

An Optimization Based Bidding Process: A New Framework for Shipper-Carrier Relationships

by

Christopher George Caplice

B.S. Civil Engineering
Virginia Military Institute, 1984

M.S. Civil Engineering
University of Texas at Austin, 1990

Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in Transportation and Logistics Systems

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1996

© Massachusetts Institute of Technology
All rights reserved.

ARCHIVES
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

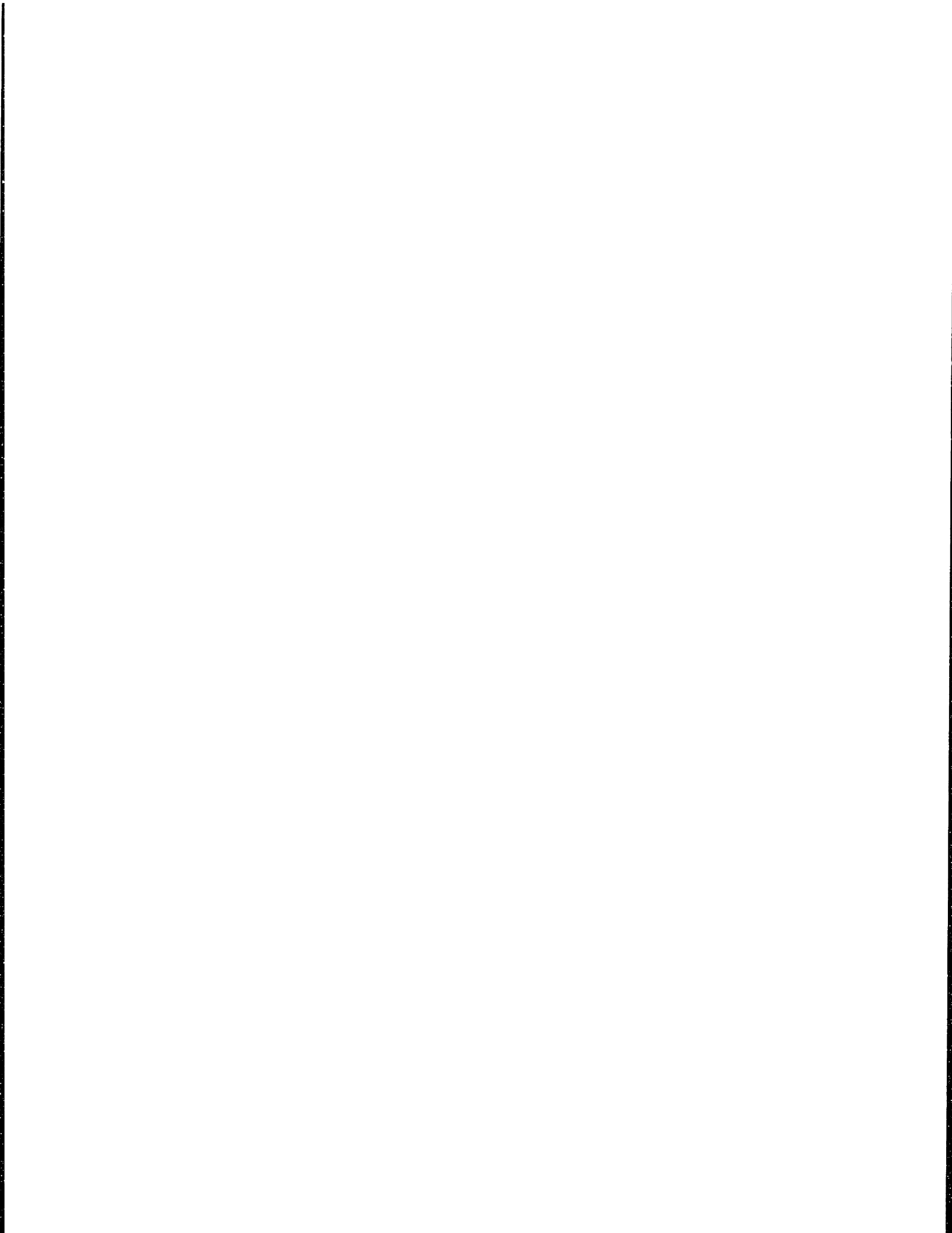
JUN 05 1996

LIBRARIES

Signature of Author
Department of Civil and Environmental Engineering
May 24, 1996

Certified by.....
Yosef Sheffi
Professor, Department of Civil and Environmental Engineering
Thesis Supervisor

Accepted by.....
Joe Sussman
Chairman, Departmental Committee on Graduate Studies



An Optimization Based Bidding Process: A New Framework for Shipper-Carrier Relationships

by

Christopher G. Caplice

Submitted to the Department of Civil and Environmental Engineering
on May 24, 1996 in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in Transportation and Logistics Systems

ABSTRACT

This thesis addresses how shippers (buyers) should procure transportation services from truckload (TL) motor carriers (suppliers). TL carriers operate over irregular routes moving directly from origin to destination without any intermediate stops. A significant portion of a TL carrier's costs is due to the repositioning of empty vehicles (deadheading) from the destination of one load to the origin of the follow-on load. Connection costs are never known with certainty by the carrier due to the short dispatching lead times and overall spatial and temporal variability of shipper demand. The uncertainty in connection costs creates lane interdependencies for TL carrier operations, that is, the cost of serving a lane (origin-destination pair) is affected by the other lanes that a carrier is serving. Thus, TL carriers exhibit economies of scope.

Current practice and research has concentrated on applying generic procurement strategies, such as supplier reduction and certification programs, to TL transportation. While beneficial in many respects, these initiatives focus on achieving economies of scale for suppliers while ignoring economies of scope. This thesis takes a new approach to TL procurement by investigating how a shipper can modify its contracting, bidding, and assignment processes in order to assist TL carriers in achieving the benefits of economies of scope. Reducing the carriers' costs can, in turn, lower the shipper's transportation costs.

This thesis makes four contributions. First, the effect of economies of scope on shipper-carrier relationships are analyzed and a network-based justification for the use of contracting in the seemingly perfectly competitive market for TL trucking is made. Second, a decision framework for designing TL carrier bids is developed, to include the use of combinatorial auctions. Conditional bids are introduced as a method of reducing carrier hedging due to connection uncertainty. Third, a methodology for identifying potential efficient aggregations of lanes is developed and applied to an actual shipper's network. Finally, optimization models incorporating conditional bids and system restrictions are developed and used to assign carriers to lanes in a shipper's network. These models are applied to data from an actual TL bid.

Thesis Supervisor: Yosef Sheffi

Title: Professor, Department of Civil and Environmental Engineering



Acknowledgments

I would like to acknowledge and thank the following people for their help during my time here at MIT:

My advisor, Professor Yosef Sheffi for his advice, support, and friendship over the last four years. I learn something new and insightful every time we have a discussion. I am constantly amazed by Yossi's unlimited energy, his boundless imagination, and his unfailing (and sometimes misplaced) faith in my abilities. But of all of Yossi's traits, the one I admire the most is his correct viewpoints on most political issues.

My committee members:

Professor Gabriel Bitran for his insightful suggestions and comments that led this thesis on a far different (but much more interesting) path than originally intended.

Professor Cynthia Barnhart for providing patient and unerring advice on matters both technical and personal. Cindy's concern and empathy for the students is matched only by her drive and capabilities for research. I would also like to thank Cindy for supporting (and feeding) the Large Scale Optimization Group from which I have benefited over the last several years.

Other professors and faculty at MIT: Dr. Jonathan Byrnes for introducing me to a wider view of logistics and business in general and for always being there for support and guidance; Professor Joe Sussman for being a good political sparring partner and for clinging to his clearly outdated liberal views; Peter Metz for his sage advice and for always being in a good mood, and Tom Humphrey for supporting me during my last semester at MIT.

Dr. Scott Smith for his insights into transportation and trucking in general. This dissertation would not have been possible without Scott's earlier work on this topic.

Fellow students at CTS, past and present, who have made life at MIT a little less dreary and allowed me to claim ownership over the Pentiums in the computer lab for months at a time. I would not have made it through MIT without the continuous help (computer and otherwise), guidance, and occasional bottle of Scotch from my office mate Daeki Kim. I would also like to especially acknowledge Rajesh Sheno, Francisco Jauffred, Jai Kue Park, and Max Kiesling for their help over the last few years.

Finally, I would like to thank my family for their support and love.

Dedication

This dissertation is dedicated to my wife Kristin for more reasons than could possibly fit in this dissertation, much less on this page.

You are allowed to shoot me if I want to get another degree.

Table of Contents

| | |
|--|-----------|
| ACKNOWLEDGMENTS | 5 |
| DEDICATION | 6 |
| TABLE OF CONTENTS | 7 |
| LIST OF FIGURES..... | 11 |
| LIST OF TABLES..... | 13 |
| CHAPTER 1 INTRODUCTION..... | 15 |
| 1.1 RESEARCH OBJECTIVE..... | 16 |
| 1.2 TRANSPORTATION SYSTEM FUNDAMENTALS | 17 |
| 1.2.1 <i>Core Activities</i> | 18 |
| 1.2.2 <i>Classification of Transportation Systems</i> | 21 |
| 1.2.3 <i>Carrier Economics</i> | 26 |
| 1.3 TRUCKLOAD PROCUREMENT PROBLEM | 30 |
| 1.3.1 <i>Procurement Model</i> | 30 |
| 1.3.2 <i>Review of Literature</i> | 31 |
| 1.4 OUTLINE OF THESIS | 33 |
| CHAPTER 2. MOTOR CARRIER INDUSTRY OVERVIEW | 37 |
| 2.1 REGULATORY AND DEMAND FACTORS | 39 |
| 2.1.1 <i>Regulatory Factors</i> | 39 |
| 2.1.2 <i>Shipper Trends</i> | 42 |
| 2.2 TRUCKLOAD (TL) CARRIERS | 44 |
| 2.2.1 <i>Specialty versus General Freight Carriers</i> | 44 |
| 2.2.2 <i>Carrier Tiers</i> | 45 |
| 2.2.3 <i>Concentration in the TL Market</i> | 51 |
| 2.2.4 <i>Competitive Environment</i> | 52 |
| 2.3 OTHER MOTOR CARRIER SEGMENTS..... | 54 |
| 2.3.1 <i>Private Fleets</i> | 54 |
| 2.3.2 <i>Less-Than-Truckload (LTL) Carriers</i> | 56 |

| | |
|--|-----------|
| 2.3.3 Package Delivery Carriers..... | 57 |
| 2.4 CLOSURE..... | 57 |
| CHAPTER 3 SHIPPER-CARRIER RELATIONSHIPS..... | 59 |
| 3.1 TRANSACTION COST ANALYSIS | 59 |
| 3.1.1 Total Transaction Costs | 60 |
| 3.1.2 Behavioral Assumptions..... | 61 |
| 3.1.3 Characteristics of the Exchange..... | 62 |
| 3.1.4 Governance Structures..... | 65 |
| 3.2 GOVERNANCE STRUCTURES IN TL TRUCKING | 69 |
| 3.2.1 Vertical Integration..... | 69 |
| 3.2.2 Spot Market | 71 |
| 3.2.3 Contractual Agreements..... | 75 |
| Summary..... | 78 |
| 3.3 SELECTION OF GOVERNANCE STRUCTURE | 78 |
| 3.3.1 Review of Previous Research | 78 |
| 3.3.2 Network Based Justifications..... | 83 |
| 3.3.3 Other Transportation Modes | 89 |
| 3.3.4 Mixed Governance Structures | 90 |
| 3.4 CLOSURE..... | 91 |
| CHAPTER 4 AUCTION DESIGN FOR TL BIDS..... | 93 |
| 4.1 GENERAL PROBLEM AND NOTATION | 93 |
| 4.2 FUNDAMENTAL AUCTION CONCEPTS | 96 |
| 4.2.1 Reservation Value..... | 97 |
| 4.2.2 Efficient Allocation..... | 101 |
| 4.2.3 Equilibrium Conditions..... | 102 |
| 4.3 AUCTION DESIGN DECISIONS..... | 112 |
| 4.3.1 Order of Auction: Sequential versus Simultaneous | 113 |
| 4.3.2 Types of Bids: Simple versus Conditional..... | 117 |
| 4.3.3 Number of Rounds: Single versus Multiple..... | 123 |
| 4.3.4 Pricing Mechanism: First versus Second Price | 129 |
| 4.4 COMBINATORIAL AUCTIONS..... | 131 |
| 4.4.1 Combinatorial Auction Concepts | 131 |
| 4.4.2 Applications and Experiments..... | 141 |
| 4.4.3 Comparisons and Relevance to TL Bids | 152 |

| | |
|---|------------|
| 4.5 CLOSURE..... | 153 |
| CHAPTER 5 PRE-BID NETWORK ANALYSIS..... | 155 |
| 5.1 STEP 1: EXPANDING THE NETWORK..... | 158 |
| 5.2 STEP 2: CONSTRUCTING THE NETWORK..... | 161 |
| 5.2.1 <i>Shipping Point Classification</i> | 162 |
| 5.2.2 <i>OD Flow Classification</i> | 165 |
| 5.2.3 <i>Identification of Zones</i> | 167 |
| 5.3 STEP 3: ASSESSING MATCHING OPPORTUNITIES..... | 169 |
| 5.3.1 <i>Network Topology</i> | 169 |
| 5.3.2 <i>Flow Distribution</i> | 170 |
| 5.4 STEP 4: IDENTIFYING POTENTIAL EFFICIENT AGGREGATIONS..... | 177 |
| 5.4.1 <i>Reloads</i> | 178 |
| 5.4.2 <i>Open Tours / Threads</i> | 181 |
| 5.4.3 <i>Closed Tours</i> | 184 |
| 5.4.4 <i>Local Tours</i> | 187 |
| 5.4.5 <i>Application to the Data</i> | 189 |
| 5.4.6 <i>Practical Concerns</i> | 189 |
| 5.5 STEP 5: DESIGNING THE BID..... | 190 |
| 5.5.1 <i>System Restrictions</i> | 191 |
| 5.5.2 <i>Repetitive Bids</i> | 192 |
| 5.5.3 <i>Separable Effects</i> | 193 |
| 5.5.4 <i>Contention Level</i> | 193 |
| 5.5.5 <i>Pre-Processing</i> | 194 |
| 5.6 CLOSURE..... | 195 |
| CHAPTER 6 CARRIER ASSIGNMENT PROBLEM..... | 197 |
| 6.1 GENERAL FORMULATION..... | 197 |
| 6.1.1 <i>Variations of Carrier Assignment Model</i> | 199 |
| 6.1.2 <i>Previous Research</i> | 200 |
| 6.2 CARRIER-SET FORMULATION..... | 203 |
| 6.2.1 <i>Formulation</i> | 203 |
| 6.2.2 <i>Implementation Strategies</i> | 206 |
| 6.2.3 <i>Summary</i> | 212 |
| 6.3 BID-SET FORMULATIONS..... | 213 |
| 6.3.1 <i>Single Primary Carrier per Lane</i> | 214 |

| | |
|---|------------|
| 6.3.2 <i>Multiple Classes of Bids</i> | 215 |
| 6.3.3 <i>Additional Considerations and Constraints</i> | 217 |
| 6.4 IMPLEMENTATION MODELS | 219 |
| 6.4.1 <i>Modification of Data</i> | 220 |
| 6.4.2 <i>Solution Strategies</i> | 223 |
| 6.4.3 <i>Business Implications</i> | 226 |
| 6.4.4 <i>Multiple Round Auctions</i> | 229 |
| 6.5 CLOSURE..... | 231 |
| CHAPTER 7 SYNOPSIS, CONTRIBUTION, AND EXTENSIONS | 233 |
| 7.1 SYNOPSIS | 234 |
| 7.2 CONTRIBUTIONS | 238 |
| 7.3 EXTENSIONS | 239 |
| REFERENCES | 243 |

List of Figures

| | |
|---|-----|
| FIGURE 1.1 SCHEMATIC OF CONSOLIDATED TRANSPORTATION OPERATIONS..... | 23 |
| FIGURE 1.2 SCHEMATIC OF THE CONNECTION BETWEEN LINE-HAULS..... | 24 |
| FIGURE 1.3 TIME-SPACE DIAGRAM FOR A SINGLE VEHICLE. | 25 |
| FIGURE 1.4 EXAMPLE NETWORK ILLUSTRATING DIFFERENT TRANSPORT ECONOMIES..... | 27 |
| FIGURE 1.5 FIVE-STEP TRANSPORTATION PROCUREMENT PROCESS. | 30 |
| FIGURE 1.6 OVERVIEW OF THESIS..... | 34 |
| FIGURE 2.1 TRUCKING INDUSTRY SEGMENTATION BASED ON SHIPMENT SIZE (POUNDS)..... | 37 |
| FIGURE 2.2 BREAKDOWN OF 1994 REVENUES AND MARKET SIZE FOR MOTOR CARRIER INDUSTRY SEGMENTS | 38 |
| FIGURE 2.3. GROWTH OF TRUCKING MARKET SINCE DEREGULATION..... | 40 |
| FIGURE 2.4 CONCENTRATION OF GENERAL FREIGHT TL SEGMENT | 52 |
| FIGURE 2.5 OVERVIEW OF COMPETITIVE ANALYSIS OF TL TRUCKING SEGMENT..... | 53 |
| FIGURE 3.1. SIMPLE CONTRACTING SCHEME (WILLIAMSON 1985, 33)..... | 67 |
| FIGURE 3.2. CONTINUUM OF GOVERNANCE STRUCTURES. | 69 |
| FIGURE 3.3 SHIPPER’S NETWORK FOR ECONOMIES OF SCOPE EXAMPLE..... | 84 |
| FIGURE 4.1. SAMPLE DISTRIBUTION NETWORK..... | 98 |
| FIGURE 4.2. SAMPLE 3 NODE NETWORK..... | 103 |
| FIGURE 4.3. BOUNDING CONSTRAINTS FOR LANE BID PRICES..... | 107 |
| FIGURE 4.4. TRAJECTORY OF BIDS - CASE 1: NO INTERDEPENDENCIES..... | 107 |
| FIGURE 4.5. BOUNDING CONSTRAINTS FOR EXAMPLE - CASE 2: MINOR INTERDEPENDENCIES..... | 109 |
| FIGURE 4.6. BID TRAJECTORY - CASE 2: MINOR INTERDEPENDENCIES..... | 110 |
| FIGURE 4.7. INFEASIBLE CORE. CASE 4: STRONG INTERDEPENDENCIES..... | 120 |
| FIGURE 4.8. BID TRAJECTORY. CASE 4: STRONG INTERDEPENDENCIES..... | 121 |
| FIGURE 4.9. SAMPLE 3 NODE NETWORK..... | 132 |
| FIGURE 4.10 EXAMPLE OF NESTED AND NON-NESTED BID STRUCTURES..... | 134 |
| FIGURE 4.11. TYPES OF COMBINATORIAL BIDS FOR A LOW-BID-WINS AUCTION..... | 139 |
| FIGURE 4.12. TIME-SPACE DIAGRAM FOR RAIL TRACK SLOTS..... | 147 |
| FIGURE 4.13. CONFLICTS DUE TO ADDING A CONDITIONAL BID..... | 148 |
| FIGURE 4.14. CONFLICTS DUE TO ADDING A CONDITIONAL BID..... | 148 |
| FIGURE 5.1 EXAMPLE OF LANE AND ZONE IDENTIFICATION..... | 162 |
| FIGURE 5.2 CONCENTRATION OF FLOWS THROUGH POINTS..... | 163 |
| FIGURE 5.3 CLASSIFICATION OF MTMC SHIPPING LOCATIONS BY VOLUME AND PERCENTAGE OF OUTBOUND LOADS HANDLED..... | 164 |

| | |
|---|-----|
| FIGURE 5.4 CLASSIFICATION OF ABC SHIPPING LOCATIONS BY VOLUME AND PERCENTAGE OF OUTBOUND LOADS HANDLED..... | 164 |
| FIGURE 5.5 PLOT OF OD ANNUAL VOLUME (LOADS/YEAR) VERSUS DISTANCE (MILES) FOR 5-DIGIT ZIP CODES FOR THE ABC DATA..... | 167 |
| FIGURE 5.6 THE RELATIVE NUMBER OF IB TO OB LANES FOR THE HIGH-DENSITY NETWORK FOR MTMC DATA..... | 170 |
| FIGURE 5.7 SEASONALITY OF TRAFFIC ACROSS HIGH DENSITY LANES FOR MTMC..... | 171 |
| FIGURE 5.8 VARIATION OF DAY OF THE WEEK ACROSS THE SYSTEM..... | 172 |
| FIGURE 5.9 DISTRIBUTION OF INBOUND AND OUTBOUND LANES AT STOCKTON YARDS, CA..... | 173 |
| FIGURE 5.10 WEEKLY DISTRIBUTION OF INBOUND AND OUTBOUND TRAFFIC AT STOCKTON YARDS, CA | 173 |
| FIGURE 5.11 DAY OF THE WEEK DISTRIBUTION OF INBOUND AND OUTBOUND TRAFFIC AT STOCKTON YARDS, CA..... | 174 |
| FIGURE 5.12. EXAMPLE NETWORK FOR THREAD IDENTIFICATION..... | 182 |
| FIGURE 5.13 SEARCH TREE FOR NODE 1 AS ROOT NODE..... | 183 |
| FIGURE 5.14 EXAMPLE OF FOUR LANE THREAD WITH ONE DEADHEAD LEG..... | 183 |
| FIGURE 5.15 EXAMPLE NETWORK FOR CLOSED TOUR IDENTIFICATION..... | 186 |
| FIGURE 5.16 SEARCH TREE FOR NODE 1 AS ROOT NODE..... | 186 |
| FIGURE 5.17. FLOWCHART OF AUCTION DESIGN CHOICE..... | 191 |
| FIGURE 5.18 FLOW BETWEEN THREE MAJOR SHIPPING POINTS IN MTMC NETWORK..... | 193 |
| FIGURE 5.19 EXAMPLE OF CONTENTION IN THE MTMC NETWORK..... | 194 |
| FIGURE 6.1 EXAMPLE FOR CARRIER-SET FORMULATION..... | 207 |
| FIGURE 6.2 BID-SET NETWORK FOR CARRIER A IN EXAMPLE..... | 210 |
| FIGURE 6.3 DUALITY GAP COMPARISON BETWEEN AGGREGATED AND DISAGGREGATED MODELS FOR PARAMETRIC ANALYSIS ON CARRIER SYSTEM LIMITS..... | 224 |
| FIGURE 6.4 DUALITY GAP COMPARISON BETWEEN AGGREGATED AND DISAGGREGATED MODELS FOR PARAMETRIC ANALYSIS ON LOCATION LIMITS..... | 224 |
| FIGURE 6.5 PERCENTAGE OF INTEGER CARRIER COLUMNS COMPARISON BETWEEN AGGREGATED AND DISAGGREGATED MODELS FOR PARAMETRIC ANALYSIS ON CARRIER SYSTEM LIMITS..... | 225 |
| FIGURE 6.6 PERCENTAGE OF INTEGER LOCATION COLUMNS COMPARISON BETWEEN AGGREGATED AND DISAGGREGATED MODELS FOR PARAMETRIC ANALYSIS ON LOCATION LIMITS..... | 225 |
| FIGURE 6.7 COMPARISON OF MARGINAL LOSS TAKEN FROM DUAL OF CARRIER SYSTEM CONSTRAINT AND COEFFICIENT OF CARRIER COLUMN..... | 228 |

List of Tables

| | |
|--|-----|
| TABLE 2.1. REVENUE VERSUS LENGTH OF HAUL FOR TL. | 43 |
| TABLE 2.2. FOUR TIERS OF CARRIERS IN THE TRUCKLOAD SEGMENT. | 46 |
| TABLE 4.1. RESERVATION VALUES FOR CARRIERS. | 102 |
| TABLE 4.2. CURRENT BIDS FOR CARRIERS. | 102 |
| TABLE 4.3. CARRIER RESERVATION VALUES. CASE 1: NO INTERDEPENDENCIES. | 103 |
| TABLE 4.4. AUCTION SUMMARY - CASE 1: NO INTERDEPENDENCIES. | 107 |
| TABLE 4.5. RESERVATION VALUES FOR CARRIERS. CASE 2: MINOR INTERDEPENDENCIES. | 108 |
| TABLE 4.6. AUCTION SUMMARY - CASE 2: MINOR INTERDEPENDENCIES. | 110 |
| TABLE 4.7. RESERVATION VALUE FOR CARRIERS FOR SINGLE LANE. | 113 |
| TABLE 4.8. RESERVATION VALUES. CASE 3: MINOR INTERDEPENDENCIES BY SEQUENTIAL AUCTION. | 114 |
| TABLE 4.9. RESERVATION VALUES. CASE 4: STRONG INTERDEPENDENCIES. | 119 |
| TABLE 4.10. SAMPLE AUCTION. CASE 4: STRONG INTERDEPENDENCIES. | 121 |
| TABLE 4.11. RESERVATION VALUES. CASE 5: STRONG INTERDEPENDENCIES USING CONDITIONAL BIDS. . | 122 |
| TABLE 4.12. CURRENT CARRIER BIDS FOR 3 NODE NETWORK - COMBINATORIAL AUCTIONS. | 132 |
| TABLE 4.13. STANDING BIDS FOR A COMBINATORIAL AUCTION. | 139 |
| TABLE 5.1 SUMMARY OF SHIPPING POINT VOLUME. | 163 |
| TABLE 5.2 ORIGIN-DESTINATION FLOW FOR MTMC DATA. | 166 |
| TABLE 5.3 ORIGIN-DESTINATION FLOW FOR ABC DATA. | 166 |
| TABLE 5.4. SUMMARY OF HIGH AND LOW DENSITY MTMC NETWORKS. | 169 |
| TABLE 5.5 RESULTS OF PORTION OF ESTIMATIONS FOR LANE PROBABILITIES. | 175 |
| TABLE 5.6. SYSTEM RESTRICTION EXAMPLE - CURRENT BIDS. | 192 |
| TABLE 6.1 CHARACTERISTICS OF THE COST FUNCTION FOR THE GENERAL MODEL. | 199 |
| TABLE 6.2 CONSTRAINT MATRIX OF CARRIER-SET FORMULATION. | 207 |
| TABLE 6.3 COMPARISON OF DATA SET SIZE. | 221 |
| TABLE 6.4 COMPARISON OF DATA SET SIZE. | 221 |
| TABLE 6.5 ANALYSIS OF ABC TL BID. | 222 |
| TABLE 6.6 COMPARISON OF PROBLEM SIZES. | 223 |

Chapter 1

Introduction

The motor carrier industry is in a period of transformation. The once simple decision by shippers (buyers) to select which carriers (sellers) to service their distribution network has become increasingly more complicated. The competitive environment that was ushered in by the Motor Carrier Act of 1980 deregulating inter-city trucking is still vibrant. Shippers and carriers are constantly searching for innovative ways to remove the inefficiencies and improve the effectiveness of the transportation process.

Motor carriers have made substantial investments in computer and communication hardware and software and are now some of the most sophisticated users of geographical information systems (GIS) and intelligent transportation system (ITS) technologies. Service levels have improved while costs have decreased in all segments of the motor carrier industry. A measure of these efficiencies is the fact that from 1982 to 1992 the spending on trucking as a percentage of GDP decreased by 5% while total tonnage hauled increased by over 50% (American Trucking Association 1994, 5-7).

Shippers, too, are placing greater emphasis on their transportation operations. Transportation has played a major role in many recent manufacturing/production initiatives, such as just-in-time (JIT) inventory policies (requiring smaller, more frequent, and highly reliable deliveries), Supply Chain Integration (requiring real-time control and tracking capabilities), and Quick Response/Efficient Consumer Response (requiring increased scheduling flexibility). In all of these programs, motor carrier transportation plays a critical role. Shippers, by placing greater reliance on the transportation link, are now searching for ways to guarantee both the availability and the quality of trucking services.

While transportation is important to these, and many other, management-led initiatives, the manner in which transportation services are procured is only very recently starting to change. Senior transportation and logistics managers are grappling with how and from whom to procure transportation services. In the last few years, companies have begun to experiment with different forms of shipper/carrier relationships. Core carrier systems (where the

transportation supplier base is dramatically reduced), dedicated fleets (where operational control remains with the shipper but all the equipment, drivers, maintenance, and other support services are leased), and single sourcing of transportation management to third party providers are just three of the many ways that shippers are searching for better methods of procuring transportation services.

1.1 Research Objective

This thesis explores how shippers, or their agents, should procure transportation services. Specifically, we investigate how shippers currently do and actually should: (1) contract with, (2) let bids out to, and (3) assign traffic to truckload (TL) motor carriers.

The key insight, and the primary contribution, of this thesis is to recognize and incorporate the underlying economics of TL carrier operations into the way in which the transportation services are procured. While often characterized as the ultimate commodity, the "box on wheels," TL trucking is not a homogenous service. The structure of the shipper's network, the level and reliability of the traffic volume, the degree to which the shipper's traffic complements a carrier's service network, as well as the competitive environment, all influence the TL carrier's cost structure. Understanding these influences can lead to better procurement methods. In most companies, however, the current system for procuring TL services ignores these opportunities and instead applies the same methods used to buy most other products - leveraging volume and treating each lane (origin-destination pair) as an independent item.

The problem with this approach is that TL carriers are more sensitive to economies of scope, where the cost of serving a lane depends on the other lanes being served, than to economies of scale. The interdependency between lanes is a result of the cost of making a connection between subsequent loads. This usually requires moving a truck with an empty trailer to a different location (deadhead) and having a driver wait (dwell time). Additionally, there is a large amount of uncertainty involved with securing a follow-on load. Because the carriers receive very little advance notice (usually just hours) from shippers, the repositioning costs are never known with certainty. Carriers, then, need to estimate these repositioning costs and modify their pricing to hedge against the uncertainty of the cost involved with finding a follow-on load.

By requiring bids (or spot prices) on individual lanes, the carrier is forced to place a value on each separate truck movement. The additional cost of identifying and repositioning the vehicle for an acceptable follow-on load can only be incorporated into the price as a hedge against the uncertainty. This hedging is exacerbated when carriers have to submit prices and capacity commitments for future services. Because the carriers' estimates are dependent on the

accuracy and detail of the shipper's estimates, an additional hedge against poor forecasts is added by the carriers.

If a shipper can significantly increase the probability of securing a (good) follow-on load for a carrier at the completion of a delivery, the carrier's total costs will be reduced. By identifying opportunities for reducing this uncertainty, the shipper can lower the carrier's costs and, in turn, reap a benefit in the form of lower transportation rates.

Shippers can assist TL carriers in achieving economies of scope by allowing all-or-nothing packages of lanes to be submitted as conditional bids. By containing the interdependencies of traffic lanes within a set of lanes, the shipper reduces the carrier's need to hedge against uncertainty. Whereas shifting market power enables one party to exploit the other, the use of conditional lane bids can reduce overall transportation costs by removing a significant portion of the uncertainty from the process. How these lane packages should be formed, what the real value of a package is, what factors influence package value, how packages of lanes affect the bidding process, and how to determine which package is superior to another when lanes overlap are all open research questions that this thesis addresses.

This thesis makes three contributions related to the procurement of TL transportation services. First, the need for contracts in the seemingly perfect spot market for TL services is explained due to the carrier's underlying economies of scope. Second, the method for competitive bidding for TL services is formalized and a framework is proposed incorporating conditional bids. Third, these concepts are applied to actual data and are shown to be useful in analyzing the TL network.

The remainder of the chapter is organized as follows. Section 1.2 introduces the fundamentals of transportation systems and explains why TL carriers exhibit economies of scope and why this is important to shippers. Section 1.3 reviews the transportation procurement process as performed by shippers in practice and discussed in the literature. Finally, Section 1.4 outlines the remainder of the thesis.

1.2 Transportation System Fundamentals

This section introduces the fundamental concepts common to all transportation systems. First, the core activities that transportation firms must perform are presented. Second, the distinction between direct and consolidated transportation modes is made. Finally, the underlying economics of direct and consolidated transportation modes are explored and the importance to shippers is discussed.

1.2.1 Core Activities

The objective of any transportation system is to deliver items from one location to another. An item is the basic unit or thing that a shipper wants to send. A shipment is a collection of items with a common origin and destination that travel together all the way from origin to destination. A load is a collection of shipments that are moving in the same vehicle for a portion, but not necessarily the entire length, of a their total trip. There are four core activities inherent in each delivery: (1) local movement, (2) line-haul movement, (3) loading/unloading, and (4) sorting. Additional activities that do not fit into these four core activities are referred to as accessories. The purpose of this discussion is to provide background on the transportation process in order to be able to distinguish and classify different systems. A more complete discussion and analysis of the costs involved in logistics and transportation systems can be found in Daganzo (1991).

Local Movement

A local movement is a vehicle trip which contains intermediate stops for pick up and delivery and typically begins and ends at the same location. Local movements are usually made over defined and regular routes under set schedules. The factors that influence the cost of performing local movements are: (1) the number of stops to be made in the trip, (2) the processing time required at each stop, (3) the distance required to move between each stop, and (4) the distance between the terminal and the zone containing these stops. The processing time at each stop is largely influenced by the number and handling characteristics of the items in the shipments because the freight must be combined with other shipments on the vehicle. Consequently, the density, fragility, weight, shape, and overall stowability of each shipment affects the processing time and, thus, the cost of the local movement.

Line-Haul Movement

A line-haul movement is a vehicle trip with no intermediate stops which begins and ends at different locations. A line-haul can be made over a regularly scheduled route, such as a standard run between two terminals, or an irregular route, such as from a plant to a specific customer location. The factors influencing the cost of performing line-hauls are: (1) the distance of the trip, (2) the processing time at origin and destination, and (3) the cost of balancing the equipment.

Distance is the most influential cost factor because the major operating costs of trucking (fuel, tires, equipment wear and tear, and driver wages) are all variable with distance. Any costs associated with processing at the origin or destination act as a fixed cost for the line-haul

movement. Thus, line-hauls exhibit decreasing unit costs (cost per unit of distance) as a function of distance.

The second cost factor, processing time (mainly due to loading and unloading), affects costs by increasing the driver's time per trip. In order to decrease these costs, line-haul operations are generally performed using articulated vehicles in which the power unit (tractor) is separable from the storage unit (trailer or container), permitting decoupling of the line-haul movement from the loading and unloading operations. Unlike local movements, the specific characteristics of the shipment generally do not affect the costs of a line-haul since the load is not handled or combined with other loads during the trip.

The third cost factor is equipment balancing. Because line-hauls do not begin and end at the same location, equipment (both power units and containers) will tend to collect at net sink locations while needed at net source locations. This requires moving empty equipment between locations. Balancing is required for both power and storage equipment and the movement is referred to as deadheads, empty moves, or repositioning. The net effect of equipment balancing is the introduction of location-specific costs to the line-haul operation and a significant amount of uncertainty as to exactly where the equipment will have to be placed in order to handle future shipments.

Handling

The handling activities for transportation primarily consist of the loading and unloading of shipments from the vehicles. The factors that influence the cost of loading and unloading are the number and physical characteristics of the shipments in the vehicle. The more "standard" the items, the easier they are to handle and the lower the cost involved. For example, items on pallets can be moved by forklift as opposed to "broken" or partial pallet of loose items which need to be moved by hand. Increased containerization, however, tends to require more intensive capital rather than labor investment. For some forms of intermodal transportation, for example, overhead cranes are required to unload the shipping containers from the railroad's flatcars onto the motor carrier's flatbed trailers.

Sorting

Sorting involves the collecting, reordering, sequencing, and categorizing of shipments and can be performed at any stage in the transportation process. In-vehicle sorting of shipments during a local movement serves as a pre-sort for the main sorting which is done at the terminals. Sorting at terminals range from elaborate mechanized systems to a person tossing items into different bins. Essentially, the activity splits the incoming freight and reclassifies it according to

its next destination (zone or region), and places it into like-destined vehicles. As with handling, the more standardized the shipment, the easier, and faster, the sorting can be performed.

Accessorials

Accessorials are those services required by the shipper that do not fall into one of the four core transportation activities. Some of the more standard accessorial charges include:

Detention / Additional Wait Time - A carrier is required by a shipper to wait beyond a reasonable time for a load at the shipper's location. The definition of 'reasonable' differs between shippers but is generally about 4 hours.

Intermediate Stops, Split Pickups/Deliveries, or Stop-Offs - A carrier in a line-haul movement is required to make an intermediate stop to pick up or deliver a portion of the shipment.

Escort Vehicles - A shipper has a load which requires additional escort vehicles for safety or security reasons.

Extra Drivers - A shipper requests that an additional driver is used to provide faster transit time.

Oversized/Overweight Loads - A load from a shipper exceed the standard size and/or weight restrictions of the normal shipments.

Pickup and Delivery Outside of Business Hours - A shipper requests early morning, late evening, or weekend deliveries.

Inside Delivery - A shipper requests that a load be delivered beyond the traditional dock, into the shipper's premises.

Pallet Return - A shipper requests that all pallets sent in the headhaul loads are returned.

Storage - A shipper requests that a carrier pick up a load immediately but delay actual delivery until a certain time - forcing the carrier to store the shipment.

Accessorials create problems for several reasons. First, the definitions differ from one shipper to another. For example, an inside delivery for one shipper might be placing the load just past the shipping dock door, while another could require going up 3 flights of stairs. Second, the accessorials seldom apply to every shipment that moves on that lane, so the charge cannot always be added to a base rate. For example, just 10% of the shipments between A and B might require a weekend delivery. Third, carriers have different costs for the various accessorial activities so the price is not always standard. A carrier with an extensive amount of oversized trailer equipment would not need to charge as much for its use as a carrier that would need to lease special equipment.

Accessorials are primarily a practical concern affecting the pricing, bidding, auditing, and carrier assignment decisions. They do not affect the theoretical or fundamental aspects of transportation and therefore will not be addressed further in this section.

1.2.2 Classification of Transportation Systems

Transportation systems can be classified by the type and degree of consolidation strategies used. Consolidation is the process of combining different items or shipments into loads on a single vehicle. The extent of consolidation performed is a prime determinant of how the system makes trade-offs between level of service (time) and cost. Consolidation is considered a strategy because the types of consolidation used by a system dictate virtually all other aspects of a carrier's operations. Several different forms of consolidation have been identified and classified by Sheffi (1986), Hall (1987), and others. The three that are most pertinent for this discussion are time, vehicle, and terminal consolidation.

Time consolidation is where a shipper (or carrier) holds items at a location until a sufficient volume is accumulated before arranging for movement. The items do not necessarily need to have the same destinations. For example, an office will typically arrange for a single pickup of express mail at the close of each business day rather than scheduling pickups throughout the day whenever a letter is ready to be sent. Similarly, if items which have the same destination can be combined (say several items can be placed on a single pallet rather than sent as a "broken" pallet), then the cost of shipping will be lower.

Vehicle consolidation is where a single vehicle makes multiple stops to pickup (deliver) shipments from shippers that have the same destination (origin). The multiple shipments are combined into a single load in the vehicle itself. Examples of this include the postal service or the express mail carriers (FedEx, UPS, etc.) which perform regularly scheduled pickup and delivery runs. Local movements are an example of vehicle consolidation.

Terminal consolidation is where shipments from multiple origins are sent to a common facility where they are unloaded, sorted, and reloaded into vehicles heading to common destinations. The next movement out of a terminal need not be to each item's final destination, however. A shipment may be routed through multiple terminals in its path to its final destination. The hub and spoke system used in the majority of passenger airline operations is an example of terminal consolidation. People arrive at the origin airport via some mode (car, bus, transit) and "sort" themselves into their proper gates for the flight out to a hub. At the hub terminal, the passengers unload and "sort" themselves to the gate for the next leg of their trip.

Different transportation systems use different types of consolidation strategies. Indeed, some may use more than one and others do not use any.¹ Transportation systems can, of course, use multiple types of consolidation. The hub-and-spoke airline system described above, for example, uses both *time* and *terminal consolidation*.

¹ Some could argue, correctly, that driving on roads is a form of consolidation because the movements are channeled onto a single path and drivers share the fixed costs of the infrastructure. We will ignore such fine points in this thesis.

As with most strategies, there are trade-offs involved with consolidation decisions. The basic trade-off is between lower unit costs and faster delivery time. *Time consolidation* lowers the unit cost of the local and the line-haul movement by spreading out the fixed cost of the movement, but increases the time by delaying the initial departure or release date for a portion of the items.² *Vehicle consolidation* lowers unit costs by spreading out the fixed cost of the local movement among multiple shipments at the expense of additional travel time due to circuitous routing. Finally, *terminal consolidation* lowers unit costs by combining multiple shipments into a single line-haul, thus spreading out the fixed costs. This is done, however, at the expense of additional time and cost involved in handling, sorting, and circuitous travel to route items through terminals. *Terminal consolidation* also requires substantial investment in facilities which must be justified by sufficiently high volumes on the line-hauls connecting them.

Shipment size and level of service (transit time) are the critical factors influencing the degree of consolidation. The smaller the size of the shipment, the more consolidation required in order to achieve the lane density necessary to justify the costs of a line-haul movement. Shipments with tighter transit time requirements have fewer opportunities for consolidation and will tend to move by direct mode.

Direct versus Consolidated Systems

The classic examples of consolidated and direct transportation systems are busses and taxis. Busses operate according to a schedule over well-defined and regular routes which rely upon all three types of consolidation. *Time consolidation* occurs when passengers must wait at the stops until a scheduled pickup is made. The busses perform *vehicle consolidation* along their set routes both picking up and dropping off passengers. These routes generally lead to a terminal where the passengers sort themselves and either exit the system or load themselves onto other buses headed for different locations.

Taxis, on the other hand, are a direct form of transportation in that they move from point to point over irregular routes. A person usually arranges for a taxi in real time (hailing from the street) or with a small lead time (calling, say, five minutes out). The taxi goes to the person's point of origin and moves him directly to his destination; there is no vehicle or terminal consolidation.³ The movement is direct in terms of both distance and time.

² Somewhat paradoxically, *time consolidation* can sometimes lead to faster transit times. For example, by delaying the departure of a shipment, additional items might become available which increase the shipment size to the point where it justifies bypassing a break terminal and moving directly to its destination, therefore actually gaining time compared to the original alternative!

³ Actually, vehicle consolidation in the form of a milk run could be made where a "shipment" of multiple "items" (people) move from a common origin to different destinations with the taxi dropping off each passenger at their own location.

The distinctions between direct and consolidated transportation systems can be explained in terms of time and costs. Bus passengers spend time waiting for the scheduled bus arrival while taxi users must wait only as long as it takes for a taxi to reposition itself. Busses move in circuitous routes which might require transferring to additional busses at terminals while taxis move directly from origin to destination. Bus fare, however, is a fraction of taxi fare and is usually not distance-related. Taxi fares, on the other hand, typically include a fixed cost, a mileage rate, and additional charges for accessories, such as wait time, tolls, airport fees, or special help with packages.

In the motor carrier industry, the major distinctions in consolidation strategies occur between LTL (consolidated) and the TL (direct) carriers. The next two sections discuss the operations of each.

LTL: Consolidated, Regular-Route Carriers

Less-than-truckload carriers operate like the busses in the previous example. They are consolidated, regular-route carriers that operate over networks as shown in Figure 1.1. Shippers with freight to be moved either call or arrange for regularly scheduled pickups. A local truck visits the shipper's location and loads the shipment(s). The truck then visits the remainder of the stops on the local movement. Once the tour is completed, the vehicle returns to the end of line (EOL) terminal where it is unloaded. The shipments are consolidated onto a line-haul truck for the movement to a consolidation terminal where the packages are unloaded, sorted, and reloaded onto vehicles heading to other EOL terminals or continuing on longer line-hauls to more distant terminals. This line-haul may be made by rail or, for package delivery carriers, by air. The operations are reversed at the other end of the network; the vehicle is unloaded, sorted into movements to other EOL terminals, transported to these terminals, unloaded and sorted into appropriate tours for the local deliveries, and delivered to the final destination.

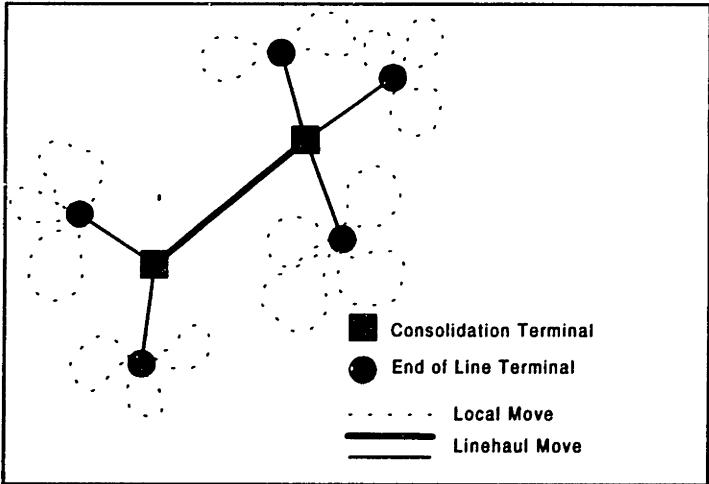


Figure 1.1 Schematic of consolidated transportation operations.

The preceding description is of a generic LTL operation. In practice, however, LTL carriers deviate from this generic description in the routing of certain shipments. LTL carriers will move directly from one EOL terminal to another if there is a sufficient volume; for some carriers 70% is sufficient, others require 90%. This is simply a form of time consolidation where the benefits of shipping directly are balanced by the size of the load needed to justify the cost of the direct line-haul. Another strategy, which appears to be based upon the same principle, but in reality is not, is sending a direct shipment (EOL to EOL) at a set time regardless of its volume. Bypassing a consolidation terminal decreases the transit time and lowers the exposure for loss and damage which usually occurs at the terminals. Even with increased direct movements, most long-haul, national LTL shipments go through 3 to 4 terminals in a trip. LTL carriers also differ in terms of the number and type of terminals they operate, the average length of haul, and the geographic regions over which they operate.

TL: Direct, Irregular Route Carriers

Truckload (TL) carriers operate like taxis in that they perform direct line-hauls from origin to destination. As mentioned previously, a primary concern is finding acceptable follow-on loads. The follow-on load does not necessarily need to take the vehicle back to its origin. While some smaller carriers and private fleets operate on a headhaul-backhaul network, most larger carriers combine different movements into more complicated tours because a vehicle will rarely be able to secure a follow-on load precisely at the previous load's destination. Rather, the vehicle typically will have to deadhead a certain distance in order to secure a load. One of the main concerns for direct carriers, therefore, is to minimize the empty (unloaded) miles and dwell time (time the tractor and driver are left idle) spent at the connections between loads.

Figure 1.2 shows the chain of events associated with the connection between two loads in a time-space diagram.

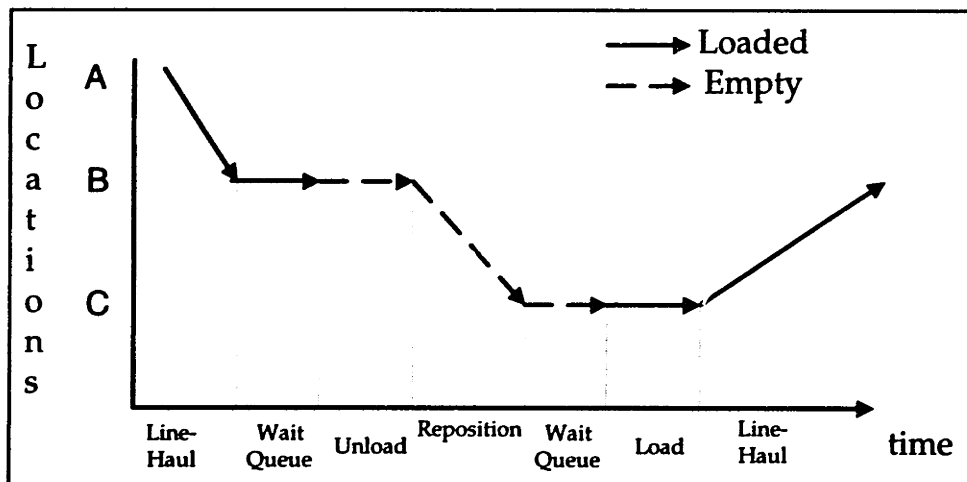


Figure 1.2 Schematic of the connection between line-hauls.

The vertical axis in Figure 1.2 represents discrete locations, in this case locations A, B, and C, while the horizontal axis represents time. The solid and dashed lines represent the movement of a vehicle through both time and space. The solid line indicates a paid (loaded) movement whereas a dashed line indicates unpaid moves (deadheads or waiting in place). The total time between successive line-haul moves consists of several components: waiting in a queue to be processed at the destination, unloading of the vehicle, repositioning to a new location, waiting in a queue to be processed at the next origin, and, finally, loading of the vehicle.

The carrier tries to reduce the time for these activities or to eliminate them altogether. For example, the unloading and loading components can be removed by using trailer pools. In a trailer pool, the shipper pre-loads trailers and has them waiting for pickup when the load is tendered to a carrier. Consequently, the carrier does not have a truck and driver sitting idle while the trailer is being loaded. Pooled operations reduce the dwell time and increase equipment utilization. Another way to reduce the dwell time is to improve the coordination between the shipper and the carrier, thereby lessening the time required for processing. Better scheduling can eliminate, or at least reduce, the time spent in a queue at a loading dock.

This thesis, however, focuses mainly on reducing the repositioning costs. The problem is that the follow-on load generally is not known in advance of the previous load's tender, and therefore, the pricing of a line-haul move must cover or hedge for the uncertainty involved in securing a follow-on load. Figure 1.3 illustrates the problem faced by carriers using a slightly different time-space diagram.

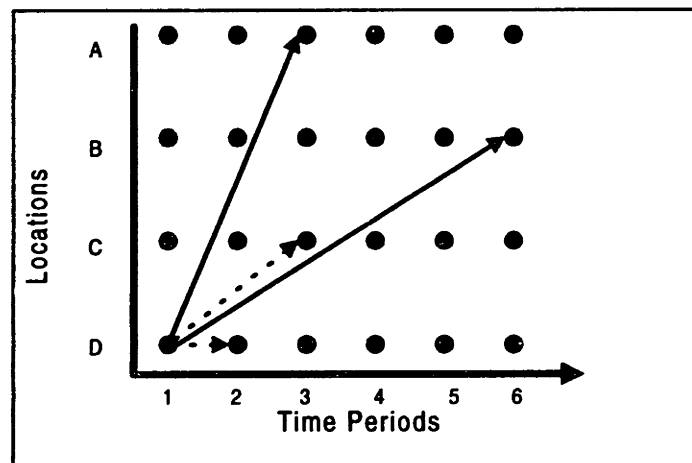


Figure 1.3 Time-space diagram for a single vehicle.

Again, the horizontal axis represents time while the vertical positions represent different locations. The figure depicts the opportunities for a single truck currently positioned at D at the start of time Period 1. At the start of time Period 1, the carrier has two possible loaded movements (to B arriving in Period 6 and to A arriving in Period 3) or the carrier may decide to

deadhead to a different location in order to pick up a load that might develop. Two possible deadheads are shown in Figure 1.3: moving to C by Period 3 or staying in place at location D until Period 2.

The carrier's dynamic fleet assignment problem, determining which load to take in order to maximize total profits, has been well-studied in the literature. Traditionally, carriers made decisions in order to minimize the empty miles accrued by accepting known loads. A better approach, described by Powell, Sheffi, and Thiriez (1984), Powell (1987), and Powell, Sheffi, Nickerson, Butterbaugh, and Atherton (1988), is to consider the profitability of each load not only in terms of the direct costs and revenues, but also in terms of the marginal contribution due to having an additional vehicle at the new point in time and space. The contribution of having a vehicle at this new point is never known with certainty since the potential loads have not been secured yet. Instead, this value is a probabilistic estimate of the expected earnings that a truck and trailer would generate if they were located at that destination location at the specific time. The expected contribution is based on the probability of a load being generated, and the direct profitability of that load. Thus, the value of repositioning is considered on its own merits rather than being arbitrarily split between other loaded moves.

By reducing the uncertainty in finding a load from a location, the carrier has in effect increased the contribution of having a vehicle at that location. This thesis addresses how the carrier can accomplish this by introducing conditional bids which let the carrier take a more active hand in designing their service network.

1.2.3 Carrier Economics

The objective of this section is to investigate how a shipper can influence a carrier's cost structure. In most vendor-buyer relationships, the only influence the buyer has over the vendor's economics is the volume of the business offered.⁴ The reason for this is that the production of most products and services can be described quite accurately as single output processes. Transportation, however, produces a multi-dimensional output. Jara Diaz (1981) defines the output of a transportation system as a 4 dimensional flow vector, Y_{ij}^k , describing the flow of commodity k between origin destination pair $i-j$ at time period t . Assuming that the timing and commodities are fixed, the shipper can control two things: (1) the total volume of traffic it offers to each carrier and (2) the placement of the traffic on the network that it offers to each carrier.

The shipper is interested in determining whether the carriers' costs are subadditive, that is:

⁴ While the timing of the deliveries will influence the costs, the volume of the total order affects the total costs of production.

$$C(Y) < \sum_{k=1}^K C(Y^k)$$

where:

- C(Y): Cost function for carrier producing output vector Y,
- Y: Vector of outputs such that $Y = \sum_k Y^k$, and
- K: Number of producers.

In the transportation context, the presence of subadditivity means that it is more efficient for a single carrier to serve the traffic on the given lanes at the given volumes than any combination of two or more carriers. Jara Diaz (1983) shows that subadditivity implies two facets: (1) increasing the total volume *proportionally across all lanes* (economies of scale) and (2) combining different lanes together as separate "output bundles" (economies of scope). These correspond to the two aspects that the shipper can control.

To help illustrate these concepts, the simple transport system shown in Figure 1.4 is used as an example, where:

- X_i : Number of shipments on lane i ,
- $C(X_1, X_2)$: Total cost of a carrier to serving X_1 shipments on lane 1, and
- λ : Scale parameter.

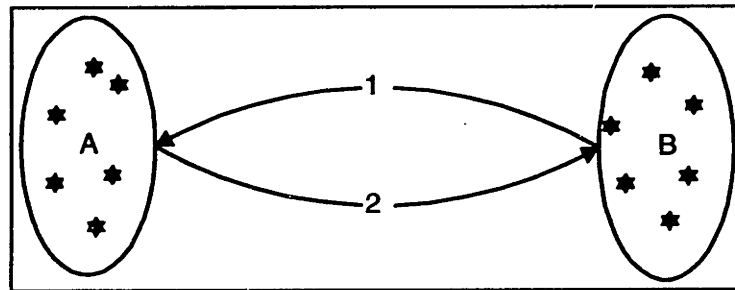


Figure 1.4 Example network illustrating different transport economies.

Economies of scale are present if the unit costs for serving over a network decrease if the volume on all of the lanes increases *in the same proportion*, that is:

$$C(\lambda X_1, \lambda X_2) < \lambda C(X_1, X_2)$$

For example, suppose that the monthly flow on the network in Figure 1.4 is 30 shipments for lane 1 and 10 for lane 2 and that a TL carrier handles 50% of the traffic on each. If the carrier exhibits economies of scale, then the cost per shipment would decrease if the carrier were assigned 100% of the shipper's traffic on these two lanes. Economies to scale presuppose that some fixed cost can be allocated over a larger base, thus lowering the unit costs. Because TL carriers have very low fixed costs and are more sensitive to the balance of the loads, they tend to have slight diseconomies to scale. In the example, doubling the flow would only serve to increase the carrier's load imbalance.

Economies of scope are present if the total cost of a single carrier serving a set of lanes is lower than the cost would be if multiple carriers served these lanes. It should be noted that, technically, economies of scope only capture the effect of an all-or-nothing assignment to a lane:

$$C_A(X_1, X_2) < C_A(X_1, 0) + C_B(0, X_2)$$

rather than the level of traffic on each lane:

$$C_A(X_1, X_2) < C_A(X_{1A}, X_{2A}) + C_B(X_{1B}, X_{2B})$$

where

$$X_1 = X_{1A} + X_{1B} \quad X_2 = X_{2A} + X_{2B}$$

The latter relationship is the definition of subadditivity for the network in Figure 1.4. Economies of scope, then, imply strict orthogonal subadditivity (Jara Diaz 1983, 431). The effect that an additional unit of flow on a specific lane has on the cost of serving that lane (or some other lane) is referred to as cost complementarity. While there are technical distinctions between economies of scope and cost complementarity, this thesis will usually refer to the placement of volume on specific lanes as economies of scope.

In addition to economies of scale and scope, a carrier's costs are also influenced by the *density* of shipments, both per customer location and per area. Economies of Density are present if increasing either the customer density or the number of shipments per customer location, while holding the total system volume (number of shipments) constant, results in the reduction of unit delivery costs. That is,

| | |
|-------------------------------|------------------------------------|
| Economies of Customer Density | $\partial c / \partial \delta < 0$ |
| Economies of Shipment Density | $\partial c / \partial \gamma < 0$ |

where:

- c: Average cost per shipment,
- δ : Average customer location density (number of customers per area), and
- γ : Average shipment density (number of shipments per customer location).

Economies of density apply primarily to consolidated carriers. For the network shown in Figure 1.4, economies of density exist if the unit cost per shipment of a carrier serving location A decreases if, holding the total number of shipments and the area constant, either (1) the number of shipments each customer makes increases or (2) the number of customer locations decreases. The costs per shipment decrease in these two cases because regardless of the number of shipments to be picked up at a location⁵, the cost to get there is unchanged and can be thought of as a fixed cost for that shipment location. The marginal cost of loading an additional package at a location is small and serves to allocate the delivery's fixed costs over a larger base.

⁵ Assuming that the capacity of the vehicle is not exceeded.

Thus, economies of density for a consolidated carrier are essentially economies of scale for their local movement.

There is a considerable literature investigating whether motor carriers actually exhibit economies of scale, density, or scope. Studies specifically looking for economies of scale have overwhelmingly found that the motor carrier industry as a whole shows constant or slightly decreasing returns to scale.⁶ The lack of economies of scale was actually used by many experts to justify that deregulation of the motor carrier industry would not result in natural monopolies. Unfortunately, practice did not match the theory's predictions as many trucking companies (mainly LTL) merged immediately after deregulation legislation was passed. Jara Diaz (1981) argues that the spatial and commodity aggregation of transportation output into a single dimension (ton-miles) were the cause of most of these inaccurate predictions. Studies looking for economies of density in motor carrier operations have had greater success. Friedlaender and Spady (1981), Caves and Christiansen (1988), Keaton (1993), and Kiesling (1995) found economies of density to exist in trucking, various consolidated transportation modes, LTL trucking, and package delivery systems, respectively. Finally, Jara Diaz (1981, 1982b, 1983, and 1988) demonstrates that economies of scope are exhibited by railroads and trucking firms.

TL carriers exhibit significant economies of scope. This is because the cost of serving an origin-destination pair is strongly affected by the probability of finding a follow-on load out of that destination location. Reducing the uncertainty in connection costs can lower the carrier's overall costs. LTL carriers, on the other hand, exhibit strong economies of both customer and shipment density. Increasing the density of shipment locations or the concentration of shipments per location lowers the unit costs of the local movement as it is amortized over a larger base of shipments.

Different procurement strategies arise from these observations. For LTL shipments, a shipper should try to concentrate its business with as few carriers as possible. Splitting the LTL shipments from a single origin between different LTL carriers reduces the density and increases the unit costs for each LTL carrier. In fact, setting aside industry-particular concerns (such as union carriers), there is no reason a shipper should have more than one LTL carrier from a common geographic location, assuming that that carrier has the sufficient coverage and service levels.

For TL shipments, a shipper has a much richer set of choices available. Using a single TL carrier is rarely the best option since different portions of the network are better suited to different carriers. Ideally, a shipper should try to award or assign those combinations of lanes

⁶ See, for example, Chiang and Friedlaender (1984); Daughety, Nelson, and Vigdor (1985); and Friedlaender and Spady (1980). Grimm, Corsi, and Jarrell (1989) contains a summary of the literature.

to carriers that best match that carrier's own cost structure. Unfortunately, the shipper has no way of knowing which lanes and in which combinations each carrier should optimally be assigned. Therefore, the shipper needs to provide the carriers with a method of determining between themselves which carriers should receive which combination of lanes. This thesis addresses the problem by proposing the use of conditional bids within a combinatorial auction framework. This thesis examines the implications and implementation issues associated with using conditional bids.

1.3 Truckload Procurement Problem

This section discusses the state of TL procurement in practice and briefly reviews the relevant literature.

1.3.1 Procurement Model

In practice, most shippers follow a general five step procedure to procure transportation services, as illustrated in Figure 1.5. The steps are arranged in order of the type of decision being made, from strategic (carrier screening) to tactical (information exchange and carrier assignment) to operational (load tendering and performance review). While shippers differ in the details of the individual steps, the overall process is robust.

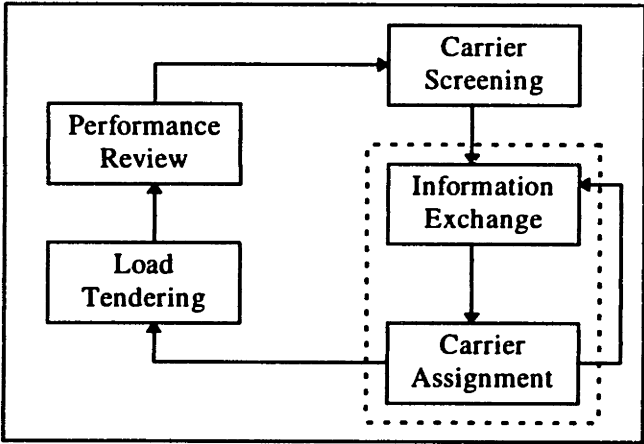


Figure 1.5 Five-step transportation procurement process.

The process begins with the strategic decision of selecting the carrier base. This typically involves the shipper using some sort of screening mechanism to reduce the number of potential carriers to a "reasonable", and hopefully higher quality, subset. The most common criteria for screening carriers from further consideration are financial stability, geographic coverage, electronic data interchange (EDI) capability, and equipment availability. Carrier screening serves two purposes. First, it reduces the complexity and cost of the final selection process by making a first cut on the candidate carriers from thousands to hundreds, or even just

dozens. Second, the initial screening of carriers allows the shipper to ensure at least a minimum quality level in the carrier base.

The next step is an information exchange. In this step, the shipper provides the reduced set of carriers specific details on the nature of the network being bid out. The content, amount, and level of detail of the information passed differs between shippers. For example, some companies give the carriers daily records of historical freight movements while other firms provide only total tonnage moved within the network. The lack of standard and useful network information is a major concern for carriers since they need to develop bids and, sometimes, volume commitments, based on this information. The exchange of information and prices between shippers and carriers may be an iterative procedure.

The next step is for the shipper to assign carriers to the network. Typically, the shipper selects carriers for each lane based on the submitted bid price for hauling freight on each individual traffic lane. This lane-by-lane assignment is performed either by a competitive bidding process for the purpose of forming a contract, or in the spot market. The end result is that carriers are assigned the right to haul freight over a lane, either as the sole carrier or with others.

The final two steps, load tendering and performance review, are operational in nature. Load tendering is the process of selecting which carrier to use for each load as it becomes ready to ship. While the assignment process provides some guidelines as to which carrier to call for certain lanes, there is often need for real-time choices to be made between alternative carriers. The final step in the process is a review of each carrier's performance. This consists of tracking carrier refusal rates, on-time rates, and other measures and assessing whether it meets established standards. As Figure 1.5, shows, the performance review can suggest corrective action at the carrier screening decision, such as to add a new carrier or remove one from the system. Thus, the entire five step process is a closed loop.

1.3.2 Review of Literature

The literature related to the transportation procurement problem consists primarily of descriptive studies identifying best practices among shippers. The object of most of that research is to understand the shippers' cost/service trade-off and to identify the criteria used for selecting carriers. That is, it mainly concentrates on the first step in the procurement process as outlined in Section 1.2.1. This section briefly reviews the literature.

In a survey of shippers, Bardi, Bagchi, and Raghunathan (1989) find the top four determinants in carrier selection to be transit time reliability, door-to-door transport costs, door-to-door transit time, and willingness to negotiate rate changes. The effect of deregulation is seen to have made carriers more responsive to both rate and service negotiations and more interested

in core carrier programs. McGinnis (1989) applied mode choice models to the selection of carriers of the same mode. He found that while freight rates were the primary criteria for selecting carriers to add to a core carrier base, service (defined as reliability, total transit time, and damage rate) was of almost equal importance.

In a set of surveys, Gibson, Sink, and Mundy (1993) and Gibson, Mundy, and Sink (1995) found the primary reason why shippers use core carrier programs is to ensure conformance to standards and reduce the carrier base. Reduced costs were a distant fourth choice. They also found that the most prominent criteria used is a measure of willingness to meet service expectations and a history of outstanding performance. Again, cost did not figure in the screening criteria. Moskal (1989) suggests that core carriers should have: (1) a strong financial position, (2) well maintained and market responsive equipment, (3) a sufficient quantity of quality drivers, (4) equipment density in required markets, and (5) electronic data interchange (EDI) capabilities. Gentry (1991), in a study examining how an organization's purchasing function's involvement in all phases of transportation, found that the predominate factor used to select carriers for any core program is on-time delivery followed by the rates, and geographic coverage, and time in-transit.

Abshire and Premeaux (1991) and Foster and Strasser (1991) investigate the 'shipper/carrier paradox' where carriers and shippers have differing opinions of which criteria are important. Shippers state that they place service over cost while carriers perceive that cost is still the primary criterion. Baker (1984) found that the acceptance criteria for adding new carriers differed from that used to eliminate carriers. Adding a new carrier required significant service and price superiority over the existing carrier base, but very little to eliminate.

Finally, La Londe and Cooper (1989) conducted a large survey of both shippers and carriers. They found that the most critical selection factors were equipment availability; consistency/reliability of pickups, deliveries, and transit time; speed of transit; and availability of information technology, such as EDI capability and shipment tracing. These differed from the criteria used for ongoing evaluation of carrier's performance which stressed flexibility and responsiveness to problems.

These studies clearly established the importance of selecting carriers based on both service expectations and low cost. Unfortunately, they do not address how shippers should assign carriers within the distribution network. In this respect, these studies treat transportation as any other input to a production process. The impact of carrier economics and network characteristics on the procurement and carrier selection decision has not been covered in the literature.

The only paper which specifically addresses the assignment decision down to the network level and considers the carrier's economics is Moore, Warmke, and Gorban (1991),

referred to as MWG. The paper describes the development and implementation of a sophisticated transportation management system at Reynolds Metal Company. The thrust of the work was on the load tendering step in the procurement process. The objective was to install a centralized dispatching system so that Reynolds could achieve continuous moves in real time. A continuous move is the matching of a carrier on an inbound load with an outbound load from the same location. A continuous move reduces a carrier's costs by eliminating the need for finding a follow-on load. MWG developed a real time matching heuristic which would identify potential incoming loads for the tendering of upcoming loads from the same location.

In order to support the real time load matching, MWG developed an integer program (IP) optimization model to provide an initial assignment of carriers to locations. The model considered carrier capacity, required commitments at shipping locations, and the total number of carriers. Carriers were asked to submit bids for both single moves (the hauling of a load from point A to point B) as well as continuous move rates (a declining mileage rate based on the distance of the total tour). Only the single move rates, however, were used in the assignment optimization model. A follow-on simulation model was used to suggest a final assignment. A more detailed analysis of the MWG model is included in Chapter 6.

MWG is the only paper to take an optimization based approach towards carrier assignment. It did not, however, consider the carrier's economics when actually assigning carriers within the network - only in the tendering process afterwards. Lanes were still bid on as independent items. While the assignment minimized total system costs by taking other implications into consideration, it did not allow carriers to select bundles of lanes in order to secure their economies of scope. In this respect, our research is taking a much more fundamental approach to the problem than MWG.

1.4 Outline of Thesis

The objective of the thesis is to analyze how shippers procure transportation services from truckload motor carriers. We refer to this as the truckload procurement problem. Starting from a fundamental view of the economics that drive motor carrier transportation, we investigate the contractual, bidding, and assignment decisions that a shipper must make in order to secure sufficient TL trucking services. The research approach is outlined in Figure 1.6 and the thesis is organized as follows.

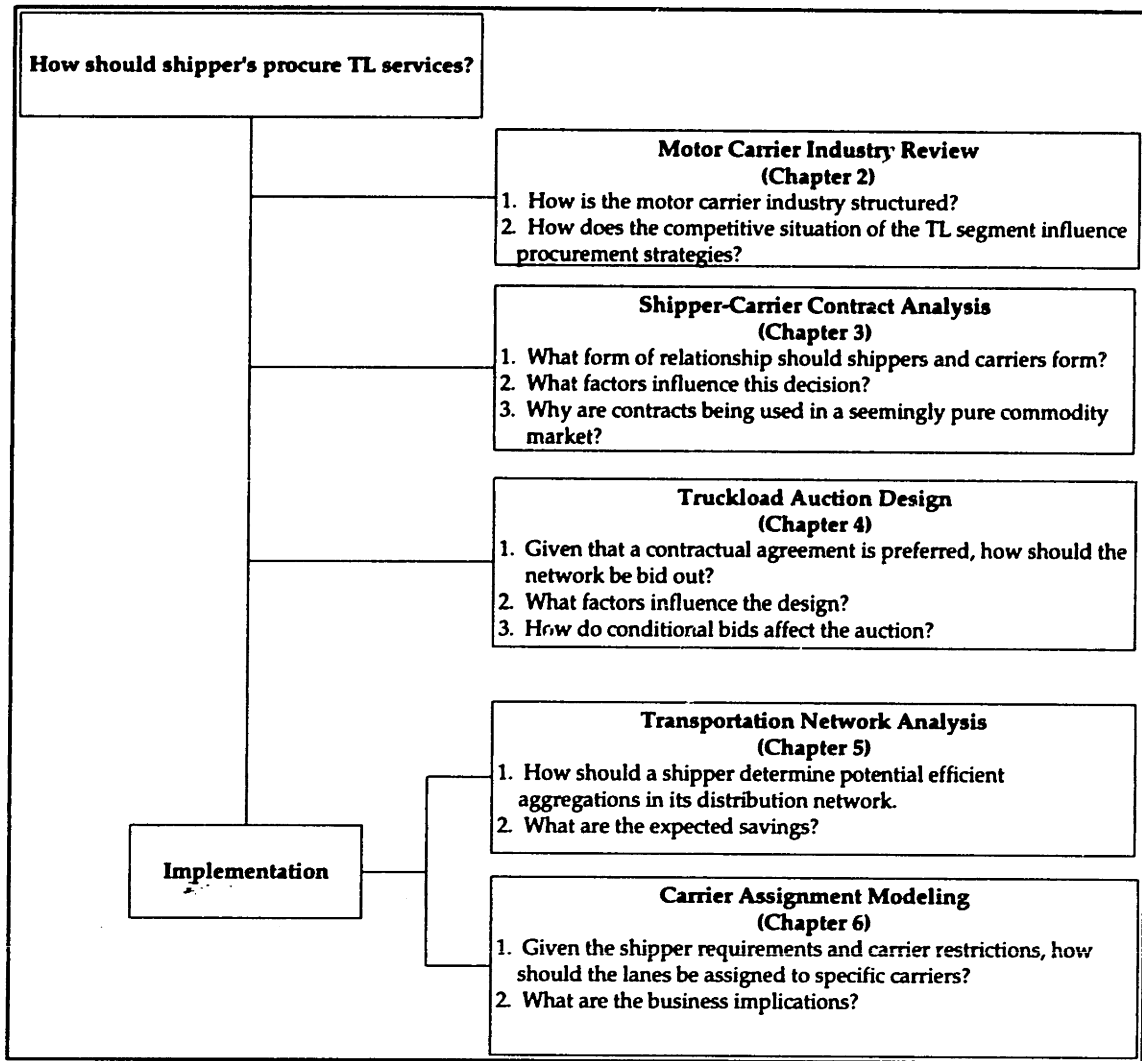


Figure 1.6 Overview of thesis.

