Operating Costs -- Despite problems with union work rules, railroads reduced the number of employees from 1,557,000 in 1947 to 885,000 in 1960 to 662,000 in 1968[5]. Although total opera-

---

[5]Ibid, p. 22.

ting costs rose from 7.6 to 9.1 billion dollars over the same period, this 20% increase is remarkably less than the 80% increase in the index of charge-out prices and the 208% increase in the index of wage rates.

3. Raise Rates -- Railroad rates were originally regulated by the Inter-State Commerce Commission to limit destructive inter-railroad competition for traffic and to prevent excessive costs and monopolistic profits at the expense of transport sensitive industries. Low rates, often not even covering short run marginal costs, were set for low-value, nationally important, bulk commodities such as agricultural products and coal. Profitability was assured by setting excessive rates on high value products whose transport costs had a minor effect on final price and demand[6]. With the rise of the trucking industry, railroads lost their monopoly on high value freight transportation, but pressure from truckers and the need to generate revenue has kept many rail rates at uncompetitive levels[7]. At the same time, competition with water and pipeline modes has kept rail rates low on bulk commodities. It is certainly possible that higher total revenues could be received by lowering the rates on high value commodities and raising those on low value commodities, especial-

---

[6]Nelson, Railroad Transportation and Public Policy, pp. 328-330.

[7]Ibid, pp. 65-66, 111. For a general background into the history of ICC regulation, see Freidlander, Dilemma of Freight Transportation; Nelson; or Pegrum, Transportation and Public Policy.

ly since recent rate increases merely returned the average rate per ton-mile to Korean War levels.

> With competition surrounding them on all sides, the railroads can re-adapt themselves to profitable operations only to the degree that they learn their costs in detail and price their product accordingly (with due account taken of competitors costs).[8]

4. Eliminate Unprofitable Traffic -- As a minimum pricing policy, rates could be raised on traffic which does not cover out-of-pocket costs. Nationalization of passenger service through the creation of AMTRAK was a giant step in this direction so far as the railroads are concerned (although the deficit has merely been shifted to the government.)

5. Abandon Unprofitable Branch Lines and Facilities -- Since 1947, abandonment of over 20,000 miles of branch lines has led to a 9% reduction in system mileage. However, a recent DOT study has indicated that nearly 80,000 additional miles may be suitable candidates for future abandonments[9]. Political and local economic pressures make branch lines notoriously difficult to abandon, but bankruptcy may make the process easier for some railroads.

6. Change Financial and/Or Organizational Structure -- Federal loans to bankrupt railroads, investment tax credits for

---

[8]Lyne, in Introduction to Poole, Costs - A Tool for Railroad Management.

[9]DOT, National Transportation Study, 1972

railroads and diversification by means of railroad holding companies are examples of actual changes. The proposed Surface Transportation Act, ASTRO, and recent DOT bills offer additional alternatives.

7. Undertake Substantial Investments in Improved Equipment and/or Facilities -- Electronic hump yards, Central Traffic Control, continuously welded track, and widespread use of electronic data processing are just some of the major projects undertaken by the industry to improve its operations. Future possibilities include nationwide installation of Automatic Car Identification systems in conjunction with better computer facilities, consolidation of the numerous freight yards that exist in some cities, and development of more durable rolling stock.

B. The Modal Choice Decision

The extent to which these possibilities have been implemented suggests that the industry has been neither complacent nor unimaginative. However, the fact remains that railroads have been losing the competition for inter-city freight and freight revenues. Central to this competition is the shipper's modal choice decision which can be modelled in terms of four characteristics[10]:

[10] Landow, "The Measurement of Service,"; Sillcox, "Stand-by and Second-Best"; and Kolsen, The Economics and Control of Road-Rail Competition, p.67. Direct expenses include loading and unloading charges and the costs associated with shipment size as well as the rate. Other factors may at times be very important including in transit privileges, availability of special equipment, and reliability in placing empties.

- direct expense
- trip time
- trip time reliability
- loss and damage

According to this model, traffic moves by the mode which minimizes the sum of the direct expenses and the indirect expenses caused by the specified service characteristics. An available mode with higher rates will be utilized whenever the rate differential is more than offset by the perceived advantages of the service differential. For example, since railroads are not currently competing with trucking in terms of transit times or reliability, rail rates must be lower than truck rates to attract traffic. The more important service is to the customer and the greater the service or rate differential, the less likely that he will use rail transportation. Thus the inflated rates on high value commodities noted above divert profitable traffic to trucks. Similarly, an investment policy and a rate structure biassed toward high tonnage shipments[11 and 12]

---

[11]Since 1940, railroad investment policy has increased the average capacity of all freight cars from 50 to 65 tons and the average capacity of new cars from less than 60 to 80 tons (Railroad Facts 1971, p.66). Due to the low marginal cost of hauling additional tonnage, pricing policy places the lowest rates per hundredweight on carload or multi-carload shipments. As a result of these policies, the average rail shipment increased from 28 to 45 tons over the last 30 years or to nearly double the average capacity of a truck. (Moody's Transportation Manual, p.22)

divert a great deal of the low tonnage traffic which accounts for the bulk of intercity freight revenues[13]. On the other hand, improved rail service at the same rates or lower rates at the same level of service would help to retain current traffic while attracting new and future traffic from other modes. Although cost reductions are always welcome, it is unlikely that they alone will lead to profit maximization:

> "It may be impossible to meet price-quality competition by changes only in price policy. A better understanding of this aspect would help rail in its fight to retain freight service."[14]

In any case, the precarious state of the rail industry necessitates exploring the possibility of improving railroad service and any other actions that might aid rail profitability.

---

[12]A 15 ton shipment may move under a truck-load rate by truck rather than under an expensive less-than-carload (LCL) rail rate. Due to the unprofitability of LCL traffic, railroads may not be unhappy to lose this traffic. However, it may also happen that a 30 ton shipment moves as two truckloads rather than as a single carload or two piggy-back shipments, either of which might be very profitable.

[13]Kullman estimates that only $4 billion of the total $32 billion inter-city truck revenues are received from truckload shipments moving more than 25 miles. Kullman,

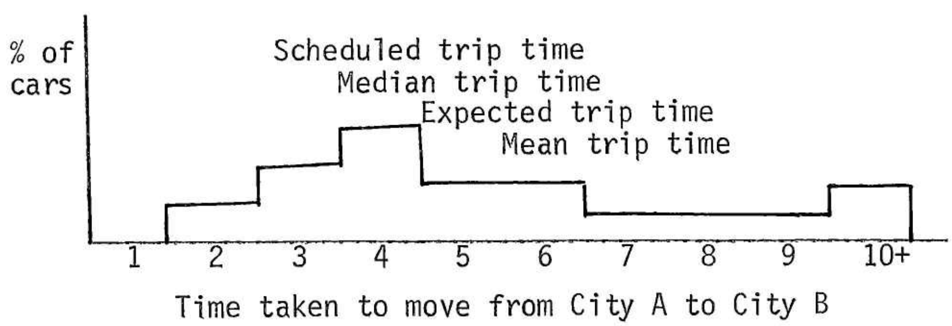
[14]Kolsen, p. 74.

C. Focus of the Thesis

This thesis is prompted by the magnitude of the impacts of rail service characteristics on the industry's competitive position. Although rates, availability, loss and damage, shipment size, and in-transit privileges are important aspects of transportation service, this study will focus on the specific characteristics associated with the trip time distribution.

Both transit time and unreliability have been cited as important determinants of modal split decisions. However, the terms are seldom defined and are often confused. Transit time has been or can be used to mean average trip time, median trip time, expected trip time, or scheduled trip time -- all of which may be different (Figure 1.1).

FIGURE 1.1



Even more confusion exists with reliability. Shippers often seem to equate it with "on time" delivery, railroads associate it with the number of movements which meet acceptable standards, while analysts define it as a characteristic (usually the standard deviation) of the trip time distribution.

In this thesis, unreliability is defined as variability in the trip time distribution which arises because the length of time necessary for a car to travel between a particular origin and destination pair is unpredictable. Such unreliability definitely exists and both shippers and railroad officials act in a manner to offset its impact. Railroads are reluctant to publish actual freight schedules to which they might legally be held or to advertise a level of service without including a safety factor for delays beyond scheduled time. Shippers likewise put more faith in their experience with railroads than in promises railroad officials might make concerning expected trip times. Hence railroad standards and shippers ordering and inventory policies make allowance for some variability in trip times. Much of the confusion associated with unreliability occurs because the nature of unreliability is not clearly distinguished from its effects. For example, "on time" performance can always be improved by adding a day or two to the definition of on time (a shipper might do this after stocking additional inventory). Reliability is never improved without changing the actual trip time distribution.

Defining unreliability as variability in the trip time distribution does not indicate how unreliability should be measured. Great care is necessary in selecting performance measures because they are the means that will be used to relate actual performance to desired goals. One such goal might be to achieve a better competitive position with respect to other modes of transportation by improving rail service characteristics. This is a qualitative goal that is not operationally effective without specific quantitative performance measure that relate current to past performance. Only with such measure will railroad decision-makers be able to evaluate past decisions as a guide to future decisions. If these measures do not reflect those aspects of rail service which are in fact relevant to shippers' modal choice decisions, then operating decisions that cause improvements in measured service may not in fact improve the railroad's competitive position. Thus, choice of performance measures must be preceded by an analysis of shippers' needs.

Care is also needed to insure that the performance measures will be appropriate with respect to problems in measuring trip times and to the nature of railroad trip time distributions. For instance, performance measures should not be overly sensitive to common types of data errors. Also, measures which have very useful properties when applied to specific types of distributions (z.b. normal or exponential) may be quite useless when applied to railroad trip time distributions.

This thesis will examine logistics systems to determine the impact rail unreliability has on shippers. After identifying the characteristics of the trip time distribution which seem most important, a number of potential measures will be evaluated before choosing specific measures to reflect these characteristics. These measures will then be used in analyzing car movement data from several railroads in an attempt to determine the causes of railroad unreliability.

D. Content of the Thesis

Chapter 2 outlines railroad operations in detail sufficient for this study. Components of the origin-to-destination (O-D) trip and of the railroad network are defined; schedules, priorities, and policies are considered as important determinant of railroad service.

In Chapter 3 the effects of unreliability are studied as a prerequisite to selecting performance measures. Impacts on shippers logistics systems are considered in some depth.

After discussing a set of evaluation criteria, the strengths and weaknesses of a number of performance measures are analyzed in Chapter 4. Separate measures are chosen for the average trip time, variability in trip time, and the extent of very long delays.

Chapter 5 presents an analytical framework for studying the causes of unreliability and offers a number of hypotheses which are tested using data from Railroads A, B, and C. Mileage, intermediate yards, and traffic volume are tested as possible determinants of

O-D reliability.

Finally Chapter 6 takes up the problem of improving reliability. The way this is done depends to a great extent on the type of improvement sought and on the specific characteristics of the railroad network involved; therefore emphasis is placed on the overall impacts rather than on evaluation of alternatives.

E. Summary

As a result of increased competition from other modes of transportation, the railroads' share of intercity freight ton-miles has fallen from 62 to 40% over the last 30 years. The impacts on freight revenues and profitability have been severe, particularly for the railroads in the northeast. Improving service characteristics is one of a number of ways that railroads can attempt to reverse their declining profitability. The industry would benefit from improved service by retaining traffic that would otherwise have been diverted to other modes, by attracting traffic currently carried by other modes, and by obtaining a greater percentage of future increases in total inter-city revenues. This thesis will focus on the nature and causes of those service characteristics associated with the origin-to-destination trip time distribution which are important to railroad customers.

Reliability is defined as variability in the trip time distribution but this does not indicate how unreliability should be measured. Specific measures can be selected only after consideration of the nature of actual trip time distributions, problems in measuring trip times, and the long and short run impacts of unreliability on railroads and on shippers. When an appropriate set of service performance measures are chosen, actual data can be analyzed to determine the current level of service and the economic benefits of improvements in that service.

CHAPTER 2

A BRIEF INTRODUCTION TO RAILROAD OPERATIONS

Unfortunately, the complexity of rail operations and the characteristics of the available data sets cause some confusion in precisely defining the terms that will be used in this thesis. In the first part of this chapter, the time periods between important events associated with the origin-to-destination trip of loaded freight cars are defined as they will be used in later chapters. It is hoped that this section will serve as a useful reference whenever the meaning of any particular term is not clear from the context in which it is used.

The second part of this chapter deals with the work that is done in railroad yards. Before a train can be run, a number of freight cars must be assembled in the desired order through a process called classification. The many reasons that this yard process can be highly unreliable are discussed in order to provide some background for the data analyses described in Chapter 5.

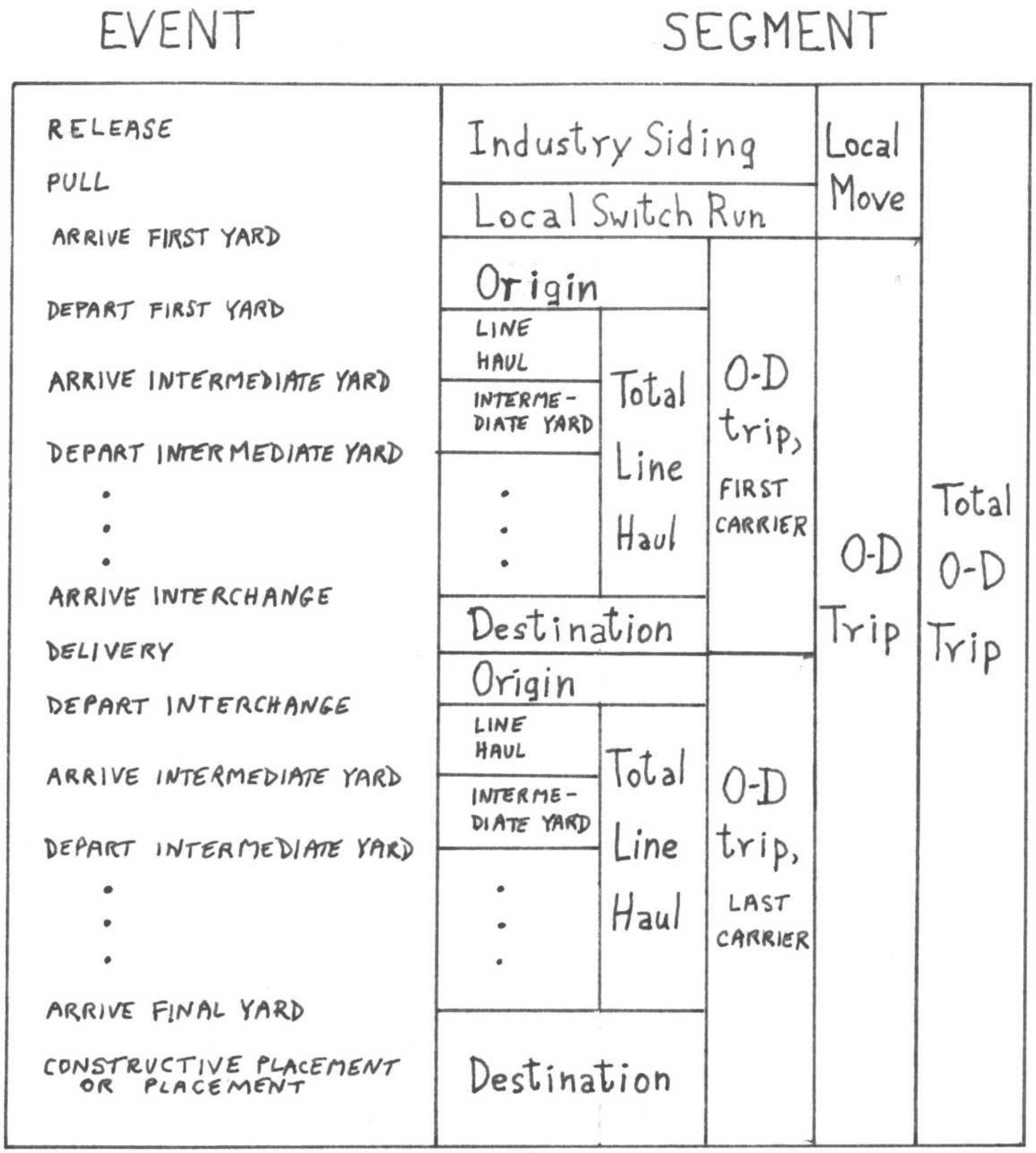
Since shippers are more concerned with performance than with railroads, they may find this chapter of little interest. However, it is essential that rail service be discussed in the context of rail operations if there is to be any hope of improving that service.

B. Definitions of Trip Segments

Railroad responsibility for a shipment begins when the shipper (consignor) notifies the railroad that it is ready to be picked up (pulled). Some time later, a local switch engine pulls the car and brings it to a railroad yard where it is assembled into a group or block of cars having the same immediate destination. The car then proceeds to its destination in a series of line haul trips between classification yards and interchanges. When it arrives, the railroad contacts the firm which is to receive the shipment (the consignee) for delivery instructions. The consignee may ask that the car be delivered immediately or that it be held in the yard at his expense until he is ready to receive it. In the first case, rail responsibility ends when the car is placed at the consignee's siding (placement); in the second, when the consignee assumes responsibility for it (constructive placement).

The time from release to placement as described above determines what is often called dock-to-dock performance or what is shown in Figure 2.1 as the Total O-D Trip. For several technical and theoretical reasons this is not always the most useful

FIGURE 2.1


EVENT
SEGMENT
RELEASE
PULL
ARRIVE FIRST YARD
DEPART FIRST YARD
ARRIVE INTERMEDIATE YARD
DEPART INTERMEDIATE YARD
ARRIVE INTERCHANGE
DELIVERY
DEPART INTERCHANGE
ARRIVE INTERMEDIATE YARD
DEPART INTERMEDIATE YARD
ARRIVE FINAL YARD
CONSTRUCTIVE PLACEMENT OR PLACEMENT
Industry Siding
Local Switch Run
Local Move
Origin
LINE HAUL
INTERME-DIATE YARD
Total Line Haul
O-D trip, FIRST CARRIER
Destination
Origin
LINE HAUL
INTERME-DIATE YARD
Total Line Haul
O-D trip, LAST CARRIER
Destination
O-D Trip
Total O-D Trip

trip segment to analyze even though it is the trip time perceived by the shipper. As indicated in the diagram, the Total O-D Trip is composed of a Local Move and an O-D Trip. The Local Move, the time from release until arrival at the first yard, is the responsibility of the local yard whereas the O-D Trip is subject to network operating procedures and conditions. Therefore it is reasonable to study the nature and causes of unreliability in the Local Move independently of the unreliability in the O-D Trip. By distinguishing between these two trip segments, moves from all shippers served by each local yard can be aggregated to provide a larger data base for studying the O-D Trip.

In addition to this theoretical reason for not analyzing the Total O-D Trip is the practical fact that most railroad data systems -- including those used in this thesis -- do not retain the release time, but instead record the arrival time at the first yard or the pull time as the beginning of a trip. A second technical difficulty is that most shipments involve more than one railroad, each of which records data on the shipment only for the time that the individual road is responsible for it. Thus, there are many instances where data for either the Total O-D Trip or the O-D Trip are not available.

These data limitations make it necessary to focus on the portion of the O-D Trip involving a single railroad. In practice this means that receiving a car from another railroad must be considered a valid opening move and that delivering a car to another road must be a valid closing move. Of necessity, the data analysis in this thesis thus involves what will be called the O-D trip (not capitalized), that part of the O-D Trip (capitalized) which involves the railroad being studied.

Figure 2.1 also defines the individual trip segments. Origin Time and Destination Time refer to the total time spent by a car at each terminal. Intermediate Yard Time refers to the time spent at a single intermediate yard; in the general case, an intermediate yard could be an interchange, but the data limitations discussed above mean that an interchange will always be either an origin or a destination of an O-D trip in the analysis undertaken in this thesis. The final trip segment identified is the Total Line Haul which includes all over-the-road operations as well as all intermediate yards.

The large number of trip segments and the limitations of railroad data systems necessarily cause some confusion. Again, it is hoped that the meaning of terms used will always be clear either from the context or by reference to this section.

C. The Classification Process

At a typical railroad yard, inbound trains are inspected in an area called the receiving yard prior to classification, a process which sorts cars onto tracks associated with particular destinations. The block of cars on each track is often carried by only one specific outbound train per day. Therefore, if the car for some reason does not depart on this train, it will be delayed a full 24 hours before it has another chance to depart. A short time before the scheduled departure time of an outbound train, the appropriate blocks are coupled and transferred to one of a series of departure tracks where the cars are again inspected. When a locomotive, caboose, and crew have been supplied, the train departs, leaving behind any cars that had not yet been classified or that had been found in need of repair. The process is the same if the block is carried by a train which stops at rather than originates from the yard in question.

Considerable unreliability is associated with each step of this process. Inbound train delays, yard congestion, classification delays, scheduling and priority policy, and resource constraints will be considered below as reasons that a car might not make the earliest connection.

1. Inbound Train Delays: If an inbound train is delayed for any reason, there may be insufficient time available for some cars on that train to make their scheduled connections. Belovarac reports that variances in arrival times of trains at their destinations on Railroad D range from 1 to 5 hours due to late departures,

variations in running time, intermediate yard stops, and various delays en route[1]. Such erratic arrival patterns were found by Reid to cause as many as 10% of all cars moving through one yard on Railroad B to miss their connections[2].

2. Yard Congestion: If an unusually high number of cars are in the yard at one time, long queues may develop in the inspection, classification, and assembly processes. This in effect increases the processing time for each car and could delay many beyond their scheduled departure.

3. Classification Delays: Accidents and other unexpected events can slow the classification process and create yard congestion effects.

4. Scheduling Policy: Reid and O'Doherty have shown that cars with greater scheduled yard time suffer fewer missed connections[3].

5. Priority Policy: Assigning high priority to certain trains, cars or blocks may cause other cars to miss connections or to be delayed excessive amounts of time. High priority can ameliorate the effect of late arrivals or assure that a block is not dropped due to tonnage constraints. Low priority blocks

---

[1]Belovarac, Determinants of Unreliability in Line Haul Operations.

[2]Reid, Yard Unreliability in Rail Freight Movement, p. 16.

[3]Reid; O'Doherty, Classification Yard Effects in Rail Freight Movement Reliability.

may be the first to be excluded due to tonnage limits and may in fact be "bumped" in favor of higher priority blocks. In the latter case, the term missed connection does not apply since a block may arrive at a yard on train XN (day 1), be bumped and depart on train XN 9day 2).

6. Maximum and Minimum Train Length Policies: Train speed considerations or physical power constraints force a limitation on the tonnage carried by trains. If this maximum is close to the average tonnage available, variability in the available tonnage makes it likely that the train length constraint will often delay cars. On the other hand, railroads often set minimum length or tonnage restrictions on trains in order to maximize operating efficiency as measured by crew costs or ton-miles per train mile. If these restrictions are not met, the train will be cancelled causing delays to cars over a period of several days due to the phenomenon of persistent delay. In Figure 2.2, a hypothetical situation is depicted in which 60 to 120 cars originate daily for the same outbound train. Given a tonnage restriction which limits the train to 120 cars, each day's traffic could be handled without delay if that were the only restriction. However, the yard has a policy not to run trains less than 70 cars in length; hence the train on day 1 which would carry only 60 cars is cancelled. The next day 100 additional cars originate to make a total of 160 cars available for the train which means that 40 must be delayed.

EFFECTS OF TRAIN CANCELLATIONS

Day	Cars Arrived	Delayed From Day Before	Total Available	Train Length	Cars Delayed
1	60		60	Cancelled	60
2	100	60	160	120	40
3	90	40	130	120	10
4	120	10	130	120	10
5	110	10	120	120	
6	95		95	95	
7	110		110	No Power	110
8	85	110	195	120	75
9	105	75	180	120	60
10	65	60	125	120	5
11	100	5	105	105	
Totals	1070	370			

Average Number of Cars: 97

Average Train Length: 115

Percentage Cars Delayed: 35%

Due to no Train: 15%

Due to no Capacity: 20%

Figure 2.2: Capacity constraints mean that delays are persistent; train cancellations cause departure queues that may last several days.

On days 3 and 4, 20 cars are delayed for the same reason. This is not the only effect of the cancellation; the fact that the train was not run on day 1 could cause a local power shortage at some other yard which in turn could cause another cancellation. In this example, when the train is cancelled on day 7 for just that reason, extensive delays result.

Consider what the policy used in this example has accomplished. The average train length was increased from 97 to 115 for an 18% improvement in car-miles/train-miles. In addition, crew costs were also cut 18% by running fewer trains. However, the policy had severe consequences for service -- fully one third of all cars carried on the train were delayed an extra day causing an increase in both the mean and variability of yard times.

7. Hold/No Hold Policy: Frequent inbound train delays or classification delays create situations in which fewer cars will miss connections if the outbound departure is delayed. At times, the outbound must be delayed because insufficient cars, crews, or power units are available due to inbound delays. More often there is a real choice between delaying the train a few hours and delaying some cars at least a day. Folk has shown that the specific policy used is an important determinant of both yard and O-D performance[4].