Chapter 2 presents an overview of the motor carrier industry. The competitive situation between motor carrier segments and within the TL sector are discussed and analyzed. The importance of the industry structure in developing procurement strategies is also explained.

Chapter 3 explores the contractual aspects of shipper-carrier relationships. Different forms of relationships are discussed, including spot markets, contractual agreements, and private fleets, and each is analyzed in terms of its use and relevance for TL trucking. The main contribution of this chapter is to explain the growing importance of contractual agreements between shippers and TL carriers in a seemingly pure spot market environment. It is shown that contracts are needed for the carrier to achieve economies of scope. Additionally, the limitations of electronic markets for TL carriers are discussed.

Chapter 4 addresses how shippers should assign truckload carriers to the traffic lanes in their distribution network. The chapter reviews auction design theory and discusses the need for and the difficulties associated with bidding on both individual lanes and on packages of

lanes using conditional bids. The primary contribution of this chapter is to synthesize the large auction design theory literature and adapt it to an analysis of the TL bidding problem. Additionally, the effect that carrier hedging has on the bidding process and the inadequacy of lane-by-lane bidding methods are demonstrated.

Chapter 5 details the pre-bid network analysis needed to identify potential spatial aggregations within distribution networks. Using data from different firms, we develop a series of algorithms which locate different forms of efficient aggregations within a shippers network. The savings potential of each type of aggregation is approximated. The contribution of this chapter is to formalize the analysis of a shipper's distribution network specifically with TL carrier economics in mind. Also, a methodology to estimate the probability of load matching within a shipper's network is developed and implemented.

Chapter 6 presents the assignment model used to award traffic lanes to carriers. We develop an integer programming model which incorporates shipper and carrier considerations, to include conditional bids, limits on the number of carriers in the system or at a location, the use of alternates and primaries within a system, and overall system coverage. Methods of incorporating service capabilities and accessorials in the bidding process and analyze the business implications of different carrier assignments are also developed. Formulation issues and solution strategies are discussed and compared. The main contribution of this chapter is to demonstrate the practicality and usefulness of optimization-based carrier assignment for large and contentious traffic networks.

Finally, Chapter 7 provides a synopsis of the thesis, summarizes the major contributions, and suggests extensions to this research.

Chapter 2

Motor Carrier Industry Overview

This chapter reviews the motor carrier industry in general, and the truckload trucking segment in particular. The objective is to describe the competitive forces acting on TL carriers and discuss the strategies that they are following. The relevance to the procurement problem is addressed throughout.

In 1994 the motor carrier industry accounted for \$300 billion (78%) of the nation's total freight bill of \$385 billion (Standard and Poor's 1995). This compares to 9% for railroads, 6% for water carriers, 3% for pipelines, and 4% for air freight carriers. Motor carriers are usually divided into three operational segments based on the range of the shipment sizes they handle and the technologies used: truckload (handling large loads in excess of 10,000 pounds), less-than-truckload (handling shipments between 150 and 10,000 pounds), and package delivery (handling shipments less than 150 pounds).

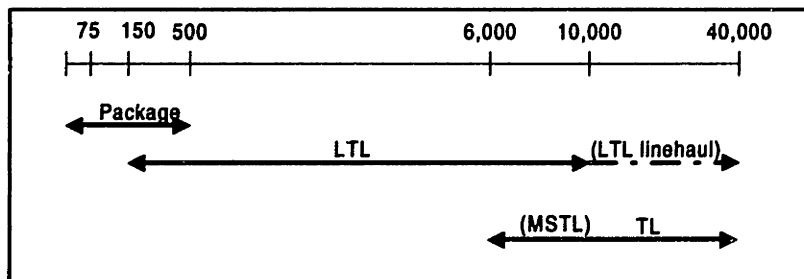


Figure 2.1 Trucking industry segmentation based on shipment size (pounds).

These shipment size ranges represent the markets in which the different trucking segments operate. The boundaries, however, are constantly shifting due to competitive pressures. TL carriers are accepting loads as low as 6,000 pounds and combining them with one or two similar sized shipments. At the other end of the spectrum, the package delivery carriers are constantly increasing their upper weight limits, typically restricted by union work rules, in order to capture some of the traditional LTL market. LTL carriers are being squeezed in the middle, but are expanding aggressively into both the small package and full truckload markets.

The net effect of all of this competition is lower prices to the shippers and no protected market for any of the motor carrier segments.

While the TL, LTL, and parcel package segments are the most commonly referred to, they represent less than a third of the total motor carrier revenues. The remaining two thirds (~\$205 billion in 1994) is accounted for by the approximately 48,000 private fleets operating in the United States.⁷ A private fleet is a fleet of vehicles operated by a shipper primarily for the movement of its own goods. The previously described TL, LTL, and package delivery segments are for-hire carriers. A for-hire carrier is a company whose primary business is the transportation of goods for other companies for a fee. Figure 2.2 breaks down the total motor carrier industry and reports the estimated revenues along with the approximate size of each segment.

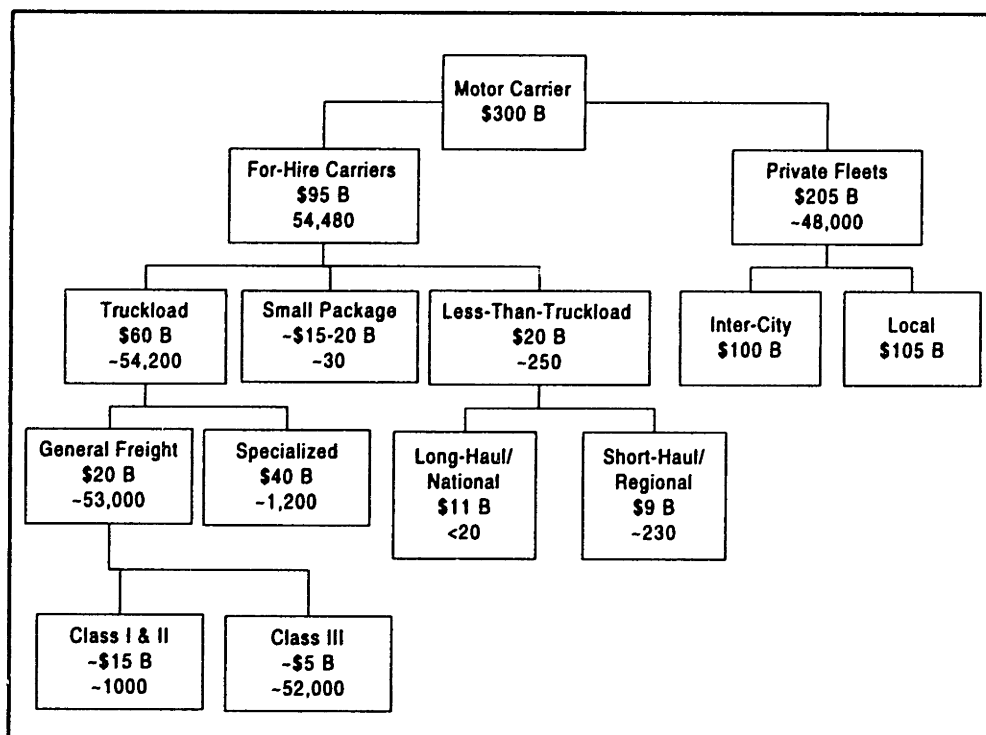


Figure 2.2 Breakdown of 1994 revenues and market size for motor carrier industry segments. Sources: Standard and Poor's (1995) and American Trucking Association (1994).

The ICC categorizes motor carriers as being Class I, II, or III based on annual revenues. Up until 1981, Class I carriers earned more than \$3 million, Class II earned between \$0.5 and \$3 million, and Class III earned less than \$0.5 million. In 1981 these breakpoints were changed to \$5 and \$1 million and in 1994 they were changed again to \$10 and \$3 million.

⁷ Due to the dearth of information on some of the segments in the motor carrier industry, it is difficult to get accurate numbers and revenue figures. Unless noted otherwise, the reported values throughout this chapter came from Standard and Poor's (1995) and American Trucking Association (1994).

The chapter is organized as follows. Section 2.1 discusses regulatory and demand factors which are affecting the motor carrier industry. Section 2.2 examines the TL segment in detail focusing on procurement implications. Section 2.3 briefly discusses the other motor carrier modes and, finally, Section 2.4 closes the chapter.

2.1 Regulatory and Demand Factors

This section discusses the regulatory and demand environment that the motor carrier industry faces. The emphasis is on identifying how these external factors are influencing TL procurement strategies for shippers.

2.1.1 Regulatory Factors

Three major acts of deregulation have had a large impact on the trucking industry. These are the Motor Carrier Act (MCA) of 1980 which deregulated the interstate trucking industry, the Trucking Industry Regulatory Reform Act (TIRRA) of 1994 which eliminated the filed rate doctrine, and the Airline Improvement Act (AIA) of 1995 which deregulated intra-state trucking.

Motor Carrier Act (MCA) of 1980

Prior to the Motor Carrier Act (MCA) of 1980, trucking firms were required to obtain authorization for hauling by both commodity and route.⁸ The process to obtain these authorizations was both costly and time consuming. Additionally, shippers with private fleets were not allowed to haul other shippers' freight and for-hire carriers were restricted to being either contract or common carriers. Common carriers could offer transportation to the general public but were required to charge the same 'reasonable' rates (tariffs) to customers with similar freight, that is, they could not discriminate between shippers. Contract carriers, on the other hand, could serve specific customers but were not allowed to carry general freight from other shippers. Also, the number of customers a contract carrier was allowed to serve was limited to eight (the rule of eight). The net effect of these rules, was the protection of the existing motor carriers through extensive barriers to entry.

The MCA deregulated the inter-state motor carrier industry primarily by removing many of these barriers. Specifically, the authorization process was liberalized to include only insurance coverage and safety standards; private fleets were granted authority to haul additional freight; for-hire motor carriers were allowed to operate dually as common and

⁸ For example, carriers were sometimes granted authority in only one direction, so that hauling from A to B was allowed, but the backhaul from B to A was not necessarily permitted.

contract carriers; and the "rule of eight" was removed. Contract carriers were now allowed to set individual rates for specific shippers as long as the rates were filed at the ICC.

Upon enactment of the MCA, there was an almost immediate entrance of small, entrepreneurial, and primarily non-union carriers into the marketplace. The number of carriers registered with the ICC rose from 16,874 in 1980 to 54,480 in 1994. While the total number of carriers dramatically increased, the growth was not uniform across all industry segments. In fact, the explosive growth masked an opposite trend. As shown in Figure 2.3, while the number of Class III carriers increased by over 400% from 1980 until 1994, the number of Class I and II carriers actually decreased by 51%.

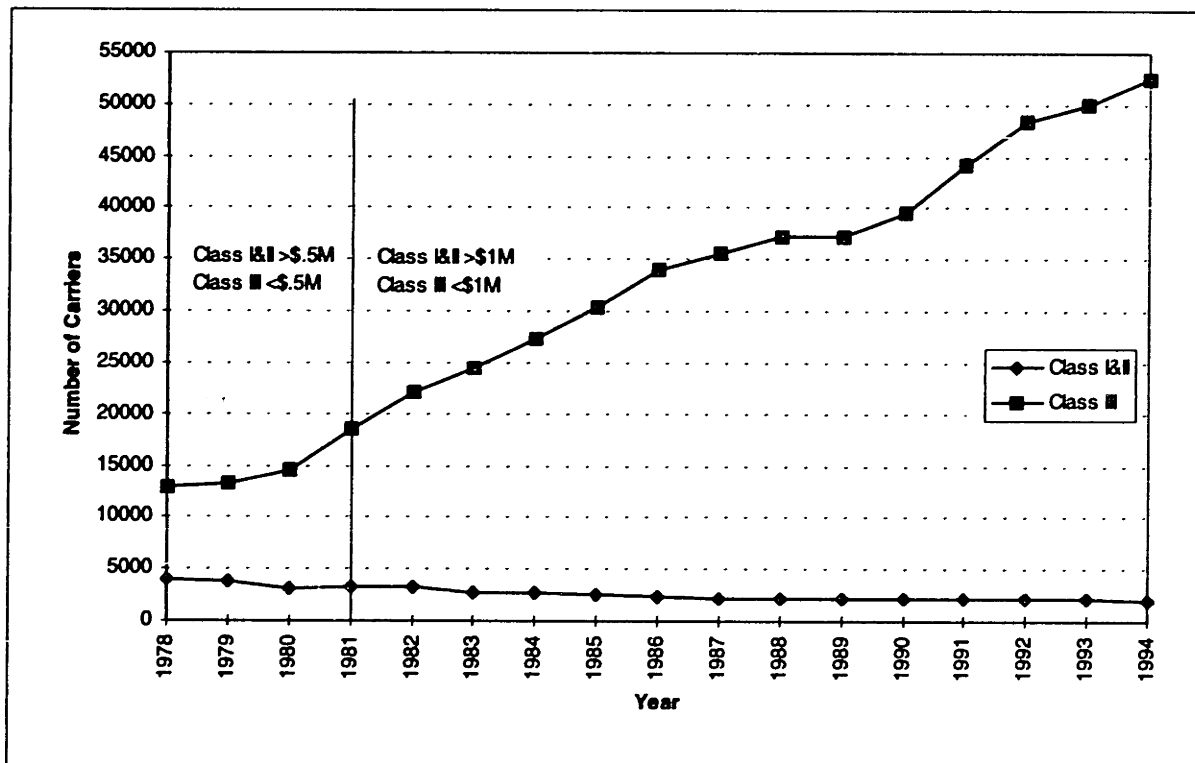


Figure 2.3. Growth of trucking market since deregulation. Note the redefinition of the classes starting in 1981. Source: American Trucking Association (1994).

As of 1994, there were approximately 2,580 Class I and Class II carriers and 51,900 Class III carriers. Virtually all of the new entrants to the industry since the passage of the MCA have been TL carriers (due to the low capital costs of direct transportation operations). The exiting Class I and II carriers, on the other hand, were a combination of both TL (general freight and specialty) and LTL carriers. Most of the reduction in the Class I and II carriers was through mergers and bankruptcies among LTL carriers.

The major effect of deregulation on transportation procurement was to reintroduce competitive forces to an industry which had been protected for close to 50 years. Existing carriers now had to compete with not only these new entrants, but also with other established

carriers ready to expand into other carriers' geographic markets. The general response of carriers across the board was to cut rates. While an apparent boon to the shippers, these rate wars lead to decreased profit margins and bankruptcies for many carriers. Adjusted for inflation, the estimated revenue per ton for LTL carriers decreased by 24% from 1980 to 1994 (Standard and Poor's 1995, R41). Finally, the distinction between common and contract carriers became irrelevant so that more carriers were allowed to enter into contractual agreements with shippers. This increased the importance of negotiation in the transportation procurement process.

Trucking Industry Regulatory Reform Act (TIRRA) of 1994

The Trucking Industry Regulatory Reform Act (TIRRA), enacted in August 1994, effectively repealed the long-standing filed rate doctrine which had required carriers to file all tariffs (rates) with the ICC at least one day prior to their enactment. A separate filing was required each time a new discount was offered to a shipper resulting in more than a million tariffs filed annually. While the tariffs were technically supposed to be available for all shippers, the criteria for most tariffs were specifically tailored for an individual customer so as to make it fit that shipper's freight alone. This tailoring effectively defeated the purpose of a tariff filing and essentially made it a spot rate rather than a common rate.

An even more troubling aspect of the filed rate doctrine was the fact that a tariff was not considered valid unless it was filed at the ICC - even if it was agreed upon by both the shipper and the carrier as part of a contract. It was the carrier's responsibility to file these rates but oftentimes, due to mismanagement or neglect, it was not done. This rule became important in 1990 when lawyers representing creditors of bankrupt carriers successfully argued that the applied rates were invalid if the same rate had not been properly filed at the ICC. Essentially, they argued, the shippers had been undercharged and should have to repay the difference between the applied rate and the higher (and technically valid) filed rate. The Supreme Court held in *Maislin vs. Primary Steel* that shippers were responsible for paying these undercharge claims which typically covered thousands of individual shipments going back several years. The total amount of undercharges that shippers paid is estimated to have exceeded \$2 billion (Standard and Poor's 1995, R42).

While the immediate effect of the passage of the TIRRA on transportation procurement was the containment and ultimate reining in of the undercharge issue, the larger and more lasting implication is the simplification and increased flexibility that carriers could now employ in setting rates. By eliminating the need to file any rates at all, the TIRRA placed the responsibility of enforcement solely within the shipper-carrier contract. Additionally, by

removing the public filing requirement, carriers enjoyed greater confidentiality for the rates they offered to different shippers.

Airline Improvement Act (AIA) of 1995

While the MCA deregulated inter-state motor carrier transportation, intra-state transportation was still highly regulated. As of 1994, forty one states still regulated motor carriers in terms of authorized routes, rates, and operating authorities. As was the case nationally prior to 1980, these states tended to have inflated rates and entrenched and inefficient carriers. For example, shipping jeans within Texas from El Paso to Dallas cost 40% more than transporting the identical shipment from Taiwan to Dallas. In California, shipping 2,000 pounds the 15 miles from Oakland to San Francisco cost as much as shipping the 200 miles from Reno, NV to San Francisco (Harper 1994). The Airline Improvement Act (AIA) of 1995 deregulated the intra-state transportation.

The main intent of the AIA was to remove the air carriers' exemption from the intra-state regulations for the ground movement portion of their networks. Because they did not move all of their freight by ground, air freight carriers were not considered motor carriers and were, thus, exempt from the intra-state regulation. This allowed them to price their ground transportation services below that of the competing LTL carriers. The LTL industry successfully removed this exemption and, as a side benefit for other national carriers, eliminated rate, route, and service regulations for all intra-state transportation.

2.1.2 Shipper Trends

In addition to the regulatory environment, transportation, as a derived demand, is very sensitive to changes in how shippers operate. Since the mid-1980's, firms have become much leaner in their operations which has affected how motor carrier transportation is being used. Four trends present in the mid-1990s are discussed: the increase in time-definite freight, the increased frequency and smaller size of shipments, closer proximity of suppliers to plants, and the widespread use of core carrier programs.

Increase in Time-Definite Freight

Time-definite freight is defined as any shipment that is required to arrive within very tight time windows, sometimes as short as 10 minutes. Penalties can be assessed for both late and early arrivals. This is the case in many just-in-time (JIT) manufacturing environments where the buffer inventory has been reduced to the point where any transportation delay has the potential to stop the production line. There have been cases where a production line was stopped at a plant because a JIT delivery truck broke down on the highway.

Time-definite freight increases the minimum required service standards expected of the motor carriers. The transactional cost of the carrier, as captured in the rate, is dwarfed by the potential cost of a line shutdown due to a late arrival. The cost of poor service is both transparent and very high. Carriers that are able to provide reliable, on-time deliveries can increase their prices and work to secure long-term contracts which can shield the carrier from the commodity market of general freight.

Time-definite freight will also tend to move on direct transportation since routing freight through terminals for increased consolidation and increased equipment utilization, will add to the potential of a delivery delay. As noted above, the cost of a potential delay is typically much larger than any incremental savings due to consolidation.

Smaller, More Frequent Shipments

Related to the increase of time-definite freight is the trend towards smaller and more frequent shipments. This is a result of JIT initiatives where the objective of the replenishment system is the continuous flow of material into a plant. While this would seem to suggest an increase in LTL use, the trend is actually for an increase in direct movements - which benefits TL carriers. TL carriers are also becoming better at coordinating multi-stop truckload runs which consolidate two to three loads onto a single trailer for delivery at several stops. Regional LTL carriers which operate networks with more direct movements are also benefiting from this trend by avoiding the more time consuming hub-and-spoke moves through multiple terminals.

Suppliers Locate Closer to Plants

A third trend, also related to leaner production techniques, is the relocation of suppliers to within a day's travel (<500 miles) of their primary customer's plants. Close proximity reduces the lead time required for both expedited and standard shipments. The result of these relocations is the rapid growth in the short-haul or regional markets. Table 2.1 shows that over 70% of TL revenues are generated by loads less than 500 miles in length. Because shippers are moving their suppliers closer to their plants, the length of haul is decreasing for TL carriers. This means that the competitiveness of the short-haul market is increasing.

| Length of Haul | <100 Miles | 100 to <500 | 500 to <1000 | 1000 to <1500 | 1500 to <2000 | 2000 to <2500 | >2500 |
|----------------|------------|-------------|--------------|---------------|---------------|---------------|-------|
| % TL Revenues | 45.2% | 26.7% | 10.1% | 7.8% | 5.4% | 3.6% | 1.2% |

Table 2.1. Revenue versus length of haul for TL. Source: Alex Brown & Sons (May 1995).

Increase in Core Carrier Programs

The fourth, and most important, trend resulting from leaner production, is the increase in the number of core carrier programs instituted by shippers. By reducing the number of carriers that they do business with, shippers are able to offer larger and more reliable volumes of traffic to carriers. In return, carriers are expected to offer lower rates. Additionally, core carrier programs simplify a shipper's transportation management, guarantee capacity, and allow shippers to achieve wider coverage through fewer carriers. Coverage is actually a form of *network externality*⁹ where the value of a service conducted over a network increases as the expected total number of locations served increases.

The justification for reducing the number of carriers that a shipper uses is based on the quality movement in the late 1980s which emphasized having fewer suppliers held to more rigorous standards. In transportation this translated into the widespread use of core carriers, a leading force in hastening the consolidation of the truckload segment. In a 1989 survey of large manufacturers, Crum and Allen (1990) found that 57% of respondents reported a decrease in their carrier base from 1984 to 1988 with the average number of carriers used dropping from 158 to 108.

The importance of core carrier programs for TL carriers is that they allow for differentiation from the commodity market. That is, core carrier programs allow motor carriers to compete on service and not solely on price. By forming closer alliances with a particular set of shippers, the carrier also raises the shipper's cost of switching carriers. For TL carriers, core carrier programs allow an opportunity to secure more stable and reliable loads.

2.2 Truckload (TL) Carriers

While the TL segment is often treated as a homogenous industry, it is actually quite differentiated both by size and by specialty. For example, over 90% of the total TL revenue was earned by fewer than 4% of the carriers in 1994. The regulatory and demand factors discussed in the previous section can have very different impacts on carriers of different size and focus. This section discusses the different types of TL carriers, identifies four tiers of carriers based on operational capabilities and competitive vulnerabilities, examines the growing concentration of the TL market, and, finally, comments on the overall competitive situation that TL carriers face.

2.2.1 Specialty versus General Freight Carriers

TL carriers can be divided into general freight and specialty carriers. General freight carriers are those with the majority of their traffic volume moving by dry van. A dry van is the

⁹ This is also referred to as having a positive consumption externality, see Economides (1994).

standard closed trailer that is 45, 48, or 53 feet in length. A specialty carrier moves the majority of its volume in non-standard trailers, such as temperature control vans, flatbeds, tank trucks, household good vans, hazardous material haulers, and heavy haulers. Additionally, there are extremely specialized trailers for specific industries, such as auto-haulers, plate glass vans, and steel haulers. While most analysis considers the general freight market to be the TL market, the specialized carrier market actually has twice the revenues (\$40 billion) with a fraction of the number of carriers.

The specialty carriers are still considered TL carriers since they move directly from point of origin to point of destination without terminal or vehicle consolidation. While many carriers specialize in these niche markets, most of the very large general freight carriers also contain some specialized equipment, usually temperature controlled vans and flat bed trailers.

Specialized carriers generally work within niche markets where the competition can be quite fierce. The average revenue per mile for a specialized carrier is higher than that of the "average" general freight carrier, but the costs are also higher. While the problem of balance is faced by both types of carriers, the empty mileage for specialized carriers (even for temperature control carriers) is almost always higher than general freight carriers. The more specialized the type of equipment, the more difficult it is to find acceptable follow-on loads. The specialized carriers tend to work in one-directional lanes and it is very important to find some sort of acceptable load for at least partially covering the backhaul costs. This tends to depress the rates in some backhaul lanes for the general freight carriers since these specialized carriers have higher fixed costs and will take almost any load.

2.2.2 Carrier Tiers

While the underlying economics for TL carriers are the same for general freight as for specialty carriers, the general freight carriers face the greater threat of being "commoditized." A dry van trailer is almost the perfect description of a commodity product. With more than 50,000 participants, the general freight TL segment appears to be a purely competitive market where carriers compete solely on price while profits go to zero. However, as Figure 2.3 showed, the TL market seems to be dividing into two markets: Class III carriers and the Class I and II carriers.

Rakowski, Southern, and Jarrell (1993) suggest that this dichotomy has created two extreme markets. The first consists of a very large number of very small carriers that act like the perfectly competitive market where there is frequent entry and exit, profit margins are very thin, and competition is almost exclusively price based. The second market consists of the larger firms that have responded by differentiating themselves by their service offerings.

While valid, this two-market view tends to hide some key dynamics working within the TL industry. There is actually a continuum of carrier sizes from the 32,000 Owner/Operators

(OO) with one or two trucks to the super-carriers with 10,000 or more tractors. In between these extremes lie over 20,000 carriers of differing fleet size. Unfortunately, there is little information available regarding carriers not at the extremes.

In order to get a better understanding of the dynamics of the industry, we have identified four general groups or tiers of carriers with different characteristics, operating patterns, and strategies. Table 2.2 lists the approximate number of carriers and fleet size for each category. These are each discussed in turn.

| Category | Approx. Number | Approx. Fleet Size |
|----------------------|----------------|--------------------|
| Owner/Operators | ~32,000 | 1-3 trucks |
| Lower Tier Carriers | ~20,000 | trucks <30 |
| Middle Tier Carriers | ~1,000 | 30 < trucks < 500 |
| Upper Tier Carriers | ~200 | >500 trucks |

Table 2.2. Four tiers of carriers in the truckload segment. Source: American Trucking Association (1994), Standard and Poor's (1995), and Griffin and Rodriguez (1992).¹⁰

Owner Operators

Owner/Operators (OOs) represent the smallest type of carrier within the trucking industry. Recent surveys by the Owner Operator Independent Drivers Association, Inc. (OOIDA) and the Upper Great Plains Transportation Institute (UGPTI) found that OOs own an average of 1.4 trucks; tend to haul dry van trailers (36%), flatbeds (27%), or temperature control vans (18%); operate their trucks slightly over 100,000 miles annually with an approximate empty mile percentage of just under 20%; and earn less than \$100,000 per tractor per year. In 1993, they spent an average of just under 200 nights on the road per year, operated in over 20 states on average, and spent as much on washing their truck as they did on lodging! The average annual compensation for OOs was \$27,800.

Rather than working directly for shippers, the vast majority of OOs lease themselves out to larger carriers, typically for a month at a time. Over 80% of the surveyed owner-operators reported that they generally lease to carriers rather than shippers. Less than a quarter of the respondents to the OOIDA survey reported that they trip lease - this is where a truck is leased for a single load. About half of the OOs report that they had not used brokers at all the previous year.

While working long hours, OOs are relatively unsophisticated in both financial management and fundamentals of TL economics.¹¹ OOs generally do not have strategic plans,

¹⁰ In addition to these published sources, the size of these tiers were estimated based on conversations with Owner Operator Independent Drivers Association officers, American Trucking Association analysts, transportation managers at brokers, and various other shippers and carriers.

and tend to operate on a cash flow basis. They are small businessmen who, in the words of Michael Porter, "are satisfied with a subnormal rate of return on their invested capital to maintain the independence of self ownership, whereas such returns are unacceptable and may appear irrational to a large publicly held competitor," (Porter 1980, 19). Some OOs work on a revenue threshold rather than a profit maximizing framework. One transportation manager for a large shipper noted that an initiative to improve service by paying owner-operators a higher rate backfired. With higher rates, the OOs were able to earn their threshold revenue level earlier in the month. These OOs simply stopped accepting loads and instead devoted their energies to other activities - such as fishing or side businesses. This made it more difficult for the shipper to find carriers at the end of each month.

While unsophisticated, the vast number of OOs serve as the slack in the TL system. As variable capacity for the TL industry, OOs allow the larger carriers to be more flexible in their load and capacity planning. Because they are more like inputs to the larger carriers rather than direct competitors, the OOs' willingness to work at less-than-sustainable rates does not depress the overall TL rate structure.

Lower Tier Carriers

The next group of carriers have fleets of less than, perhaps, 30 trucks. There are even fewer data on these small carriers than for owner-operators. The majority of this discussion is culled from conversations with brokers and TL carrier representatives. The strategy for many of these carriers is to operate in a headhaul-backhaul network. Small fleet carriers generally have one or two local customers around which their service network is centered. They are able to generate loads outbound from this home location, but then need to locate other loads back to their home base. This is primarily done through brokers. The brokers are used to "get them back home" rather than for some sort of floating through the nation from load to load. Freight brokers are discussed in Section 3.2.

Lower tier carriers can offer the small shippers that they serve extremely personalized service. These shippers are too small to attract the larger carriers, but are large enough to maintain a small carrier fleet. Lower tier carriers have a fairly secure market since they are willing to accept slightly lower rates of return than the larger carriers. They have very low overhead and make up for the lack of advanced communication technology with highly personalized service.

" For example, a newsgroup on the Internet for truckers (misc.transport.trucking) recently had the posting, "There is no reason that a load of lumber from Portland, OR to LA should pay \$1.30 per mile and a load of steel from LA to Portland should only pay \$.85. The drivers wages are the same coming [sic] back north, the cost of the fuel is the same, and so is the cost of the fuel and mileage taxes. It still takes a buck a mile to run the truck! If trucking companies would let this 'backhaul' freight sit and rot on the docks the rates would come up where they belong!!!", posted 13 July 1995.

Middle Tier Carriers

These carriers are the most vulnerable of the four tiers. This tier is comprised of those carriers with fleets between, say, 30 and 500 trucks. They are generally too large to be sustained by a small set of local shippers but too small to provide the national coverage that larger shippers are demanding. Their regional operating strategies and short haul markets have been invaded by both regional LTLs and the national TL carriers. While expected to compete with the large TL firms for spots in the core carrier programs, they are not able to generate the sufficient revenues to keep up with the necessary technology investments. These mid-range carriers are caught between a captive market of very small shippers and the service intensive shippers operating nationally or within a region for time-definite freight.

These middle tier carriers include both Class I and Class II carriers. They are more sophisticated than the smaller Class III carriers in that they do not operate solely over head-haul back-haul networks. They have more extended networks and include triangulation of loads and increased tours. While many of these carriers may operate efficiently, they do not provide the large national coverage that many of the larger shippers require for entry into their core carrier programs.

Upper Tier Carriers

Upper tier carriers have large fleets in excess of 500 tractors. These carriers generally have national coverage and sufficient size to achieve economies of scale in maintenance, driver training and retention, and communications and technology investments. Because of their size, they create network externalities which make them primary candidates for shippers' core carrier programs. Even within this group there are dramatic size distinctions, though. For example, the 1994 revenue of the largest carrier was four times larger than that of the fifth largest carrier.

Lane (1987) coined the term Advanced Truckload Firms (ATLFs) to describe the carriers able to provide shippers with high service levels and wide coverage. Corsi and Stowers (1991) narrowly define an ATLF as a truckload firm that concentrates primarily in "long-distance medium- to high-density corridors with balanced flows" and competes based on service rather than price. They state that one of the key characteristics of ATLFs is their strategy of not using long-term owner/operators because OOs are incompatible with "system wide efficiency and overall effectiveness." Defining a set of about a dozen carriers as being ATLFs, they found that empty mileage for these carriers was around 6 to 8 percent while traditional TL firms using owner/operators averaged in the 15% to 20% range. While portions of this description are true for most of the carriers within this upper tier, it paints a rather uniform portrait. Actually, there

are several different strategies being pursued. Three strategies are discussed below: reducing fixed costs, partnering with railroads, and entering niche markets.

Reduce Fixed Costs

The third largest TL carrier, LandStar successfully follows a strategy which goes completely against the grain of ATLFs. They have concentrated on keeping their costs as variable as possible in order to respond quickly to fluctuations in the market. Over 90% of LandStar's 10,000 tractors are owned and operated by OOs. Additionally, the sales, customer service, and dispatching staff are compensated primarily as a percentage of load revenue. The end result is that 82% of LandStar's cost structure is variable as compared to 40-60% for other TL carriers.

While Landstar has taken this strategy to the limit, many other upper tier carriers use OOs to some degree. The reasoning is the same - OOs provide flexible capacity. Because they can be leased for relatively short periods at a time (monthly is common), the carrier can adjust its total effective fleet quite quickly.

Partnership with Railroads

Another strategy being followed by many of the upper tier carriers, is forming partnerships with railroads to provide intermodal service. In this context, intermodal transportation refers to movements where the local transportation is provided by a trucking firm (drayage) and the line-haul is performed by rail. Railroads are the most inexpensive mode for long-hauls (>1,000 miles) while trucks can travel virtually anywhere. Using intermodal can cut a trucker's line-haul costs by 25% while trimming days off regular rail or truck transit times. Additionally, using railroads for the line-haul portion of the trip reduces the time away from home for many of the truck drivers. Driver turnover is an acute problem in the TL segment.

The type of the intermodal move is classified by the type of container and technology that is used to make this transfer. The most common forms of intermodal movements are trailer-on-flatcar (TOFC), container-on-flatcar, (COFC), double-stack, and RoadRailer. The different technologies provide different trade-offs. Double stack is much more efficient on the line-haul RR portion but requires a crane for the rail-to-truck transfers, as does COFC. Both COFC and double-stack use containers which require special trailers for the trucking portion. TOFC does not require a crane for unloading or special trailers, but it is less efficient on the line-haul since a single trailer is placed on each railcar. The RoadRailers are special trailers which can be used over the road but also have steel wheels which are lowered for rail movements. While slightly heavier than standard over-the-road trailers, RoadRailers are the easiest and

fastest to transfer from rail to road: The RoadRailers sacrifice utilization in the line-haul with extremely high efficiency in the yard.

While virtually all of the largest twenty TL carriers had some sort of an agreement with railroads for intermodal service by 1994, they did not all follow the same form or use the same technology. The two largest TL carriers, for example, are taking opposite tacks. JB Hunt, the pioneer of intermodal operations, has committed to using containers. Forming an alliance with Santa Fe, they began offering the "Quantum" intermodal service in 1990. Hunt has since formed agreements with eight other railroads in the US and Canada. They purchased 14,000 containers in the early 1990's and have invested over \$200 million in intermodal equipment. The goal is to have just 3,000 over-the-road drivers by the year 2000. As of 1994, 30% of Hunt's revenue was generated from intermodal movements.

Schneider has also gotten involved in intermodal operations, forming partnerships with six railroads. Schneider derived about 20% of its revenue from intermodal in 1994. While they do perform some double stack, COFC, and TOFC intermodal movements, they have begun to take a different tack by investing heavily in the RoadRailer systems. These allow for immediate drayage from the rail terminals and can operate in much smaller yards. Schneider purchased 500 of the RoadRailers in 1995 and is expected to procure more. Schneider has stated that their goal is to have approximately 20,000 drivers in their system by the year 2000 (Schulz 1995a).

Enter Niche Market

Another strategy followed by some successful carriers is to position themselves into a niche market based on commodity, route, service, or some other factor. For example, a handful of carriers have concentrated on that segment of the market deemed "dead" by intermodal - cross-country line-hauls. One of the biggest challenges for intermodal operations is the different standards that railroads and motor carriers use for on-time performance. Being on-time 80% of the time is considered high for railroads but is unacceptable for a TL carrier that is used to delivering 95%. The added variability of using railroads has opened a niche for carriers using team drivers to serve the cross country market. These carriers provide consistent service while shaving off at least a day from the intermodal transit time. Covenant, for example, focuses exclusively on long-haul lanes with an average length of haul of 1,800 miles. Two thirds of their drivers operate as teams so that loads move, for example, from Los Angeles to Atlanta in 2.2 days compared to 4 days for a solo driver and 6-7 days for intermodal. Their average tractor utilization is 150,000 miles per tractor per year with less than 6% empty miles.

Another niche market is dedicated transportation. This is where for-hire carriers replace a company's private fleets and completely dedicate selected tractor assets to a shipper. The largest player in this field is the Ryder Dedicated Logistics unit of Ryder System with more

than \$850 million in revenues in 1995 and an estimated 10,000 trucks. These dedicated contracts are being used more frequently for JIT and highly time-sensitive freight where the availability of capacity is of prime importance. Because it requires a high degree of coordination, the barriers to entry are increasing. The expertise in handling these operations creates a strong brand presence which can bolster their margins while the tight contractual ties with customers shields the carriers from soft cycles in demand.

2.2.3 Concentration in the TL Market

The reason for discussing the different tiers of carriers, is to understand how the TL market is evolving. The vast number of owner-operators will not disappear in the near future since they play a valuable role by providing variable capacity to the TL market. The lower tier carriers are also fairly secure in their smaller, geographically constrained markets. They are typically serving smaller shippers that do not generate the volumes needed to attract larger carriers. Likewise, these carriers provide a more intimate service component than the large brokers can provide to the small shippers. On the other end of the market, the upper tier carriers are able to provide both wide coverage and high service levels. Despite the different strategies that are being followed, all of the top tier carriers can leverage their size in lowering their administrative costs. Additionally, they can invest more in communication and information technologies that further improve their service capabilities.

The middle tier carriers, however, are vulnerable. They are too large to be sustained on small clusters of customers, but cannot provide the coverage or service to satisfy the larger shippers. By being excluded from core carrier programs, middle-tier carriers continue to lose volume to the benefit of the larger, upper tier carriers. Due to network externalities of TL transportation, carriers with larger service coverage are more desirable to shippers. Thus, the bigger carriers keep getting bigger while the middle tier carriers continue to get squeezed out.

This is fueling increased concentration in the general freight TL carrier segment. In 1994, the top ten TL carriers generated 25% of the total general freight TL revenues compared to 21% in 1993 and just 10% in 1990. Figure 2.4 shows the number of TL carriers required to capture different market share of the TL general freight segment. The biggest gains do not appear to be at the very top, where the top three carriers' combined market share has remained fairly constant, but at the 30%, 40% and 50% market share levels. Just 39 carriers accounted for 50% of the general freight TL segment in 1994 compared to 50 in 1993, and just over 60 in 1992.

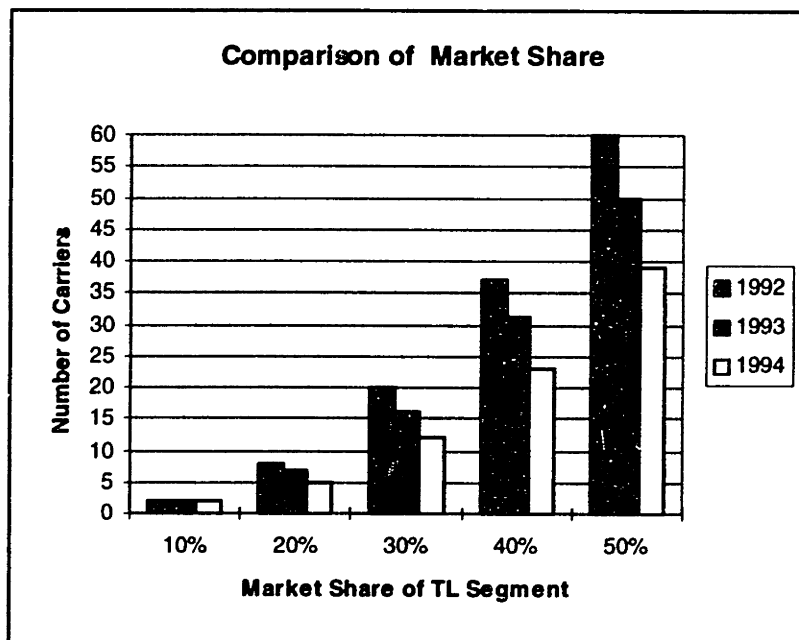


Figure 2.4 Concentration of general freight TL segment.
Source: Hoffman (1993, 1994, and 1995)

While the market is becoming more concentrated, it is still very far away from any anti-trust problems, thanks to the huge number of Class III TL carriers. The Herfindahl Index is sometimes used to determine the concentration of an industry. The index is defined as:

$$HI = 10,000 \sum_i S_i^2$$

where:

HI: Herfindahl Index and
S_i: Market Share of the ith firm.

A monopoly would have an index of 10,000 while a purely competitive market has a value of essentially zero. An index above 1800 is generally considered cause for concern of a possible concentration of the market. The HI for the TL segment is less than 150.¹² This index has doubled over the last two years, however. Still, even if the market was highly concentrated, the low entry barriers make this a very contestable market.¹³

2.2.4 Competitive Environment

The competitive environment for the TL market can be illustrated using Porter's structural analysis model. Figure 2.5 summarizes the analysis along the five major forces which drive an industry's profitability: (1) barriers to entry, (2) rivalry among competitors, (3) substitutes, and (4) buyer and (5) supplier bargaining power. This section provides background

¹² It is difficult to measure this exactly since the revenues for all TL carriers are not available. However, carriers below a market share below 0.5% add less than 0.25 to the HI. The 40th top carrier's revenue was less than a 0.5% market share.

¹³ A contestable market is where although there is not direct competition, the threat of entry is high enough so that the potential competition is enough to suppress prices to a competitive level.

to understand the various strategies carriers are taking and the impact that better transportation procurement practices might have on the industry.

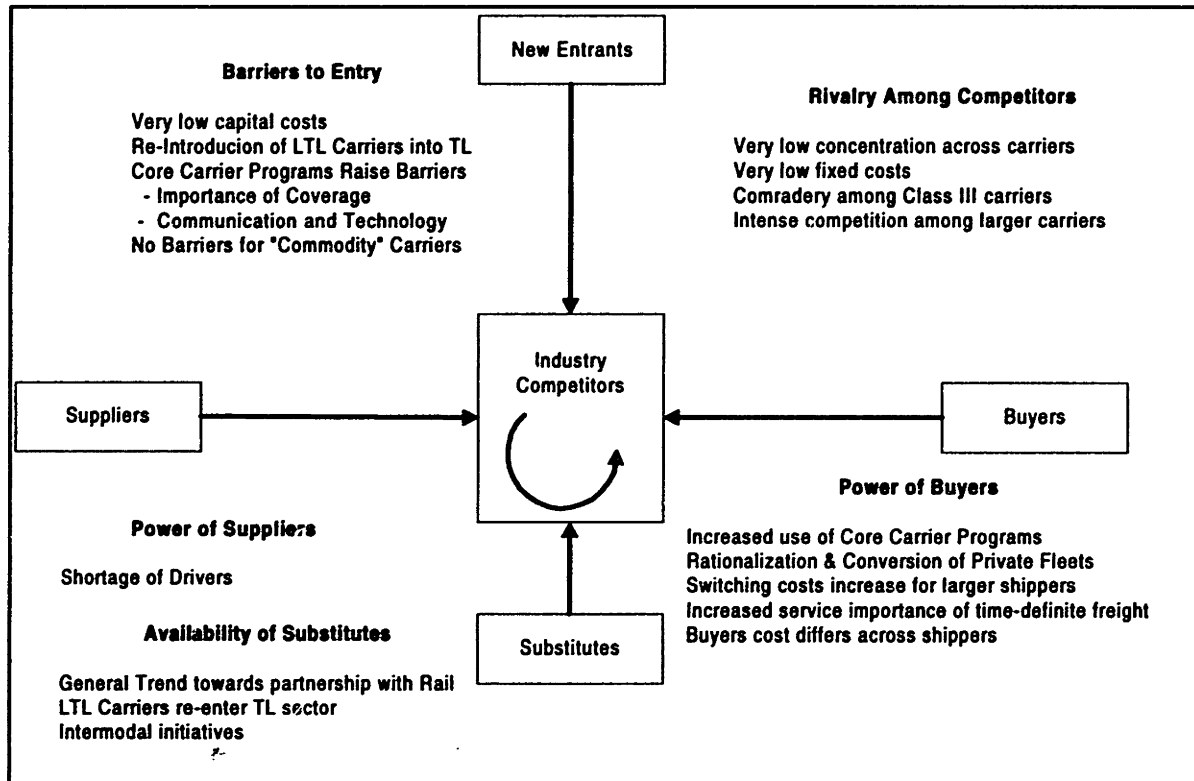


Figure 2.5 Overview of competitive analysis of TL trucking segment.

While there are many factors affecting the TL market, four are of special interest. First, the buyers (shippers) are becoming much more powerful. This is seen in the growing use of core carrier programs. As discussed previously in the chapter, these programs are placing an increasing percentage of freight into a decreasing percentage of carriers. Inclusion in a core carrier program is a way for a carrier to gain more stable and predictable flows. Because shippers are including additional requirements to join these programs, such as equipment coverage and technological capabilities, the barriers to entry are increasing.

The second factor is an increase in supplier power brought about by the driver shortage. This was a critical problem in 1994 when there were severe capacity problems. The next year, however, saw slower growth and an excess capacity in the TL sector. While the driver shortage was not as pressing during 1995, it appears to be a long term problem due to generally low wages, tough working conditions, a decreasing pool of potential drivers, and strict federal safety regulations. In response to this shortage, many carriers are using newly developed software packages which enable drivers to switch loads enroute allowing them to return to their home base while the load continues on to its final destination. Working in real time and connected to sophisticated tracking and communication technologies, these systems are reducing the number

of days a driver needs to spend away from home. Intermodalism and regional TL fleets are other ways of keeping drivers local.

The third factor is barrier to entry for the TL market. Actually, the threat of entry differs significantly between the upper and lower tiers. There are essentially no entry barriers to the OO or lower tiers. The rate of return at this level is insufficient to foster growth beyond these small levels, however. At the other end of the market, there are fairly high barriers. These derive, in large part, to the coverage and service requirements being placed on carriers by shippers. While the price of these technologies continue to decline as they become more widespread, the investment costs for communication and information technology is daunting for smaller carriers.

Finally, different modes can serve as substitutes for different portions of the traditional TL segment. Intermodalism, for example, has made the long-haul market very competitive. Railroads are the lowest cost form of transportation for these longer lanes, so TL carriers competing need to compete based on service differentiation rather than price. In the short-haul market, regional LTL carriers are providing more direct movements which compete directly with TL services. Additionally, many of the larger LTL carriers are moving back into the TL market, beyond just covering backhauls.

2.3 Other Motor Carrier Segments

This section briefly reviews the other motor carrier modes. The objective is to understand how they interact with and influence the procurement of TL transportation services.

2.3.1 Private Fleets

As shown previously in Figure 2.2, private carriage is, by far, the largest segment of the motor carrier freight industry capturing approximately two-thirds of the \$300 billion market in 1994. A total of 48,000 private fleets consisting of 2.7 million vehicles are currently operating in the United States in a wide variety of industries. These fleets are very diverse in terms of size, level of outsourcing, and usage. The number of vehicles in a fleet ranges from one truck to several hundred. The 1995 *Private Fleet Directory* shows that 86% of the fleets have less than 50 vehicles and just 11% have more than 100. AT Kearney (1993) found that 81% of the private fleets used company drivers, 68% outsourced the maintenance and 64% leased at least a portion of their equipment.

While it is widely accepted that private fleets are not as efficient as for-hire carriers, there is debate as to the degree of these inefficiencies. Efficiency is usually measured in terms of per mile costs and equipment utilization (the percentage of empty miles and the annual number

of miles per tractor). AT Kearney (1993) reports that, for 1993, private carriers had an average cost per mile of \$1.46, operated tractors an average of 100,100 miles, and averaged 24.4% empty miles. For-hire carriers' average cost per mile ranges between \$0.90 to \$1.10 with tractors hauling between 100,000 and 120,000 miles annually with between 5% to 15% empty miles. Direct comparisons are difficult to make since the values are culled from several reports and studies of differing sample sizes and levels of accuracy. However, even the president of the National Private Truck Council has conceded that shippers should outsource their transportation needs in situations where private fleets cannot compete in terms of cost and service (Melbin 1995).

Private fleets have never been more efficient than for-hire carriers, even during regulated times. Miller (1973) and ICC (1977) showed that empty miles for private carriers were 150% to 200% higher than for-hire carriers. Prior to deregulation, though, the inefficiencies of the private fleets were less than the inflated profit margins of the for-hire carriers. Since deregulation, though, the TL market has become more competitive and private carriers' inefficiencies have become more glaring.

Today, shippers justify their private fleets either by higher (reported) levels of service or as a "strategic advantage." The perceived service improvements arise from reduced loss and damage, increased reliability of shipment deliveries, and guaranteed truck capacity during peak seasonal periods. Additionally, a private fleet with uniformed company drivers, is thought to produce goodwill for a company and serve as a competitive edge in establishing a brand presence. The obvious example for this is the Frito-Lay fleet which has a strong tradition of drivers serving as on-site marketing representatives. Finally, shippers often maintain private trucks in order to provide leverage in negotiating transportation contracts with for-hire carriers.

A more cynical justification of private fleets, in light of their much lower efficiencies, is the protection of managerial turf. Private fleet managers are the most aware of the inefficiencies in private fleets but are not keen on removing their own position. The evaluation of the relevance of private fleets is often performed by the same executives that would be replaced if the fleet was found to be inefficient. Additionally, cost evaluations of private fleets often compare the internal costs of the high volume lanes to the rates on the low volume, poor quality lanes that are bid out to for-hire carriers. Extrapolating the cost of these poor quality lanes onto the entire network inflates the comparable for-hire carrier costs and makes the private fleet look more efficient than it actually is.

In the light of these inefficiencies, private fleet operations have come under greater scrutiny in the early 1990s. Alex. Brown & Sons (May 1995) estimate that \$15 billion of the private fleet market is open for conversion to for-hire TL carriers before the year 2005. While a

great deal of this conversion will be for total fleets, there should also be a move to rationalize and convert portions of the fleets - most likely the long haul segment suited for TL carriers.

2.3.2 Less-Than-Truckload (LTL) Carriers

In 1994 there were approximately 250 LTL carriers operating in the United States earning revenues of \$20 billion. The LTL segment consists of two types of carriers: nationals and regionals. National carriers are predominately long-haul and unionized while regional carriers are largely short-haul and non-unionized.

The national market is highly concentrated with the largest three national LTL carriers generating over \$7 billion in revenues in 1994, representing a third of total (national and regional) LTL carrier revenues.¹⁴ While national carriers earned higher revenues than the regional carriers, their operating ratios (defined as operating expenses divided by revenues) have been among the worst in the entire motor carrier industry during the 1990s. Additionally, regional LTL carrier revenue growth increased by 14% in 1994 while national LTL decreased by 1.3% (Schulz 1995b).

The national carriers tend to operate in large hub and spoke networks requiring substantial investment in terminals and sorting facilities. The number of terminals for some LTL carriers reached over 700 in the late 1980's. By the mid-1990s, the number of terminals had been reduced for most of the nationals to between 300 to 400. The national carriers' inefficiencies are primarily due to a heavy reliance on consolidation (for better equipment utilization) while the market has moved to more time-sensitive freight. Additionally, the national carriers have very high union labor wages and restrictive work rules. Wages constitute 60-70% of total costs for unionized LTL carriers (Alex Brown & Sons 1995b).

Regional LTL carriers tend to operate over short haul lanes, typically less than 300 miles in length. Rather than operating over hub and spoke networks, they send more shipments directly between end of line terminals or with, perhaps, one intermediate sorting terminal. By sending more direct shipments, regional carriers lower handling and sorting costs but increase the line-haul costs due to lower utilization. By providing direct transportation, the regional carriers are attracting shippers of time-definite freight which tends to have higher margins. In 1994, the margins for regionals were about 6% higher than nationals - due to both lower labor rates and operational efficiencies in shorter hauls.

Because of the increased competition from the regional LTL carriers, the nationals are shifting to more direct shipments and relying less on their elaborate hub-and-spoke systems, discounting heavily to increase their lane densities, and forming alliances of non-union regional

¹⁴ This is lower than in previous years primarily due to a teamsters strike in the Spring of 1994.

carriers to parallel their unionized national network. Additionally, some national carriers are moving back into the TL segment using their non-union subsidiaries. Prior to deregulation, close to 60% of the traffic volume for most LTL firms was actually TL. This number was reduced to less than 20% in the early 1990's.

The regional LTL carriers, in turn, are forming alliances with other regional LTLs and TL carriers in order to increase their available coverage - the primary strength of the nationals. By aligning themselves with other regionals, these LTLs can form "super regionals" and, with combined markets, can achieve the coverage of a national, while maintaining the flexibility of a regional. Regional LTLs are also forming alliances with TL carriers in order to extend their service regions and provide better interline service to non-adjacent, allied regional LTLs. These alliances, while involving multiple carriers, appear seamless to shippers.

2.3.3 Package Delivery Carriers

The ground package delivery segment is still dominated in the mid 1990s by the carrier which founded it, the privately held United Parcel Service (UPS). UPS earned \$14 billion in revenues in 1994 with just under \$1 billion in net income. The only other major player in the ground package delivery market is Roadway Package System (RPS) which had revenues of \$1.2 billion in 1994 for an estimated net income of \$180 million. RPS has made inroads into UPS's market by concentrating on gaining market share on UPS's highest density lanes. Without the burden of providing service to all of the hinterlands that UPS serves, RPS can provide very competitive rates. The remainder of the players in the package delivery segment are smaller (less than \$100 million in revenues) regional carriers which provide niche services within well defined corridors. For example, one carrier provides next day parcel delivery within 600 miles for parcels dropped off before midnight (Schulz 1996).

Package delivery carriers are expanding into the LTL market by increasing the allowable shipment size. An additional threat to the ground parcel post carriers is the increased use of trucks by FedEx. Standard and Poor's (1995, R36) estimate that over half of FedEx's air freight never leaves the ground but instead travels via its growing truck fleet.

The net effect of the ground package carriers is that they are placing increased competition on LTL carriers. By squeezing LTL carriers at the bottom end they are providing incentives for the LTLs to compete elsewhere, namely in the TL segment.

2.4 Closure

The objective of this chapter was to provide general background on the motor carrier industry and a more detailed analysis of the TL segment.

We first examined how trends in the regulatory and demand environments have influenced motor carrier procurement. The combined effect of the deregulatory legislation passed in the 1980's and 1990's was shown to have given the carriers (1) greater flexibility in setting price/service options, (2) the ability to form specialized contracts, and (3) the ability to maintain confidentiality in agreements. On the demand side, the volume of time-sensitive freight was shown to be increasing in importance while shippers are reducing the size of their carrier bases through "core carrier programs." The net effect of these demand trends is a growing use of TL and regional LTL carriers and increased competition between carriers of the same mode to be included in shippers' core-carrier programs.

The analysis of the TL segment argued that the seemingly homogenous 50,000 carrier strong TL market is, in fact, quite diverse. There are definite differences in the operations, types of customers, and strategies followed by different TL carriers. These differences are primarily based on the size of the carriers' fleets and the type of service network that they can support. The Owner-Operators (<2 trucks) were shown to act primarily as variable capacity for the larger carriers while the next larger carriers (<30 trucks) generally serve small headhaul-backhaul networks. The concentration of the TL market was shown to be occurring primarily to the detriment of the middle-tier carriers (50 < trucks < 500). Finally, the upper tier carriers (>500 trucks) were shown to be using a variety of strategies all aimed at differentiating themselves from the rest of the TL market.

Even with this stratification of carriers, though, there are thousands of TL carriers that are competing for the same general-freight, dry-van market. The competition within the TL segment was shown to be rather intense - especially in the middle and upper tiers. It does not seem clear why shippers are choosing to restrict the number of carriers with which they do business. Core carrier programs require contractual agreements and are generally being initiated by shippers. It would seem that shippers should prefer to use the spot market to take advantage of the carrier's competition rather than be tied to a contract. This is discussed next.

Chapter 3

Shipper-Carrier Relationships

This chapter investigates how and why shippers and carriers use different forms of buyer-seller relationships. The main contribution of the chapter is to explain why a seemingly perfectly competitive market for general freight, dry van truckload trucking is becoming increasingly procured using long-term contracts. It is shown that contracts are necessary for a TL carrier to achieve economies of scope which in turn minimizes the transaction costs.

The chapter is organized as follows. Section 3.1 reviews transaction cost analysis (TCA) and economic organizational theory as they relate to TL transportation. Section 3.2 describes the current use of the different types of relationships formed between shippers and TL carriers. Section 3.3 proposes an explanation for the use of contracts for TL carriers. Finally, Section 3.4 closes the chapter with a summary of the major points.

3.1 Transaction Cost Analysis

Any exchange between two parties requires some sort of governance structure to guide the exchange. Williamson (1985) defines a governance structure as the mechanism used by buyers and sellers to ensure that an exchange is carried out successfully. The governance structure guides the way in which the exchange is "initiated, negotiated, monitored, adapted, enforced, and terminated" (Palay 1984, 265). The three basic types of mechanisms used by buyers and sellers are competitive market forces (e.g., the spot market), contractual agreements (e.g., core carrier programs or dedicated fleets), and administrative controls (e.g., private fleets).

Interestingly, all three types of governance structures are present in the TL carrier segment. The TL spot market is used by virtually all shippers and carriers to some extent - either as a primary method of procurement or to fill in needed capacity. At the other end of the spectrum, vertical integration is still the dominant governance structure with more than 48,000 private fleets generating over two thirds of all motor carrier revenues. Contractual relationships between shippers and carriers, covering the vast territory between spot markets and vertical

integration, are becoming more popular. La Londe and Cooper (1989, 19) estimate that use of contract carriage will increase by as much as 44% in TL over the next several years.

Transaction cost analysis (TCA) examines “the *comparative costs of planning, adapting, and monitoring task completion under alternative governance structures*” (Williamson 1985, 2). This section reviews TCA by addressing the total costs involved in a transaction, the behavioral assumptions being made, the specific characteristics of the service being exchanged, and the governance structures that arise out of these characteristics. These tools will be used in the later sections to explain why shippers and carriers use different forms of governance structures.

3.1.1 Total Transaction Costs

TCA theory is based on the assumption that companies select the governance structure that minimizes the total cost of a transaction or exchange between a buyer and a seller. For the TL procurement problem, the basic transaction is the shipment. The main point of TCA is that the cost of arranging this movement (whether by a for-hire carrier or by private fleet) is more than the flat or per-mile freight rate charged. The total cost of the movement consists of the sum of three components: *ex ante* costs, transportation costs, and *ex post* costs.

Transportation costs are the most widely recognized costs of the transaction and are oftentimes considered the sole cost of an exchange. They include the variable or direct costs accrued by the physical movement of the products from origin to destination in accordance with the required service standards. For a TL shipment, this is typically a flat charge based on the distance, backhaul opportunity at the destination, and any accessorials required.¹⁵

Ex ante costs include the planning and negotiating conducted before the transaction for the purpose of “safeguarding an agreement” (Williamson 1985, 20). At one extreme, all possible outcomes can be addressed and all contingencies planned for. This is an extensive task and would be quite costly. At the other extreme, the planning could consist of a general agreement between parties to address deviations from the plan as they arise during the transaction.

Ex post transaction costs occur when remedying deviations from the agreed upon *ex ante* plans. These are the costs associated with carrying out any after-the-fact “haggling” or renegotiating of the agreement. Anytime there is a discrepancy between the actual and the originally planned outcome, there is a need for *ex post* activities to correct and assign responsibility.¹⁶ If the discrepancy is specified in a legal contract, then the *ex post* costs would be the legal costs involved in settling the dispute.

¹⁵ Recall that an accessorial is any service performed by the carrier above and beyond the standard load, unload, and transport activities. These include inside deliveries, specialized packaging, and intermediate stops.

¹⁶ The cost of the non-performance is not included in the definition of the *ex post* costs. For example, suppose that a carrier overturns a trailer and ruins the entire shipment. The *ex post* costs are those costs required to settle the contract, determine who was at fault, and collect payment for the damaged goods. The cost of the damaged freight is not an *ex post* cost.

The *ex ante* planning and the *ex post* correction costs are somewhat inversely related. The former screens for problems before the actual transaction occurs while the latter serves as a mechanism or safeguard to correct actions when things do not go quite as planned.¹⁷ While there is a degree of trade-off between the *ex ante* and *ex post* costs, we cannot entirely eliminate the *ex post* costs by increasing the *ex ante* planning. This form of comprehensive contracting requires omniscience. Conversely, relying entirely on the *ex post* mechanisms to correct for any and all deviations would lead to increased *ex post* costs since the exchange would be fraught with ambiguity and vagueness without any pre-shipment agreements.

The overall objective of selecting a governance structure, then, is to minimize the total (*ex ante*, transportation, and *ex post*) transaction costs involved with a specific exchange. The reason why we need to consider both the pre- and the post- transaction costs are due to the behavior of the players and the characteristics of the transaction itself.

3.1 2 Behavioral Assumptions

Transaction cost analysis assumes that people are both boundedly rational and opportunistic. Bounded rationality assumes people to be '*intendedly* rational, but only *limitedly* so,'" (Simon 1961, xxiv). That is, while people are assumed to *want* to know all future contingencies, it is *impossible* for them to foresee exactly. This assumption appears valid for transportation procurement since neither party can foresee how the changes in the market will develop.

Opportunism is defined by Williamson (1985, 47) as "self-interest seeking with guile." That is, TCA assumes that people take actions to improve their own condition (self-interest), but they may break (or simply not honor) commitments if it is in their best interest.¹⁸ Opportunism also includes the suppression or omission of information when requested (Kreps 1990, 745). The opportunism assumption appears to be valid for transportation procurement since it can be assumed, for example, that a carrier will not (nor be expected to) reveal the absolute lowest cost they are currently charging in a lane to prospective customers. By the same token, shippers will choose to keep private information (such as what previous carriers charged per lane) to themselves. This is not to say that carriers and shippers are necessarily dishonest and unethical. They are simply using their private information in a strategically wise and clever fashion during the negotiation process in order to best achieve their own self interested ends. To do otherwise

¹⁷ These two costs are similar to the concepts of adverse selection and moral hazard in the insurance and economics literature.

¹⁸ For example, suppose that after accepting a tender from shipper A, a carrier receives a call from shipper B with a much more profitable load to be picked up at the same time. If B is a good customer and A is an unknown, then the carrier might break the commitment to A in order to haul B's load. If both A and B are regular customers, then it might be in the carrier's long-run best interest to honor the commitment to shipper A - thus trading a short-term decrease in profit for a potential higher long-run profit.

would actually be bad business practice. Additionally, though, this assumption implies that people may 'bend the rules' at times if the potential benefits justify the risk of being caught.

With these two assumptions, planning is now incomplete (bounded rationality), information is not always shared, and promises are not always kept (opportunism).¹⁹ Both assumptions are required for transaction costs to become relevant. Without bounded rationality, comprehensive contracting - planning for every possible outcome - would be feasible. Without opportunism, we could assume that everyone "plays by the rules" and thus use rule-governed contracts which provide general rules to guide behavior to lead to a maximum joint profit rather than covering every contingency.

3.1.3 Characteristics of the Exchange

Even with the behavioral assumptions, if the service or product being exchanged is a commodity, then TCA theory predicts that a competitive market will minimize the *ex ante* and *ex post* costs²⁰. A commodity is defined as any product or service that could be sold to one of many buyers or procured from one of several suppliers without any significant differences in quality or pricing. The need for specialized, complex contracts arise only when the exchange of a product or service becomes specific to either the supplier or the buyer. When this occurs, the competitive market is no longer an efficient mechanism for safeguarding the exchange.

We next discuss the four characteristics of an exchange which make the traditional competitive market unacceptable as a governance structure: asset specificity, temporal specificity, transaction uncertainty, and transaction frequency.

Asset Specificity

Asset specificity describes how easily a resource can be converted to another use. If the difference in the value of a good or service for its intended use is much higher than the value in its second best use, then there is high asset specificity.²¹ Kreps (1990, 747) notes that asset specificity is present if one side of an agreement "becomes more tied in to and in the 'power' of the other side." These "lock-in effects" occur when one of the parties has monopoly power over the other due to some specific characteristic of the exchange (Williamson 1985, 53). As such,

¹⁹ Kreps (1990, 747) provides a clear overview of how TCA and the assumptions of bounded rationality and opportunism fit in with other economic theories.

²⁰ Consider, for example, the case of mailing a document overnight using an express parcel service such as Federal Express, the United Parcel Service, or even the United States Postal Service. The *ex ante* and *ex post* costs have been minimized by market pressures. The standards of and adherence to the agreement are enforced by the fact that if you are not satisfied, you will call another carrier next time.

²¹ For example, a refrigerated van is designed for hauling temperature controlled freight but can easily be used for hauling general dry van freight. The difference in the rates are not dramatic, so there is not a large degree of asset specificity.

there is the potential for abuse of this power (opportunism). Thus, without the provision of opportunism, asset specificity would not affect the selection of governance structure.

Williamson (1985, 95-6) identified four major types of asset specificity: site, physical, human, and dedicated. Each of these describe a situation where the good or service being exchanged is no longer a commodity and requires some sort of special governance structure.

Site Specificity occurs when investment is made in a facility (such as a warehouse) that either cannot be moved or has a high relocation cost. The large transfer cost decreases the value of the facility for any other use, thus 'locking-in' the investors. For example, Schneider National built a terminal at Kimberly-Clark's plant specifically to provide better service. Without Kimberly-Clark's business, the value of the terminal to Schneider (or other potential renters) is low, thus, the terminal has high site specificity (Schulz 1995c). Another example is the construction of a rail loading terminal at a customer's location. The rail head has no value except to that specific customer.

Physical Asset Specificity occurs when the physical features or characteristics of some of the equipment is specialized for one customer to the point where the value of its second best use is very low. Autoracks and specialty tank cars for railroads are classic examples of physical asset specificity. Often times the autoracks are designed for only one manufacturer's vehicles which makes the asset extremely specific to that shipper.

Human Asset Specificity occurs when there is substantial 'learning by doing' or problems involved in moving or retraining teams. A company can come to rely on a supplier to perform certain operations to the point where removing them would require extensive re-training of their own personnel. The JIT II concept pioneered at Bose, where suppliers have representatives working within the Bose factories and becoming de facto co-workers, operates on this principle (Dixon and Miller 1994). The suppliers' representatives become intimately involved with Bose's operations; creating human asset specificities due to the large learning curve that a new supplier would require. Also, the personal connections (or face time) made between workers form a very strong bond that further increases the specificities.

Dedicated Asset Specificity occurs when investment in additional equipment or facilities is based on behalf of a particular buyer. Without that buyer's business, the seller would have excess capacity and the equipment would be idle. For example, suppose a carrier purchased a large number of tractors in order to handle expected traffic from a new customer. Once these trucks are purchased, the carrier is relying on the shipper to actually come through with the agreed upon volume. At this point, the shipper might be able to leverage the carrier's excess capacity by demanding lower rates, even though the equipment is not specific to that customer.

Each of these forms of asset specificity exhibit the potential for one of the parties to exploit the other in order to extract monopoly rents. When assets become specific to either the

buyer or the seller they become 'locked in' so that the bargaining power is no longer equal. In these cases, the ability of the market to enforce the contract is lost since one of the parties faces a hazard in finding an alternate partner in the exchange.

Temporal Specificity

Temporal specificities occur when the timing of the exchange affects the bargaining power between the buyer and the seller. First recognized by Masten, Meehan, and Snyder (1991, 9), temporal specificity differs from asset specificity in that the product or service being exchanged need not be specific to either the buyers or the sellers. Instead, *temporal specificity* occurs when timeliness of performance is important and any delaying of an activity may become a potentially effective strategy for exacting price concessions. The original example concerned large construction projects where sub-contractors could gain price concessions knowing that the general contractor had severe penalties for time delays and that their sub task was a prerequisite for other tasks. Thus, they had monopoly power over the completion of the entire project due to their placement on the critical path rather than the specific nature of their task.

Transaction Uncertainty

Transaction uncertainty limits the use of the competitive market as a governance structure. Because we are assuming bounded rationality, uncertainty in the transaction will always be present and the ability to adapt an existing agreement to changing and unforeseen situations is required. The selected governance structure must be able to cope with *ex post* changes in order to preserve the exchange.

Uncertainty has many origins. It can be attributed to unpredictable events such as changes in the markets, economies, or customer demand. It can also be due to those things that can be anticipated, but are deemed too expensive to be entered into the formal contract or to outcomes about which only one party has information (Kreps 1990, 749). The idea of uncertainty is not simply randomness; it can also appear in a purely deterministic setting. Simon's classic example is a chess game which, while wholly deterministic, has uncertainty due to the computational inability of the players to determine the environment with all contingencies exactly, that is, approximation is required in decision making.