---

[4] Folk, Models for Investigating the Unreliability of Rail Freight Shipments.

As shown by the above discussion of the causes of variability in yard times, unreliability can result as much from seemingly rational railroad policies as from inbound train delays and other unforeseen events. Excessive concentration on cost reduction in an industry confronted with highly variable local demands (i.e., the number of cars to be moved at the individual yard) may have led to an unnecessarily low level of service.

D. Summary

This chapter defines the terms that will be used throughout the rest of the thesis. Figure 2.1 summarizes the inter-relationships between the various trip segments that are of interest. Local Moves are not included in most railroad data systems and in any case would be analyzed separately as an important indpendent component of the Total O-D Trip. Since most data systems are also limited to moves over a single railroad, the O-D trip used for the analyses in this thesis starts with either arrival at the first yard or receipt from interchange and ends with either placement or constructive placement.

Railroad cars proceed to their destination in a series of line haul trips between classification yards and interchanges. Congestion, line haul unreliability, operating policies, and other factors often cause delays to cars in these yards. Since a car typically makes a connection with a particular outbound train that departs only once a day from a classification yard, it will be delayed 24 hours if it misses a connection for any of these reasons.

CHAPTER 3

THE EFFECTS OF UNRELIABILITY ON SHIPPERS

A. A Logistics Framework[1]

1. Introduction

A shipper's logistics system is that business activity which supports the production and marketing functions by assuring that materials and products are available when and where they are needed. The quality of transportation services can affect the costs associated with many of the other components of that system. These components and the relevant inventory policies are discussed below in order to exhibit the extensive framework within which transportation decisions are made. Notice that the mean and variability of transit times directly affect the "lead time" which is a major determinant of inventory costs. This relationship will be explored in the remainder of this chapter in order to identify those characteristics of the trip time distribution which are most important to shippers. In Chapter 4, performance

---

[1]The following section is largely based on material from Heskitt, Ivie, and Glaskowsky, Business Logistics, and Shycon, "Designing the Distribution System."

measures will be selected which best reflect these characteristics.

2. Logistics Activities and Policies

The purpose of this section is to define the various aspects of the logistics framework.

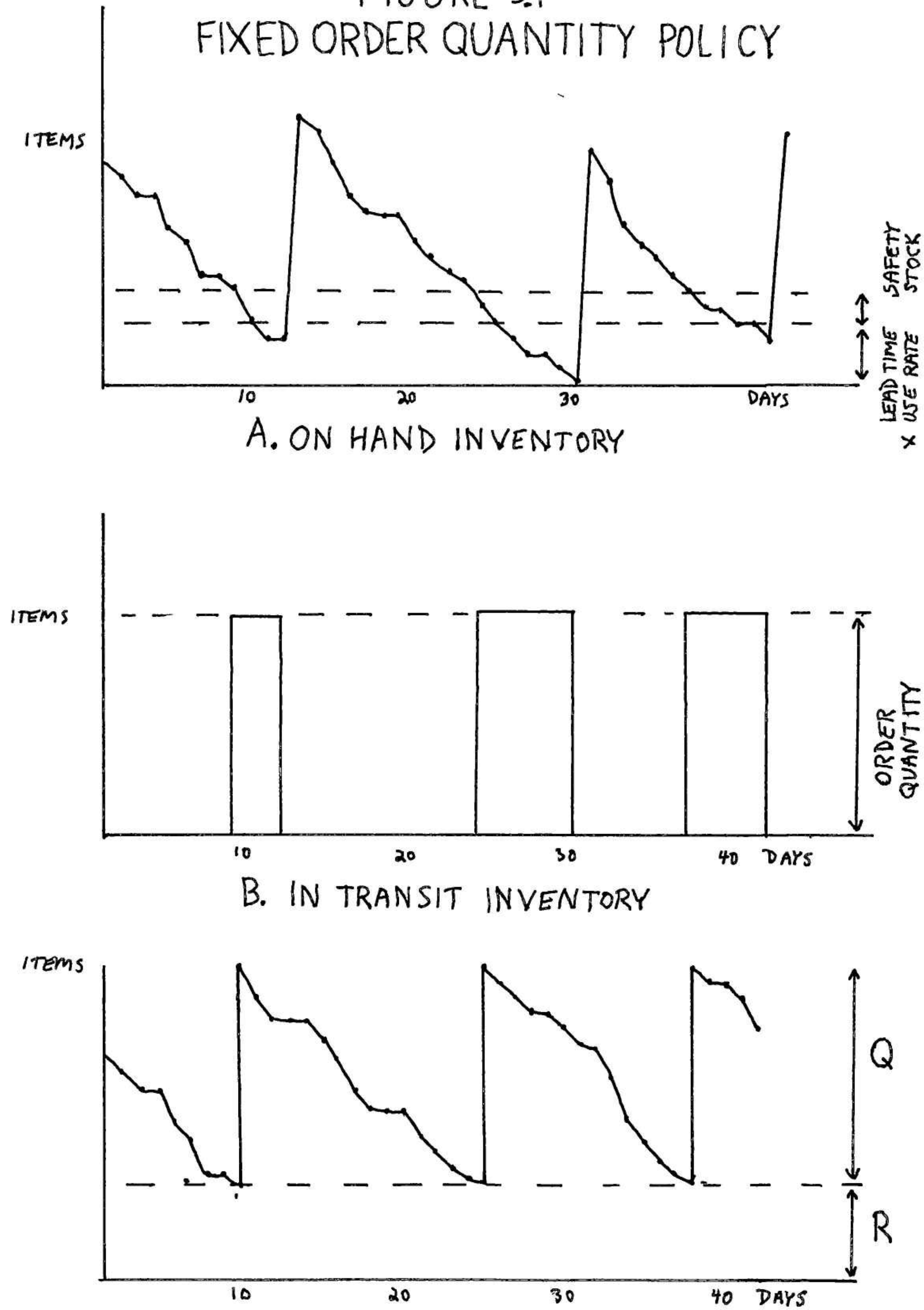
a. Transportation: Choice of mode to transport inputs from their point of origin to the production facility and outputs to storage facilities or to the customer. As discussed in Chapter 1, mean trip times, reliability, costs, and loss and damage all influence this decision.

b. Storage: Locations where inputs and outputs are held prior to use or shipment. Warehousing operations are established near demand centers to facilitate final distribution of goods and to facilitate response to variations in final demand. The number, size, and location of storage facilities are all decision variables for the logistics manager. The quality and cost of transportation is instrumental in these decisions.

c. Inventory: Goods held in storage (inventory on hand) or in the distribution pipeline en route to or from the firm (in transit inventory).

d. Fixed Order Quantity Policy: An order for a pre-determined number of items is placed whenever inventory on hand falls below a specified level called the Reorder Point (Figure 3.1).

FIGURE 3.1
FIXED ORDER QUANTITY POLICY



C. TOTAL INVENTORY

e. Fixed Order Period Policy: An order for a sufficient number of items to bring total inventory up to a pre-determined level is placed at specified time intervals (Figure 3.2).

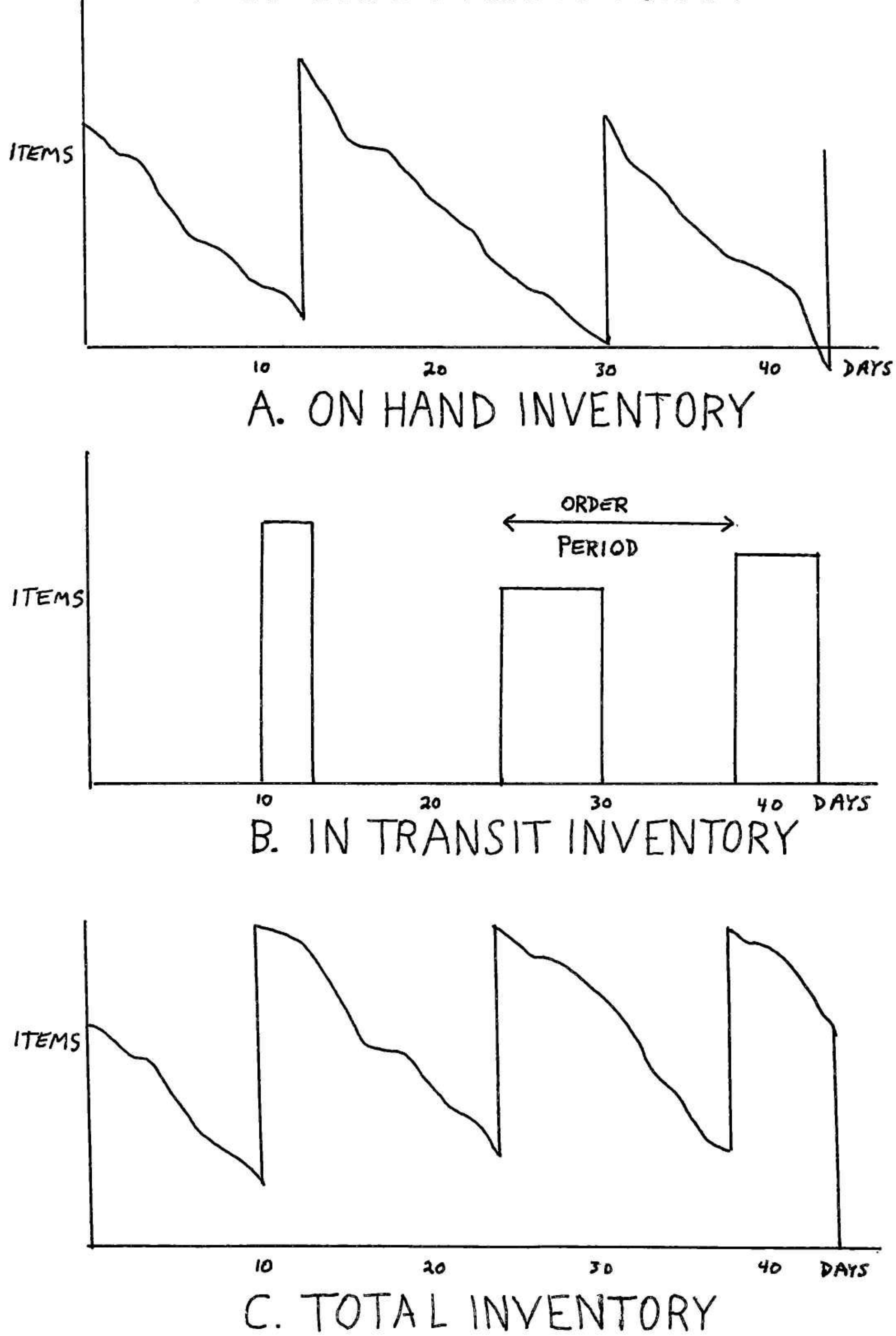
f. Lead Time: The period between the time that an order is placed and the time that it actually arrives. Lead time is a function of communications, ordering, and materials handling procedures as well as the trip time distribution of the mode of transportation that is used.

g. Safety Stock: The reorder point defines the amount of inventory available to service demand after an order is placed but before the shipment arrives. Safety stock is that portion of this inventory which protects against running short of inventory (stocking out) due to either a longer than average shipment time or greater than average demand during that time. Were both the shipment time and the use rate constant, no safety stock would be necessary.

h. Taxes: Real estate tax on facilities, personal property tax on inventory, and state business taxes add to the expense of the logistic inventory system. In so far as transportation affects the amount of and the location of storage facilities, taxes will increase with unreliability.

i. Materials Handling: Movement of materials within the firm; loading and unloading operations.

FIGURE 3.2
FIXED ORDER PERIOD POLICY



j. Communications: Means used to place orders and to ask questions about delivery or other problems. Telephone, mail and computer-based tracing systems are examples. Unreliable transportation can increase the need for communications facilities and employees.

k. Order Processing: Action taken to ensure delivery of and accounting for an order once it is received. Note that this can be a non-trivial element of the lead time.

l. Packaging: Protective measures taken to limit loss and damage during transit.

3. Logistics Management

Logistics management involves coordination of logistics activities and policies, processing information, and evaluating costs and trade-offs. Numerous possibilities exist for increasing the cost of one component of the system in order to decrease the costs of others. For example, using fast, reliable, but expensive transportation reduces the lead time, the total inventory required, the expectation of stock-outs, and perhaps overall logistics costs. It could also result in better customer service, a more efficient location of warehousing facilities, and even more efficient production processes.

The nature and knowledge of the costs of such tradeoffs and hence management policy vary greatly between industries and between firms. For most large firms, determining the optimum logistics system with respect to profitability of the firm is probably an unreasonable goal due to the complex relationships between service, production, costs, and final demand[2]. Understanding the cost and service implications of using a given logistics system is a more realizable goal. The question that can be asked is "Given fixed production and distribution facilties and therefore fixed demand for transportation, what mode of transportation and what inventory policy will provide the lowest total logistics costs?" This question can be answered by a computer program developed by Roberts[3] if the following information is supplied:

1. the nature of the item and the level of demand for the item
2. the length of haul
3. characteristics of the transportation possibilities
4. all relevant cost information including stock-out costs

This approach is based on the ideas that minimum system costs are not achieved by minimizing a single component cost such as that of transportation and that any of the component costs is likely to be

---

[2]Heskitt, Ivie, and Glaskowshy, pp. 499-501.

[3]Roberts, "The Logistics Management Process as a Model of Freight Traffic Demand."

very important for some commodities. In particular, improving rail service may be more important than lowering rail rates.

4. The Impact of Unreliability

In the context of this research, the effect of transport unreliability (defined here as the variability in transit times) on the consignee's costs is of particular interest. Clearly this is but one factor affecting basic logistic decisions and it is difficult to isolate costs that can be directly attributed to unreliability. For a given firm, the fact that a shipment does not arrive when expected could have any of the following consequences[4]:

- No effect at all other than "executive heartburn"
- Extra demurrage charges
- Stock outs of various durations (with the possibility of a loss of sales or of a customer)
- Shut down of the production process
- Loss of value (perishables, Christmas cars arriving in January)
- Placement of expedited orders using a faster mode
- Disruption of the loading-unloading process
- Special communications (attempting to locate the shipment)

---

[4]Heskitt, Ivie, and Glaskowshy, pp. 433-434; Riestrup, "On the Selling of Service"; Smith, "Service Problems in Communication"; and Quibble, "Consumer Products."

It should be emphasized that the cost of these effects, especially stockouts, can be extraordinarily difficult to quantify. They are, however, very important and a major logistics function is to avoid them. Increasing inventory, changing to a more expensive mode, hiring more people for materials handling, foregoing production processes which are critically dependent on timely delivery of inputs, and merely allowing more time for delivery are only some of the ways a company may seek to eliminate stockouts. The nature of the industry, the company, and the commodity would influence action taken in any specific instance.

The cost of unreliability thus includes long-run system costs such as those associated with inventory and hard-to-measure short-run costs such as those associated with stock-outs. Firms with identical transportation requirements and service will not perceive identical short-run costs of unreliability except in the unlikely event that their logistics systems are the same. Hence, looking at short-run costs alone will not give the total costs of unreliability; it may not even be possible to distinguish these costs from those caused by fluctuations in demand or by unreliability in the other components of the lead time. In any case there is no reason to derive a measure of unreliability which reflects either the short- or long-run costs alone.

B. An Examination of Inventory Requirements as a Function of Transportation Characteristics

Although decidedly not the only aspect of logistics affected by transportation, inventory policy must make explicit trade-offs between the short- and long-run costs of unreliability. Therefore it may provide some insight into the relative and absolute importance of various characteristics of transportation service. It is necessary to identify the most important of these so that appropriate performance measures can be chosen in Chapter 4.

In order to clarify the terms that will be used, consider a firm with a fixed order quantity inventory policy which orders Q items whenver its inventory falls to the reorder point R. The firm's total inventory as a function of time will be similar to Figure 3.1C. The firm's average total inventory T is

$$T = \tfrac{1}{2}\,(\text{order quantity}) + (\text{minimum inventory level})$$
$$= Q/2 + R$$

The order quantity Q equals the total annual use divided by the number N of shipments made per year. The reorder point R is a function of both the daily use rate and the trip time distribution. If these are constants K and T, then the reorder point R = KT provides exactly enough inventory to last until the shipment arrives. More commonly, the daily use and the trip time will be random variables and R must be greater than the product of their means in order to decrease the probability of a stockout. The critical issue in inventory policy is what reorder point will best resolve the trade-off between inventory and stock-out costs. Often this is not or cannot be done explicitly and an acceptable level of stockouts is defined by management. In general, this level will depend on the value and the use of the item, storage costs, and the consequences of a stockout as well as the transportation level of service.

Kullman has developed a simple model which can be used to test the effect of the use and trip time distributions on the reorder point given an acceptable level of stockouts[5]. Each point in the graphs in Figure 3.3 represents the reorder point necessary to achieve the indicated stockout probability for a particular trip time distribution. Both the mean trip time and the compactness of the distribution effect the magnitude of the reorder point. As shown in 3.3B, increasing the compactness of the distribution even without changing the mean transit time will lower the reorder point necessary to achieve any of the stated stockout levels[6]. An alternative to lowering R is reducing the desired stockout level; for example in 3.3B, improving the compactness without changing the mean decreases this probability from 10% to 1% if R remains at 1800

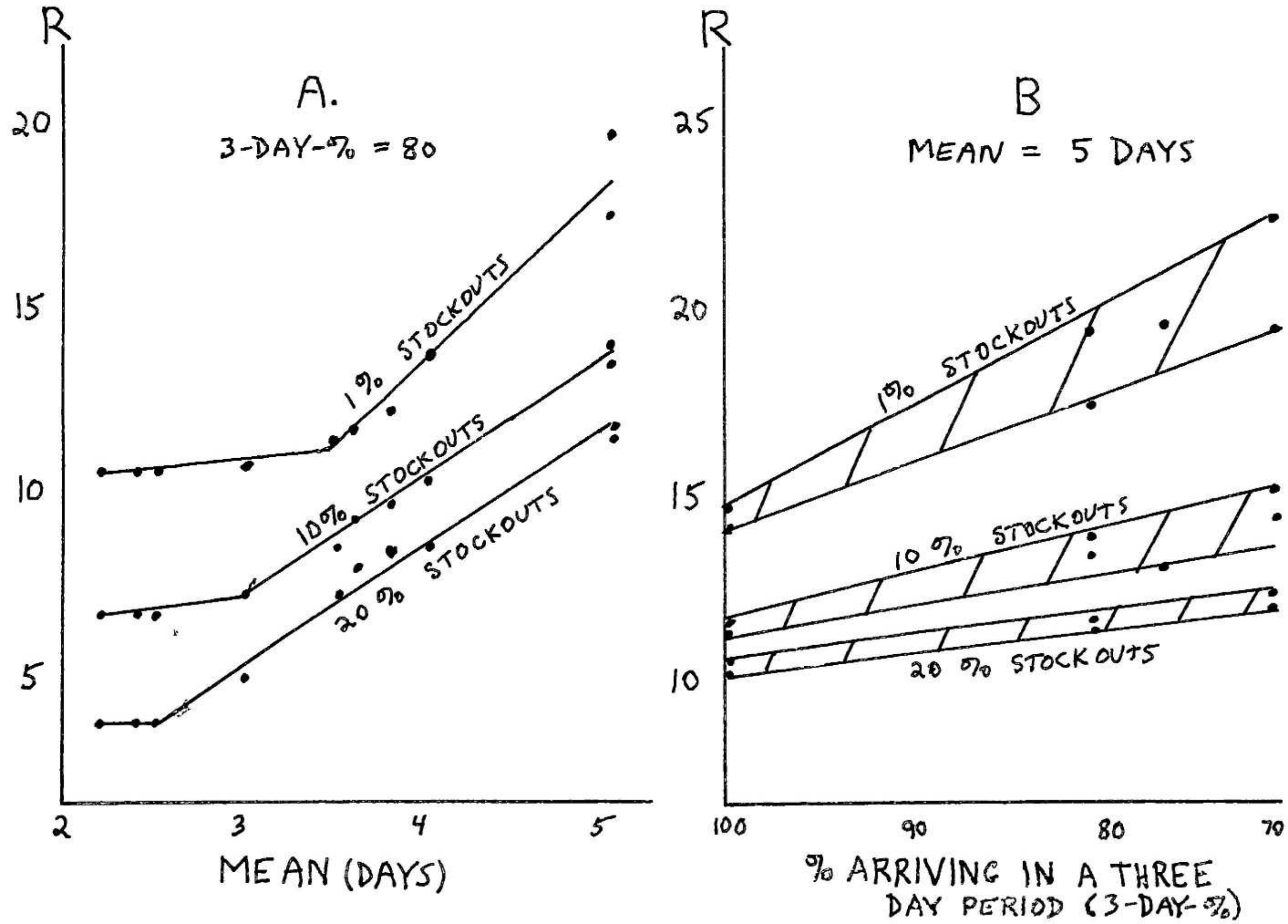
---

[5]M.I.T. Dept. of Civil Engineering, Progress Report on the Study of Reliability in Railroad Network Operations, January 1972, Task 1.

[6]An implicit assumption is that the daily use distribution remains more reliable than the trip time distribution.

FIGURE 3.3

REORDER POINTS NECESSARY TO ACHIEVE DESIRED LEVELS OF STOCKOUT PROBABILITY FOR LEVELS OF TRANSPORT SERVICE



units. At any rate, the compactness of the distribution is one important characteristic of the trip time distribution.

Decreasing the mean trip time can also lower the reorder point. If the entire distribution is shifted one day to the left, the reorder point can be reduced by the average amount used daily and the probability of a stockout will remain approximately the same. The change in R can be either greater or less than this if, as is usually the case, other characteristics of the trip time distribution are also affected. The importance of the mean is depicted in Figure 3.3A. Notice, however, that continued improvements in the mean do not lower the reorder point, particularly when a maximum of 1% of stockouts is specified. The reason for this phenomenon is that enough inventory must be held to supply demand in the event that the trip time is exceptionally long. For example, if there is a greater than 10% chance that the shipment will take at least 6 days to arrive, then at least 6 days of inventory must be held for the probability of stockouts to be kept below 10%. The extent of these long delays is a third characteristic of the trip time distribution critical to inventory policy because such delays account for most stockouts if the use rate is fairly constant or predictable. Performance measures should reflect each of these three characteristics.

This analysis demonstrates that improved rail service can reduce the amount of safety stock that must be held. However, since T = Q/2 + R, the importance of reductions depends on the initial magnitudes of Q and R. Q, the order size, is inversely proportional to the number N of orders made per year; therefore, the higher N or the higher the initial reorder point, the more important reductions in R will be. Table 3.1 gives representative reorder points from Figure 3.3 as a fraction of the average total inventory for various shipment frequencies. For example, if the mean trip time in Figure 3.3A is reduced from 5 to 3 days, the reorder point falls from 16 to 8 units. This 8 unit reduction is always an 8 unit reduction in T also, but it is only a 12.5% reduction in T if shipments are made monthly (case 3, Table 3.1). On the other hand, it is a 45% reduction in T if shipments are made daily (case 9) and a 29% reduction with weekly shipments (case 6). Generalizing from this example, improvements in transportation service will be most important for shippers who desire low levels of stockouts (i.e., with high reorder points) or who make frequent shipments of quantities small relative to annual use.

C. Summary

Transportation is a key element in a shipper's logistics system. Minimizing transportation costs will not necessarily minimize

TABLE 3.1

SAFETY STOCK AS A PERCENTAGE OF TOTAL INVENTORY FOR VARIOUS REORDER POINTS AND ORDER QUANTITIES, BASED ON AN AVERAGE DAILY USE RATE OF 3K

The values of the reorder point are representative of those in figure 3.3.

Case	N	Q	R	T	R/T
1	12	90K	5K	50K	.10
2			10K	55K	.18
3			15K	60K	.25
4	52	21K	5K	16K	.31
5			10K	21K	.48
6			15K	26K	.58
7	365	3K	5K	7K	.72
8			10K	12K	.84
9			15K	17K	.89

N = orders per year
Q = order quantity
3K = average daily use
R = reorder point = safety stock
T = average total inventory = ½ order quantity + safety stock = Q/2 + R

logistics costs, nor will the minimum cost logistics system necessarily be optimum with respect to the profitability of the firm as a whole. Transportation unreliability can cause direct expenses associated with early and late shipments, but it also occasions additional fixed costs to limit the negative effects of such shipments. Were transportation perfectly reliable, the firm's profitability might also be increased by restructuring the entire logistics, production, or marketing systems. It is important to emphasize that shippers are concerned with actual rather than scheduled performance.

Although only one aspect of operations affected by transportation unreliability, inventory control must explicitly confront the tradeoff between short- and long-run costs. The mean trip time, the compactness of the Total Trip Time distribution, and the extent of very long delays are three characteristics of rail service which seem to have significant impacts on inventory levels. Improving these characteristics will give the greatest percentage reductions in inventory for firms which ship frequently or which desire a low probability of stock-outs.