While both shippers and carriers are uncertain as to actual future demand for services, which rely on the greater market, the sources of uncertainty differ between the two. For shippers, the uncertainty lies in the availability of trucks and the service level of the delivery. For carriers, the uncertainty is in not only the availability of future loads, but also the temporal and spatial characteristics of the loads. As discussed in Chapter 1, TL carriers need to balance their equipment between net sink and net source locations which can cause significant costs.

Carriers are more sensitive to connection uncertainty than shippers in this respect. The distinction is important in that carriers have a greater incentive to control the amount of uncertainty in the system.

Transaction Frequency

Transaction frequency refers to both the magnitude and distribution of the individual transactions. The larger the total number of transactions occurring between buyer and seller, the easier it is to recoup the costs of any transaction or asset specific investments. In this respect, it is similar to a fixed set up cost that is allocated across all units produced; the burden per unit is reduced as the utilization increases. If a special form of contract is needed, it is easier to justify the additional costs of a more complex governance structure if there are a sufficient number of transactions to absorb the costs. High frequency by itself is not enough to warrant a complex contractual agreement, but, when asset specificity is present, higher transaction frequency will make the use of complex contractual agreements more palatable to shippers and carriers.

The distribution of these transactions (over both time and space) also affects the choice of governance structure. A steady, reliable stream of balanced shipments, for example, is more attractive to a TL carrier than the same total number of shipments spread out across the entire network.

Summary

Of the four characteristics of the exchange, the asset and temporal specificities are considered to be the most important. Without the opportunity for one of the parties to extract monopoly rents, TCA theory predicts that the competitive market will be sufficient to safeguard the transaction. If asset or temporal specificity exists, then these other factors exacerbate the need for more complex governance structures. Higher uncertainty increases the need for handling unforeseen events while increased transaction frequency (magnitude) justifies the cost and makes that stream of transactions (shipments) more attractive to the carrier.

3.1.4 Governance Structures

The degree to which asset and temporal specificity, uncertainty, and frequency occur within a transaction determines the most suitable form of governance. This section shows how they interact through a simplified example adapted from Williamson (1985, 32-35) and an example of a complex contractual relationship found in practice.

Simple Example

Suppose that a shipper is interested in obtaining standard TL services from a carrier which requires no special investment at all. That is, the equipment and the entire arrangement has no asset specificity to it. The carrier and the shipper could, conceptually, enter into this one-time agreement for a single shipment using the spot market with an agreed upon sale price of p . By theory, the rate p is the clearing price for the market and is, thus, the break-even cost for the carrier. At the completion of the shipment, if either party is dissatisfied they can walk away from any future transactions and find new buyers or sellers with which to exchange. Thus, the spot market minimizes the *ex ante* and *ex post* costs.

Now suppose that the carrier could modify its service by making an investment, of k dollars in, say, a special feature customized to that shipper's needs. The increased value of this service would be totally lost if not used in this transaction for this specific shipper. This could be a specially designed trailer or agreeing to wait beyond the scheduled pick up time if the load is late. The first case is an example of physical asset specificity while the second is a temporal specificity. In each case, the shipper could conceivably leverage the carrier, since both the specialty trailer and the extra wait time are, by assumption, of no value to other shippers. The carrier, assuming all of the risk of potential lost investment (in dollars and time), would require a higher price, $p' > p$, if it would offer it at all. Again, p' is the break-even price for the carrier when offering this new service. This price would need to cover the loss of investment if the shipper uses its advantage of being the only useful buyer of this service. Due to the opportunistic behavior assumption, the carrier has no guarantee that the shipper will actually pay the higher amount, p' . This is an unstable arrangement if enforced using only market forces since the carrier stands to lose the invested time or money. Being able to walk away after the transaction will not minimize the *ex post* costs for the carrier in this situation since the investment cost, k , cannot be recouped.

Now, suppose that the shipper, instead of using the spot market, is willing to offer some sort of *ex ante* safeguards to the carrier that lessen the carrier's risk. These safeguards can be in the form of financial penalties (hostages) if the shipper does not fulfill the agreement,²² some agreed upon arbitration scheme where a third party might be brought into to settle claims, or a guarantee of continuity of business (i.e., increased frequency and decreased uncertainty). This could also include the use of administrative controls, i.e., vertical integration. If the value of the safeguard, s , is sufficient, then the carrier would be willing to enter into a contract and would offer a price of, say, p'' .

²² The safeguard for waiting is essentially an accessorial clause in a contract.

Figure 3.1 illustrates the concept behind how asset specificity determines the form of the governance structure. The break-even price for the Case II ($k>0$ and $s=0$) will always be higher than that of Case III ($k>0$ and $s>0$), that is, $p' > p''$, since the carrier is absorbing all of the investment risk in the second case. However, the spot market price of Case I is not necessarily the lowest cost alternative. If $p < p''$ then the spot market is better suited for the transaction since the price of providing the specialized equipment is not justified even with the safeguards provided by the shipper. However, if $p > p''$ then a contract with specified safeguards ($s>0$) are preferred.

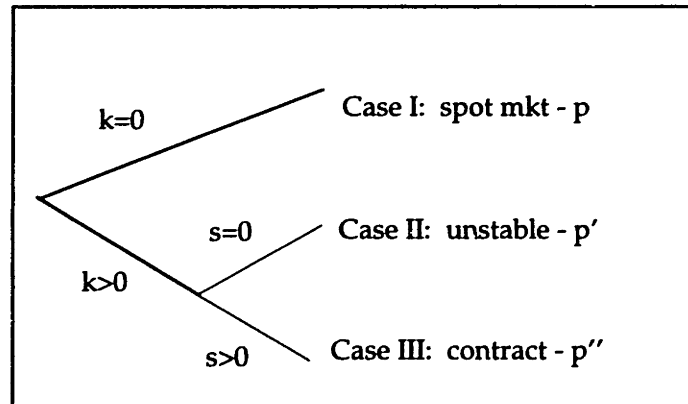


Figure 3.1. Simple contracting scheme (Williamson 1985, 33).

The specificity (both temporal and asset) is reflected in the value of k . As the specificity increases, the need to contract (or use administrative controls) also increases in order to spread out the investment risk. Increased transaction frequency, if guaranteed to the carrier, is manifested in a higher safeguard value, s , which can simply be a guarantee of future business. This lowers the threshold for deciding whether to contract. Finally, greater uncertainty increases the investment of k and raises the value of the safeguard s needed to maintain a stable relationship.

Extended Example

An actual example of how these characteristics interact in determining details of governance structures, is the agreement set up between Libby-Owens-Ford, LOF, (an industrial glass manufacturer), Schneider National (the largest TL carrier in the U.S.), and Wabash National (a trailer manufacturer) (Distribution 1993).

LOF, like most plate glass manufacturers, moves product via "A-frame" style trailers which reduces their backhaul opportunities, increases the number of empty miles, and causes poor vehicle utilization.²³ While they had approached numerous carriers about developing a specialized trailer which could easily be converted to a conventional flatbed for hauling general

²³ The percentage of empty miles for specialty trailers is typically 3 to 5 times that of dry vans and can run as high as 40%.

freight on backhauls, carriers had always responded that the investment was too high. Ultimately, Schneider was approached concerning the specialized trailer, along with the possibility of becoming LOF's single transportation source. This served as a safeguard (continuity of business) for the carrier to enter into the agreement.

Schneider brought in Wabash (which is their single source trailer provider) and discussed the production possibilities. LOF wanted the specialty trailer to be a competitive advantage and requested that it be developed exclusively for them. In exchange for Schneider making 120 trailers available and building a new multi-million dollar terminal at their plant, LOF guaranteed a continuation of business for two years beyond the normal depreciation period of the equipment. Also, Schneider requested, and received, freight rates slightly higher than LOF's original rates.

The higher transportation rates were more than offset by lower logistics costs due to increased allowable shipment size, better communications, and added scheduling flexibility. Wabash received nine patents for the trailer design and agreed to not build them for any other customers for the life of the contract between LOF and Schneider. While LOF could have demanded exclusive use for the life of the patents (17 years), this was not proposed since Schneider needed to be compensated for limiting their opportunities by granting exclusivity to LOF.

There are several interesting points here. First, the trailers are specialized but would still be usable by other glass manufacturers. LOF artificially increased the physical asset specificity of the investment by restricting Schneider and Wabash from using them for any other customers. Additionally, LOF required a highly site specific terminal to be built near their plant. In return, Schneider received safeguards in the form of extended contracts (depreciation life of trailer plus 2 years) and increased frequency (sole transportation provider). Wabash is restricted from using this trailer elsewhere - but retains the patents. Also, as single source to Schneider, Wabash receives safeguards in the form of continuity of business (higher frequency and lower uncertainty).

Because each party had significant incentives to exploit at least one of the other's weaknesses, these specialty trailers would not have been developed in the spot market. It would be even more expensive in a vertically integrated setting since the network is not balanced and LOF would need to absorb excessive empty miles or locate its own backhauls; a task that Schneider with its extensive service network is far better suited. Thus, the complex contracts between the three parties with specific agreements minimized the total transaction costs.

3.2 Governance Structures in TL Trucking

The three governance structures discussed (spot markets, contractual agreements, and vertical integration) have thus far been treated as separate and discrete. In actuality, though, the distinctions between these governance structures are not sharp and, instead, the different forms can be thought of as points along a continuum. As shown in Figure 3.2, the range of possible structures are anchored by pure spot markets on one end and pure vertical integration (private fleets) at the other. Spanning all points in between are contractual arrangements between shippers, carriers, and, sometimes, third parties, such as brokers or non-asset based logistics companies.

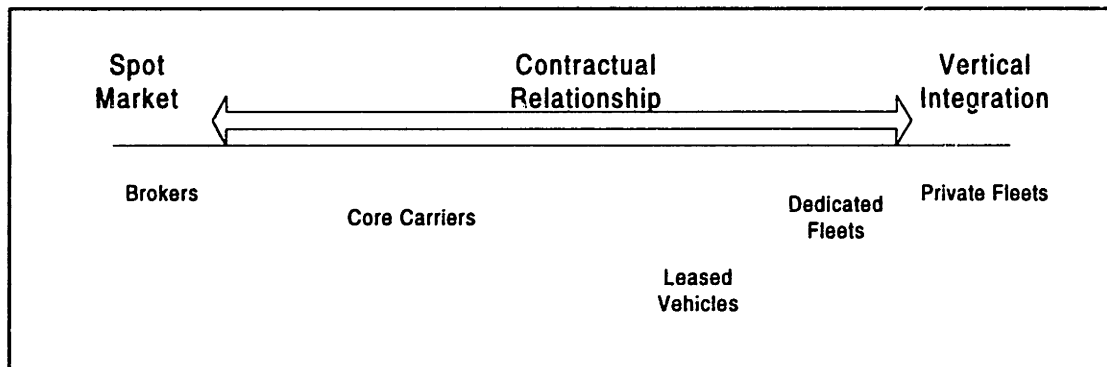


Figure 3.2. Continuum of governance structures.

The distinctions between the governance structures are becoming increasingly blurred. For example, many shippers with private fleets lease their equipment. Does this make it a contractual agreement? What if they also outsource their maintenance operations, drivers, or routing and scheduling guidance? At what point does the private fleet become a completely contractual arrangement? At the other end of the spectrum, shippers often form “contracts” with a single broker rather than maintain its own extensive list of spot carriers. Is this a contractual agreement or is it still a spot market? Again, there are no hard and fast lines demarcating where each structure begins and ends.

The remainder of this section summarizes how and to what extent each type of governance structure is being used by shippers and TL carriers.

3.2.1 Vertical Integration

The MIT Dictionary of Modern Economics defines vertical integration as occurring when “the activities of a firm extend over more than one successive stage in the production process of transforming raw materials into final goods.” For transportation, vertical integration implies the shipper uses a private fleet. The traditional incentives, performance measures, and administrative controls used within a firm are applied here to secure the required service and

efficiency levels. The use of private fleets was discussed in depth in Section 2.3.1. This discussion is primarily just a summary.

In its purest form, a private fleet consists of a shipper owning all of its equipment, using only company drivers, and performing all of the ancillary services required with running a fleet of vehicles. This includes: maintaining and procuring the trailers, tractors, and other equipment; hiring, training, and supervising all drivers and administrative workers; scheduling, routing, and tendering of all truck movements; and billing and auditing for all freight payments. Thus, for total vertical integration, the shipper both owns the equipment and controls all the operations.

In practice, though, most shippers outsource at least a portion of their private fleet operations. While the majority of private fleets still use their own drivers, AT Kearney (1993) reports that close to two thirds outsource their maintenance and equipment. Other shippers are outsourcing (1) specific activities associated with private fleet operations (such as having a third party firm identify and schedule 'milk runs' for them) or (2) geographic portions of their network (such as the line-haul portion between plants). The primary benefit of outsourcing personnel, equipment, or certain tasks, is that it decouples the ability to control the operations from the need to own all aspects of it. The important characteristic of private fleets is that the shipper maintains control of operations, regardless of the ownership of the equipment or drivers.

There are several reasons for using private fleets. Private carriage is perceived by many shippers to improve goodwill through better service, increase name recognition with customers, and to guarantee equipment availability. Additionally, a private fleet can provide bargaining leverage for shippers when negotiating with for-hire motor carriers and can allow the shipper to maintain total control over the transportation process.

On the negative side, private fleets are, in general, less efficient than for-hire carriers. Investment in a private fleet is also a diversion of capital and managerial attention. Another deficiency of vertical integration in general is the lack of innovation that competition in an open market normally brings. Williamson (1985, 161) notes "the transfer of a transaction out of the market into the firm is regularly attended by an impairment of incentives. It is especially severe in circumstances where innovation (and rewards for innovation) are important." This bears directly on TL trucking where innovations over the last decade in communications equipment, computer technology, and operational software have had a huge impact on the service and efficiency levels that for-hire carriers can provide.

3.2.2 Spot Market

A pure competitive or spot market enforces the terms of an exchange through the threat of alternative sellers or buyers. Simply put; there is always someone else with whom to do business. The exchanges in a spot market need to be well defined and uniform and imply no additional relationship beyond the transfer of the goods or services. The traditional characteristics of a perfectly competitive market should apply, to wit:

1. large number of firms each producing the same homogenous product,
2. each firm attempts to maximize profits,
3. each firm is a price taker: its actions have no effect on market price,
4. no substantial barriers to entry or exit,
5. prices are assumed to be known by all market participants: information is perfect, and
6. transactions are costless: buyers and sellers incur no costs in making exchanges.

The dry-van TL market satisfies most of these criteria. With more than 50,000 general freight TL firms, each individual carrier's impact on influencing price is minimal. The barriers to entry and exit are almost nonexistent, consisting primarily of minimal safety and financial standards. The fifth and sixth criteria are not as easily satisfied, however. Rates are not always known by all parties and the exchange is not quite 'frictionless.' The Trucking Industry Regulatory Reform Act (TIRRA) of 1994 removed the requirement for carriers to publicly file their rates. Additionally, in a transportation spot market, there is a great deal of information uncertainty concerning the availability and price of transportation to a shipper. Likewise, the carrier is equally uninformed about availability, characteristics, and rates for potential follow-on loads. This lack of readily available information leads us to expect the use of brokers to facilitate the information exchange.

The use of a broker in a competitive market does not necessarily change the dynamics and, in fact, can lower the transaction costs by reducing the search costs. If the broker charges a per-unit fee (per-shipment in our case), then there may be a slight dampening of trading volume with the broker's fee being divided between the shippers and carriers in relation to their price elasticities.²⁴ Of course, this ignores the increase in volume that the broker might bring to the market by facilitating the exchange of information. While not totally costless, using brokers does not destroy dynamics of the competitive market.

Because the use of brokers is so widespread, the freight broker industry is discussed in greater depth in the next section. Additionally, the growing use of alternative electronic spot markets is addressed.

²⁴ If the shipper's demand is perfectly elastic, then the carriers would absorb the brokerage fee while, conversely, if the demand is perfectly inelastic then the shippers would absorb the fee. Essentially, the party which is less sensitive to changes in price will end up paying a larger portion, if not all, of the fee (Nicholson, 1992, 440).

Freight Brokers

A freight broker's primary role is to match shippers' loads to carriers' capacities. Additionally, brokers serve as a quality screening service for both shippers and carriers. Earning between 5-10% of the freight bill, brokers usually pay the carrier, collect from the shipper, and serve as creditors by offering cash advances to carriers for driver expenses.²⁵ Brokers traditionally assume no liability for the loads brokered and have never been required to publish tariffs. This has provided them a great deal of flexibility in pricing and has allowed them to avoid, at least the immediate, pecuniary risk of a lost or damaged shipment. Of course, a broker's long-term reputation is hurt if it does not maintain a low loss and damage rate.

Along with the increase in the number of TL carriers in the early 1980's, there was a parallel increase in the number of ICC registered freight brokers; from 70 in 1975 to more than 6,000 in 1988 (Brown 1990b). Standard & Poor's (1994, R35) estimated that approximately \$5 billion in freight revenue was handled by brokers in 1993, representing over 6% of the total for-hire market. Like the TL segment, though, freight brokers are a "barbell" industry consisting primarily of very small regional brokers and very large national brokers.

Information on regional brokers is sparse, incomplete, and mostly anecdotal. Brown (1984, 1990a, 1990b, and 1991) examines regional brokers using interviews, surveys, and some statistical analysis. In 1987, the typical regional freight broker operated out of one office, employed less than 6 workers, tendered about 2,500 TL and 1,000 LTL loads annually (with length of hauls ranging from 500 to 1,500 miles), generated annual revenues of \$1.8 million, used close to 200 carriers, and served slightly under 100 shippers.²⁶

National brokers, while operating on much larger scales, perform the same activities as the regional brokers. C.H. Robinson (CHR), for example, processed over 700,000 shipments generating \$1.3 billion in sales in 1994. CHR relies on having a large stable of small carriers. In 1994, over 15,000 carriers were used, of which only 1,000 hauled more than 100 loads (less than two a week).²⁷ Just under two thirds of these carriers are under contract with CHR. The contracts, however, are designed almost exclusively to reduce *ex post* costs; they delineate liability for numerous contingencies but only guarantee the carrier 3 loads per year.

The main users of both regional and national brokers are the small shippers. These are typically manufacturing firms with fewer than 100 employees that ship primarily TL quantities

²⁵ Beilock and Shell (1992, 64) note that while many small carriers rank this service as being very important, these cash advances are "in effect, unsecured loans at usurious interest rates." Often times the rates exceed 100% for payment within 21 days. These are mostly owner-operators rather than fleet carriers.

²⁶ These are averages with some of the ranges being quite large. For example, revenue ranged from under \$1,000 to over \$92 million with a median of \$700,000. These figures from Brown (1990b, 53) are only meant to provide background.

²⁷ Conversation with Paul Wegner, Manager Logistics Services, C.H. Robinson Company, 22 February, 1996.

with a frequency of less than 20 shipments per week across their entire network (Brown 1991, 46). The products shipped are generally lower valued and time-insensitive in nature.

On the supply side, virtually all carriers (both for-hire and private) use brokers to some degree. Even the Advanced Truckload Firms (ATLFs) described in Chapter 2 use brokers, or act as brokers themselves, for portions of their service network. Different types of carriers, however, use brokers for different purposes. Private fleets and larger for-hire carriers generally use brokers on an infrequent basis in order to fill an odd backhaul. In the mid-1980's many large shippers with private fleets acquired broker licenses and tried to fill their own backhauls.²⁸ A large carrier can actually be both a buyer and a supplier for a broker. Upper tier carriers often times broker out certain lanes that they win from a shipper because the lanes do not fit into their service network. The heaviest users of freight brokers, though, are the lower and middle tier carriers. These carriers tend to use brokers as their marketing departments allowing them to operate over geographically dispersed areas or to just "get them back home" (Brown 1990a, 35).

By using the spot market, shippers are purchasing a commodity. They benefit from using brokers in that a single call can provide access to multiple carriers and thus can most likely find a carrier to haul their load. It provides the shipper the perception of having a carrier with a large vehicle fleet without actually needing to contract with an upper tier carrier - which might not consider that shipper a very attractive customer, anyway.

Having a large carrier base is both the broker's strength and weakness. The large base allows a broker to provide greater coverage and capacity to its shipper customers. However, since this relies on the presence of many individual carriers, the broker achieves wider coverage at the cost of lower service consistency. Thus, shippers that value service over cost, will probably not use the spot market for time-sensitive freight since often the penalty for poor service can outweigh the actual cost of transportation.

Electronic Markets

Two forms of electronic markets can be envisioned. In the first system, opportunities for matching individual loads from shippers to available carriers are identified and rates are negotiated directly between the two parties, thus replacing intermediaries. In the second system, the same information is made available, but intermediaries (brokers) are needed to formalize the matchings, similar to the system used by travel agents to arrange flights for airlines.

Beilock and Shell (1990) examine the potential for using both types of electronic marketing systems (EMS) in the produce trucking segment. They estimate that the benefits due

²⁸ Xerox, for example, used its broker status to fill-empty backhauls for its for-hire carriers in the hopes of reducing the rates (Distribution 1983).

to electronic tendering and marketing could reach ~\$100 million for the produce segment where the transaction costs for electronic marketing would be about 1-2% of freight rate as compared to 10% for broker fees. They suggest that the latter form of electronic market be used because the former suffers from asymmetric information and an inability to assess the other player's reliability or qualifications.

In late 1995, however, a handful of different load matching services (generally of the first type of system) were launched using the World Wide Web . Three systems that seem to be the most heavily referenced are Internet Trucking Information Exchange (ITIE), MassMotion, and the Internet Truckstop.²⁹ While all three of the services are designed to provide easier information exchange between shippers and carriers, they differ dramatically in their details.

Internet Trucking Information Exchange (ITIE) is a free system with totally unrestricted access to all postings. Shippers post information on loads available for pickup while carriers post a vehicle's current location and desired destination. If the price is not listed on the shipper's posting of a load, then it is negotiated off-line. ITIE is essentially an electronic bulletin board replacing the physical load board found in trucking stops. The second system, MassMotion is free to shippers, but carriers pay a monthly fee and a service charge for each truck posted. While carriers post when and where their trucks will be available, shippers cannot post their available loads. Instead, shippers can only perform searches across the carriers' postings for a suitable truck. The searching screen requires exact information (complete city and state) and is therefore unable to determine that a load in Boston can be served by a carrier with a truck in Cambridge! Shippers cannot search on any larger geographic level. The price for the load, if matched, is determined off-line. The third system, the Internet Truckstop, is unlike the other two systems in that it is primarily aimed at serving brokers and carriers, rather than shippers directly. Carriers pay a monthly fee to be allowed to post available trucks as well as search for loads posted by brokers. While brokers can post loads for free, they must pay a monthly rate to be able to search for available trucks. The searching mechanism recognizes proximity and uses a map interface.

Of the three services, the Internet Truckstop is the most active, receiving an average of about 1000 loads and 200 trucks posted each day.³⁰ The majority of the carriers signed up are owner/operators and lower tier (less than 50 vehicle fleet) carriers. Regional brokers have posted most of the loads with very little direct shipper or national broker interest.

²⁹ The universal resource locators (URL) are: <http://www.internet-trucking.com/>, <http://www.massmotion.com/>, and <http://www.truckstop.com/>, for Internet Trucking Information Exchange (ITIE), MassMotion, and the Internet Truckstop, respectively. This is not a comprehensive list. These three web-sites were the most heavily cross-referenced on trucking related web-sites.

³⁰ Conversation with co-owner Internet Truckstop, 11 March 1996.

The success of the different systems is related to their underlying structure. ITIE provides full information but, because it is free and accessible, it "commoditizes" both shipments and carriers. MassMotion, on the other hand, ignores the fact that transportation is a derived demand by not allowing carriers to search for loads. The Internet Truckstop, on the other hand, allows for searching by all parties and recognizes proximity of loads. All three of these services are aimed at the commodity end of the TL segment and are essentially replacing the use of faxes or phone calls to truckstops.

All electronic markets used for load matching have potential temporal specificity problems which open up opportunities for haggling in the price negotiation. If the posting date is old, but the ship date is fast approaching, then a carrier might be able to leverage this information to negotiate a higher rate. From the shipper's perspective, if there are a large number of carrier postings for an area, then they can use this information to bargain for lower rates. In other words, the information passed in the system can provide indicators as to the bargaining power of the other player and the strength or weakness of the market.

The long term use of electronic matching systems seems to be for brokers rather than directly for shippers. Because the service only matches loads to carriers, it performs just a portion of the broker's job. Shippers using electronic matching systems have to deal with a different carrier for each load. With a broker, however, they can coordinate the majority of the activities through a single point of contact. Thus, electronic matching systems do not seem to be a threat to brokers, but rather a better tool for them to use.

In summary, the spot market, is used by virtually all carriers and shippers but is relied upon by small shippers and the lower tier carriers, typically using freight brokers. In all cases, though, spot markets treat each load as a separate and independent transaction with no *ex post* guarantees beyond the pickup, line-haul, delivery, and payment. Electronic markets do not change the load-by-load characteristic of the spot market - they only improve the information flow. Interestingly, this improved information flow can, in fact, lead to a further "commoditization" of the TL services themselves as the shippers can shop between carriers more easily.

3.2.3 Contractual Agreements

The third general form of governance structure consists of all agreements lying between the extremes of vertical integration and pure competitive markets. Essentially, a contract is involved any time a shipper procures transportation services outside its own organization and the exchange is tempered or modified to better suit one (or both) of the two parties. A contractual agreement between a shipper and a carrier can take many different forms and can include a wide variety of stipulations. While contracts are legal documents and are thus entitled

to settling *ex post* disputes in court, most try to avoid such recourse through other enforcement mechanisms. For example, contracts can specify price penalties, contract lengths, arbitration procedures, volume adjustments, cost sharing, and many other methods with which to settle differences.

There has been a substantial increase in both the number and the diversity of contractual agreements entered into by shippers and carriers. This growth is being attributed to increased interest in forming 'value added partnerships' between shippers and carriers. Partnerships or alliances of this sort have varying degrees of formality depending on the tastes and trusts of the parties involved. Examples of new types of relationships include evergreen contracts (where the length of the contract is unspecified and subject to automatic renewal), dedicated fleets (where a carrier leases all aspects of the trucking operation to the shipper who maintains operational control), JIT II relationships (where the carrier has the responsibility for its own tendering), and single source providers (where a carrier takes over all aspects of the transportation process).

Like most contractual relationships, however, there are very few public data on contract particulars. Shippers and carriers have a vested interest in keeping contractual information confidential, especially concerning rates, service levels, and capacity commitments. The majority of the information on contracts reported in the academic and trade journals is either anecdotal or based on relatively small surveys. These are briefly reviewed.

In a review of close to 100 contracts between shippers and carriers negotiated during the period 1978 to 1984, Cavinato (1984) listed and described 21 key components of transportation contracts.³¹ Meant to be more of a primer for transportation managers than a review of practices, Cavinato provides the structure that contracts usually take. The components fall into three general categories: guarantees to the carrier (volume requirements, scope of transportation services to be provided, disclosure of goods to be hauled, and method of transport, etc.), guarantees to the shipper (performance standards, operational standards, etc.), and *ex post* mechanisms to correct discrepancies (loss and damage, escalation procedures, force majeure, failure to comply clause, dispute resolution methods, etc.).

La Londe and Cooper (1989) surveyed slightly over 300 shippers and carriers in 1987 and reported that the use of contract carriage was expected to increase by over 100% for LTL and 44% for TL. In a survey of 40 large manufacturers with JIT programs, Lieb and Miller (1988) reported that 78% of the firms had reduced their carrier base through the use of contractual agreements and 73% had negotiated specific contracts with carriers for their JIT

³¹ These are: Disclosure of Goods, Responsibility of Goods, Method of Transportation, Responsibilities of Instrumentalities of Transportation, Term and Termination, Escalations, Volume Requirements, Scope of Transportation, Performance Standards, Operational Standards, Force Majeure, Billing and Collecting, Government Regulation, Applicable Law, Assignability, Title to Goods, Method of Changing the Contract, Notice, Effect of Failure to Comply with Contract, Method of Resolving Disputes, and Confidentiality.

operations. Additionally, 51% and 27% of respondents reported increasing their use of contract carriage for outbound and inbound traffic, respectively.

Crum and Allen (1990 and 1991) report on surveys conducted of both shippers and carriers. They found that 82% of shippers surveyed had engaged in contracting over the last year with 40% reporting that more than half of their for-hire trucking expenditures were for contract carriage. Close to three quarters of these contracts were of 1-2 years in duration, 45% contained stipulations of some level of dedicated equipment, and over half of all shippers that used contracts did so with Class I carriers. Crum and Allen also found that over 80% of the carriers reported some level of contracting activity with the amount of contracting being related to the size of the carrier: 53% of the Class I TL carriers reported earning more than half of their revenues from contracts compared to just 38% of the Class II carriers.

A 1993 poll of shippers by *Transportation and Distribution*, a trucking trade journal, showed that 45% of the respondents used contracts to define their relationship with transportation companies (Richardson 1993). This is a lower percentage than reported in the academic journals and probably reflects a survey sample with smaller, less sophisticated shippers than those responding to multiple page surveys from academics. Perhaps, though, it is more representative of the actual conditions.

While these studies report on carriers partially involved in contracting, some carriers and third party providers are moving into contracts specifically to avoid the transaction based business altogether. Ryder has set as an overall strategic goal to increase contractual relations and totally leave the transaction-oriented end of the business in order to increase stability in both traffic flows and in pricing. As of 1995, 82% (\$3.7 billion) of Ryder's annual revenues was generated under contract (Schulz 1995d).

Finally, Rinehart (1989) surveyed 104 carriers and 260 shippers that were actively engaged in contracts. He found that over 80% of all contracts concerned two parties, one third had a duration of between 7 months to a year, just under half had a duration of 1 to 5 years, and a price based bidding process was used for two thirds of all contracts.³²

In summary, contractual agreements are fast becoming the dominant governance structure used between shippers and TL carriers. The literature suggests that contracting is primarily occurring between shippers and the larger, Class I TL carriers. The terms of the contracts seem to include longer periods of time but still concentrate on detailing *ex post* liabilities. Guarantees of capacity by the carriers seem to be more prevalent than volume guarantees made by the shipper.

³² Interestingly, rather than seeing bidding as the best method of allocating lanes to carriers, Rinehart opposed competitive bidding noting that "[f]ailure to circumvent the bidding system may lead to an agreement which reduces the potential benefit to the shipper and the carrier."

Summary

Relations between shippers and TL carriers cover the full range of governance structures. While still the predominant structure, private fleets are becoming more difficult to justify in light of the increased efficiency and effectiveness of for-hire carriers. The for-hire market is also better at identifying and implementing innovations, further widening the efficiency gap. The spot market is utilized by all players, but it is most heavily used via brokers by the smaller shippers and carriers. Finally, contractual agreements between shippers and carriers are becoming increasingly popular across all segments of the TL market.

Economic theory, however, does not provide any convincing arguments as to why contracting is being used more frequently for general freight carriers. The spot market still appears to be the most efficient mechanism for procuring TL services. There are a large number of both shippers and TL carriers, the service is fairly uniform, and brokers enhance the information exchange making it almost frictionless. So, the question remains, why are shippers and carriers forming contracts for dry-van freight? The next section addresses this question.

3.3 Selection of Governance Structure

This section discusses why shippers and carriers are using contracts when the conditions for a spot market are so pronounced. A review of previous research is presented first, followed by three proposed network based justifications.

3.3.1 Review of Previous Research

Palay (1984) explored how railroads and shippers select governance structures. The analysis took place in the fall of 1979, was subject to regulated rates, and consisted of field interviews with representatives from both shippers and railroads. Essentially, he found that the more specialized the investment, the lower its value was in its next best use or quasi-rent. He found that the quasi-rent for autoracks, for example, is essentially the value of the scrap metal!

Palay classified different governance structures based on five elements: method of contract enforcement, how adaptations to changed circumstances were carried out, types of adjustments that were made, whether information for long-term planning was exchanged, and whether structural planning was attempted. He found that the critical component was the method of enforcement. These methods included having the ability to exit from the exchange (as in a competitive market), maintaining an implicit threat to withdraw future service or volume from the agreement, having the desire to continue a reputation of fair play and honesty, being reluctant to exploit short-term advantages to the detriment of long-run relations, and, finally, the bureaucratic controls of vertical integration. The middle three enforcement

mechanisms are common to contractual agreements. Overall, Palay confirmed TCA theory and found that in all exchanges reviewed, the type of governance structure hinged on the degree of asset specificity. As assets become more specific, the contract becomes more tailored to the two parties involved.

Investigating the decision to use for-hire TL firms or private carriage, Bienstock (1994) examines how the cost of motor carrier transportation of finished goods is affected by both the governance structure (contractual and vertical integration) and the characteristics of the transportation activities. Simulating a distribution channel's transportation system, she tested for the effect that physical asset specificity, uncertainty, and frequency would have on transportation costs. The dependent variable was the mean total cost (transportation and transaction) for a full year (250 days) of shipments. Four independent variables were controlled for in the simulations: governance structure (private fleet versus contract), asset specificity (refrigerated trailers versus dry vans), uncertainty (variability of transit time using two levels [high/low] based on information from a practitioner), and frequency (number of shipments per year was varied between high, 1200, and low, 350, where the shipment size was kept constant so that higher frequency levels implied higher overall volume).

Overall, the simulations did not show strong evidence that vertical integration resulted in lower total costs when asset specificity was present. The overall cost was slightly lower for the contract carriers rather than for private fleets in every instance (dry van versus refrigerated trailer, high versus low volume, and predictable versus unpredictable loadings). The difference between contract and private fleets did decrease as volume increased - suggesting that increased frequency encourages vertical integration.

Beier (1989) suggests and develops the idea that there often exists a significant learning curve between dispatchers and transportation managers at the shipper-carrier interface. Increased familiarity with each other improves their interaction and the operation of the entire process. Any change in this arrangement decreases the efficiency and requires additional time to train the new party. This qualifies as human asset specificity and a contract would be expected - especially since both parties are 'locked-in' to a degree.

Empirically testing Beier's hypothesis of human asset specificity between shippers and carriers, Maltz (1993) examined the shipper's decision between private fleets and for hire motor carrier transportation. He used the results of a survey of 138 companies of which 28 used private carriage, to some degree. The use of contracts with for-hire carriers was not included as a specific option. He concluded that vertical integration is the preferred form of governance if there are human and dedicated asset specificities present. He defined human asset specificity as existing if respondents stated that costs would increase during a transition to another carrier or if it would take another carrier a "great deal" of time to learn to serve the customer. Dedicated

asset specificity was defined by agreeing to the statement “the delivery carrier must customize its service.”³³ Maltz concluded that if a shipper believes that “changing carriers will be an expensive or lengthy process the company will tend to use the private fleet to service that customer” Although there are some definition problems, the research contributes to the idea that when human asset specificity exists, a shipper will try to economize the transaction costs by entering into a more complex governance structure: contracts or vertical integration.

Pirrong (1993) used temporal specificity in order to explain why bulk ocean shipping contracts are formed. Traditional transaction cost theory predicts that only spot markets should exist for bulk shipments since there are none of the traditional specificities (site, physical, dedicated, or human). Shipping contracts (charters) can be classified as spot if the cargo, ship, origin, destination, loading and discharge dates, and payment due to carriers are specified in a very short time, between a day and two weeks, before loading. Pirrong proposes that *temporal specificities* exist where “spatial/temporal proximity is a form of relationship-specific capital” which can be thought of as a form of site-specific capital where the “spatial linkage is fleeting” rather than fixed and durable as in the traditional sense (Pirrong 1993, 943).

Pirrong uses as an example a port with a number of possible carriers available, each a specific time away from the port. Assuming costs are related to travel time (distance), the shipper at the port should form a spot contract with the closest carrier in order to ship the cargo immediately. However, there may be some large quasi rents involved here if (1) delays in shipment cause great harm to the shipper (e.g., opportunity costs, additional setup costs, etc.) or (2) the carrier’s next best cargo shipping opportunity is sufficiently distant (requiring the carrier to incur direct and indirect costs of an empty transit). In order to capture these quasi rents, both the shipper and carrier may expend valuable resources. Letting d be the cost to the shipper for each day of delay in shipping, e be the cost to the carrier for empty transit to the next available load, and c be the variable cost of shipping, the range of negotiation is $[c-e$ to $c+d]$. The carrier will go no lower than the cost minus the empty transit cost while the shipper will pay no more than the delay cost incurred while waiting for the next ship to arrive. If the negotiation range is large, then there may be haggling between the parties to determine the actual price. The size of the negotiation range is due primarily to the fact that “ships cannot move instantaneously and costlessly from point to point” (Pirrong 1993, 943)

When these temporal specificities exist, forward contracting tightens the negotiation range. The carrier must compete against all other carriers that can serve the load within the lead

³³ While the former is reasonable, the definition of dedicated asset specificity is not accurate. Williamson defines dedicated asset specificity as “investment in generalized . . . production capacity that would not be made but for the prospect of selling a significant amount of a product to a specific customer.” (Williamson 1985, 95). This is clearly not the same.

time while the shipper cannot leverage a carrier's close proximity since they will have more opportunities to select from. Forward contracting requires both parties to enter an agreement prior to all of the relevant information being known concerning the value and costs of transportation services. Thus the trade-off between reduced flexibility and avoidance of market exposure must be made.

In summary, previous research has shown that asset specificity is needed in order to necessitate the use of more complex governance structures in the transportation field. Palay and Bienstock showed that *physical asset specificity* can influence shippers and carriers to form some sort of contractual agreements in order to economize transaction costs. Beier and Maltz explored the effect that *human asset specificity* had on contractual agreements and found that it too can provide incentives for some form of complex governance structure - private fleet or contractual. Finally, Pirrong adapts *temporal specificities* to transportation to explain the use of contracts in time-sensitive freight market. Temporal specificities can exist where either the shipper or the carrier can exploit some monopoly power on the other player if one is highly sensitivity to time delays.

Past research has not, however, provided justification for the use of contractual agreements between shippers and carriers for the general-freight dry-van market. The equipment is generic (no *physical asset specificity*), the service is routine (no *human asset specificity*), and the freight is not hyper-time sensitive (no apparent *temporal specificities*). While the next section proposes three network-based explanations, in the remainder of this section we briefly discuss four explanations that are frequently mentioned in practice.

One possible explanation is that contracts are used to hedge against future price changes. While both shippers and carriers may have incentives to contract in order to secure long-term price guarantees, these incentives do not coincide. The shipper has an incentive to lock-in current prices only when the freight market is going to tighten while carriers are interested in contracting when the market is softening. A related justification is the idea that contracts stabilize prices which allows better planning. Because most contracts include some form of price escalation clause for fuel and other inputs, a contract is not a guarantee to freeze rates altogether anyway. While price hedging or stabilizing might explain a portion of the increased use of contracting, it does not seem compelling enough to explain the overall trend. The price fluctuations do appear to influence the timing of when contracts are let out to bid, however.

A second possible explanation of the rise in the use of contracting is the increase of time-definite freight. This can create the *temporal specificities* described by Pirrong (1993). At first glance, though, the large number of available motor carriers, their ease of movement, and the fact that no single carrier is on the "critical path" of a shipper's transport decision appear to

make temporal specificities irrelevant for TL trucking. This is not the case, however, whenever the shipper is hyper-sensitive to delays in shipments, e.g., JIT supplier to a manufacturer. The time scale is now minutes instead of days. While the shipper is always able to locate another possible carrier to handle the load, it involves an additional amount of search time to locate another carrier. Because the time frame is so tight, this additional search time can provide the carrier with a potential for leverage.

This leverage can be seen by a simple example. Let c be the variable transportation cost, e be the carriers deadhead cost in empty miles to next load, d be the cost of delay to the shipper in minutes, and s be the searching time required to locate another carrier. Then the range of negotiation for the shipper and carrier is $[c-e \text{ to } c+ds]$, where the carrier will go no lower than the variable cost minus the cost of the empty miles to a competing load and the shipper will go no higher than the variable cost of the load plus the cost of delay due to having to search for a new carrier. Because the time scale is so tight, even an additional phone call might cause a costly delay at the production facility. The magnitude of a delay penalty usually dwarfs the freight rate. Contracting ahead of time for these services will negate this potential monopoly rent. If these are frequent trips, then there is even more incentive to form some sort of contractual agreement with a carrier as the overhead cost is allocated over a larger transaction base. In these types of contracts, the shipper is guaranteed capacity while the carrier is provided a safeguard in the form of future business.

A third justification used in practice is that contracts are required to hold the carriers to better levels of service - such as on-time delivery or transit reliability. Requiring a high level of service, however, is not a sufficient reason by itself for the formation of contracts. For example, passenger airlines maintain very high service levels and actually compete on on-time delivery and baggage handling accuracy without relying on contracts with passengers.³⁴ The standards of the service are self contained within each transaction and are enforced by the travelers' ability to use alternate carriers. The same is true for package delivery firms. For freight carriers, most of the more standard 'value-added' services, such as in-transit tracking, monitoring of products, or pallet return, are not generally shipper specific and therefore do not require contractual relationships to be performed. A threshold system-wide volume level, however, might be required to justify the costs of implementation of these 'value-added' services.

Finally, a fourth explanation frequently given is that shippers simply have greater market power and are forcing these contracts on the carriers. While the buyer usually has a greater degree of power in most relationships (excepting monopolistic supplier markets), the literature shows that contracting is occurring with the largest TL carriers, seemingly the ones

³⁴ Airlines rely mainly on incentive programs (frequent flier miles) rather than contractual agreements.

with the most clout. The power of large buyers' comes from the effect that their high volume orders have on the suppliers' operations. Thus, leveraging large volume orders pre-supposes the supplier being sensitive to economies of scale. As Chapter 1 showed, TL carriers exhibit only mild economies of scale. Proportionately increasing the volume of traffic awarded to a carrier across the shipper's network will exacerbate any current imbalances. This also explains why carriers are moving away from offering system-wide volume discounts. Market power in TL procurement seems to alternate year to year between shippers and carriers depending on the capacity demand in the market.

While temporal specificity explains why an important segment of the TL market is using complex contracts, it still does not explain why an increasing amount of the general freight, dry van segment is being procured through contracts rather than in the spot market. Hedging, level of service, and market power do not seem to be sufficient answers by themselves.

3.3.2 Network Based Justifications

This section proposes three, somewhat related, reasons for why contracting is the best (transaction cost minimizing) form of governance structure for certain segments of the dry van, general freight TL trucking segment. First, contractual agreements are required for carriers to achieve and reap the savings from economies of scope. Second, bundling a mix of profitable and unprofitable lanes (from the carrier's perspective) within a network can, in some circumstances, lower the total transaction costs. Third, in order to offer both sufficient coverage and consistent service desired by certain shippers, a carrier must have a sustainable network. Such a network cannot be formed in the spot market. Each justification is discussed in turn.

Economies of Scope

Transaction cost analysis theory predicts that as the economies of scope for the transaction increase, a buyer is more likely to rely on the market rather than vertical integration or contractual agreements. In most situations, the theory suggests, market suppliers are better able to combine diverse demands for products in order to achieve economies of scope. Williamson shows that strong economies of scope actually dampen the influence of asset specificity so that market forces are acceptable for goods and services with higher levels of asset specificity if economies of scope or scale are present.³⁵

Economies of scope can actually have the opposite effect in transportation networks - especially when the complimentary lanes are self-contained within a shipper's network. Unlike traditional goods or services, economies of scope in TL trucking can encourage rather than

³⁵ Williamson notes that " Economies of scale and scope thus favor market organization over a wider range of asset specificity values" (Williamson 1985, 93).

hinder the use of contractual arrangements or even vertical integration. Balanced, high volume lanes are usually the last lanes that are converted from private to for-hire fleets. Frito-Lay, for example, still uses its long-haul private fleet to cherry pick the extremely reliable closed tours between plants and distribution centers.³⁶ By using private fleets, a shipper does not have to share the savings realized by economies of scope with a for-hire carrier. However, for-hire carriers will almost certainly have lower labor, equipment, and maintenance costs as well as more efficient overall operations. Thus, vertical integration is not the governance structure which minimizes transaction costs for TL movements.

The spot market becomes increasingly less efficient as the degree of complementarity between the lanes increases. Because the savings are achieved by reducing uncertainty of connection costs, economies of scope requires serving a set of specific lanes. Spot markets, however, treat each load as an independent transaction. Thus, there is no guarantee that a carrier will secure all of the lanes required to realize these savings - even if the shipper is willing to tender out a set of lanes together. This is illustrated through an example.

Suppose that a shipper has a TL network consisting of lanes 1 and 2 moving between locations i and j, as shown in Figure 3.3. Lane 3 is a deadhead move while lane 4 is part of another shipper's network. Lanes 1, 2, and 4 are 500 miles while the deadhead lane, 3, is 200 miles. Carrier A operates over the lanes at a cost of \$1.00 per mile of linehaul without considering any connection costs.

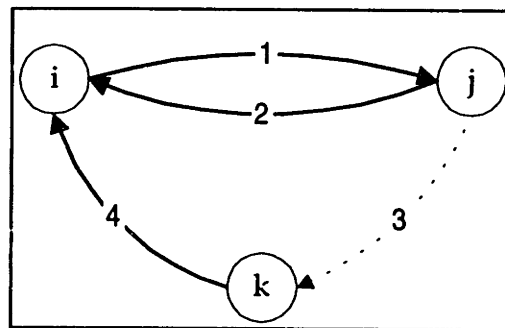


Figure 3.3 Shipper's network for economies of scope example.

Suppose that the shipper tenders a load on lane 1 and Carrier A knows that the next day a load will be available on lane 2 (from j to i) as well as on lane 4 (from k to i). Carrier A knows that it will get at least one of these loads. If both lanes 1 and 2 are tendered to Carrier A, its costs will be \$1,000 (due to a zero connection cost at node j) or \$1.00 per loaded mile. If lanes 1 and 3 are tendered to Carrier A, however, its costs will be \$1,200 (due to the 200 mile deadhead from j to k) or \$1.20 per loaded mile. Rather than considering bidding strategies (as will be done

³⁶ Private conversation with Joseph Dagnese, Director of Transportation, Menlo Logistics.

in Chapter 4), assume that the shipper tenders a load on lane 1 to Carrier A at a price of \$1.10 per mile.

If Carrier A accepts the load, then the series of loads will either generate a profit of \$100 or a loss of \$100, depending on which lane is tendered.³⁷ Because these are transactions in a spot market, each load is tendered independently of the other. So, unless the carrier is willing to risk some potential exposure, the loads will not be accepted at this price. Or, the shipper, unable to secure a load, will increase its rate until a truck is found. Thus, the spot market is not supporting the more efficient assignment and the transaction costs are not minimized.

Suppose now that the shipper jointly tenders out a series of 2 loads (one each on lanes 1 and 2) at the same time for a price of \$1.10 per mile and Carrier A accepts the loads. At the completion of the first leg, however, Carrier A has become "locked-in" to the tour. To see this, assume that another carrier, Carrier B, is also available at location j at the same time and is willing to take the load on lane 2 for \$1.00 per mile. The shipper, being opportunistic, may be willing to tender the load to Carrier B in order to save the \$50 difference. Without a contractual agreement, Carrier A is not safeguarded against the shipper's opportunistic tendencies. The shipper has gained an economic advantage over Carrier A and can either tender the load to Carrier B, or leverage Carrier B's offer to re-negotiate Carrier A's price to, say, \$0.95 per mile. Essentially, by taking the first load, Carrier A has exposed itself to potential opportunistic behavior by the shipper without any safeguards. This is equivalent to the unstable Case II situation shown in Figure 3.1.

Conversely, the shipper could be exposed to these lock-in effects by the carrier. Suppose, now, that the shipper has no alternative carriers available at location j. After completing the load on lane 1, Carrier A is in the position to extract monopoly rents from the shipper. The range of the rates which can be charged are bounded by the delay cost facing the shipper.³⁸

The end result is that when the spot market is used the most efficient form of assignment is not achieved, the potential for costly and unproductive haggling is increased, and the transaction costs are not minimized. Even if tendered as a pair, the time lag between the tender of the first leg and the tender of subsequent legs in a set introduces the potential for opportunistic behavior by both shippers and carriers.

³⁷ This assumes that the price will be paid on all lanes.

³⁸ A distinction should be made between the first case of separate tenders for each lane and the second case of joint tenders of a set of lanes. In the first case, the transaction costs are not minimized because the carrier will not be able to achieve the efficient aggregation of lanes. For the second case, there is potential for opportunistic behavior by both the shipper and the carriers. This is different from temporal specificity in that the lock-in effects are caused by the carrier having served the first leg of the tour.

In practice, the opportunistic behavior is often manifested by the inclusion of additional accessorials which were not disclosed at the time the load was tendered. For example, a shipper might require the carrier to make an inside delivery. Because the carrier has already invested time and effort in the line-haul, it has few options and is thus "locked-in" to the shipper. Conversely, if the freight is highly perishable, the carrier could simply try to out-wait the shipper. Again, this introduces room for unproductive haggling that can be removed by the use of a contract.

While the spot market is unable to realize the economies of scope within a network, the shipper can realize these economies by designing a contract which assigns sets of lanes to a specific carrier. The carrier provides the shipper with a lower freight rate, related to a fraction of the savings due to the economies achieved and, in return, the shipper provides safeguards to the carrier. Safeguards can be in the form of guaranteed volume on these, and possibly other, lanes, or the shipper may simply pay for the carrier's empty miles on this tour. This latter arrangement is typical of dedicated carriers where the shipper maintains operational control of a carrier's vehicle and driver. Dedicated fleets are essentially one step away from having a private fleet.

Interestingly, because there are more hauling opportunities for general freight, the presence of economies of scope is more likely in generic dry van fleets rather than the more asset specific equipment. Thus, increasing the asset specificity of a trailer actually makes it more difficult to achieve economies of scope and, in turn, less likely to be transacted under a contract using the economies of scope argument! Of course, increased asset specificity brings up other factors which encourage the use of contracts.

Lane Bundling

The second explanation of why contracting is becoming more prevalent in situations without asset specificity is the presence of lanes of mixed profitability within a shipper's network. The profitability of a traffic lane refers to whether the rate that the shipper is willing to pay will actually cover the carrier's costs of serving that lane. It is caused by the presence of a gap between shipper and carrier expectations of the transportation costs on a lane. If the shipper's expectation for transportation cost is *higher* than the carriers', then the lane is considered profitable. Conversely, if the shipper's expectation for transportation cost is *lower* than the carriers', then the lane is considered unprofitable.

Gaps between shipper and carrier expectations are primarily caused by the use of cross-subsidies by shippers. While carriers base their cost estimates for each lane on actual practice (probability of reloads, expected empty miles required, etc.), shippers often use planning costs

developed across the firm.³⁹ Because these planning costs are often based on financial accounting costs rather than activity or opportunity based costs, they introduce distortions which will favor (subsidize) some lanes while penalizing (taxing) others. It is these cross-subsidies, then, which distort the shippers' expectations.

Unprofitable lanes are typically associated with poor reload or follow-on load potential - something that cost accounting systems do not capture. For example, an isolated plant with sporadic shipments might require a lengthy deadhead in order to pick up a load. Other examples include plants with excessively long load/unload times or freight that requires special treatment either before pickup or after the delivery (such as washing the trailer or charging the refrigerator unit). Loads on these lanes are considered to be unprofitable because the shipper's planning cost is lower than that of the carriers' actual costs.

When a load over an unprofitable lane is tendered on the spot market, carriers will often delay accepting it. A shipper may have to offer the same load to a number of brokers sequentially and the brokers will typically broadcast this rate to their stable of carriers and owner/operators. At truckstops, it is common to see drivers waiting for the same load to be offered by different brokers because each time it is re-offered the rate per mile increases. Thus, the drivers are exploiting the shipper's need to get the load off of its dock.⁴⁰

While the spot market works, the price eventually rises to the point where a carrier accepts the load, the transactional costs to the shipper are quite high. The repeated calls to different brokers takes time and effort. This effect is magnified if multiple loads need to go to these locations at the same time. One transportation executive noted that each additional truck requires an 'exponential' number of phone calls.⁴¹

Whenever price distortions (caused by cross-subsidies within the shipper's transportation costs) are present, contracts can lower the overall transaction costs by requiring the carriers to bundle the unprofitable and profitable lanes.⁴² The shipper guarantees the carriers the right to haul on several profitable lanes in the network in exchange for providing coverage on the less or un-profitable lanes. The carrier is being guaranteed a system-wide volume level (or perhaps a percentage of volume) as well as some measure of the quality level of the traffic (some guarantee that they won't receive all of the traffic on the unprofitable lanes).

³⁹ These planning costs for transportation are typically used to estimate plant and product profitability and thus are heavily intertwined in the shipper's decision making. Because these planning costs are developed firm wide, changing them is considered to be beyond the scope of the transportation manager.

⁴⁰ This is the same leverage that an electronic market may provide the carriers, depending on the amount of information provided.

⁴¹ This example is taken from a series of interviews with the transportation department of USX, Pittsburgh, PA, June 8, 1994. Conversation with George Bokelberg, USX, June 8, 1994.

⁴² Often times, the reverse procedure is used to justify the use of private fleets. All of the bad lanes are unbundled from the good lanes and offered on the spot market separately. The cost of serving these lanes, with no additional guarantees, will naturally be quite high. These costs are then compared against the private fleets' costs on the good lanes.

Bundling of lanes essentially brings the freight rates closer to the shipper's own cross-subsidized rates. Thus, it will add costs to the lanes with high reload probability while reducing the costs to the lanes that have poor follow-on potential. Of course, a more reasonable method would be to price and cost the lanes properly - without system-wide cross-subsidies - but this is often out of the transportation manager's scope.

Shippers benefit from the bundling because it reduces the search cost required to secure carriers to cover the unprofitable lanes. Even if the rate on the bad lanes is close to the rate paid in the spot market, the search costs are reduced by using a contract, lowering the total transaction costs. For carriers, however, it appears to be more of a concession made in order to secure the additional 'good' freight.⁴³ In order to cover a mix of profitable and unprofitable lanes, carriers essentially charge some average rate and manage the subsidy.

Expansion of the TL Service

A third explanation for increased contractual arrangements between shippers and TL carriers is that the coverage and consistency of service desired by shippers are simply not able to be produced by carriers in the spot market. Carriers need to be able to exploit what Rakowski (1990, 500) refers to as 'economies of markets.' These economies allow the carrier to satisfy the majority of the shippers that are interested in obtaining "a multipoint transportation system that fills a variety of continuing and diverse geographic shipment or travel needs; not a simple point-to-point movement." Essentially, Rakowski describes a form of network externality where the value of a carrier to a shipper increases as the number of locations that carrier serves is increased. This increased demand for wider coverage is one factor contributing to the concentration of the TL market.

The question is, which governance structure is capable of ensuring both system-wide coverage and consistency of service? The spot market is capable of providing sufficient coverage through the use of a broker, but the service will differ based on which carrier the broker calls. Thus, the spot market trades off consistency with increased coverage. Vertical integration, on the other hand, is able to provide consistent service, since the private fleet can handle all of the shipper's traffic. Sufficient coverage can also be provided, but only at a high cost to cover the inefficiencies of a private fleet. This is especially true if the fleet is relied upon to cover the shipper's entire network. This leaves contractual agreements.

⁴³ Interestingly, though, in one core carrier program the carriers were the staunchest defenders of bundling good and bad lanes. The Vice-President of Sales for one carrier in the program stated that by having the shipper determine the allocation of lanes across the system, the carriers in the program have an almost "live and let live" attitude. They do not have to compete against each other and thus can avoid a pricing frenzy which would lower all of their margins.

In order to be able to provide both consistency and coverage in an efficient manner, a carrier must have a sustainable (balanced) service network. A sustainable network implies that there is some degree of certainty in securing follow-on loads at a large number of geographically dispersed locations. By using contracts, the carrier can secure sufficient volume at these locations to achieve economies of scope. The consistency and coverage that shippers desire is only achievable by carriers that have a sustainable network which, in turn, can only be produced by carriers relying on contractual agreements. Thus, contractual agreements are needed to provide both consistency and coverage to shippers.

Summary

The three network based justifications for contracting all argue that a carrier's costs are driven by the network design and that contractual agreements are the only way to guarantee that the network is maintained. It was shown that contracts are needed for the carriers to reap any savings due to economies of scope. The placement of the volume on the network is more critical than the total volume awarded.

Ironically, while shippers are demanding more use of contracts, primarily in conjunction with core carrier programs, it is the carriers that are accruing the majority of the benefits in terms of operational efficiencies. Carriers have more to gain by contracting than shippers. While shippers do accrue benefits, they are not as critical as the carrier's. Consider, for example, which party offers the incentive to form a contract. Carriers are expected to lower their rates when entering into a contract. While some carriers quip that "partnership is just another word for rate reduction," this is actually very true. Carriers do accrue a greater degree of the benefit by having a more secure and reliable source of loads and are forced to share these efficiencies with a shipper in order to entice them into a contract.

If contracts truly benefited shippers more than, or even just as much as, carriers, then carriers would be able to maintain the same rates or even demand a slight increase in rates for forming partnerships. Interestingly, the Schneider-LOF-Wabash contract described in Section 3.1.4 is the exception that proves the rule since in this contract the per mile transportation prices actually went up for LOF!⁴⁴

3.3.3 Other Transportation Modes

In order to better understand why procuring TL transportation is different from most other modes, in this section we briefly compare and contrast procurement strategies for both consolidated and direct carriers.

⁴⁴ The transportation and logistics rate per glass unit, however, went down.

We have argued that contracts are required in order to guarantee that a TL carrier actually hauls the requisite traffic needed to reap the operational efficiencies gained from economies of scope. TL carriers are not alone in this situation; all direct forms of transportation are sensitive to economies of scope. Thus, any direct mode of transportation is expected to use contractual agreements with shippers for the same reasons. Transaction costs are minimized when economies of scope are realized and contracts are required to ensure that the carrier can actually achieve these economies.

Consolidated carriers, however, are different. Because consolidated carriers generally operate in accordance to schedules, the uncertainty of finding follow-on loads loses its relevance. Thus, economies of scope do not have a significant influence over consolidated carriers. Instead, these carriers are very sensitive to shipment density in the pick up and delivery operations. Unit costs for an LTL or parcel delivery carrier, for example, decreases as the number of shipments per customer location or the number of customers per area increases. This is because the main cost driver for a stop at a location is the time and distance required to arrive there - not the time at the stop. The marginal cost of loading an additional package at a location is small and serves to allocate the delivery's fixed costs over a larger base. Similarly, if a carrier can increase the number of shipping locations in a constrained area (such as different buildings in a college campus), then their unit costs will also decrease since the distance between customers will decrease. Again, the fixed cost of traveling between buildings is small compared to driving to a more distant location for a pick up or delivery.

The differences in the economics driving the different modes determines the strategies that the shipper should follow during procurement. For direct transportation modes, the shipper should encourage the carriers to establish economies of scope by allowing them to select which combinations of lanes to serve. For consolidated modes, on the other hand, shippers should leverage the consolidated carriers' economies of density on the local movements. Consolidated carriers have the economic incentive to serve all of the traffic that a shipper generates. Shippers, then, can only influence a consolidated carrier's costs by increasing the density offered. Again, it is the carrier that accrues the majority of the benefits from these contracts in the form of more efficient operations.

3.3.4 Mixed Governance Structures

A shipper will often need to use more than one type of governance structure across their TL network. There is no single form of relationship which makes sense for all shippers or even for all segments of a single shipper's network. The shipper, then, should select the governance structure which matches the different components of the network. Chapter 5 specifically examines how the characteristics of a distribution network influences the contracting, bidding,

and assignment procedures. A discussion and analysis of which type of governance structure to employ in these situations is delayed until then.

Interviews with a number of shippers⁴⁵ revealed that some leading companies are employing a variety of contractual agreements within their networks. For example, PPG utilizes a dedicated fleet for the most heavily concentrated lanes and a core carrier base for the remainder. James River also operates tiered contracting system with a set of “continuous move carriers” that operate on rates which decline with the total length of a dispatch, a larger set of base carriers, and a list of excess carriers which are only called in emergencies. The base carriers negotiate fixed volume commitments for locations where James River will pay for the trucks if they are not used. Similarly, Reynolds Metals contracts with its carriers to provide two different types of capacity at locations; dedicated (which is paid for if not used) and floating, which is not. The duration of the contracts also vary according to the type of relationship. Dedicated fleets might have (evergreen) contracts with no specific termination date while core carriers might need to re-bid for their lanes annually.

3.4 Closure

The objective of this chapter was to provide background on shipper-carrier relationships and to explain why contractual agreements are becoming the dominant governance structure over vertical integration (private fleets) and competitive markets (spot markets).

Private fleets, while still accounting for over half of the TL market (measured in imputed revenue), are difficult to justify anymore. Private carriage lost its cost advantage after deregulation and any perceived service advantage is being wiped out by the increased competition and service differentiation strategies being pursued by some of the upper tier carriers. Several carriers are specifically targeting private fleet conversion as a prime strategy. This conversion process typically transforms private fleets (owned and controlled by the shipper) into dedicated fleets (owned by a carrier and at least partially controlled by the shipper). Essentially, the conversion decouples ownership from control. This form of contractual agreement is just a step away from private carriage. There is no clear line demarcating where private carriage stops and contractual agreements begin.

At the other end of the governance structure spectrum, spot markets, too are actively being used in the TL segment. Fueling the spot market are a very large number of very small carriers and brokers that provide smaller shippers the basic TL commodity service. Additionally, larger shippers use the spot market as a back up for their own private fleet or for

⁴⁵ These included Circuit City, Baxter Medical, Ethyl, Gillette, Goodyear, Hamilton Beach, Home Depot, IBM, James River, PPG Industries, Reynolds Metal, and USX. The bidding and assignment processes for these companies were already described earlier in Chapter 2

their carrier base. This market will continue to operate as long as there are small carriers willing to accept a low return on their investment. Additionally, this channel provides a way for both shippers and large carriers to add variable capacity. In the words of one analyst, the spot market and the small carriers within it serve the same purpose as the temporary help industry does.⁴⁶

Much of this spot market, however, is entering into more of a contractual arrangement as well. The larger brokers are forming contracts, albeit quite weak ones, with favored carriers while shippers are contracting out for carriers to provide backup for the system as a whole. These contracts are just a step away from the spot market with minor liability distinctions and minimal volume guarantees (usually 3 loads per year).

Traditional Transaction Cost Analysis (TCA) theory correctly predicts the governance structure for many situations between shippers and TL carriers. For example, TCA theory predicts that vertical integration is not an adequate governance structure for industries that require a significant amount of innovation - such as TL since deregulation. Other examples include using contracts where highly asset specific equipment is used (e.g., Schneider-LOF-Wabash) and for hyper-time-sensitive freight under JIT programs (due to temporal specificity). TCA theory does not, however, justify the use of contractual agreements for the general-freight dry van market - where the use of contracts is steadily increasing. The dry van market has all of the markings of a perfectly competitive commodity market.

We proposed a series of network based explanations as to why this contracting occurs. Essentially, we argue that carriers are the prime beneficiaries of contractual agreements since they allow carriers to achieve economies of scope. The spot market cannot efficiently support the procurement of sets of lanes due to the temporal nature of the follow-on loads. That is, the carrier must perform each delivery in sequence so that they can potentially become "locked-in" to a series of lanes and be in a position to be exploited by the shipper, or vice versa. Other reasons provided for contracting include the bundling of different quality lanes and in cases where wide and consistent coverage is required.

Contracting, then, is shown to be the governance structure which minimizes the total transaction costs for both shipper and TL carrier. The next chapter examines how the bidding process should be designed to take advantage of the carriers' underlying cost structures.

⁴⁶ Phone conversation with Carolee Harlin, Transportation Technical Services, Fredericksburg, VA, 14 August 1995.

Chapter 4

Auction Design for TL Bids

The objective of this chapter is to critique current TL bidding methods and to develop a framework for shippers to use in selecting how to bid out their truckload network. Because shippers' networks differ in terms of size, traffic volume, lane characteristics, and potential for achieving economies of scope, a single type of auction designed for all TL bids is not practical. Instead, the auction should be designed to match the network being bid out. This means that a variety of types of auctions need to be considered. The contribution of this chapter is to present and condense a rather large body of literature and adapt it specifically to the TL bidding problem faced by shippers. This is important because the way in which shippers have traditionally conducted their bids has not been rigorously examined in this light.

The chapter is organized as follows. Section 4.1 restates the problem in terms relevant to the bidding and assignment decision. The problems of incentive compatibility and interdependency and the role of auctions in mitigating these problems are discussed. Section 4.2 addresses three key concepts in auction theory which are particularly relevant to the TL bidding problem: reservation value, allocational efficiency, and equilibrium conditions. These concepts are needed to understand the design trade-offs. Section 4.3 discusses four auction design decisions that are relevant to the TL bidding problem. These are the order in which the lanes should be bid out, whether to use simple or conditional bids, how many rounds should be used in the auction, and how the final lane rate for each lane should be determined. Section 4.4 addresses the details of combinatorial auctions, which use conditional bids, to include a discussion of general concepts and examples of several cases of combinatorial auctions. Finally, Section 4.5 closes the chapter by summarizing the major points.

4.1 General Problem and Notation

The objective of a TL bid is to find the minimum cost assignment of carriers to traffic lanes within the shipper's distribution network. The problem is how to partition a set of lanes, X , among K different carriers such that the total cost of serving the distribution network is a

minimum. The optimal assignment of lanes to carriers, X^* , is the lowest cost allocation of lanes to carriers. It is in the shipper's best interest to assign carriers that have the best cost structure to their network. The assignment, X^* , solves the optimization problem (PC):

$$\begin{aligned} \text{Min}_{X_k} Z_{PC} &= \sum_{k=1}^K C_k(X_k) \\ \text{s.t.} & \\ (1) \quad \cup_{k=1}^K X_k &= X \\ (2) \quad X_i \cap X_j &= \emptyset \quad \forall i, j \end{aligned} \quad \text{[PC]}$$

where:

Indices:

k : Carriers bidding on lanes, $k = 1, \dots, K$

Decision Variable:

X_k : Set of lanes assigned to carrier k

Data:

$C_k(X_k)$: Cost function for carrier k given the assigned lanes

X : Set of all lanes to be assigned

Objective Function Value:

Z_{PC} : Total cost for allocation of lanes

The objective function minimizes the sum of the cost functions for each carrier. Due to economies of scope in carrier operations, the cost functions are generally not linear. Constraint (1) ensures that the partitioning of the lanes, X_k , for each carrier is collectively exhaustive while constraints (2) ensure that the partitions are mutually exclusive. In this model, each lane is assigned to one carrier, but a carrier may serve multiple lanes. In practice, this could be modified by setting a separate carrier requirement for each lane, or by assigning carriers a number of loads per lane. For this chapter, we will only consider the case of assigning a single carrier to each lane.

Unfortunately, the shipper does not know the cost functions for the different carriers. The carriers themselves generally do not know their own costs for serving multiple lanes in a network. Instead, the shipper has to use the bids submitted by the carriers as proxies for their cost structures and solve problem (PB):

$$\begin{aligned} \text{Min}_{x_{ik}} Z_{PB} &= \sum_{i=1}^N \sum_{k=1}^K b_{ik} x_{ik} \\ \text{Subject to:} & \\ (1) \quad \sum_{k=1}^K x_{ik} &= 1 \quad \forall i \\ (2) \quad x_{ik} &= \{0,1\} \quad \forall i, k \\ (3) \quad x &\in W \end{aligned} \quad \text{[PB]}$$

where:

Indices:

- i : Lanes in the network, $i = 1, \dots, N$,
 k : Carriers involved in the bid, $k = 1, \dots, K$,

Decision Variable:

- x_{ik} : = 1 if carrier k is awarded lane i , = 0 otherwise,
 x : Allocation of lanes to carriers,

Data:

- b_{ik} : Bid for lane i submitted by carrier k ,
 W : Set of all feasible solutions - all other constraints,

Objective Function Value:

- Z_{PB} : Total shipper cost for assignment

Problem [PB] is a set partitioning problem. The objective function minimizes the sum of the individual lane bid rates. Constraints (1) ensure that only one carrier is assigned to each lane; (2) ensure integrality; and (3) contains all side constraints for the TL problem, such as a limit on the number of carriers in the system, carrier capacity requirements, and minimum carrier coverage within a system. If the side constraints designated by W are ignored, this problem is trivial to solve since each bid covers only one lane. A simple ranking of the bids submitted for each lane determines the low cost carrier for each lane, which is how the auction is typically performed in practice.

While solving PB is straightforward from a mathematical programming perspective, there are two additional complications which make it interesting for auction design. First, there is an *incentive compatibility* problem. Because the model uses bid rates submitted by carriers, as the coefficients in the objective function, the solution to [PB] is only as good as the bid rates submitted. Simply asking carriers for their true valuations of each lane, however, does not mean that they will actually provide them. In fact, carriers usually have an incentive to submit bids higher than cost-plus-margin depending on the relative strength of the shipper and the competing carriers. The auction should provide the carriers with an incentive to report their true reservation values for the traffic lanes.

Second, there is an *interdependency* problem due to economies of scope in TL trucking. While the problem [PB] uses a separate bid for each lane, the actual cost for a carrier to serve a lane depends on the other lanes being served. Because the value of a lane is dependent on the other lanes, it is difficult, and sometimes impossible, to obtain an efficient allocation using simple bids as shown in [PB]. Allowing conditional bids which treat package of lanes as all-or-nothing packages increases the difficulty of solving [PB] and complicates the auction process.

The incentive and the interdependency problems can be reduced by careful design of the auction. Before discussing the specific design decisions, some fundamental concepts in auction theory are first introduced.

4.2 Fundamental Auction Concepts

An auction is a method of valuation and allocation. The valuation or 'price discovery' process is most useful when the market value of an item is not established or is not suited for a single fixed price. Examples include the value of a fresh catch of fish which is very time dependent, or rare art.⁴⁷ The ability of an auction to allocate items is an offshoot of its ability to determine value. As such, auctions are typically used for exchanges of items with uncertain value (e.g., livestock, tulips, government bonds, television rights to an event, and professional sports contracts), scarce supply (e.g., landing slots at an airport), or both (FCC auctions for spectrum licenses).

The TL bidding problem is primarily an allocation problem in which the shipper awards traffic lanes to carriers according to some mechanism. There are several ways that a shipper can do this, besides holding an auction. For example, the shipper could randomly assign carriers to lanes or use a lottery for carriers to select which lanes to serve. A more realistic alternative is an administrative process whereby a traffic manager or transportation committee selects carriers based on direct negotiation. All three of these allocation mechanisms are used in other areas. Random assignments are often used in classroom assignments, lotteries are frequently used for housing and dorm room assignments, and committees are frequently used to allocate resources in government programs.

Experiments conducted at the California Institute of Technology⁴⁸ have directly compared different types of allocation mechanisms and have found that a well designed auction consistently outperforms these other mechanisms in achieving efficient allocations, especially for more complicated problems. In an efficient allocation, the items being auctioned are awarded to those bidders that value them the most. For the TL bidding problem, this means to assign traffic lanes to the carrier with the lowest cost structure for those lanes.

This discussion of auctions for the TL bidding problem differs from that typically found in the literature in two respects. First, while most of the auction design literature discusses cases where the auctioneer is a seller awarding items to a bidder (buyer) for the *highest* buying price bid, the TL bidding problem is the letting of a contract. The auctioneer (the shipper), then, is searching for the *lowest* bid for each item (lane). Second, the TL bidding problem establishes an ongoing relationship rather than a discrete transaction of goods. This auction determines from which carriers the shipper will procure transportation services over an extended period of time. There is a value to maintaining goodwill with the contracted carriers. So, finding the absolute

⁴⁷ McAfee and McMillan (1987, 701) ask "how can one discover the worth of an original copy of Lincoln's Gettysburg Address except by auction method?"

⁴⁸ See, for example, Ledyard, Noissair, and Porter (1995), Brewer and Plott (1995), Olsen and Porter (1994), Ledyard (1993), or Banks, Ledyard, and Porter (1989).

minimum cost carriers is not always desired. The whole idea behind using conditional bids is to remove inefficiencies from the carrier's operations, not to erase each carrier's margins.

For the remainder of the chapter, it is assumed that lanes are the items being auctioned, the shipper is the auctioneer, and the bidders are carriers. While third parties, such as non-asset based contract logistics providers, are not excluded from these bidding activities in practice, for the purpose of this discussion they are an unnecessary complication.

This section discusses three fundamental concepts important to auction design for the TL bidding problem: (1) carriers' reservation value of the lanes, (2) efficient allocations, and (3) equilibrium conditions.

4.2.1 Reservation Value

A carrier's reservation value for a specific traffic lane is the lowest bid that that carrier is prepared to make to acquire that lane from a shipper, given the current information. This concept is used widely in auction design theory. While analytically convenient to treat the reservation value as a known scalar, it is actually more of a statistical distribution. The reservation value for a lane, for example, is a composite of several factors, none of which are known with great certainty. For this discussion, we can break down a carrier's reservation value for a lane into two components: operational and strategic. The operational component is related to the actual cost of hauling traffic over the lane while the strategic component deals with the profit margin for that lane. The carriers' total reservation value for a traffic lane is the sum of these components.

The shipper should be concerned about the carrier's reservation values for each of the lanes since the submitted bids will be based on these values. By definition, the submitted bid for a lane will not be lower than the reservation value for that lane.⁴⁹ In designing the TL bidding auction the shipper should assist the carriers in achieving the lowest possible reservation values for their lanes. Additionally, the auction design should encourage the carriers to bid their true reservation values. These are not necessarily the same goals - the former concerns enabling carriers to determine the true reservation values for themselves while the latter concerns having them actually use these values in the bidding process. This is not to say that the lanes will be awarded at the reservation value prices or that a carrier is guaranteed to win every lane it submits its reservation value for, just that it is more efficient to allocate the lanes using 'true' reservation values than using biased ones.

The design of an auction affects these two components differently. Operational factors are significantly influenced by the characteristics of the shipper's network. Strategic factors are

⁴⁹ The winner's curse, introduced later in this section, applies when the reservation value is incorrectly estimated - too low for low-bid-win auctions and too high for high-bid-wins auctions.

influenced more by external or environmental effects, such as the strength of the competing carriers, the shipper's negotiating credibility, and the importance of the shipper's business to the carrier. The operational and the strategic factors are each discussed in turn.

Operational Factors

The operational component of reservation values reflects the cost of TL operations on a traffic lane. Recall from Chapter 1, the primary costs for TL carriers are the line-haul and repositioning costs. The shipper can improve the accuracy of the carrier's estimation of these costs by providing accurate and detailed information on each lane. For carriers, better information means providing both spatial and temporal distributions for the traffic flows in the network as well as details on the exact requirements for accessorials. The more vague or aggregated the information is, the more the carrier will have to assume.

As an example, consider the distribution network depicted in Figure 4.1. The distribution center serves customers by six outbound lanes and is supplied by plants via two inbound lanes. The shipper is bidding out the six outbound lanes and can present the information to the carriers in several different ways, to include:

- Annual number of loads leaving the distribution center,
- Annual number of loads going to 3 digit ZIP code regions,
- Annual number of loads per lane,
- Average monthly/weekly number of loads per lane, or
- Expected daily distribution of loads per lane.

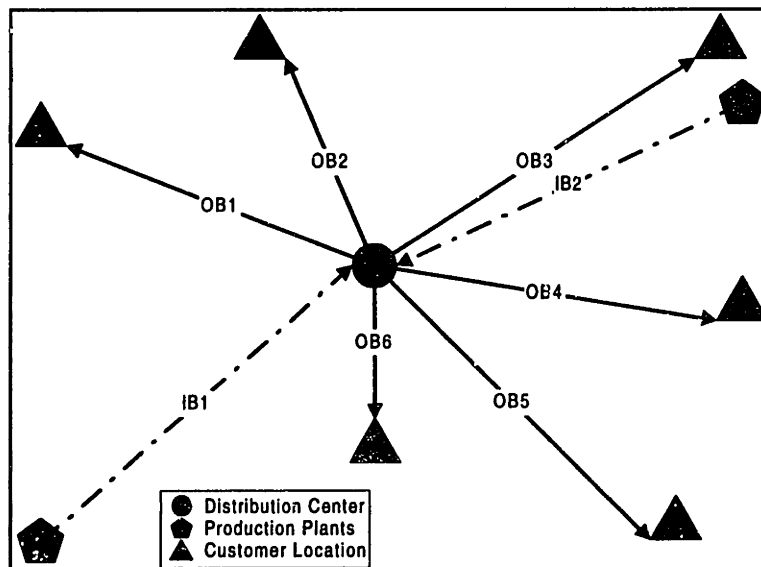


Figure 4.1. Sample distribution network.

Suppose that the shipper provides the carriers with just the annual number of loads originating from the distribution center. For the carrier to estimate a rate to charge for hauling

on a lane, it would need to make a spatial assumption on the split between the lanes as well as a temporal assumption on the timing of these loads. While even the shipper does not know the split with certainty, the shipper is in a much better position to estimate it (or provide historical values) than the carriers are. The net effect of these assumptions is that the carrier will need to hedge its reservation value in order to be reasonably sure that their costs will be recovered.

Continuing the example, suppose the shipper instead provides the carriers with the aggregated number of loads per year for each lane. While better than the previous case, the carrier still needs to make a temporal assumption on the seasonality of the loads. The carrier can assume, at one extreme, that the loads are distributed uniformly across the entire year, or at the other extreme, that all loads move in one month. The former assumption is overly optimistic and can lead to underbidding while the second assumption can lead to an overbid so that the carrier does not win the lane at all.

Carriers that are unfamiliar with the shipper's operations that make overly optimistic assumptions will end up winning a lane, but incurring a loss. This is referred to as the "winner's curse"⁵⁰ where the carrier that wins the lane in an auction is the one that most underestimated its value. Put another way, the carrier submitting the lowest bid: (1) is working with the thinnest profit margin, (2) truly has a lower cost structure than the other carriers, or (3) has underestimated its reservation value of that lane. In cases where carriers are unfamiliar with the shipper's network or the information is not detailed, the third outcome is very likely.

It is not in the shipper's best interest to have carriers bid below their reservation prices and operate at a loss. While bidding out the lanes at a very low cost might appear to be initially beneficial to the shipper, it is a short term gain. Having carriers serving at a loss can lead to future service problems, a lack of priority by the carrier, and a likely default in service later on which most likely will be more costly than any rate savings.

Better information leads to better informed bids, reduces the likelihood of the winner's curse and lessens the incumbent carrier's advantage over new carriers. The incumbent carriers' main advantage is that they do not need to make as many assumptions about the shipper's operations.

Providing more information, then, will not necessarily lower the submitted bids. In fact it might lead to higher bids from carriers that would have made overly optimistic assumptions as to the spatial and temporal distribution of the loads. It will lead to bids that are more closely linked to the carriers' cost structures, and, in the long run, should save money by avoiding costly service failures.

⁵⁰ The concept and term winner's curse was introduced by Capen, Clapp, and Campbell (1971).

Simply providing accurate and detailed information on the network will not, however, ensure that carriers will achieve economies of scope by securing specific lane packages. Unless a set of lanes can be guaranteed, a carrier will need to hedge their bid to cover for uncertainties in the connection costs associated with each lane. Hedging for uncertainties in finding follow-on loads can be reduced by the use of conditional bids. Other methods that can work under certain conditions are discussed later in the chapter.

Strategic Factors

The goal of decreasing the operational elements of the reservation value was to reduce the hedging for uncertainty (spatially, temporally, and for follow-on loads) in the traffic flow. For the strategic factors, however, shippers want to minimize the carriers' ability to inflate their margins. While the shipper should not be trying to drive the carriers' margins to zero, it should, through the auction design, remove any additional padding. Two sources of padding that can be reduced by auction design are the carrier's evaluation of the shipper's strategic worth and the shipper's credibility.

The strategic worth of the account is based on perceptions of the competitive strengths between the carrier and the shipper. As such, it influences the margin a carrier considers. The margin on a lane can range from contributing to overall profitability, recovering fully allocated costs, recovering just the operating costs, or even operating at a loss as part of strategic underpricing for certain lanes. While the shipper cannot readily influence the carriers' estimations of its strategic worth, it can allow enough flexibility in the auction so that carriers can re-adjust their positioning as needed. This means that carriers could lower the acceptable margin on a lane in the middle of an auction in order to win it.

A carrier has the incentive to pad its reservation value based on a shipper's lack of credibility in negotiation. If a shipper is known to re-negotiate with carriers after the auction, then a carrier will add a buffer to its reservation value in order to protect it from the expected price concession. The lower a shipper's credibility, the higher the initial bids will be.

Reservation Value versus Bid Price

The shipper should design the auction so that carriers minimize the operational hedging and the strategic padding of reservation values. To do this, the shipper should encourage 'truth-telling' on the part of the carriers, by providing incentives for them to bid their 'true' reservation values. To do this, the auction has to be designed such that each carrier's dominant strategy is to bid their true reservation value rather than a strategically inflated estimate. A dominant strategy is defined as a course of action by a bidder that yields better results than any other alternative regardless of what any other bidder does.

Not all auction designs have this characteristic. The strategy for a carrier in a single-round, sealed-bid auction, for example, is to choose a bid price, b , that maximizes the expected value of the payoff.⁵¹ The payoff can be expressed as the cumulative probability that b will be lower than all other n bids multiplied by the profit margin of $(b - r)$, where r is the reservation value for serving that lane. Carriers in this situation have offsetting incentives. Increasing their bid, b , improves the contribution if the bid is won, but decreases the likelihood of winning. There is no incentive for a carrier to bid its true reservation value for the lane. In fact, there is an incentive to bid above his true reservation price as long as it is just slightly below the next highest bidder's price. The carrier's payoff is maximized at this point. Because the carriers do not know the other carriers' reservation values, bids in a single-round, sealed bid will not generally reflect the true reservation values and determining an efficient allocation will be more difficult.

The importance of providing incentives for carriers to submit their true reservation values rests on the revelation principle in game theory. The revelation principle states that if the dominant strategy for players is to "tell the truth," then no other mechanism can produce better results for either the shipper or the carriers (Kreps 1990, 700).

4.2.2 Efficient Allocation

The final outcome of an auction is an efficient allocation if each item is awarded to the bidder which valued it the most. For the TL bidding problem, an efficient allocation is one in which carriers are awarded those lanes where they have the lowest reservation value. Note that this is not necessarily the same as awarding lanes to the carriers with the lowest bids. In practice, it is impossible to determine whether an auction outcome is the efficient allocation since shippers will never know the carriers' reservation values.⁵² However, the concept is important in testing the design of an auction using sample data. If an auction design is incapable of achieving an efficient allocation in a small test sample, then it will probably not be able to do so in practice.

In experimental studies, the efficiency of an allocation is measured by comparing the reservation values of the actual assignment to one where the items are awarded to the lowest cost bidders. The allocational efficiency of a "low-bid-wins" auction is defined as:⁵³

⁵¹ This discussion is intended to provide a general guideline and makes several assumptions such as all bidders being risk neutral, symmetric bidders, etc.

⁵² Realistically, even the carriers may not know their own reservation values.

⁵³ In the literature, this is usually shown for "high-bid-wins" auctions and the measure is the inverse. In either case, $E=1$ for an efficient allocation.

$$E = \frac{\sum_k R_k(X_k^*)}{\sum_k R_k(X_k^-)}$$

where:

E: Efficiency of the auction,

$R_k(X_k^-)$: Reservation value for carrier k to serve the lanes actually assigned, and

$R_k(X_k^*)$: Reservation value for carrier k to serve the lanes where it is the low cost carrier.

The efficient allocation is the assignment of lanes to carriers with the lowest cost structures in those lanes, that is, where $E=1$. For example, suppose that two carriers (A and B) are bidding on two lanes (1 and 2) with the reservation values shown in Table 4.1.

| | Carrier A | Carrier B |
|--------|-----------|-----------|
| Lane 1 | \$10 | \$5 |
| Lane 2 | \$6 | \$8 |

Table 4.1. Reservation values for carriers.

The efficient allocation is for Carrier B to win lane 1 and Carrier A to win lane 2, $\sum R_k(X_k^*)=\$11$. Suppose, however, that the final bids are as shown in Table 4.2 with Carrier B winning both lanes.

| | Carrier A | Carrier B |
|--------|-----------|-----------|
| Lane 1 | \$14 | \$9 |
| Lane 2 | \$10 | \$9 |

Table 4.2. Current bids for carriers.

The total cost to the shipper is \$18 but the total reservation value of this allocation is $R_k(X_k^-)=\$13$. The efficiency is $E = 11/13 = 0.84$. Efficiency is measured using the reservation values, not the current bids.

By the revelation principle, if the auction induces the carriers to be truthful in their bidding, then the efficient allocation results in the allocation which can achieve the lowest cost to the shipper. There is a distinction between the allocation and the total cost - the same allocation can yield very different total costs depending on the design of the auction. In other words, just because the allocation is efficient does not mean that the price the shipper will pay is equal to the winning carriers' reservation values. But, since the winning carriers are the best equipped to handle these lanes, the shipper will be better off in the long run.

4.2.3 Equilibrium Conditions

An allocation is said to be in equilibrium if no carrier has an incentive to submit a lower bid for any of the lanes. That is, no carrier can unilaterally increase their profitability by submitting a lower bid on any of the lanes. If a current allocation is not in equilibrium, then the bidding will continue since a carrier can improve its net utility or profitability. The equilibrium

concept is important because it tells us whether certain desired outcomes are actually achievable by the auction. Bykowsky, Cull, and Ledyard (1995) proposed very similar conditions for an ascending bid auction. They defined a local Nash equilibrium as “a stationary point in the auction at which no agent can unilaterally improve its net utility.” The methodology in this section follows their analysis. These equilibrium conditions are also closely related to the Nash equilibrium conditions required for optimal contract design.⁵⁴ We will simply refer to these conditions as equilibrium conditions. The equilibrium concepts can be shown by the way of an example which will be used throughout this chapter.

Example - Case 1: No Interdependencies

Suppose a distribution network has three nodes and three lanes as shown in Figure 4.2. There are three carriers, A, B, and C, bidding on the lanes. Table 4.3 lists their reservation value for individual lanes as well as the value for certain efficient aggregations or packages of lanes. For this first example, the value for the packages are equal to the sum of the stand alone reservation values. Carrier A, for example has a reservation value of 11 for lane 1 while Carrier C has a reservation value of 12 for the same lane. There are 6 total potential bid packages. Packages 1, 2, and 3 are simply the stand alone lanes while package 4 represents awarding lane 1 and 2 together, package 5 represents 2 and 3 together, and package 6 represents lanes 1 and 3 awarded together. These are the recognized efficient aggregations although for this first example, we are showing no savings. Note that Partition #7 = {1,2,3} is not used here.

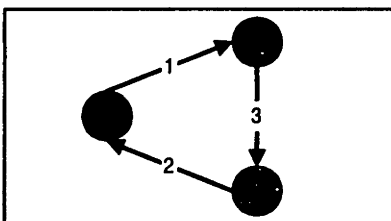


Figure 4.2. Sample 3 node network.

| Carrier Bid Package Res. Value (r _i) | A | | | B | | | C | | |
|--|----|---|----|---|----|----|----|----|----|
| | 1 | 2 | 4 | 2 | 3 | 5 | 1 | 3 | 6 |
| Lane 1 | 11 | 8 | 19 | 9 | 11 | 20 | 12 | 12 | 24 |
| Lane 2 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| Lane 3 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 |

Table 4.3. Carrier reservation values. Case 1:
No Interdependencies

⁵⁴ See, for example, Kreps (1990, 684-91).

The end result of the auction is that the lanes will be allocated to the different carriers. We can represent the potential allocations as three arrays $A_{ij} = \{P_i, S_{ij}, R_{ij}\}$ consisting of a partitioning array, P_i , an assignment array, S_{ij} , and a reservation value array R_{ij} . The partition describes in which packages the lanes are awarded. There are four feasible (mutually exclusive and collectively exhaustive) partitions of bid packages in this example: $P_1 = \{1,2,3\}$, $P_2 = \{1,5\}$, $P_3 = \{2,6\}$, and $P_4 = \{3,4\}$. The numbers in the brackets represent the packages that are awarded. Each partitioning of the lanes can be assigned to different carriers. For this example, there are eight possible assignments for the first partition and two for each of the other three partitions. From Table 4.3, recall, that all of the carriers did not bid on all of the lanes. Finally, each assignment within a partition has an associated cost array representing the reservation values (or bids when used in practice). All outcomes of the auction can be represented by the allocations, shown below.

$$\begin{array}{cccc}
 A_{11} = \begin{bmatrix} 1 & 2 & 3 \\ A & A & B \\ 11 & 8 & 11 \end{bmatrix} & A_{12} = \begin{bmatrix} 1 & 2 & 3 \\ A & A & C \\ 11 & 8 & 12 \end{bmatrix} & A_{13} = \begin{bmatrix} 1 & 2 & 3 \\ A & B & B \\ 11 & 9 & 11 \end{bmatrix} & A_{14} = \begin{bmatrix} 1 & 2 & 3 \\ A & B & C \\ 11 & 9 & 12 \end{bmatrix} \\
 A_{15} = \begin{bmatrix} 1 & 2 & 3 \\ C & A & B \\ 12 & 8 & 11 \end{bmatrix} & A_{16} = \begin{bmatrix} 1 & 2 & 3 \\ C & A & C \\ 12 & 8 & 12 \end{bmatrix} & A_{17} = \begin{bmatrix} 1 & 2 & 3 \\ C & B & B \\ 12 & 9 & 11 \end{bmatrix} & A_{18} = \begin{bmatrix} 1 & 2 & 3 \\ C & B & C \\ 12 & 9 & 12 \end{bmatrix} \\
 \\
 A_{21} = \begin{bmatrix} 1 & 5 \\ A & B \\ 11 & 20 \end{bmatrix} & A_{22} = \begin{bmatrix} 1 & 5 \\ C & B \\ 12 & 20 \end{bmatrix} \\
 A_{31} = \begin{bmatrix} 2 & 6 \\ A & C \\ 8 & 24 \end{bmatrix} & A_{32} = \begin{bmatrix} 2 & 6 \\ B & C \\ 9 & 24 \end{bmatrix} \\
 A_{41} = \begin{bmatrix} 3 & 4 \\ B & A \\ 11 & 19 \end{bmatrix} & A_{42} = \begin{bmatrix} 3 & 4 \\ C & A \\ 12 & 19 \end{bmatrix}
 \end{array}$$

Allocation A_{31} , for example, consists of Carrier A winning lane 2 and Carrier C winning lanes 1 and 3 while allocation A_{32} uses the same partition but Carrier B wins lane 2 instead of Carrier A. The first subscript designates the partition and the second specifies the assignment. From inspection, we can see that A_{11} and A_{11} are the lowest cost or efficient allocations. Because there are no interdependencies in this example, there is no difference between the two. Carrier A wins lanes 1 and 2 while Carrier B wins lane 3 for a combined total reservation value of 30 units.

Let us assume that the auction is being conducted in multiple-rounds where carriers submit a bid for interested lanes and the shipper announces the current winning price for each lane in between rounds. Note that bids are only allowed to cover a single lane at a time - that is bids for multiple lanes are not allowed.

At each successive round, the carriers consider the current bids for each lane and decide whether to submit a lower bid or not. Conceptually, the carriers compare the profitability of the current allocation to all possible other allocations at the current bid rates:

$$\sum_{i \in X_k} b_i - R_k(X_k) \leq \sum_{i \in X_k^-} b_i - R_k(X_k^-)$$

where:

- X_k : Current allocation of lanes to carrier k,
- X_k^- : All other possible allocations of lanes for carrier k,
- b_i : Standing bid for lane i, current minimum of all bids submitted for lane i,
- $R_k(X_k)$: Reservation value for carrier k for current lane allocation X_k and
- $\pi(X_k)$: $= \sum_{i \in X_k} b_i - R_k(X_k)$. Expected payoff for carrier k if awarded lanes in partition X_k .

If the condition holds, then some other allocation, say, X_k' , would yield Carrier k a higher payoff. Thus, Carrier k would be able to bid up to the difference, $\pi(X_k') - \pi(X_k)$, in order to win those lanes contained in the allocation X_k' . If the condition does not hold, then Carrier k cannot improve its payoff by further bidding. Because this is an auction, the bids will be monotonically decreasing so that once a carrier cannot improve its payoff, it will not be able to do so in any future rounds either. The only exception is if the carrier decides to change its reservation value in the middle of the auction; this change will always be negative. As long as one carrier has an incentive to bid, the auction will continue. At the point where no carrier can improve their payoff by bidding lower, an equilibrium point is reached. There is no guarantee, however, that this is a unique equilibrium.

The next question is whether the efficient allocation is sustainable at an equilibrium point. If so, the following equilibrium conditions must hold for all carriers:

$$X^* = X^o \quad \text{if}$$

$$\sum_{i \in X_k^*} b_i - R_k(X_k^*) \geq \sum_{i \in X_k^-} b_i - R_k(X_k^-) \quad \forall k, X^-$$

where:

- X^* : Efficient allocation of all lanes,
- X^o : Equilibrium allocation of lanes, not necessarily unique,
- X_k^* : Efficient allocation of lanes to carrier k,
- X_k^- : All other possible allocations of lanes for carrier k,
- b_i : Current bid for lane i, and
- $R_k(X_k)$: Reservation value for carrier k for current lane allocation X_k .

By using these equilibrium conditions, we can find the set of lane bids (b_1 , b_2 , and b_3), where the efficient allocation, $A^* = A_{41}$, is also an equilibrium point, A° , and is thus achievable. For our example, the following conditions must hold in order for the efficient allocation A_{41} , (Carrier A wins lanes 1 and 2 and Carrier B wins lane 3) to be an equilibrium point:

Carrier A:

| | | | |
|------|--------------------------------|------------------------------------|---------------------|
| (1a) | $b_1 - 11 \leq b_1 + b_2 - 19$ | if not, A will bid on lane 1 alone | $8 \leq b_2$ |
| (1b) | $b_2 - 8 \leq b_1 + b_2 - 19$ | if not, A will bid on lane 2 alone | $11 \leq b_1$ |
| (1c) | $0 \leq b_1 + b_2 - 19$ | if not, A will not bid at all | $19 \leq b_1 + b_2$ |

Carrier B:

| | | | |
|------|--------------------------------|-------------------------------------|--------------------|
| (1d) | $b_2 - 9 \leq b_3 - 11$ | if not, B will bid on lane 2 alone | $2 \leq b_3 - b_2$ |
| (1e) | $0 \leq b_3 - 11$ | if not, B will not bid at all | $11 \leq b_3$ |
| (1f) | $b_2 + b_3 - 20 \leq b_3 - 11$ | if not, B will bid on lanes 2 and 3 | $b_2 \leq 9$ |

Carrier C:

| | | | |
|------|-------------------------|-------------------------------------|---------------------|
| (1g) | $b_1 - 12 \leq 0$ | if not, C will bid on lane 1 alone | $b_1 \leq 12$ |
| (1h) | $b_3 - 12 \leq 0$ | if not, C will bid on lane 3 alone | $b_3 \leq 12$ |
| (1i) | $b_1 + b_3 - 24 \leq 0$ | if not, C will bid on lanes 1 and 3 | $b_1 + b_3 \leq 24$ |

which defines the following region:

| | | |
|------|-----------------------------------|------------------------------|
| (1j) | $11 \leq b_1 \leq 12$ | from (1b) and (1g) |
| (1k) | $8 \leq b_2 \leq 9$ | from (1a) and (1f) |
| (1l) | $11 \leq b_3 \leq 12$ | from (1e) and (1h) |
| (1m) | $19 \leq b_1 + b_2 \leq 21$ | from (1c) and (1f)+(1g) |
| (1n) | $30 \leq b_1 + b_2 + b_3 \leq 33$ | from (1c)+(1e) and (1f)+(1g) |

The region defined by these constraints is the set of all feasible equilibrium points for the auction where Carrier A wins lane 1 and 2 and Carrier B wins lane 3. It is referred to as the core of the assignment and is shown in Figure 4.3. The core is comprised of those possible outcomes that are "coalitionally stable," (Shubik 1985, 254). That is, any outcome not contained in the core can be improved upon by some group of carriers. The specific location on the core depends on the details of the bidding process, such as the minimum bid increments. However, for the efficient allocation to occur, it must be bounded by the core, or else some carrier has an incentive to continue bidding.

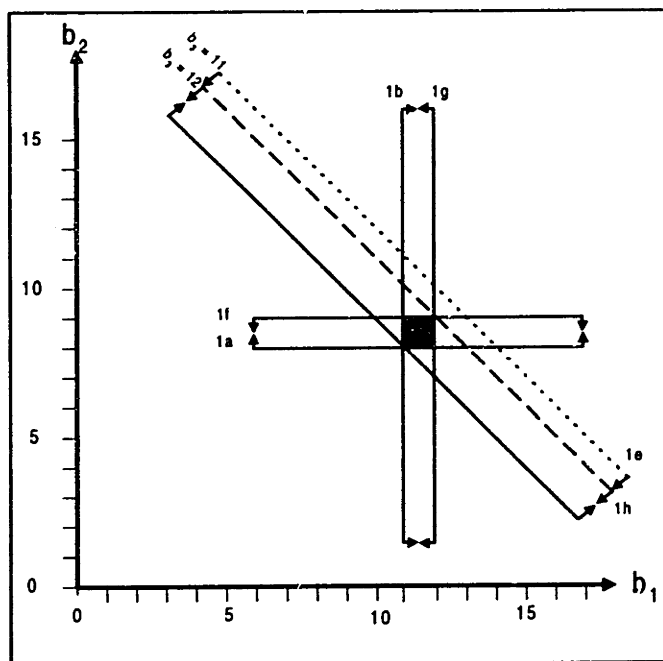


Figure 4.3. Bounding constraints for lane bid prices. Shaded area represents core of assignment problem. Case 1: No Interdependencies

When there are no interdependencies, the carriers do not need to assume any risk in arriving at the efficient allocation. To see this, consider a possible auction as shown in Figure 4.4 and Table 4.4. Table 4.4 lists the winning bids for each lane for five rounds. Figure 4.4 shows the trajectory of the total of the bids received at the end of each round, R_i .

| | Lane 1 | Lane 2 | Lane 3 |
|---------|--------|--------|--------|
| Round 1 | | | |
| Bids | 15 | 13 | 15 |
| Carrier | C | A | C |
| Round 2 | | | |
| Bids | 14 | 12 | 14 |
| Carrier | A | B | B |
| Round 3 | | | |
| Bids | 13 | 11 | 13 |
| Carrier | C | A | C |
| Round 4 | | | |
| Bids | 12 | 10 | 12 |
| Carrier | A | B | B |
| Round 5 | | | |
| Bids | 12 | 9 | 12 |
| Carrier | A | A | B |

Table 4.4. Auction summary - Case 1: No Interdependencies.

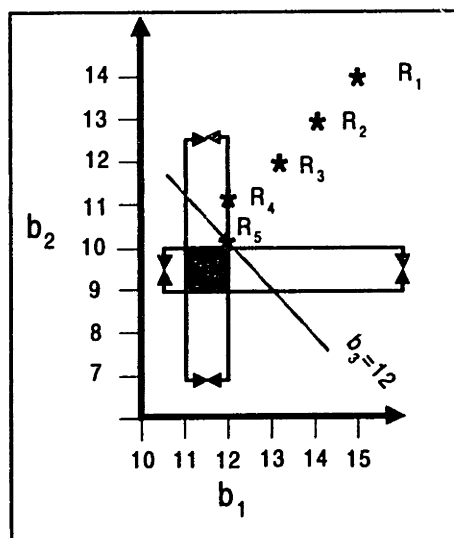


Figure 4.4. Trajectory of bids - Case 1: No Interdependencies.

During each round, a carrier makes a new bid only if it will improve their profit - assuming that all other bids will remain the same. For example, in Round 3, Carrier C wins lanes 1

and 3 by bidding a total of \$26 for an expected profit of \$2. In the next round, the lanes are won by lower bids by Carrier A and B for \$12 each. Because Carrier C cannot improve its current profit level, it does not bid any further and the solution is at point R₃ in Figure 4.4.

The final winning bid for a lane, then, is the reservation value of the next lowest (second-price) bid. At that point no other carrier has an incentive to bid lower and still improve their profitability, thus, the bidding stops when the surface of the core is reached. Recall that the core is defined as those outcomes where no carrier is able to improve its profitability. For this example, Carrier A receives profits of \$2 and Carrier B profits of \$1.

Example - Case 2: Minor Interdependencies

The same example can be made more interesting by introducing slight interdependencies. Table 4.5 lists the new reservation values. All of the values are the same except for lane packages 4, 5, and 6 which have been reduced by 10 to 15 percent below the sum of their respective individual stand alone rates.

| Carrier Bid Package | A | | | B | | | C | | |
|------------------------|----|---|----|---|----|----|----|----|----|
| | 1 | 2 | 4 | 2 | 3 | 5 | 1 | 3 | 6 |
| Res. Value | 11 | 8 | 17 | 9 | 11 | 19 | 12 | 12 | 21 |
| Lane 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| Lane 2 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| Lane 3 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 |

Table 4.5. Reservation values for carriers. Case 2: Minor Interdependencies.

The same allocations are feasible and the efficient allocation is once again for Carrier A to win lanes 1 and 2 and Carrier B to win lane 3 for a combined total cost of \$28. The following equilibrium conditions need to hold for the efficient allocation to be at equilibrium:

Carrier A:

- (2a) $b_1 - 11 \leq b_1 + b_2 - 17$ if not, A will bid on lane 1 alone $6 \leq b_2$
- (2b) $b_2 - 8 \leq b_1 + b_2 - 17$ if not, A will bid on lane 2 alone $9 \leq b_1$
- (2c) $0 \leq b_1 + b_2 - 17$ if not, A will not bid at all $17 \leq b_1 + b_2$

Carrier B:

- (2d) $b_2 - 9 \leq b_3 - 11$ if not, B will bid on lane 2 alone $2 \leq b_3 - b_2$
- (2e) $0 \leq b_3 - 11$ if not, B will not bid at all $11 \leq b_3$
- (2f) $b_2 + b_3 - 19 \leq b_3 - 11$ if not, B will bid on lanes 2 and 3 $b_2 \leq 8$

Carrier C:

- (2g) $b_1 - 12 \leq 0$ if not, C will bid on lane 1 alone $b_1 \leq 12$
- (2h) $b_3 - 12 \leq 0$ if not, C will bid on lane 3 alone $b_3 \leq 12$
- (2i) $b_1 + b_3 - 21 \leq 0$ if not, C will bid on lanes 1 and 3 $b_1 + b_3 \leq 21$

which gives the following core which is shown in Figure 4.5:

- (2j) $9 \leq b_1 \leq 10$ from (2b) and (2i)-(2e)
- (2k) $6 \leq b_2 \leq 8$ from (2a) and (2f)

- | | | |
|------|-----------------------------------|------------------------------|
| (2l) | $11 \leq b_3 \leq 12$ | from (2e) and (2h) |
| (2m) | $17 \leq b_1 + b_2 \leq 18$ | from (2c) and (2f)+(2i)-(2e) |
| (2n) | $28 \leq b_1 + b_2 + b_3 \leq 29$ | from (2c)+(2e) and (2f)+(2i) |

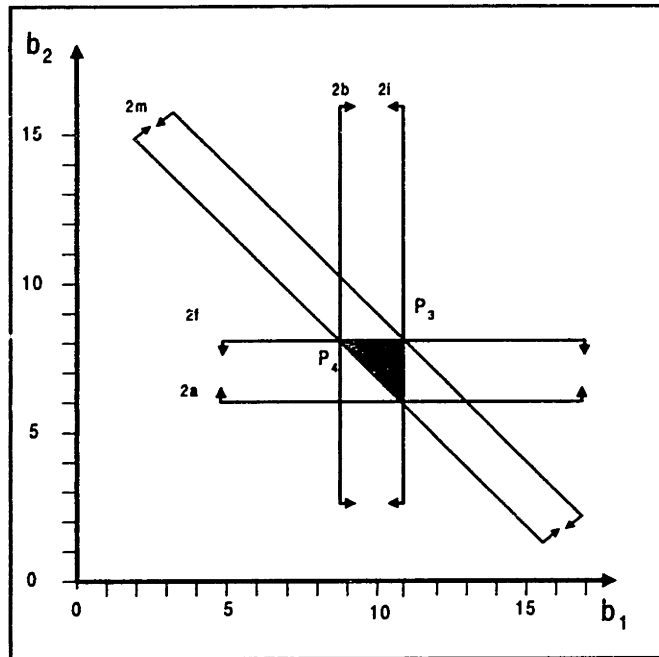


Figure 4.5. Bounding constraints for Example - Case 2: Minor Interdependencies.

An important difference between this situation and Case 1, where there were no interdependencies, is that carriers may risk "exposure." A carrier risks being exposed when under-bidding on a portion of the lanes in a set in the hopes of acquiring the entire set of lanes at a lower total reservation value. If the carrier fails to collect all of the lanes in the set, then they run the risk of winning lanes below their reservation value and incurring the winner's curse. The same potential exposure risk was shown in Chapter 3 to exist in the spot market whenever loads were tendered sequentially. The situations are analogous. For example, consider the auction shown in Table 4.6 and Figure 4.6.

| | Lane 1 | Lane 2 | Lane 3 |
|---------|-----------|-----------|-----------|
| Round 1 | | | |
| Bids | 14 | 12 | 15 |
| Carrier | A | A | C |
| Round 2 | | | |
| Bids | 13 | 11 | 14 |
| Carrier | C | B | C |
| Round 3 | | | |
| Bids | 12 | 10 | 13 |
| Carrier | A | A | B |
| Round 4 | | | |
| Bids | 11 | 9 | 12 |
| Carrier | C | B | C |
| Round 5 | | | |
| Bids | 10 | 8 | 11 |
| Carrier | A | A | B |

Table 4.6. Auction summary -
Case 2: Minor Interdependencies.

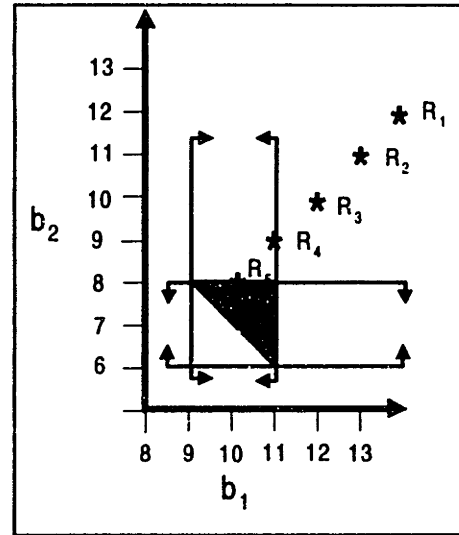


Figure 4.6. Bid trajectory -
Case 2: Minor Interdependencies.

In this example, for the efficient allocation to also be an equilibrium allocation, Carrier A needs to be able to obtain lanes 1 and 2. In round 4 of the auction, Carrier A had to bid below the stand alone reservation cost for lane 1 in order to win both lanes 1 and 2 and achieve the savings. With small levels of interdependency, it is still possible for carriers to achieve an efficient allocation with simple (one lane per bid) bids, albeit while assuming a small degree of risk. The risk is due to the carrier having to bid separately on the two lanes in the hope that they will obtain both lanes in that lane package.

If Carrier A were risk averse, then, perhaps, they would not have bid below the stand alone reservation value for lane 1. This would prevent the efficient allocation from being formed. The result is that bidding would end with a total cost one dollar higher with Carrier C winning lanes 1 and 3 and Carrier A winning lane 2. The loss in efficiency due to the positioning costs is one dollar, in this example. As shown in Section 4.3.2, conditional bids can be used to guarantee that efficient aggregations are formed.

Before leaving the topic of equilibrium conditions, a couple points should be made concerning the core and the relationship to the linear programming model PB, shown again below.

$$\text{Min}_{x_{ik}} Z_{PB} = \sum_{i=1}^N \sum_{k=1}^K b_{ik} x_{ik}$$

Subject to:

- (1) $\sum_{k=1}^K x_{ik} = 1 \quad \forall i \quad \text{[PB]}$
- (2) $x_{ik} = \{0,1\} \quad \forall i, k$
- (3) $x \in W$

where:

Indices:

i : Lanes in the network, $i = 1, \dots, N$,
 k : Carriers involved in the bid, $k = 1, \dots, K$,

Decision Variable:

x_{ik} : = 1 if carrier k is awarded lane i , = 0 otherwise,
 x : Allocation of lanes to carriers,

Data:

h_{ik} : Bid for lane i submitted by carrier k ,
 W : Set of all feasible solutions - all other constraints,

Objective Function Value:

Z_{PB} : Total shipper cost for assignment

Observation 1. Assuming that an efficient allocation is supported by an equilibrium and the bidding is made at very small (infinitesimal) increments, then the final solution will always end at an extreme point of the core.

The bidding will only decrease while there are at least two participants. The bidding will stop when the second to the last bidder's reservation value is reached making that equilibrium constraint tight. A point in the middle of a face of the core can always be improved by some carrier submitting a lower bid on one of the lanes. Thus the final solution will always be at an extreme point.

Observation 2. The location of the final solution on the core determines how the total profits or welfare is distributed.

Recall that the auction, while ending up at different extreme points of the core, will always consist of the same efficient allocation. The extreme points of the core that are closer to the origin, then, will have the lowest total value. This means that the shipper recoups most of the profits. On the other end, at the points farthest from the origin (the "North Eastern" points) carriers receive all of the profits.

Observation 3. A currently winning bid in the auction is equivalent to a basic variable in the PB.

This is by definition. The reduced cost of a variable in PB is $\underline{b}_{ik} = b_{ik} - \sum \mu_i$, where μ_i is the dual value of lane i . The reduced cost of the basic variables is 0, so, this simply means that the bid rate is equal to the clearing price for each lane. The auction can be thought of as a column generation problem with PB as the main problem and the carriers acting as the sub-problem. Before each round, the carriers generate new bids or columns which are then added to the main problem and solved generating new market clearing prices for the lanes. For simple bids, the bid rate is the clearing price while for conditional bids it becomes more complicated.

4.3 Auction Design Decisions

The design of an auction can significantly affect the outcome. Simply by changing the rules slightly, auctions of identical items and bidders may result in different allocations and valuations of the items.

As an example, consider the case where a single lane is being bid out to three potential carriers using one of two different types of auctions: Dutch and English. In a Dutch auction, the auctioneer begins by calling out a low price and continually raises it until one of the bidders accepts it, stopping the auction. This process can also be automated where a large digital readout shows the increasing asking price and the first bidder who signals acceptance wins at that price. This form of auction originated in the Netherlands as a method of selling lots of cut flowers. The critical features of the Dutch auction are that the asking price is known by all bidders and increases continuously, but no information on other bidder's valuations are provided. Thus, each bidder submits bids based solely on their own internal or private valuations.

In an English auction, the auctioneer opens the bidding at a high value and accepts increasingly lower bids until there is only one remaining bidder. The item is awarded to the last remaining bidder at the current bid. This is sometimes called an oral auction since bids were traditionally called out or announced by the auctioneer. It can also be performed with bidders submitting prices electronically or silently either in a continuously or in a series of rounds. English auctions are the most familiar form of auction and they are widely used for artwork and antiques. The critical features of the English auction are that the price is continually descending and that at any point in time all of the bidders know the level of the bidding. Thus, each bidder receives market information from the other participants which can be used to complement their own private valuations. Additionally, if the identification of the bidders is also known then different degrees of credence can be placed on each bidder's valuations.

Continuing the example, suppose the three carriers have a range of reservation values as shown in Table 4.7. Carrier A is the incumbent carrier and having detailed knowledge of the shipper's operations can estimate a fairly tight range for their reservation bid. Carrier B is less familiar and has a larger range. Finally, Carrier C is new to this shipper and has a very large variance for its reservation value due to a lack of knowledge of the shipper's operations. With more information this range could be tightened. For example, the low end of the estimate might be based on the assumption that the shipper's unload time is as they report it, say, less than an hour, while in reality it may average over 4 hours. This is something that the incumbent carrier, Carrier A, already knows. Thus, Carrier C's estimate of \$0.85 is based on a false and optimistic assumption, however, they have no way of figuring this out before the auction.

| | Carrier A | Carrier B | Carrier C |
|-----------------------|-----------|-----------|-----------|
| High Reservation Cost | \$1.10 | \$1.10 | \$1.15 |
| Low Reservation Cost | \$1.05 | \$0.95 | \$0.85 |
| Status | Incumbent | Familiar | New |

Table 4.7. Reservation Value for carriers for single lane.

The results of the two different auctions might be as follows. If the Dutch auction were used, the initial asking price would start very low, say \$0.75, and increase in small increments until a carrier accepted the asking price. In this case, Carrier C, using only its own internal values, would most likely be the winning bidder at either \$0.85 or, if they wanted to wait a little longer, \$0.90. If an English auction were used, the bidding would start at a high asking price, say, \$1.50 and would decrease in small increments based on each carriers' bids. Carriers have no incentive to submit bids more than the minimum required increment below the previous bid. After several rounds, as the bid prices approached \$1.05, Carrier C might notice that the bidding slows down and that even the incumbent does not bid lower. This might make Carrier C reconsider its valuation of the lane. It could even make them not bid below Carrier B's final bid of \$0.95. Thus, one result could be that Carrier B wins the lane at a bid price of \$0.95. Of course, if Carrier C's low reservation value is based on a real savings in their cost structure, rather than uncertainty of the shipper's operations, then they would be willing to bid lower.

The simple example shows that the auction design can have significant influence over the final allocation and value of the items. The Dutch and the English auctions are two standard types of auctions which have many counterparts. The Dutch auction is similar to a single-round, sealed-bid auction typically used in TL bids. The English auction is similar to a multiple-round, sealed-bid auction, provided that some amount of information is provided between rounds. This provides market information to the other bidders which can be used to adjust their own private valuations. The primary difference between these two general auction designs is that the latter allows for market information while the former does not.

This section discusses the major design decisions for the carrier bidding process. While there are many design decisions for an auction, four were thought to be the most important for the TL bidding problem: (1) the ordering of the lanes, (2) the type of bids to allow, (3) the number of rounds to use, and (4) how to determine the final lane rate or the pricing mechanism.

4.3.1 Order of Auction: Sequential versus Simultaneous

The order of the auction determines the sequence in which the traffic lanes are bid out. The two extreme options are to auction off each lane separately (sequential) and to auction off all of the lanes at once (simultaneous). Between these two extremes are different hybrids which combine various elements of each. Each form is discussed in turn.

Sequential Auctions

In a pure sequential auction, each lane is bid on separately. The typical art auction is sequential; the bidding for one painting does not start until the bidding on the previous painting has closed. The main appeal of sequential auctions is that they simplify the auction process. Instead of worrying about all of the lanes at the same time, the carriers can concentrate on one lane at a time. This allows for more detailed and accurate bids.

There are several negative points, though. If there are many lanes to be bid out, it would be extremely time consuming to auction off each lane by itself. This can be mitigated by grouping lanes as discussed later in this section. A far more damaging problem, though, is the increased risk of overexposure. This is relevant whenever interdependencies between the lanes exist. Interdependencies between lanes can be either complements (where the cost of serving a set of lanes is lower than the costs of different carriers serving each lane separately) or substitutes (due to a capacity constraint for the carrier). When lane interdependencies exist, the specific ordering of the lanes will determine the value placed on each lane and the risk assumed by the different carriers. This is best shown by an example.

Example - Case 3: Minor Interdependencies by Sequential Auction

Using the same example from Section 4.2.3, as shown in Table 4.8, suppose that the lanes are to bid out sequentially in the order 3, 2, 1. The bidding on lane 3 will most likely be quite intense since both Carriers B and C can use it to achieve an efficient aggregation. If Carrier B wins lane 3, then it has the possibility to pair it with lane 2 to achieve a savings of \$1 under the stand alone reservation bids for lanes 2 and 3. Similarly, Carrier C can pair lane 1 with lane 3 to achieve a savings of \$3 below the sum of the stand alone reservation bids.

| .Carrier | A | | | B | | | C | | |
|-------------|----|---|----|---|----|----|----|----|----|
| Bid Package | 1 | 2 | 4 | 2 | 3 | 5 | 1 | 3 | 6 |
| Res. Bid | 11 | 8 | 17 | 9 | 11 | 19 | 12 | 12 | 21 |
| Lane 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| Lane 2 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| Lane 3 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 |

Table 4.8. Reservation values. Case 3: Minor Interdependencies by Sequential Auction

Since both Carriers B and C need lane 3 in order to achieve these efficient aggregations, they would most likely be willing to bid slightly below their stand alone reservation values. Suppose that Carrier C wins the lane with a bid of \$10. Carrier B cannot justify bidding below \$10 for lane 3 even if paired with lane 2. Carrier C is willing to go as low as \$10 since it still has \$1 of combined savings left if it also wins lane 1. So, lane 3 is awarded to Carrier C. Lane 2 is put up to bid, next. Since lane 3 is not available any more Carrier B cannot go below its stand alone reservation price and Carrier A can win the lane for \$9. Finally, when lane 1 is auctioned,

Carrier A, having already secured lane 2, can bid as low as \$9 and still be profitable. Carrier C, on the other hand, can only bid as low as \$11 to reach break-even, or they might be willing to bid as low as \$9 for lane 1 so that their total loss would still be \$2. Regardless of whether lane 1 is awarded to Carrier A or C at \$9, the end result is that Carrier A has a profit of \$1 and Carrier C has a loss of \$2.

Because of the auction ordering, Carrier C was forced to concentrate the potential savings of winning lanes 1 and 3 on the first lane that came up for bid; lane 3. This increases the risk of overexposure. Once lane 3 was won, Carrier C needed to win lane 1 in order to try to break even. Likewise, by winning lane 3, it significantly decreased the competition for lane 2 - since Carrier B no longer had significant interest in it. This exacerbated Carrier C's problem since Carrier A could now pick up lane 2 at a profit. Note that if the sequencing was changed to 1, 2, 3 an entirely different outcome would most likely arise.

While this is a contrived example, the problem is very real. In a pure sequential auction, the value of the lanes are very dependent on the specific ordering due to the interdependencies. Similarly, a capacity constraint creates problems because it forces carriers to treat lanes as substitutes. For example, suppose a carrier has a hierarchy of lane preferences (lane 2 is preferred to lane 1) and a capacity of a certain number of lanes that it can serve. Suppose that at some point in the auction, the carrier has the capacity for just one more lane and lane 1 is the next to be bid. If the carrier does not win lane 1, it will bid on lane 2, but if it wins lane 1, it will not bid on lane 2. Thus, they act as substitutes and the lane valuations are dependent on the ordering.

Simultaneous Auctions

In a pure simultaneous auction, all lanes are bid on at the same time. The primary benefits from holding a simultaneous auction are the increased flexibility and decreased risk involved. By bidding on all lanes at the same time, a carrier does not need to commit to one lane at a time. If there are multiple rounds to the auction, the carrier can switch strategies in the middle of the auction if one of the key lanes of an efficient aggregation falls below some threshold value. This lowers the risk of committing prematurely to a lane. While a carrier will still have substitute lanes due to, say, a budget constraint in the form of total capacity available, the simultaneous auctions do not require that the carrier commit for one lane before the other.

Also, by being able to bid on all of the lanes in a package at the same time, the risk can be spread out. If an efficient aggregation consisting of n lanes has a total savings from the sum of the stand alone rates of x dollars, the carrier can reduce its bids on each lane by up to x/n instead of committing some portion of x to one lane at a time, as required in sequential auctions.

The previous example, as worked out in Section 4.2.3, illustrates that a simultaneous auction can achieve an efficient allocation. However, it does not remove all of the inherent risk.

The only downside to simultaneous auctions is the potential complexity. If the shipper has a lot of lanes open for bid, a carrier may find it overwhelming to bid on all of the lanes at the same time. This could reduce the attention paid on each lane and opportunities may be lost. This is especially true if the auction has multiple rounds.

Hybrids

Both sequential and simultaneous auctions have positive and negative attributes. Sequential auctions allow carriers to concentrate on limited numbers of lanes while simultaneous auctions provide carriers greater flexibility and decreased risk. In order to capitalize on the strengths of each, two hybrid orderings can be used.

Hybrid 1: Group Lanes then Bid Out Sequentially

If the interdependencies between certain lanes are easily contained in only one reasonable aggregation, then it might make sense to group these lanes together and treat them as a single item. Additionally, if these groups of lanes do not interact, and there are not too many of them, they can be auctioned sequentially. For example, suppose a shipper has facilities located in four separate regions across the United States with minimal intra-facility shipments. Suppose further, that while the total volume out of each facility is sufficient, no individual lane has a reliable volume level. In this case, the shipper might want to group together all lanes by region and bid out each region as a single item. The regions would then be bid out in sequential order.

Hybrid 2: Simultaneously Bid Out Groups in a Sequential Order

This alternative differs from Hybrid 1 in that after grouping together lanes that might have interdependencies, all lanes within each group are bid out simultaneously. This is a sequential ordering of simultaneous auctions. Using the same example, the shipper would run 4 simultaneous auctions (one for each region) in sequential order. This is a reasonable course of action if the individual lanes within a region have more reliable lanes and can be bid out separately.

Another time this ordering might be used is when different trucking segments are auctioned out. Thus, simultaneous auctions are run for refrigerated, flatbed, and dry van lanes separately in some sequential order, even though some carriers will bid on all three segments. The majority of the carriers are single mode, so that adding in the other modes may only affect a

small percentage of the larger carriers. The interdependencies involved in common segments will usually outweigh the added complexity and confusion of including all segments in one bid.

An exception to this is when the different mode networks have strong complements. For example, one shipper that had previously bid out its temperature controlled network separately from its dry van network noticed that very high volume closed tours could be formed with a dry van forehaul and a temperature controlled backhaul. The tours were included in the temperature controlled carrier bid since they could handle both legs of the tour.

4.3.2 Types of Bids: Simple versus Conditional

In all of the discussions thus far, the bids have been assumed to be simple, that is, a separate bid was required for each lane. This is reflected in the objective function of the model PB shown in Section 4.1. Another option is to allow the use of conditional bids.⁵⁵ As the name implies, a conditional bid includes multiple lanes which are awarded as an-all-or-nothing package. The bid price is conditioned on the awarding of all of the lanes within the package, not just a portion of them. Whenever conditional bids are used, the auction is referred to as a combinatorial auction since there are so many different ways to combine the individual lanes into packages.

A bid on a group of lanes as described in Section 4.3.1 is not a conditional bid since those lanes are only considered in one grouping. In the example, all of the lanes in a region were bid on as a whole and not allowed to be repackaged into any different form, such as half of the lanes from region one combined with half of the lanes in region two. For conditional bids, the lanes involved need to be offered in at least two separate packages. If there is no conflict over how to combine the lanes, then the lanes should simply be grouped and treated as one lane using simple bids. However, when the same lane can be used in several different tours or as a stand alone lane, then the shipper cannot *a priori* determine to which package it should belong. In this case, a bid covering that lane would be termed conditional since there are alternative ways in which it can be combined with other lanes.

This seemingly simple twist (allowing multiple lanes on a bid) creates serious difficulties for the bidding process and therefore conditional bids should only be used if there are strong enough interdependencies to warrant the added complexity. The pros and cons of both simple and conditional bids are discussed in turn.

⁵⁵ These type of bids are also referred to as package bids, complex bids, or combination bids in the literature. The term conditional bids portrays the critical distinction, that it is an-all-or-nothing assignment. It also allows for other types of conditions to be included such as bids based on achieving a total volume or serving certain areas. Thus, it is a more general term.

Simple Bids

Virtually all auctions use simple bids. They are straightforward for the shipper to use and easy for the carriers to interpret. Finding the lowest cost carrier for each lane is trivial with simple bids; one need only sort the bids and find the lowest in each lane. If the auction has multiple rounds, the carrier knows immediately the amount their bid needs to be reduced to beat the current winning bid. Overall, simple bids should be used as long as economies of scope are weak.

Whenever interdependencies are strong, however, simple bids can have several negative qualities. First, simple bids still require that carriers take risks in acquiring a package of lanes by bidding on each separately, as seen in Case 2 of the example. While the risk can be spread across all of the lanes within the package, assuming it is a simultaneous auction, there is still an element of exposure risk associated with being caught with only a portion of an efficient aggregation and not having control over which portion. In most efficient aggregations for TL carriers, the value of a partial tour is just the sum of its stand alone components, that is, it loses almost, if not all, of its savings. The actual amount of the risk depends on which lanes are won. For example, suppose the tour is four lanes in a closed tour and that the total savings is 15% if all of the lanes are secured. If only two lanes are won and they are adjacent then the savings due to having a reload might be 5%. If the two lanes are not adjacent, then there would be no “back-up” configuration for other savings.

Second, when there are strong interdependencies there is no guarantee that simple bids will result in the efficient allocation. This is due to simple bids being unable to support equilibrium conditions when lane values are strongly influenced by other lanes.⁵⁶ This problem was identified by Koopmans and Beckman (1957) in the context of a facility location problem. They found that when the cost of locating a facility became interdependent on the locations of the other facilities, that no individual rents were able to promote decentralized decision-making. This can be shown through a continuation of the example introduced in Section 4.2.3.

Example - Case 4: Strong Interdependencies

The table of reservation bids for the three carriers for the three lanes and various combinations are shown in Table 4.9.

* The same concept also explains why using a spot market for certain TL lanes is not efficient.

| | | | | | | | | | |
|-------------|----|---|----|---|----|----|----|----|----|
| Carrier | A | | | B | | | C | | |
| Bid Package | 1 | 2 | 4 | 2 | 3 | 5 | 1 | 3 | 6 |
| Res. Values | 11 | 8 | 15 | 9 | 11 | 16 | 12 | 12 | 19 |
| Lane 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| Lane 2 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| Lane 3 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 |

Table 4.9. Reservation values. Case 4: Strong Interdependencies

The reservation values for the conditional bids are now just 20% below the sum of the respective stand alone values. This means that Carrier A achieves a savings of \$4 for pairing lanes 1 and 2, Carrier B saves \$4 for pairing lanes 2 and 3, and Carrier C saves \$5 if lanes 1 and 3 are paired. Even with this level of interdependency, simple bids fail to reach an equilibrium that supports the efficient allocation.

From inspection, the efficient allocation is once again Carrier A winning lanes 1 and 2 and Carrier B winning lane 3 for a combined total value of \$26. Setting up the equilibrium conditions, we obtain the following set of constraints:

Carrier A:

- (3a) $b_1 - 11 \leq b_1 + b_2 - 15$ if not, A will bid on lane 1 alone $4 \leq b_2$
(3b) $b_2 - 8 \leq b_1 + b_2 - 15$ if not, A will bid on lane 2 alone $7 \leq b_1$
(3c) $0 \leq b_1 + b_2 - 15$ if not, A will not bid at all $15 \leq b_1 + b_2$

Carrier B:

- (3d) $b_2 - 9 \leq b_3 - 11$ if not, B will bid on lane 2 alone $2 \leq b_3 - b_2$
(3e) $0 \leq b_3 - 11$ if not, B will not bid at all $11 \leq b_3$
(3f) $b_2 + b_3 - 16 \leq b_3 - 11$ if not, B will bid on lanes 2 and 3 $b_2 \leq 5$

Carrier C:

- (3g) $b_1 - 12 \leq 0$ if not, C will bid on lane 1 alone $b_1 \leq 12$
(3h) $b_3 - 12 \leq 0$ if not, C will bid on lane 3 alone $b_3 \leq 12$
(3i) $b_1 + b_3 - 19 \leq 0$ if not, C will bid on lanes 1 and 3 $b_1 + b_3 \leq 19$

which results in the following infeasible core

- (3j) $7 \leq b_1 \leq 8$ from (3b) and (3i)-(3e)
(3k) $4 \leq b_2 \leq 5$ from (3a) and (3f)
(3l) $11 \leq b_3 \leq 12$ from (3e) and (3h)
(3m) $15 \leq b_1 + b_2 \leq 13$ from (3c) and (3f)+(3i)-(3e)
(3n) $26 \leq b_1 + b_2 + b_3 \leq 24$ from (3c)+(3e) and (3f)+(3i)

These constraints defining the core are shown in Figure 4.7. It is obvious that an equilibrium which supports the efficient allocation does not exist. The problem can be seen most clearly by noting that the total cost of serving the three lanes must satisfy the constraints: $26 \leq b_1 + b_2 + b_3 \leq 24$. This indicates that while the efficient allocation is \$26, in order to maintain an equilibrium the total cost for the three lanes must also be below \$24. Thus, the core is infeasible and the efficient allocation will not be reached. Note that if Carrier C's reservation bid for package 6 (lanes 1 and 3) were increased by just \$2 to \$21, then the core would be the point

at the intersection of constraints (3f) and (3c). This point (8, 7, 11) would constitute an equilibrium supporting the efficient allocation.

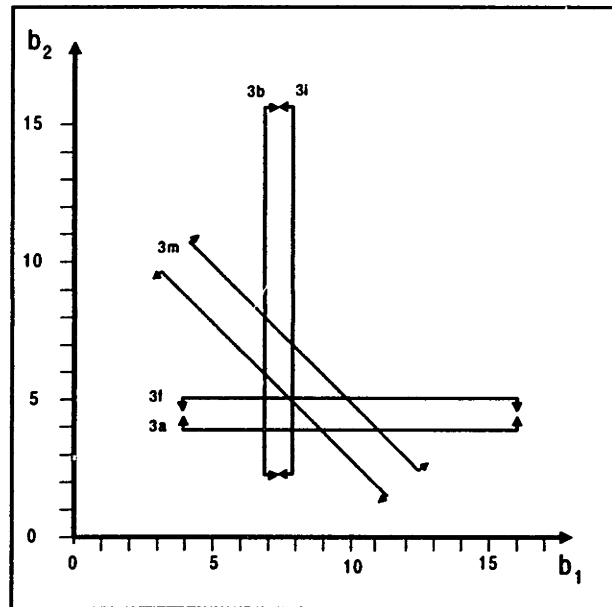


Figure 4.7. Infeasible core. Case 4: Strong Interdependencies.

The consequences of not achieving an equilibrium can be illustrated through a simple (simultaneous, multiple-round) auction as shown in Table 4.10 and Figure 4.8. The bidding starts out normally and each carrier makes bids that strictly improve their expected profitability. The problem occurs in Round 4 when Carrier B, being the leading bidder for lane 2, bids on lane 3 below the stand alone reservation value in order to pair them. However, it loses lane 2 to Carrier A's bid and is stuck holding lane 3 at a loss. Carriers A and B continue to trade for lane 2 for the remaining rounds - each time bidding only to improve their profitability. However, this leads to a final assignment where there is a loss for at least one carrier. In this case, both Carriers A and B lose money.

The problem for both Carriers A and B is that they are stuck with partial packages with each needing the same lane. Bykowsky, Cull, and Ledyard (1995) refer to this as mutually destructive bidding. Using simple bids forces carriers to bid on each lane separately and risk committing to portions of a tour with no guarantee of getting it all. This situation is similar to the Escalation or Entrapment Game⁵⁷ where two individuals bid on a single dollar bill with the only stipulation being that the runner up has to also pay their last bid. In most situations, the dollar bill is sold for much more than a dollar since it is in each individual's best interest to bid higher to cover their losses.

⁵⁷ For interesting discussion, see Raiffa (1982, 85-90).

| | Current Winning Bids | | | Current Profits | | | Incentives for Next Round of Bids |
|---------|----------------------|------|------|-----------------|----|---|---|
| | Ln 1 | Ln 2 | Ln 3 | A | B | C | |
| Round 1 | | | | 0 | 1 | 5 | |
| Bids | 12 | 12 | 12 | | | | A - bids on 1&2 - E(profit) = 7 B - bids on 3 - E(profit) = 5 C - no bid - cannot improve |
| Carrier | C | B | C | | | | |
| Round 2 | | | | 7 | 0 | 0 | |
| Bids | 11 | 11 | 11 | | | | A - no bid - cannot improve B - bids on 2 - E(profit) = 5 C - bids on 1&3 - E(profit) = 1 |
| Carrier | A | A | B | | | | |
| Round 3 | | | | 0 | 1 | 1 | |
| Bids | 10 | 10 | 10 | | | | A - bids on 1&2 - E(profit) = 2 B - bids on 3 - E(profit) = 3 C - no bid - cannot improve |
| Carrier | C | B | C | | | | |
| Round 4 | | | | 3 | -2 | 0 | |
| Bids | 9 | 9 | 9 | | | | A - no bid - cannot improve B - bids on 2 - E(profit) = 1 C - no bid - cannot improve |
| Carrier | A | A | B | | | | |
| Round 5 | | | | -2 | 1 | 0 | |
| Bids | 9 | 8 | 9 | | | | A - bids on 2 - E(profit) = 1 B - no bid - cannot improve C - no bid - cannot improve |
| Carrier | A | B | B | | | | |
| Round 6 | | | | 1 | -2 | 0 | |
| Bids | 9 | 7 | 9 | | | | A - no bid - cannot improve B - bids on 2 - E(profit) = -1 C - no bid - cannot improve |
| Carrier | A | A | B | | | | |
| Round 7 | | | | -2 | -1 | 0 | |
| Bids | 9 | 6 | 9 | | | | A - bids on 2 - E(profit) = -1 B - no bid - cannot improve C - no bid - cannot improve |
| Carrier | A | B | B | | | | |
| Round 8 | | | | -1 | -2 | 0 | |
| Bids | 9 | 5 | 9 | | | | A - no bid - cannot improve B - no bid - cannot improve C - no bid - cannot improve |
| Carrier | A | A | B | | | | |

Table 4.10. Sample auction. Case 4: Strong Interdependencies.

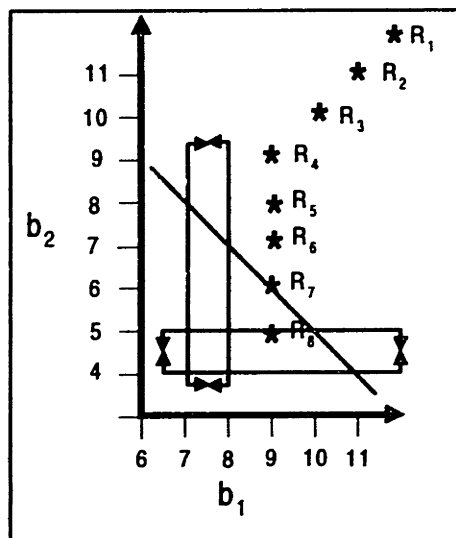


Figure 4.8. Bid trajectory. Case 4: Strong Interdependencies.

Of course, these are purely myopic bidding decisions. If a carrier is risk averse, they will not bid as aggressively, the efficient allocation will not be reached, and the total cost will be higher for the shipper. This means that some lanes might be won relatively easily if a carrier feels that they are in an escalation game situation. Because no other players know the others'

true reservation values, this is a very real possibility. So, when interdependencies exist and are sufficiently large, simple bids are not adequate to sustain the efficient allocation. The two major problems are (1) carriers must take unnecessary risks in achieving efficient allocations and (2) it is possible to have mutually destructive bidding situations.

Conditional Bids

The primary benefit of conditional bids is that they allow the carriers to achieve efficient aggregation without risk. Since carriers can bid on all of the lanes within an efficient aggregation as an all-or-nothing package, they can place all of the value of the interdependency in a single bid. The carriers do not have to divide and strategically assign the potential savings to the individual lanes within an efficient aggregation. Additionally, conditional bids guarantee that the efficient allocation is supported by an equilibrium. Because of this, conditional bids remove the possibility of mutually destructive bidding. This can be shown using the same example as before, but this time allowing conditional bids.

Example - Case 5: Strong Interdependencies - Conditional Bidding

Table 4.11 contains the same reservation values as in Case 4. The efficient allocation is again Carrier A wins lanes 1 and 2 and Carrier B wins lane 3 for a combined total value of \$26.

| Carrier | A | | | B | | | C | | |
|-------------|----|---|----|---|----|----|----|----|----|
| Bid Package | 1 | 2 | 4 | 2 | 3 | 5 | 1 | 3 | 6 |
| Res. Vals. | 11 | 8 | 15 | 9 | 11 | 16 | 12 | 12 | 19 |
| Lane 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| Lane 2 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| Lane 3 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 |

Table 4.11. Reservation values. Case 5: Strong Interdependencies using Conditional Bids.

Unlike the last case, however, we can now obtain a feasible set of constraints. Conditional bids are indicated by double subscripts so that bid r_{12} is a conditional bid for lanes 1 and 2 as an all-or-nothing package. The equilibrium conditions form the following set of constraints:

Carrier A:

- (4a) $b_1 - 11 \leq b_{12} - 15$ if not, A will bid on lane 1 alone $4 \leq b_{12} - b_1$
- (4b) $b_2 - 8 \leq b_{12} - 15$ if not, A will bid on lane 2 alone $7 \leq b_{12} - b_2$
- (4c) $0 \leq b_{12} - 15$ if not, A will not bid at all $15 \leq b_{12}$

Carrier B:

- (4d) $b_2 - 9 \leq b_3 - 11$ if not, B will bid on lane 2 alone $2 \leq b_3 - b_2$
- (4e) $0 \leq b_3 - 11$ if not, B will not bid at all $11 \leq b_3$
- (4f) $b_{23} - 16 \leq b_3 - 11$ if not, B will bid on lanes 2 and 3 $b_{23} - b_3 \leq 5$

Carrier C:

| | | | |
|------|----------------------|-------------------------------------|------------------|
| (4g) | $b_1 - 12 \leq 0$ | if not, C will bid on lane 1 alone | $b_1 \leq 12$ |
| (4h) | $b_3 - 12 \leq 0$ | if not, C will bid on lane 3 alone | $b_3 \leq 12$ |
| (4i) | $b_{13} - 19 \leq 0$ | if not, C will bid on lanes 1 and 3 | $b_{13} \leq 19$ |

which gives the following feasible core:

| | | |
|------|--------------------------------|------------------------------|
| (4j) | $15 \leq b_{12} \leq 21$ | from (4c) and (4e)-(4d)+(4g) |
| (4k) | $11 \leq b_3 \leq 12$ | from (4e) and (4h) |
| (4l) | $26 \leq b_{12} + b_3 \leq 25$ | from (4c)+(4e) and (4d)+(4i) |

The benefits of conditional bids come at a high price, however. Conditional bids make both the shipper's and the carriers' jobs more complicated. The shipper can no longer simply sort the bids by lane in order to find the lowest cost allocation. They need to solve a set partitioning problem, which is not trivial. The carriers also have a much more difficult time determining how to bid in subsequent rounds. Conditional bids are not as transparent as simple bids. For example, suppose that in some intermediate round Carrier A is low bidder for lanes 1 and 2 with a conditional bid of \$17 and Carrier C is low bidder for lane 3 with a simple bid of \$13. What should Carrier B bid for lane 2 in the next round? Can Carrier B bid low enough on lane 2 so that it wins both lanes 2 and 3? If so, how much is required? There is no clear solution here. The shipper could assist the carriers by providing additional information - but what information can the carriers use?

Because combinatorial auctions are so complicated, yet offer significant opportunities for achieving efficient aggregations, Section 4.4 is dedicated to discussing them in greater detail. At this stage we note only that simple bids are adequate and preferable for auctions where the interdependencies between lanes are sufficiently low. When the interdependencies are strong, it is better to switch to conditional bids in order to gain the benefits of more efficient allocations.

4.3.3 Number of Rounds: Single versus Multiple

The third auction design decision is whether to use a single round, multiple rounds, or some hybrid, such as a single round auction followed by an additional "negotiation" period or a multiple round auction which switches from conditional to simple bidding during the auction. Each of these are discussed in turn.

Single Rounds

A pure single round auction consists of the shippers sending out lane information, the carriers submitting back lane bids, and the shipper then selecting the low cost carriers based on the bids submitted. Essentially, they solve the model PB once. Even with conditional bids, the shipper will only need to solve the set partitioning problem once.