CHAPTER 4

RAILROAD RELIABILITY PERFORMANCE MEASURES

A. Evaluating Performance Measures

1. What is a Performance Measure?

Chapter 1 presented the concept that railroads sell freight transportation in competition with other modes and that trip time unreliability is one of four important factors in shippers' modal choice decisions. Unreliability, defined as "variability in trip times" in Section 1.C, was shown to have two important characteristics in Section 3.C: The probability of very long trip times and the compactness of the Total O-D Trip Time distribution. If railroads adopt the goal of improving rail reliability in an attempt to increase their market share, they must be able to quantify exactly what is meant by these characteristics. Otherwise it will not be possible to evaluate intelligently alternative means of achieving improvements.

For example, a hypothetical firm ships 21 tons of an item per week or 1100 tons per year. It desires less than 1% stockout and its inventory costs are $200 per ton per year.

Considering the following trip time probability distributions:

	Trip Time in Days									
	1	2	3	4	5	6	7	8	Mean	3-day-%
A	0	0	.1	.4	.2	.1	.1	.1	5.0	70
B	0	0	0	.25	.5	.25	0	0	5.0	100

The mean and the 3-day-% are the measures used in Figure 3.3 to estimate the reorder point for a number of trip time distributions. Assuming the use rate implied by that figure is equivalent to that of the firm in this example, Figure 3.3B indicates that at a 1% stockout level, distribution A requires a reorder point of about 24 tons while B requires only 17 tons. At $200 per ton per year, the firm would save $1400 per year in inventory costs by receiving the level of service represented by B. This is equivalent to $1.20 per ton shipped and could be a critical factor in a modal choice decision. At a constant rate level, the firm would clearly prefer not to receive that level of service represented by A. Now consider a railroad that measures its performance by the mean trip time and the percentage of cars that arrive "on time" relative to the appropriate freight schedules. If in this case the scheduled trip time is 4 days, the on time performance is as follows:

Distribution A -- 50% on time

Distribution B -- 25% on time

Railroad management thus perceives distribution A as providing a superior level of service whereas it actually increased one shipper's reorder point nearly 50% and cost him $1400 per year in inventory costs.

As demonstrated by this example, performance measures occupy a critical position in a management information and control system because they

1) determine how basic car movement data will be aggregated for presentation to management, and

2) form the basis on which standards of performance are set

Standards which define acceptable levels of measured performance and thereby influence inumerable operating decisions are only as good as the underlying performance measures. If these performance measures do not in fact reflect management's goals, then wrong decisions will be made and scarce resources will be misallocated. In particular, if the measures do not reflect service as perceived by shippers, then measured improvements in service may actually be meaningless.

Railroads will likely want an integrated set of performance measures reflecting network and sub-network as well as O-D performance. Measures reflecting the effect of unreliability on shippers' logistics will not be sufficient for these overall purposes however well they are suited to marketing. Additional measures will be necessary to identify the causes of unreliability and to insure that decisions made at sub-network levels will not have unintended repercussions on network performance. Specific measures of O-D performance will be evaluated after general criteria applicable to performance measures and special problems of railroad data sets are reviewed.

2. Criteria for Evaluating Performance Measures

A.S. Lang has suggested that performance measures should satisfy a number of criteria which relate to the way in which performance measures will be used[1]. These are reviewed below in relation to their implications for railroad reliability measures.

Explicitness, Understandability, and Reproduceability: Measures should be precisely defined and clearly relate actual performance to desired goals.

Applicability to Both Present and Future Operations: Measures should not be linked to standards which are dependent on current operating policy. For example, "on time performance" is not a measure of reliability because on time is a function of current schedules. Measures must be based solely on the relevant trip time distribution, not on current perceptions of the important aspects of that distribution.

Relevance: Measures must be relevant to management goals. In chapter 3, shippers were found to be interested in 1) the Total O-D Trip as defined in Chapter 2, b) actual rather than

[1]These criteria were suggested in a lecture for Transportation Systems Analysis III at M.I.T.

scheduled performance, c) mean trip times, d) very long trip times, and e) the compactness of the trip time distribution. Hence if management desires to improve the competitive position of railroads by increasing reliability, it must use measures which reflect these considerations. In particular, railroads should in this case measure very long trip times and the compactness of the trip time distribution in addition to the mean O-D trip time.

Controllability: Presumably a railroad can affect O-D performance by changing schedules or policies or by re-allocating men and equipment. Since it cannot normally affect performance over other railroads, it must measure performance of cars only on its own network. However, the previous criterion implied that the Total O-D Trip is of interest to the shipper, not just the portions on an individual railroad. Hence the railroad industry should consider ways to monitor joint performance when shipments are handled by more than one carrier.

Inexpensive and Accurate Observability: In particular, the measures should not be overly sensitive to the types of errors that are prevalent in railroad data sets (see below, Section 4.A.3).

Independence: Mean trip time, very long trip times, and the compactness of the trip time distribution have each been hypothesized to be important determinants of modal choice decisions. These three aspects of rail service should be measured by three independent measures rather than by one or two composite measures.

Consistent Operational Effectiveness: As mentioned above, trip time measures must be related to both network and sub-network measures. Line haul and yard measures should promote operating decisions which have the desired impacts on trip and network performance.

3. Problems in Measuring Rail Service

Aggregation, data errors, weekends, and routing all cause special problems in measuring rail service. Since these problems are inherent to the industry, data sets and performance measures should be constructed which minimize the distortions that they induce. These problems are discussed in some detail below because they give valuable insights into the causes of measured reliability.

Aggregation: Aggregating data from a number of shippers to obtain yard-to-industry performance has already been found necessary for practical and theoretical reasons in Chapter 2. However, further aggregation of data by city or by region as is sometimes done will have an adverse impact on measured reliability. Figure 4.1 shows that combining moves from city B to yards 1 and 2 of city A would obscure the fact that one move was fairly reliable and that the other was fairly unreliable. Figure 4.2 indicates how the same aggregation could turn two reliable distributions into a less reliable one.

FIGURE 4.1

EFFECTS OF AGGREGATION (DESTINATION)

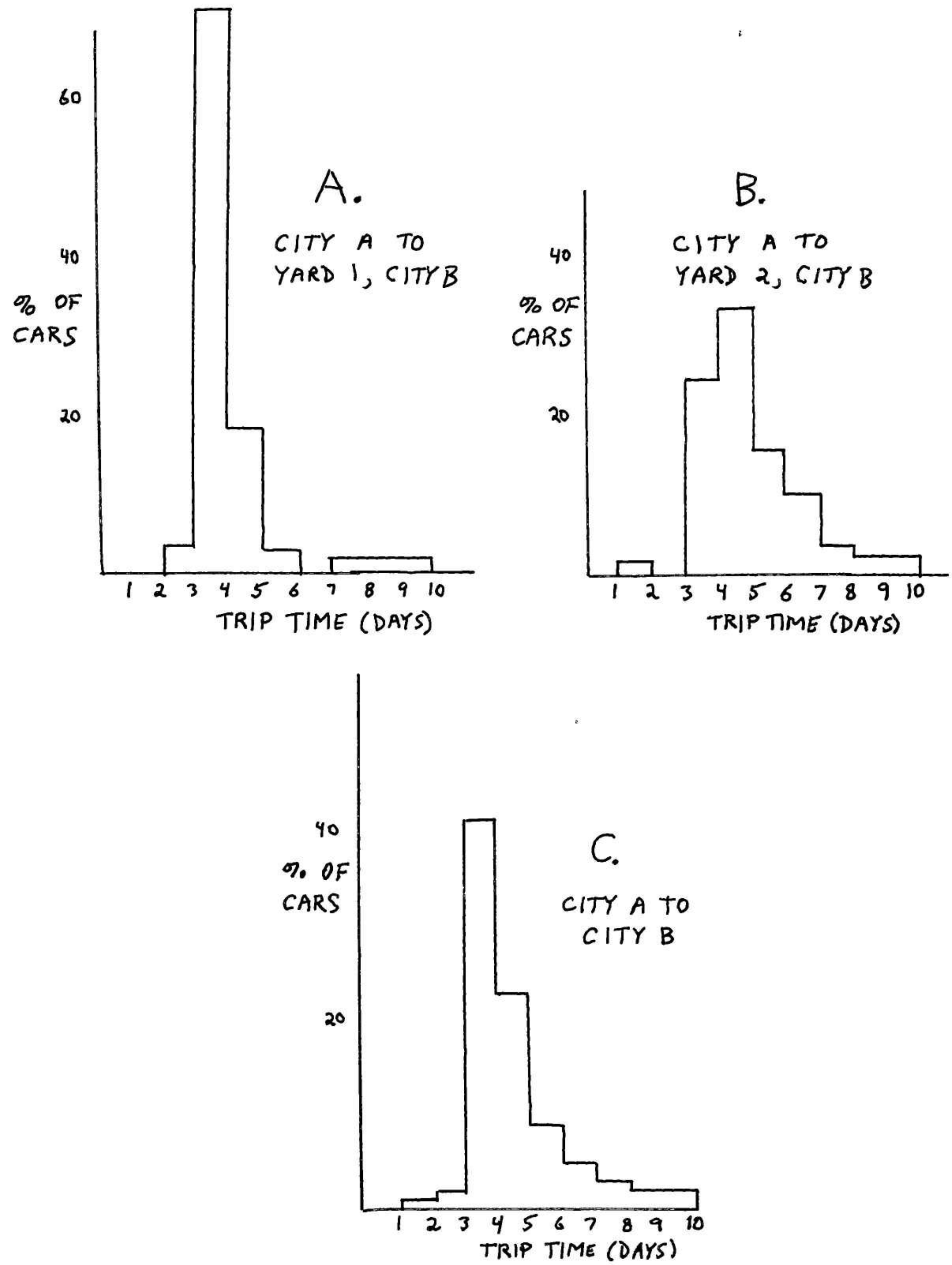
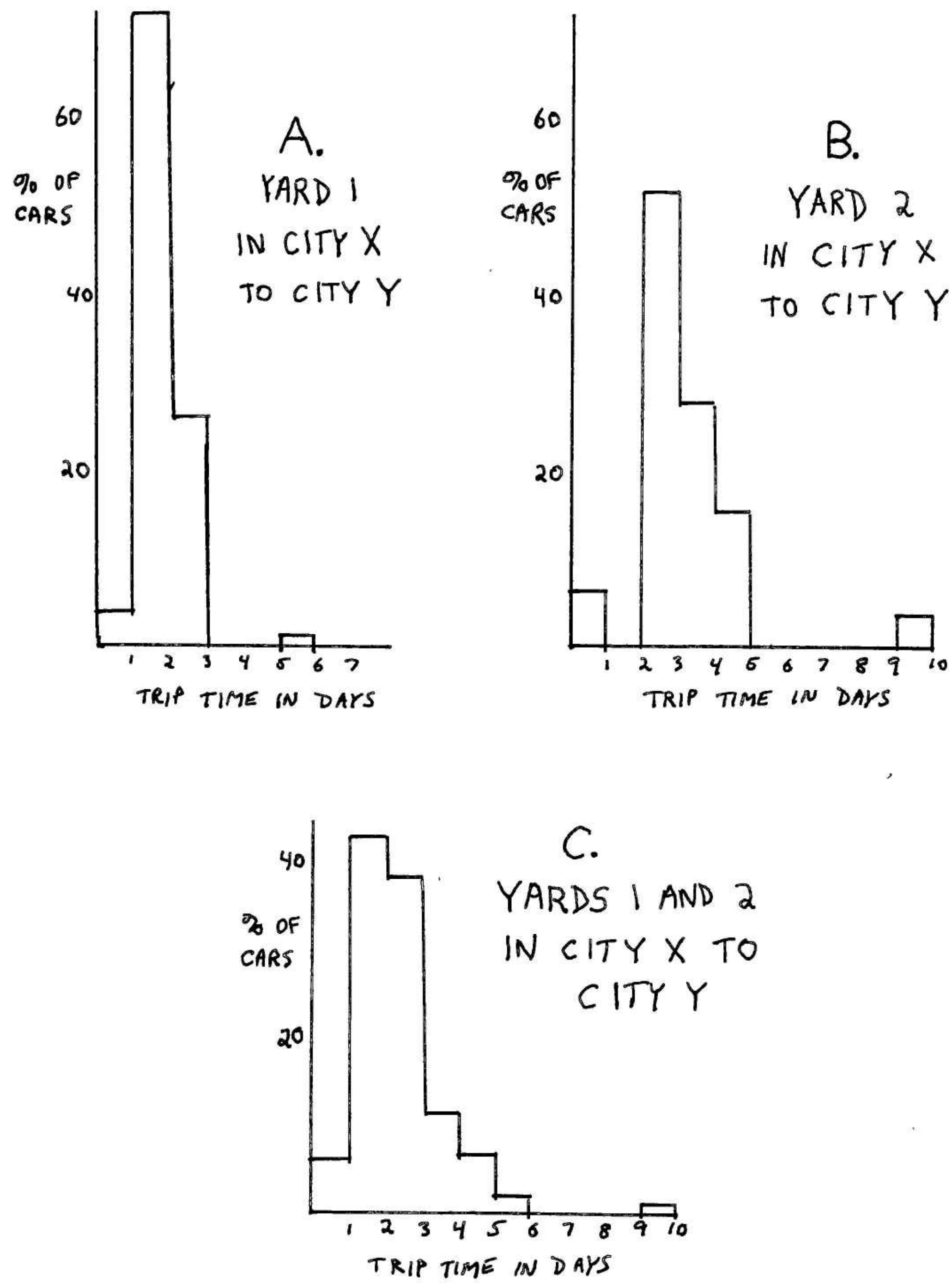


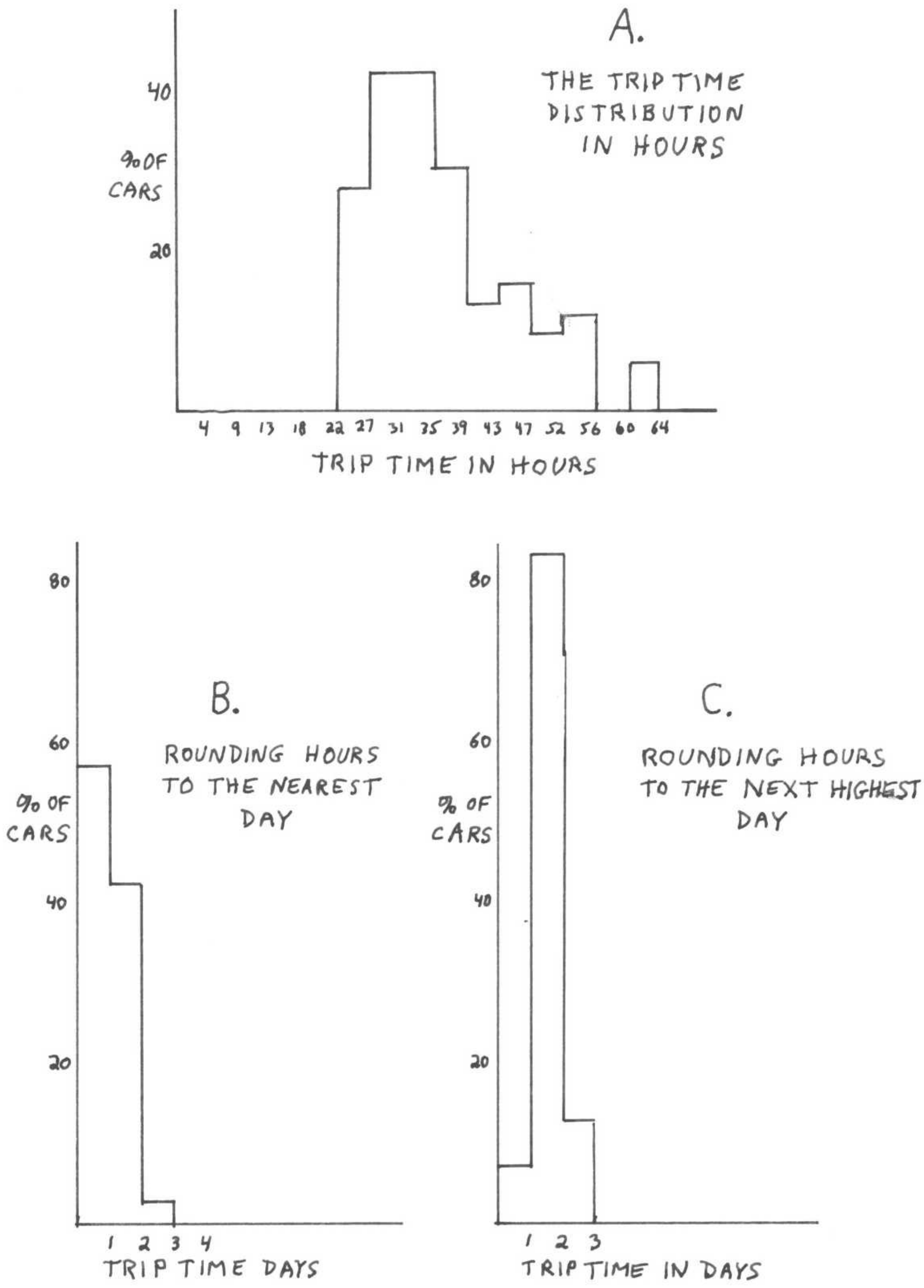
FIGURE 4.2
EFFECTS OF AGGREGATION



Thus, it may be possible tp speak of service from yard A to yard B but not of service from City 1 to City 2. In the latter case, the _measured_ service is unlikely to be relevant to the _actual_ service received by a shipper in a particular sector of the city -- and it is precisely this service that must be the ultimate concern of the railroad industry. This type of aggregation necessitated eliminating several O-D pairs from the analysis for data from Railroad B.

Most of the distributions in this report are based on trip times measured in days. The manner in which we convert time in hours to time in days is another form of aggregation which affects the shape of the resulting distribution. This is shown in Figure 4.3. The actual distribution of trip time from interchange to constructive placement for a particular move is given in 4.3A. Figure 4.3B gives the distribution in days if the time in hours is rounded to the nearest day. In this case, 57% of the cars arrive between 12 and 35 hours or 1 day, 40% between 36 and 59 hours or 2 days, and 3% between 60 and 83 hours or 3 days. The distribution appears to be markedly improved when the time in hours is rounded to the next highest day. In Figure 4.3C, 7% arrive between 0 and 24 hours which is considered one day, 80% arrive in 2 days, and 13% in 3 days. Note that if the original distribution were shifted twelve hours in either direction, the distributions in B and C would be reversed, implying that neither way of converting

FIGURE 4.3


A.
THE TRIP TIME DISTRIBUTION IN HOURS
40
% OF CARS
20
4 9 13 18 22 27 31 35 39 43 47 52 56 60 64
TRIP TIME IN HOURS
B.
ROUNDING HOURS TO THE NEAREST DAY
80
60
% OF CARS
40
20
1 2 3 4
TRIP TIME DAYS
C.
ROUNDING HOURS TO THE NEXT HIGHEST DAY
80
60
% OF CARS
40
20
1 2 3
TRIP TIME IN DAYS

to days is inherently superior to the other. The implication of this exercise is that performance measures will always be subject to some distortion due to the manner in which the movement time distribution has been constructed from the detailed movement data.

Data Errors: Railroads B and C both indicated that very high and very low values of trip time are likely to be the result of erroneous input data. Thus, care must be taken in dealing with ranges.

Routing -- Unique routings for cars between an origin and a destination do not always exist. Suppose that cars going from X to Y go by way of V or W with equal probability. If the trip XVY always takes 2 days while the trip XWY always takes 3, then a shipper will see 50% of his cars arriving in 2 days and 50% in three days, even though each routing is itself perfectly reliable. Hence, performance measures should be based on all moves between an O-D pair, not only moves with the same routing.

Extreme Values -- Not all extreme values are errors or even bad moves. Cars re-routed in transit, partially loaded or unloaded en-route, or shipped with no ultimate destination specified will have long trip times which reflect the flexibility of rail service rather than its unreliability. The moves which actually reflect poor service may result from a single special delay - misroute, no-bill, bad-order - or a concatenation of delays resulting from normal network operations. This study has not attempted to differentiate between these classes of extreme values; instead, it has concentrated on the network unreliability induced by normal operating policies, resources, and demands. Long delays, identified as an important factor in stockout probabilities, are a suitable subject for future research.

Weekends -- A problem with the destination terminal time is that if consignees' firms are closed on weekends, cars will not be put out to placement or to constructive placement until Monday morning. If no correction is made for this possibility, then data which uses placement as the termination of the trip will overstate actual trip time. On the other hand, data which uses arrival at the last yard as the termination of the trip will overstate unreliability if railroads give low priority to traffic scheduled to arrive Saturday or Sunday that cannot be placed until Monday.

Figure 4.4 shows the kind of distortion introduced into trip time distributions by weekends. Five cars per day are shipped from various locations in City A to a single location in City B. For illustrative purposes, assume that one of these always takes two days to get to a yard at B, three take three days, and one takes four. Figure 4C shows the distribution that would result if there were seven day delivery from B. If there were no deliveries on Saturday and/or Sunday, some cars would be held in the yard an extra day after their arrival for delivery on Monday (Figures 4A and 4B). The actual service provided from industry to final yard

FIGURE 4.4

DISTORTIONS IN RECORDED MOVEMENT TIMES DUE TO WEEKENDS (HYPOTHETICAL DATA)

A. NO SUNDAY DELIVERIES

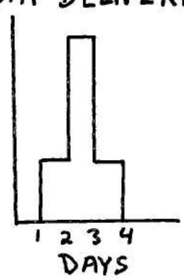
DAY SHIPPED	DAY RECEIVED W	T	F	S	S	M	T
M	1	3	1				
T		1	3	1			
W			1	3		1	
T				1		4	
F						4	1

B. NO WEEKEND DELIVERIES

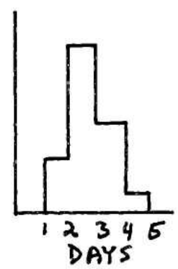
DAY SHIPPED	DAY RECEIVED W	T	F	S	S	M	T
M	1	3	1				
T		1	3			1	
W			1			4	
T						5	
F						4	1

TRIP TIME DISTRIBUTIONS:

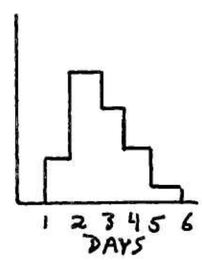
C.
7 DAY DELIVERY



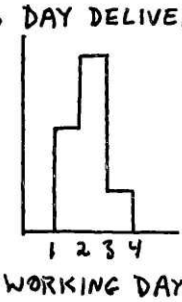
D.
6 DAY DELIVERY



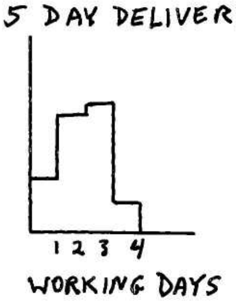
E.
5 DAY DELIVERY



F.
6 DAY DELIVERY



G.
5 DAY DELIVERY



is given in C, an average of 3 days, and 60% arriving in one day. Weekends reduce the measured level of service from industry to industry to 3.2 days and 52% arriving on a single day (closed Sundays) or 3.8 days and 40% (closed all weekend). It is conceivable that the consignee measures service by working days in which case the weekends act to improve trip time and to hurt reliability. With 6 day delivery (4F) the average is 2.8 days and the reliability is 56%; for five days delivery, the average is 2.2 days and reliability is 40%. In short, Figure 1 demonstrates how sensitive reliability is to the way it is measured and to the way it is perceived.

B. Strengths and Weaknesses of Possible Performance Measures

A number of possible reliability measures are illustrated by the hypothetical trip time distribution in Figure 4.6. Each may be useful in some situations, but many have critical weaknesses with respect to criteria developed in the previous section and summarized in Figure 4.5. In this section, these measures will be evaluated with respect to these criteria and specific measures will be chosen for the two characteristics of unreliability, very long delays and the compactness of the trip time distribution.