The advantages to single round auctions are primarily economic. Single round auctions are very inexpensive; both for the shipper and for the carriers. It does not require any extensive re-evaluation of bids by the carriers and the decision rule is straightforward for the shipper. No additional rules for the auction are required and the process is simple for carriers to understand. Additionally, single round auctions do not require that the shipper provide the carriers any rate information; not even for those lanes that they bid and lost on. This is important if rates are to be kept confidential.

Single round auctions have three major deficiencies, though. They do not (1) provide carriers with any market information - aside from the won/lost signal at the completion of the auction, (2) address the incentive compatibility problem, or (3) allow for any flexibility in carrier bidding strategy. Like the Dutch auction discussed in Section 4.2, the carriers are making their bids with no external information. This may be important if conditional bids are allowed since carriers are not currently in the practice of bidding on packages of lanes and may not have a good feel for the savings from these interdependencies. A single round auction will provide no external information that they can use to benchmark their estimations.

Similarly, with no market information there is no market pressure to encourage carriers to bid their true valuations. Instead, carriers will tend to submit bids based on their estimation of the relative strength of the other carriers in the auction - which will add strategic padding to the reservation values.

Finally, by assigning lanes based on a one time selection, the carriers have very little flexibility. As was discussed in Section 4.2.1, the reservation values are not fast and firm values - they are composites of numerous sources and include probabilistic estimates of traffic volumes and lane assignments, many of which that are out of the carrier's control and depend on the outcome of the bid! By assigning lanes based on a one time estimation, the carriers are not given the opportunity to adjust their reservation values. For example, a carrier might reconsider the acceptable margin on a lane if that lane completes a closed tour. Thus, if multiple rounds are used, the carrier can shift their priorities, not just their cost estimates, according to the developing situation. So, carriers are deprived of the opportunity to reconsider their reservation values during the auction.

Multiple Rounds

Using multiple rounds in an auction increases the likelihood of achieving an efficient allocation. The two primary benefits of multiple round auctions is that they give carriers an opportunity to adjust their bidding strategies and an incentive to bid based on their reservation values. The increased flexibility arises from carriers being able to adjust their bids in-between rounds to change priorities or re-evaluate their previous estimates and concerns. The incentive

compatibility problem is solved because it is always in the carrier's best interest to bid according to their true valuation. This does not mean, however, that the carrier should actually bid their true reservation value. Instead, by following their true reservation values, the carriers will end up with those lanes where they have the lowest cost structure, but the winning bid will be at the next highest carrier's bid - that is, the lowest losing bid rate.

To see this, consider the case where there are just two carriers, A and B, left bidding on a lane. Carrier A has a reservation value of \$1.00 and Carrier B has a reservation value of \$0.95 and the current bidding is at \$1.10. The carriers will alternate making bids, decreasing their bid by the minimum allowable increment (which we assume here to be infinitesimal), until it reaches the higher of the two reservation values. In this case, the bidding will stop at \$1.00 since Carrier A has no incentive to bid any lower.⁵⁸ Carrier B will win the lane at Carrier A's (the lowest losing bidder) reservation value. Thus, the winning bidder is awarded the lane at the level set by the valuation of the next highest carrier, not by its own valuation. Each carrier's dominant strategy, then is to always bid down to their true reservation value but not to go below it. There is no benefit to bidding below their true reservation value (or else there is a chance of winning the lane at a loss) or not bidding when the current bid is above the carrier's reservation value (or else the carrier will lose the lane and all potential profits).

Of course, this analysis is dependent on the minimum allowable increment being very small and the cost of an additional round being costless to the carriers. The general principle still applies, though, in that the dominant strategy is for a carrier to continue bidding until the price falls below their reservation value. This will lead to an efficient allocation.

The negative side of multiple round auctions is the high cost and the added complexity. Running an auction for more than one round is expensive and time consuming - especially for very large distribution networks. This is true for both the shipper and the carriers. Having to re-evaluate and resubmit bids requires additional effort by the carriers. This effort is not warranted if the value of the item or the expected benefit of a re-bidding is low.

Some of the time and dollar costs can be reduced by using recent information technological advances. For example, instead of having carrier representatives gather at a single location to make bids, the bidding could be made via modem. This has been used successfully in experiments at the California Institute of Technology.⁵⁹ The bidding can take place over several days at predetermined round increments providing adequate time for carriers to reconsider the latest round developments.

⁵⁸ Of course, if at this point, Carrier A reconsiders the value of attaining this lane and adjusts their reservation bid to, say, \$0.92, then the bidding would continue until the next highest reservation bid is reached.

⁵⁹ See, for example, Banks, Ledyard, and Porter (1989).

The added complexity of multiple round auctions comes from the management of the auction itself. The shipper needs to establish (1) the type of information to provide the carriers between rounds, (2) an activity rule which specifies how carriers remain in the auction, (3) stopping rules to know how to end the auction, and (4) withdrawal rules.

The type of information passed between rounds will influence the bidding in the next round. If no information is passed, then the auction reduces to a series of single bid auctions. If the carriers are provided with just the winning bids in each lane, assuming just simple bids for the moment, then they can adjust their bids accordingly. If the identity of the leading bidders in each lane is also provided, then the carriers can use this additional information as well. This could lead to collusion between carriers, however. If conditional bids are allowed, then the type of information to pass becomes much more difficult. The details are discussed in the next chapter.

Activity rules provide incentives to keep carriers that eventually want to bid on a lane, actively involved in the bidding process throughout the auction. The earlier and more frequent bids are made in the process of a multiple round auction, the better the information that is provided and the faster the auction will converge. If there is no minimum requirement for staying active, carriers have an incentive to wait for as long as possible until the market has settled down and values are emerging. The longer a carrier waits, the more stable the prices become and the likelihood of overestimating the value of the lane, thereby underbidding, decreases. By waiting, a carrier benefits from the other carriers' valuations while not providing any additional information itself. Activity rules can be in the form of a minimum number of bids submitted per round, percentage or absolute values of bid decreases, or several other forms.

Stopping rules specify when and how the auction should end. Stating ahead of time that an auction will go for exactly x rounds essentially invalidates the first $x-1$ rounds since all of the carriers know that the last round is the only one that matters. Thus it is reduced to a single round auction. Instead, the stopping rule needs to be based on the bidding itself. A common rule is to close a lane, or set of lanes, if the bid value does not change for, say, two rounds.

The simplest closing sequence is to close all markets simultaneously - when the last item meets whatever stopping criteria was set. Another method of sequencing the stopping is to force markets to close in a prescribed order. For example, the minor markets might not be allowed to close until the major markets finish. For the TL problem, one could also close by related sets where lanes that are closely related are closed together, i.e., do not close a lane without closing its backhaul. Generally, unless there is some over-riding reason, it is best to let the lanes, or perhaps sets of correlated lanes, close independently. Forcing all of the lanes to remain open until the very last is closed increases the likelihood of strategic misrepresentation

and opportunism. For example, a carrier could feign disinterest in a set of lanes, knowing that other lanes in the auction will keep the auction open for several more rounds. Then, it can suddenly reactivate a previously 'settled' portion of the auction.

Withdrawal rules specify when a carrier may recall a bid from consideration. Generally, this is not an advisable feature since a carrier could use this option to deflate the bids on certain lanes in order to drain the resources of a competitor for the purpose of acquiring a totally different set of lanes. Penalizing bid withdrawals discourages insincere bidding. Another problem is that a chain reaction can occur whereby withdrawing a bid forces the second highest bidder to allocate resources to a bid it had thought it had already lost. These resources might have been committed elsewhere and could cause that bidder to default elsewhere in the system and so on.⁶⁰

Hybrids

Overall, single round auctions are more economical and are faster than multiple round auctions, but they do not guarantee efficient allocation. Multiple round auctions are more expensive and have added complexities which are often quite daunting, and unwarranted in some situations. In between these two extremes are two hybrid forms.

Hybrid 1: Single-Round with post-Auction Negotiation

In this alternative, the shipper holds a simultaneous, single-round, sealed-bid, auction and after the lanes have been awarded, approaches some or all of the carriers to see what if any rate concessions can be made. This is quite common in practice. Some companies re-negotiate only with the winners of the auction with the hidden message that to remain the winning carrier, their rates need to be reduced by some percentage. Other shippers will approach carriers about lanes that they did not win to see if lower rates could be negotiated. Some shippers refuse to do any *ex post* discussion though to avoid "whip-sawing" the carriers.

While on the surface, this seems to be an effective method for getting lower rates, it actually exacerbates the incentive problem. If carriers know that the shipper is going to come back after the negotiation, they will tend to pad their bids to protect themselves from the forced concessions. Additionally, since a single round is used, market forces do not force the prices down. So, instead of providing a carrier with an incentive to bid their reservation value, this method gives carriers an incentive to adjust their bids based on their assessment of the competing carriers' strengths.

⁶⁰ This occurred in a major spectrum auction in New Zealand.

Re-negotiation also tends to lessen the goodwill between the shipper and the carriers. It is not advantageous for a long term relationship. Carriers will be less forthcoming to share cost information with a shipper that uses that information to squeeze their rates. While switching from this practice is advisable, it is very difficult since the shipper has lost credibility. Carriers need a credible commitment that if they actually bid according to their true reservation values, that they will not be punished in post-auction negotiations.⁶¹ It is very difficult to regain credibility on this front. One possible method would be for a shipper to employ a third party which has the existing level of credibility to conduct the bidding procedure. The shipper then is hiring the third party's credibility with carriers as well as their bidding skills. Another possibility is when a major restructuring is occurring and the shipper can credibly communicate a sea-change in policy and management of the entire transportation program. A good time to do this is after major mergers or acquisitions when the shared transportation systems are being developed.

The urge to re-negotiate is not the sole property of shippers, however. Losing carriers will often approach the shipper after lanes have been awarded and offer various price concessions in order to secure traffic lanes. This reduces the credibility of the carrier and adds to the shipper's perception that all carriers inflate their bids and can afford to reduce their bids after the auction. An auction should be designed to lower the carriers' regret in not winning certain lanes by providing opportunities within the auction to offer additional bids.

Despite all of these problems, re-negotiation does have one very beneficial property. It allows the shipper to specifically identify and target those lanes and carriers to work with. If other restrictions are important, such as a limit on the total number of carriers, then the shipper might be able to identify the opportunities for meeting these requirements better than the auction would.

Hybrid 2: Single Conditional Bid Round followed by Multiple Simple Bid Rounds

The idea of this hybrid form is to let the carriers determine in the first round (or rounds) which combinations of lanes makes the most economic sense. Once this has been determined, the current partitions are fixed, and the auction is continued treating the efficient aggregations as single lanes.

The advantages to this system are that it speeds up the auction and decreases the complexity of the later rounds. Simple bids are easier to process and interpret than conditional bids so that both the shipper and the carriers would require less time between rounds. The

⁶¹ Kreps (1990, 677-9 and 696-7) discusses the same credibility problem in the situation of insurance companies selecting high and low risk policies. Without a credible commitment against re-negotiation, there is no incentive to be truthful.

initial conditional-bid rounds provide some level of flexibility for carriers to determine which lanes are best but not as much as a full conditional bid auction.

Another way to incorporate this same effect, is to simply aggregate or consolidate any conditional bid that does not change partitions within, say, two rounds. That is, if the same three lanes are maintained in an efficient aggregation, even if between rounds different carriers bid on it, then it becomes fixed and those lanes are now treated as a single entity and cannot be pulled into a competing package. This would add slightly more complexity, but is a compromise between closing all conditional bids at the same time and keeping them all open. It lets the market determine which ones to close and when.

4.3.4 Pricing Mechanism: First versus Second Price

The pricing mechanism determines the rate at which the winning carrier will charge the shipper. In all of our discussions, we have assumed that the carrier will charge the price that they bid. This is referred to as the first-price mechanism. Another option is to let the bids determine which carrier wins a lane, but the rate which they win it at is determined by the runner up bid. This is called the second-price mechanism. In order to be able to distinguish between these mechanisms, assume for this section that instead of one carrier being assigned to a lane, the shipper wants x carriers assigned. For simple bids this means that the x lowest bids are accepted and for conditional bids an extra constraint is added to ensure a carrier doesn't win a lane more than once. This section discusses first-price, second-price, and a hybrid type pricing mechanism in turn.

First-Price Mechanism

For each lane in first-price auctions, the x carriers with the lowest bids win that lane at the rate they submitted. This means that when tendering the loads, different carriers will have different rates for hauling over the same lane. This auction is sometimes referred to as a discriminatory or multi-price auction since the same lane is being served by different carriers at different prices.

This is the most common pricing mechanism for TL bidding in practice. It is easy to understand and winning carriers are paid at the rate that they actually bid. In tendering, it can sometimes be a problem if the local dispatcher is measured according to transportation cost since the tendency would be to tender loads only to the low cost carrier - which could violate any volume guarantees. This adds some minor complexity to the local tender decision. Still, this is a simple and commonsensical approach. It does not provide any incentives to carriers to reveal their true valuations, however.

Second-Price Mechanism

For each lane in second-price auctions, the x carriers with the lowest bids still win that lane but the rate charged is the price of the $x+1^{\text{th}}$ carrier's bid, that is, the lowest losing bid. This type of auction is sometimes referred to as uniform, or nondiscriminatory pricing since all carriers will receive the same rate for serving that lane. This is also referred to as a auction after the economist who first classified and analyzed these schemes, (1961).

While not widely used in practice, second-price auctions provide an incentive for carriers to bid their true reservation values. Like the multiple round auction, the dominant strategy for carriers in second-price auctions is to follow their true reservation values. The basic idea is that a carrier can never do better, and will often do worse, by bidding above or below its true reservation value. In order to see this, suppose that in a auction there are just two carriers, A and B, bidding for one lane. They have submitted bids b_A and b_B and have reservation values of r_A and r_B . If Carrier A bids its reservation rate and wins, then the payoff is the difference between Carrier B's bid (the lowest losing bid) and its reservation value, $b_B - r_A$. If Carrier A bids below its reservation value and wins, then the outcome will be the same (payoff of $b_B - r_A$). However, if Carrier A's reservation value is actually greater than Carrier B's, $r_B < r_A$, then Carrier A will win the lane, but the payoff of $b_B - r_A$ will be *negative* since $b_B = r_B < r_A$. Finally, if Carrier A bids above its reservation value and wins, the outcome will be the same (payoff of $b_B - r_A$) unless the bid falls below Carrier B's reservation value in which case Carrier A loses the auction altogether and receives no payoff. Thus, a bidder can do no better and can possibly do much worse by not bidding their true reservation value.

The result is identical to the multiple round auction. The winning carrier is assigned a rate determined by the lowest losing carrier's bid. While there is no theoretical restriction to multiple-round, second-price auctions, it is overly complicated and does not make the auction more efficient. As with multiple round auctions, though, second-price auctions do not transfer to conditional bids very easily. It is not readily evident what the "next-best" bid for a lane contained in a package would be. This problem is discussed in the next chapter.

Hybrid Pricing Mechanism

A hybrid type of pricing mechanism is one where the shipper determines a target price for service on a lane and then identifies carriers that are willing to accept that price. USX used this strategy in the mid-1980's to set the freight rates for their traffic lanes. The rate could be determined from previous contracts or from some selected data source. The advantage is that the local transportation managers have a single rate for each lane thus lowering the complexity of their decisions. Additionally, target-pricing is very easy to manage and is often used for

quick bids when a small number of new lanes or territories are added to a distribution network. The disadvantage is that the shipper may be missing out on opportunities by not using the market to determine an equilibrium rate. This form of mechanism is not recommended.

4.4 Combinatorial Auctions

This section discusses combinatorial auctions. A combinatorial auction is any auction which allows the use of conditional bids. There are four general benefits to using conditional bids. First, conditional bids can ensure that an efficient allocation is reached when simple bids cannot. Second, conditional bids lower the threat of exposure faced by carriers in achieving efficient combinations of lanes. Third, conditional bids allow carriers to place the synergy value arising from a set of lanes onto that set as a whole, instead of allocating it across the individual lanes. Finally, conditional bids provide both the shippers and carriers more control over the allocation process.

Along with these benefits come four general drawbacks. First, the implementation of a combinatorial auction is more complicated for both the carriers and the shipper. The increased complexity is mainly caused by the very large number of potential packages, and the need for the carriers to be able to price out each of these combinations. Second, the underlying problem that the shipper must solve is the set partitioning problem which is NP Complete. In most realistic auctions, several additional side constraints need to be added which increases the solution difficulty. Third, there is a communication problem involved in combinatorial auctions if multiple rounds are used. The shipper needs to pass some sort of information to the carriers in between rounds, but there is no consensus on what the correct information is. Finally, the whole area of combinatorial auctions is very new, both in the literature and in practice.

The section introduces standard terminology and explores four basic concepts of combinatorial auctions: the bid packet structure, the contention level of an auction, the bid types, and the threshold problem. Finally, four examples of where combinatorial auctions were considered, tested, or actually applied are presented and compared.

4.4.1 Combinatorial Auction Concepts

This section introduces the basic characteristics of a combinatorial auction. Before describing these concepts, a standard terminology is first introduced. Throughout this chapter, the following notation and terminology will be used to describe the facets of combinatorial auctions. This necessitates that the terms used here will not match all of the current literature.

| | |
|----------------|---|
| <i>Lane</i> | The smallest unit being auctioned off, |
| <i>Package</i> | A set of lanes which are combined into an all or nothing bid. |

- Conditional Bid** A bid which covers a package where the bid price is conditioned on receiving all of the lanes within the package.
- Singleton:** A conditional bid for a package which only contains a single lane.
- Partition** The current set of packages which, if the auction was stopped, would be assigned to the different bidders.
- Standing Bid** The current winning bid for a package. During the auction, every package has a standing bid regardless of whether it is included in the partition.
- Active Bid** A standing bid which is currently in the partition. Standing bids for packages not in the partition are inactive.
- Assignment** The set of bidders that submitted the current standing bids for those packages that are in the partition.
- Allocation** The current partition, assignment, and corresponding bids.
- Synergy Value** The savings accrued by combining certain lanes which exploit the underlying interdependency between the lanes.

The following notation will be used:

- i : subscript for lanes to be auctioned,
 j : subscript for packages, each consisting of specific lanes,
 k : subscript for bidders,
 p_j : set of items contained in package j ,
 r_{jk} : reservation value of package j by bidder k ,
 b_{jk} : bid price for package j from bidder k ,
 B_j : standing bid for package j , thus, $B_j = \{\min b_{jk}; \forall k\}$
 P : partition of packages that are in current allocation,
 P_k : partition of packages awarded to bidder k , thus $\cup_k P_k = P$,
 $S(P)$: assignment of bidders to packages in the current partition, and
 $A(P, S, B)$: allocation of packages to bidders at specified bids.

These terms are illustrated by an example. Suppose during an auction, the current bidding for the three lane network in Figure 4.9 is as shown in Table 4.12. The table also shows which lanes are included in which packages. For example, Carrier A's bid for package 4 (lanes 1 and 2) is $b_{4a} = \$15$ while Carrier B's bid for package 2 (lane 2 as a singleton) is $b_{2b} = \$10$.

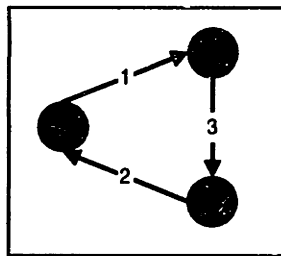


Figure 4.9. Sample 3 node network.

| Bid Package | 1 | 2 | 3 | 4 | 5 |
|-------------|----|----|----|----|----|
| Carrier A | 11 | 13 | 14 | 15 | 22 |
| Carrier B | 15 | 10 | 13 | 14 | 18 |
| Lane 1 | 1 | 0 | 0 | 1 | 0 |
| Lane 2 | 0 | 1 | 0 | 1 | 1 |
| Lane 3 | 0 | 0 | 1 | 0 | 1 |

Table 4.12. Current carrier bids for 3 node network - combinatorial auctions.

Three items are to be auctioned: lanes 1, 2, and 3. These three lanes have been aggregated into five packages:

$$p_1 = \text{Lane 1} \quad p_2 = \text{Lane 2} \quad p_3 = \text{Lane 3} \quad p_4 = \text{Lanes 1 and 2} \quad p_5 = \text{Lanes 2 and 3.}$$

The standing bids for each package are in bold in Table 4.12, they are:

$$B_1 = \$11 \quad B_2 = \$10 \quad B_3 = \$13 \quad B_4 = \$14 \quad B_5 = \$18$$

The partition consists of those packages which form the lowest cost combination which covers all of the lanes exactly once. The partition, $P = \{p_i, p_j\}$. The partition for each of the carriers is: $P_a = \{p_i\}$ and $P_b = \{p_j\}$. The assignment is, $S(P) = \{A, B\}$. While the current allocation is simply $[P, A(P), B(P)]$ or:

$$A = \begin{bmatrix} 1 & 5 \\ A & B \\ 11 & 18 \end{bmatrix}.$$

The remainder of this section addresses four defining characteristics of combinatorial auctions: the package structure, the level of contention, the types of bids, and the threshold problem.

Bid Package Structure

The bid package structure refers to how the different items are grouped into packages. While the form of the package should be dictated by the prevailing interdependencies, it can differ on many dimensions. The structure, for example, could be such that bid packages only contain two items or conversely, that they all must contain at least n items. For the TL bidding problem, the interdependencies rise from the matching of inbound and outbound traffic lanes so the bid packages will generally, but not always, consist of adjacent lanes in a tour. The packages which represent closed tours, for example, will most likely not contain more than four or five items (lanes) since it becomes unmanageable to control in real-time tendering. Methods of identifying various forms of lane packages are discussed in Chapter 5. Other applications have very different interdependencies which result in very different structures of packages.

The structure of the bids is important because the underlying problem of assigning lanes to carriers is the set partitioning problem which is NP Complete. The structure of the bids can provide some insight into developing a faster algorithm for solving the problem. As a trivial example, if the packages are all singletons, where $|p_i|=1$, then a simple sorting of bids for each item is sufficient to solve the assignment problem. The structure of the packages, not just the number of packages, influences the complexity of the auction. Thus, restricting the number of packages that a carrier can bid on generally will not make the problem easier to solve.

Rothkopf, Pekec, and Harstad (1995), referred to as RPH, identified three types of structures which can significantly affect the solvability of the combinatorial auction. These are nested structures, cardinality-based structures, and geometry-based structures. Each is briefly discussed in turn along with its applicability for TL bidding.

A nested structure occurs if, for every possible pair of packages that have one or more items in common, one of the packages is a subset of the other. This is shown graphically in Figure 4.10 for an auction of three lanes: (1, 2, 3). Figure 4.10a shows the situation where five packages are available {1; 2; 3; 1-2; 1-2-3} along with the current standing bids for each package. This is a nested structure. If the bid packages can be formed into a tree, then they have a nested structure. If a sixth package {2-3} is included, however, the structure is ruined, as shown in 4.10b.

To see why this is important, consider the problem of finding the current partition for case (a). RPH suggest a simple $O(|P|)$ algorithm which takes advantage of the tree structure. Starting at the leaf nodes (single item packages in this case), the algorithm compares the standing bid for each parent node to the sum of the standing bids for all of its children. For example, since, $B(1-2) = \$18 < B(1) + B(2) = \20 , package 1-2 is included in the temporary partition. Next, comparing $B(1-2-3) = \$28 > B(1-2) + B(3) = \27 , we see that packages 1-2 and 3 are included in the partition while packages 1, 2, and 1-2-3 are not. This greedy algorithm does not hold for case (b), since a package may be the child of multiple packages. It is no longer sufficient to find which packages dominate. Starting at the singleton, or leaf node, bid it is unclear how package 2 should be handled.

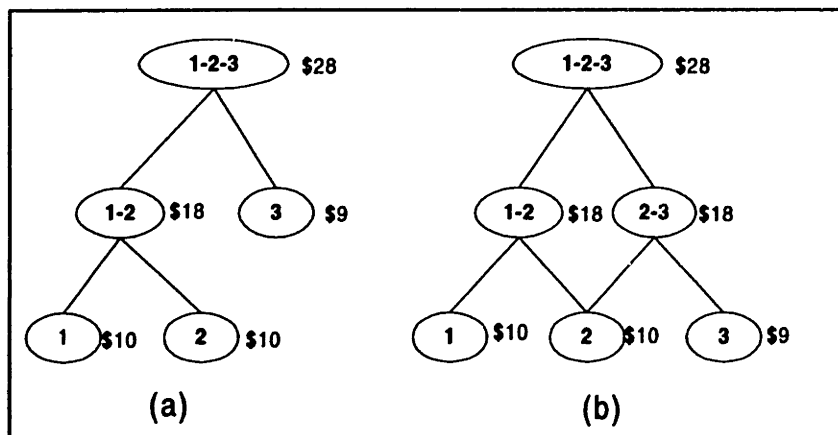


Figure 4.10 Example of nested and non-nested bid structures.

Nested bid structures can occur in TL bidding situations whenever the network has very few tours or if they are relatively isolated. If the competition for a tour of lanes are only the stand alone lanes, as opposed to other overlapping tours, then RPH's simple algorithm can be employed quite easily. An additional property of any nested structure could be exploited by

noting that the constraint matrix possesses the “consecutive one’s” property which, with minor manipulation, becomes totally unimodular. This implies that solving the modified LP of this problem would result in an integer solution.

The other two structures identified by RPH have less relevance to TL bidding. The cardinality-based structure applies if the number of items allowed within a package is two or less, $|p_i| \leq 2$. This allows for solving the problem as a maximum weight matching algorithm. Finally, geometry based structures occur if the interdependencies are related to proximity. While this appears to be valuable for TL bidding, the algorithmic advantages are only achieved when the proximity is in a single dimension, such as towns in a corridor. However, if this is increased to two or more dimensions, the algorithmic efficiencies disappear.

An additional decision concerning the package structure, is who can designate the bid packages. If this is restricted to the auctioneer, some level of interdependencies which are not visible or thought to be relevant will be missed. However, unrestricted combinatorial auctions could theoretically generate

$$\sum_{i=1}^N \binom{N}{i}$$

potential bid packages where N is the total number of items in the auction. In practice, for the TL bidding problem, this is not a realistic concern. As shown in Chapter 6, conditional bids were found to be worthwhile for approximately 10% of the lanes within the TL network.⁶² Since the bid packages reflect synergies due to matching complementary lanes, the opportunities for these packages are more limited. Because tour proliferation is not a large concern, allowing carriers to form their own bid packages is recommended. Carriers have access to many more opportunities to achieve synergies to which individual shippers do not have access.

Level of Contention

The contention level of an auction is a measure of the difficulty of the allocation problem. This has been recognized informally by several researchers. Rassenti, Smith, and Bulfin (1982) referred to this as “combinatorial complexity” and identified two influencing factors:

- the amount of package repetition between bidders and
- the amount of item repetition between a bidder’s packages.

Bykowsky, Cull, and Ledyard (1995) called this the “fitting complexity” and added two more factors:

- the amount of packages which act as substitutes and
- the size of the package interdependencies.

⁶² Of course, the traffic on these few lanes represented the majority of the traffic in the network.

Essentially, the 'difficulty' of an auction increases with the amount of package repetition between bidders, increases with the amount of item repetition between each bidder's own package bids, and decreases with the number of substitute packages available. The influence of the size of the package interdependencies is more relative. If the differences between the interdependencies of competing packages is very close, then the complexity is thought to increase. On the other hand, if the synergy value for a package for one bidder is much higher than the others, the complexity should be lowered.

Olson and Porter (1994) formalized this concept by defining the contention level of an auction for auctions where each bidder receives no more than one item and each item is awarded to no more than one bidder. This auction is formulated as:

$$\begin{aligned} \text{Max}_x W &= \sum_{i \in N} b_{ik} x_{ik} \\ \text{Subject to:} \\ (1) \quad \sum_{i \in N} x_{ik} &\leq 1 && \forall k \in K && \text{[OPP]} \\ (2) \quad \sum_{k \in K} x_{ik} &\leq 1 && \forall i \in N \\ (3) \quad x_{ik} &\geq 0 && \forall k \in K, i \in N \end{aligned}$$

where:

Indices:

i: Items to be auctioned, $i = 1, \dots, N$,
k: Bidders, $k = 1, \dots, K$,

Decision Variable:

x_{ik} : = 1 if bidder k wins item i , =0 otherwise,

Data:

b_{ik} : Bid for item i submitted by bidder k , and

Objective Function Value:

W : Total Welfare Value.

The problem [OPP] minimizes the total welfare value, W , defined as the sum of the winning bids for each item. Constraints (1) ensure that each bidder wins only one item while (2) make sure each item is awarded to only one bidder. Only simple bids were used in this example. Koopmans and Beckman (1957) showed that the clearing prices (dual variables, μ_k and π_i) that support an efficient allocation will always exist for this problem. Solving [OPP], then, always produces an integer solution. The dual for constraint (1), μ_k , represents bidder k 's surplus or rent. The dual for constraint (2), π_i , is the clearing price for item i . For all accepted bids, the sum of these duals will be equal to the item bid price. The metric for contention level is defined as:

$$C = \frac{\sum_i p_i}{W^*}$$

where:

- C: Contention level of auction,
 p_i : Vickrey price or minimum dual price for item i , and
 W^* : Total welfare of outcome efficient auction.

The Vickrey prices, p_i , are found by solving the 'pseudo-dual' problem:

$$\begin{aligned} & \text{Min}_p \sum_{i \in N} p_i \\ & \text{Subject to:} \\ & (1) \quad \mu_k + p_i \geq b_{ik} \quad \forall k \in K, i \in N \quad [\text{OPD}] \\ & (2) \quad \sum_{i \in N} p_i + \sum_{k \in K} \mu_k = W \\ & (3) \quad \mu_k, p_i \geq 0 \quad \forall k \in K, i \in N \end{aligned}$$

where:

Indices:

- i : Items to be auctioned, $i = 1, \dots, N$,
 k : Bidders, $k = 1, \dots, K$,

Decision Variable:

- p_i : Vickrey price for item i ,

Data:

- μ_k : Dual variable from carrier constraint in [OPP],
 b_{ik} : Bid for item i submitted by bidder k , and

Objective Function Value:

- W : Total Welfare Value.

The Vickrey prices, p_i , are the minimum value dual variables which still support the efficient allocation.⁶³ The bidder surplus duals, μ_k , are taken from the original problem. Larger μ_k values indicate a wider gap between the bids for the items. This implies less competition between the carriers. If $b_{ik} = p_i$ at the optimal solution, then these prices are extracting all of the rent from the market and the competition for that item is very intense.

Unfortunately, most combinatorial bidding problems do not have the same convenient totally unimodular constraint structure as [OPP]. Thus, the LP solution will rarely yield duals which offer any insight. Ledyard, Noussair, and Porter (1995) applied the same measure of contention to a combinatorial auction, but their problem, as it turned out, was totally unimodular, by design. Another drawback of this measure is that one needs to solve the

⁶³ Note that these formulations assume that the auctions allow infinitesimal bidding increments.

problem in order to tell how difficult the problem will be! Thus, this formal definition of contention is mainly used to classify and compare different controlled auction experiments.

In practice, knowing the contention level is most helpful prior to having to solve the problem. Three simple metrics used for comparisons in this chapter are:

- the median number of packages an item belongs to,
- the percentage of the packages which are singletons, and
- the total number of items divided by the total number of packages being offered.

The smaller the median number of packages per item, the lower the contention. The higher the percentage of singletons, the lower the contention. Finally, the contention increases as the number of packages greatly exceeds the number of items. These three measures provide a quick comparison across different environments and do not require solving any linear programs.

Bid Types

Unlike simple auctions where each bid influences only one item, a bid submitted in a combinatorial auction can affect the status of many items. In order to better understand this effect, Brewer and Plott (1995) identified five types of bids.⁶⁴ These bid types are only relevant, however, if the auction has multiple rounds. Recalling that,

| | |
|------------|---|
| p_j : | set of items contained in package j , |
| b_{jk} : | bid for package j from bidder k , |
| r_{jk} : | reservation value for package j for bidder k , |
| B_j : | standing bid for package j , thus, $B_j = \{\min b_{jk}; \forall k\}$ |
| P : | partition of packages that are in current allocation, |
| P_k : | partition of packages awarded to bidder k , thus $\cup_k P_k = P$, |

the five types of bids are:

Pivotal Bids: A bid, b_{jk} , that improves the bidder's profit ($B_j > b_{jk} > r_{jk}$) and is also low enough to change the current allocation,

Strong Neutral Bids: A bid, b_{jk} , that could potentially improve the bidder's profit ($B_j > b_{jk} > r_{jk}$) but does not do so immediately because it is not low enough to change the current allocation, $B(P)_{p \in P} > B(P)_{p \notin P}$

Dominated Bid: A bid, b_{jk} , that would lower the bidder's profitability ($B_j > r_{jk} > b_{jk}$) if the auction stopped immediately after the bid was submitted. The bid is low enough to change the current allocation.

Dominated Neutral Bids: A bid, b_{jk} , that could potentially lower the bidder's profitability ($B_j > r_{jk} > b_{jk}$) but is not low enough to change the current allocation.

Null Bid: A bid that is higher than the current standing bid for a package, ($b_{jk} > B_j$).

⁶⁴ While Brewer and Plott (1995) analyzed a "high-bid-wins" auction, these bid types are defined in terms of "low-bid wins" auctions, such as in TL bidding.

Two comments are in order. First, the dominated bids are not relevant to the TL bidding problem because they are not expected to happen in practice. In experiments, Brewer and Plott (1995) noted that bidders did not submit dominated bids at all. Second, because pivotal bids can change the allocation in two very different ways, we further divided them into major and minor pivots. A *minor pivotal bid*, is one which only changes the assignment but not the partition. That is, some of the packages already in the partition simply change hands between bidders. A *major pivotal bid* is where the partition changes as well as the assignment. This necessarily involves the shifting of some previously strong neutral bids into the partition. The flowchart in Figure 4.11 illustrates the differences between the different types of bids.

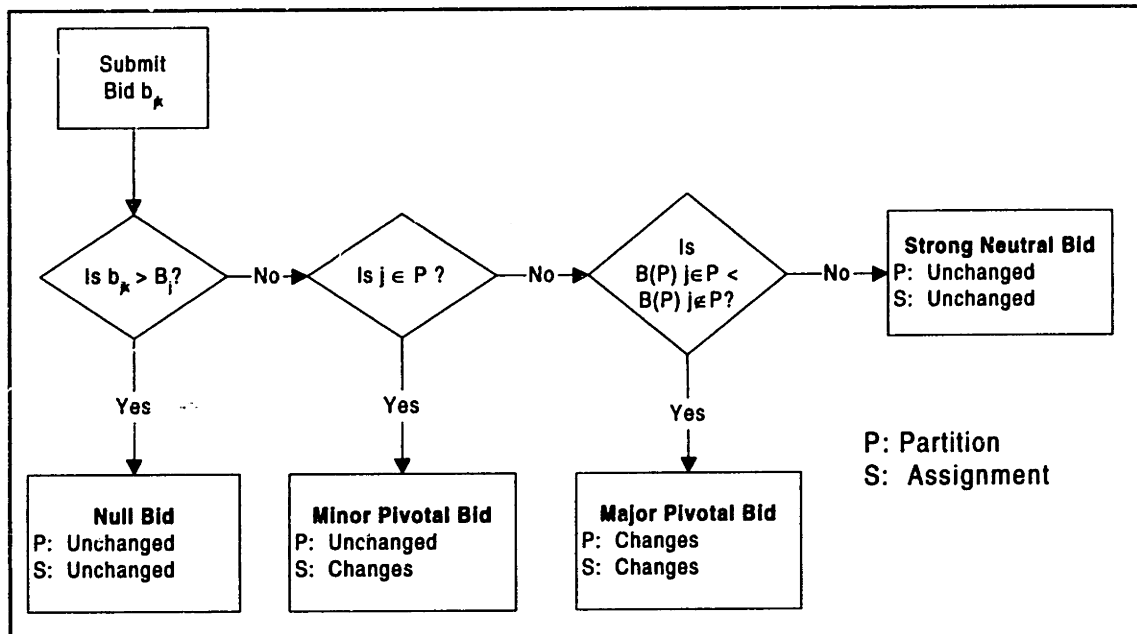


Figure 4.11. Types of combinatorial bids for a low-bid-wins auction.

The different types of bids can be illustrated with an example. In an auction of three lanes, three carriers, and six packages:

- | | |
|----------------|-----------------------|
| $p_1 =$ Lane 1 | $p_4 =$ Lanes 1 and 2 |
| $p_2 =$ Lane 2 | $p_5 =$ Lanes 2 and 3 |
| $p_3 =$ Lane 3 | $p_6 =$ Lanes 1 and 3 |

suppose the standing bids are as shown in Table 4.13.

| Package | 1 | 2 | 3 | 4 | 5 | 6 |
|--------------|----|----|----|----|----|----|
| Standing Bid | 13 | 12 | 14 | 26 | 27 | 24 |
| Carrier | A | B | C | A | B | C |

Table 4.13. Standing bids for a combinatorial auction.

The partition is $P = (p_2, p_6)$, the assignment is $S = (B, C)$, and the total bids are $B = \$36$. Using the notation (Carrier; Package Number; Bid Price), then we can classify the following bids.

Null Bids - having no effect on the current allocation or the standing bids.

(C, 1, \$15) - Carrier C's bid is above the standing bid of \$13 from Carrier A.
(A, 2, \$16) - Carrier A's bid is above the standing bid of \$12 from Carrier B.

Strong Neutral Bids - having no effect on the current allocation, but changing the standing bids.
(A, 3, \$13) - Carrier A's bid is below the standing bid for lane 3, but $B(p_1, p_2, p_3) = \$38 > \36 .
(B, 1, \$12) - Carrier B's bid is below the standing bid for lane 1, but $B(p_1, p_2, p_3) = \$38 > \36 .

Minor Pivotal Bid - changing the assignment of carriers to packages, but not the partition.
(A, 2, \$11) - Carrier A becomes standing bid for lane 2 replacing Carrier B.
(B, 6, \$22) - Carrier B becomes standing bid for package 6 (lanes 1 & 3) replacing Carrier C.

Major Pivotal Bid - changing both the assignment of carriers to packages and the partition itself.
(B, 5, \$22) - Carrier A submits low bid for conditional bid for lanes 1 and 2. The new partition becomes $P = \{p_1, p_5\}$ while the assignment is $S = \{A, B\}$.

Both of the previous packages in the partition, p_2 and p_6 , were removed while the standing bid for lane 1 (\$13 from Carrier A) is suddenly reactivated! This can create problems for a bidder if a bid which was once considered to be dead suddenly becomes valid again.

Strong neutral bids should be encouraged since their value often will determine the level or threshold of a bid needed to change the partition. The auction needs to provide some kind of mechanism to transmit the standing bids to other bidders. *Minor pivotal bids* are identical to simple bids since they only affect a single package at a time. *Major pivotal bids*, on the other hand, are the cause of most of the problems associated with combinatorial bidding.

Threshold Problem

The threshold problem occurs when a bidder is interested in bidding on a package that is currently not in the partition and is only relevant for multiple round auctions. In order for this bid to be successful, it must past some threshold level which will push out the other packages which have overlapping items. The problem is that the amount that a bidder needs to bid is not obvious - even if for a single item.

An example will illustrate the problem. Suppose that there are two lanes to be bid out with the reservation value of three of the carriers being:

- Carrier A reservation value lane 1 = \$0.90
- Carrier B reservation value lane 2 = \$0.90
- Carrier C reservation value Lanes 1 & 2 = \$1.85

Suppose that at some point in the auction, Carrier C is leading with a bid of \$1.85 for the package of lanes 1 and 2 while the current standing bids are \$1.00 each for lanes 1 and 2 as singletons by some other carriers, say, D and E. Even though a more efficient allocation is to assign Carrier A to lane 1 and Carrier B to lane 2, neither Carrier A nor B, acting unilaterally, can win their respective lanes. If Carrier A bids its true reservation value of \$0.90, it will still not capture lane 1. Carrier A needs Carrier B in order to reach the threshold to overcome Carrier C's conditional bid. Thus, each carrier has a vested interest in the other's bid.

While conditional bids incorporate interdependencies between lanes awarded to the same bidder, they cannot be used across different bidders. The threshold problem addresses how to coordinate the efforts of independent bidders with a shared interest in overcoming another party's bid. The key to mitigating the threshold problem, then, is to encourage some sort of communication between carriers. This can be done by allowing direct coalitions, or by the auctioneer issuing some sort of signals to the bidders. Of course, this raises additional problems of collusion between bidders.

4.4.2 Applications and Experiments

The literature concerning combinatorial bidding, and multiple item auctions in general, is extremely scarce. The complexities of auction analysis limited the theoretical attention to more tractable, highly stylized, games. Recently, however, applications suited for combinatorial auctions have arisen, which are now driving the development of new auction theory. Because the design of an auction is heavily influenced by the characteristics of the bidders, the auctioneer, and the items being auctioned, the details of each use of combinatorial bidding differs. It is important to understand how the environment influences the design factors for these auctions.

This section presents four applications where combinatorial auctions were either implemented, tested, or seriously considered: airport landing slot allocations, resource scheduling on space satellites, railroad routing, and radio spectrum license assignment.

Airport Take-Off/Landing Slot Allocations

Rassenti, Smith, and Bulfin (1982) designed and tested a combinatorial, single-round, sealed-bid, second-price auction for allocating airport runway slots to airlines. Due to congestion at many major airports, these slots are considered scarce resources and have to be divided among the different airlines in some fashion. Often times, this is allocated by committee. The slots exhibit interdependencies due to the need for both a take-off and a landing slot within a reasonable time period. The synergy value is very high for the packages since the value of a landing slot without a corresponding takeoff slot is zero. Additionally, a "budget" constraint was enforced by some of the airlines limiting the number of certain combinations of slots. Thus, the slots behaved as both complements and substitutes.

A single round auction was used due to the frequency of the need to allocate slots. In order to induce the airlines to reveal their true valuations for the different slots, a uniform or second price pricing mechanism was used. It was also thought that while an after-auction market where airlines could trade slots would exist, it would not be very efficient to rely too heavily on it.

Airlines submitted conditional bids for packets of landing/take-off slots. Simple bids for a single slot were not allowed. The primary auction allocated slots by solving the problem:

$$Max_x = \sum_{j=1}^J \sum_{k=1}^K b_{jk} x_{jk}$$

Subject to:

$$(1) \quad \sum_{j=1}^J \sum_{k=1}^K a_{ijk} x_{jk} \leq s_i \quad \forall i = 1, \dots, I \quad [P]$$

$$(2) \quad \sum_{j=1}^J d_{jk} x_{jk} \leq e_j \quad \forall k = 1, \dots, K$$

$$(3) \quad x_{jk} = \{0,1\} \quad \forall j = 1, \dots, J; k = 1, \dots, K$$

where:

Indices:

i : Slots to be auctioned,
 j : Packages, each consisting of specific slots,
 k : Bidding airlines,

Decision Variable:

x_{jk} : = 1 if bidder k is allocated package j , =0 otherwise

Data:

b_{jk} : bid for package j from airline k ,
 a_{ijk} : coefficient; = 1 if slot i is in package j for airline k ; =0 otherwise
 d_{jk} : coefficient; = 1 if slot i is in package j for airline k ; =0 otherwise
 s_i : supply of landing slot i ,
 e_k : some integer ≥ 1 ,

The objective function maximizes the sum of the bid prices submitted by the airlines for the various packages. Constraints (1) ensure that the slots are not over allocated; (2) enforce any logical requirements on the carriers - typically a budget constraint on the number of packages assigned; and (3) enforce integrality and non-negativity.

Problem (P) finds the highest valued allocation of airlines to slots. The price paid for these slots is determined by solving the following two "pseudo-dual" programs, D_A and D_R :

$$Min_y = \sum_{jk \in R} y_{jk}$$

Subject to:

$$(1) \quad \sum_i w_i a_{ijk} \leq b_{jk} \quad \forall jk \in A \quad [D_R]$$

$$(2) \quad y_{jk} + \sum_i w_i a_{ijk} \geq b_{jk} \quad \forall jk \in R$$

$$(3) \quad y_{jk} \geq 0 \quad w_i \geq 0$$

$$Min_y = \sum_{jk \in A} y_{jk}$$

Subject to:

$$(1) \quad \sum_i v_i a_{ijk} \geq b_{jk} \quad \forall jk \in R \quad [D_A]$$

$$(2) \quad -y_{jk} + \sum_i v_i a_{ijk} \leq b_{jk} \quad \forall jk \in A$$

$$(3) \quad y_{jk} \geq 0 \quad v_i \geq 0$$

where:

Indices:

- i: Slots to be auctioned,
- j: Packages, each consisting of specific slots,
- k: Bidding airlines,

Decision Variable:

- y_k : surplus value of bid by airline k for package j,

Data:

- A: set of all accepted conditional bids,
- R: set of all rejected conditional bids,
- w_i : lower bound for price on slot i,
- v_i : upper bound for price on slot i, and
- b_k : bid for package j from airline k.

The solutions to problems D_A and D_R define the bounds for the slot prices, that is the package bids can take one of three cases:

- Case 1. if $b_k > \sum_{i \in j} v_i$ then $x_k = 1$ Bid always accepted at price $\sum_{i \in j} v_i$
- Case 2. if $b_k < \sum_{i \in j} w_i$ then $x_k = 0$ Bid is always rejected
- Case 3. if $\sum_{i \in j} w_i \leq b_k \leq \sum_{i \in j} v_i$ then $x_k = 0$ or 1 Cannot guarantee either.

The final price charged to all airlines winning a package is the sum of the lower bound prices for all of the slots within that package. This is the lowest possible price for a package which could have possibly been allocated. Thus, every airline pays the same price for identical slots.

Eight experiments consisting of six successive periods were run to compare this combinatorial auction against a simple bid, second-price auction. Each experiment used the same slot values for each of the six periods, but the values differed between participants. The participants were provided with a table listing their redemption values for different combinations of slots. Each of the periods consisted of two markets. The first, or primary, market was a combinatorial, single-round, sealed-bid auction. Once the slots were allocated and priced, the second market opened where participants could buy and sell slots between the other participants. Thus, the auction had an essentially "free" or frictionless after-market. The participants were paid based on their profits.

The tests were run under both high and low contention scenarios. While seven slots were auctioned off in each experiment., the number of combinations of slots differed from 10 in

the low contention auctions to 20 in the high contention auctions. These slot combinations formed the packages for conditional bids. For the low contention auction, the mean number of packages that each bid was a part of was 4 compared to 9 for the higher contention auction. There were no singleton packages in either form. The ratio of items to packages was 0.7 for the low contention auction and 0.35 for the high. Overall, both types of auctions had a high contention level. Additionally, the participants were divided into experienced and inexperienced bidders.

The experiments showed that the efficiency of the combinatorial auction exceeded that of the simple-bid auction in all trials. Also, while the simple bid auction relied heavily on the after market (increasing the efficiency by up to 40% in some trials), the combinatorial auction only improved its efficiency by less than 5% in the after-market. The reliance on the after market was more pronounced for the inexperienced bidders. Experience was found to be less important in the combinatorial auctions than the simple-bid auctions.

Scheduling Resources on Space Satellite

In the late 1980's, the National Aeronautics and Space Administration (NASA) funded research into methods for allocating slots on the Deep Space Network satellite. A slot is a unique pairing of a time period and a satellite resource. The satellite has three types of antennae located at three different locations. The bidders are NASA engineers and researchers who want to use different pieces of the equipment at different time periods for different purposes. Resources, however, cannot be shared during a single time period.

The interdependencies between the slots arise from the different uses of the satellite's antennae. These interdependencies determined the types of packages (set of slots). Four primary types of packages based on different forms of interdependencies were identified:

- Contiguous* The same resource is needed for more than one time period in a row. Thus, the slots in contiguous packages are complementary.
- Periodic* The same resource is desired for multiple time periods separated by fixed intervals. For example, a satellite might need to have its communications equipment readjusted every 36 hours.
- Array* Multiple resources are required during the same time period. For example, to increase accuracy of data transmission, two antennae are used simultaneously.
- Maintenance* Single slots are desired since timing is not important.

Because of the variety of interdependencies, no simple bid structure was identified. However, since this is the scheduling of assets, every slot did not have to be covered. It was fully expected that some of the slots would not be used. This differs from the TL bidding problem where every lane must be covered.

As a result of this project, the Adaptive User Selection Mechanism (AUSM) was developed.⁴⁵ The AUSM auction works as follows. At any time during the auction, bidder k may submit a bid for a package j . The bid is accepted if $b_k > \sum_i B_i \forall p_i \cap p_j \neq \emptyset$. That is, for a new bid to be accepted as the new standing bid for a package, its value must be greater than the standing bids of all packages that will be displaced. Since this is a scheduling problem, it is not required that any newly vacated slot be covered.

The decision rule can be shown by example. Suppose there are three slots (1, 2, & 3) which can be bid on separately or in one of six different packages:

$$\begin{array}{ll} p_1 = \text{Slot 1} & p_4 = \text{Slots 1 and 2} \\ p_2 = \text{Slot 2} & p_5 = \text{Slots 2 and 3} \\ p_3 = \text{Slot 3} & p_6 = \text{Slots 1 and 3} \end{array}$$

Suppose, the current standing bids are $B_1=B_2=B_3=\$10$. A new bid, b_4 , for the package of slots 2 and 3 must be greater than the sum of all standing bids for packages containing slots 2 or 3. Thus, $b_4 > B_2 + B_3 = \$20$. Suppose the bid is for \$21. The new partition is $P = \{p_1, p_4\}$ with the respective standing bids $B_1=\$10$ and $B_4=\$21$. Note that the total value will always increase since each new bid must be higher than the standing bids it is replacing. A complication arises when we consider the next round, however. Suppose a bid is made for package p_4 for slots 1 and 2. To be accepted, $b_4 > B_1 + B_3 = \$31$. A bid of \$32 is accepted and the new partition is $P=\{p_4\}$ with the respective standing bid $B_4=\$32$.

The AUSM does not automatically reassign the original bid for slot 3 back into the allocation. The decision rule to accept a new bid is based on one package at a time - it does not consider the effect of reactivating older bids or the implicit aggregation of separate bids. This gives rise to the threshold problem where packages with more slots are inefficiently included in the final allocation. For example, suppose that bidder A values slot 1 at \$20, bidder B values slot 2 at \$20, and bidder C has the standing bid covering both slots at $B_{12}=\$32$. These new bids will not overturn the standing bid since $b_1=\$20 < B_{12}=\32 and similarly $b_2=\$20 < B_{12}=\32 . Obviously, the more efficient allocation is for the slots to be won individually. The AUSM, however, requires that each bid unilaterally overcomes all conflicting bids.

To overcome this problem, a standby queue which allows the bidders to communicate was added. Essentially, AUSM with Queue makes the standby bids available for other bidders to combine it with their own bid. A standby bid can be withdrawn without penalty and a bidder is allowed only a set number of standby bids. Continuing the small example, if bidder A places a bid of \$17 for slot 1 in the standby queue, then bidder B can combine this with a bid of

⁴⁵ See, for example, Olson and Porter (1989), Banks, Ledyard, and Porter (1989), and Ledyard, Noussair, and Porter (1995)

\$16 to overturn the standing bid of $B_{12}=\$32$. The addition of the standby queue makes the process much more efficient.

Ledyard, Noussair, and Porter (1995) conducted 12 experiments, each lasting between eight and twelve rounds, of the AUSM with Queue. As opposed to the combinatorial auction for the landing slots, this was a multiple round auction. Again, high and low contention auctions were tested. Comparison of contention level is difficult, though. They classified the auctions by modifying the formal measure developed by Olson and Porter for use in the assignment problem [OPP]. The high contention auction had a contention level of 0.55 measured it as the ratio of the sum of the covering constraint dual variables to the optimal objective function value.⁶⁶ Because the variables in their problem were bounded between 0 and 1, the basic variables had positive reduced costs. So, their metric of contention actually captured the ratio of the duals for the upper bounds to the duals of the covering constraints. If these upper bounds are removed, since they are redundant, their measure of contention would go to 1.

Using the simple contention measures, the high contention auction appears to not be very contentious. The mean number of packages that an item belongs to is 1, the percentage of singleton packages is 44%, and the ratio of items to packages is 0.74.

Bykowsky, Cull, and Ledyard (1995) report that an "AUSM-like mechanism" was used for allocating logistics services (trucking) for 850 connected routes for a shipper. Spreadsheets were used to transmit in-between round information to the carriers and to receive bids from carriers. They report that the auction saved the firm an estimated \$10 - \$15 million on a total annual freight bill of \$150 million. No information is available on the details of the auction since the work was proprietary.

Swedish Railroad Scheduling

As part of the Swedish Central Rail Administration's plans to partially privatize the railroad system, research was conducted in developing a decentralized method for allocating track routings. While the ownership of the rail would remain with the state, access to the track, in the form of routings, would be made available to private companies. Brewer and Plott (1995) developed and tested the Binary Conflict Ascending Price (BICAP) mechanism to investigate how well a decentralized auction could allocate track routings to independently bidding firms or agents.

The BICAP mechanism operates on the principle that only certain combinations of items are feasible. Any infeasible combination of items can be reduced to a conflict between two

⁶⁶ In Olson and Porter (1994), the Vickery prices rather than the duals were used in measuring contention.

items; thus there is always a binary conflict at the core of all conflicts in general. The most obvious infeasibilities are potential collisions between trains. This can be seen in the time-space diagram shown in Figure 4.12 for eight potential train slots between locations 1 and 3 with a potential siding at location 2. The diagonal lines marked A through H represent the time and locations of the routings. For example, routings D and B cannot both be awarded since they occupy the same track at the same point in time. Routings C and D are feasible, though, since C is in a sidetrack at the time of D's passing. Another restriction is safety where a certain distance is required between trains. For example, routings E and G may not both be awarded since they do not provide enough time between trains. Figure 4.13 shows the conflict graph. All directly connected nodes represent a binary conflict. These infeasibilities form the basis of the interdependencies between the different routings.

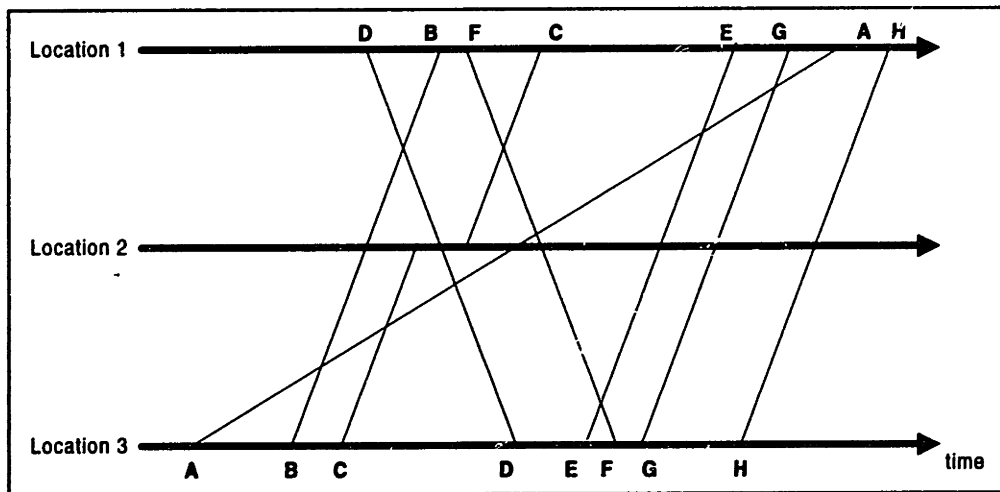


Figure 4.12. Time-Space diagram for rail track slots.

The BICAP auction functions as follows. At the beginning of the auction, the standing bids for each routing is set to zero, $B_j=0$. Firms may submit bids continuously in the form b_{jk} where j is the route desired and k is the bidder's identification. If $b_{jk} > B_j$ then $B_j = b_{jk}$, that is, the standing bid is set to the new high bid for that route. The identification of the firm with the standing bid is also made public. If a bid is accepted as the new standing bid for a package, the allocation is recomputed. The new allocation is found by solving the "track value" optimization problem which finds the feasible set of standing bids that have the highest bid value. For the rail routing problem this is achieved by looking at the set of conflicted pairs of routings, thus the name of the auction. No details are provided on the procedure used to solve this problem. Once the new allocation is reported, the next bid is processed. If the auction is continuous, then it will close after no bids are received in a set period of time. If discrete rounds are used, then the auction closes after no bids are received for a set number of rounds.

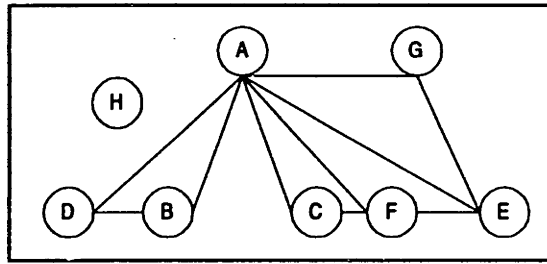


Figure 4.13. Conflicts due to adding a conditional bid.

While the tests that have been run to date do not allow for bids on sets of routings⁶⁷, many of the ideas are readily transferable to conditional bidding. In one sense, a bid on a routing can be thought of as a bid on all of those routings that it is in binary conflict with. So, a bid for routing B in Figure 4.12 could be considered a combinatorial bid for the package {A, B, and D} since none of these routings are allowed if routing B is awarded. Also, the mechanism could easily include explicit conditional bids by modifying the conflict graph. For example, suppose that for the situation shown in Figure 4.12, an agent wishes to bid on routings B and F together as an all-or-nothing package. Because the model works by considering conflicts, this can be included in the conflict set as shown in Figure 4.14. The newly constructed conditional bid of routings B and F adds the additional conflicts.

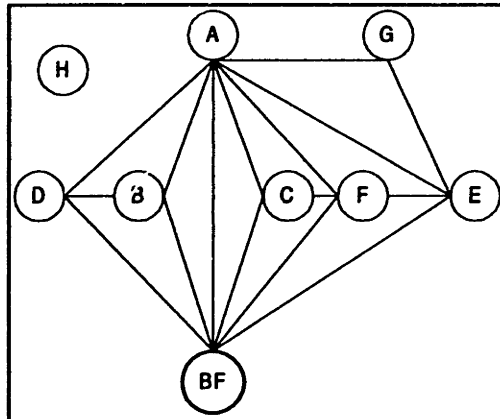


Figure 4.14. Conflicts due to adding a conditional bid.

The BICAP Mechanism was tested in a series of experiments involving students submitting bids via computer terminals for a stylized rail network similar to that shown in Figure 4.12. The bidding was run as a continuous time auction whereby each bidder could submit a bid, (bid price, train route), at any time during the auction rather than at discrete intervals.

The test-bed experiments achieved an average efficiency of 97% with 18 of the 21 trials achieving 100% efficiency. Additionally, no agents submitted *dominated* or *dominated neutral bids*

⁶⁷ They define a set along the lines of a conditional bid, where a set can be “interpreted as a request that all bids be accepted or none be accepted” (Brewer and Plott 1995, 10).

and generally did not allow the bidding to stop as long as *pivotal bids*, or even *strong neutral bids*, remained. That is, participants would continue bidding until they had no possibility of becoming the standing bid for any of the packages.

Essentially, the *strong neutral bids* indicate where agents are trying to “negotiate” a way to introduce currently non-partitioned train routes into the allocation. Brewer and Plott note that “making bids on unallocated trains an agent is contributing to the ‘public good’ of defeating the current allocation.” This is a method of revealing the value of the unallocated routing rather than letting the market close down. The whole process essentially lets the market “wander” from potential allocation to potential allocation until the optimal one is discovered. The signals to move are contained in the *pivotal bids*, but also, and more importantly, the *strong neutral bids*.

The contention of the auction was not discussed. But, if we consider a bid for a routing actually being a bid for all conflicted routings, the contention appears to be quite high. There were 8 total packages with the median number of packages an item belonged to equal to 4.4. There were just 22% singleton packages and the ratio of items to packages is one.

FCC Frequency Spectrum Auctions

In the summer of 1993, the Federal Communications Commission was given the authority to use competitive bidding procedures to award initial licenses for Personal Communication Services in the 900 MHz band. This grouping of frequencies is commonly referred to as ‘narrowband PCS’ and was being made available for use by mobile and portable radio communications services, interactive video data services, private paging, and other forms of personal communications for both businesses and individuals. Each license specified a frequency and a geographic coverage area. A total of 3,554 narrowband PCS licenses were to be issued: 11 for national coverage, 30 licenses across the 5 regional areas, 561 licenses across the 51 major trading areas (MTAs), and 2,952 licenses across the 492 basic trading areas (BTAs). While these services are expected to be in very high demand in the near future, the actual value of each license was unknown.⁶⁸

PCS is a new service with many inherent uncertainties and very high interdependencies - both as substitutes and complements. Complementary licenses allow bidders to achieve economies of scope, facilitates roaming over large areas, and decreases the interference at license boundaries. Other licenses acted as substitutes in providing similar services or connecting the same regions through alternative routes.

The use of competitive bidding was a change for the FCC. It had traditionally awarded initial licenses to interested parties by random lottery or committee hearings which took several

⁶⁸ Estimates of their value ended up being orders of magnitude off of the final prices paid.

years to grant a license.⁶⁹ By using a competitive bidding scheme, it was hoped that significant revenues could be generated for a previously 'free' good. Additionally, the auction was thought to lead to faster deployment, promote economic opportunity and competition, recover a portion of the value of the public spectrum, and encourage the efficient and intensive use of the electromagnetic spectrum.

The FCC used as its general guiding hypothesis the idea that "licenses generally should be awarded to those who value them most highly" (FCC 2nd Order, 71). Some additional concerns that the bidding needed to address were facilitating the efficient aggregation of licenses, awarding licenses rapidly, and avoiding excessive implementation costs and complexity.

Based on these concerns, a spirited⁷⁰ debate began between the numerous academicians working for the players. The sheer magnitude and complexity of the process invited high powered analysis. The commission received hundreds of comments, reply comments, and *ex parte* presentations relating to the auction design. The most prestigious names in game theory and economics were hired by the major telecommunications companies in order to help design the auction and to improve their strategy in the actual auction itself. The interaction between academia and industry provided a great opportunity to test some of the more theoretical aspects of auction design on an unprecedented scale.

The primary debate was whether to hold simultaneous multiple-round ascending-bid auctions *with* or *without* combinatorial bidding. The vast majority of the economists and game theorists involved were opposed to combinatorial auctions. The objections fell into five categories: bias towards large bidders, computational difficulty, potential for gaming, non-transparency for bidders, and lack of real benefits.

The bias towards large bidders was based on the threshold problem. It was thought that some of the national bidders would make large conditional bids which would be difficult for a coalition of smaller, and more efficient, bidders to break. This problem was initially referred to as a version of the free rider problem where bidders have a disincentive to overcome a public good. The counter arguments focused on auction designs which allowed forming of coalitions, such as AUSM.

⁶⁹ "These licenses were granted in record time," said Regina Keeney, Bureau Chief of the Wireless Telecommunications Bureau. "Before, when the Commission granted licenses under the lottery system, it took more than a year from initial application to license grant. When licenses were granted by comparative hearings, it often took several years." (Keeney 1995).

⁷⁰ For example, in response to a statement that the bidding would take excessively long since only one bidding round was to be held per day, one professor replied (in the official responses, "Prof ___ fails to distinguish between auctions and multi-vitamins: only the latter need to be taken once-a-day." In response to another experts critique, one author exclaimed "we are at a loss to explain this confusion." and sums up the situation by stating that "experience has shown that even economics Ph.D. students have trouble understanding the description."

The large number of potential combinations that could be made were also thought to require computational proficiency that was beyond many of the bidders capabilities. Or, that reliance on using newly developed and untested software could lead to questions as to whether the computer had been programmed correctly and licenses awarded properly. The counter arguments were that the seemingly simple auctions without conditional bids actually required more complicated stopping and activity rules.

The third complaint was that combinatorial auctions present a large opportunity for bidders to 'game' the system. Bidders, it was thought, would, by bidding on complicated combinations designed specifically to confuse opposing bidders, block other bidders' potential aggregations, or simply "muddy the water." The counter arguments were that some amount of gaming was going to occur and that even the simple auctions will be gamed to some extent.

A fourth perceived drawback was the lack of transparency for the bidders. The threshold problem poses difficulties in knowing when and how to form coalitions with other bidders in overcoming larger bids. Also, it is difficult for a bidder to determine how to counter-bid when facing a combinatorial bid since it is competing against only a portion of the bid. Again, the counter arguments were that the auction design can facilitate the forming of coalitions and provide information for bidders to formulate new bids.

The fifth drawback with combinatorial auctions, was the lack of real benefits. Several critics of combinatorial auctions discounted the 'exaggeration' of interdependencies on a national level commenting that the interdependencies are "not an all-or-nothing proposition" since the total value of a combination of licenses is not lost if one, or even several, of the licenses is not won. Other critics thought that coalitions of regional winners could just as easily achieve what a single owner could due to competitive pressures. Nalebuff and Bulow argue that supposed gains in flexibility due to combinatorial bidding are "somewhat of an illusion" since combinatorial bids have more of a 'lock-in' effect, especially if bid withdrawal is limited.

The commission ended up rejecting the use of combinatorial bidding for use in the PCS auctions. While acknowledging the benefits of combinatorial bidding "especially in terms of efficient aggregation of licenses, we concluded that simultaneous multiple round auctions offer many of these same advantages without the same degree of administrative and operational complexity and without biasing outcomes in favor of combination bids" (FCC 3rd Order, 24).

The final auction was designed as a sequential series of simultaneous auctions moving from the largest to the smallest markets. Simultaneous multiple-round auctions were held for the largest and most valuable licenses while single-round, simple sealed-bid auctions were used for the smaller licenses.

As of this writing, four auctions have been held. The auction for the 10 national licenses ran from 25-29 July 1994, lasted 47 rounds, and raised a total of \$617 million. The auction for the

30 regional licenses (6 each in 5 regions) was run from 26 October to 8 November 1994, lasted 104 rounds, and raised a total of \$395 million. In addition to these narrow band spectrum auctions, the FCC held separate auctions for broadband licenses and IVDS (interactive video and data services) licenses. The IVDS auction brought in \$214 million. The broadband auction was held in March of 1995, lasted 112 rounds, and generated \$7.02 billion! The FCC estimated that the revenues were equivalent to \$35 for each US citizen or \$98 for each household.⁷¹ By all accounts, the auctions were an extremely successful way of pulling close to \$9 billion dollars literally out of the air.⁷²

4.4.3 Comparisons and Relevance to TL Bids

In the four cases presented above, three allowed conditional bids for packages of items while the fourth opted for simple bids. Each of the three combinatorial auctions proposed a different mechanism. The airport landing slots case used the uniform price single-round auction proposed by Rassenti, Smith, and Bulfin (RSB). Allocating resources on a deep space satellite used the Adaptive User Selection Mechanism with Standby Queue (AUSMQ). The scheduling of railroad routings used the Binary Conflict Ascending Price (BICAP) Mechanism. The fourth case of allocating frequency licenses for the FCC used a simultaneous multiple-round simple-bid mechanism. This section briefly compares and contrasts these cases and their respective mechanisms along three topics: value revelation, handling rejected bids, and characteristics of interdependencies.

Each of the four mechanisms were designed to provide bidders an incentive to reveal their true valuations. The AUSMQ, BICAP, and FCC auctions employed multiple rounds while the RSB mechanism used second-pricing. An additional reason for using multiple rounds is to introduce additional market information into each bidder's valuations. Because airport landing slots are routinely auctioned, the value is not completely unknown. The FCC licenses and the space satellite, on the other hand, are items whose values are not known very well. Multiple round auctions assist in determining the value of items in addition to allocating them.

A multiple round auction has to handle rejected bids. If simple bids are used, then the current standing bid replaces all other bids for an item. For conditional bids, however, only a portion of the standing bids for packages will be active, that is, in the current partition. The primary difference between the AUSMQ and BICAP mechanisms is on this point. The AUSMQ posts the rejected standing bids for each package providing an opportunity for other bidders to

⁷¹ These figures as well as very detailed, round by round bids, are available on the FCC's World Wide Web page: <http://www.fcc.gov/>.

⁷² On 21 March 1994, the FCC announced that it received \$1.4 billion down payment from the winning bidders. They marked this event as the "largest one-time transfer of wealth from the private sector to American Taxpayers." This is just the 20% down payment required within 5 days of auction closing.

combine them. The BICAP mechanism finds the winning bid for each package and then finds, for the given set of packages, the combination which provides the highest possible value. The key difference is that BICAP can potentially activate an inactive standing bid while the AUSM requires that the bidder specifically place a bid in the standby queue in order to be combined by another bidder. Both mechanisms provide full information to the potential bidders.

Finally, the nature of the interdependencies places a significant role on the auction design. The interdependencies between the spectrum licenses in the FCC auction, for example, were not considered to be strong enough to warrant combinatorial auctions. Airport landing slots, on the other hand, lose all of their value if the package is partially allocated. The resources for a satellite were shown to have several different forms of interdependencies so that the auction needed to be able to handle different forms of packages. Generally, it seems that the value of the interdependencies, combined with the level of degradation of this value if the package is not complete, determines whether a combinatorial auction is worthwhile.

These four cases appear to be the only publicly available examples of combinatorial auctions used in practice or experiments. As such, they provide a baseline of how to design a combinatorial auction for TL bids. The biggest lesson is that the majority of the problems associated with combinatorial auctions are caused by having multiple rounds. The threshold, transparency, and gaming problems identified by the critics of the FCC auction are only relevant if the auction has multiple rounds. The main contribution of both the AUSMQ and the BICAP auctions was the development of a new way to handle bids for the multiple rounds. Overall, then, the use of multiple rounds for combinatorial auctions should be approached with caution. Some of the other concerns that were raised against combinatorial bids for the FCC auction are not applicable to the TL problem. For example, the bias towards larger bidders should be encouraged in the TL problem. Additionally, the after-market is generally allowed in TL bids. In some cases, though, a winning carrier is allowed to broker out the lanes they won, but the originally winning carrier still maintains responsibility over the service.

4.5 Closure

The objective of this chapter was to analyze the TL bidding process and develop a framework for designing an auction for different TL networks.