FIGURE 4.5

EVALUATION CRITERIA FOR RELIABILITY PERFORMANCE MEASURES

A. General Criteria

1. Explicit, Understandable and Reproducable
2. Applicable to Both Present and Future Operations
3. Relevant to the Effects of Unreliability
4. Controllable at Decision-Making Levels
5. Readily, Inexpensively, and Accurately Observable
6. Independent
7. Consistent and Operationally Effective
8. Insensitive to Extreme Values Which May be Errors or Valid, Reliable Moves
9. Insensitive to Small Distortions Caused by Weekends and Necessary Aggregations

Requirements for Data Sets

1. Should include O-D trip times as defined in Chapter 2
2. Should include all moves for each O-D pair, not merely those with similar routing
3. Should not aggregate moves from more than one origin yard or to more than one destination yard

FIGURE 4.6

EXAMPLES OF POSSIBLE PERFORMANCE MEASURES

	Trip Time in Days											
	1	2	3	4	5	6	7	8	9	10	11	12
% Arriving on Specified Days	0	4	15	27	18	10	6	4	2	4	6	4

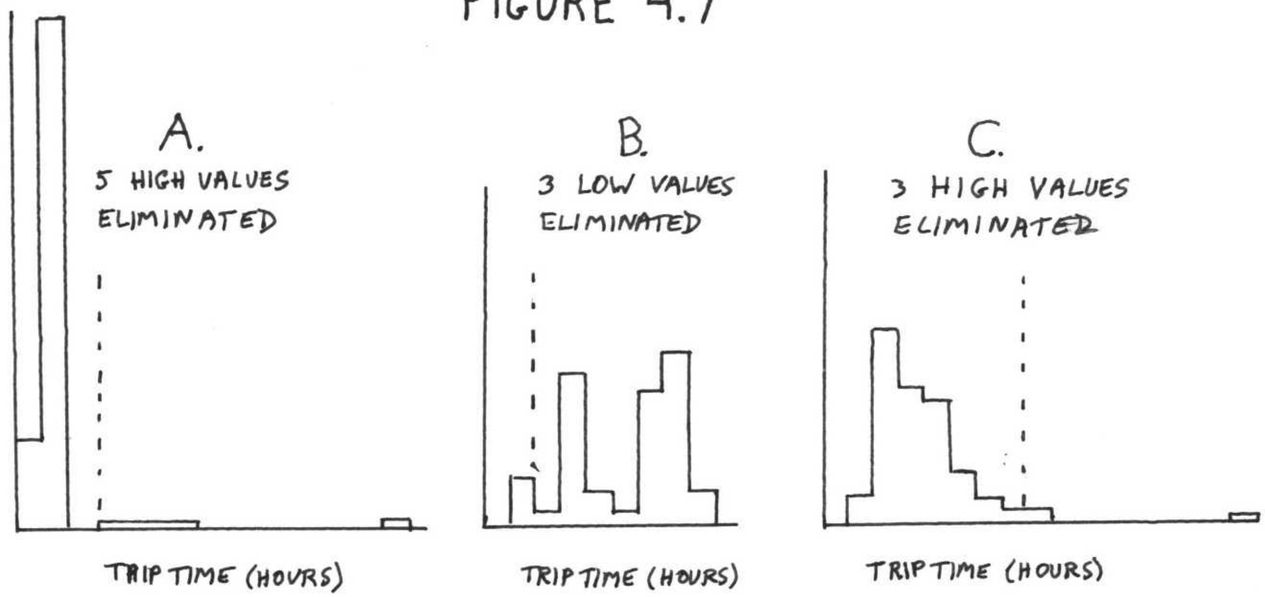
Schedules:	3 days
Mode:	day 4
Median:	day 5
Mean:	5.3 days
Standard Deviation:	3 days

A.	Time by which 90% of cars arrive	day 10
B.	Range in which 90% of cars arrive (starting with first arrival)	day 2 - day 10
C.	Shortest interval in which 80% of cars arrive	day 2 - day 7
D.	Maximum % of cars arriving in any three day interval (day 3 - day 5)	60%
E.	% of cars arriving after that three day interval (day 6 - day 12)	36%
F.	% of cars arriving within 2 days of scheduled arrival (day 1 - day 5)	64%
G.	% of cars arriving beyond 7 days	20%
H.	% of cars arriving 2 or more days beyond average (rounded to nearest day) (day 7 to day 12)	30%

1. Standard Deviation: Relative to normal distributions, rail freight trip time distributions are generally skewed to the left as a result of 1) minimum trip times, and b) frequent long trip times. The preponderance of extreme values alone destroys the usefulness of this statistic for comparative purposes. In Figure 4.7, the 3 distributions have essentially the same standard deviation although A intuitively appears to be the more reliable. Since Penn Central and Southern officials have indicated the highest values are likely to be data errors, the standard deviation will give a meaningless measure of performance. Even without data errors it is likely to reflect the existence of extraordinary values rather than the compactness of the distribution. The measure is rejected because it is not understandable, independent, relevant, or insensitive to extremes.

2. Standard Deviation -- Exclusion of Extreme Values: Eliminating the longest trip times before computing the standard deviation vastly increases its value as a performance measure (Figure 4.7). However, the standard deviation still does not give a reliable estimate of the compactness of a skewed distribution. That is, the normal interpretation of the standard deviation does not hold with the type of distributions encountered in studies of railroad unreliability. Thus the measure fails to reflect reliability as viewed by the shipper and must be rejected because it is

FIGURE 4.7



	A.	B.	C.
σ_{LH} :	25	25	26
σ_{LH}^{*} :	3	24	16
MEAN :	23	67	41
MEAN* :	17	70	37

ONLY AFTER ELIMINATING THE HIGHEST VALUES AND ANY UNREASONABLY LOW ONES DOES THE STANDARD DEVIATION GIVE A RANKING OF THESE DISTRIBUTIONS THAT REFLECTS THE OBVIOUS DIFFERENCES IN COMPACTNESS not relevant, understandable, or independent.

3. Variance, Exclusion of Extreme Values: Although the variance suffers from the same shortcomings as the standard deviation, it can be used to obtain a measure of Total O-D Trip reliability given only the reliability over the networks of each of the railroads involved in the Total O-D Trip. Assuming that performance over each road is independent of performance over all other raods, the variance of the Total O-D Trip will be the sum of the variances of the O-D trips measured by each railroad. In all other cases, however, the measure must be rejected for the same reasons as the standard deviation.

4. Ranges in Which X% of All Cars Arrive: Since this interval is critically dependent on the earliest arrival time which is in turn subject to be a data error, this measure must be rejected.

5. Percentage of Cars Arriving Within N Days of a Predetermined Standard: This measure is in fact used by railroad B to give a measure of "on time" performance. However, it is not a measure of variability in trip times, it is subject to change as a result of functionally meaningless changes in the standard, and it is not based solely on the trip time distribution. Management will undoubtedly want to set standards for reliability, but such standards must be based on measures of variability in trip times or of very long trip times. Rather than compare each individual

movement to a standard, statistics of the trip time distribution for a given period of time will be compared to a standard. Reject this measure because it is not relevant, independent, or applicable to future operations.

6. Percentage of Cars Arriving Within N Days of the Scheduled Arrival Time: This is a special case of the preceding measure and is equally unsatisfactory.

7. Shortest Interval in Which X% of All Cars Arrive: This measure has the advantage of being relatively independent of distortions caused by extreme values and of not being tied to the mean trip time. Is is not as meaningful for distributions which are given in days as for those which are given in hours because of the sizeable magnitude of the discontinuities. Consider the following percentage distributions:

	Trip Time in Days							
	1	2	3	4	5	6	7	8
A:	0	20	20	20	20	10	10	0
B:	0	20	20	20	19	11	10	0

Although the two are nearly identical, the 80% interval for A is 4 days while that for B is 5 days; the measure reflects a difference that is no difference. However, this does seem to be a useful measure if the trip time distribution is given in hours.

8. Maximum Percentage of Cars Which Arrive in an N Day

Period (N-Day-%): This measure is equivalent to the previous one except that it is not affected by discontinuities. Both measures are understandable, independent of standards, related to the shippers perception of compactness, insensitive to data errors, and unrelated to the mean trip time. In addition, they provide a means of comparing distributions that is not greatly biassed by the effects of aggregation, weekends, or routing. In addition, the percentage in one case and the length of the interval in the other can be chosen to reflect the degree of unreliability observed. For instance, in this thesis, a three day interval is used for analyzing actual trip time data because the observed 3-day-% is seldom greater than 95 or less than 50. Clearly, if the majority of O-D pairs have 3-day-%'s greater than 90, a two day interval would be more appropriate[2].

9. % Late: If cars arriving during the three day period identified by the 3-day-% are considered "on time," then cars arrriving after than period would be late. Notice that this definition of on time and late depends only on the actual trip time distribution and not on any standards or schedules. This measure is most useful when used on conjunction with the 3-day-%; otherwise it is hard to understand.

---

[2] A 3-day-% of 90 was cited by W.K. Smith of General Mills as a standard for transportation reliability in "Service - Problems in Communication," in RSMA, The Measure of Railroad Freight Service.

10. Percentage of Movements Taking Longer Than N Days: Another measure used by Railroad B, the percentage taking more than N days satisfies all of the criteria with the exception of independence. Since this measure would be used in conjunction with the mean, it should reflect only those characteristics of the trip time distribution not described by the mean. However, this measure will in fact show more cars arriving late as the mean trip time increases and is useless if the mean is greater than N. Also, if the Total O-D Trip involves more than one railroad, the values of N which are reasonable for each individual railroad would be unreasonable if applied to the Total trip time distribution. This is a critical weakness and the measure must be rejected.

11. Percentage of Movements Taking N Days Longer Than the Median or the Mean (%-N-Days-Late): By using a characteristic of the trip time distribution itself, this measure resolves most of the difficulties associated with the previous measure. As with all measures of long delays, it will be sensitive to data errors, but it does satisfy the other evaluation criteria.

12. Time by Which X% of all Cars Arrive: The Xth percentile of the distribution expressed in days from the start of the trip also gives some idea of the number of very long trip times when used in conjunction with the mean. However, it may reflect differences that are no differences as was the case with (7), the shortest interval in which X% of all cars arrive. Hence the %-N-day-late is

probably a more generally appropriate measure.

C. Summary

Performance measures are necessary to translate goals into quantifiable concepts. They are a critical part of a management information and control system because they determine the nature of the information that is conveyed to management as a basis for numerous operating and policy decisions.

These measures should reflect the level of rail service as perceived by rail customers. In particular, they must be based on the Total O-D Trip Time distribution as much as possible. Carriers thus should jointly monitor performance from origins on one line to destinations on another[3]. However, they will normally be able to monitor only on-line performance from release or receipt from interchange until placement, constructive placement or delivery to interchange. Some aggregation will be necessary to develop meaningful trip time distributions; in Chapter 2 several reasons were given for monitoring local moves and O-D trips separately. Distortions in measured reliability will be minimized

---

[3]Railroads F and C recently developed the capability to measure service between various points on their networks.

if this aggregation is limited to moves between the same two terminal yards which receive similar priority. Neither combining several yards into a single origin or destination nor including perishables, TOFC, unit trains, or movements with in-transit privileges with general freight will give true measures of reliability.

O-D performance measures are necessary for each of the 3 characteristics of the trip time distribution which were found to be important to shippers in Chapter 3. After consideration of a number of possible measures and a set of evaluation criteria, the maximum percentage of cars that arrive in an N day interval (the N-day-%) was chosen as the best measure of compactness and the percentage of cars that arrive after this and the percentage of cars that arrive N days beyond the mean were chosen as acceptable measures of very long trip times. These two reliability measures together with the mean trip time describe those characteristics of the trip time distribution which are of most importance to shippers.

Measures of network reliability are readily derived from these measures of O-D reliability. The network 3-day-% is the percentage of all cars that arrive in an N-day interval for the appropriate O-D pair. Likewise, the % late for the network is the percentage of all cars that arrive after the appropriate N-day period and the network %-N-days-late is the percentage that arrive N or more days beyond the appropriate mean.

CHAPTER 5

ANALYSIS OF ORIGIN TO DESTINATION TRIP TIME DATA

A. Introduction

Chapters 1 and 3 explored the reasons that the level of rail service is in fact a problem for railroads and shippers while Chapter 4 chose a set of performance measures of the O-D trip time distribution to describe that level of service. In this chapter, data from 3 railroads are used to analyze:

1. the characteristics of the O-D trip time distribution
2. O-D unreliability
3. trip segment unreliability
4. mean O-D trip times, and
5. the relationship between unreliability and mean O-D trip times

The goal of these analyses is to explain the observed variations in O-D performance using a minimum amount of information about the O-D trip. Hopefully, railroads can use the results of this and similar studies to evaluate alternative means of improving rail freight service.

In the sections that follow, an attempt is first made to explain service as a function of general trip characteristics such as distance and the number of yards. Although a reasonable regression equation (R^2=.55) is obtained for mean trip times (Section 5E), the equation obtained in Section 5C for the 3-day-% is totally inadequate (R^2=.11). Hence it is necessary to look at more detailed information concerning the O-D trip to find the causes of unreliability. In Section 5D, the average contribution of each trip segment to O-D unreliability is studied to discover if either the origin, destination, or total line haul segment is the primary cause of unreliability. In general, both the reliability of each segment and its contribution to O-D unreliability vary extensively between O-D pairs. Furthermore, terminal reliability varies considerably even when only O-D pairs with the same terminal are considered. This suggests that the factors contributing to yard unreliability (see Chapter 2) must also be considered in order to explain O-D unreliability.

B. The Trip Time Distribution and Network Unreliability

Some general characteristics of trip time distributions became apparent during extensive analyses involving O-D data. For a single origin and destination, trip times measured in days typically form a uni-modal distribution skewed toward low values (Figures 5.1 and 5.2). Since extreme values, if any, are concentrated on the high side of all observed distributions, the mode and the median are nearly always lower than the mean trip time. In all of these distributions, each of these measures falls within the three day interval determined by the 3-day-%.

The 3-day-% and the related measures, the %-early and the %-late, divide moves between a single O-D pair into three distinct groups:

1) the 3-day-%: those moves arriving within the interval determined by the 3-day-% for that O-D pair
2) the %-early: those arriving before this interval
3) the %-late: those arriving after this interval

Every car movement will fall into one and only one of these groups for the appropriate O-D pair. Network performance can therefore be readily measured as the percentages of all moves which fall into each of these three groups. Typical measures are given in Table 5.1 for sets of O-D pairs from two railroads:

TABLE 5.1

NETWORK PERFORMANCE

Railroad	O-D Pairs	%-Early	3-day-%	%-Late	Moves
A	124	3	87	8	7000
B	129	4	81	13	16000

FIGURE 5.1: SAMPLE O-D DISTRIBUTIONS

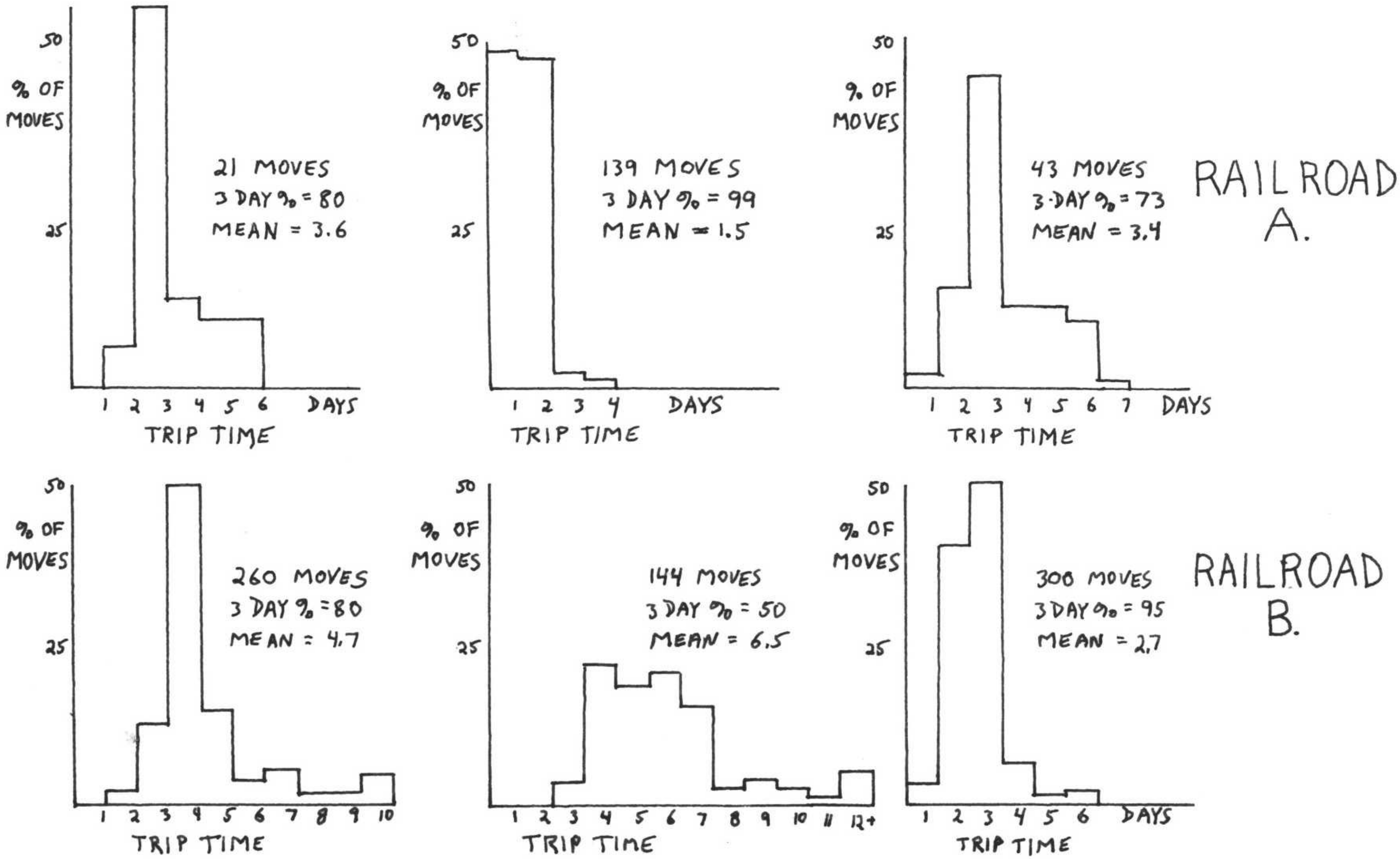
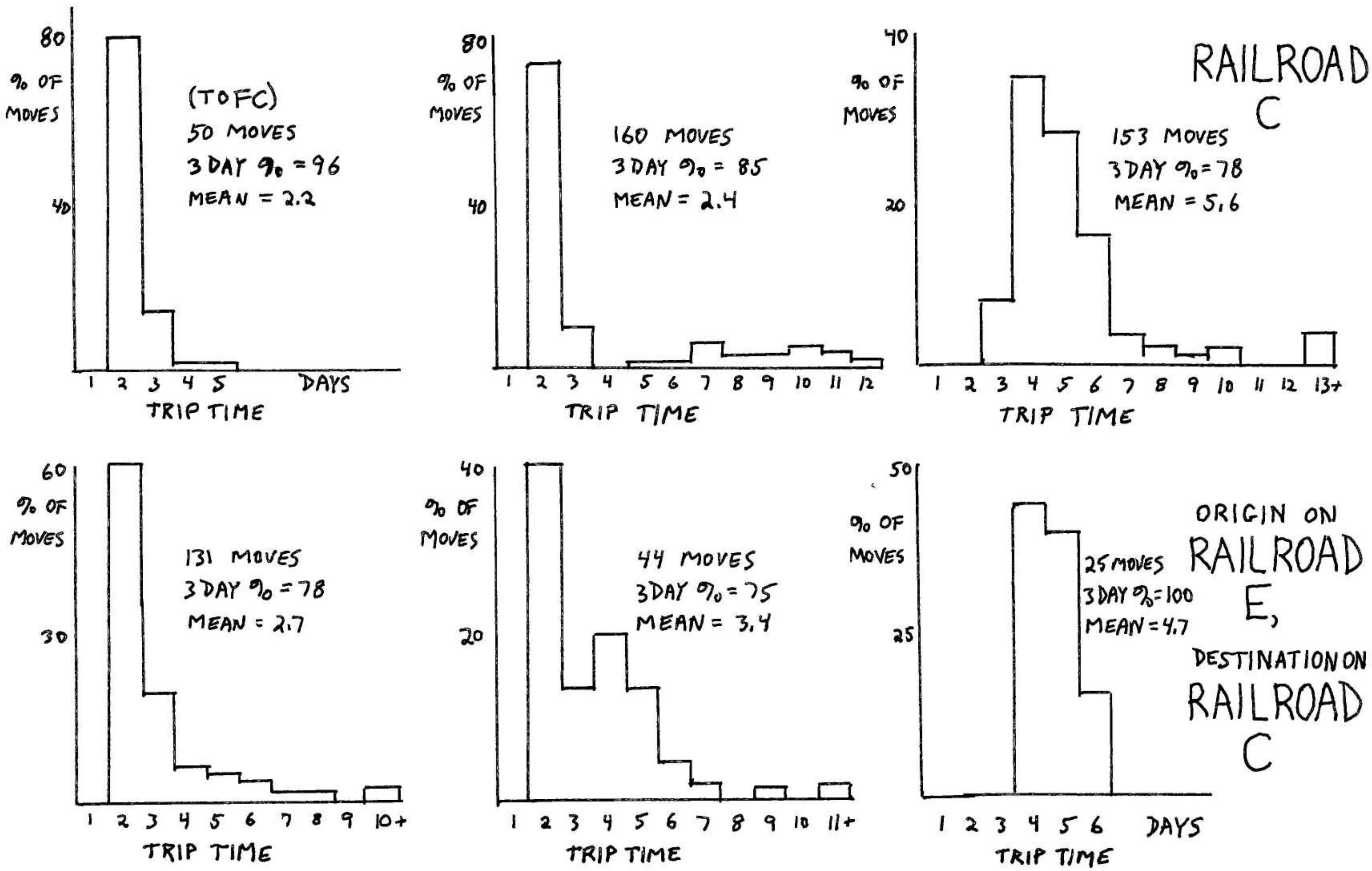


FIGURE 5.2: SAMPLE O-D DISTRIBUTIONS



Notice that the skewed nature of the O-D distributions is reflected in the large percentage of cars arriving late relative to the percentage arriving early.

Particular O-D pairs exhibit reliability much different from these network averages. As shown in Figures 5.3 and 5.4 the 3-day-% may be as low as 50 or as high as 100 and the %-late ranges from 0 to greater than 34. Clearly, a wide range of O-D unreliability exists; some O-D pairs receive very good service, while others receive relatively poor service. The remainder of this chapter will analyze the reasons for these observed fifferences.

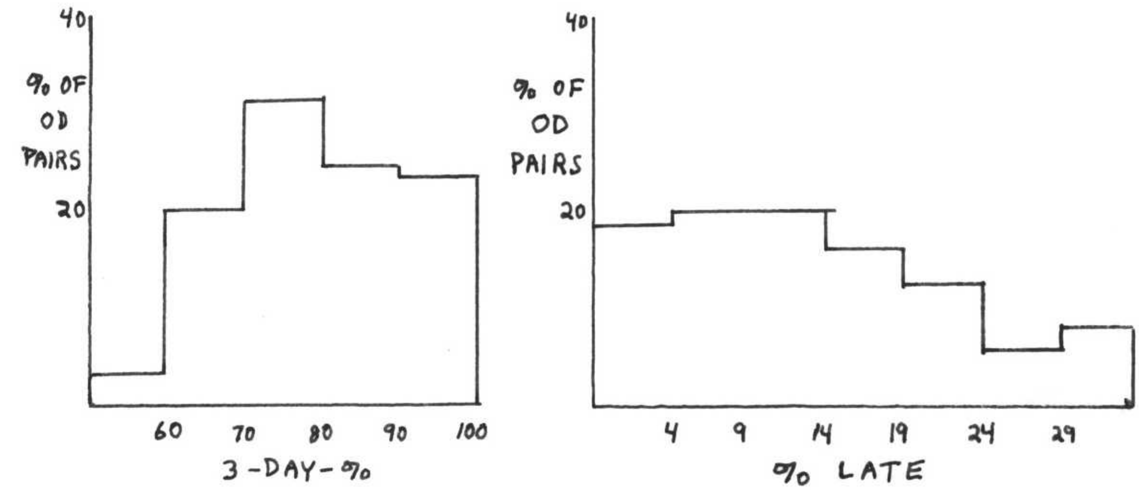
C. A Regression Model of Unreliability

General characteristics of the O-D trip such as mileage, the direction of travel, and the number of yards have been proposed as possible explanations of railroad unreliability. For instance, DeHayes developed a regression model of unreliability to be used by shippers to predict variability in trip times[1]. His model included the following variables: Northeast territory (0 or 1), north/south direction of travel (0 or 1), log (rail distance),

---

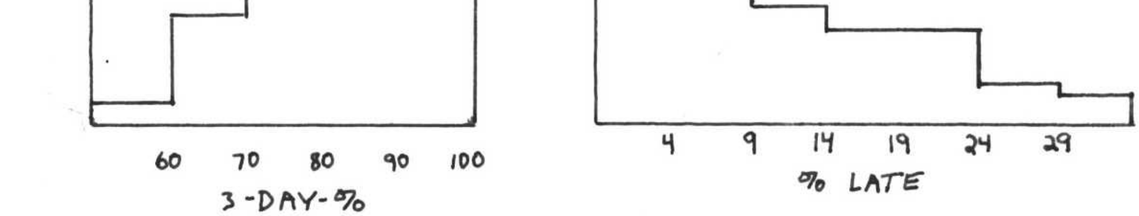
[1]DeHayes, "Industrial Transportation Planning: Estimating Transit Time for Rail Carload Shipments."

FIGURE 5.3



RAILROAD B, 129 O-D PAIRS

FIGURE 5.4



RAILROAD A, 124 O-D PAIRS

month, % grain shipments, same geographical territory (0 or 1), and the average number of carriers. The model differs from the one proposed below in several respects:

a. Variability is measured by the standard deviation of trip times

b. Only grain shipments were considered

c. A smaller number of moves and O-D pairs were used

d. Geographic and demographic variables were included

e. The Total O-D Trip Time was used rather than the O-D trip (see Chapter 2)

The regression equation had a multiple correlation coefficient of .41.

This thesis, however, is concerned with the operating factors that cause unreliability, not with factors that are related to unreliability. The first hypothesis that must be tested is that unreliability is caused by the general trip characteristics mentioned at the beginning of this section. In order to test this hypothesis, the following set of data was obtained from Railroad B:

1) The trip time distribution (from arrival at a specific origin yard to placement or delivery to interchange from a specific destination yard) for 129 O-D pairs having at least 30 loaded moves in the three week period covered by the data.

2) The number of hump, flat, and interchange yards where cars were handled for each of these O-D pairs

3) The rail distance between each origin and destination

4) The total number of loaded and empty moves for each O-D pair.

The 3-day-% was computed for each O-D pair and plotted against the numbers of each type of yard, the distance, the total number of moves and the percentage of loaded moves. These plots showed the relationship, if any, between each type of yard and unreliability to be linear, and that between the other variables and unreliability to be closer to log-linear. Since in addition, flat and hump yards appeared to have a similar relationship to the 3-day-%, they were combined into a single variable called "yards." Therefore the following model of unreliability was proposed[2]:

$$\text{3-day-\%} = b_0 + b_1(\text{yards}) + b_2(\text{interchanges}) + b_3\ln(\text{miles}) + b_4\ln(\text{moves}) + b_5\ln(\%\ \text{loads}) + \text{error}$$

When this equation was fitted to the data using a linear-least-squares regression technique, the number of interchanges and the natural logarithm of distance were the only variables signifi-

---

[2]Each of these independent variable is discussed in some detail in Section 5 of the M.I.T. Department of Civil Engineering Progress Report on the Study of Reliability in Railroad Network Operations.

cant at the .2 level. The resulting equation and the standard errors of the coefficients were

$$\text{3-day-\%} = 90.5 + \underset{(.5)}{5.4}\ (\text{interchanges}) - \underset{(1.3)}{1.7}\ \ln\ (\text{miles}) + \text{error}$$

where

Mean (error) = 0

Standard deviation (error) = 10.6

This model gives an unsatisfactory explanation for the observed differences in reliability; the multiple correlation coefficient, $R^2=.11$, indicates that only 11% of the sum of the squared variations from the mean are explained by the regression equation. An analysis of correlation coefficients offered evidence that the reasons for the low R^2 are the low correlations between the 3-day-% and the independent variables and the relatively high correlations between some of the independent variables.

This regression was undertaken to determine if the general characteristics of the O-D trip adequately account for the observed variations in O-D performance. Since the analysis showed that these characteristics have very little individual or joint impact on reliability, it will be necessary to look at the specific characteristics of each O-D trip to find the factors contributing to perceived levels of service. More detailed information concerning the individual trip segments is needed to explain the wide range of O-D reliability. Analysis of these segments in the next section will hopefully demonstrate why this regression failed and what additional information is needed.

D. Trip Segment Performance

1. The Relative Importance of the Trip Segments

A close look at segment performance will help show why the general characteristics of the O-D trip do not satisfactory explain unreliability. In Chapter 2, the O-D trip was divided into three major segments -- the origin yard time, the total line haul time including intermediate yard times, and the destination yard time. Railroad C provided these segment times for all loaded general purpose box car movements for a one month period in 1971. The 177 O-D pairs for which at least 40 moves were reported are analyzed below.

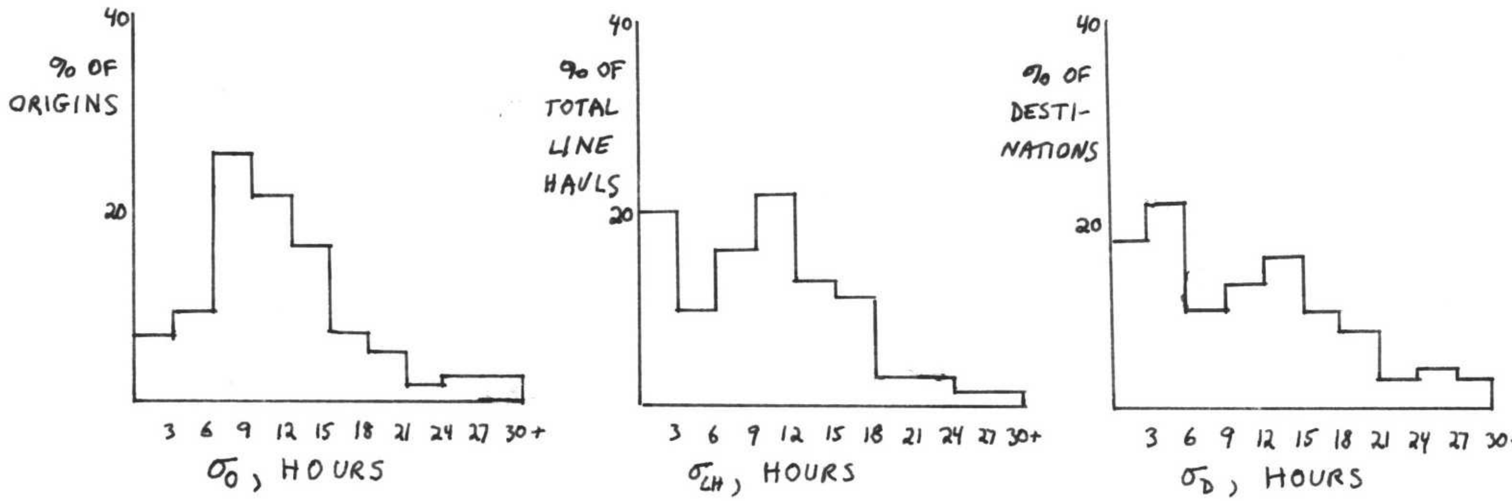
As shown in Figure 5.5, there is a wide range of reliability in the O-D trip segments. The standard deviations[3] for each of

---

[3]The highest 5% of segment times were eliminated prior to computing the standard deviation because these times are apt to be either unrelated to normal operations (i.e., due to in-transit privileges) or caused by data errors. In addition, low total line haul times were eliminated if they implied an average train speed greater than 50 mph or if they were 10 or more hours less than the next lowest line haul segment time.

FIGURE 5.5

STANDARD DEVIATIONS
OF SEGMENT TIMES
(RAILROAD C)



A. ORIGIN
177 O-D PAIRS

B. TOTAL LINE HAUL
157 O-D PAIRS

C. DESTINATION
146 O-D PAIRS

these range from a low of less than 3 hours to a high of more than 24 hours. The fact that yards had so little impact on the regression of 3-day-% is not surprising in light of this variability in terminal performance. An O-D pair with no intermediate yards but with an unreliable origin or destination could easily have a lower 3-day-% than an O-D pair with one or more intermediate yards. Since yard performance is itself so variable, the mere number of yards is not sufficient to explain O-D unreliability.

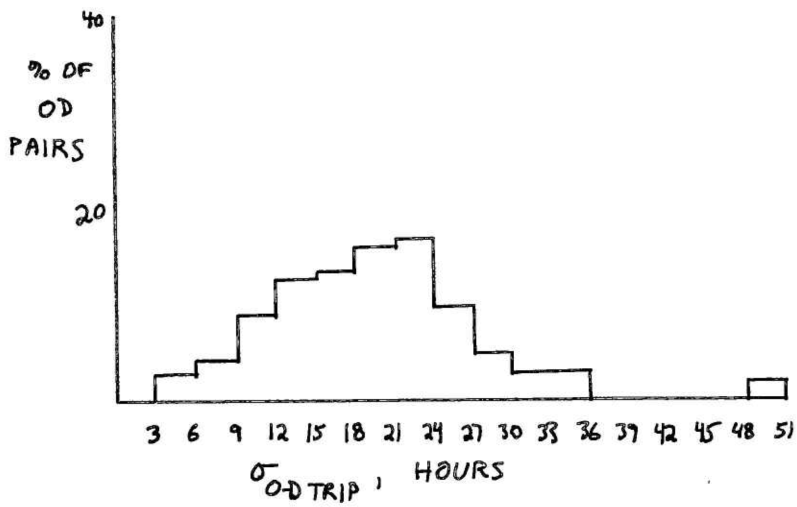
By assuming that trip segment times are independent[4], it is possible to obtain a first order approximation of the variance and the standard deviation of the O-D trip time (Figure 5.6). Although these are not generally useful performance measures as discussed in Chapter 4, they are computed here because the relative importance of each segment can be measured by its percentage contribution to the trip time variance. In Figure 5.7, this contribution has been calculated for each segment of the O-D trip for the 136 O-D pairs which include all three segments[5]. The overall patterns are remark-

---

[4]Preliminary analysis of a number of O-D pairs indicates that this is a reasonable assumption to make. In most cases, the sum of the segment variances is within 15% of the actual O-D trip variance.

[5]For some O-D pairs, either the origin or the destination yard time is under the control of an independent railroad company. In such cases, the measured trip starts or ends when the train hauling the car arrives on or leaves Railroad C's tracks and there is no terminal time reported. In other instances, Railroad C is itself performing the terminal function and only a single terminal time is reported.

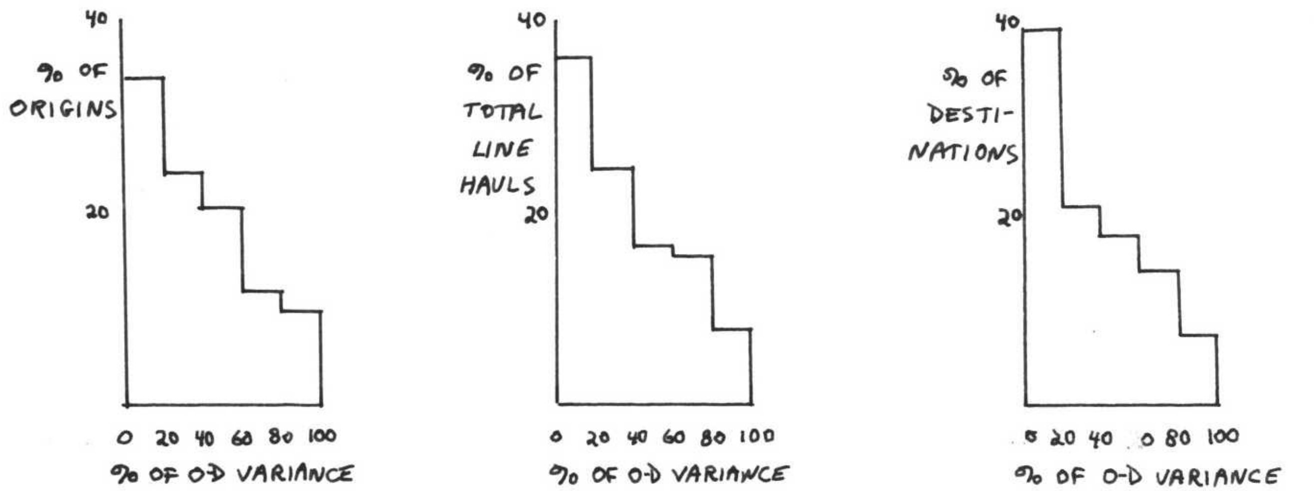
FIGURE 5.6



STANDARD DEVIATIONS OF O-D TRIPS

171 O-D PAIRS

FIGURE 5.7



A. ORIGIN B. TOTAL LINE HAUL C. DESTINATION

CONTRIBUTIONS OF SEGMENT VARIANCES TO TOTAL VARIANCES

133 O-D PAIRS

ably similar. No segment can be identified as the primary cause of unreliability and for a particular O-D pair any segment may account for either all or no part of the total variance.

Although the segments seem to have similar characteristics, distinguishing between reliable and unreliable trips -- where all trips with a variance greater than 350 are considered unreliable -- reveals a significant difference (Figure 5.8). In order to quantify this difference, the average contribution of each segment to the total variance was computed. The total contribution of each segment was determined by summing the variances of the segments represented in each distribution. Adding the three appropriate sums gave the total variance; the contribution of each segment was then readily quantifiable as its percentage contribution to this total (Table 5.2).

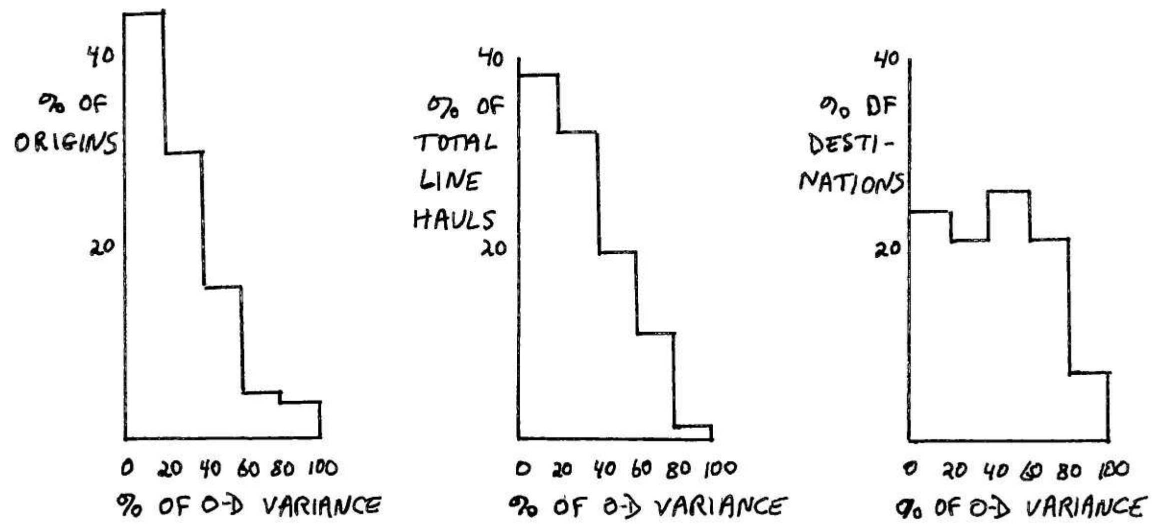
TABLE 5.2

	Number of O-D Pairs	Segment Variance as a % of O-D Variance: Origin	Total Line Haul	Destination
All O-D pairs	134	34%	33%	35%
Reliable O-D pairs	51	41	36	23
Unreliable O-D pairs	83	27	30	43

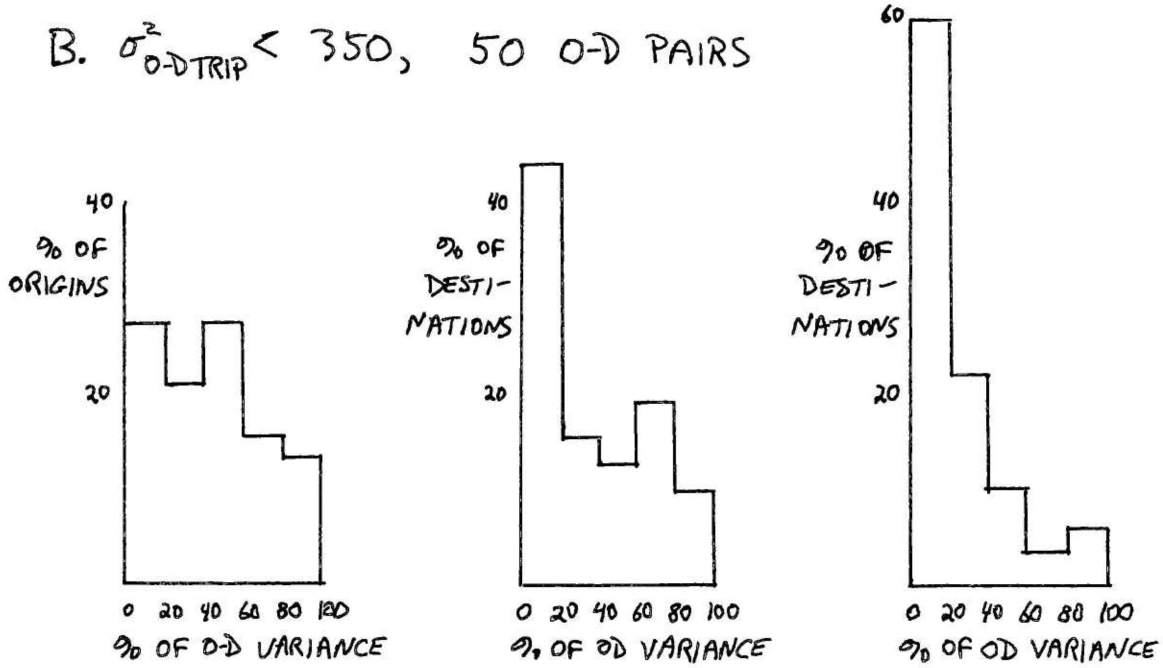
The destination accounts for 43% of the total variance in the average O-D pair with a trip time variance greater than 350 $hours^2$. The importance of destination unreliability is evident in Figure 5.8 even if it is not overwhelming. Whereas the destination variance accounts

FIGURE 5.8
CONTRIBUTIONS OF SEGMENT VARIANCES TO TOTAL VARIANCES

A. $\sigma^2_{O\text{-}D\ TRIP} \geq 350$, 83 O-D PAIRS



B. $\sigma^2_{O\text{-}D\ TRIP} < 350$, 50 O-D PAIRS



for more than 40% of the trip variance for over half of the unreliable trips, it accounts for that much of the total in only a fifth of the unreliable O-D pairs. On the other hand, the origin and line haul variances each account for more than 40% of the total variance in less than a third of the O-D pairs. Thus, variability in the time spent at the final yard is the largest single contributor to O-D unreliability. This is important because this time reflects system performance much less than it reflects local yard policy and performance. The origin yard time and the total line haul time are more directly affected by train performance and schedules. Although some of the destination variance may reflect delays in putting cars out to constructive placement due to either weekends or customer requirements, the extent of such delays is limited by ICC regulation and is hypothesized to be overshadowed by the variance due to infrequent local deliveries. In any case, this analysis implies that improvements in local movements at the destination yard may have as significant an impact on measured O-D performance as improvements in network reliability. Additional research will be needed on measured destination yard performance to determine how much it actually reflects unreliability in putting cars out to placement or delivering them to interchange.

2. Terminal Performance

Figure 5.5 showed wide ranges in the standard deviations of origin and destination yard times. One reason for this may be that interchanges may perform differently from other terminals. Since they are equivalent to an intermediate yard under the control of two railroads, interchanges may well be more reliable than the average terminal. The origin and destination distributions in Figure 5.5 were separated into interchange/non-interchange categories[6] to test this hypothesis. As shown in Figure 5.9, the standard deviations of the interchange yard times are on the whole lower than those for the non-interchange yards. Although there is considerable overlap in these distributions, the vast majority of origins and destiations with standard deviations of more than 15 hours are not interchanges.

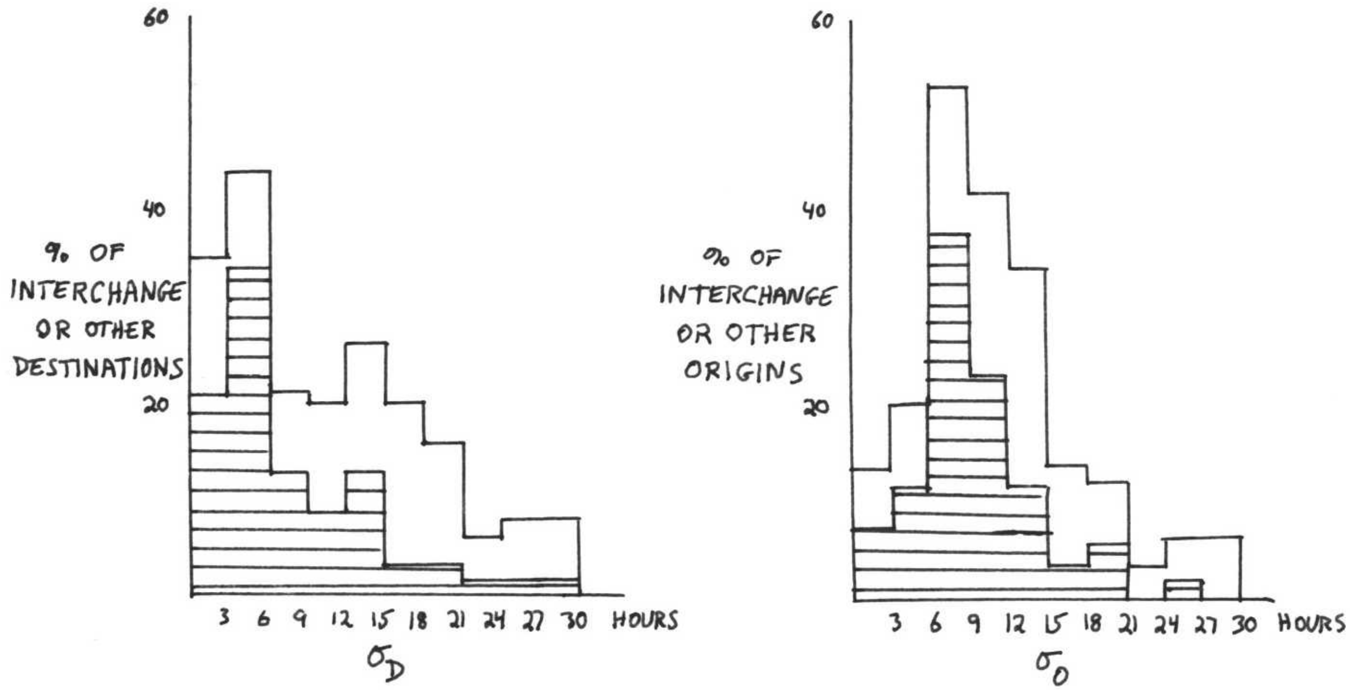
TABLE 5.3

Group	Number	Yard Time Standard Deviation > 15 hrs. Number	% of Group Total
Interchange Origins	95	10	10
Other Origins	82	24	30
Interchange Destinations	68	7	10
Other Destinations	78	34	46

[6] The data used in this section identifies cars recived from or delivered to interchanges. Although the O-D pair as defined in Chapter 2 aggregates local and interchange traffic, many O-D pairs in fact have primarily only local, overhead, received, or forwarded traffic. In all but a few cases, at least 80% of the moves through terminals are local or interchange. Thus it is usually quite possible to specify an origin or destination as either an interchange or a local yard.

FIGURE 5.9

COMPARISON OF RELIABILITY FOR INTERCHANGES ▤ AND OTHER TERMINALS ☐



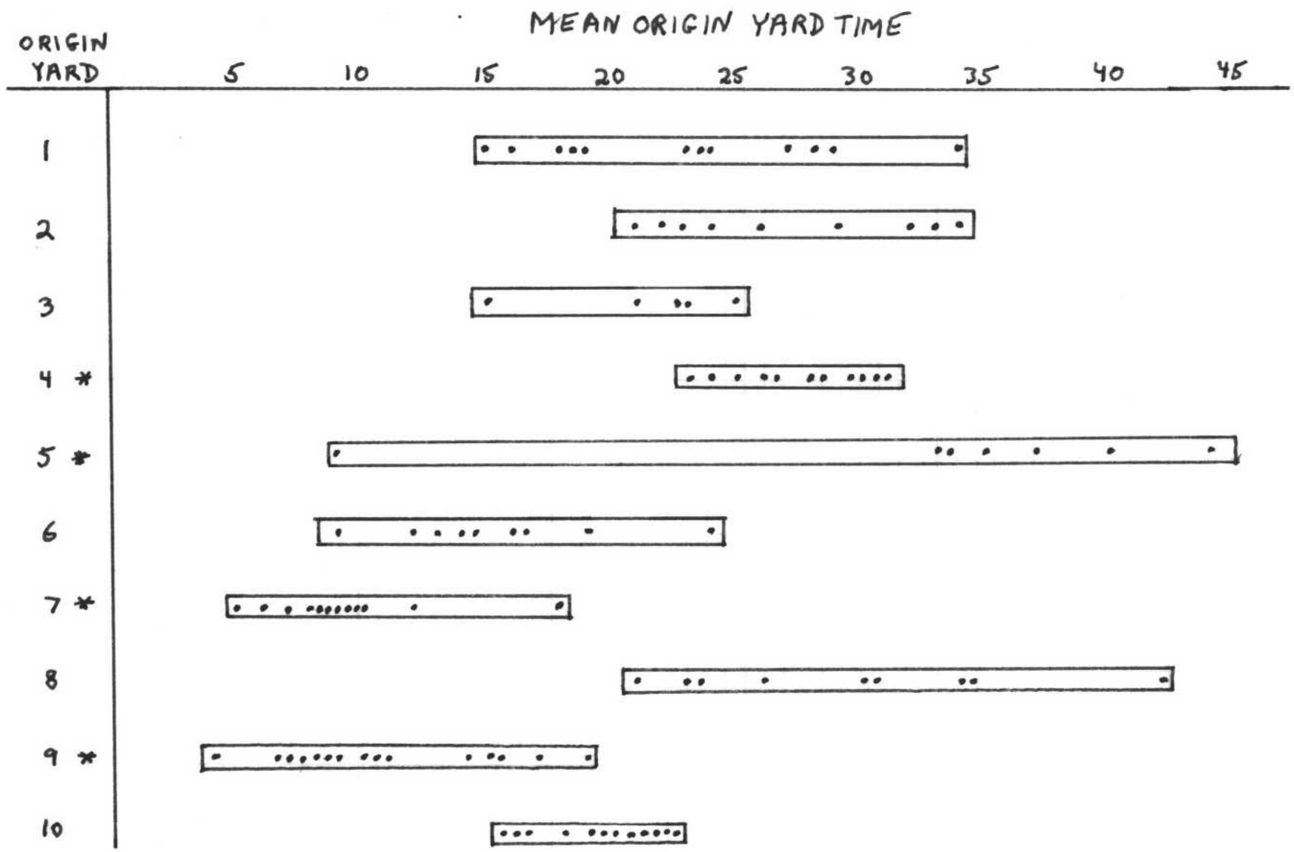
This result agrees with the regression analysis which showed that O-D pairs are more reliable if the terminals are interchanges.