The chapter provided an exhaustive review of auction design theory with specific emphasis on the TL bidding problem. The analysis showed that the current practice of single-round, sealed-bid auctions using simple bids (single lane per bid) rarely finds an efficient allocation if there is any degree of interdependency between the lanes caused by economies of scope, capacity constraints, and shipper-specific restrictions. The need for carriers to assume

exposure risk when trying to form efficient aggregations in a simple bid auction was demonstrated through examples.

The effect of carrier hedging on the bidding process was discussed. It was shown that two major sources of a carrier's heading are uncertainty in (1) the accuracy or reliability of the estimated flow information and (2) the connection cost. The shipper can reduce the need for carrier hedging of the former type by providing detailed (disaggregated) spatial and temporal information on the expected traffic flows. Reducing the hedging of the latter type requires the use of conditional bids.

Finally, combinatorial auctions (auctions that require the use of conditional bids) were introduced. Combinatorial auctions were shown to be very effective at arriving at an efficient allocation when the items (lanes) exhibit interdependencies. The drawbacks associated with combinatorial auctions (complexity, lack of transparency, etc.) were shown to be critical only when multiple rounds are used.

The next two chapters apply the auction concepts discussed in this chapter to the particulars of the TL bidding problem.

Chapter 5

Pre-Bid Network Analysis

The major theme of this dissertation is that the shipper's behavior has the power to influence TL carriers' economics. We have argued that the characteristics of the network should be used to determine the form of the contract between shippers and carriers as well as the bidding procedure to use. This chapter presents a methodology to be used by shippers to determine those characteristics needed to make TL procurement decisions. The contribution of this chapter is the development of the pre-bid analysis.

The chapter is organized chronologically along the five steps of the analysis: (1) Expanding the Truckload Network, (2) Constructing the Network, (3) Assessing Matching Opportunities, (4) Identifying Potential Efficient Aggregations, and (5) Designing the Truckload Bid. Section 5.1 discusses six different ways that a shipper can increase both the size of its distribution network and its attractiveness to carriers. Section 5.2 describes how the origin-destination (OD) flows should be aggregated into lanes and locations that are ultimately presented to the carrier for bidding. Section 5.3 examines the temporal distribution of the loads in greater detail and demonstrates how to estimate the probability for matching inbound and outbound lanes. Section 5.4 presents heuristics for identifying the relevant types of efficient aggregations of lanes within the network. While carriers are expected, and should be encouraged, to identify their own conditional bids (perhaps with lanes from other shippers) the shipper can assist the carriers by identifying and suggesting those bundles which are self contained within its network. Section 5.5 addresses how the bid should be designed based on the network characteristics and, finally, Section 5.6 summarizes and closes the chapter.

Before presenting the pre-bid analysis, the two sets of data used for the analysis in this and the following chapter are briefly described. The purpose of the data examples is to illustrate the concepts - not to provide a full analysis of the different TL networks.

Military Traffic Management Command (MTMC) Truckload Network Data

The first data set consists of 1.45 million freight bill records for the period from August 1992 to July 1993 for domestic shipments across all modes for the Military Traffic Management Command (MTMC). The MTMC has authority over all military traffic shipped via commercial (non-military) vehicles. While TL shipments were not specifically identified in the freight bill records, all shipments using motor carrier that weighed 10,000 pounds or more were considered to be TL shipments. This created a set of 191,555 TL shipping records.

The network consists of just under 5,000 shipping points, 25,000 origin-destination flows, and a total annual freight bill of approximately \$242 million. Over 80% of the traffic moved on less than 15% of the lanes. More details of the network and traffic flow are presented later in the chapter as examples to the pre-bid analysis.

More than 1,200 different TL carriers were used throughout the year. One reason for such a large number of carriers is that the MTMC is encouraged to offer business to small carriers in regions where military bases are located. This is primarily politically driven. An attempt to rationalize the carrier base was discontinued due to pressure from congressional members concerned with small trucking firms in their constituencies.⁷³

Each record contains the information for a single shipment and includes: Origin, Destination, Carrier, Trailer Type, Weight, Total Cube, Charges Paid, Pickup and Delivery Date, Ton Miles, and Carrier Mode. The distances for the shipments were not provided directly in the freight records but instead were calculated as the weighted average of the Ton Miles divided by the Weight for shipments on common point to point movements.

The origins and destinations were specified by Standard Position Location Codes (SPLC) rather than the more familiar United States Postal ZIP Codes. While both systems are essentially easy ways to define a zone or geographic region, there are differences between the two systems. A short comment on the difference between the two is in order.

Postal ZIP codes are based on population densities and political boundaries. They differ widely in the amount of area covered. For example, the state of Connecticut (~5,000 square miles and population density of 677 people per square mile) has ten 3-digit ZIP codes while North Dakota (~70,000 square miles and population density of 9 people per square mile) has just nine. Both states have a single 2-digit ZIP code. So a 3-digit ZIP covers, on the average, between 500 and 7,700 square miles!

SPLCs have 6 digits, are finer grained, and are better suited for geographic use. The first position specifies one of 8 regions across the U.S., the second specifies the state (or intra-

⁷³ Phone conversation, Patty Maloney, Division Chief, Transportation Services, Negotiations, Military Transportation Management Command, 6 March 1995.

state sections for the larger states, e.g., North and South California) within the region, the third and fourth specify the county, and the fifth and sixth identify a specific shipping point - often a railroad station. While the size of counties (4-digit SPLCs) do vary across states and regions, the difference is not as extreme as for ZIP codes. For example, Connecticut has 10 counties while North Dakota has 38. Thus, the area per 4-digit SPLC ranges between 500 and 1,800 square miles - a much tighter range than ZIP codes.

There are three major shortcomings of the MTMC TL network data set. First, because it was all hand entered, the sample is quite 'dirty.' The primary correction made was to ensure that the SPLC matched the city-state designation. The city state designation had priority over the SPLC codes. The second shortcoming is the lack of detail in pickup and delivery times. For more detailed analysis, the specific time of shipment is preferred since the time scale for trucking connections is in hours while the data is presented only in daily buckets - and the accuracy of these entries is questionable. For the lack of any better empirical evidence, the distribution of loads across a day is assumed to be uniform.⁷⁴ The final shortcoming is that neither the advance notice time nor the desired pickup and delivery windows are provided. A better analysis, especially concerning the probability of matching loads, could be made if the warning time provided to carriers and the earliest and latest pickup and delivery windows were included in the data. This information is not typically provided in freight bill records, however.

ABC Truckload Bid Data

The second data set is from ABC, the disguised name of a firm specializing in conducting transportation bids. The data used in this research consist of a joint TL bid held for the 1992-1993 time period for two dozen customers (shippers) bidding simultaneously. The data also included information on the network lanes, shipping locations, and the bids submitted by the carriers. The ABC data do not contain any individual freight bill information - only aggregated annual lane volumes.

The ABC network consists of 1369 shipping locations at 5-digit postal ZIP code level, 1,743 origin-destination flows, and a total estimated freight bill of \$40 million. Three hundred and fifty carriers submitted a total of 82,843 bids for lanes on the network. More details of the network and traffic flow are presented later in the chapter as examples to the pre-bid analysis.

⁷⁴ This is actually a conservative assumption since transportation activity is usually concentrated in the afternoon, towards the end of the business day.

5.1 Step 1: Expanding the Network

The first step in the pre-bid analysis is to expand the TL network as much as possible. The larger the TL network, the more potential there is for forming efficient aggregations of lanes. The TL network can be increased by either converting traffic moving on other modes to TL traffic or taking over TL procurement responsibilities from other players in the supply chain (different companies or different departments within the shipper's own company). This section discusses six areas that can be exploited to increase the TL network size: adding inter-facility, inbound, and outbound moves, extending along the supply chain, combining different modes, and partnering with other shippers.

Inter-Facility Movements

Inter-facility movements are usually the most stable and predictable traffic flows in a shipper's network. These movements include all shipments between a shipper's own facilities; typically from a plant to a distribution center (DC) or from plant to plant. Because the shipper controls both ends of the lanes, these movements will most likely go in truckload quantities. A private fleet, if the shipper has one, would be used on these lanes. The opportunity for expansion on inter-facility moves is primarily to convert from private fleet to for-hire carriers.

Inbound Movements

The second target for TL expansion is the inbound traffic from suppliers to the shipper's production and assembly plants. In many companies the transportation for these movements are procured and managed by a separate purchasing function rather than the transportation department. Historically, the inbound movement of the raw materials was treated separately from the outbound of the finished product. This makes sense if the inbound and outbound shipments move on different modes, such as for some steel mills where all inbound is by rail and outbound is by flat bed truck. In many companies, though, the inbound is procured through a totally different channel than the outbound, even though they share modes.

Procuring inbound and outbound transportation separately is inefficient for three reasons. First, the network volume is divided which lessens the shipper's buying power. Second, by bidding out disjoint segments of a distribution network, potential opportunities for synergies between lanes are lost. The inclusion of inbound shipments to a traditional outbound network could significantly increase the potential for matching at the production plants. Third, procurement departments typically do not possess the same level of expertise in transportation practice as does the transportation department. For example, a procurement department might

search and negotiate for the lowest possible LTL rates for inbound delivery of small shipments without considering forming multi-stop TL milk runs to bypass the LTL carriers altogether.

A related situation is where a supplier procures the inbound transportation and includes it in the price of the product with no explicit separation. This makes it difficult to monitor how much the transportation actually costs or the terms under which it was procured. Small suppliers are most likely not able to achieve the buying and bargaining power that the shipper could apply. The question of who should control the transportation for movements between a supplier and a buyer is often debated. Gentry (1991, 18) argues that purchasing transportation free-on-board (FOB) origin⁷⁵ allows the shipper to monitor and control the inbound in the same manner as the outbound movements. Larger shippers, then, will sometimes require that all shipments be delivered FOB origin.

Outbound Movements

A third area for TL expansion opportunities is the outbound movement from DCs to customer locations. This could include (1) the conversion of traditional LTL lanes to TL lanes and (2) taking responsibility for TL lanes that were previously controlled by the customer - making these movements FOB destination.

Certain LTL lanes from a DC to distant customer locations can be consolidated into fewer, and less expensive, linehaul movements to a common regional LTL carrier's location. The idea is to consolidate the LTL moves and perform the linehaul portion of the movement independent of the LTL carrier. This reduces the distance traveled at the higher priced LTL rates. By performing the linehaul directly, local LTL carriers can be used at the distant customer location.⁷⁶ This strategy can also speed up the transit time since the shipper is removing all but one of the consolidation terminals from the movement.

Extend Along Supply Chain

A fourth area for TL expansion is to include those movements in-between either suppliers or assembly locations that are not specifically owned by the shipper. For example, Addison-Wesley, a book publisher, subcontracts the printing and binding processes to small, specialized firms. In order to retain control over the process, however, Addison-Wesley procures, manages, and pays directly for all of the transportation of the product along each

⁷⁵ FOB origin means that the buyer takes control of the shipment at the moment it leaves the supplier's location. FOB destination means that they do not take control of the shipment until it reaches their own dock - and thus they do not control or manage the inbound transportation. The payment of the transportation services can be included in the price, broken out separately but paid by the supplier, or paid directly by the buyer. Practices differ tremendously between shippers.

⁷⁶ This practice is called 'zone-skipping' and is used quite often in the parcel package segment.

stage of the process, even though they do not own the specific origins or destinations.⁷⁷ Essentially, they have expanded their TL network along the logical flow of the product beyond the traditional legal confines of their own distribution network.

Combine Different Modes

A fifth area for TL expansion is to combine compatible modes into the same auction. Usually, separate bids are held for the flatbed, dry freight, and refrigerated TL segments. It is possible, though, that the movements on the different modes can complement each other. For example, one shipper found that the inbound lane to a production plant for dry freight could be matched to the outbound lane for refrigerated vans. Thus, a two lane closed loop tour was formed where a carrier could serve both lanes using a refrigerated van. Other opportunities include combining flat bed and dry van movements.

Partner with Other Shippers

A sixth area for potential TL expansion is to conduct a joint bid with other shippers. While a joint bid does not actually increase the size of the shipper's own TL network, it increases the buying power. A joint bid is more attractive to carriers because it allows them to make conditional bids across a wider selection of lanes. The tendering processes for each company do not need to be combined, but using a centralized load dispatching operation would improve the probability of matching across the networks.

Potential candidates for joint bids should have complimentary TL networks and be non-competitors. That is, the effect of joining the different shippers' networks should be to improve the balance of the traffic flow; not just add volume. Perhaps the best choice of partners for a joint bid are buyers and suppliers. Joint bids have been conducted by a health and beauty manufacturer paired with a packaging company as well as an aluminum manufacturer and a brewery. Joint bids are also being used by third parties - such as the ABC data set.

The end result of the network expansion is to create the largest possible set of potential shipment demand and supply points for TL traffic. For example, the transportation department of one office supply company has responsibility for all inter-facility movements from the company's 10 plants to its 3 DCs. These lanes are bid out annually. They also control outbound shipments from DCs to customers, which are all moved essentially by LTL. All inbound movements from suppliers to plants, however, were managed by the purchasing department. The inbound transportation was arranged by the supplier and the cost was included in the

⁷⁷ Phone conversation, Lisa Myers McKinley, Manager of Logistics Services, Addison-Wesley Longman Publishing Company, 13 November 1995.

purchase price of the materials. While the existing TL network for this company showed no potentials for matching, since the plants only originated TL lanes and the DCs only received incoming TL shipments, the network was expanded in two ways to increase matching opportunities. First, by incorporating the inbound movements to the plants, the potential for matchings at the plants are increased. Also, because the suppliers are widely dispersed, there are additional opportunities to match lanes in regions where suppliers are located near a DC. Second, by clustering customers into regions, the company initiated a TL line-haul lane from the DCs to the larger clusters. A local trucking company is used for break bulk and delivery of the shipments to the individual customers. By expanding the TL network, more opportunities for matching lanes will usually emerge.

5.2 Step 2: Constructing the Network

The second step of the pre-bid analysis is to determine how to represent the network to the carriers for the TL bid and is essentially a first cut at examining the TL flows using the static monthly or annual aggregated flows. The analysis involves examining the OD flows from the expanded network and determining how to aggregate them into lanes. Rather than deciding on the classification of lanes *a priori*, the network characteristics are used to dictate the granularity and form. Additionally, the shipper's priorities are used to guide the process in order to construct a network that achieves a specific objective. In our case, the objective is to design the network that minimizes the carriers' need to hedge their reservation values (and hence their bids) due to either incomplete information or uncertainty in making connections.

The following terminology is used throughout the rest of this discussion. A *point* is a specific shipping location, either an origin or destination. In the MTMC data, a point corresponds to a shipping location at the 6-digit SPLC level while for the ABC data it is a 5 digit ZIP code. A *zone* is any set of shipping locations which are aggregated into a common region and treated as a single entity. Thus, four types of lanes are possible: *point to point*, *point to zone*, *zone to point*, and *zone to zone*.

The form of the network is ultimately determined by how the shipping points are aggregated into zones. Once the zones are formed, the lanes are essentially defined. Of course, the lane characteristics influence the zone aggregation. Figure 5.1 shows the three basic options for a network with two plants and ten destination points. The shipper can: (a) treat each origin-destination flow as a separate lane, (b) aggregate some of the points into zones while leaving others as points, or (c) aggregate both shipping and destination points into zones.

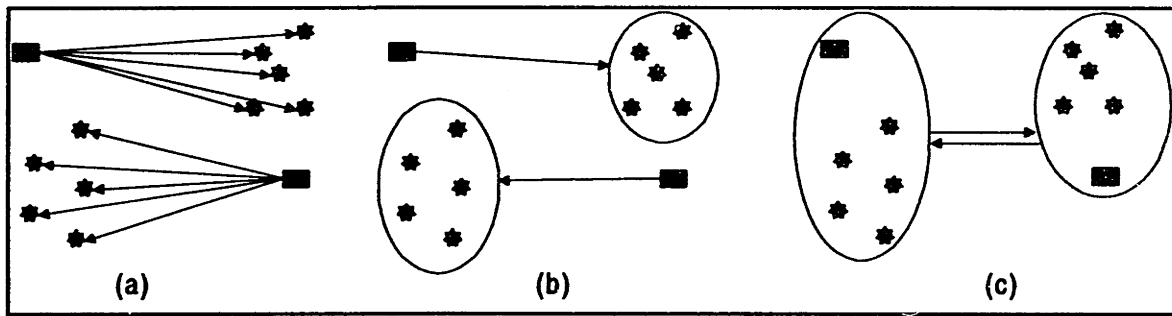


Figure 5.1 Example of lane and zone identification.

While there are several factors involved in forming the zones, the basic trade-off is between lane density and precision. Treating every OD flow as a separate lane, as in Figure 5.1 (a), will result in very low lane volumes, but there will be no ambiguity in the lane definition. Conversely, aggregating into large zones, as in Figure 5.1 (c), will increase the lane densities, but will add to the uncertainty of the actual origins and destinations within the zones.

The shipper needs to make this trade off based on their priorities. For example, if a shipper wanted to simplify the tendering decision for the plant dispatchers, the network could be configured to treat each plant as a separate point with all destinations from each point consolidated into a single zone. Another objective could be to gain universal coverage, in which case the shipper could simply aggregate all of the points into, say, 3-digit ZIP Code lanes. The MTMC, when designing their network, originally selected 11 zones for both inbound and outbound destinations for the 180 military shipping bases. They later increased the number of origin zones to 34 due to complaints from the carriers over the zones being too large.

As stated earlier, the over-riding objective for the design of the network is to minimize the carriers' need to hedge their bids. The way in which the network is represented to the carriers will influence their bidding behavior. As detailed in Chapter 4, a carrier will hedge its reservation values (and in turn its bids) based on (1) the accuracy of the network information and (2) the uncertainty in making connections. The need for accuracy means that the information should be precise enough to reduce carrier hedging due to uncertainty in lane distances and volumes, but still have sufficient lane density to be reliable. Hedging against the connection uncertainty suggests that the design of the network should enhance the carriers' ability to identify and form conditional bids on lane bundles.

5.2.1 Shipping Point Classification

Traffic flow through most distribution networks is concentrated through a small number of shipping points. The general form of the OD flows in a shipper's network will be closer to the network in Figure 5.2 (a) than (b). Because of this, we can focus on first identifying

the critical shipping points. Shipping points can be classified according to the total volume of loads handled at each point and the balance - the ratio of inbound to outbound loads. The objective is to identify those shipping locations to be treated individually rather than combined into a zone.

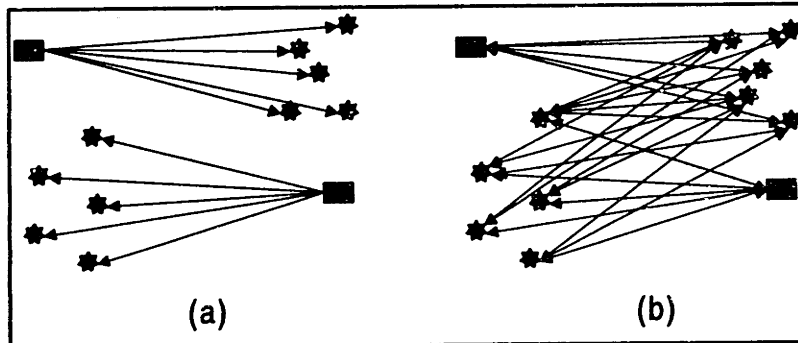


Figure 5.2 Concentration of flows through points.

The MTMC network consists of 4,925 distinct 6-digit SPLC shipping points (27% pure outbound, 42% pure inbound, and 31% with both inbound and outbound). Over 80% of all of the loads in the MTMC system were handled through fewer than 8% of the points. Aggregated at the county (4-digit SPLC) level, there are 2,356 counties originating or receiving loads. The movements originated and terminated in 1,600 and 1,900 counties, respectively, and all states, except Hawaii. The ABC network consists of 1,369 5-digit ZIP shipping points (45% pure outbound, 28% pure inbound, and 27% with both inbound and outbound). Over 80% of all of the loads in the system were handled through approximately 15% of the points. There are 527 3-digit ZIP shipping points. Table 5.1 summarizes the traffic at the shipping points for both the MTMC and the ABC data at their finest level of detail.

| Volume | MTMC Data | | ABC Data | |
|--------------------|--------------|----------|--------------|----------|
| | Total Points | % Points | Total Points | % Points |
| loads ≤ 1 | 1843 | 37% | 0 | 0% |
| 1 < loads ≤ 10 | 1613 | 33% | 504 | 37% |
| 10 < loads ≤ 100 | 976 | 20% | 591 | 43% |
| 100 < loads ≤ 1000 | 425 | 9% | 248 | 18% |
| 1000 < loads | 68 | 1% | 26 | 2% |
| Total | 4925 | | 1369 | |

Table 5.1 Summary of shipping point volume at the 6-digit SPLC and 5-digit ZIP level for the MTMC and ABC data, respectively.

Comparing the balance of the shipping points for the two TL networks sheds light on the network differences. The plot of total annual shipping volume versus the percentage of outbound loads is shown in Figures 5.3 and 5.4 for the MTMC and ABC data, respectively. Shipping locations located at the far left of these figures are pure inbound points, the locations at

the far right are pure outbound points, and those in the middle handle both inbound and outbound loads.

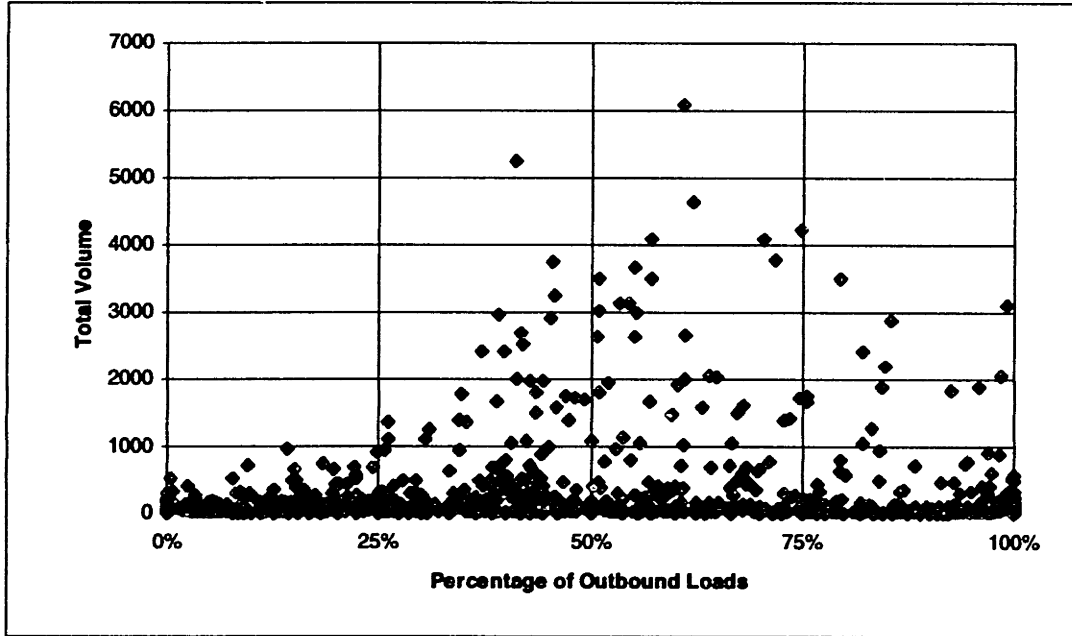


Figure 5.3 Classification of MTMC shipping locations by volume and percentage of outbound loads handled.

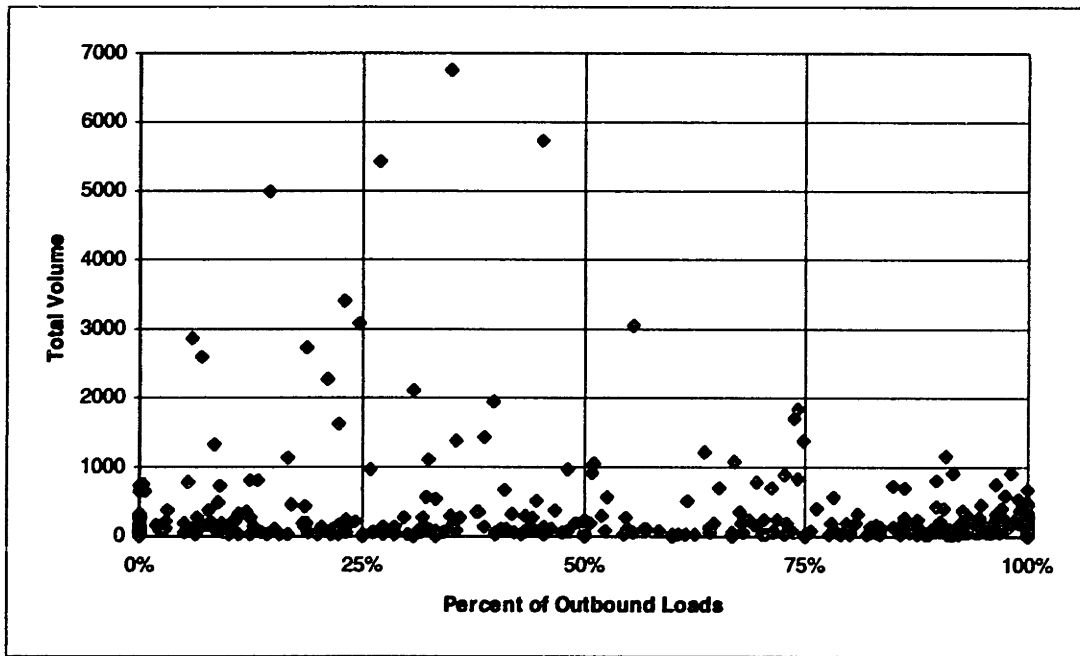


Figure 5.4 Classification of ABC shipping locations by volume and percentage of outbound loads handled.

As Figure 5.3 shows, the MTMC network is primarily for outbound distribution. This is suggested by the large number of small-volume, inbound and large-volume, outbound shipping

points. This is just the opposite of the ABC network, Figure 5.4, where the larger shipping points handle mainly inbound loads. ABC, then, is primarily an inbound oriented network.

The shipping location analysis serves two purposes. First, the analysis identifies the critical shipping locations. Second, the general orientation of the network (inbound or outbound) is shown. This suggests which types of lanes to form. For inbound oriented networks, the lanes are aggregated zone-to-point while for outbound oriented networks the lanes are aggregated point-to-zone.

5.2.2 OD Flow Classification

The origin-destination flows between shipping points can be classified according to distance and volume. The purpose of this step is to identify different types of lanes based on the linehaul distance and the flow characteristics.

The distance is important because it dictates the approximate time required to service a load. A truck can travel approximately 450-500 miles per work day. The length of the lane influences how much advance notice the carrier has to find a follow-on load as well as the acceptable amount of empty miles at the conclusion of the load. Four approximate distance categories are identified. While the specific distances used to distinguish the types of lanes are approximations, the types are common across networks.

Short Hauls: The short hauls are those lanes in which a vehicle can service the load and return to the origin in the same day. The approximate maximum short haul lane distance is 250 miles.

Regional Hauls: Regional hauls are lanes which travel for the better part of a day. They are useful in combining into tours. Regional hauls are generally between 250 and 500 miles in length.

Mid-Range Hauls: The mid-range lanes require approximately two days. They are between 500 and 1000 miles.

Long Hauls: The long haul lanes require more than two days to complete and are candidates for driver teams. Team drivers can keep a truck moving continuously since one driver sleeps while the other is driving. These are worthwhile for the long haul lanes where the amount of time saved can be quite substantial. These lanes are at least 1000 miles in length.

The total volume of the OD flow has two effects. First, higher volume makes an OD pair more attractive to carriers. Higher volumes increase the reliability of the OD flow, makes the capacity planning easier, and improves vehicle utilization. If the volume is sufficiently high, a vehicle can be dedicated solely to that lane, or set of lanes.

Second, the volume of traffic on a lane, and in the system as a whole, can determine whether the auction will be by location, lane, or load. All of the discussion in the previous chapter has assumed that carriers will be assigned traffic lanes. While we have mainly discussed the case where a single carrier is assigned to each lane, it is easy to assign any number

of carriers to a lane. If the volumes are very low, however, it might be beneficial to bid out entire portions of the distribution network. For example, all traffic leaving a certain plant. On the other hand, if the volumes are very high, then a single carrier might not be able to handle a lane by itself. Instead of arbitrarily determining a set number of carriers to assign to each lane, each load could be auctioned off. Bidding by load requires each carrier to submit either a percentage or an absolute number of loads that they are willing to haul for each lane. The proportion of loads won by a carrier on a certain lane could be used in the tendering decision.

Four classes of lane volume were used.

Daily: These are the high volume lanes with at least 250 loads/year. The number of work days within a year differs between companies. For the analysis of the MTMC network, we assumed 250 work days per year.

Weekly: Weekly loads are considered less reliable and are therefore generally aggregated into zones, if possible. The range is between 52 and 250 loads per year.

Monthly: Monthly loads are considered even less reliable than the weekly lanes. The threshold is between 12 and 52 loads per year.

Infrequent: These are the lanes with fewer than 12 loads for the entire year. These are extremely unreliable and are best aggregated into the largest zones possible.

Tables 5.2 and 5.3 show the distribution of both distance and volume for the OD flows for the MTMC and ABC data at their finest levels of detail.

| | Daily | Weekly | Monthly | Infrequent | Percentage | (miles) |
|------------|-------|--------|---------|------------|------------|--------------------|
| Short Haul | 26 | 227 | 496 | 3584 | 18.2% | distance < 250 |
| Regional | 10 | 130 | 374 | 2738 | 12.9% | 250<distance ≤500 |
| Mid Range | 3 | 104 | 564 | 5561 | 24.6% | 500<distance ≤1000 |
| Long Haul | 1 | 81 | 746 | 1039 | 44.3% | distance > 1000 |
| Percentage | 0.2% | 2.1% | 8.6% | 89.1% | | |
| (lds/yr) | ≥250 | ≥52 | ≥12 | <12 | | |

Table 5.2 Origin-destination flow for MTMC data.

| | Daily | Weekly | Monthly | Infrequent | Percentage | (miles) |
|------------|-------|--------|---------|------------|------------|--------------------|
| Short Haul | 11 | 66 | 170 | 290 | 30.8% | distance < 250 |
| Regional | 15 | 51 | 169 | 248 | 27.7% | 250<distance ≤500 |
| Mid Range | 10 | 60 | 177 | 288 | 30.7% | 500<distance ≤1000 |
| Long Haul | 1 | 18 | 90 | 79 | 10.8% | distance > 1000 |
| Percentage | 2.1% | 11.2% | 34.8% | 51.9% | | |
| (lds/yr) | ≥250 | ≥52 | ≥12 | <12 | | |

Table 5.3 Origin-destination flow for ABC data.

The tables show that there are relatively few OD pairs, in either network, which justify a point to point lane. However, the majority of the traffic volume moves over these few high volume OD flows. The vast majority of the OD flows in both networks are infrequent short-hauls. Figure 5.5 shows this distribution for ABC data.

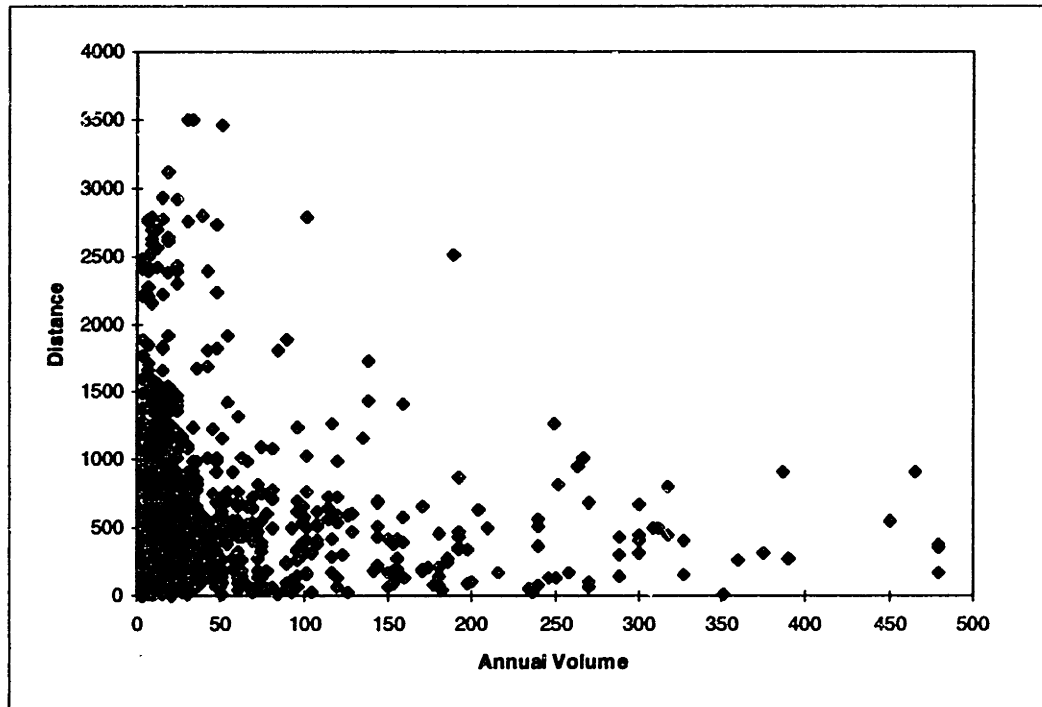


Figure 5.5 Plot of OD annual volume (loads/year) versus distance (miles) for 5-digit ZIP codes for the ABC data

Classifying the shipping points and their respective OD flows allows the shipper to identify those high volume lanes which should be treated differently from the other lanes. The next step looks specifically at how to select which shipping points to aggregate into zones and which to treat as individual shipping points.

5.2.3 Identification of Zones

Once the shipping points and the OD flows have been examined, the zones can be identified. Because the objective is to make this network as "carrier-friendly" as possible, the zones should be designed to promote the use of conditional bids.

This step determines which shipping locations to aggregate into zones. A zone can be defined by standard boundaries (e.g., SPLCs or ZIP Codes) or uniquely by the shipper. Standard boundaries are well understood by carriers and comparable across shippers. Additionally, both SPLC and ZIP codes can be aggregated simply by reducing the number of digits in the code. Shipper specific zones can be formed using clustering techniques which consider the inclusion of each point into a zone using some criterion, such as minimum expected

empty miles per zone. These are not comparable across shippers, however. Additionally, the same algorithms can be used with the standard zones with the only exception being that the shipping locations need to be treated as sets rather than individual points.

This analysis used standard boundaries to define the zones. Three levels of aggregation are available in the MTMC data set: *counties* (4-digit SPLC), *states*⁷⁸ (2-digit SPLC), and 8 *regions* (1-digit SPLC). Likewise, in the ABC data, the zones can roll up into 3 and 2 digit ZIP codes. Lanes are identified based on whether the origin and destination are points or zones.

Zones can be aggregated in many different ways depending on the priorities of the shipper and the characteristics of the network. Because we are interested in identifying as many matching opportunities as possible, the method used for the MTMC and ABC data emphasized the use of points as much as possible. We refer to this as the volume approach because, essentially, the lanes are identified based on whether they reach a minimum threshold of volume. If a single point to point pair generates a sufficient volume, then it is treated as a lane by itself. In order to increase the volume on a lane, either the origin or the destination, or both, can be expanded into zones consisting of multiple shipping points. This method finds the lower bound on the size of the zones. That is, zones are only as big as is necessary to pass some threshold value. The method is described below.

Volume Threshold Method

Identify Point to Point Lanes

Select all OD flows with at least daily level (>250 loads/year) as being point to point lanes. Remove these OD flows from consideration.

Identify the Point to Zone (Zone to Point) Lanes

Scan for all shipping points with at least 1,000 outbound loads per year.

For each shipping point, *i*, determine if it is predominately inbound or outbound

If outbound, identify those destinations which can be served by a common zone.

Create a point to zone lane if volume is at least twice weekly (100 loads/year).

If inbound, identify those destinations which can be served by a common zone.

Create a zone to point lane if volume is at least twice weekly (100 loads/year).

Remove these OD flows from further consideration.

Identify the Zone to Zone Lanes

For any remaining OD flows not yet assigned, aggregate both origin and destination by zone until all geographic areas are covered.

This simple procedure was used to create a hierarchy of lanes. The high density lanes consisted of all point to point and any other lanes which had at least twice weekly flow at the county level (2nd level of aggregation for SPLCs). The remainder of the OD flows consisted of infrequent shipments and were identified as low density lanes. Table 5.4 shows the characteristics of each network.

⁷⁸ New York, Pennsylvania, Virginia, Kentucky, Michigan, Wisconsin, Indiana, Illinois, North Carolina, Tennessee, Georgia, Iowa, Missouri, Kansas, Arkansas, Oklahoma, Louisiana, Montana, Wyoming, Colorado, New Mexico, and California are divided into two sections. Alaska is divided into three and Texas into four. There are a total of 89 2-digit SPLC regions.

| | Total System | High Density | Low Density |
|--------------------------------|-------------------|------------------------------|---------------------|
| Nodes | 4,925 | 213 (45 point, 178 zones) | ~200 (all zones) |
| Lanes | 25,000 | 265 | ~1600 |
| Loads (percentage of total) | 191,000 (100%) | 53,000 (28%) | ~140,000 (62%) |

Table 5.4. Summary of high and low density MTMC networks.

The high density network consists of those lanes and shipping locations that have potential for forming efficient aggregations. This analysis showed that approximately 30% of the traffic travels on just slightly over 1% of the OD flows.

5.3 Step 3: Assessing Matching Opportunities

In order to determine if there are any opportunities to match loads in a shipper's network, a detailed analysis should investigate the network topology and the lane distribution of traffic over time.

5.3.1 Network Topology

The topology of the network refers to its structure. The critical factor is the degree of the nodes defined as the total number of lanes connected to that node. The indegree of a node is the number of lanes entering while the outdegree is the number of lanes exiting a node. Nodes where both indegree and outdegree are greater than zero are potential candidates for matching inbound and outbound lanes. The nodes can be points or zones. Only the higher-density lanes are examined since they are the most reliable lanes.

Some networks have no potential for the matching of inbound and outbound traffic lanes even if they have large total volumes. For example, a certain mini-mill's distribution network consists of five primary traffic lanes all emanating from a single facility in Oregon. The mill serves five markets (British Columbia/Washington, Eastern USA, eleven Western States, California, and Oregon) exclusively by flatbed TL trucking. Each of the markets, though, is independent and does not interact with the others. All inbound deliveries of raw steel are by rail. Because their distribution network is radial there are no matching opportunities.

Figure 5.6 shows the distribution of inbound to outbound lanes for the high density MTMC network. Just under half of the nodes have both inbound and outbound lanes. These are potentials for matching opportunities. For example, the location with close to 4,000 loads per year has approximately 40% outbound lanes and 60% inbound.

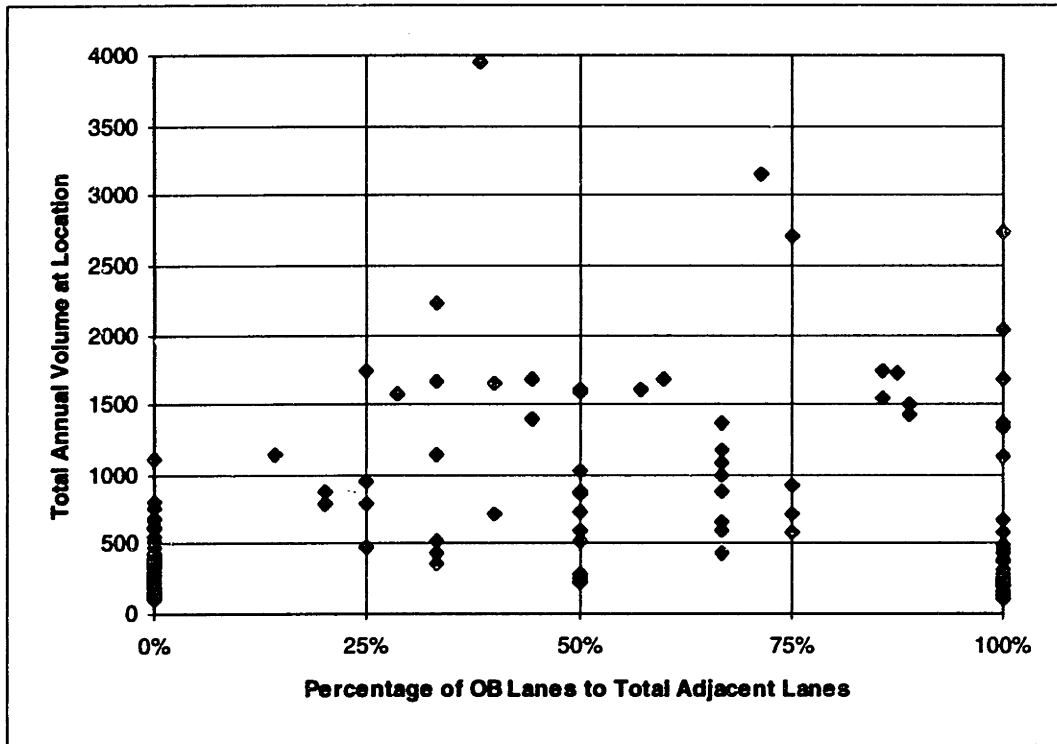


Figure 5.6 The relative number of IB to OB lanes for the high-density network for MTMC data.

5.3.2 Flow Distribution

The distribution of the flow refers to the variability and seasonality of the traffic on a lane and at a location. For a set of lanes to be good candidates for efficient aggregations, they must share certain characteristics. In addition to lane distance, total volume, and proximity for connections, the distributions must be compatible. While an inbound and an outbound lane having similarly high volumes might appear to be good candidates for matching, if they exhibit different seasonalities (say one peaks in winter and the other in the summer) then they are not compatible.

This section first discusses the seasonality of freight movements and how it affects both the total volume levels and the balance at locations. Second, the traffic flow on a lane is shown to often times follow a Poisson distribution. Finally, the probability of matching at a location, given the estimated traffic probabilities is calculated.

Seasonality

Seasonality of traffic flows is the variation in traffic volumes which exhibits a predictable pattern. This pattern can be repeated over the course of a year, days of the week, or even time of day. Seasonality can affect both the volume in a lane or at a location as well as the balance (inbound versus outbound loads) at a node.

The seasonality of the flows across the network can be critical. Reviewing historical data allows for identification of potential peaks in the shipments. While reliance on historical data is not perfect, it does give some insight into true seasonality.

Annual seasonality can be caused by artificial, as well as, natural causes. The peak of agriculture movements during the late months of the summer are due to the growing season. Other industries have more artificial peaks, such as, Halloween for candy manufacturers, the Back-To-School-Season for office product companies, and Christmas for most retailers and manufacturers.

When the seasonalities are severe, a shipper might want to have a stable of carriers to rely upon. In order to have sufficient capacity at these peaks, however, they might need to allocate loads to carriers throughout the year. This would reflect in the assignment of more carriers per lane, or in the system as a whole, than would normally be justified.

Likewise, there is seasonality across days of the week in most companies. Shipping on weekends is rare and there is an incentive to clear the docks by the end of the week. This can create a higher demand for vehicles on Fridays. Additionally, shipping on Friday can give the shipper two 'free' days, in terms of transit time since the carrier can use the weekend to make the delivery by the next working day.

The monthly pattern for the MTMC data across the high density lanes is shown in figure 5.7. The monthly shipping volume was depressed in the winter months - probably due to a slow down in military training activities.

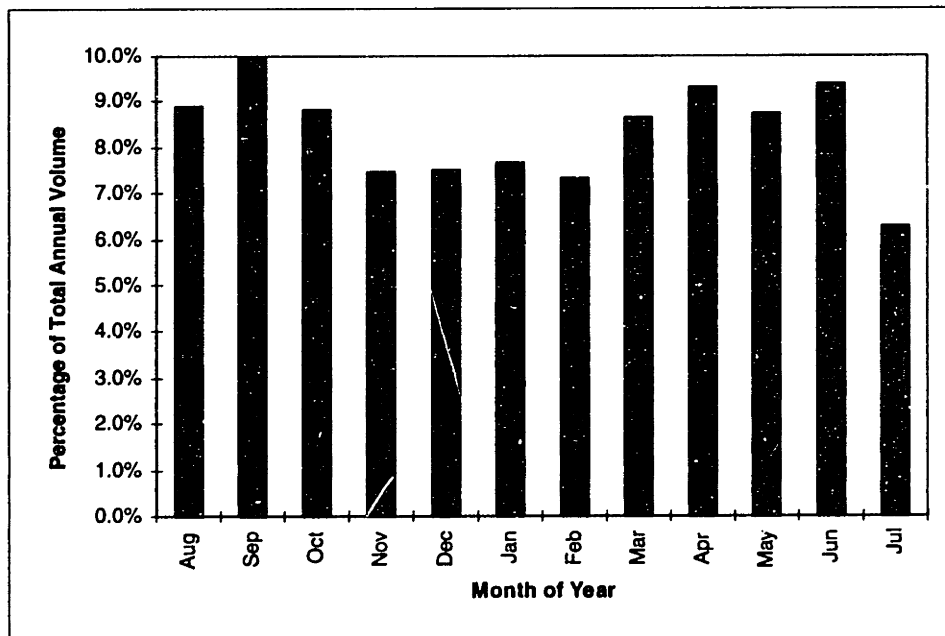


Figure 5.7 Seasonality of traffic across high density lanes for MTMC.

The day of the week patterns across the entire MTMC system are shown in Figure 5.8. As expected, there is very little activity on the weekends. Other than this, however, there did not appear to be any great seasonalities based on the day of the week, for either volume or balance of flow.

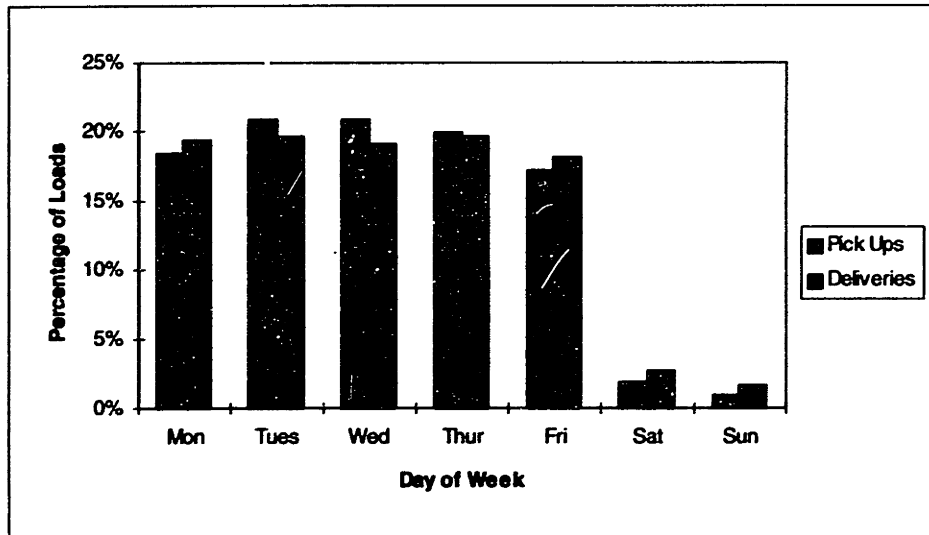


Figure 5.8 Variation of day of the week across the system.

Analysis of the system as a whole is interesting, but individual lane and location seasonality is more important for carriers. To illustrate the importance, we will look briefly at a single shipping location, Stockton Yards in California.

Stockton Yards handled over 3,800 shipments in the year covered. The loads were evenly balanced between inbound and outbound traffic. Figure 5.9 shows the monthly variation of the flow through the yard. Notice that not only do the shipping volumes change month by month - closely echoing the system-wide pattern - the balance between inbound and outbound also changes. The ratio of IB to OB imbalance ranges from 5:3 to 4:5. Such wide swings make it difficult to estimate probabilities of a match.

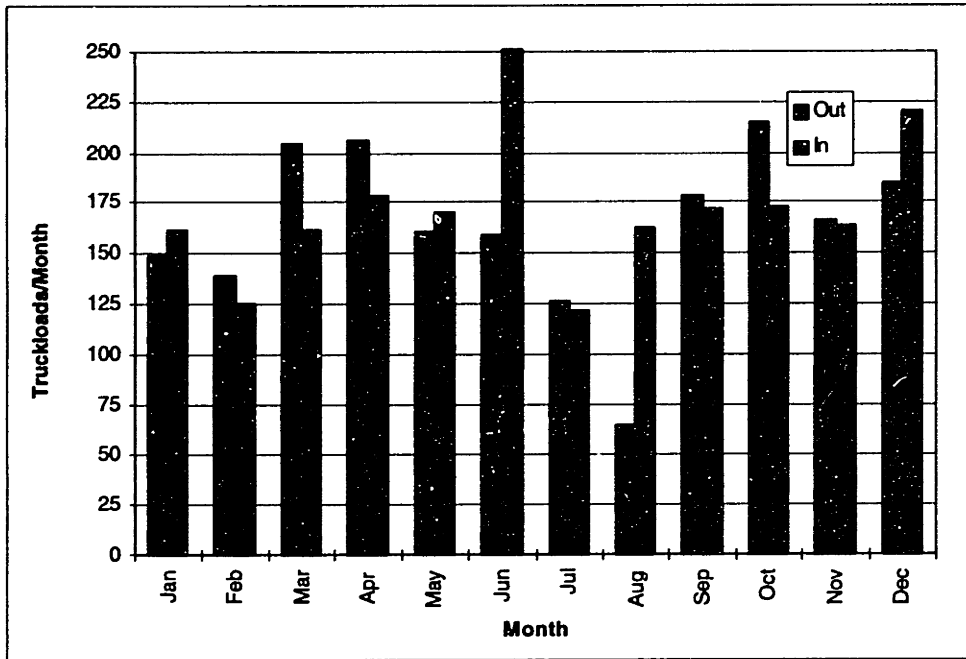


Figure 5.9 Distribution of Inbound and outbound lanes at Stockton Yards, CA

The weekly inbound/outbound distribution is shown in Figure 5.10. Again, there are wide swings in both the balance and the total volume of traffic. The ratio of inbound to outbound on a weekly basis changed from 10:1 in mid-August (week 33) to 1:2 in October (week 42). Except for the trough in August, however, there appears to be a minimum threshold level of 20 loads per week for both inbound and outbound traffic.

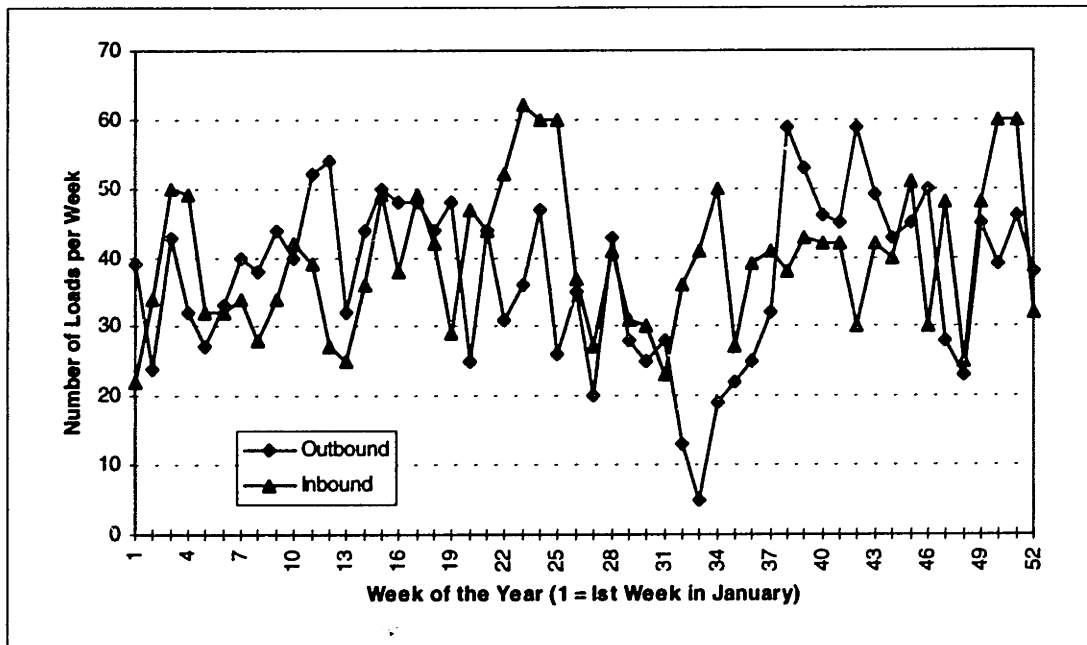


Figure 5.10 Weekly distribution of inbound and outbound traffic at Stockton Yards, CA

The average number of shipments per day of the week are shown in Figure 5.11. Again, weekend traffic is substantially lower than the other days of the week. There is also a slight rise in activity at the beginning of the week - especially for the inbound traffic.

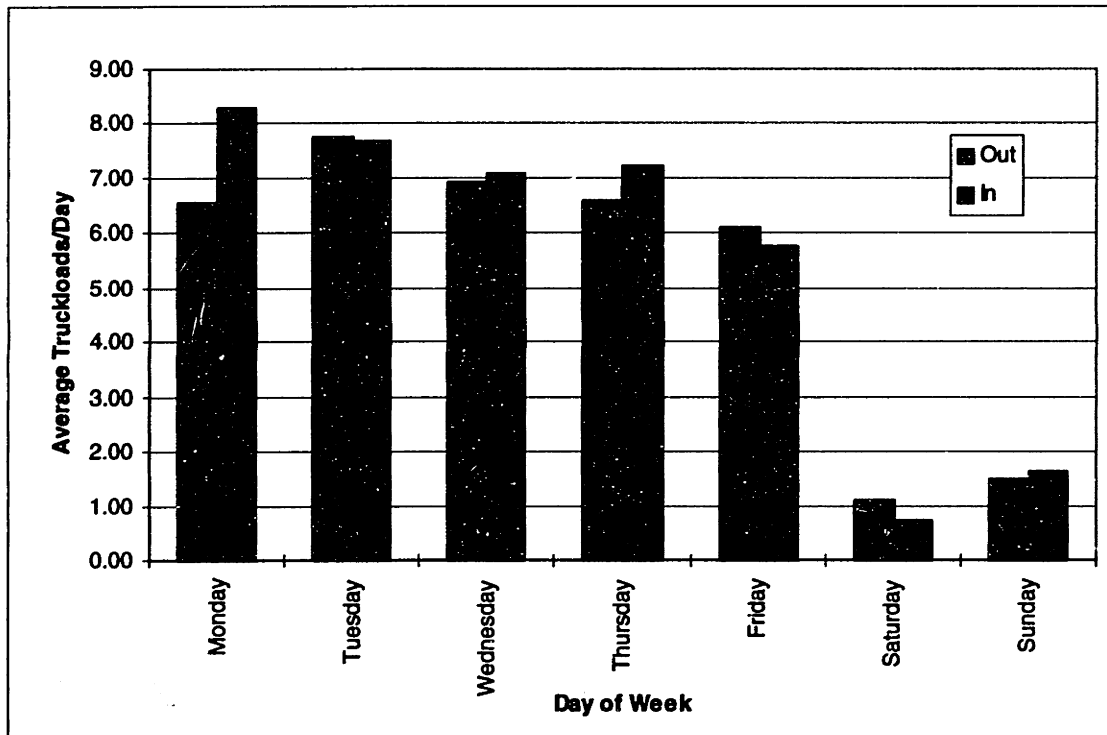


Figure 5.11 Day of the week distribution of inbound and outbound traffic at Stockton Yards, CA

The seasonality of the traffic is important for the carrier in order to plan capacity. The shipper can pass on information about seasonality to the carriers in a number of ways. One way is to provide the carriers with historical, or estimated, traffic volumes over each week of the year. Another is to provide a minimum and maximum volume on each lane over the course of the year. For example, a minimum weekly value of 20 reloads per week could be applied to the Stockton Yard example.

Lane Flow Distribution

The volume in the system and on individual lanes is important for the carrier in order to perform capacity planning for trucks, trailers, and drivers. An additional concern, critical for the forming of efficient aggregations, is the probability of a carrier connecting an inbound load to an outbound load. Without a high probability of this occurring, the economies of scope are not achievable.

The first step in determining the probability of a connection, either at a point or within a zone, is to characterize the daily flow on a lane. While knowing the average number of inbound and outbound loads is helpful, a better characterization is if the traffic could be fit to a known

probability distribution. The Poisson distribution suggests itself as being appropriate since it is discrete, takes on only non-negative values, and allows for observations of zero value.

In order to check the validity of this assumption, lane traffic on a sample of lanes within the MTMC data were fitted. The number of loads generated per each weekday over the year were used as data. The weekend shipments were not included since the traffic on weekend days did not seem to be independent of the day of week. This resulted in a 250 day year. All days which had no recorded shipments were assumed to have zero. This is subject to some data entry distortions since there were some incomplete records.

The Poisson density function is defined as:

$$\text{Probability [} k \text{ arrivals in time } t] = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

where:

- λ : Average number of arrivals (loads tendered) per unit time,
- t: Time period, and
- k: Number of arrivals of interest.

The value of the parameter λ for each traffic lane was estimated using maximum likelihood methods.⁷⁹ The rate was assumed to be in loads per day. The fit of traffic to the Poisson distribution was generally quite good. The majority of the traffic lanes achieved a goodness of fit⁸⁰ above 0.90. The lanes which did not fit the distribution, however, were generally very bad. Table 5.5 shows the estimated values for five point to point lanes (3 inbound and 2 outbound) which are adjacent to Virginia Beach, Virginia.

| Lane | Ann. Vol. | Est. λ | t-stat | L(B) | L(0) | ρ^2 |
|------|-----------|----------------|--------|--------|------|----------|
| IB1 | 788 | 2.63 | 27.18 | -27.01 | -250 | 0.89 |
| IB2 | 680 | 2.72 | 26.09 | -0.04 | -250 | 0.99 |
| IB3 | 203 | 0.81 | 2.96 | -245 | -250 | 0.02 |
| OB1 | 221 | 0.89 | 1.83 | -2.48 | -250 | 0.99 |
| OB2 | 390 | 1.46 | 7.28 | -226 | -250 | 0.10 |

Table 5.5 Results of portion of estimations for lane probabilities.

The lanes which did not fit well appeared to have many days with no arrivals and a few large spikes of volume throughout the year. These lanes might need to be screened out and treated differently - perhaps deterministically. As a first approximation, however, the Poisson distribution seems to fit the majority of the high volume lanes adequately.

⁷⁹ See Sheffi (1985, 62) for a description of using maximum likelihood for the Poisson distribution.

⁸⁰ The goodness-of-fit index used measures the fraction of an initial log-likelihood value that is explained by the model. See, Ben Akiva and Lerman (1985, 91).

Probability of a Load Matching

After estimating the parameters for the different lanes, we can now determine the probability of matching of an inbound to an outbound load. Consider the case of a node with a number of inbound lanes and a number of outbound lanes. Assume the arrivals on each of the inbound lanes and departures on the outbound lanes are independent Poisson processes with rates $\lambda_{IB1}, \lambda_{IB2}, \dots, \lambda_{IBi}$ and $\lambda_{OB1}, \lambda_{OB2}, \dots, \lambda_{OBj}$ respectively. The combined inbound rate is $\lambda_{IB} = \sum \lambda_{IBi}$ and the combined outbound rate is $\lambda_{OB} = \sum \lambda_{OBj}$.

We are concerned with two probabilities. First, given that an inbound load just occurred, what is the probability that an outbound load will occur within the next t time periods? Second, given that an outbound load just occurred, what is the probability that an inbound load arrived within the previous t time periods?

The reason for separating these two probabilities is that the carrier is concerned with connecting to a follow-on load only if the previous load existed. The first probability is applicable to locations or zones with more inbound than outbound lanes while the second is for locations with more outbound than inbound lanes. The probability of the less likely event is being determined.

The probability that an inbound vehicle will be matched to an outbound load within a maximum acceptable dwell time, t_M , is simply the inter-arrival time between outbound departures. If we assume that the rates are Poisson, the inter-arrival times are distributed exponentially. Also, because the exponential distribution is memoryless, we do not have to factor in the driving time of the previous trip, t_{drive} . The probability that an event occurs between 0 and t_M minutes is the same as the probability that an event occurs between $t_{drive}+0$ and $t_{drive}+t_M$. Finally, the memory property also lets us use a single rate for the traffic on a lane rather than using separate rates for arrivals and departures.

Thus, the probability that a matching occurs within t_M time units is given by:

$$\text{Prob[OB occurs within } t \mid \text{IB occurred]} = \int_{t=0}^{t=t_M} \lambda_{OB} e^{-\lambda_{OB}t} dt$$

which reduces to:

$$\text{Prob[OB occurs within } t_M \mid \text{IB]} = (1 - e^{-\lambda_{OB}t_M})$$

where,

- λ_{IB} : Rate of loads on the inbound lane,
- λ_{OB} : Rate of loads on the outbound lane, and
- t_M : Maximum acceptable dwell time for the connection.

Likewise, the probability for the second case is the same expression except using the inbound arrival rate, λ_{IB} .

The probability of a match can be determined at the aggregated level or for any combination of inbound and outbound lanes. In the Virginia Beach example, the probability that any inbound is matched to any outbound (since there are close to 3 inbound lanes to each outbound) is equal to $(1 - \exp[-(0.88 + 1.46)(0.5)]) = 0.69$. Thus, there is a 70% probability that an outbound load can be matched to an inbound load within half of a day ($t_{1/2} = 0.5$ days).

There are several practical limitations to using this simple model. First, the data generally do not exist. The arrival times and the release times for departures are required. While the actual arrival *dates* are recorded in freight records, generally the actual time is not. Similarly, while the departure date is recorded, the time that the load becomes available is not usually recorded. The data required are not just the release date for the shipment but both the earliest possible and the latest release time windows. The IB shipment can be matched to an OB shipment's earliest release time or the OB can be delayed for an incoming IB.

Second, the assumption that the arrivals and departures are distributed in a Poisson fashion may not hold. The release of the shipments as well as the different arrivals may not be independent of each other. Also, on some of the sparser lanes, the Poisson distribution does not seem to fit that well.

Third, there is great seasonality in traffic distributions which makes the stability of the distribution quite unpredictable. The rate will only apply for certain time periods such as summer or peak/off-peak.

Fourth, the upcoming year's distribution may change from the detailed distribution for the historical period (last year) used to fit the data. The changes in distribution patterns may also change the split between the IB and OB shipments.

Even with these limitations, this type of analysis is beneficial in exploring the potential for matchings. The local plant managers also have some influence on encouraging IB and OB matchings in that they can hold certain OB shipments (i.e., time consolidation) so that they can be matched to a slightly later IB shipment. Thus, this simple model is a conservative estimate of the probability of a matching.

5.4 Step 4: Identifying Potential Efficient Aggregations

The lanes within a distribution network can be combined in different ways to form efficient aggregations. An efficient aggregation is a bundle of lanes which, if serviced by a single carrier, has potential to achieve operational cost efficiencies. The shipper cannot identify all potential aggregations *a priori* since carriers have different service networks and lane

combinations have different values. However, the shipper can at least identify some aggregations that are fully contained within the network.

This section discusses four forms of lane bundling: reloads, open tours, closed tours, and local tours. The conditions for each is described, algorithms for identifying the opportunities within the network are presented, and potential savings are estimated.

5.4.1 Reloads

A reload is the matching of an inbound trip to a shipping location with an outbound trip from that same location. Because the origin and destination are the same, there are no empty miles associated with a reload. This is its primary attraction. In practice, this consists of a truck unloading at one dock and picking up another load at another dock door at the same, or adjacent, facility. It could also occur at a trailer pool operation where a load is dropped into the pool and a follow-on load is picked up immediately from the pool.

While the empty miles for a reload are essentially zero, the dwell time may not be. The driver, for example, may have to wait for the trailer to be unloaded and then for the new load to be readied before departing. Pool operations reduces the dwell time as well as the empty miles since the outgoing trailer should be waiting for pickup at the time of arrival.

Reload opportunities can consist of a single inbound and a single outbound lane or of a combination of inbound and outbound lanes. For example, if a plant with a total of 3 inbound and 10 outbound lanes could bundle together all of the inbounds with a portion of the outbounds. The condition could be, if the carrier handles at least, say, 80% of the outbound lanes, they will get all of the inbound traffic.

Proposed Algorithm for Reload Identification

Identifying potential reload opportunities is straightforward. The algorithm identifies pair-wise matchings between inbound and outbound lanes at selected high volume nodes and only considers lanes which originate or end at a point, as opposed to a zone. The objective is to locate *opportunities* for reloads - not all possible combinations of lanes. The carriers will have the opportunity for creating their own sets of lanes if they so desire. Those lanes which do not, by themselves, meet the minimum threshold level for being a stand alone lane are aggregated together and are included as a group in the proposed reload opportunities list.

Notation:

- P: Set of all high volume shipping locations where $V_i > \text{Threshold}$ for $i \in P$,
- L: Set of all lanes $(ij) \in L: i \in P$ or $j \in P$,
- V_i : Total shipping volume (inbound plus outbound) handled at node i ,
- V_{MIN} : Minimum acceptable volume on lane to be included as reload,

v_i : Volume on arc from node i to node j ,
 d_i : Distance of the arc from node i to node j ,
 OB_i, IB_i : Set of lanes to be considered for reload from/to node i ,
 OB'_i, IB'_i : Set of lower volume lanes to be considered for reload from/to node i , and
 $RELOAD_i$: Set of paired lanes for potential reloads.

Algorithm

Algorithm: FIND_RELOADS

Step 1 (Initialization): Create reload network consisting of all of the high volume shipping locations, P , and all lanes which either originate or end at a high volume shipping location.

Step 2 (Search Nodes): Select a node, i , in P .

Scan all inbound arcs (ji) . If $v_j \geq V_{MIN}$, place (ji) into IB_i . Otherwise, place (ji) into IB'_i .

Scan all outbound arcs (ij) . If $v_i \geq V_{MIN}$, place (ij) into OB_i . Otherwise, place (ij) into OB'_i .

Step 3 (Identify Reloads): Place each pair of lanes in IB_i with lanes in OB_i into the set of reloads $RELOAD_i$. If total volume from $IB'_i \geq V_{MIN}$, pair it with lanes in OB_i and place it in $RELOAD_i$. If total volume from $OB'_i \geq V_{MIN}$, pair it with lanes in IB_i and place it in $RELOAD_i$.

Step 4 (Identify Reloads): Repeat step 2 until all nodes in P have been searched.

Potential Savings

The savings created by a reload is both carrier and location specific. The value of the reload is due to the elimination of deadheading to find a follow-on load and hopeful reduction in the dwell time. In order to estimate the savings, the expected empty miles needed to find a follow-on load and the expected dwell time need to be compared against the expected values under the reload conditions.

The reload opportunities do not guarantee that there will always be a reload. It just increases the probability. The expected cost of finding a connection at a location i without a reload opportunity is the expected dwell time and the expected deadhead miles multiplied by their respective cost coefficients. It is given by:

$$C_S = c_d E[d_i] + c_e E[e_i]$$

where:

C_s : Cost of making a connection for a follow-on load with a single lane at node i ,
 c_d : Cost per hour for dwell time,
 c_e : Cost per mile for empty miles,
 $E[d_i]$: Expected dwell time at location i , and
 $E[e_i]$: Expected empty miles required to connect to follow-on load from i .

The expected cost for a connection if both lanes are won is the sum of the original connection cost times the probability of not making the connection and the expected dwell time costs multiplied by the probability of making the reload connection. It is given by:

$$C_{Ri} = (1 - \text{Prob}[\text{Reload}_i]) (C_{Si}) + (\text{Prob}[\text{Reload}_i]) (c_d E[d_i | \text{Reload}])$$

Assuming Poisson arrivals, the probability of a reload within a certain maximum dwell time, t_{MAX} is:

$$\text{Prob}[\text{Reload}_i] = \int_0^{t_M} \lambda e^{-\lambda t} dt = 1 - e^{-\lambda t_M}$$

where:

- λ : Rate of outbound loads from location i (loads/hour) and
- t_M : Maximum acceptable dwell time for carrier at the location.

The expected dwell time conditioned on a reload occurring within the maximum acceptable dwell time, is:

$$E[d_i | \text{Reload}] = \int_0^{t_M} t \lambda e^{-\lambda t} dt = \frac{1}{\lambda} - \left[e^{-\lambda t_M} \left(t_M + \frac{1}{\lambda} \right) \right]$$

So that the connection cost for a reload at location i, C_{Ri} is:

$$C_{Ri} = (e^{-\lambda t_M}) C_{Si} + (c_d) \left(1 - e^{-\lambda t_M} \right) \left(\frac{1}{\lambda} - \left[e^{-\lambda t_M} \left(t_M + \frac{1}{\lambda} \right) \right] \right)$$

The savings for a reload at a node is the difference between C_R and C_S which reduces to:

$$\begin{aligned} \text{Savings} &= C_{Ri} - C_{Si} \\ &= (1 - e^{-\lambda t_M}) \left(c_e (E[e_i] - 0) + c_d \left(E[d_i] - \frac{1}{\lambda} + \left[e^{-\lambda t_M} \left(t_M + \frac{1}{\lambda} \right) \right] \right) \right) \end{aligned}$$

As the load rate on the outbound lane increases, the savings will increase since the probability of a match increases.

A simple example will help demonstrate. At a given location the expected deadhead is 50 miles and average dwell time is 4 hours. The cost of each empty mile is \$1.00 and an hour of dwell time costs \$10.00. The connection cost for the lane alone is $(50)(1) + (4)(10) = \$90$. Suppose that the rate of outbound loads from the location, λ , is 0.5 loads/hour and the maximum acceptable dwell time, t_M , is 2 hours. The savings due to a reload becomes $(0.67)[(1\$)(50 - 0) + (\$10)(4 - 2 + (.33)(2 + 2))] = \56 per inbound load. If the connection cost of the load represents about 15% of the rate of the incoming load, then the savings due to a reload is about 10% of the total lane rate.