Reid, O'Doherty, and Jennings have done extensive research on the causes of unreliability in individual yards and found that erratic train performance, schedules, and congestion are all impor-tnat[7]. This thesis will not add to their work other than by noting that variations in reliability are substantial between, as well as within, yards. That is, the ranges of the mean and standard deviations of terminal yard times, for O-D pairs with the same terminal, are smaller than the overall ranges of terminal means and standard deviations. In Figures 5.10 through 5.13, each slash represents a single O-D pair; slashes for O-D pairs with the same origin (5.10 and 5.11) or the same destination (5.12 and 5.13) are enclosed in rectangles. The fact that many of these rectangles do not even overlap means that total yard performance varies greatly from yard to yard. However, the fact that there is also a great deal of variation within many of the rectangles means that the overall performance of terminals will not always give a good indication of terminal performance for a particular O-D pair. The work on yards mentioned above is essential to a deeper understanding of O-D performance

---

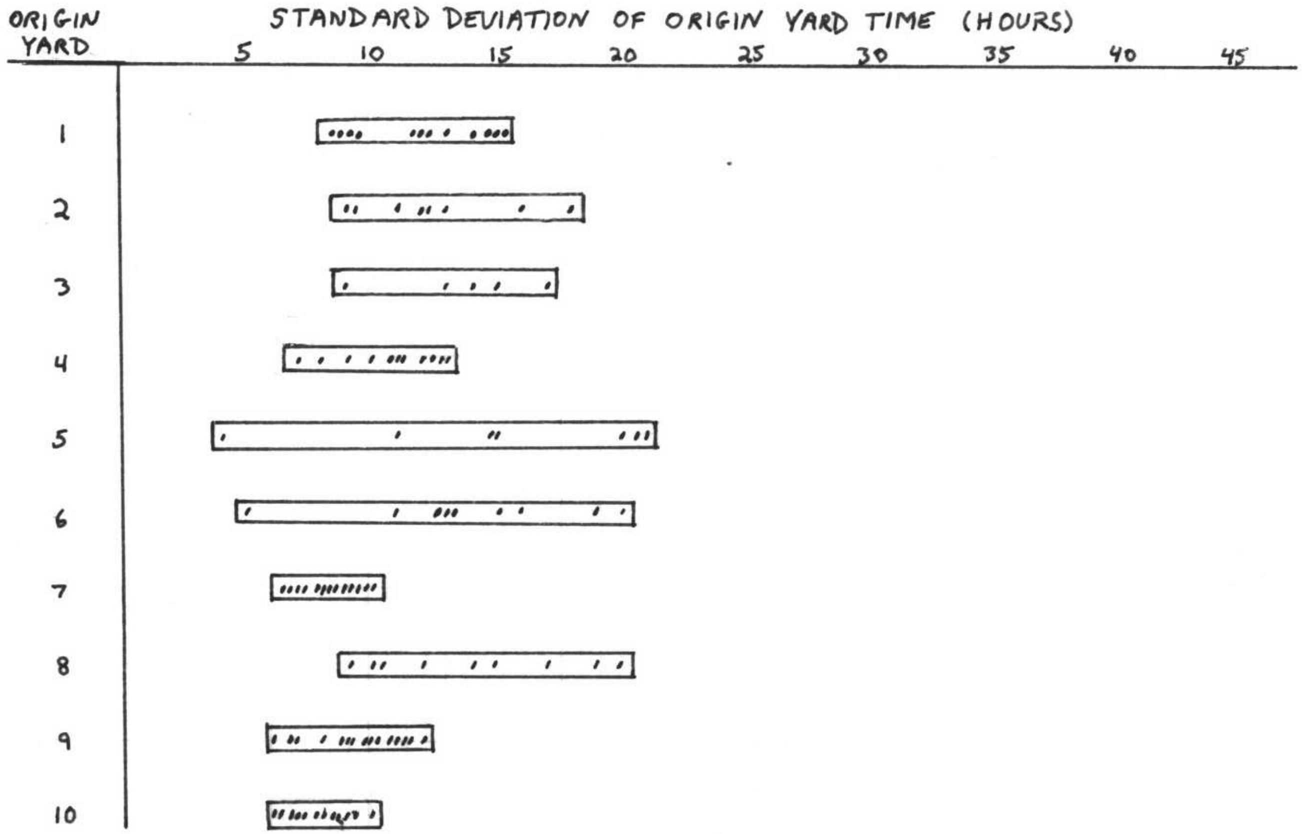
[7] Reid, Yard Unreliability in Rail Freight Movement; O'Doherty, Classification Yard effects in Rail Freight Movement Reliability; Jennings, The Effects of Train Length on the Reliability of Operations of Railroad Yards.

FIGURE 5.10



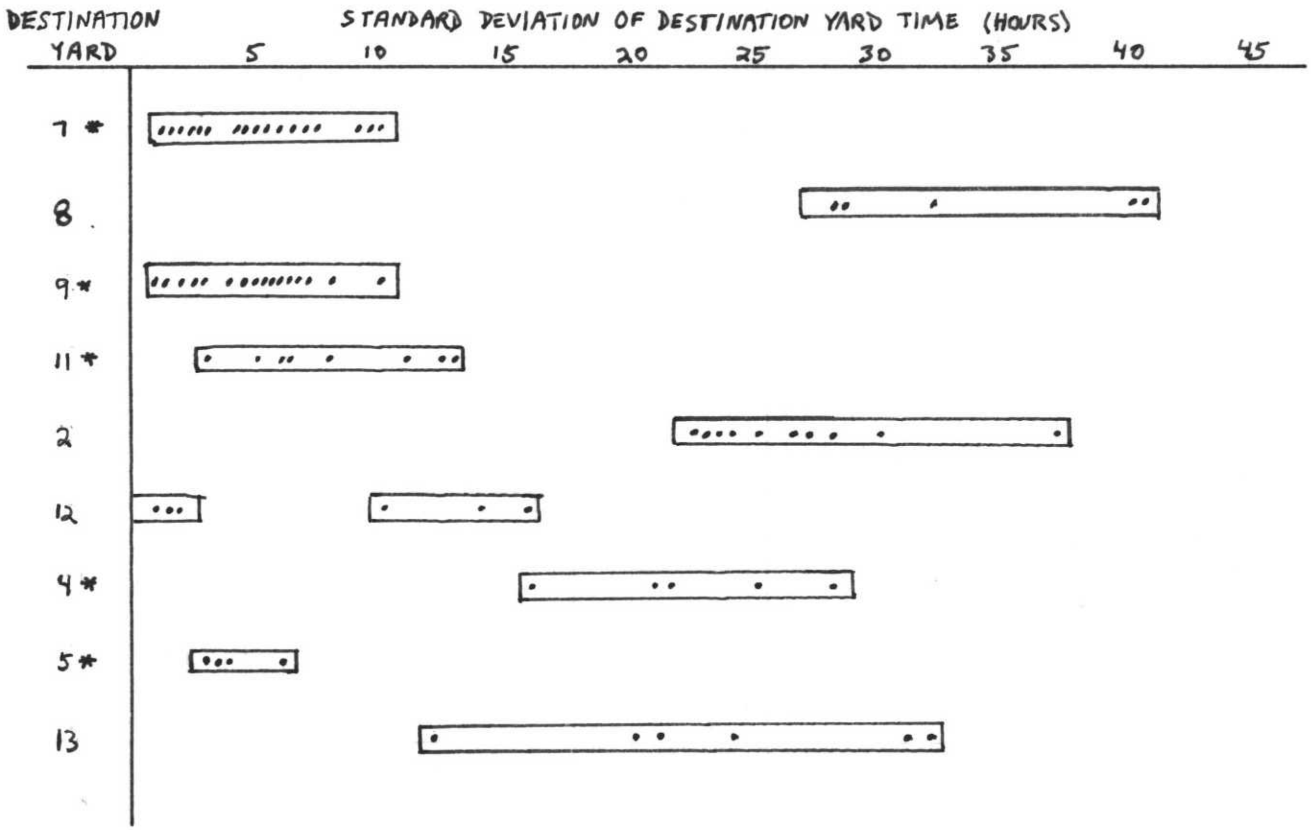
* DENOTES INTERCHANGE

FIGURE 5.11



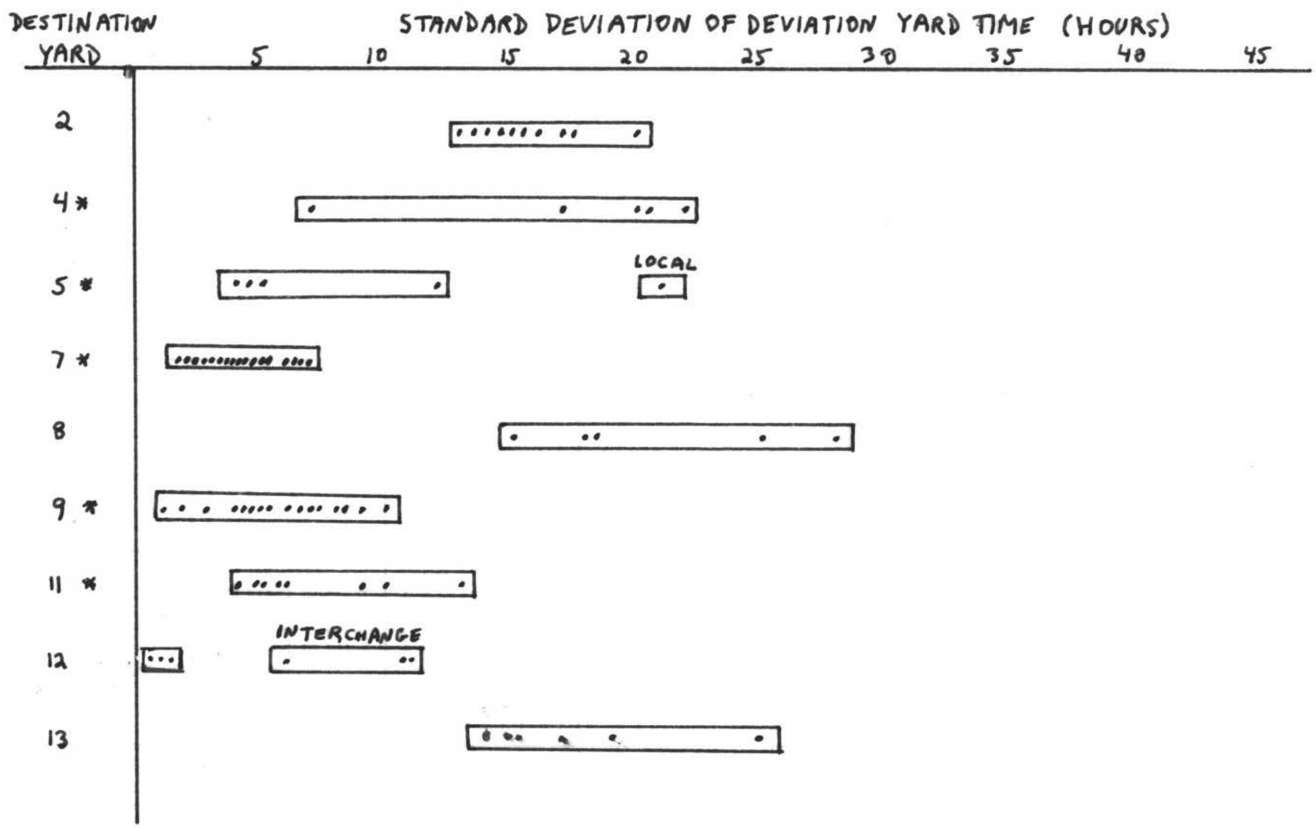
* DENOTES INTERCHANGE

FIGURE 5.12



* DENOTES INTERCHANGE

FIGURE 5.13



* DENOTES INTERCHANGE

3. Total Line Haul Performance*

The natural logarithm of distance was one of two significant independent variables in the regression equation of the 3-day-% that was described in Section 5C. In that equation, the distance term is -4.6% at 100 miles and -11.7% at 1000 miles. The difference between these opposite extremes is only 4%, less than the effect of a single interchange on the 3-day-%. One possible explanation of this small impact might be that the variations in terminal performance noted in Figure 5.5 obscure the effect of distance on O-D performance. If so, mileage should have a greater impact on the total line haul segment which includes only intermediate yards. Using the same set of data from Railroad C, the standard deviations of the total line haul segments were plotted against the percentage of O-D pairs for 3 mileage groups (Figure 5.14). Clearly the unreliability increases with distance:

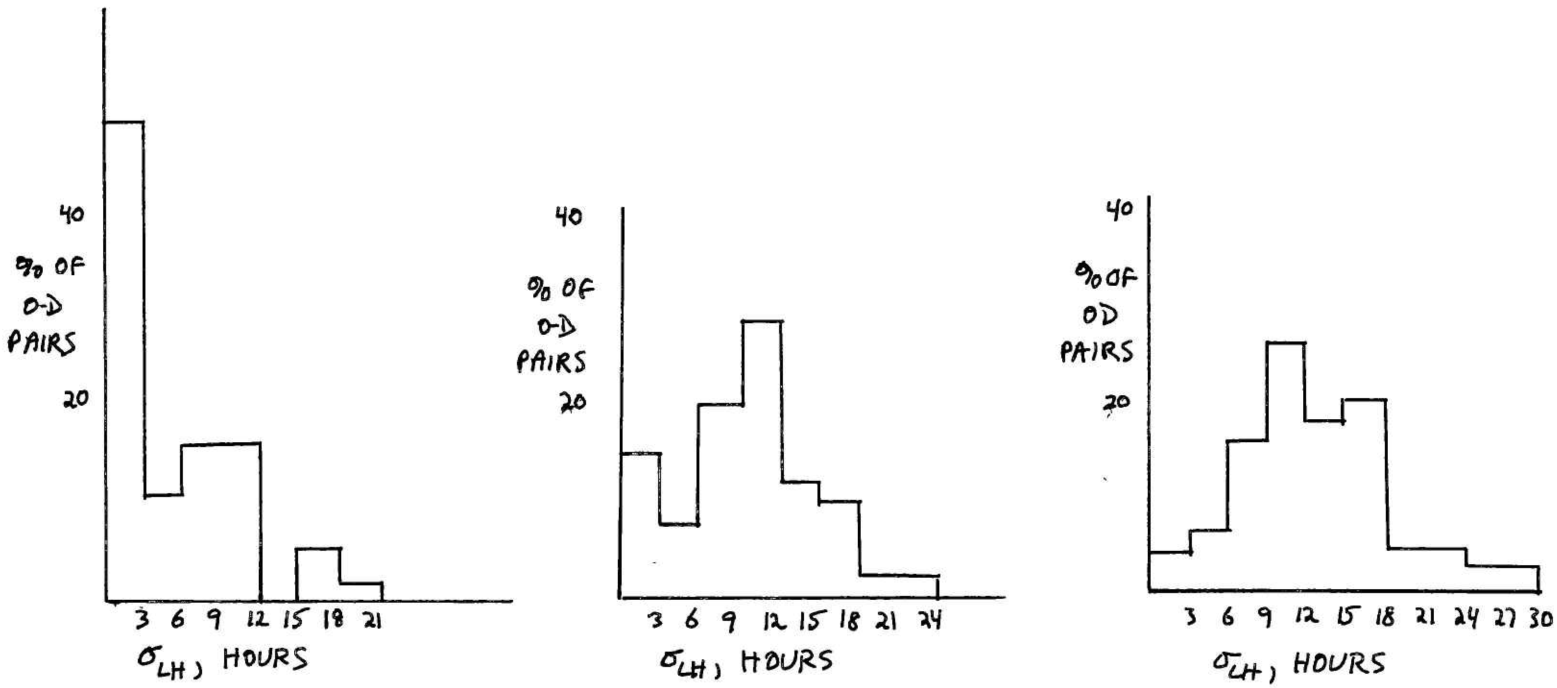
TABLE 5.4

	Distance		
	0-300	300-600	600-1300
Number of O-D pairs	37	60	47
Percentage of pairs with $\sigma_{lh} \leq 9$	76%	43%	25%
Median σ_{lh}	3 hrs.	10 hrs.	12 hrs.

---

*The total line haul segment includes intermediate yards.

FIGURE 5.14



0-300 MILES   3-600 MILES   6-1300 MILES

However, these differences are related to distance rather than caused by distance. As dicussed in Chapter 2, the yard time needed to switch a car from one train to another may well be highly variable, while Belovarac has found standard deviations of the total line haul time to be under 5 hours in cases where that segment reflects time spent as part of a single train[8]. Therefore, the number of times a car is switched at intermediate yards is likely to be a much more important cause of total line haul unreliability than the total distance that it travels. To test this hypothesis, the total line haul data from Railroad C was partitioned by the number of intermediate yards in the O-D trip. As shown in Figure 5.15 and in Table 5.5, the total line haul standard deviation increases markedly with the number of intermediate yards.

TABLE 5.5

UNRELIABILITY INCREASES WITH THE NUMBER OF INTERMEDIATE YARDS

Number of Intermediate Yards	Number of O-D pairs	Median σ_{1h}	% O-D Pairs With $\sigma_{1h} \leq 9$
0	31	2 hours	100
1	72	10 hours	42
2 or more	41	16	7

[8]Belovarac, Determinants of Railroad Line Haul Unreliability.

Again, however, the range in unreliability for trips with the same number of intermediate yards is substantial. In Figure 5.15, a third of the O-D pairs with 2 or more intermediate yards have total line haul standard deviations less than 12 hours, while more than a quarter of those with only 1 intermediate yard have standard deviations greater than 12 hours. The work done by the authors cited earlier delves into the causes for these differences in intermediate yard reliability.

In order to demonstrate the predominance of yard work over distance, reliability graphs are given in Figure 5.16 for O-D pairs with the same number of intermediate yards and either short (0-600 miles) or long (600-1300) mileage. Unreliability is clearly more sensitive to increases in the number of intermediate yards (left to right) than to increases in distance (top to bottom).

E. A Regression Model of Mean O-D Trip Times

1. The Regression Model

Mean trip times were analyzed because they are a second service characteristic important in modal choice decisions. As was the case with the 3-day-% and the %-late, large differences were noted in mean O-D trip times. In Figure 5.17, for instance, O-D pairs

FIGURE 5.15

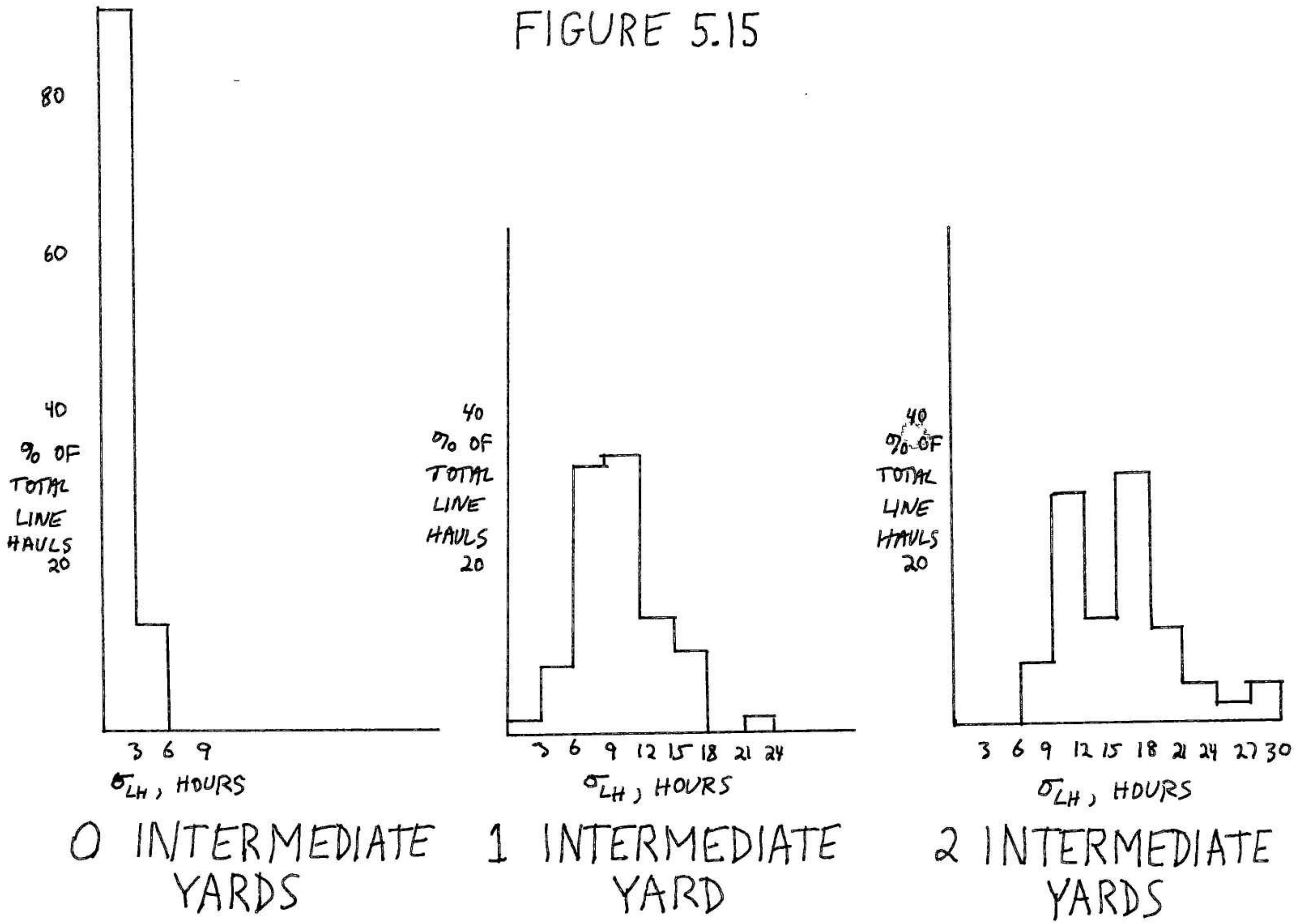

80
60
40
% OF TOTAL LINE HAULS
20
3 6 9
σLH, HOURS
0 INTERMEDIATE YARDS
40
% OF TOTAL LINE HAULS
20
3 6 9 12 15 18 21 24
σLH, HOURS
1 INTERMEDIATE YARD
40
% OF TOTAL LINE HAULS
20
3 6 9 12 15 18 21 24 27 30
σLH, HOURS
2 INTERMEDIATE YARDS

FIGURE 5.16: THE PREDOMINANCE OF THE NUMBER OF INTEMEDIATE YARDS OVER DISTANCE AS A FACTOR AFFECTING TOTAL LINE HAUL UNRELIABILITY