In practice, the rule of thumb is that a reload can save about 5% on a load. Another way of looking at the savings is not as a percentage of the inbound or outbound lanes, which are related to the distance, but as a separate rebate. So, if a shipper is able to provide the carrier with a reload, a deduction of, using the numbers from the previous example, half of the savings,

$[\$90 - (10 \text{ \$/hr})(2 \text{ hrs})]/2 = \35 , is applied to the freight bill. This makes auditing more difficult but does separate the reload discount from the different lanes.

5.4.2 Open Tours / Threads

An open tour, or thread, is a collection of lanes together which begin and end in different locations. Unlike a reload, a thread may include more than two lanes, contain “deadhead” legs, and use zone to zone type lanes.

In practice, the linking of lanes in real-time is referred to as a continuous move. At this stage, however, the objective is to identify potential lanes in the network which might lend themselves to continuous moves in the tendering. As with reloads, the objective is to identify potential opportunities for high probability of connection, not a guarantee.

Proposed Algorithm for Thread Identification

The algorithm finds threads by first adding deadhead arcs between all points in the network. These arcs have a known distance. The algorithm works by performing a depth first search in a tree consisting of all deadhead and actual arcs. If a candidate thread is identified, it is sent to the list of potential threads and a validity check is performed. The algorithm stops after each node has been the root node.

Notation:

| | |
|-----------------------|--|
| P: | Set of all high volume shipping locations (points), |
| Z: | Set of all zones of shipping locations, |
| L: | Set of all loaded lanes, |
| D: | Set of all deadhead lanes, |
| d_{ij} : | Distance of arc from node i to node j, |
| v_{ij} : | Volume on arc from node i to node j, |
| T_{MAX}, T_{MIN} : | Maximum and Minimum number of lanes in thread, |
| d_{MAX}, d_{MIN} : | Maximum and Minimum length of each lane in thread, |
| v_{MIN} : | Minimum acceptable volume on lane in thread, |
| e_{MAX} : | Maximum percentage of empty miles allowed in thread, and |
| THREAD _i : | Set of lanes for the i th potential thread. |

Algorithm

Algorithm: FIND_THREADS

Step 0 (Initialization): Create augmented network $N(P \cup Z, L \cup D)$ with the nodes consisting of both point and zone locations and loaded and deadhead arcs.

(Screen Out Unacceptable Arcs): Scan all loaded arcs, (ij) in network. If $v_{ij} < v_{MIN}$, $d_{MAX} < d_{ij}$ or $d_{MIN} > d_{ij}$ then remove (ij) from the network. Stop if no arcs remain in the network.

(Initialize Search Tree): Select a root node, $i \in (P \cup Z)$, set depth=0, active_node = root, and active_arc = NULL.

Step 1 (MAIN_LOOP):
 Go to SELECT_ARC.
 If active_arc = NULL go to BACKTRACK,
 Otherwise depth = depth + 1,
 active_node = destination of active_arc
 Add active_arc to active_thread.
 Go to CHECK_THREAD.
 If (active_node = Root or depth = T_{MAX}) then goto BACKTRACK.
 }

Step 2 (SELECT_ARC): Select first unmarked outbound arc from active_node as the new active_arc. Mark it as visited.
 If active_thread[depth-1] is a deadhead ignore all deadhead arcs.
 If no unmarked arcs from active_node exist, set active_arc = NULL.

Step 4 (BACKTRACK):
 If active_node = Root then select a new root, set depth = 0
 Otherwise, set active_node = origin of active_arc.

Step 5 (CHECK_THREAD):
 Calculate percentage of empty miles, E, in active_thread.
 If ($E < e_{MAX}$, $T_{MIN} < \text{Depth} < T_{MAX}$, and active_arc \neq deadhead)
 then add active_thread to THREAD,

Example

Suppose a network consist of 4 nodes and 4 loaded arcs as shown in Figure 5.12a. The algorithm first adds in deadhead arcs (shown as dashed lines) between each node, as in Figure 5.12b.

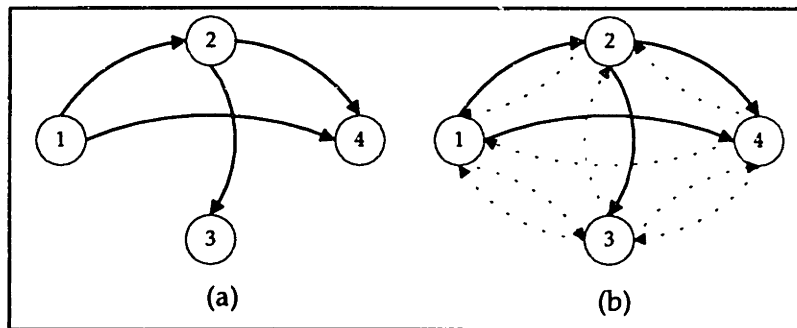


Figure 5.12. Example network for thread identification.

For each root node selected, the algorithm creates a tree of all possible threads, shown in Figure 5.13 for node 1 as the root. A depth first search in this tree identifies the threads which are acceptable. The result of the search in this case would be the following 3 threads:

- {(1,2)-(2,3)}
- {(1,2)-(2,4)}
- {(1,4)-(4,2)-(2,3)}

Thread 2 contains the deadhead arc (4,2). The tree is pruned wherever there would be two consecutive deadhead arcs or when a node in the thread is revisited. Also, no thread is kept if the last arc is a deadhead.

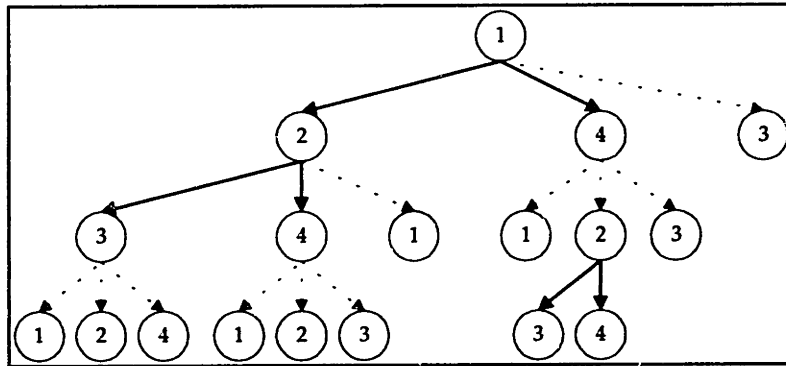


Figure 5.13 Search tree for node 1 as root node.

Potential Savings

The method of estimating the savings due to a thread is the same as for reloads with the exception that there are now empty miles included in either connecting deadhead lanes or in lanes based on zones rather than points.

A reload is essentially just a two lane thread with a point as the middle location. The probability of a thread being completed requires that each connection be made. For example, for the four lane thread shown in Figure 5.14, three connections need to be made in order to complete the thread.

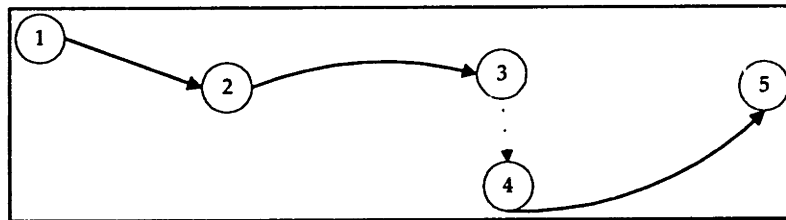


Figure 5.14 Example of four lane thread with one deadhead leg.

If we again start with the idea that the first load has occurred, then we are interested in finding the probability that the connections are made at nodes 2 and 4. The probability of making a connection at node 2 is the same as a reload:

$$\text{Prob}[\text{Load from 2 occurs within } t_M \mid \text{Load from 1}] = (1 - e^{-\lambda_2 t_M})$$

After the load is accepted, the truck would travel from 2 to 3, (t_{23}), unload ($t_{\text{DWELL}03}$), and then drive empty to node 4 (t_{34}). Due to the memoryless property of the Poisson distribution, however, we do not need to determine the probability that a load is available to leave node 4 at

$(t_{23} + t_{\text{DWEELL03}} + t_{34})$. Instead, we need only to find the probability of making that specific reload, which is:

$$(1 - e^{-\lambda_i t_M})$$

The probability for completing a thread becomes:

$$\text{Prob}[\text{Completing Thread with } i \text{ loaded lanes}] = \prod_{i=1}^I (1 - e^{-\lambda_i t_M})$$

The savings can be calculated in the same manner as for the reloads where the cost of connecting at each node j within a thread is:

$$\text{Savings}_j = (1 - e^{-\lambda_j t_M}) \left(c_e (E[e_j] - e_{IB_j}) + c_d \left(E[d_j] - \frac{1}{\lambda_j} + \left[e^{-\lambda_j t_M} \left(t_M + \frac{1}{\lambda_j} \right) \right] \right) \right)$$

where:

- c_d : Cost per hour for dwell time,
- c_e : Cost per mile for empty miles,
- $E[e_j]$: Expected empty miles at node j without thread option,
- $E[d_j]$: Expected empty miles at node j without thread option,
- e_{IB_j} : Empty miles required on inbound lane to node j , and
- λ_j : Rate of outbound lanes from node j .

The first term is the probability of missing the connection at j multiplied by the expected cost of making a different connection at node j . The second term consists of the costs of making the thread connection, in terms of both expected dwell time and the deadhead mileage, multiplied by the probability of making the thread connection at node j .

The total cost of connection for a thread should be compared to the expected costs of connecting at each location separately. The total savings is just the sum of the savings at each node for all nodes within the thread.

5.4.3 Closed Tours

A closed tour is a combination of lanes together which begin and end at the same location. By being self contained, the closed tours essentially remove all of the uncertainty of finding follow-on loads for the carrier. That is, if the connections are guaranteed, then the carrier does not need to include any connection costs. Closed tours are typically given to the dedicated fleets where the shipper is paying for the empty miles. The volume of the loads on the lanes must be sufficiently large in order to keep a vehicle in continuous motion.

Proposed Algorithm for Closed Tour Identification

The algorithm FIND_CLOSED TOURS is almost identical to FIND_THREADS, except that after each iteration, the root node is removed from future consideration. Because the closed tours are cycles, the closed tour can begin at any point.

Notation:

| | |
|----------------------|--|
| P: | Set of all high volume shipping locations (points), |
| Z: | Set of all zones of shipping locations, |
| L: | Set of all loaded lanes, |
| D: | Set of all deadhead lanes, |
| d_{ij} : | Distance of arc from node i to node j , |
| v_{ij} : | Volume on arc from node i to node j , |
| T_{MAX}, T_{MIN} : | Maximum and Minimum number of lanes in tour, |
| d_{MAX}, d_{MIN} : | Maximum and Minimum length of each lane in tour, |
| v_{MIN} : | Minimum acceptable volume on lane in tour, |
| e_{MAX} : | Maximum percentage of empty miles allowed in tour, and |
| TOUR $_i$: | Set of lanes for the i^{th} potential thread. |

Algorithm

Algorithm: FIND_CLOSED TOURS

Step 0 (Initialization): Create augmented network $N(P \cup Z, L \cup D)$ with the nodes consisting of both point and zone locations and loaded and deadhead arcs.

(Screen Out Unacceptable Arcs): Scan all loaded arcs, (ij) in network. If $v_{ij} < v_{MIN}$, $d_{MAX} < d_{ij}$ or $d_{MIN} > d_{ij}$ then remove (ij) from the network. Stop if no arcs remain in the network.

(Initialize Search Tree): Select a root node, $i \in (P \cup Z)$, set depth=0, active_node = root, and active_arc = NULL.

Step 1 (MAIN_LOOP):

Go to SELECT_ARC.

If active_arc = NULL go to BACKTRACK,

Otherwise depth = depth + 1,

active_node = destination of active_arc

Add active_arc to active_tour.

Go to CHECK_TOUR.

If (active_node = Root or depth = T_{MAX}) then go to BACKTRACK.

}

Step 2 (SELECT_ARC): Select first unmarked outbound arc from active_node as the new active_arc. Mark it as visited.

If active_tour[depth-1] is a deadhead ignore all deadhead arcs.

If no unmarked arcs from active_node exist, set active_arc = NULL.

Step 4 (BACKTRACK):

If active_node = Root then {

select a new root,

set depth = 0, and

remove root

```

}
Otherwise, set active_node = origin of active_arc.

```

Step 5 (CHECK_THREAD):

Calculate percentage of empty miles, E , in active_tour.

If $(E < e_{MAX}$ and $T_{MIN} < \text{Depth} < T_{MAX})$ then add active_tour to TOUR,

Example

Suppose a network consist of 4 nodes and 4 loaded arcs as shown in Figure 5.15a. The algorithm first adds in deadhead arcs (shown as dashed lines) between each node, as in Figure 5.15b.

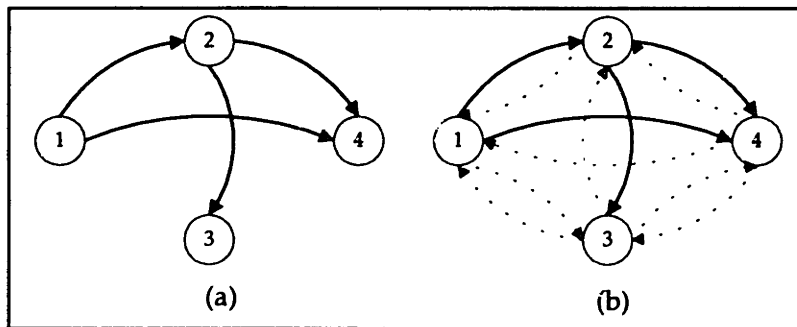


Figure 5.15 Example network for closed tour identification.

For each root node selected, the algorithm creates a tree of all possible tours, shown in Figure 5.16 for node 1 as the root. A depth first search in this tree identifies the closed tours which are acceptable. The result of the search in this case would be the following three closed tours:

$\{(1,2)-(2,3)-(3,1)\}$

$\{(1,2)-(2,4)-(4,1)\}$

$\{(1,4)-(4,2)-(2,3)-(3,1)\}$

The tree is pruned wherever there would be two consecutive deadhead arcs or when a non-root node in the tour is revisited.

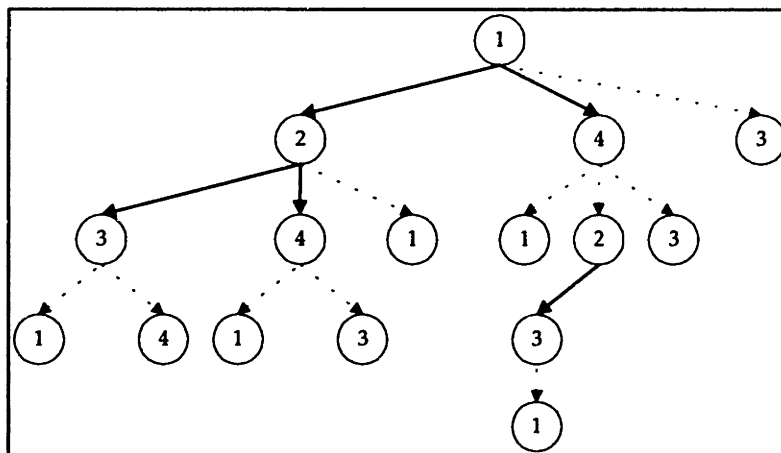


Figure 5.16 Search tree for node 1 as root node.

Potential Savings

If a closed tour is of sufficient volume, the vehicle can be kept operating continuously. Because the need for identifying connections for follow-on loads has been removed, the cost of the service is purely a function of the distance driven and the dwell time at each shipping location. The savings, then, is the difference between the current connection cost at each location within the closed tour and the cost of connection using the closed tour, which is zero. In practice, closed tours have been estimated to save between 10 and 15 percent of the total freight rate. This is close to the average empty miles of a large TL firm - so the savings appears to be based on the elimination of the empty miles.

5.4.4 Local Tours

Local tours are a series of short haul lanes with a common origin and/or destination. In practice this is where a plant or distribution center makes a large number of deliveries or pickups within half a day's drive. The length of each lane needs to be short enough so that a driver can return to the original location at the end of each day. Thus, these are trips which can be repeated daily, or are a series of lanes which a single vehicle can serve.

These occur quite frequently at distribution centers supplying nearby manufacturing facilities. While the volume on a single short haul lane might not be sufficient to justify a dedicated vehicle, the aggregated effect might be sufficient to keep several vehicles occupied.

The difference between this form of efficient aggregation and the other three forms of efficient aggregations is that the empty miles will almost always be over 50%. Short haul lanes, in general, are not highly desirable by TL carriers since the line haul is too short to recoup the lost time at the pick up and delivery. Also, a trip that is less than a day, or only a couple of hours, requires that multiple loads be tendered for these vehicles each day.

Thus, rates for short haul lanes are much higher on a per mile rate than long or even medium length tours. In the ABC data, for example, average cost per loaded mile is \$3.26 for lanes under 250 miles and \$1.52 for lanes between 250 and 500 miles. A large component of this cost is due to the added complexity of managing loads that turn over within one day, to include the fixed cost of the dispatch. Bundling together several short hauls into a local tour lowers this complexity cost since the need for finding multiple loads for each day is reduced. While the relative empty miles for each load in these tours still might exceed 50%, on an absolute scale each load is completed within a single day.

Proposed Algorithm for Local Tour Identification

The algorithm is similar to FIND_RELOADS. It searches each shipping point above a certain threshold volume. For each point, it identifies all inbound and outbound lanes with at

least a certain volume and a distance less than a threshold distance (250 miles). The output is a set of nodes with potential short haul opportunity and the lanes which support it.

Notation:

- P: Set of all high volume shipping locations where $V_i > \text{Threshold}$ for $i \in P$,
- L: Set of all lanes $(ij) \in L: i \in P \text{ or } j \in P$,
- V_i : Total shipping volume (inbound plus outbound) handled at node i ,
- V_{MIN} : Minimum acceptable volume in total local tour,
- v_{MIN} : Minimum acceptable volume on lane to be included in local tour,
- v_{ij} : Volume on arc from node i to node j ,
- d_{MAX} : Maximum distance of arc to be included in local tour,
- d_{ij} : Distance of the arc from node i to node j ,
- LTOUR: Set of paired lanes for potential reloads.

Algorithm

Algorithm: FIND_LOCAL TOURS

Step 1 (Initialization): Create short haul network consisting of all of the high volume shipping locations, P , and all lanes which either originate or end at a high volume shipping location and are less than 250 miles in length.

Step 2 (Search Nodes): Select a node, i , in P .
 Scan all inbound arcs (ji) and outbound arcs (ij).
 If $(v_{ij} \geq v_{\text{MIN}} \text{ and } d_{ij} \leq d_{\text{MAX}})$ or $(v_{ji} \geq v_{\text{MIN}} \text{ and } d_{ji} \leq d_{\text{MAX}})$,
 then place the arc into LTOUR.

Step 3 (Identify Reloads): If the total volume of LTOUR, $V_{\text{LTOUR}} > V_{\text{MIN}}$, include it in the list of local tour opportunities.

Potential Savings

By dedicating a vehicle, or vehicles, to a set of short haul lanes, a carrier can keep a equipment and personnel operating continuously. The repetitiveness of the loads will compensate for the relatively high empty miles. The savings to the carrier, then, accrue due to the time the vehicle is kept in motion. The total time in a vehicle is used by a local tour can be estimated as:

$$\sum_{ij \in \text{LocalTour}} v_{ij} \left(\frac{d_{ij}}{250} \right)$$

where:

- v_{ij} : Annual volume on lane ij , and
- d_{ij} : Distance of lane ij .

The total time is the sum of the total number of loads per year multiplied by the portion of a day that each load should take. The value of 250 miles per day makes the conservative assumption of no backhauls. The cost of serving the local lanes can be estimated by comparing the

utilization of vehicles winning a portion of the lanes within the local tour, and the utilization from the above estimate.

5.4.5 Application to the Data

These four algorithms were applied to the MTMC data. The effectiveness of the algorithms is dependent on the topology of the network, the traffic volume, and the different parameters used. Of the 45 points, 30 were identified as having reload potential.

Thirty one open tours were identified: 21 two lane, 8 three lane, and 2 four lane. The parameters for the algorithms were set as follows: minimum threshold volume, $v_{\text{MIN}} = 100$ loads per year; number of lanes in a thread [$1 \leq \text{lanes} \leq 4$]; and length of lanes = [$100 \leq \text{distance} \leq 1000$ miles].

Fifteen closed tours were identified: 12 two lane and 3 three lane tours. The parameters for the algorithms were set as follows: minimum threshold volume, $v_{\text{MIN}} = 100$ loads per year; number of lanes in a thread [$1 \leq \text{lanes} \leq 4$]; and length of lanes = [$100 \leq \text{distance} \leq 500$ miles].

Seven locations with potential local tours were identified. Each location used a threshold of at least 100 loads per year per lane, maximum distance of 250 miles, and a total tour volume of at least 1000 loads per year. Each location used a least 6 individual lanes.

5.4.6 Practical Concerns

While efficient aggregations of traffic lanes are economically beneficial to both the shipper and the carrier, two practical concerns should be considered: auditing of rates and tendering requirements.

Auditing Problems

One problem with forming efficient aggregations of lanes is that they may create difficulties in the auditing process. The audit is typically needed to validate the correct rates on the tens and hundreds of thousands of transportation transactions that occur annually at a firm. If the price on a lane is dependent on future or past actions, then the audit is much more complicated and may not be supportable by many trucking firms. For example, if the rate for a load is \$1.00 per mile from B to C, but \$0.95 if the truck is coming from A to B results in a reload at B, then the freight bill for the outbound load needs to indicate which rate applies. That is, the rate for a lane is no longer independent of other lanes.

An alternative to having rate interdependence is to request a single rate from the carriers based on their understanding of the probability of being able to make the connections inherent in the efficient aggregation. The rate from B to C, then, might be \$0.97 regardless of where the inbound vehicle to B is coming from since the carrier knows that, say, 50% of the time

it will be a reload. Readjustment of the rates based on the true percentage of reloads could be included in the standard quarterly meetings.

Tendering

The benefit of the efficient aggregations rely on the connection between different lanes being made. While the contract and bidding process might specify that a carrier is the primary carrier on both the inbound and outbound lanes at a point, only the real time tendering process will guarantee that the carrier will receive both of these loads. This is especially true for closed tours which rely on the carrier being kept in continuous motion.

The shipper needs to ensure that the tendering matches the contracted decisions. The minimum acceptable tender rate could be included in the contract as a means of enforcement. Additionally, software or different systems might be required at the operational level in order to quickly identify and match the inbound to the outbound loads. The tendering decision needs to consider not only which carrier to assign to an outbound load, but which type of load to match it to.

5.5 Step 5: Designing the Bid

The last step in the Pre-Bid analysis, is to determine the auction design best fitted to the network in question. The design is situation specific. The shipper should select the most economical auction design to allocate carriers to lanes. There is a trade-off between the cost of running the auction and the benefits of reaching an efficient allocation. Not all TL auctions require simultaneous, multiple-round, conditional-bid auctions. Likewise, not all auctions should be single-round, sealed-bids by default. The flow chart in Figure 5.17 illustrates the thought process and the factors critical to the auction design selection.

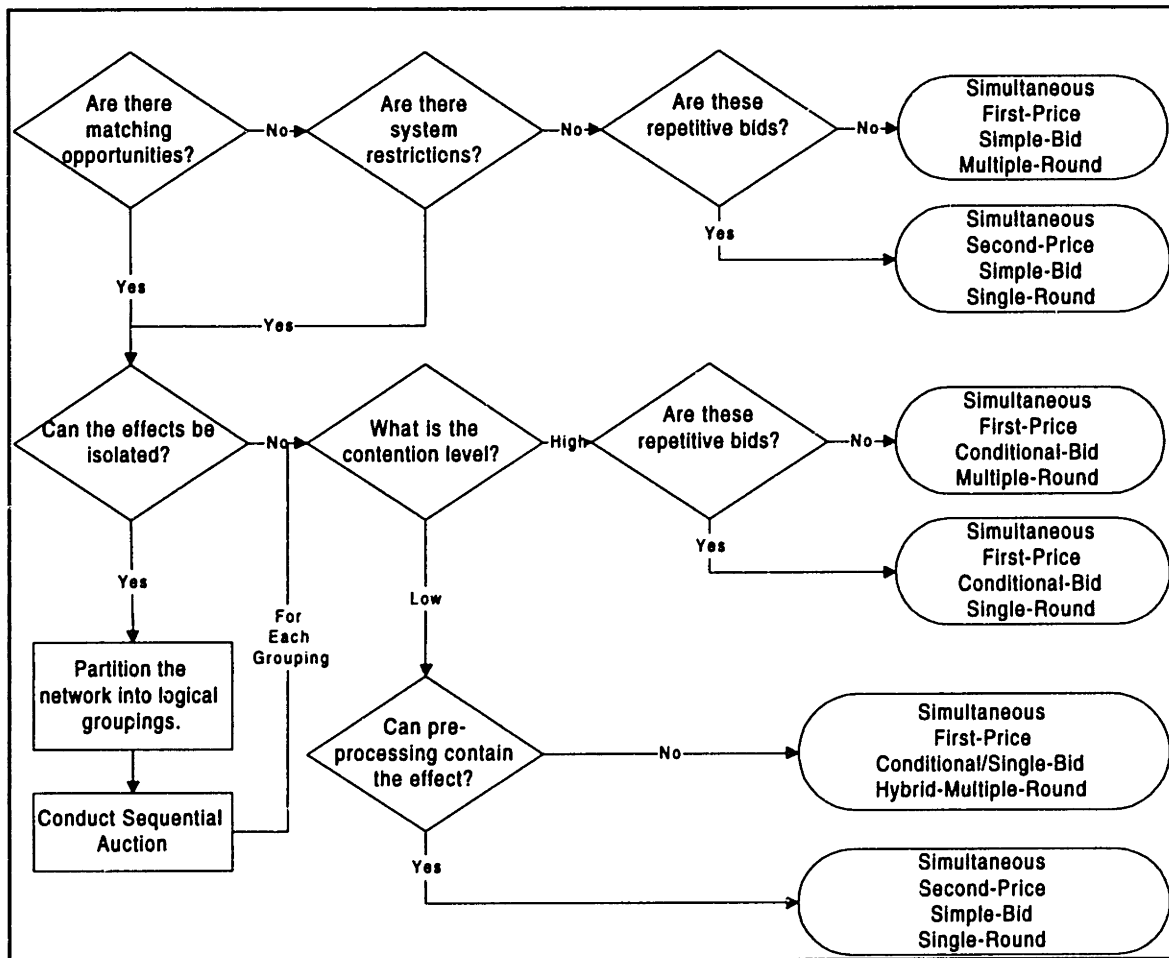


Figure 5.17. Flowchart of auction design choice.

Six decision points are identified: presence of matching opportunities, use of any system restrictions, repetitiveness of the bids being let out, how well portions of the network can be partitioned, level of contention, and ability to pre-process portions of the network. The latter five points are briefly discussed in this section.

5.5.1 System Restrictions

The typical auction does not place restrictions on the outcome. For example, at an art auction, the auctioneer does not care who wins which paintings or how many paintings each person wins. For TL bidding, however, the shipper often desires to enforce certain requirements on the outcome. For example, a shipper might want to restrict the number of carriers in the system or serving a particular location while ensuring that each carrier in the system wins a certain threshold volume.

Restrictions are difficult to enforce in an auction because they create different, and sometimes perverse, interdependencies. Take for example an auction where the shipper wants

no more than two carriers. Suppose that at the end of one of the rounds of bidding for three lanes the situation is as shown in Table 5.6 with the current winning bids in bold.

| Carrier | A | B | C |
|---------|------------|----------|-----|
| Lane 1 | 1 | 2 | 2.5 |
| Lane 2 | 3 | 2 | 3 |
| Lane 3 | 2.5 | 3 | 2 |

Table 5.6. System restriction example - current bids.

Since the shipper is restricting the number of carriers to two, the leading bidders are Carrier A for lanes 1 and 3 and Carrier B for lane 2. Note that Carrier C has bid below the leading bid for lane 3 and still loses. For Carrier B to win lane 3, it needs to submit a bid just below \$2.5, say \$2.4. Carrier C, however, would have to submit a bid of \$1.4 for lane 3 in order to win it! This would result in Carrier B winning lanes 1 and 2 at \$2 each and lane 3 winning lane 3 at \$1.4.

Even more interesting, if Carrier B bids \$1.4 on lane 1, then Carrier C will also win lane 3! This would result in Carrier B winning lanes 1 and 2 at \$1.4 and \$2, respectively, and Carrier C winning lane 3 at \$2! Through no action of its own, Carrier C wins the lane! The restriction of having only two carriers adds an additional cost to carriers not currently in the system.

The restricted bids reflect a cost of being in the system. This is a form of interdependency. Instead of being based on lane interactions, these interdependencies are based on the carrier as a whole. The incremental margin that a carrier needs to win a lane depends on whether it has won any other lanes. Restrictions on an outcome, then, require combinatorial auctions. The problem, even with simple bids, requires the solving of a set partitioning problem. The bids are conditional on whether a carrier has won any lanes, rather than a specific lane. Also, the interdependency applies to all lanes for that carrier.

5.5.2 Repetitive Bids

If the shipper bids out the lanes on a regular basis, then the carriers will begin to get an idea of the range of bids being accepted. If there are not significant changes in either the patterns of the shipments or the lanes, then the repetitiveness of the auctions will help reduce the likelihood of incurring the winner's curse. Annual auctions essentially serve the purpose of previous rounds in a multiple-round auction. Therefore, a single-round auction might be sufficient.

The market information that would normally be developed by having multiple rounds already exists in the market - it is common knowledge. For trucking, there is a tremendous amount of market information concerning very similar items (loads going over the same routes

or into/out of the same areas) which means the range of valuation is very tight for standard lanes. Thus, the probability of overvaluation is lowered, but certainly not eliminated.

Both the MTMC and the ABC networks are bid out every year, so they are repetitive. Even more interesting, in the MTMC bids, all submitted rates are made public after submission. This tends to reduce the uncertainty among bidders but can lead to implicit collusion and price setting by the carriers.

5.5.3 Separable Effects

Sometimes, the interdependencies between lanes can be eliminated by consolidating the lanes into a group and treating it as a single entity. If it is possible to separate the auction into simpler sections, then it should be done. This also applies to different trucking segments if they are incompatible, such as tank cars and flatbeds. If the lane interdependencies are separable, the lanes should be bid out in a sequential series of simultaneous auctions.

The MTMC holds separate bids for lanes from each different major shipping point. By letting out each major shipping point in separate bids misses the opportunity for creating long haul tours, either open or closed. For example, the San Francisco Bay, Philadelphia, and Virginia Beach/Newport News areas are all major shipping locations where bids are let out separately. The traffic between these three locations though is not only of fairly high volume, it is closely balanced as well. Figure 5.18 shows the annual traffic volume for these lanes.

By bidding out each shipping location separately, the synergies of forming long haul tours between locations are lost. In the MTMC data, no single carrier operated on more than one of these legs! Clearly, separating the bids out by shipping location is not efficient in this case.

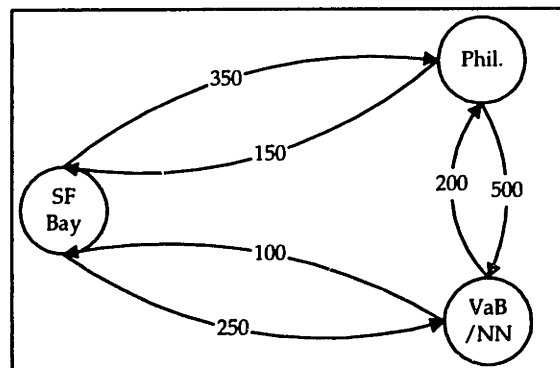


Figure 5.18 Flow between three major shipping points in MTMC network.

5.5.4 Contention Level

The contention level of an auction is a measure of the amount of competition between the items. For the TL bidding problem, it is based on the amount of overlaps of interdependent

lanes, the number of different combinations that these lanes can be put in, and the number of carriers interested in these different combinations. The key point is, the higher the contention for different lanes, the more intense the competition in the auction. High contention auctions imply that the different interdependencies have a great deal of overlap and therefore require conditional bids. Low contention auctions, where, for example, the number of carriers participating is less than the maximum number desired by the shipper, may not require conditional bids. Also, a single round auction might be sufficient if the contention is low - even if the interdependencies are quite high.

The contention level will most likely be quite high for any TL bid where both tours and reloads are offered. Figure 5.19 shows a highly contentious segment of the MTMC network in the central California region.

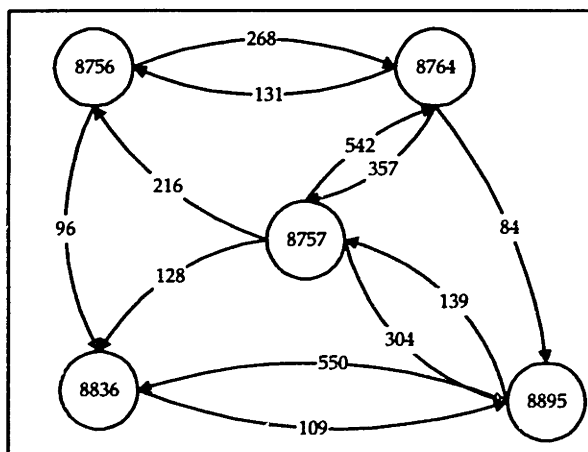


Figure 5.19 Example of contention in the MTMC network.

The same high density lanes will tend to be candidates for a number of different tour combinations. In the MTMC data, only 48 lanes were used to fill the 107 lane 'slots' identified in the 46 closed and open tours. In the ABC data, 597 closed and open tours were identified involving a potential of 1769 lane slots. These slots were actually covered by just 141 lanes. While the median number of lanes per package was 0 across the entire network, for those lanes in any tour, the median was 6 with an average of 12 lanes per package. One of the lanes was used in 57 different tours. So, while contention in the TL networks appears to be low over the entire network, it is quite high over certain high-volume lanes. Contention makes the assignment problem more difficult and is discussed in greater depth in Chapter 6.

5.5.5 Pre-Processing

If the auction has a low level of contention, then, it might be possible to pre-process some of the lanes into single entities. This is different from dividing the network into larger segments - such as for flatbed or dry van. Instead, if a two lane (there-and-back) tour has very

high volume and there are no alternative pairings of the lanes that would result in more savings, then it makes sense to offer this as a single entity from the start.

For example, a local tour from a location might be of barely sufficient volume for a dedicated vehicle unless all of the lanes are included. If there are no other contending tours for these lanes, then it makes sense to bundle them ahead of time prior to the bid and let them out as a single item.

If it is not possible to pre-process some of the lanes, but the contention is still low, then perhaps conditional bids can be used in the first round of the auction to identify the 'best' efficient aggregations. After these have been identified, they can be locked in and simple bids can be used for the remainder of the auction. Pre-processing is used to try to use simple bids as much as possible.

5.6 Closure

The objectives of this chapter were to understand why the different network characteristics are important, how they can be identified, and how to estimate the potential savings. A pre-bid analysis was used to formalize the process by which a shipper can analyze their own network.

A key observation that the analysis brought out, is that the traffic flow is not uniform across lanes or locations within a network. The majority of the shipping locations will have very low traffic density. Only a small set of lanes and locations within a distribution network will be able to achieve predictable economies of scope using efficient aggregations. These lanes can be pulled out and treated as a separate network for the purpose of identifying efficient aggregations. Because the efficient aggregations can occur over such a small portion of the network, there will be a high level of contention between the high-density lanes.

Four types of aggregations were identified, each with its own level of savings and probability of occurrence: reloads, threads, closed tours, and local tours. Reloads require a high level of traffic both into and out of a shipping location. The savings accrue due to the removal of empty miles from the connection between loads. Threads are simply a chain of reloads with the exception that deadhead lanes are allowed and the connections can be made within zones - not just at specific shipping locations. The savings accrue due to decreased expected dwell time and empty miles and are generally lower than reload savings. Closed tours are a chain of lanes that together begin and end at the same location or zone. In order to maintain a closed tour, the truck has to be essentially dedicated to the shipper so that the savings is accrued by the removal of all connection uncertainty costs. The savings from closed tours are the highest of the efficient aggregations. Finally, local tours are a collection of short haul lanes with a common origin or

destination. The savings accrue due to the repetitiveness of the less-than-one-day trips. Also, the shipper benefits directly by having reliable capacity on a sometimes undesirable lane.

The final outcome of the pre-bid analysis is to identify a set of *potential* efficient aggregations to offer to the bidding carriers. These are certainly not the complete set of all possible aggregations. Carriers may (and should be encouraged to) identify aggregations that involve other customers' lanes. The shipper can only report what is contained within its own network and let the bidding carriers take them or form their own through conditional bids.

Chapter 6

Carrier Assignment Problem

The carrier assignment problem determines which carriers should be assigned to which portions of the shipper's distribution network. The form of the problem is dependent on the type of auction and the contracting form used. In simple-bid auctions, for example, the assignment of carriers to lanes is done by sorting the bids by value for each lane and assigning the lowest cost carrier. When conditional bids are used, or system restrictions are in force, the assignment problem involves making more difficult trade-offs and requires the solution of an integer mathematical program.

This chapter discusses the formulation, solution, and implications of various carrier assignment models and demonstrates the practicality and usefulness of optimization-based carrier assignment for large and contentious traffic networks. In addition to assigning carriers to lanes, the models can be used to provide insight into business practice and quantify business decisions.

The chapter is organized as follows. In Section 6.1, a general carrier assignment model is introduced and the various ways it can be customized to reflect different carrier and shipper restrictions are discussed. The distinction between bid-set and carrier-set models are explained later in the chapter, but essentially, bid-set models capture economies of scope while carrier-set models can potentially capture economies of density as well. Section 6.2 presents a proposed carrier-set assignment model formulation and discusses solution approaches. Section 6.3 introduces a set of lane-based, bid-set models which are used for the analysis within the chapter. Section 6.4 discusses the implementation of these models, solution strategies, and business implications. Finally, Section 6.5 closes the chapter.

6.1 General Formulation

The carrier assignment model selects which carriers should serve the shipper's TL network. The objective is typically to find the lowest cost assignment. A very general carrier assignment model can be formulated as:

$$\min_x \sum_k C^k(x^k)$$

Subject to

[GCA]

$$(1) \quad \sum_k x^k = D$$

$$(2) \quad x^k \in X \quad \forall k$$

Where:

- x^k : Vector of volume for each lane (or segment) assigned to carrier k ,
- $C^k(x^k)$: Cost function for carrier k to serve the vector of lanes (segments) x^k ,
- D : Vector of expected volume demanded on each lane (segment), and
- X : The set of all feasible volume allocations and lane (segment) assignments.

The objective is to minimize the total cost to the shipper of having carriers serve the expected traffic volumes. Constraint (1) is a covering constraint ensuring that each lane (segment) is served by a sufficient number of carriers. Constraints (2) specify that the allocated lanes and respective volumes are feasible to both the carrier and the shipper.

The cost function, C^k , and the set of feasible assignments, X , are determined jointly by both the shipper and the carrier. The form of the cost function dictates the type of problem being solved. Three examples of cost functions are shown below:

$$C_1(x_{ij}^k, y_i^k, z^k) = \sum_k F^k z^k + \sum_k \sum_i F_i^k y_i^k + \sum_k \sum_{ij} c_{ij}^k x_{ij}^k$$

$$C_2(v_{ij}^k) = \sum_k \sum_{ij} C(v_{ij}^k) v_{ij}^k = \sum_k \sum_{ij} \left(a_{ij}^k v_{ij}^k + \frac{b_{ij}^k}{v_{ij}^k} \right)$$

$$C_3(v_{ij}^k) = \sum_k \sum_{ij} C(v_{mn}^k) v_{ij}^k = \sum_k \sum_{ij} \left[\sum_{mn} (d_{ij}^{mn}) v_{ij}^k \right]$$

where:

- x_{ij}^k : =1 if carrier k is assigned all lane ij , =0 otherwise
- v_{ij}^k : = volume of traffic awarded to carrier k on lane ij
- y_i^k : =1 if carrier k serves origin i , =0 otherwise
- z^k : =1 if carrier k is in the system at all, =0 otherwise
- F^k : Cost of including carrier k into the system and
- F_i^k : Cost of including carrier k serve location i .
- a_{ij}^k : Cost per load for carrier k to service lane ij
- b_{ij}^k : Marginal effect to lane ij of a unit of flow
- c_{ij}^k : Total annual cost for carrier k to service lane ij
- d_{ij}^{mn} : Marginal effect to lane ij of a unit of flow on lane mn

The cost function C_1 is divided into a variable component for the volume and fixed costs for the lane assignment and system assignment. With cost function C_1 , [GCA] becomes a network design problem where each lane (element in array D) has a set of k potential or candidate arcs which can be added to make the connection. Each candidate arc (one for each carrier offering to service that lane) has fixed costs, variable costs, and, perhaps, capacity requirements. Having multiple arcs per OD pair is allowed and the objective is to find the

minimum cost design. Similarly, part of the cost function may be determined by the carriers' bids and another part specified by the carrier based on, say, administrative costs or service quality. The cost function C_2 introduces a non-linear function with respect to the volume on each lane but the cost of serving each lane is *separable* from the activity on the other lanes. The cost is now concave. Cost function C_3 incorporates the volume on other lanes into the cost of serving a lane. The costs are still *additive* across the lanes and carriers, however. These are only a few examples, there are many possible cost function forms that can be used.

6.1.1 Variations of Carrier Assignment Model

We have shown that TL carriers' costs are affected by economies of scope and density which in turn are influenced by how the shipper assigns lanes and allocates traffic volume. These costs can be reflected in the assignment model in the form of the objective function. Table 6.1 shows six different ways that the relationship between the cost of serving a lane, the volume allocation, and the lane assignment can be captured.

| | | Economies of Scope | | |
|----------------------|---------------|--|--|--|
| | | Separable | Non-Separable | |
| Economies of Density | Independent | (1) Lane specific, volume independent. | (2) Rate independent of volume but affected by all lane assignments. | Lane rate is independent of volume on any lane. |
| | Separable | (3) Lane specific, rate is variable wrt lane volume. | (4) Rate affected by lane volume and all lane assignments. | Lane rate is affected by volume in that lane alone. |
| | Non-Separable | (5) Lane specific, rate affected by <u>system wide</u> volume. | (6) Rate affected by lane assignments and volume allocations across entire system. | Lane rate is affected by volume throughout the system. |
| | | Lane rate is affected by assignment of this lane alone. | Lane rate is affected by all lane assignments. | |

Table 6.1 Characteristics of the cost function for the general model.

The most common form of cost function used in TL carrier assignment in practice is class (1), in the upper left corner, which assumes that the bid rate from a carrier is independent of the volume carried on a lane and is unaffected by any other lanes awarded to that carrier. In auction design terms, this implies that the lanes exhibit no interdependencies so that simple-bids are sufficient. At the other extreme is class (6), in the lower right corner, which assumes that the cost of serving an additional load on one lane is influenced by the loads being served on every other lane in the system. Class (6) models require a non-linear cost function taking the vector of all lane assignments and volume allocations as arguments. The cost function C_2 discussed previously is of this type.

While desirable from a theoretical point of view, a model with a class (6) objective function is not practical. First, the full relationship between bid rates, lanes, and volumes is unlikely to be known in theory, much less found in practice. Second, even if this was quantifiable, it would be non-linear and extremely difficult to solve.

A better solution is to use conditional bids. Conditional bids capture the salient aspects of *economies of scope* that drive a carrier's economics but are not highly dependent on the details of the network volume estimates. Rather than trying to capture the effects of every lane on every other lane's cost structure through an involved non-linear cost function, conditional bids let each carrier specify the relationships between costs, volumes, and lane assignments on small, self-contained packages of lanes. An additional benefit is that conditional bids can be used to incorporate both economies of scope while still maintaining a linear cost function. A carrier submitting a conditional bid for a lane package essentially makes the cost of serving each of these lanes non-separable and incorporates economies of scope.

6.1.2 Previous Research

The only published account of any optimized based carrier assignment model is Moore, Warmke, and Gorban, (1991), referred to as MWG. They describe how the Reynolds Metal Company centralized their entire transportation management system. A critical step in this process was revamping how Reynolds bid out lanes of traffic and assigned carriers to lanes, locations, and the system as a whole. They developed a mixed integer program (MIP) model which minimizes freight costs (using only single-move, flat rates) by assigning carriers to specific shipping locations and traffic lanes and allocating traffic volumes. The results of the MIP were then used in a simulation model of daily freight operations which took into account service problems with the fluctuating shipping requirements and other considerations. The simulation led Reynolds to make the final decisions on the number of carriers, core carrier selection, carrier assignment to shipping locations, fixed and variable commitments⁸¹, expected shipping volumes and commitments, and estimated costs and savings. The carrier assignment model was only a small part of their larger centralized dispatch system which looked for continuous moves in real-time.

⁸¹ MWG defines a fixed commitment as a requirement for a number of vehicles at a location each day where the carrier is paid whether they are used or not. Variable commitments are the same except the carrier is only reimbursed if the vehicles are used.

They modeled the carrier assignment as:

$$\min_x \sum_i \sum_j \sum_k R_{ijk} x_{ijk}$$

Subject to

- (1) $\sum_k x_{ijk} = D_{ij} \quad \forall i, j$
- (2) $x_{ijk} - y_{ik} D_{ij} \leq 0 \quad \forall i, j, k$
- (3) $y_{ik} - z_k \leq 0 \quad \forall i, k$
- (4a) $\sum_k y_{ik} \leq MX_i \quad \forall i$
- (4b) $\sum_k y_{ik} \geq MN_i \quad \forall i$
- (5a) $\sum_k z_k \leq MXT$
- (5b) $\sum_k z_k \geq MNT$
- (6) $\sum_j x_{ijk} - y_{ik} P_{ik} \leq 0 \quad \forall i, k$
- (7) $\sum_i \sum_j x_{ijk} - z_k PT_k \leq 0 \quad \forall k$
- (8) $\sum_j x_{ijk} + y_{ik} MNSV_{ik} \geq 0 \quad \forall i, k$
- (9) $\sum_i x_{ijk} \geq OVMN_{ik} \quad \forall i, k$ [MWG]
- (10) $\sum_i \sum_j x_{ijk} \geq OVMNT_k \quad \forall k$
- (11) $x_{ijk} \geq 0, y_{ik}, z_k = \{0,1\}$

Where:

Indices:

| | |
|-----------|-------------------------------|
| <i>i:</i> | Shipping location origin |
| <i>j:</i> | Shipping location destination |
| <i>k:</i> | Carrier identification |

Decision Variables:

| | |
|-------------|--|
| x_{ijk} : | Number of shipments by carrier <i>k</i> from <i>i</i> to <i>j</i> |
| y_{ik} : | =1 if carrier <i>k</i> serves origin <i>i</i> at any level, =0 otherwise |
| z_k : | =1 if carrier <i>k</i> is in the program at all, =0 otherwise |

Data:

| | |
|----------------|---|
| D_{ij} : | Demand (# of shipments) from <i>i</i> to <i>j</i> |
| MN_i, MX_i : | Minimum, maximum number of carriers allowed at origin <i>i</i> |
| MNT, MXT : | Minimum, maximum number of carriers allowed in the program |
| $MNSV_{ik}$: | Minimum award level for carrier <i>k</i> if it serves origin <i>i</i> |
| P_{ik} : | Available trucks for carrier <i>k</i> at origin <i>i</i> |
| PT_k : | Total system-wide trucks available from carrier <i>k</i> |
| $OVMN_{ik}$: | Override minimum award level for carrier <i>k</i> at origin <i>i</i> (a priori) |
| $OVMNT_k$: | Override minimum total award level for carrier <i>k</i> (a priori) |
| R_{ik} : | Rate charged by carrier <i>k</i> to ship a load from <i>i</i> to destination <i>j</i> |

The objective is to minimize the variable transportation costs for an average day's demand. Constraints (1) ensure all demand is met, (2) assign a carrier to an origin if it is awarded any loads from that origin, (3) assign any carrier to the program if they are assigned to any origin, (4) keep the number of carriers at each origin within a preset minimum and maximum, (5) keeps each carrier within minimum and maximum limits for the entire system, (6) ensure the number of loads assigned to a carrier at each origin does not exceed the available trucks at that origin, (7) ensure the number of loads assigned to a carrier system-wide does not exceed the total number of trucks available, (8) ensure that if a carrier is assigned to an origin then it must have a certain minimum number of loads assigned to it, (9) are override constraints which specify the minimum number of shipments that a carrier has to be awarded from an origin, thus forcing its assignment to that origin, is met and finally (10) are more override constraints which specify that a carrier has a minimum total number of loads in the system is met.

The model was solved on a reduced set of carriers from the total bid. MWG removed all "short lanes" (representing 30% of the network but carrying less than 3% of the freight bill) and included the top 19 carriers based on the percentage of their bids that were in the lower 50th percentile for each lane. This reduced set of carriers was used in the assignment model. A typical model run (for 19 carriers over a network of 120 origins and 5,000 destinations)⁸² consisted of over 5,000 constraints, 9,000 columns, 26,000 non-zero elements, and 200 binary variables. Using an IBM 3090 computer and an MPSX IP solver, the runs took two to three CPU minutes. These models solved very quickly and required no special solving techniques.

The MWG matrix is very sparse - less than 0.05% non-zero elements. This is mainly due to their distribution network being very spread out with distant and isolated plants served by regional carriers. This sparse nature may help explain why the model was solved so quickly. Shippers with dense networks and more lanes, will most likely have much slower times. The number of columns will grow significantly if more national carriers that serve the majority of traffic lanes were included. Also, if the carriers were primarily regional, then the number of shipping origins they are willing to be assigned to is quite small. This is reflected in the model only having about 200 binary variables, 19 of which are the z_k variables leaving only 181 y_k variables for assigning carriers to the origin locations. Once these location variables are assigned, the model becomes a network flow problem.

A few of comments are in order to describe the differences between our proposed models and the MWG model. First, economies of scope are not captured in the MWG model.

⁸² From the reported size of the matrix, the number of actual lanes used in the model could not have exceeded a few hundred. The shipping origins and destinations had to have been aggregated into zones.

Even though the final objective is to identify continuous move potentials, only single lane bids were used and traffic lanes were assigned independently of each other. Thus, the carriers still have an incentive to hedge against the uncertainty of finding a follow-on load. This will inflate the bids offered. Second, the actual carrier selection was based on a simulation made after the optimization model was run. In fact, the model was run mainly as a preliminary to the simulation. Third, the problem size that we consider is much larger and is being solved on a desktop PC environment. Finally, we examine the use of the optimization model within a larger auction framework while the MWG model was used exclusively as generating input for the simulation model.

6.2 Carrier-Set Formulation

This section presents the carrier-set [CS] formulation of the carrier assignment model. It is called a carrier-set model because it treats the entire allocation of lanes and traffic volume to a carrier as a single decision variable. Considering a carrier's economics as a whole allows us to include various volume discount pricing strategies while maintaining a linear IP. The formulation is designed to be solved using column generation.

This section specifically addresses the methodological concerns with carrier-set models rather than any business implications. Readers interested only in the practical concerns of the transportation procurement problem can skip this section without any loss of continuity.

6.2.1 Formulation

The carrier-set assignment model [CS1] finds the lowest cost assignment of carriers to the shipper's distribution network subject to hard restrictions on the number of carriers in the system and at each location. It is formulated as:

$$\min_{CS_s^k} \sum_s \sum_k R_s^k CS_s^k$$

Subject to

$$\begin{aligned} (1) \quad & \sum_s \sum_k a_{ij}^{sk} CS_s^k \geq D_{ij} & \forall i, j & \quad (\pi_{ij}) \\ (2) \quad & \sum_s CS_s^k \leq 1 & \forall k & \quad (-\sigma_k) & \quad \text{[CS1]} \\ (3) \quad & \sum_s \sum_k \delta_i^{sk} CS_s^k \leq K_i & \forall i & \quad (-\rho_i) \\ (4) \quad & \sum_s \sum_k CS_s^k \leq K & & \quad (-\theta) \\ (5) \quad & CS_s^k = [0,1] & & \end{aligned}$$

where:

Indices:

- i*: Origin shipping location, $i = 1, 2, \dots, m$
j: Destination shipping location, $j = 1, 2, \dots, n$
k: Carrier identification
s: Carrier-set identification

Decision Variables:

CS_s^k : =1 if carrier *k* is assigned Carrier-set *s*; =0 otherwise

Data:

- R_s^k : Cost of shipping all loads contained in the carrier-set *s* using carrier *k*
 D_{ij} : Demand for truckloads required for lane *i* to *j*
 a_{ij}^k : Number of loads assigned to lane *i* to *j* for carrier *k* under Carrier-set *s*
 K : Maximum number of carriers allowed in the system
 K_i : Maximum number of carriers allowed to serve location *i*
 δ_{is}^k : =1 if carrier-set *s* for carrier *k* has a lane originating from location *i*; =0 otherwise

The objective function for [CS1] minimizes the costs of the carrier-sets. Constraints (1) ensure demand is satisfied for each lane; (2) is a convexity constraint which states that no more than one carrier-set is selected for each carrier; (3 & 4) are forcing constraints limiting the number of carriers serving a location and the system, respectively; and (5) define the variable type. Note that the hard constraints do not require additional location or system variables. The dual variables are in parentheses next to their respective constraint.

Alternatively, the carrier-set model can be formulated using 'soft' constraints, as in [CS2], below.

$$\min_{CS,y,z} \sum_s \sum_k R_s^k CS_s^k + \sum_i F_i y_i + F z$$

Subject to

- (1) $\sum_s \sum_k a_{ij}^{sk} CS_s^k \geq D_{ij} \quad \forall i, j \quad (\pi_{ij})$
 (2) $\sum_s CS_s^k \leq 1 \quad \forall k \quad (-\sigma_k) \quad \text{[CS2]}$
 (3) $\sum_s \sum_k \delta_{is}^k CS_s^k - y_i = 0 \quad \forall i \quad (\omega_i)$
 (4) $\sum_s \sum_k CS_s^k - z = 0 \quad (\Psi)$
 (5) $y_i, z = \text{integer}, \quad CS_s^k = [0,1]$

where:

Decision Variables:

- y_i : Number of carriers serving location *i*
 z : Number of carriers in the system

Data:

- F_i : Fixed cost of adding carrier *k* to serve location *i*
 F : Fixed cost of adding carrier *k* to the system

The objective function for [CS2] minimizes the total cost to include the fixed costs for carriers and locations. Constraints (1) ensure demand is satisfied for each lane; (2) are convexity constraints for each carrier; (3 & 4) are logical constraints that assign a carrier to a location and the system if any loads are carried; and (5) defines the variable types. Note that location and

carrier columns are now required although they may be desired for implementation in order to tighten the LP bounds. The dual variables are shown in parentheses next to their respective constraint. We will not consider this formulation any further in this section.

The decision variables for these formulations consist of all assigned lanes and allocated volumes for a carrier within the entire system. A column generation framework is suggested in order to efficiently create candidate carrier-sets to use in [CS1]. Designating [CS1] as the master problem, we can develop optimality conditions to determine if candidate columns generated from a subproblem are worth considering.

The optimality conditions can be taken directly from the dual of the problem, shown below.

$$\max_{\pi, \sigma, \rho, \theta} \sum_i \sum_j D_{ij} \pi_{ij} - \sum_k \sigma_k - \sum_i \rho_i K_i - \theta K$$

Subject to

$$(1) \quad \sum_i \sum_j (a_{ij}^{sk} \pi_{ij}) - \sigma_k - \sum_i (\delta_i^{sk} \rho_i) - \theta \leq R_s^k \quad \forall k, s \quad \text{[DCS1]}$$

$$(2) \quad \sigma, \rho, \theta, \pi \geq 0$$

The objective function of [DCS1] maximizes the shadow prices paid for the resources used in [CS1]. Constraints (1) ensure that the sum of the shadow prices for the resources used for each carrier-set do not exceed the amount bids, R_s^k , and (2) ensure the correct signs for the variables.

We consider a candidate carrier-set for the master problem only if it is dual feasible, that is, if the reduced costs are negative. Thus, we are optimal if the reduced cost of all potential carrier-set columns are non-negative. The reduced costs for [CS1] are:

$$\begin{aligned} \bar{R}_s^k &= R_s^k - \sum_i \sum_j (a_{ij}^{sk} \pi_{ij}) + \sigma_k + \sum_i (\delta_i^{sk} \rho_i) + \theta \geq 0 \\ R_s^k - \sum_{i \in (sk)} \left[\sum_j (a_{ij}^{sk} \pi_{ij}) - \rho_i \right] + \sigma_k + \theta &\geq 0 \\ R_s^k - \sum_{i \in (sk)} \left[\sum_j (a_{ij}^{sk} \pi_{ij}) - \rho_i \right] &\geq -(\theta + \sigma_k) \end{aligned}$$

The complementary slackness conditions state:

$$CS_s^k \left(R_s^k - \sum_{i \in (sk)} \left[\sum_j (a_{ij}^{sk} \pi_{ij}) - \rho_i \right] + \sigma_k + \theta \right) = 0$$

along with feasibility and non-negativity requirements. These imply that if carrier-set sk is used in [CS1] the reduced cost must be equal to 0. For optimality of the entire problem, all of these conditions need to hold.

6.2.2 Implementation Strategies

This section describes a method for solving the carrier-set assignment problem using column generation. A restricted version of the original carrier-set model, [CSR], is used as the master problem. The system limit constraint has been removed.

$$\min_{CS_s^k} \sum_s \sum_k R_s^k CS_s^k$$

Subject to

$$\begin{aligned} (1) \quad & \sum_s \sum_k a_{ij}^{sk} CS_s^k \geq D_{ij} & \forall i, j & \quad (\pi_{ij}) \\ (2) \quad & \sum_s CS_s^k \leq 1 & \forall k & \quad (-\sigma_k) \\ (3) \quad & \sum_s \sum_k \delta_i^{sk} CS_s^k \leq K_i & \forall i & \quad (-\rho_i) \\ (4) \quad & CS_s^k = [0,1] \end{aligned} \quad \text{[CSR]}$$

where:

Indices:

- i:* Origin shipping location, $i = 1, 2, \dots, m$
- j:* Destination shipping location, $j = 1, 2, \dots, n$
- k:* Carrier identification
- s:* Carrier-set identification

Decision Variables:

- $CS_s^k:$ =1 if carrier *k* is assigned carrier-set *s*; =0 otherwise

Data:

- $R_s^k:$ Cost of shipping all loads contained in the carrier-set *s* using carrier *k*
- $D_{ij}:$ Demand for truckloads required for lane *i* to *j*
- $a_{ij}^k:$ Number of loads assigned to lane *i* to *j* for carrier *k* under carrier-set *s*
- $K_i:$ Maximum number of carriers allowed to serve location *i*
- $\delta_{is}^k:$ =1 if carrier-set *s* for carrier *k* has a lane originating from location *i*; =0 otherwise

The objective function for [CSR] minimizes the costs of the carrier-sets. Constraints (1) ensure demand is satisfied for each lane; (2) are convexity constraints limiting carriers to one carrier-set; (3) limits the number of carriers serving a location; and (4) ensures variables are binary. The dual variables are in parentheses next to their respective constraints.

The bid in Figure 6.1 is used as a numerical example for the remainder of the section. Figure 6.1 shows that 3 carriers have placed a total of 18 bids on the 5 lane, 4 node distribution network. Each bid specifies the carrier's minimum and maximum capacity and the relevant lanes. In Figure 6.1, each column is a bid-set. The volume on all lanes within a bid need to be the same. For example, Carrier A has bid \$2.85 in a conditional bid covering lanes {2-3; 3-4; and 4-2}. Carriers also have system and capacity limits, as shown in the matrix in the bottom left of Figure 6.1. Finally, Carriers A and C have offered discounts dependent on the volume awarded - these are specified in Figure 6.1.

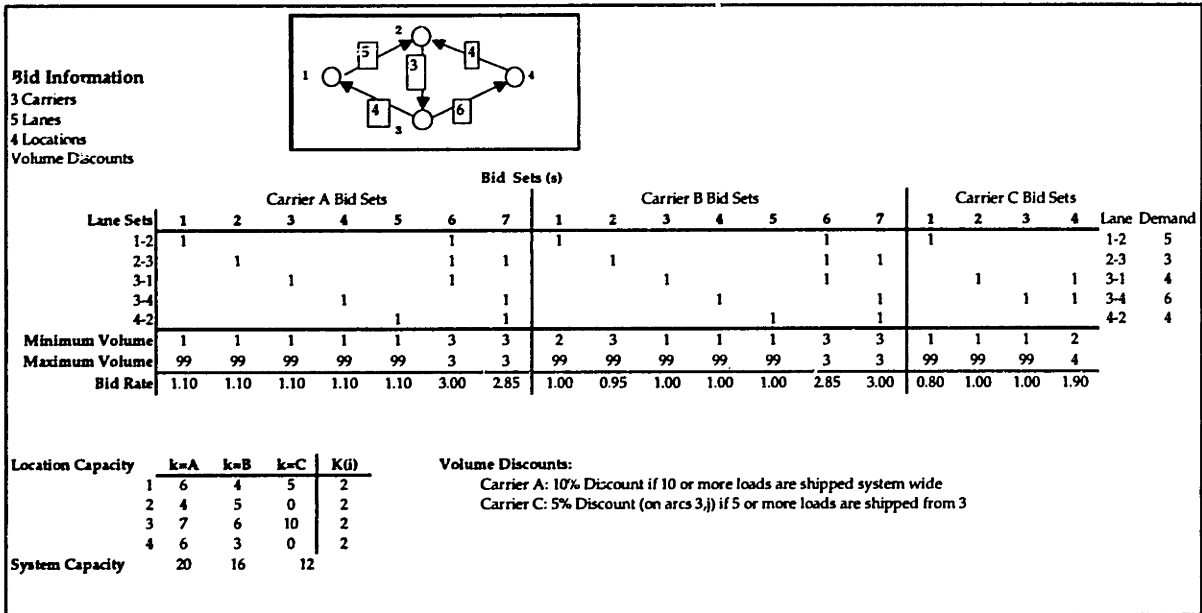


Figure 6.1 Example for carrier-set formulation.

A hypothetical constraint matrix for [CSR] is shown in Table 6.2 for the bid shown in Figure 6.1. Each column in Table 6.2 represents an assignment for a carrier. Carrier A is shown with 3 potential carrier-sets while Carriers B and C have 2 each. There are, of course, many other carrier-sets which could be formed. Recall that the coefficient, a_{ij}^{sk} , represents the number of truckloads that are carried on lane ij by carrier k . The demand constraints are greater than rather than equality constraints in order to maintain integral carrier-set solutions. The matrix is dense because of the small network where carriers serve almost every location. For realistic networks, it will be quite sparse.

| | | Carrier-sets | | | | | | | | RHS | dual |
|--------------|-----|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|-------------|------|
| (1) | 1-2 | a_{12}^{1A} | a_{12}^{2A} | a_{12}^{3A} | a_{12}^{1B} | a_{12}^{2B} | a_{12}^{1C} | 0 | $\geq D_{12}$ | π_{12} | |
| | 2-3 | a_{23}^{1A} | 0 | a_{23}^{3A} | a_{23}^{1B} | a_{23}^{2B} | a_{23}^{1C} | a_{23}^{2C} | $\geq D_{23}$ | π_{23} | |
| | 3-1 | a_{31}^{1A} | a_{31}^{2A} | 0 | a_{31}^{1B} | a_{31}^{2B} | a_{31}^{1C} | 0 | $\geq D_{31}$ | π_{31} | |
| | 3-4 | 0 | a_{34}^{2A} | 0 | a_{34}^{1B} | 0 | a_{34}^{1C} | a_{34}^{2C} | $\geq D_{34}$ | π_{34} | |
| | 4-2 | a_{42}^{1A} | a_{42}^{2A} | 0 | a_{42}^{1B} | 0 | a_{42}^{1C} | a_{42}^{2C} | $\geq D_{42}$ | π_{42} | |
| (2) | 1 | 1 | 1 | 1 | | | | | ≤ 1 | $-\sigma_1$ | |
| | 2 | | | | 1 | 1 | | | ≤ 1 | $-\sigma_2$ | |
| | 3 | | | | | | 1 | 1 | ≤ 1 | $-\sigma_3$ | |
| (3) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | $\leq K_1$ | $-\rho_1$ | |
| | 2 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | $\leq K_2$ | $-\rho_2$ | |
| | 3 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | $\leq K_3$ | $-\rho_3$ | |
| | 4 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | $\leq K_4$ | $-\rho_4$ | |
| Coefficients | | R_1^A | R_2^A | R_3^A | R_1^B | R_2^B | R_1^C | R_2^C | | | |
| Variables | | CS_1^A | CS_2^A | CS_3^A | CS_1^B | CS_2^B | CS_1^C | CS_2^C | | | |

Table 6.2 Constraint matrix of carrier-set formulation.

Carrier-set Generation

There are two problems with generating carrier-sets. First, carrier-sets consist of non-overlapping bid-sets, so we can only consider certain combinations. Working in the existing

network does not allow us to easily maintain a feasible combination of bid-sets. Second, the costing out of a candidate carrier-set cannot be done until it is completely generated, due to the various discounting schemes. Therefore, we need to split the process of generating carrier-sets into two steps: generating feasible carrier-sets and then pricing out to determine good candidate columns to add to the master problem.

Feasible Carrier-sets

Carrier-sets consist of feasible load-based and volume allocations for a single carrier. To be feasible, a carrier-set must:

- not exceed lane demand for any lane within the set,
- not exceed carrier truck supply at any location, lane, or system,
- not have overlapping bid-sets, or
- adhere to upper or lower volume bounds for any lane.

A feasible carrier-set must therefore satisfy the following constraints:

$$\begin{aligned}
 (1) \quad & \sum_p \delta_{ij}^{pk} v_p^k \leq S_{ij}^k && \forall i, j, k \\
 (2) \quad & \sum_j \sum_p \delta_{ij}^{pk} v_p^k \leq S_i^k && \forall i, k \\
 (3) \quad & \sum_i \sum_j \sum_p \delta_{ij}^{pk} v_p^k \leq S^k && \forall k \\
 (4) \quad & \sum_p \delta_{ij}^{pk} x_p^k \leq 1 && \forall i, j, k \\
 (5) \quad & UV_p^k x_p^k - v_p^k \geq 0 && \forall k, p \\
 (6) \quad & LV_p^k x_p^k - v_p^k \leq 0 && \forall p, k \\
 (7) \quad & v_p^k \geq 0, \quad x_p^k = [0,1]
 \end{aligned}$$

where:

Indices:

- i*: Origin shipping location, $i = 1, 2, \dots, m$
j: Destination shipping location, $j = 1, 2, \dots, n$
k: Carrier identification, $k = 1, 2, \dots, p$.
p: Package of lanes

Decision Variables:

- v_p^k : Number of truckloads moving in package p using carrier k
 x_p^k : =1 if carrier k is assigned package p ; =0 otherwise

Data:

- S^k : Supply of truck loads available in system for carrier k
 S_i^k : Supply of truck loads available at location i for carrier k
 S_{ij}^k : Supply of truck loads available for lane i to j for carrier k
 δ_{ij}^{pk} : =1 if package p for carrier k contained lane i to j ; =0 otherwise
 UV_p^k, LV_p^k : Upper and lower truck load volume shipped on package p for carrier k .

These constraints specify that every bid-set within the carrier-set must; (1-3) not exceed lane, location, or system capacity; (4) cover a lane no more than once; (5,6) have total traffic volume within the minimum and maximum bounds set by the carrier; and (7) have non-negative flow. The bounds, UV_p^k , LV_p^k , are set by each carrier as part of the bid, and need to agree with the shippers predetermined bounds. For a dedicated run, where there is a fixed number of loads, $LV=UV$, that is, the carrier cannot change the volume. While this requires that all lanes within a package carry the same volume, this could be easily changed by further delineating the flow variable, v_p^k , to include additional subscripts, ij , to indicate the flow on lane i to j due to package p by carrier k . We will keep the current notation for simplicity.

Two methods for finding carrier-sets are described: complete generation and clique generation.

Complete Generation

While not practical for a large network, generating every possible carrier-set is a starting point. For each carrier k , we need to

1. Define each bid-set with a specific volume as a separate carrier-set.
{max number of CSs generated = $\sum_p [(1 + \min\{D_{ij} - UV_p^k\} - LV_p^k)]$ }
2. Find all feasible combinations of bid-sets (ignoring volume):
No overlapping of lanes => Set Covering Problem
3. Define a separate CS for each possible volume level, for each feasible combination.

Of course, this creates a huge number of columns. For the example with just 3 carriers and 5 lanes, we can form over 3,000 distinct carrier-sets. Factors that contribute to the large number of carrier-sets include the range in the upper and lower volume bounds, the number of distinct bid-sets that a carrier offers, and the number of lanes that each carrier serves. Step 2 can be accomplished by a small routine that generates all possible combinations of bid-sets and throws out those that violate the single-rate constraint (i.e., $\sum_p \delta_{ij}^p \leq 1$). Assigning a volume is more difficult, since we are restricted not only by location but by the whole system as well.

Clique and Undiscounted Rate Method

Another approach is to transform the matrix representing a carrier's bid-sets into a network where each node represents a bid-set and an arc between i and j indicates that they have no overlapping lanes. This transformation allows us to work on a network of bid-sets. Using the original network of origins and destinations, we have no way of maintaining bid-sets since they may not even contain adjacent arcs. The bid-set Network shown in Figure 6.2 is for Carrier A in the numerical example.