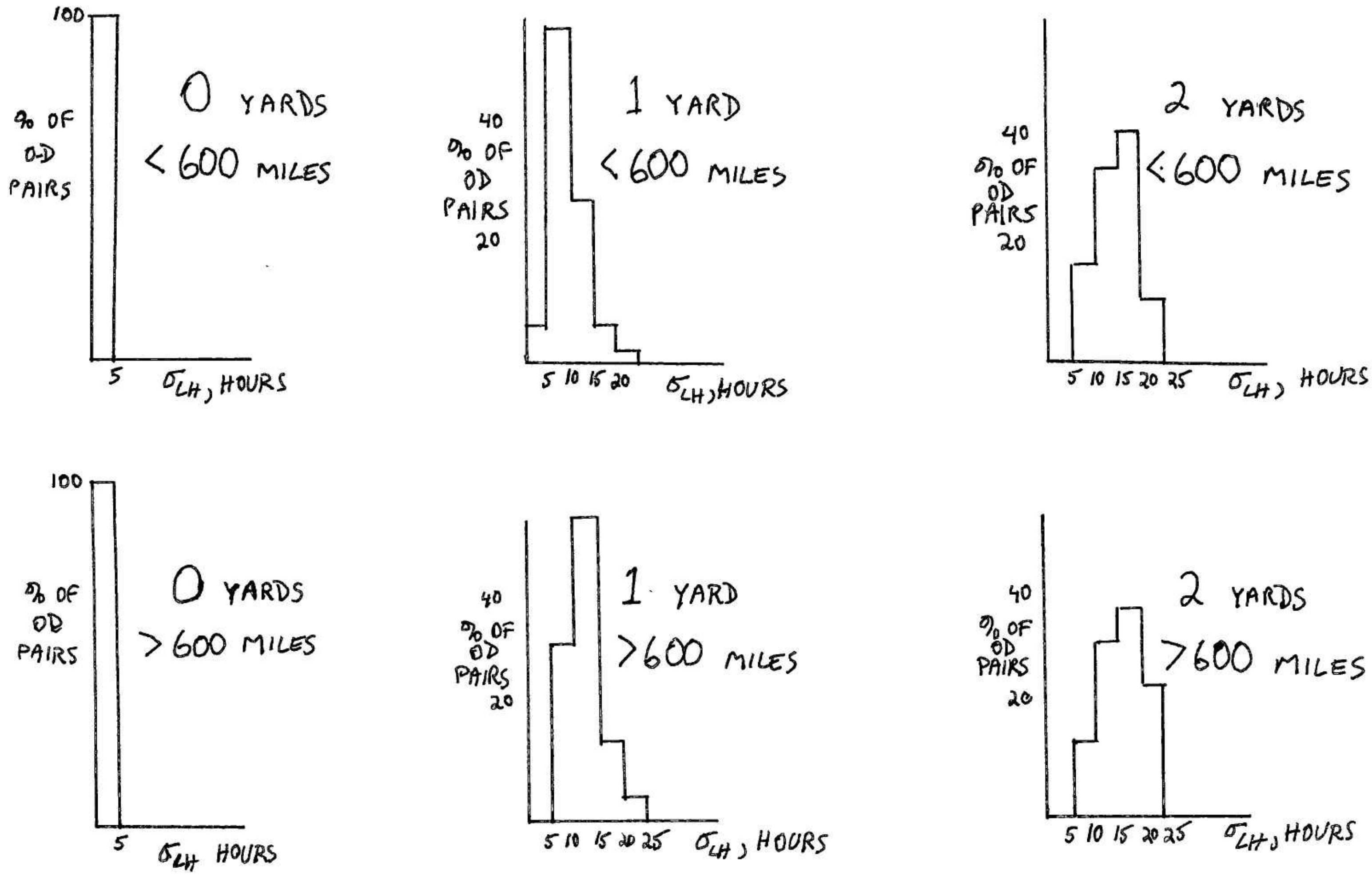
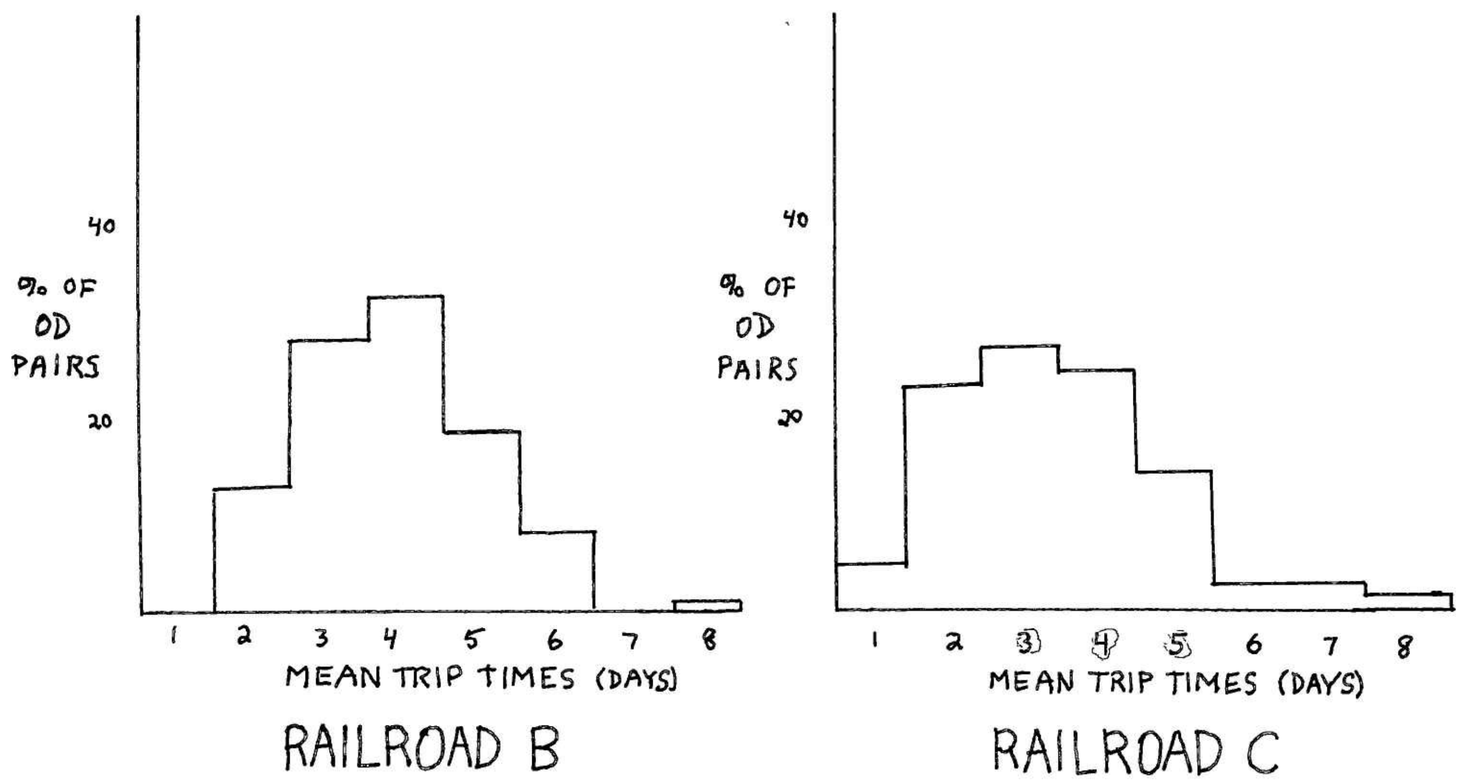


FIGURE 5.17


40
% OF
OD
PAIRS
20
1 2 3 4 5 6 7 8
MEAN TRIP TIMES (DAYS)
RAILROAD B
40
% OF
OD
PAIRS
20
1 2 3 4 5 6 7 8
MEAN TRIP TIMES (DAYS)
RAILROAD C

on Railroad C have means between 2 and 5 days. Many analysts assume these mean O-D trip times are a function only of distance. Meyer, et al, used the following model in their analysis of the determinants of modal splits[9].

Mean trip time = Distance/average train speed + 6 hours/interchange + 8 hours/intermediate yard + 48 hours for pick up/delivery

The model assumes interchanges occur every 228 miles and intermediate yards every 140 miles. Thus the model takes no account of either reliability or of the possibility for bypassing intermediate yards. In fact, it gives mean trip times as a pure function of distance.

Work by Benthe gives some evidence that this model is inadequate[10]. He attempted to find a relationship

$$\text{Mean trip time} = b_0 + b_1 \text{ distance}$$

for grain moves over each of 11 different railroads. With sample sizes varying from 100 to 1500, he was unable to obtain any R^2 greater than .19 and many were below .05.

Thus distance alone does not explain the variability in trip

---

[9] Meyer, Peck, Stenson, and Zwick in the Economics of Competition in the Transportation Industries (pp. 192-3).

[10] Benthe, Freight Transport Mode Choice: An Application to Corn Transportation.

times. DeHayes, using a model similar to the one described in Section C of this chapter, was able to obtain an R^2 of .72[11]. However, this model used demographic and geographic data which are unrelated to rail operations.

This thesis tests the hypothesis that mean trip times are a function of general O-D trip characteristics, such as mileage and yards. The following basic model of the mean O-D trip time was proposed:

$$\text{Trip Time} = b_0 + b_1\text{Distance} + b_2\text{Yards} + b_3\text{Reliability} + \text{error}$$

The mean time is composed ot the mean time needed for train movements, the mean terminal yard times, and the mean intermediate yard times. Since hump, flat, and interchange yards perform distinct functions, they were included as separate variables in the regression; thus there are three b_2 coefficients which reflect the average time spent in each type of yard. The b_1 coefficient should approximate the inverse of the average train speed, while b_0 should in some sense reflect the minimum time necessary to originate and terminate a car. These variables alone, however, will not account for the observed differences in mean trip times. The analysis in the previous section revealed a great deal of variation in the perfor-

---

[11]DeHayes, op. cit.

mance levels of individual yards that was quite unexplainable without detailed yard analyses. In particular, the regression of 3-day-% showed that the variables specified in this model do not satisfactorily explain unreliability, while independent analysis showed a clear relationship between 3-day-% and mean trip times (Figure 5.18). In the situations where cars must suffer a 12 or 24 hour delay if they miss a connection, it is clear that the more missed connections, the greater the mean trip time will be. Figure 5.19 illustrates why this is so. The bottom horizontal line is a time axis for the origin yard of an O-D trip; the other horizontal lines are time axes for each of the other yards in the trip. Cars arrive in the origin yard at time A. Each set of vertical lines represents a train scheduled between two yards*; the time between the vertical lines as measured on the yard time axis is the interval between successive trains. Hence, all broken lines starting at A and proceeding upward or to the right and ending on the topmost time axis represent feasible O-D trips. The O-D trip time is given by the length of the path taken. The path AB highlighted in the figure is the shortest O-D trip possible because there are no missed connections; it includes 32 hours of scheduled yard times

---

*Assume that the height of the line gives the time necessary for the train movement between the two yards plus the classification of the car at the second yard.

FIGURE 5.18

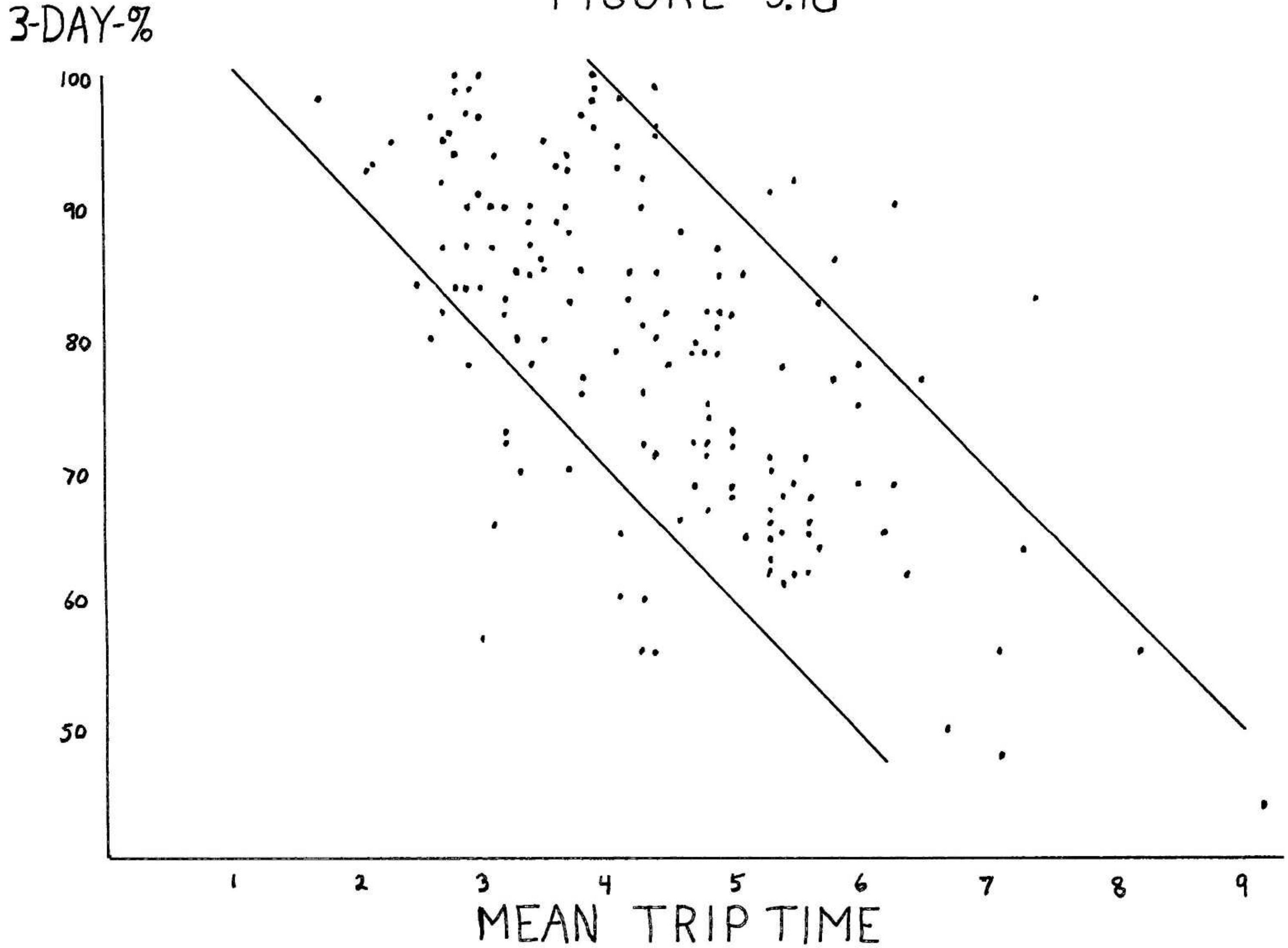
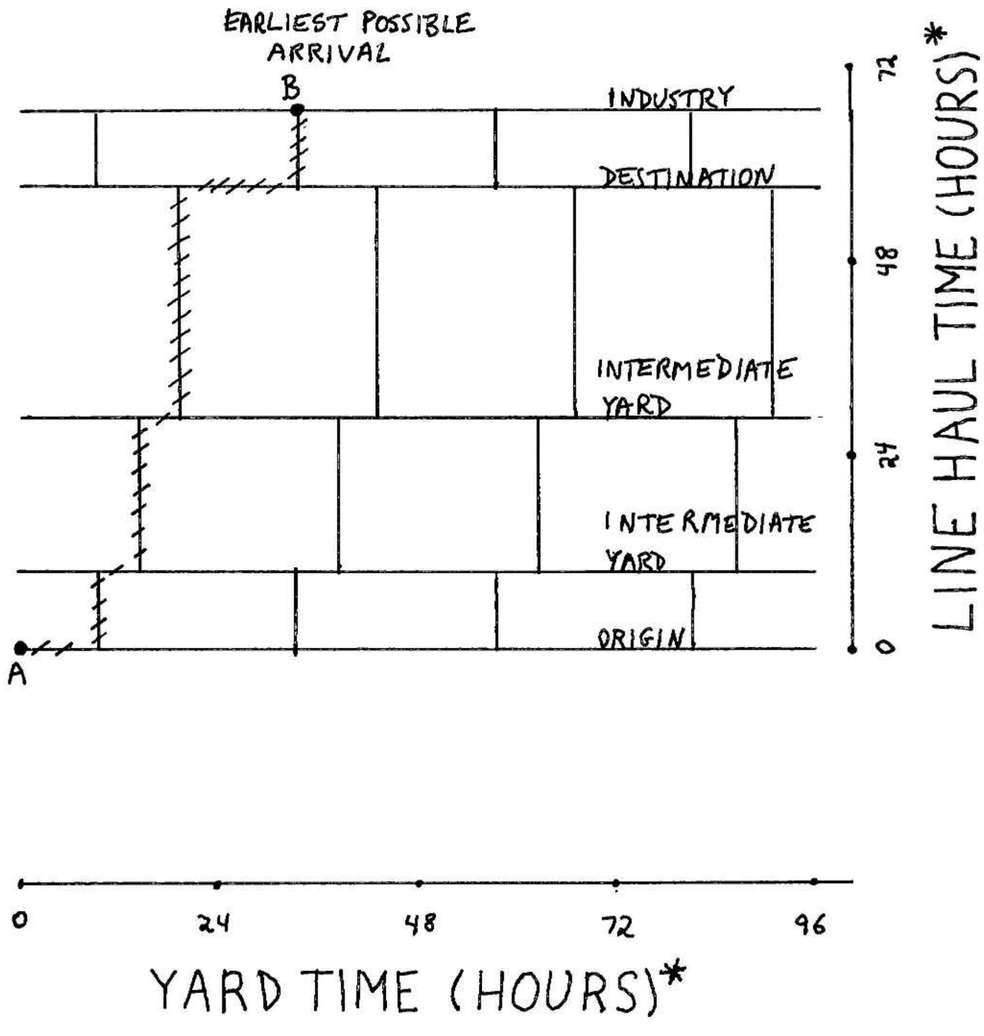


FIGURE 5.19



* LINE HAUL TIME INCLUDES THE NORMAL PROCESSING TIME NECESSARY IN YARDS

and 68 hours of scheduled line haul and processing time. Since in this case the time intervals between trains are always 24 hours and cars originate at time A each day, the O-D trip time will always equal 100 + 24 N where N is the number of missed connections. Therefore, the mean trip time will be AB + 24 E(N) where E(N) is the expected number of missed connections per O-D trip. Thus it is appropriate to use a measure proportional to the expected number of missed connections as an independent variable in this regression analysis of mean trip times. The 3-day-% is used in this study because it measures the compactness of the O-D trip time distribution which is itself a function of the probability of missed connections.

2. The Regression Equation

The model described above was fitted to the same set of data used for the regression analysis of unreliability. For 129 O-D pairs on railroad B, the following equation was obtained:

$$\text{Mean} = 1.2 + .0007M + .72H + .63F + .39I + .45U + E$$

where

			Standard Error	T value
M	=	distance/100	.00036	2.0
H	=	number of hump yards	.15	4.9
F	=	number of flat yards	.12	5.3

		Standard Error	T Value
I	= number of interchanges	.17	2.3
U	= (100 - (3-day-%))/10	.07	6.6
E	= error	.74	

F-Value = 30.3

The multiple correlation coefficient for the equation was .55 (with 124 degrees of freedom). Notice that the coefficients fit the model described above:

b_1 = .07 days/100 miles b_1^{-1} = 59 mph

b_2 = 17 hours/hump yard

b_3 = 15 hours/flat yard

b_4 = 9 hours/interchange

b_5 = 1 hour/unit reduction in the 3-day-%

This regression equation illustrates the important impact of reliability on mean transit time and hence on car utilization and shippers' modal choice decisions. According to this model, the mean trip time for an O-D pair would be reduced .5 days by an increase of 12% in the 3-day-%. The implication of this type of impact and of the other results of this chapter will be studied in Chapter 6. Notice that this regression indicates that there is not necessarily a tradeoff between mean trip times and reliability.

F. Summary and Conclusions

The typical O-D trip time distribution measured in days is unimodal and skewed toward low values. Individual distributions exhibit wide variations in performance using the performance measures developed in Chapter 4. The variation in the 3-day-% measure is not explained by a regression on general O-D trip characteristics such as rail distance, the number of yards, and the traffic volume. Analysis of trip segment reliability shows that the regression was inadequate because there are also great variations in origin, total line haul, and destination performance. Although the destination is the largest contributor to the variance of the O-D trip, any segment may be the most unreliable in a particular instance.

Analysis was presented which accounts for some, but not all, of this variation in segment performance: intermediate yards are much more important than distance in determining the total line haul variance; interchange movements are more reliable than local movements; and most single yards exhibit ranges in means and standard deviations of segment times much smaller than the overall ranges of segment performance. Work is cited which shows that erratic train performance, congestion, and yard policies are major determinants of segment performance.

The general characteristics of the O-D trip account for most of the variations in mean O-D trip times. Since reliability was not explained by these characteristics, the 3-day-% was used as an independent variable in the regression analysis of mean trip times. This relationship between unreliability and mean trip times may in fact be the most important impact of unreliability.

CHAPTER 6

IMPROVING RAILROAD RELIABILITY

A. Introduction

This thesis suggested in Chapter 1 that improving railroad reliability could be an important means of reversing the continued decline in railroad profitability. Chapter 3 identifed those aspects of service which are important to shippers and are therefore reflected in the performance measures selected in Chapter 4. Chapter 5 analyzed actual railroad data using these measures in order to understand the level of service currently provided by railroads and to explain the reasons for the observed differences in service provided individual O-D pairs. Several important results were obtained:

1. The level of service as measured by the mean trip time, the 3-day-%, and the %-late varies greatly between O-D pairs;
2. the variations in the 3-day-% are not explained by the general characteristics of the O-D trip because
3. the reliability of the individual trip segments also varies greatly between O-D pairs; and

4. unreliability increases mean trip times beyond the minimum feasible trip time for an O-D pair.

These results make it possible to clarify what is or is not meant by the problem of rail freight service unreliability. The fact that many O-D pairs have a 3-day-% greater than 90, a mean of less than 3 days, and/or a %-late of 0 means that fast, reliable service is not technologically impossible. The fact that the only impact of distance on unreliability is due to its correlation with the number of times that a car is switched and the fact that its impact on mean trip times is less than the combined impact of yards and unreliability means that faster trains are not a prerequisite to improved service. The problems lie in freight yards and specifically in the process of switching a car from one train to another. The time required for this process is highly variable at many yards. As a result, cars are often delayed beyond what might seem a minimum feasible trip time based on past experience and/or freight schedules.

The problem of unreliability is encompassed in the following questions:

1. What is the nature of O-D unreliability?
2. What are its impacts?
3. What are its causes?
4. How can it be eliminated?
5. What are the potential costs and benefits?

The performance measures selected in this thesis can be used to answer the first question as is graphically illustrated in Figures 5.3 and 5.17 of the previous chapter. Many O-D pairs receive poor service as measured by any standards.

Chapter 3 discussed the range of impacts that might pertain in specific situations. Reliable transportation may or may not be important to a shipper, depending on its relation to the rest of his logistics system. The level of rail service received by a given O-D pair must always be evaluated with respect to the needs of all firms who transport or could transport goods between that origin and destination.

The causes of unreliability are functions of the specific characteristics of the O-D trip. Congestion, low priority, train cancellations, erratic train performance, inappropriate schedules, traffic volume fluctuations, or any of a number of other causes could be a primary reason for unreliability in the yards involved in a particular O-D trip. This implies that there can be no general prescription for eliminating unreliability. Section B of this Chapter lists some of the means that might be appropriate in particular situations.

The costs of improving rail service depend on which of these means that are used and will not be discussed in this thesis. Although the benefits likewise depend on the specific instance, this thesis will estimate the magnitude of nationwide benefits resulting from an overall improvement in rail service. Section C of this chapter will demonstrate that these are sufficient to justify further research efforts and considerable railroad investment.

B. Alternative Means of Improving Railroad Reliability

A railroad can approach the problem of reliability with a variety of strategies. Management may decide to focus on those O-D pairs currently receiving the worst levels of service or on the problem of long trip times in general. The goal can be to improve measures such as the network 3-day-%, the network mean, and the network %-late or to insure that all yards meet a minimum standard of performance. The best strategy, of course, is a function of the network, current performance, the commodities carried, and the resources available. Most of the alternatives listed below can be used to help implement any of these strategies.

1. Improved Management and Information Systems -- Accurate measures of network, sub-network, and O-D performance are needed if management is to identify the extent and sources of unreliability. Specific alternatives include:

a) Include performance as one of the criteria by which operating personnel are judged. If the measures used indeed reflect interaction between sub-network and network performance, then operating personnel will promote the overall goal in the course of doing their own job. This alternative is attractive because it enlists the aid of yardmasters and superintendents--the men most intimately involved with the day-to-day activities that affect system performance[1].

b) Develop data systems which provide better information to top management. Once a situation has been identified and recognized as a problem, management may be able to take appropriate action.

c) Identify cars which have been delayed and give them higher priority for the rest of their trip[2].

2. New Marketing Techniques -- Only a continuing dialogue with individual shippers will inform railroads of their transportation needs. Specialized equipment, unit trains, and piggy-back are responses to shippers' needs that may have to be expanded. Piggy-back in particular seems to offer great potential in that low tonnage and relatively short haul markets can be tapped. Another alternative is to provide less frequent, but more reliable service.

---

[1]Evaluating operating personnel on the basis of measured terminal performance led to 24-33% improvements in the first year(Aase, R.L., "A Terminal Management System," in RSMA, The Measure of Railroad Freight Service, p. 54).

[2]In a discussion at M.I.T., Mr. Robert Wharton of the Southern Railway indicated means of implementing this alternative were being considered.

3. New Schedule and Operating Policies[3]

   a) Schedule more through trains

   b) Run shorter trains in order to provide more frequent service or to connect a greater number of yards[4]

   c) Shift workloads from congested to non-congested yards[5]

   d) Insure that cars travelling between the same origin and destination always receive the same routing; substantial variability results if the number of intermediate yards where the car is switched varies from day-to-day

   e) Expand the policy of running through trains over more than one railroad

   f) Insure that schedules allow sufficient time for all work to be completed and a reasonable amount of delay; adjust

---

[3]Folk investigates the effect of many of these policies. Folk op. cit.

[4]The outstanding example of this is the Santa Fe's Super C which will make a 2,200 mile trip with only 9 cars (Time, 19 July 1971, p. 14). The Rio Grande instituted a policy of running shorter, more frequent trains over its entire network in the early 1960's. Modern Railroads, March, 1967.

[5]This alternative came up in a very interesting discussion with Mr. R. Lacy, general Manager of Operations of the Southern Pacific. Since yards are controlled at the district level even though they perform system functions, there is a powerful incentive to improve district performance (measured primarily by costs) at the expense of network performance (measured both by costs and level of service).

schedules at particular yards in order to achieve the most efficient operation

g) Reduce the number of train cancellations by decreasing minimum train lengths

h) Insure availability of power to run extra trains, thereby limiting the extent of persistent delays caused by train cancellations or lack of power

i) Insure prompt placement, constructive placement, or delivery once a car reaches its destination.

4. Investment Possibilities

a) More motive power may be needed to implement the alternatives discussed above

b) Studies have shown that consolidating facilities in and around cities substantially reduces yard times and yard time variability. For example, A.T. Kearney & Company showed that modernization of the facilities of the Terminal Railroad Association of St. Louis would reduce the average time spent at the St. Louis gateway by 50% and the standard deviation of that time by 60%[6].

---

[6]Leilich, R.H. "Evaluation of Terminal Facilities for Service Reliability," A.T. Kearney & Company, 1972, p. 13.

C. The Economic Impact of Improved Rail Reliability

1. Assumptions

Three major types of benefits will accrue from improved rail reliability:

1) Shippers will realize savings in total (long-run plus short-run) logistics costs
2) Railroads will capture a larger share of the inter-city freight market, and
3) Railroads will receive better car utilization

This section will discuss the magnitude of these benefits that could result from a general improvement in service.

The regression model developed in the last section of Chapter 5 demonstrated the importance of the relationship between unreliability and the mean O-D trip time. That model indicated that increasing the 3-day-% by 10 would decrease the mean .45 days or 11 hours. Since the network 3-day-% for Railroad B was 81 (Section 5B), service improvements increasing this measure to 100 could conceivably result in a reduction in the network mean of nearly an entire day[7]. In the analysis that follows, it is assumed that

---

[7] Extrapolations such as this are generally dangerous. It is used only to find a rough estimate of the potential benefits of improved reliability.

the 3-day-% for all O-D pairs is improved to the extent that the mean O-D Trip Time is reduced by two days[8].

2. Savings in Total Logistics Costs

Inventory benefits are more easily quantified than other logistics savings. The Bureau of the Census estimated the total value of manufacturers', wholesalers', and retailers' inventories to be $172,000,000,000 at the end of 1967. In addition, they estimated the value of manufacturers' inventories alone to be $82,500,000,000 and of shipments from manufacturers to be $45,000,000 per month or $1,500,000,000 a day[9]. This latter figure represents the capital saving resulting from a one day reduction in the average lead time; the annual saving will range from ten to twenty-five percent of the capital savings depending on the industry. Storage, insurance, obsolescence, and loss and damage costs can be substantial, while a 10% cost of capital defines the minimum inventory costs. Thus manufacturers would receive annual savings of at least $150,000,000 and perhaps twice that much for every day that the lead time is

---

[8]Recall the definition in Chapter 2 that the O-D Trip is the time from arrival in the first yard to placement or constructive placement. It is the sum of two or more O-D trips as analyzed in Chapter 5 except in the case that a single railroad both originates and terminates the car. Since the average movement is handled by more than 2 railroads, the assumption that the mean O-D trip is reduced by 2 days implies that the mean O-D trip is reduced by less than a single day.

[9]Bureau of the Census, Manufacturers' Shipments, Inventories, and Orders 1961-1968, p. 9.

reduced.

Assume that a two day reduction in mean times also reduces the lead time two days. Hence, given the hypothetical improvement in service under consideration, shippers using rail transportation will be able to reduce their in-transit inventory by the equivalent of two days shipments. The value of this reduction is estimated in Table 6.1 by multiplying the value of monthly shipments for each by the rail market share for that industry. Summing over all industries, the value of rail shipments from manufacturers is seen to be $20,000,000,000 per month or $670,000,000 per day. Therefore, the two day reduction in lead times resulting from an overall improvement in rail reliability will reduce manufacturers in-transit inventory by $1,340,000,000. The annual savings ranging from $1,340,000,000 to $270,000,000. represent a lower bound on the total savings to industry because

1) if long delays are substantially reduced, the average lead time could be lowered more than two days
2) non-manufacturers shipping by rail will also receive benefits
3) inventory is just one aspect of logistics systems that will be affected
4) the improved service will lower the total logistics costs of all shippers who switch to rail from other transportation modes

TABLE 6.1

TOTAL VALUE OF RAIL SHIPMENTS FROM MANUFACTURERS 1967

Industry	Shipments Tons	Shipments $	Rail Percent	Rail Shipments	Finished Goods Inventory
Stone, Clay, Glass	134[10]	$1.3[11]	36%	$ .5[11]	$1.0[11]
Primary Metal	152	4.6	50	2.3	2.2
Fabricated Metal	17	2.8	23	.6	2.1
Machinery	23	4.7	28	1.3	2.6
Electrical Machinery	14	7.8	34	2.7	2.4
Transport Equipment	43	6.8	56	2.8	1.2
Instruments	2	.9	13	.1	.5
Other Durable	27	2.9	21	.6	3.7
Food and Kindred	192	7.5	50	3.7	2.6
Tobacco, Candy	40	.4	30	.1	2.1
Textile, Leather	20	1.8	10	.2	1.0
Paper and Allied	72	2.0	55	1.1	1.2
Chemicals	158	3.9	50	1.9	2.8
Petroleum, Coal	418	1.9	6	.1	1.2
Rubber, Plastics	10	1.2	25	.3	.8
Other Non-Durable		4.3	30	1.3	9.3
Totals	1384	45.7	33	19.6	26.6
Total (except petroleum, coal)	967	43.8	45	19.5	25.4

[10]TAA, _Transportation Facts and Trends_, p. 13.

[11]Bureau of the Census, _Manufacturers' Shipments, Inventories, and Orders_, p. 9.

3. Increased Rail Revenues

The fourth factor just mentioned is probably the most important because railroads would receive direct benefits in the form of increased revenues. Kullman has found the differential in transit times to be one of the strongest factors influencing the rail-truck modal split. His regression analysis suggests that in situations where rails and trucks each have 50% of the market, a one day reduction in the mean trip time raises the rails share to 60%[12]. Although no estimate is available for the total increase in revenues, it is interesting to note that a 4% increase in rail revenues amounts to nearly $500,000,000 gross income and on the order of $100,000,000 additional profit if marginal costs are assumed equal to 80% of marginal revenues. In short, railroad profits are highly leveraged on traffic volume.

4. Better Car Utilization

Railroads will profit from industry's savings only if rates are raised or if traffic volumes increase. A more tangible effect will be evident in freight car utilization. The 1,750,000 freight cars in service in the United States represent a capital investment of

---

[12] Kullman, M.I.T. Ph.D. Thesis, work in progress.

more than 10 billion dollars. Table 6.2 shows that these cars are underutilized today and that improvements in utilization since 1939, if any, have been marginal. In fact, car loadings per car year declined despite increases in ton miles per car day, total tonnage carried, and average freight train speed. Dividing miles/car-day by the average freight train speed indicates that cars spend on the average only 2.7 hours per day in trains, the majority of the remaining time is spent in railroad yards while a substantial amount is also spent at industrial sidings.

There is clearly room for improvement in car utilization, especially since freight car shortages are frequently cited as a major problem facing the industry. Since the average freight car may be worth $8,000 and the average new boxcar twice that much, the importance of car utilization cannot be over-emphasized. Reducing the average cycle time only 10% or 2.3 days would be equivalent to a costless addition of 170,000 (used) cars to the nations fleet. The capital savings to the industry as a whole would in that case be at least 1.3 billion dollars.

This section is estimating the benefits to be derived from a 2 day reduction in the mean Total Trip Time and an improvement of the network 3-day-% to 100. Under these conditions it is probable that more efficient distribution of empties could also be achieved and that shippers could better schedule loading and

TABLE 6.2

CHANGES IN RAILROAD EFFICIENCY FACTORS, 1939-1969

	1939	1969
Cars in service	1,961.705	1.794,655
Average capacity (tons)	49.7	65.6
Car Loadings	33,911,498	28,291,939
Car Loadings (tons)	901,669,000	1,476,500,000
Car Loadings/ Car year	17.3	15.8
Tons/Loading	26.9	43.5
Miles/Car-day	36.4	53.9
Cycle time (days	21.1	23.4
Average Freight train speed	16.7	20.1
Average distance/shipment (miles)	350	497
Ton miles/freight car day	610	1,426

SOURCE: Yearbook of Railroad Facts

unloading and hence reduce the time spent at industrial sidings. Also, if railroads were more reliable, shippers would not have to allow extra time for placement of empties. Thus a 10% increase in utilization is most likely an understatement of what would actually occur.

The actual savings to industry resulting from this increase in car utilization could be even greater if the need for new investment is reduced. In a recent speech, Federal Railroad Administrator John Ingram stated that "the inabaility of the carriers in their present condition to secure on reasonable terms an adequate supply of freight cars" is a critical problem facing the industry. He then estimated the need for new freight cars:

> "an average of 62,000 new and rebuilt freight cars will be needed annually for replacement purposes. Another 43,000 cars per year will be needed through 1974 to quickly enhance the level of car service and eliminate the deficit that has accrued since the 50's and 60's. Through 1980, therefore, this comes to a total of 747,000 cars at a cost of $11.1 billion."[13]

The improvement in utilization discussed above would decrease the projected need to 580,000 cars at a cost of $8.6 billion -- hence the capital savings could be as great as $2.5 billion over the next decade. At a 10% cost of capital this is equivalent to $250 million annual savings; although only 2% of total rail revenues, this is 33%

---

[13]FRA Administrator John W. Ingram in a speech delivered to the Central Eastern Shippers Advisory Board in Omaha, 27 January 1972.

of net railway operating income. The capital savings represent close to 10% of total net investment in railroads after accrued depreciation and nearly two years of capital expenditures (Table 6.3).

TABLE 6.3*

Transportation Revenues	$10,800
Net Railway Operating Income	760
Annual Capital Expenditures	1,500
Annual Capital Expenditures on Equipment	1,100
Net Investment in Railroads	27,500

*1966-69 averages. Source: Railway Facts

Assuming a 2% increase in rail revenues, the total benefits to railroads and their customers are easily greater than $500,000,000 per year. Figure 6.1 summarizes the individual components of this sum.

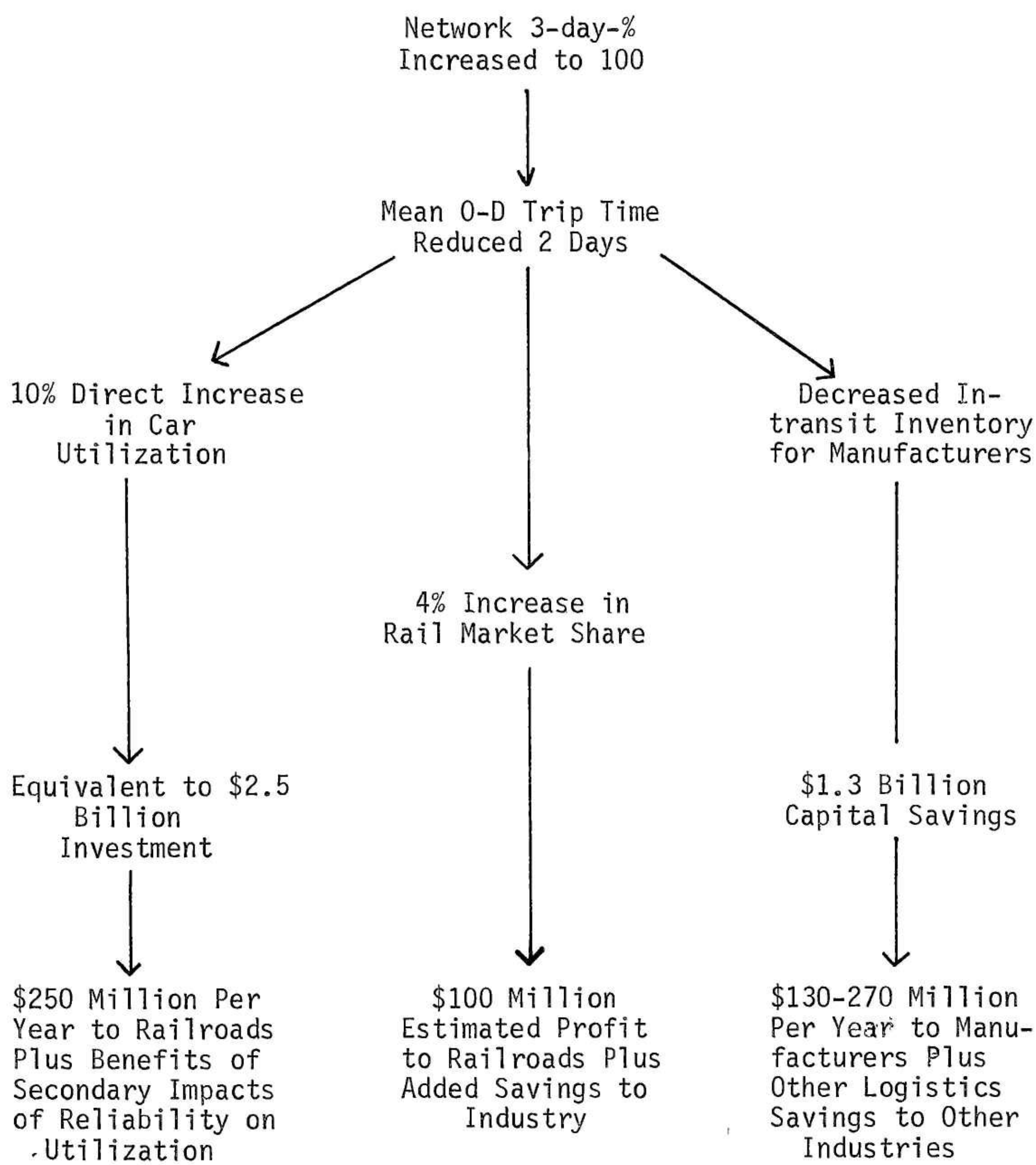
D. Summary and Conclusions

1. The Importance of Improving Rail Reliability

Increased intermodal competition has lowered railroad profitability. Although the rail industry has increased the total tonnage and ton-miles carried annually, its share of intercity freight

FIGURE 6.1

ECONOMIC IMPACT OF IMPROVED RAIL RELIABILITY



traffic and revenues has fallen considerably since World War II. Strategies for strengthening rail profitability ultimately involve either cutting costs or increasing revenues. While the industry has been successful in reducing real costs of labor and supplies, minimizing costs will not necessarily maximize profits. Hence attention must also be given to the possibility of increasing revenues by changing the rate structure or by improving service. In particular better reliability could lead to a greater demand for rail freight transportation.

2.1 Measuring Railroad Reliability

A crucial issue arises immediately -- what is meant by "reliability?" Defining it as the variability in O-D trip times is helpful, but this does not indicate how it is to be measured. A qualitative concept such as this can rationally influence operating decisions only to the extent that it is quantifiable. If railroads hope to achieve the goal of attracting more traffic by providing better service, then their measures of service must in fact be relevant to shippers. Therefore, railroads must actively consider the impact of rail service on shipper profitability before choosing performance measures.

2.2 The Impact of Reliability on Logistics Systems

The impact of unreliability on a shipper will depend to a large degree on the particular structure of his logistics system, a set of inter-related components which cannot be managed independently. Transportation has a direct impact on the lead time which is an important factor in inventory and warehousing decisions. The fact that untimely shipments are costly means that shippers will accept long term expenses in order to prevent them or to ameliorate their effects. Thus the costs of unreliability include more more than the direct costs associated with individual early or late shipments.

Inventory is but one component of logistics systems, but it provides a useful means of studying the long- and short-run impacts or unreliability. Analysis in Chapter 3 demonstrates that the mean trip times, the probability of long trip times, and the compactness of the trip time distribution all influence the amount of inventory required to achieve a given level of stock-out probability. These three characteristics are assumed to be the most important aspects of rail service to measure. However, a continuing dialogue with specific shippers will be necessary to determine exactly what kind of performance they consider essential.

2.3 Selecting Performance Measures

Chapter 4 evaluates a number of possible measures of compactness and long delays. "On time performance," a commonly used measure, is not a meaningful measure of reliability. By changing the definition of "on time," measured reliability can be changed even if the trip time distribution is not. The standard deviation is also generally inappropriate because the extreme values often destroy its interpretive usefulness. The maximum percentage of cars arriving in a three day period (the 3-day-%) is chosen as the best measure of compactness. Long delays are measured by the percentage of cars which arrive after this three day interval. Used in conjunction with the mean trip time, these measures clearly describe each of those characteristics of the trip time distribution found to be important to shippers.

3.1 Origin to Destination Reliability

The typical O-D trip time distribution measured in days is unimodal with extreme values concentrated on the high side. Hence the mean is always greater than the mode of the distribution. O-D performance measures (for moves handled by a single carrier) usually fall in the following ranges:

Mean:	2-6 days
3-day-%:	60-95
%-late:	0-25

A regression of the 3-day-% on general characteristics of the O-D trip such as distance and the number and type of yards fails to account for the observed range of values for this measure. Further study shows that large variations in yard performance may account for this negative result. In any case, this result is important because it indicates that studies of unreliability cannot deal with network characteristics alone, but must also consider specific characteristics of individual yards.

3.2 Trip Segment Performance

The O-D trip is composed of three segments: the origin yard time, the total line haul time including intermediate yards, and the destination yard time. Unreliability in each of these segments varies widely between O-D pairs and may at times contribute greatly to the trip time unreliability. Only some of this variability can be explained by general trip characteristics. Moves to and from industry cause the highest yard time standard deviations and the number of intermediate yards (not distance) raises the total line haul standard deviation. However, detailed analyses of yard operations and policies will be necessary to discover the reasons for unreliability in particular yards and between particular O-D pairs.

3.3 The Relationship Between Unreliability and Mean Trip Times

Unreliability is caused in part by cars missing connections at classification yards. As mentioned above, the extent of unreliability in O-D pairs is not directly related to the number or type of yards where a car is switched. Since the delays which cause unreliability also increase the mean trip time, reliability must be included as an independent variable in studies of mean trip time. A significant regression equation (R^2=.55) is derived in Chapter 5 which gives mean trip time as a function of distance, yards, and the 3-day-%. According to this equation, raising the 3-day-% 10 lowers the mean 11 hours. This relationship is used in Chapter 6 to investigate the consequences of a major improvement in network performance. Notice that this analysis implies that there need not be a tradeoff between reliability and the mean trip time.

4.1 Improving Rail Service

"Improved service" is a nebulous concept when applied to a railroad network. It could mean improvements in network performance measures, an increase in the percentage of O-D pairs with acceptable O-D performance, or specific changes in service received by important O-D pairs. Railroads should choose such goals only after consideration of the potential costs and benefits to themselves and to all shippers -- not just to those currently shipping by rail. Many alternatives for achieving stated goals exist in each of the following areas:

Investments in facilities and power

Management and information systems

Marketing techniques

Operating schedules and policies

These, however, must be evaluated in each individual situation.

4.2 Benefits of Improved Service

In general there are three categories of benefits that can be attained. Shippers can realize savings in total logistics costs; railroads can capture a larger share of the intercity freight market; and investments in rolling stock can be reduced due to better car utilization. Chapter 6 estimates the magnitude of these benefits resulting from a two day reduction in mean dock-to-dock trip times caused by an increase in O-D reliability.

The capital savings to manufacturers currently shipping by rail resulting from inventory savings alone could amount to nearly $3 billion. Savings to other industries shipping by rail and in other components of logistics systems would also be substnatial. In addition, as a result of these service improvements and the potential logistics savings, other shippers would be induced to change to rail transportation. Even a small (4%) increase in revenues may lead to a $100 million increase in yearly net operating income.

The third type of benefits, could, however, be the most important. A conservative estimate of a 10% increase in car utilization would substantially lower the need for new rolling stock. In fact, a $3 billion investment in cars could be avoided over the next ten years.

The sum of all benefits to railroads and to industry is easily equivalent to $500 million annually (Figure 6.1). This impact is great enough to justify not only continued study of the problem, but also actual attempts to improve rail service.

BIBLIOGRAPHY

American Association of Railroads, Yearbook of Railroad Facts, AAR, 1971.

Battelle Memorial Institute, "A Pilot Study of Shippers Modal Choice and Rail Freight Movements," Report to AAR, November 30, 1970.

Beckman, McGuire, Winsten, Studies in the Economics of Transportation (Yale Press 1956).

Belovarac, K., Determinants of Unreliability in Railroad Line Haul Operations, S.M. Thesis, M.I.T., 1972.

Benthe, M.V., Freight Transport Mode Choice: An Application to Corn Transportation, Ph.D. Thesis, Northwestern 1968.

Bureau of the Census, Manufacturers' Shipments, Inventories, and Orders, 1961-68, Series M3-1.1, (USGPO, Washington, D.C. 1968).

Cook, Emory, "Prediction and Analysis of Classification Yard Performance," presentation to Railway Systems and Procedures Association, Chicago, Illinois, 1960 Fall Meeting.

Collard, Thomas, "Industrial Quality Control Techniques Applied to Railroad Freight Service Measurement," (obtained from Freight Service Planning, Penn Central, Philadelphia, Pennsylvania), 1970.

Folk, J., Models for Investigating the Unreliability of Rail Freight Shipments, Ph.D. Thesis, M.I.T., 1972.

Friedlander, Ann, Dilemma of Freight Transportation (Brookings Institute, Washington, D.C.).

Hadley, G. and T.M. Whitin, Analysis of Inventory Systems (Prentice-Hall, Inc., Englewood Cliffs, N.J.).

Heskett, J.L., R. Ivie, and N. Glaskowsky, Jr., Business Logistics, (The Ronald Press Company, New York, New York).

Jennings, A.A., The Effects of Train Length on the Reliability of Operations of Railroad Yards, S.M. Thesis, M.I.T., 1972.

Kolsen, H.M., The Economics and Control of Road Rail Competition, (Sydney University Press, Sydney, Australia) 1968.

Kullman, B., M.I.T. Ph.D. Thesis, in progress.

Lang, A.S. and R. Reid, "Railroad Movement Reliability: A Preliminary Study of Line-Haul Operation," M.I.T. Research Report, 1970.

Leilich, R.H., "Evaluation of Terminal Facilities for Service Reliability," A.T. Kearney & Company, 1972.

Little, A.D., Study of Freight Car Shortages, Appendix A.

M.I.T. Department of Civil Engineering, Progress Report on the Study of Reliability in Railroad Network Operations (Federal Railroad Administration Contract #DOT-FR-10006) January 1972.

Meyer, Peck, Stenason, and Zwick, The Economics of Competition in the Transportation Industries (Harvard University Press) 1959.

Moody's Transportation Manual, 1971.

Nelson, James, Railroad Transportation and Public Policy (Brookings Institute, Washington, D.C.) 1959.

O'Doherty, J., Classification Yard Effects in Rail Freight Shipments, Civil Engineering Thesis, M.I.T., 1972.

Poole, E.C., Costs -- A Tool for Railroad Management (Simmons-Boardman Publishing Corporation, New York, 1962).

Reid, R., Yard Unreliability in Rail Freight Movement, M.I.T., S.M. Thesis, 1971.

Roberts, Paul O., "The Logistics Management Process as a Model of Freight Traffic Demand," (Working Paper, Harvard Business School, April 1971) HBS 71-11.

Railway Systems and Management Association, Railroad Marketing, February, 1968.

------------, The Measure of Railroad Freight Service, November, 1970.

------------, Transportation Network Analysis, October, 1969.

Sillcox, H., "Stand-By and Second-Best," presented before GSBA, Harvard, 1971.

Transportation Association of America, Transportation Facts and Trends, April, 1972.

U.S. Department of Transportation, Office of Systems Analysis and Information, National Transportation Study, 1972.