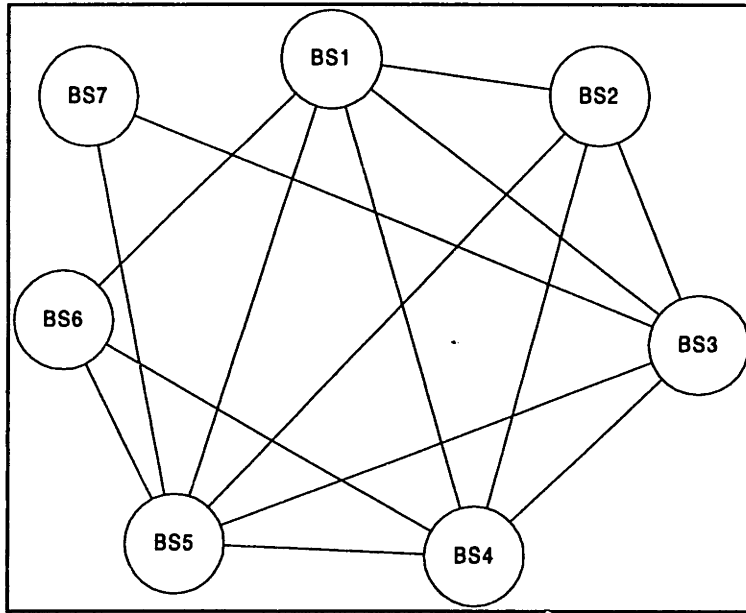


Figure 6.2 Bid-set network for Carrier A in example.

We can check whether a collection of bid-sets makes a feasible carrier-set by seeing if they form a clique. While the maximum clique problem is NP Hard, we are not particularly interested in finding the largest clique. In fact, we really want to find "mid-sized" cliques. The reason for this is that when we add these columns to [CSR] we are going to get fractional solutions. By having smaller cliques we have greater flexibility in identifying the important lane assignments than if only 'full' columns, where a carrier serves all of the lanes, are generated. This is similar to the concept of "simple cycles" used by Barnhart, Hane, Johnson, and Sigismondi (1994).

We can use the ρ_i dual variables to help identify good bid-sets. They correspond to the location capacity constraints, so that the more negative the value, the more desirable a bid-set which services that location might be. Once a feasible collection of bid-sets is found for a carrier-set, we can assign flow. For the time being, ignore all volume discounts. Defining an undiscounted carrier-set rate as $\tilde{R}_s^k = \sum_{p \in s} C_p^k v_p^k$, we can then redefine the undiscounted reduced costs for a carrier-set:

$$\begin{aligned}
 \bar{R}_s^k &= \tilde{R}_s^k - \sum_{i \in (sk)} \left[\sum_j (a_{ij}^{sk} \pi_{ij}) - \rho_i \right] + \sigma_k \\
 &= \sum_{p \in s} C_p^k v_p^k - \sum_{p \in s} \left[\sum_{ij \in p} \pi_{ij} v_p^k - \sum_{i \in p} \rho_i \right] + \sigma_k \\
 &= \sum_{p \in s} \left[\left(C_p^k - \sum_{ij \in p} \pi_{ij} \right) v_p^k + \sum_{i \in p} \rho_i \right] + \sigma_k
 \end{aligned}$$

To illustrate the reduced costs, take as an example, a new carrier-set for carrier A, CS_s^A , that will consist of x_4^A , x_5^A , and x_6^A . The arc dual variables are: $\pi_{12}=6.23$, $\pi_{23}=1.85$, $\pi_{31}=10.0$,

$\pi_{34}=8.72$, and $\pi_{42}=7.51$. The single-rate dual for carrier A is: $\sigma^A=25.64$ while the location supply duals are: $\rho_1=\rho_2=\rho_4=0$, $\rho_3=20.06$. The undiscounted carrier-set rate, as a function of bid-set volumes, becomes:

$$= [(1.10 - 8.72)v_4^A + 20.06] + [(1.10 - 7.51)v_5^A + 0] + [(3.10 - (6.23 + 1.86 + 10.0))v_6^A + (0 + 0 + 20.06)] + 25.64$$

$$= -7.62v_4^A - 6.41v_5^A - 15.09v_6^A + 65.76$$

where the volumes must satisfy the feasibility constraints. A quick way to assign the volumes and maintain feasibility is to assign the maximum flow to the bid-set with the most negative coefficient, bid-set 1 in this example. The maximum value that can be assigned is equal to the minimum of four items:

1. Demand on each arc in the x_p^k , ($=\min\{D_{ij} \text{ for all } ij \in p\}$),
2. Upper Volume Bound for the x_p^k , (UV_p^k),
3. Remaining supply of vehicles in the system, (S_k'), and
4. Remaining supply of vehicles at locations in x_p^k , ($=\min\{S_{ik}' \text{ for all } i \in p\}$).

This value must also be greater than the Lower Volume bound, LV_p , for that bid-set. So the assigned volume needs to satisfy:

$$LV_p^k \leq v_p^k = \min \{ [\min\{D_{ij} \text{ for all } ij \in p\}], [UV_p^k], [S_k'], [\min\{S_{ik}' \text{ for all } i \in p\}] \}.$$

Once a volume is assigned to the most negative bid-set, we update the next most negative one. A modification is to calculate the maximum feasible volume ($=$ minimum of the 4 values) for each potential bid-set first, and then determine which bid-set makes the most improvement. Either way, we end up with volumes assigned to each of the bid-sets contained within the carrier-set. Note that once the bid-sets to include in a carrier-set are determined, the ρ_i values do not change with the assignment of volumes. For our example, we assigned volume in the following order and amounts:

$$v_6^A = \min \{ [\min\{5, 3, 4\}], 3, 20, [\min\{6, 4, 7\}] \} = 3$$

$$v_4^A = \min \{ 6, \infty, 17, 4 \} = 4$$

$$v_5^A = \min \{ 4, \infty, 11, 6 \} = 4$$

So that the undiscounted carrier-set rate $= -7.62(4) - 6.41(4) - 15.09(3) + 65.76 = -35.63$.

In summary, the procedure is as follows for each carrier in the system:

1. Transform the given bid-set matrix into a bid-set Network by making the bid-sets the nodes and connecting them if they do not have overlapping lanes. {This is done only once}
2. Find a clique in the bid-set Network of a moderate size (say 5 nodes). This defines the bid-sets to include in the candidate carrier-set.
3. Calculate the Undiscounted carrier-set Rate using the duals from the previous iteration of the master problem [CSR].
4. Assign feasible flows to each lane set, starting with the one with most negative coefficient, by selecting the minimum of $\{D_{ij}, UV_p, \text{Remaining } S_p, \text{ and Remaining } S\}$ as the volume.

5. Repeat steps 2-4 until 'enough' columns have been generated for carrier k . Then repeat for each other carrier.

This procedure results in a carrier-set with assigned volumes. The actual (discounted) carrier-set Rate, R_s^k , can now be quickly back calculated since the discounts can be easily determined once the volumes are known. Continuing our example, we satisfy the volume discount for carrier A by hauling more than 10 truckloads and receive a 10% discount on rates. Since $\tilde{R}_s^k = \sum_{p \in S} C_p^k v_p^k$ then $R_s^k = (\tilde{R}_s^k)(d) = \sum_{p \in S} (dC_p^k v_p^k)$, where d is the percentage discount. For our case: $R_s^k = [4(\$1.10)+4(\$1.10)+3(\$3.00)] = (17.8) = \16.02 . And the reduced cost for the column becomes: $-35.36 - (.10)(16.02) = -36.97$. Note that the volume discount made this column even more attractive, as expected.

Pricing Out Columns

Once a feasible carrier-set column has been generated, it is easy to determine whether it is worth adding to the master problem. If

$$R_s^k - \sum_{i \in (sk)} \left[\sum_j (a_{ij}^{sk} \pi_{ij}) - \rho_i \right] \leq -\sigma_k + \theta$$

or for our restricted example:

$$\sum_{p \in S} \left[\left(C_p^k - \sum_{ij \in P} \pi_{ij} \right) v_p^k + \sum_{i \in P} \rho_i \right] \leq -\sigma_k$$

then we add that column to [CSR]. If no columns qualify, then we are optimal. For our example, we would add the CS,^A column to the master problem.

6.2.3 Summary

The carrier-set formulation has the potential to capture all of the economies that affect a single carrier into a single column in the master problem. The master problem would be relatively straightforward to solve. Also, it would provide a great degree of flexibility in the formation of different combinations of lane and volume assignments for the carrier.

The implementation, though, is not fully developed yet since it has some potentially difficult problems. First, the sub-problem is extremely time consuming. Finding cliques is not a fast procedure. Second, the integer solution to the master problem was not discussed - and is not trivial. Third, the number of potential carrier-sets is huge - every time the volume awarded to a carrier changes an entirely new carrier-set would be generated. So, while carrier-set formulations present a potentially powerful approach, they are not quite ready for implementation.

6.3 Bid-Set Formulations

This thesis focuses primarily on lane-based assignment models which award specific carriers to lanes as a whole as opposed to load-based models, such as [MWG] and [CS1], which assign a number of loads to each carrier. The selection between assigning by-lane or by-load affects the underlying mathematical model. Assigning carriers to a network by-load, without any system restrictions, can be solved very quickly as a linear program. Assigning by lane requires the solution of an integer program. As will be shown later, though, assignment by-lane can help in finding strong LP solutions.

The decision between assigning by-load versus assigning by-lane is determined by the shipper's strategy. If a shipper wants to have a primary or lead carrier for each lane, assigning by-lane makes sense. Additional carriers can be assigned to a lane as alternates if the shipper deems it necessary. Assigning by-load, on the other hand, delegates the decision of the number of carriers per lane to the model. Essentially, assigning by-lane results in a ranking of carriers for each lane rather than an apportionment of loads per carrier for each lane. Assigning by load separates the real-time decision of how to allocate volume per lane from the tactical decision of which carriers to assign.

All of the lane-based models use the following nomenclature.

Indices:

| | |
|-------|---------------------------------------|
| i : | Shipping location origin |
| j : | Shipping location destination |
| p : | Package of lanes in a conditional bid |
| k : | Carrier identification |

Decision Variables:

| | |
|-----------|---|
| x_p^k : | =1 if carrier k is assigned all lanes in package p , =0 otherwise |
| y_i^k : | =1 if carrier k serves origin i , =0 otherwise |
| z^k : | =1 if carrier k is in the system at all, =0 otherwise |

Data:

| | |
|----------------------|--|
| D_{ij} : | Number of carriers requested on lane i to j |
| L_s : | Maximum number of carriers allowed in the system |
| L_i : | Maximum number of carriers allowed at location i |
| M : | Large constant |
| δ_{ij}^{pk} : | =1 if carrier k 's bid package p contains lane i to j , =0 otherwise |
| c_p^k : | Total annual cost for carrier k to service lane package p |

The cost coefficient for each lane package, c_p^k , is the total estimated cost required to service all of the lanes included in the package, p , for the entire period of the contract. The cost for each lane is calculated as the product of the flat lane rate and the estimated lane volume. Alternatively, if per mile rates are submitted, the cost for each lane is the product of the mileage rate, lane distance, and estimated lane volume. A different lane rate can be applied to each lane

included within a package, but a single combined cost of serving these lanes as a package is used in the assignment model.

6.3.1 Single Primary Carrier per Lane

The initial carrier assignment model [CA1] was formulated to find the primary carriers to assign to each lane.

$$\begin{aligned}
 & \min && \sum_k \sum_p c_p^k x_p^k \\
 & \text{subject to:} \\
 (1) & && \sum_k \sum_p \delta_{ij}^{pk} x_p^k = 1 \quad \forall i, j \quad (\pi_{ij}) \\
 (2) & -My_i^k + \sum_k \sum_p \delta_{ij}^{pk} x_p^k && \leq 0 \quad \forall k, i \quad (\omega_i^k) \\
 (3) & \sum_k y_i^k && \leq L_i \quad \forall i \quad (\theta_i) \\
 (4) & -Mz^k + \sum_p x_p^k && \leq 0 \quad \forall k \quad (\lambda^k) \\
 (5) & \sum_k z^k && \leq L_s \quad (\mu) \\
 (6) & x_p^k = [0,1], y_i^k = [0,1], z^k = [0,1]
 \end{aligned}
 \tag{CA1}$$

The objective is to minimize the total cost of hauling TL traffic over the network. Constraints (1) ensure that each lane is covered by a single carrier; (2) enforce the condition that any carrier assigned a lane from an origin is assigned to that origin; (3) restrict the total number of carriers serving a location; (4) enforce the condition that any carrier assigned a lane is also assigned to the system; (5) restricts the total number of carriers assigned to the system; and (6) states that all variables are binary.

While [CA1] is a valid formulation, it produces extremely fractional LP solutions which make it very weak in solving the IP. Minimizing the constant M in constraints (2) and (4) help produce stronger bounds, but they are still quite weak. A better way to strengthen the problem is to disaggregate constraints (2) and (4). Barnhart et al (1993) have shown that, in many cases, disaggregating the model will lead to tighter bounds when solving the IP. This results in the model [CA2] shown below:

$$\begin{aligned}
& \min && \sum_k \sum_p c_p^k x_p^k \\
& \text{subject to:} && \\
(1) &&& \sum_k \sum_p \delta_{ij}^{pk} x_p^k = 1 \quad \forall i, j \\
(2) &&& -y_i^k + \delta_{ij}^{pk} x_p^k \leq 0 \quad \forall k, ij \\
(3) &&& \sum_k y_i^k \leq L_i \quad \forall i && \text{CA2} \\
(4) &&& -z^k + \delta_{ij}^{pk} x_p^k \leq 0 \quad \forall k, ij \\
(5) &&& \sum_k z^k \leq L_s \\
(6) &&& x_p^k = [0,1], \quad y_i^k = [0,1], \quad z^k = [0,1]
\end{aligned}$$

The objective function and logic of the constraints have not changed. This disaggregation relies on the values of bid variables being binary, that is, the assignment must be by-lane and not by-load.

6.3.2 Multiple Classes of Bids

A further complication is added when we consider the assignment of different classes of bids to the system. Typically, this involves the assignment of a single primary and one or more alternate carriers to each lane. It is assumed that a carrier may serve the same network as both a primary and alternate carrier, if desired, but that they cannot serve the same lane in both roles. The class of each bid is known ahead of time, because carriers will adjust their bids as to whether they are to be the primary or an alternate. In the ABC data, the alternate bids for lanes were either the same or higher than the primary bids. The formulation [CA3] is shown below.

$$\min \quad \sum_c \sum_k \sum_p c_{pc}^k x_{pc}^k$$

subject to:

- (1) $\sum_k \sum_p \delta_{ij1}^{pk} x_{p1}^k = 1 \quad \forall c=1, ij$
- (2) $-y_{i1}^k + \delta_{ij1}^{pk} x_{p1}^k \leq 0 \quad \forall c=1, k, ij$
- (3) $\sum_k y_{i1}^k \leq L_i \quad \forall i$
- (4) $-z^k + \delta_{ijc}^{pk} x_{pc}^k \leq 0 \quad \forall c, k, ij$ CA3
- (5) $\sum_k z^k \leq L_s$
- (6) $\sum_k \sum_p \delta_{ij2}^{pk} x_{p2}^k = A_{ij} \quad \forall c=2, ij$
- (7) $\sum_c \sum_p \delta_{ijc}^{pk} x_{pc}^k \leq 1 \quad \forall ij, k$
- (8) $x_{pc}^k = [0,1], y_{ic}^k = [0,1], z^k = [0,1]$

where:

Indices:

| | |
|-------|---------------------------------------|
| i : | Shipping location origin |
| j : | Shipping location destination |
| p : | Package of lanes in a conditional bid |
| c : | Class of the bid submitted |
| k : | Carrier identification |

Decision Variables:

| | |
|--------------|--|
| x_{pc}^k : | =1 if carrier k is assigned all lanes in package p of class c , =0 otherwise |
| y_{i1}^k : | =1 if carrier k serves origin i , =0 otherwise |
| z^k : | =1 if carrier k is in the system at all, =0 otherwise |

Data:

| | |
|-----------------------|---|
| L_s : | Maximum number of carriers allowed in the system |
| L_i : | Maximum number of carriers allowed at location i |
| δ_{ijc}^{pk} : | =1 if carrier k 's bid package p contains lane i to j of class c , =0 otherwise |
| c_{pc}^k : | Total annual cost for carrier k to service lane package p of class c |

The objective is to minimize the total estimated annual freight bill. Constraints (1) ensure that every lane is covered by exactly one primary (class 1) carrier; (2), (3), (4), and (5) are the same from [CA3] except that the limit on carriers at a location is applied only to the primary carriers, (6) ensure that each lane is covered by a predetermined number of alternative carriers, (7) ensure that a single carrier is not awarded two bids for the same lane (as primary and alternate or as multiple alternates); and, finally, (8) state that all variables are binary.

6.3.3 Additional Considerations and Constraints

The models shown in [CA2] and [CA3] are the basic lane-based, bid-set carrier assignment models. By bid-set, we mean that the decision variables are based on individual bids for lanes or packages of lanes. In the lane-based, bid-set models, the restrictions on the number of carriers in the system are the most prevalent type of constraints that shippers are interested in - especially in light of the core carrier movement. In many cases, the decision to reengineer the TL transportation system is premised on a major reduction of the carrier base. In addition to the limits on the number of carriers within a system or at a location, there are many other considerations which can be included in the models. Some of the more important ones are discussed below.

Service Requirement for Alternates

When the number of carriers within a system is small, it is in the shipper's best interest to retain good relations with all carriers within the carrier base. A shipper can limit the number of carriers within an entire system by requiring that all carriers act as both primaries and alternates over different segments of the system.

The restriction that any carrier serving as an alternate must be a primary somewhere in the system can be implemented with the following constraints. Class 1 are primary and class 2 are alternate bids.

$$\begin{aligned} -z_c^k + \delta_{ijc}^{pk} x_{pc}^k &\leq 0 \quad \forall \text{ classes } c, \text{ carriers } k, \text{ lanes } ij \\ z_1^k - z_2^k &\leq 0 \quad \forall \text{ carriers } k \end{aligned}$$

where:

$$z_c^k: \quad =1 \text{ if carrier } k \text{ is assigned at least one lane with a bid of class } c; =0 \text{ otherwise.}$$

Minimum/Maximum Coverage

A shipper might want to ensure that the amount of traffic that a carrier wins is within a certain bound. The shipper might not want to rely too heavily on a single carrier - thus setting a maximum coverage. Conversely, the shipper might want to give enough business to a carrier to remain a significant customer - thus setting a minimum. Coverage can be measured in terms of lanes won or in total estimated dollar value. The constraints below ensure that carrier k is within the preset volume bounds.

$$\begin{aligned} \sum_p c_p^k x_p^k &\leq \text{MaxFreightVolume}^k && \text{for some } k \\ \sum_p c_p^k x_p^k &\geq z^k (\text{MinFreightVolume}^k) && \text{for some } k \end{aligned}$$

A related restriction required by some carriers is that every carrier must be able to cover all areas served by a location. That is, for a carrier to be assigned to lane ij , it must serve every other location served from location i . This type of restriction essentially transforms all of the lanes originating from i into a single item. If a carrier does not bid on them all, then the partial bids can be screened out. Those carriers that bid on them all can transform these lanes into a single column with the cost coefficient being the sum of all of the individual lane bids multiplied by their expected annual lane volumes.

“Soft” versus “Hard” Constraints for Carrier Limit

The number of carriers in the system or at a location can be restricted by imposing a ‘hard’ constraint of a maximum upper limit or by using a ‘soft’ constraint which imposes a cost on each carrier used. The models discussed thus far, to include MWG, have only used hard constraints. Soft constraints can be used instead by simply modifying the cost function as shown below in [CA4].

$$Z(x_p^k, y_i^k, z^k) = \min \sum_k F^k z^k + \sum_k \sum_i F_i^k y_i^k + \sum_k \sum_p c_p^k x_p^k$$

subject to:

$$\begin{aligned} (1) \quad & \sum_k \sum_p \delta_{ij}^{pk} x_p^k = 1 \quad \forall \text{ lanes } ij \\ (2) \quad & -y_i^k + \delta_{ij}^{pk} x_p^k \leq 0 \quad \forall \text{ carriers } k, \text{ lanes } ij \\ (3) \quad & -z^k + \delta_{ij}^{pk} x_p^k \leq 0 \quad \forall \text{ carriers } k, \text{ lanes } ij \\ (4) \quad & x_p^k = [0,1], \quad y_i^k = [0,1], \quad z^k = [0,1], \end{aligned} \quad \text{[CA4]}$$

where all variables are the same as previous models with the addition of:

F^k : Cost of including carrier k into the system and
 F_i^k : Cost of including carrier k serve location i .

These fixed costs can be carrier and location specific, as shown in [CA4], or uniform. Essentially, these fixed costs act as penalties for each carrier added.

In practice, soft constraints are rarely used. The [MWG] formulation, for example, uses hard constraints to limit the number of carriers allowed in the system and at a particular location. The reason for this is that, in reality, no one knows the fixed cost of adding an additional carrier. The transportation quality manager at one company, James River Corp., estimated the cost of maintaining a carrier to the system to be slightly under \$5,000. This is a very small amount of money considering a carrier can haul freight worth millions of dollars per year.

Inclusion of Level of Service

While all the models minimize cost, the level of service provided by a carrier can be incorporated into the decision. The optimization model minimizes the selected bids based on the coefficient, c_p^k , which has, until this point, been considered to be the bid price submitted by the carrier for the lanes contained in package p multiplied by the estimated volume on these lanes over the period of the contract. This coefficient, however, can be modified in order to capture a specific carrier's service on a specific lane (or set of lanes). The coefficients can be modified in two different ways: by a multiplicative factor or an additive factor.

A rate multiplier based on service level can be constructed based on subjective or objective rankings of the carrier. For example, suppose a shipper defines service as consisting of three components: on-time arrivals over each lane, rate of refusal of carriers from each location, and ease of doing business with a carrier. Carriers can be ranked according to these objective and subjective criteria for each location, lane, or for the entire system. The rankings can then be combined into a composite multiplier which is applied to the appropriate bids for each carrier.

A second way of incorporating the level of service into the models is to add the known costs to the respective coefficients. For example, suppose that a carrier offers to provide 3 day service, using driver teams, on a lane which typically has a 5 day transit time standard. If the value of saving 2 days can be estimated for the aggregate shipments on this lane, then this amount (or a percentage of it) can be subtracted off of the carrier's bid. In this way, the carrier is being rewarded by providing faster service. Similarly, reliability of service can be included where the inventory carrying cost of the required safety stock can be factored into the cost coefficients. Any additional costs that can be allocated to a specific lane can be added to the respective bids. Savings or costs that can only be allocated to a location or to the carrier as a whole can be included in the objective function as fixed costs using the respective indicator variables, y_i^k and z^k . These coefficients modify the duals for the specific carrier and carrier-location in question.

Regardless of which method is used, only the assignment is based on modified coefficients. The paid rate is still the amount bid. The modification of the coefficients allows carriers to bid not just on price but also on service level - either historical or promised.

6.4 Implementation Models

This section describes the implementation of the lane-based bid-set models, [CA1], [CA2], [CA3] and [CA4], to the ABC TL bid data. The data are from a TL bid conducted in 1992 by a third party firm, referred to as ABC, for a set of about two dozen clients (shippers). The network consists of 2120 identified traffic lanes between 530 shipping locations. A total of 350

carriers submitted 73,138 simple (single lane) bids. The bids were flat rates and were submitted as either primary bids (50,158) or as alternates (22,980). Each traffic lane was assigned a single primary and one to two alternates.

Prior to using the models on the ABC network, the data had to be modified by adding potential lane packages. Section 6.3.1 describes how the data were modified and shows why the use of conditional bids is warranted for the ABC bid. Section 6.3.2 describes the solution strategies taken to find integer solutions to the problem within reasonable amount of time. Section 6.3.3 discusses the business implications of certain policies using the lane-based models and, finally, section 6.3.4 outlines how the model can be used within a multiple round auction.

6.4.1 Modification of Data

Because ABC only allowed simple TL bids, all potential efficient aggregations needed to be created. This was done by using the algorithms described in Chapter 5. Four types of efficient aggregations were created using the primary bids: reloads, open tours, closed tours, and local tours. The alternate bids were not considered for lane packages. The cost for each alternate bid was turned into an annual cost by multiplying the bid times an assumed 10% of the annual volume. This was done to instill scale between the primary and the alternate bids. In practice, the actual number of loads that the primary carrier defaulted on in previous years could be used.

Two parameters determine the type and extent of the lane packages formed: minimum criteria for forming a package and the amount of savings that the carrier would be expected to pass on to the shipper due to winning a conditional bid. The minimum criteria determines the number of packages formed so that lowering the minimum criteria for a lane package (in terms of volume or distance) increases the number of packages formed. The number of lane packages identified is only dependent on the network data and does not rely on the submitted bids. The savings value from each lane package, on the other hand, determines their desirability over the single-lane bids. Increasing the savings (percentage of total bid) for lane packages will increase the likelihood of it being assigned. By modifying these two parameters, different bidding scenarios can be formed. Contention between lanes and bids increases as the number of and the savings from lane packages increase. The high contention scenarios were used in the model runs.

Additionally, the raw bid data were screened. The variation in the bids was extremely high and many of the lanes were bid on by a single carrier. Two sets of data were created from this raw data: the complete data set and the reduced data set. Table 6.3 shows the size characteristics for both sets.

| | Reduced Data Set | Complete Data Set | Original |
|--------------------------|------------------|-------------------|----------|
| Carriers | 92 | 300 | 350 |
| Lanes | 1739 | 2099 | 2120 |
| Primary Single Bids | 25,350 | 50,007 | 50,158 |
| Primary Conditional Bids | 2,921 | 8,583 | 0 |
| Alternate Bids | 12,900 | 22,923 | 22,980 |
| Total Bids | 28,271 | 81,513 | 73,138 |

Table 6.3 Comparison of data set size.

The complete data set was created by removing lanes with 2 or fewer bids on them. Carriers were eliminated if their bids were dominated (under bid) by another carrier. The reduced data set was created by removing all of the bids which fell outside of 1.5 standard deviations from the mean for that lane. This step was repeated two times and finally, all carriers that did not have at least 50% of their bids within 1 standard deviation were removed. Additionally, any lanes with less than 5 primary bids were removed from the network. The reduced data set is actually more contentious than the complete data set due to the reduced variance between lanes. As Bykowsky, Cull, and Ledyard (1995) noted, contention increases as the relative difference between bids decreases.

In the course of the analysis, all of the lanes that were identified as potentials for threads were also found to be candidates for the two lane reloads. Because of this, all open tours were restricted to 2 lanes. Since the savings for a reload were always greater than that for an open tour, it made no sense to include both. The number of lane packages identified and the number of bids submitted on each are shown in Table 6.4. The total number of conditional bids identified was around 10% for both sets of data.

| | Reduced Data Set | | Complete Data Set | |
|------------------------|------------------|------|-------------------|------|
| | Identified | Bids | Identified | Bids |
| Closed Tours | 257 | 555 | 306 | 2534 |
| Open Tours/Reloads | 340 | 2312 | 402 | 5898 |
| Local Tours | 12 | 54 | 17 | 151 |
| Total Conditional Bids | | 2921 | | 8583 |

Table 6.4 Comparison of data set size.

Before we can justify the additional costs of using conditional bids in a combinatorial auction, we need to examine the results of a bid that used a supposedly inferior method. The next section analyzes the actual results of the ABC TL bid.

Need for Conditional Bids

The primary criteria that we can use to evaluate an auction is its efficiency. An auction should allocate the items to those buyers who value them the most. Without detailed information on the reservation prices and internal valuations of each of the involved carriers,

however, it is not possible to determine the true allocational efficiency. What we can do is observe the results and see if obvious inefficiencies occurred. Because this is a multiple item auction with interdependencies, the auction should try to achieve efficient aggregations of the various lanes. So, we can evaluate the efficiency of the auction by examining how well combinations of lanes were formed and whether they were won intact by single carriers.

An efficient auction leads to the efficient aggregation of complementary items. That is, those items which had the maximum value in combination with each other would end up being won by the same bidder. For the ABC TL bid, we identified those closed-tours in which each lane on the tour had at least 250 loads - most likely daily loads. Five such tours existed, as shown in the table below. It was thought that for these very heavily traveled traffic lanes that a single carrier would be able to cover all lanes in a tour since they constituted a closed tour. This ensured some guarantee of balance. If a single carrier won all of the lanes in a tour, this would indicate an efficient outcome. If multiple carriers won the lanes within a tour, this would indicate that the efficient aggregation of lanes was not achieved. Table 6.5 shows the outcome for these high density tours.

| Tour | Lane Numbers | Annual Volume | Winning Carriers |
|-------------|---------------------|----------------------|-------------------------|
| 1 | 377 | 267 | AAA |
| | 824 | 372 | BBB |
| 2 | 475 | 1260 | CCC |
| | 701 | 360 | CCC |
| 3 | 477 | 327 | EEE |
| | 787 | 312 | CCC |
| 4 | 477 | 327 | EEE |
| | 797 | 327 | FFF |
| | 838 | 306 | GGG |
| 5 | 797 | 327 | FFF |
| | 844 | 630 | FFF |

Table 6.5 Analysis of ABC TL bid.

Just two of the five tours resulted in an efficient aggregation. But, these tours were the most unbalanced. The two tours (3 & 4) with annual volumes within 10% of each other were, surprisingly, not aggregated.

Additionally, a set of reload and local tours were selected using very high minimum criteria standard so that only the very strongest and most obvious packages would be identified. A total of 10 single-lane to single-lane reload opportunities and 3 local tour locations were identified. Of the 10 reloads, only 2 had both inbound and outbound lanes won by the same carrier. For the 3 local tours, only 1 had the same carrier winning more than 50% of the lanes.

Thus, the simple-bid auction did not appear to encourage the formation of efficient lane packages. The 5 closed tours, 10 reloads, and 3 local tours identified here are some of the strongest potential lane packages, so this is a conservative analysis. Therefore, the use of conditional bids for the ABC TL bid seems justified.

6.4.2 Solution Strategies

The models were run on Pentium based PCs using the CPLEX Optimization Software's callable library.

Initial model runs were made solving [CA1] as an LP with just the covering constraints for the primary carriers. Even with the contention between the bid packages, the solutions were almost entirely integer. The savings values for packages had to be raised significantly to result in any fractional solutions. In these non-restricted runs with the reduced data set, all 92 of the carriers were used even though the minimum feasible number of carriers is 12. Adding the system and location limit constraints destroyed the integrality of the solutions, however. In fact, efforts to find an integer solution to [CA1] were not successful for problems restricting the number of carriers below 50. Several different branching strategies and starting heuristics were tried which improved the initial solution, but none were tremendously successful.

The disaggregated model [CA2] however, provided much tighter bounds for the IP. While the finding the solution of the root node can take much longer⁶³, in many cases the solution is integral. The size also increased the number of constraints by 400%. Table 6.6 lists the size of the constraint matrices for the aggregated and disaggregated models for both the complete and the reduced data sets. The size of the [MWG] model is also shown.

| | Reduced Data Set | | | Complete Data Set | | | MWG |
|-------------------------|------------------|---------|---------|-------------------|---------|---------|--------|
| | CA1 | CA2 | CA3 | CA1 | CA2 | CA3 | |
| Rows | 13,039 | 53,705 | 81,291 | 21,737 | 102,504 | 154,526 | 5,000 |
| Columns (integer) | 41,394 | 41,394 | 55,396 | 79,936 | 79,936 | 104,684 | 200 |
| Columns (continuous) | 0 | 0 | 0 | 0 | 0 | 0 | 9,000 |
| Elements | 117,049 | 161,767 | 245,799 | 242,712 | 338,122 | 481,065 | 26,000 |
| Carriers | 92 | 92 | 92 | 300 | 300 | 300 | 15 |
| Lanes | 1739 | 1739 | 1739 | 2099 | 2099 | 2099 | ? |
| Bids | 28,271 | 28,721 | 41,171 | 58,590 | 58,590 | 81,513 | ? |

Table 6.6 Comparison of problem sizes.

Comparisons between the aggregated [CA1] and the disaggregated [CA2] models were made by running parametric analyses on both the carrier system and location limits. Figures 6.3

⁶³ The importance of the optimization algorithm within the software was seen in the different solving times for the initial root node. Solving the Complete Data set for 60 carriers took over 28 hours with Primal Simplex and less than 2 and a half hours with the dual optimization.

and 6.4 show the duality gap for the two parametric model runs for both the aggregated and the disaggregated models. While the disaggregated model has a much smaller duality gap (in many of the cases a gap of zero) the total duality gap of even the worst case for the aggregated model is only 4%!

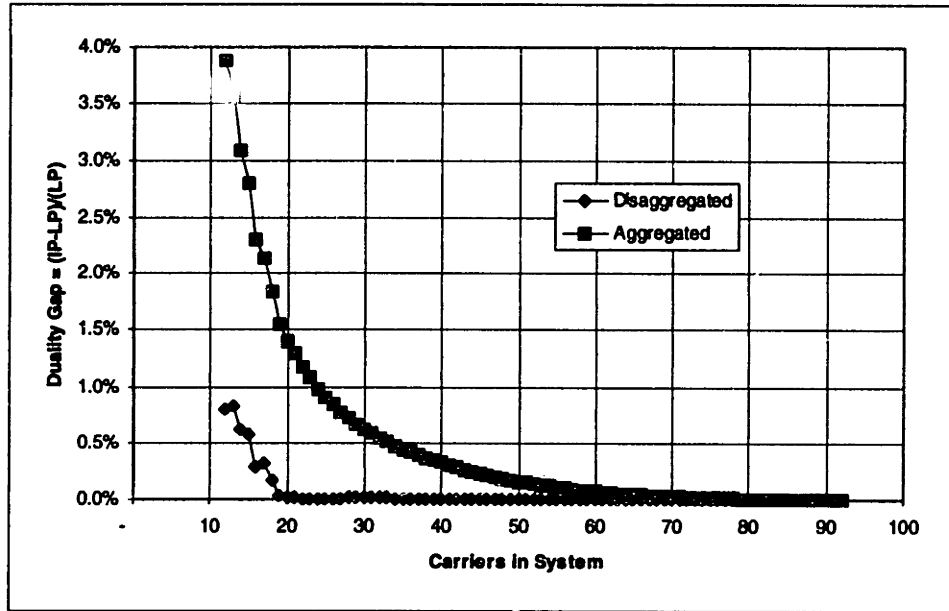


Figure 6.3 Duality gap comparison between aggregated and disaggregated models for parametric analysis on carrier system limits.

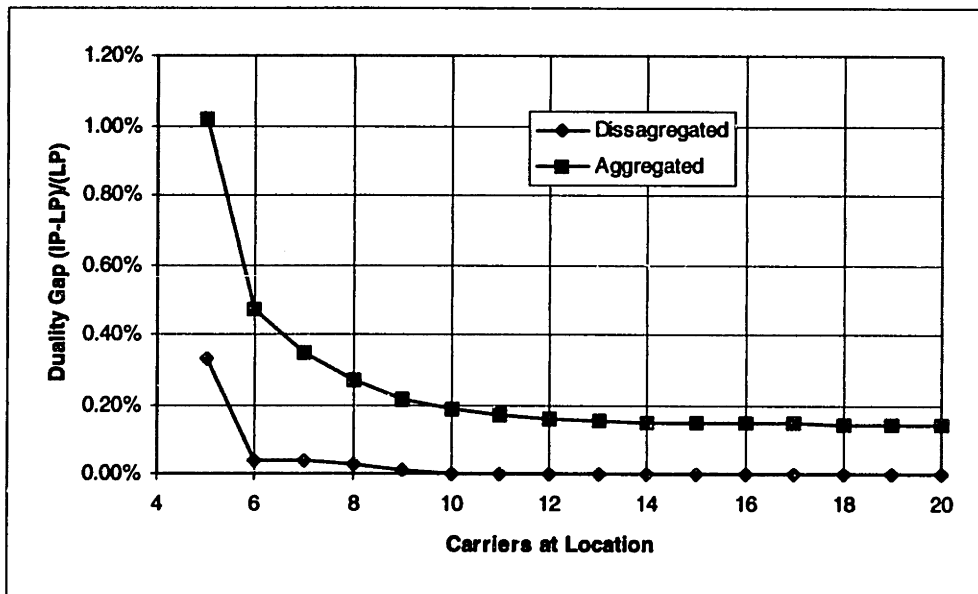


Figure 6.4 Duality gap comparison between aggregated and disaggregated models for parametric analysis on location limits.

While the duality gaps are not large for either model, the integrality of the disaggregated model is always greater and almost double in some cases. The percentage of integer columns is defined as the number of columns of a particular type (carrier, location, or

bid) that are fractional divided by the sum of all non-zero basic variables. Note that this is a more conservative measure of integrality than if all basic variables were included as being integral. Figures 6.5 and 6.6 shows the percentage of integer columns for both parametric runs for both types of models. The behavior seems to be very data specific. For example, there is a sudden drop of integrality when reducing system carriers from 51 to 50, as shown in Figure 6.5. The disaggregated model displays much higher levels of integrality than the aggregated model.

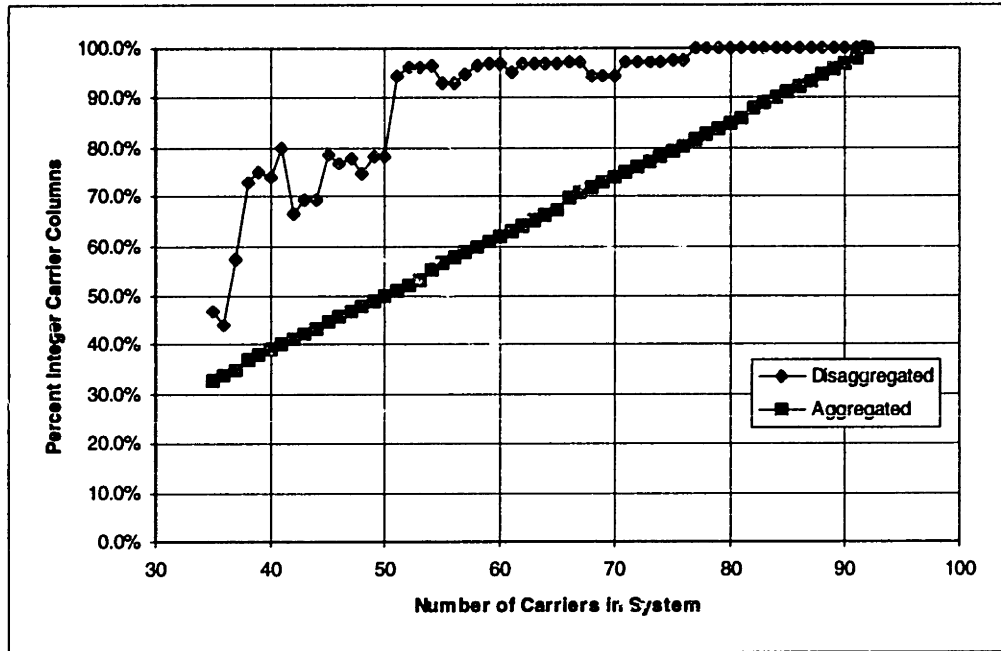


Figure 6.5 Percentage of integer carrier columns comparison between aggregated and disaggregated models for parametric analysis on carrier system limits.

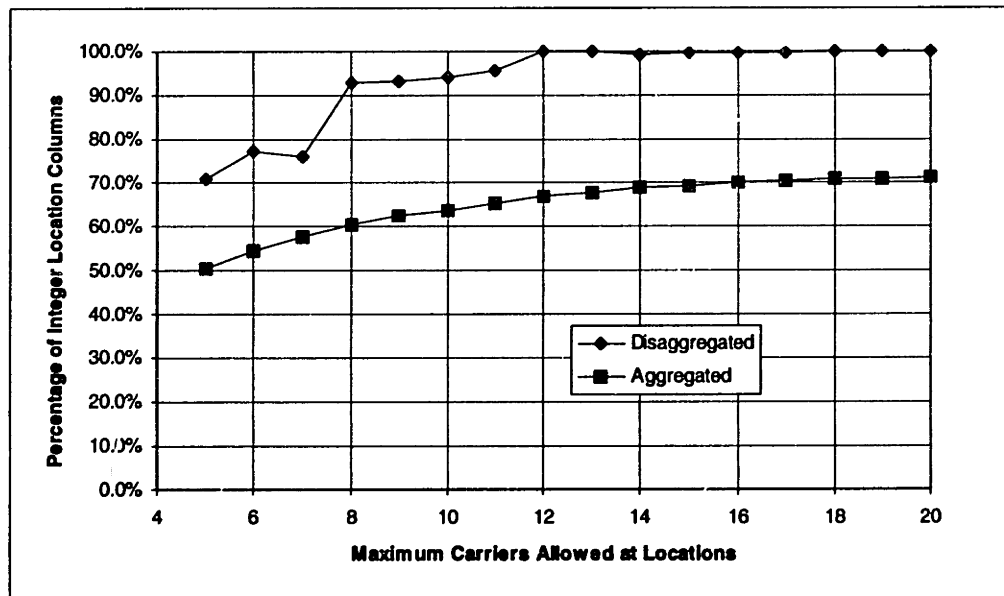


Figure 6.6 Percentage of integer location columns comparison between aggregated and disaggregated models for parametric analysis on location limits.

Even for those runs where the disaggregated model did not provide an integer solution at the root node, the optimal integer solution was usually less than 10 nodes into the branch-and-bound tree. A branching order strategy was set to branch on carrier columns first, followed by location columns, and, finally, bid columns. The initial direction was set to the lower value since it is the stronger of the two branches. Setting the value of a carrier column to zero excludes the carrier from the system while setting it to one just makes that carrier available - not necessarily assigned. Essentially, the strategy was set as a depth first search. This appeared to do quite well except as the iterations approached the minimum feasible number of carriers allowed.

Solving model [CA3] with the primary and alternate carriers did not seem to affect the integrality of the problem - in fact it increased the integrality for most runs! Running a parametric analysis restricting the number of carriers in the system, the root LP node of [CA3] was integral up until the number of carriers was reduced to 40. Even at this point the integer solution was found at the third node of the branch-and-bound tree. The solution time, however, was much longer as the problem had 60% more rows, 30% more columns, and 50% more non-zero matrix elements. Solution times for the root node LPs took on the order of 2 to 8 hours compared to 5 minutes to an hour for [CA2] with the similar system and location limits.

Overall, the disaggregated models appeared to perform quite well for even the larger problems with the full data. The formulation seems sufficient to solve problems of a fair size in an acceptable period of time.

6.4.3 Business Implications

While solving the carrier assignment problem is the primary objective of these models, they are also useful for analyzing certain business decisions. For example, the practice of reducing the carrier base has been widely embraced, but there has not been any analysis in the literature that looks at the implications or costs incurred.

Costs of Carrier Reduction

While the model cannot determine the cost of removing a carrier from the base, it can provide a threshold which the savings must exceed in order to justify its removal. The savings can be created by the rate reductions offered by the remaining carriers due to their increased volume or by lower operational costs due to less coordination of a smaller carrier base.

The loss due to unused carriers can be estimated by examining the dual of the system carrier limit constraint. This is equivalent to the fixed cost that each carrier would need to be charged in order to result in the same number of carriers in the system. To see that this is the

same value, consider the Lagrangian Relaxation of the model [CA2], shown again below along with the dual variables written in parentheses next to each constraint.

$$Z(x_p^k, y_i^k, z^k) = \min \sum_k \sum_p c_p^k x_p^k$$

subject to:

$$\begin{aligned} (1) \quad & \sum_k \sum_p \delta_{ij}^{pk} x_p^k = 1 \quad \forall \text{ lanes } ij \quad (\pi_{ij}) \\ (2) \quad & -y_i^k + \delta_{ij}^{pk} x_p^k \leq 0 \quad \forall \text{ carriers } k, \text{ lanes } ij \quad (\lambda_{ij}^k) \\ (3) \quad & \sum_k y_i^k \leq L_i \quad \forall \text{ locations } i \quad (\theta_i) \\ (4) \quad & -z^k + \delta_{ij}^{pk} x_p^k \leq 0 \quad \forall \text{ carriers } k, \text{ lanes } ij \quad (\omega_{ij}^k) \\ (5) \quad & \sum_k z^k \leq L_s \quad \text{System Limit} \quad (\mu) \\ (6) \quad & x_p^k = [0,1], \quad y_i^k = [0,1], \quad z^k = [0,1] \end{aligned} \quad [\text{CA2}]$$

The location limit (3) and the system limit (5) constraints can be relaxed to form the Lagrangian Relaxation of the problem:

$$L(\mu, \theta_i) = \min \sum_k \sum_p c_p^k x_p^k + \mu (\sum_k z^k - L_s) + \sum_i \theta_i (\sum_k y_i^k - L_i)$$

subject to:

$$\begin{aligned} (1) \quad & \sum_k \sum_p \delta_{ij}^{pk} x_p^k = 1 \quad \forall \text{ lanes } ij \quad (\pi_{ij}) \\ (2) \quad & -y_i^k + \delta_{ij}^{pk} x_p^k \leq 0 \quad \forall \text{ carriers } k, \text{ lanes } ij \quad (\lambda_{ij}^k) \\ (4) \quad & -z^k + \delta_{ij}^{pk} x_p^k \leq 0 \quad \forall \text{ carriers } k, \text{ lanes } ij \quad (\omega_{ij}^k) \\ (6) \quad & x_p^k = [0,1], \quad y_i^k = [0,1], \quad z^k = [0,1], \end{aligned} \quad [\text{LCA}]$$

The last two terms in the objective function of [LCA] are less than or equal to zero due to complementary slackness conditions. Due to weak duality, $L(\mu, \theta_i) \leq Z(x, y, z)$ and strong duality at optimality, $L^*(\mu^*, \theta_i^*) = Z^*(x^*, y^*, z^*)$. Rearranging the objective function of [LCA], we obtain:

$$\begin{aligned} L(\mu, \theta_i) &= \min \sum_k \sum_p c_p^k x_p^k + \mu (\sum_k z^k - L_s) + \sum_i \theta_i (\sum_k y_i^k - L_i) \\ &= \min \sum_k \left(\sum_p c_p^k x_p^k + \sum_i \theta_i y_i^k + \mu z^k \right) - \mu L_s - \sum_i \theta_i L_i \end{aligned} \quad [\text{LCA}']$$

So, the resulting objective function [LCA'] is equivalent to the objective function of the disaggregated model with 'soft' constraints, [CA4], shown below. The last two terms of [LCA'] are constants and therefore do not affect the solution, just the objective function value.

$$\begin{aligned}
Z(x_p^k, y_i^k, z^k) &= \min \sum_k F^k z^k + \sum_k \sum_i F_i^k y_i^k + \sum_k \sum_p c_p^k x_p^k \\
&= \min \sum_k \left(\sum_p c_p^k x_p^k + \sum_i F_i^k y_i^k + F^k z^k \right)
\end{aligned}$$

The relaxation shows that the loss due to unused carriers can be estimated by examining the duals of the system carrier limit constraint. This provides the marginal cost of removing one more carrier from the system, but it should be approached with caution, however, since the relaxation relies on continuous variables and our problem is an IP. There will most likely be a duality gap where $L_{ip}^*(\mu^*, \theta^*) \leq Z_{ip}^*(x^*, y^*, z^*)$. Additionally, duals are most likely not available. However, the problem can be fixed, that is, the values on the variables from the optimal IP problem can be set as the limits for the problem as an LP, and then solved as a linear program. This allows pseudo-dual variables to be determined which can provide some information. Dual information from these models is still suspect, however.

Another way of determining the marginal loss value is to run a parametric analysis on the cost coefficients for the carrier columns and track the number of carriers included in the solution. Figure 6.7 shows that for these data, the two are an almost perfect match.

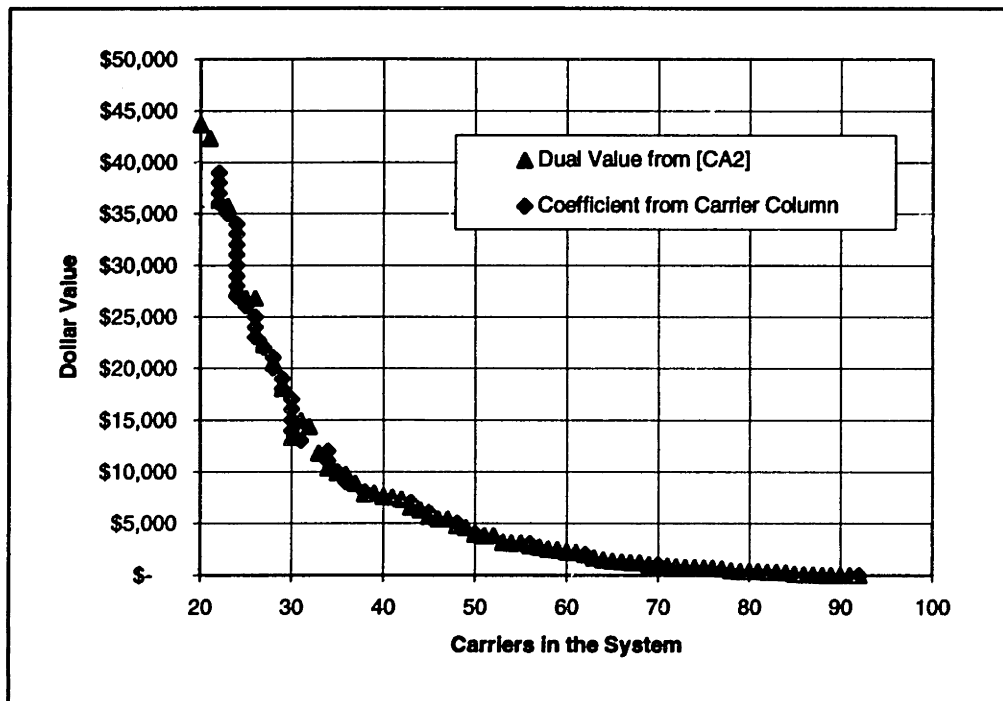


Figure 6.7 Comparison of marginal loss taken from dual of carrier system constraint and coefficient of carrier column.

The total loss due to the reduction of carriers from its 'natural' or unrestricted base size is simply the sum of these marginal values. It is the difference between the unrestricted objective function and the final model's objective function value. For the reduced data set, trimming the carrier base from its original number of 90 to the minimum feasible size of 12

results in an increased cost of \$1.4 million dollars. This is an increase of about 4%. For the complete data set, a reduction from the natural base of 240 carriers to the minimum feasible base of 34 results in an increased cost of \$4 million or about 11%.

These values can be interpreted as the foregone savings due to restricting the carrier base. The savings from a core carrier program should be larger than this value in order to justify ignoring these additional carriers. The savings could be in the form of rate reductions or improved coordination. While these numbers are very data-specific, the procedure is robust.

Effect of Conditional Bids on Assignment

While conditional bids affect the carriers' bidding behavior by reducing the need to hedge, they do not have a pronounced effect on the assignment model. In most of the models run, about 10% of the lanes were won under conditional bids. This is roughly equivalent to the percentage of efficient aggregations formed using the algorithms from Chapter 5. The distribution of the size of the winning conditional bids is also fairly stable with two lane reloads accounting for the majority of the conditional bids won. The distribution moves slightly towards larger conditional bids winning when the restrictions are tightened, but not always. The carrier winning the conditional bid does change slightly between iterations when the system carrier restrictions are tightened. These are very data specific observations, however, so no clear insights were gained. Additionally, the efficient aggregations that were used for these conditional bids were generated using the information available to the shippers. In practice, the shipper will not know the affect that the carrier's other business will have on the assignment.

Overall, the conditional bids did not appear to make the models more difficult to solve. This is most likely due to the emphasis placed on limiting the number of carriers in the entire system which creates a strong interdependency between each carrier's bids. This value dwarfs the lane interdependencies involved. If, however, the model was run on a network where all carriers that bid would be awarded some portion of the network, the importance of the lane interdependencies would most likely rise to prominence. For example, contention will be critical when reassigning lanes to the dozen carriers within a core carrier program each year.

6.4.4 Multiple Round Auctions

Multiple round auctions are useful when the value of the items, or the combined value of sets of items, are not known very well. While the individual traffic lanes within a TL bid might be well understood, the pricing of lane packages is probably not. Because the initial use of conditional bids involves carriers to price lanes in a new way, multiple round bids might be useful for carriers in that it will let them use market information to guide their pricing decisions.

For the shipper, this will reduce the carriers' regret and minimize the winner's curse, as discussed in chapter 4. Multiple rounds with conditional bids are complicated, however.

Multiple round auctions using simple bids work by providing the current winning price for each lane to the carriers at the end of each round. The winning bid for each lane contains all of the information that a carrier needs in order to submit a new bid. For simple-bid auctions, the price for a lane is assumed to be independent of (1) the prices of the other lanes, (2) the identity of the carrier submitting the bid, and (3) the other lanes that that carrier has won.

For conditional bids, however, the current winning lane price is not always obvious since a lane may be included as part of a lane package. The lane price is not necessarily unique - it has to be determined as part of the package. Also, a restriction on the number of carriers makes the winning bid carrier-specific, that is, a carrier already in the system will not need to bid as low as a carrier not included in the system. The same holds for location limits. The effect of location limits and multiple round bidding is that each carrier will have an incentive to gain as much business as possible from each location in order to amortize the cost of serving there in the first place.

The primary problem that the shipper faces is to determine what information to provide to the carriers that is relevant for their pricing decision for the next round. One approach to this is to treat the entire auction as a column generation problem. In this framework, the shipper solves the master problem, [CA2], and the carriers act as separate sub-problems.

$$Z(x_p^k, y_i^k, z^k) = \min \sum_k \sum_p c_p^k x_p^k$$

subject to:

- (1) $\sum_k \sum_p \delta_{ij}^{pk} x_p^k = 1 \quad \forall \text{ lanes } ij \quad (\pi_{ij})$
- (2) $-y_i^k + \delta_{ij}^{pk} x_p^k \leq 0 \quad \forall \text{ carriers } k, \text{ lanes } ij \quad (\lambda_{ij}^k)$
- (3) $\sum_k y_i^k \leq L_i \quad \forall \text{ locations } i \quad (\theta_i)$
- (4) $-z^k + \delta_{ij}^{pk} x_p^k \leq 0 \quad \forall \text{ carriers } k, \text{ lanes } ij \quad (\omega_{ij}^k)$
- (5) $\sum_k z^k \leq L_s \quad \text{System Limit} \quad (\mu)$
- (6) $x_p^k = [0,1], \quad y_i^k = [0,1], \quad z^k = [0,1]$

The carriers submit potential columns (bids) to the shipper. The shipper uses these bids to solve the master problem. This produces dual information on the constraints which is given back to the carriers in order to make better bidding decisions for the next round.⁶⁴ The duals can

⁶⁴ The cautions of using the psuedo-dual information apply to this approach. Additionally, there is the complication that the carriers really want information on making sets of changes, not changes on a single bid at a time - which is what this approach provides.

be used to price-out potential bids. For [CA2] a newly submitted bid by carrier k for lane package p needs to satisfy the following conditions to be considered in the next round:

$$\begin{aligned}\hat{c}_p^k &< \sum_{ij \in p} (\pi_{ij} + \lambda_{ij}^k + \omega_{ij}^k) \\ \mu &< \sum_{ij} \lambda_{ij}^k \\ \theta_i &< \sum_j \omega_{ij}^k\end{aligned}$$

These are simply the optimality conditions for [CA2], or more exactly, the conditions where a new column would make the current solution non-optimal. Essentially, this step provides carriers with information they can use to price out their bids similar to the way columns are priced out for entry into a master problem. The first condition states that the new bid from carrier k for package p must be lower the sum of the duals for the relevant constraints. For a carrier already serving within the system, the value of ω_{ij}^k will be less than or equal to zero. Similarly, for carriers serving at location i , the value of λ_{ij}^k will be less than or equal to zero. Thus, carriers already in the system or at a location receive a “rebate” of ω_{ij}^k or λ_{ij}^k which in effect lowers the price that their bid needs to be to win.

This information could be passed to the carriers in the form of current lane values and system and location rebates. The rebates are both carrier and location specific, so the information would need to be customized for each carrier. At the end of a round, the carrier could be presented with a system rebate, if they have won any lanes, and location rebates for each location they serve. To see the minimum acceptable bid that they can submit for a package, simply sum up the respective lane values, location rebates, and system rebate. Their bid must be below this to be considered.

6.5 Closure

The objective of this chapter was to formulate, solve, and interpret various carrier assignment models.

The carrier assignment model is at the heart of any optimization-based bidding process. Whenever an auction uses conditional bids or includes system considerations (such as limits on the number of carriers), an optimization model is required. While the formulations will differ in the details between shippers, the general form will remain the same. The chapter distinguished between assigning carriers by-load and by-lane where the former apportions the traffic on each lane while the latter ranks the winning carriers.

Another distinction made in the carrier assignment models is between using bid-sets and carrier-sets. Bid-set models treat each conditional bid submitted by a carrier as a unique

variable in the optimization model. Carrier-set models treat each possible lane allocation to a carrier as a single variable. The bid-set models allow for the use of conditional bids and reduces the carrier assignment problem to a set partitioning problem. The carrier-set model is much more flexible, can incorporate volume based pricing in addition to conditional bids, and is formulated as a column generation problem. Both model forms are discussed but only the bid-set models were implemented.

A series of by-lane, bid-set carrier assignment models were implemented using data from an actual TL bid. The final models were shown to produce integral LP solutions in the majority of the cases. The ability of the models to solve very large assignment problems in a PC based environment within a reasonable time was demonstrated. Finally, the models were shown to be of use in analyzing the value of 'lost bids' due to carrier base reductions.

Chapter 7

Synopsis, Contribution, and Extensions

This thesis addressed the question of how shippers should procure transportation services from TL carriers. The majority of prior research and practice in this area has focused on adapting generic purchasing strategies to transportation applications. Reducing the number of carriers used and applying rigorous certification programs for carriers, are all initiatives that were originally applied across the procurement field. Even when procurement strategies have considered specific qualities of transportation, such as transit time or delivery reliability, the decisions have been applied on the macro-level. That is, the procurement strategy assesses whether a carrier, as a whole, should or should not be used by the shipper. While beneficial in many respects, these initiatives treat all products and services alike and, therefore, ignore the unique characteristics of TL carriers.

We approached the procurement problem from a different, more micro-level, perspective to determine not just the best set of carriers to use, but the optimal assignment of carriers to lanes within the network. By examining the specific economics of the carriers (rather than just assuming economies of scale, as do most generic policies), we identified opportunities to reduce the carriers' costs which in turn can lower the shipper's costs. Recognizing that TL carriers are influenced more by economies of scope than economies of scale, we investigated how a shipper can positively influence a TL carrier's cost structure - other than by simply allocating greater volume. While there is very little that the shipper can do to reduce the carriers' line-haul costs, the shipper can lower the carriers' connection costs by decreasing the degree of connection uncertainty. Because this uncertainty causes carriers to hedge their bids, we proposed the use of conditional bids and optimization-based carrier assignment models. Essentially, we recommend redesigning the way in which shippers and carriers interact in order to remove costs from the transportation process.

This chapter is organized into three sections. Section 7.1 synthesizes the thesis and provides a synopsis of the research. Section 7.2 summarizes the contributions of the research. Finally, Section 7.3 briefly discusses some additional areas in which this research can be extended.

7.1 Synopsis

This section provides a narrative description of the thesis. Starting with the analysis of motor carrier operations, it describes how the economics of a transportation mode should influence the choice of shipper-carrier relationship, the design of the bidding process, and the formulation of the carrier assignment model.

Carrier Economics

Chapter 1 classified transportation systems as being either direct or consolidated. TL carriers, operating over irregular routes and moving from origin to destination without any intermediate stops for load consolidation, are a direct mode. Consolidated carriers, such as LTL and package delivery carriers, require the use of terminals and scheduled routes to collect smaller shipments and combine them into larger loads.

The cost structures for direct and consolidated modes are very different. The majority of the costs for direct modes can be divided into two activities: line-haul movement and connection to a follow-on load. Line-haul movement costs are mainly variable with distance (fuel, tires, wages) and are fairly stable. The costs associated with connecting to follow-on loads consist of the deadheading (movement of an empty truck from its current position to the location of a new load) and dwell time (time the driver has to stay at a location waiting for a follow-on load to be identified). The cost of making a connection is never known with certainty by a carrier due to short dispatching lead times and the overall spatial and temporal variability of shipper demand. The uncertainty in making a connection is a major element of a TL carrier's costs. Consolidated modes, on the other hand, have large fixed costs due to the need for terminals and scheduled pick up and delivery runs.

Because of the different underlying cost structures, direct and consolidated carriers are influenced by different factors. Chapter 1 showed that direct carriers exhibit economies of scope due to the uncertainty in the connection costs. That is, the cost for a single carrier to serve a two lane closed tour, for example, is lower than the cost of having two different carriers serve each lane. Consolidated carriers, on the other hand, are affected more by economies of density - both in terms of shipments per location and customers per area.

While most procurement strategies focus on helping suppliers achieve economies of scale, the real benefits for transportation services are derived from increased economies of density and scope. The thesis argued that shippers should follow different strategies for procuring direct and consolidated transportation services. For TL, the shipper should provide carriers with opportunities to select and bundle lanes together in order to reduce connection costs and, thus, achieve economies of scope. For LTL and package delivery, the shipper should provide the carriers with the highest possible lane densities in order to better amortize their fixed delivery and terminal costs and, thus, achieve economies of density.

Contracting

After establishing the primary economic characteristics of both direct and consolidated carriers, the thesis next addressed how shippers and carriers should procure the different transportation services. The form of a relationship between buyers and sellers is determined by the characteristics of the product or service being sold and the players involved and usually falls into one of three types: spot markets, contractual agreements, or vertical integration.

Chapter 2 provided a competitive analysis of the TL segment. With more than 50,000 carriers providing a generic “box-on-wheels” service, TL trucking seems to fit all of the criteria for being a perfectly competitive market and thus transacted through the spot market. In a spot market, shippers use the first available carrier to service each load (one at a time) as it arises with no guarantees of future service. Indeed, a spot market does exist within the TL sector, but Chapter 3 showed that the use of contractual agreements between shippers and carriers is becoming dominant. While economic theory can explain why contracting occurs when specialized equipment or hyper-time-sensitive freight is involved, it does not, however, explain why contracting is becoming the dominant form of relationship in the most commodity-like segment of TL trucking, general-freight, dry vans.

In Chapter 3 we proposed a series of network based explanations as to why contracting occurs. Essentially, we argue that carriers are the prime beneficiaries of contractual agreements since they allow carriers to achieve economies of scope. The spot market cannot efficiently support the procurement of sets of lanes due to the temporal nature of the follow-on loads. That is, because the carrier must perform each delivery in sequence, a vehicle can, potentially, become “locked-in” to a series of lanes which opens the possibility of being exploited by the shipper. A contract provides a guarantee of future service and introduces greater reliability into the carrier’s service network. The contract, then, lowers the carrier’s uncertainty of making connections to follow-on loads and improves the probability of achieving economies of scope. In order to entice shippers into forming contracts with them, the carriers typically have to make rate concessions.

Bidding - Auction Design

After concluding that contracting is the preferred form of relationship for most shippers and carriers, the thesis next examined how a shipper should determine which carriers to use. The objective is to assign carriers to those portions of the network where they are best suited. This requires an exchange of information between the shipper and the carriers. The best method to accomplish this information exchange when a large number of parties are involved is through an auction or competitive bid.

In practice, the most prevalent form of competitive bidding for TL services is the single round, sealed bid auction which selects the low cost carrier for each lane. Chapter 4 showed that by ignoring all of the synergies between the lanes, the typical TL bid essentially replicates a spot market! Under a forward contract the situation is even worse, however, since the carrier has to base its pricing and capacity commitments not on the situation of the current market (as for a spot bid) but over an extended period of time relying on the shipper's own projections and estimates of future traffic flows.

Chapter 4 analyzed the TL bidding problem specifically to understand what effect different auction designs have on both shipper and carrier behavior. The two primary findings are that the traditional TL bidding methods used by shippers (1) can actually increase the carrier's hedging and (2) do not improve the carrier's ability to achieve economies of scope - one of the main goals of a carrier entering into a contract in the first place.

Carrier's hedge their prices for serving a lane due to uncertainty in both the accuracy of network information (distance, volume, seasonality, etc.) and the ability to make connections between lanes. A shipper can reduce the first type of hedging simply by providing more useful (disaggregated) and detailed information to the carriers. The second type of hedging is related to the ability of the carrier to achieve economies of scope and requires the shipper to provide some form of additional guarantees to the carriers - to be enforced by the contractual agreements discussed in Chapter 3.

One way to reduce the hedging due to connection uncertainty is through the use of conditional bids. A conditional bid is a bid submitted by a carrier which covers a set of lanes and is conditioned on the carrier receiving all of the lanes within the set. Thus, if a carrier believes there is synergy between a set of lanes within a shipper's network, a conditional bid allows the carrier to price these lanes as a set, lowering the uncertainty of making a connection between lanes, and, therefore, reducing the carrier's hedge and subsequent bid price.

General procurement strategies, such as carrier reduction, will not achieve the same result as conditional bids. Limiting the number of carriers that win in an auction will increase the volume awarded to the winning carriers, but will not reduce the hedging. The carriers do

not know ahead of time the volume that they will be awarded, and, if the selection is again lane by lane, a carrier would still need to hedge its bids in order to protect itself from the possibility of winning only a small fraction of the lanes bid. Exacerbating this is the empirical observation (winner's curse) that in this type of auction the lanes that a carrier is most likely to win are the ones they most over-valued.

Assignment Process

The final area that the thesis examined, was how to implement these concepts in order to reduce the carrier's hedging. Chapter 5 addressed how to reduce carrier hedging caused by poor or overly aggregated network information. We described a five step analysis which allows carriers to both improve the attractiveness of the network to carriers and increase the detail and usefulness of the information. Methods of organizing and characterizing the spatial and temporal distribution of loads were developed, all with the goal of improving the carrier's ability to use the information in achieving economies of scale.

Chapter 5 also describes the four most common forms of efficient aggregations for TL carriers. An efficient aggregation is defined as a set of lanes which, when taken together, allow a carrier to increase the probability of finding a follow on load and reduce the need for deadheading and/or dwell time. That is, it is a set of lanes over which a carrier can achieve economies of scope. We identified four relevant types of efficient aggregations: reloads, open tours, closed tours, and local tours. Reloads are opportunities for matching inbound and outbound loads at a specific location. An open tour is a collection of lanes that form a chain which begins and ends at different locations and may contain deadhead movements. Closed tours are similar to open tours except that they begin and end at the same location. Finally, local tours were identified as a collection of lanes with a common origin or destination that require less than half a day's travel.

In Chapter 6 addressed how to reduce carrier hedging caused by connection uncertainty using conditional bids. We demonstrated the practicality and usefulness of optimization-based carrier assignment. Various model formulations were developed which allow for the inclusion of economies of scope, economies of density, and level of service. bid-set models, which permit the use of conditional bids and other system considerations, were implemented on a large TL bid sample. The formulations were shown to produce integral solutions for most cases. We show that while conditional bids can increase the model's complexity, they do not seem to make the problem more difficult to solve. In fact the system considerations pose more difficulties.

Chapter 6 showed that the carrier assignment problem for a large and contentious network can be solved in a reasonable amount of time in a PC environment. The models

provide not only an optimal lane assignment, but also tools to analyze different business practices.

7.2 Contributions

This thesis makes five contributions to the transportation procurement field:

1. Inclusion of supplier economics in TL procurement strategy

Unlike previous research which focused solely on providing suppliers with economies of scale, this thesis based the procurement strategy on the specific factors affecting TL carriers. Because TL carriers are primarily affected by economies of scope, shippers can adjust their bidding and assignment practices accordingly. The contribution is in the recognition and inclusion of the TL carriers' economies of scope into all phases of the shipper's procurement process.

2. Justification of contracts in general freight TL market

Previous research had not provided an adequate explanation as to why contracts are forming in the most generic segment of the TL trucking industry; dry-van general-freight. A network based justification for the growing use of contractual arrangements in the seemingly perfectly competitive TL market was proposed which fills this gap in the economics literature. The growing dominance of contracts was shown to be due to (1) the importance of economies of scope for the TL carriers and (2) inability of a spot market to achieve an efficient allocation in the presence of economies of scope. Carriers were shown to be the primary beneficiaries of contractual agreements - which explains their need to provide rate concessions in order to entice shippers.

3. Formal analysis and critique of TL bidding practices

TL bidding had never been formally examined as an auction before. The auction design theory literature was synthesized, adapted, and applied to the TL bid process. The analysis showed that the current practice of single-round, sealed-bid auctions using simple bids (single lane per bid) rarely finds an efficient allocation if there is any degree of interdependency between the lanes. A decision framework was developed to assist in the design of TL bids based on the characteristics of the shipper's traffic network. Additionally, the effect that carrier hedging has on the bidding process was demonstrated and conditional bids were introduced as a method of reducing carrier hedging due to connection uncertainty.

4. Development of Pre-Bid Analysis for shippers

The way in which a shipper represents its network to carriers can influence the carriers' costs. We developed a methodology which analyzes the shipper's network and determines how to (1) form traffic lanes, (2) aggregate shipping locations into larger zones, (3) represent seasonality of traffic flow, (4) expand the network to increase its attractiveness to carriers, and (5) identify different forms of potential efficient aggregations within the shipper's network. Additionally, the savings potential of each type of aggregation is approximated and a methodology of estimating the probability of load matching within a shipper's network was developed and implemented. The contribution of this analysis is in formalizing a method for shippers which includes the TL carrier's economics in its determination of how to represent the network to carriers.

5. Implementation of optimization-based assignment models

We demonstrate the practicality and usefulness of optimization-based carrier assignment for large and contentious traffic networks by implementing a carrier assignment model to data from an actual TL bid. Previous applications of carrier assignment optimization methods had only considered simple-bids and were used on smaller problems. A general model was developed which can be used to incorporate economies of scope, economies of density, and other forms of carrier costs. We also developed methods of incorporating service capabilities and accessorials in the bidding process and analyzed the business implications of different carrier assignments.

7.3 Extensions

While there are many potential directions that this research can be extended, we briefly discuss only three.

Re-Negotiation

In Chapter 6 we introduced and solved a series of optimization based carrier assignment models. The models were used for determining the actual assignment of lanes to carriers as well as for analyzing different business decisions, such as limiting the number of carriers in the system. Another useful purpose that these models could serve is to suggest a course of action for potential re-negotiation after a single round auction. Specifically, we are interested in determining which carriers to approach, in what order, and for which lanes. Additionally, the analysis should provide some guidance as to how much a carrier needs to

reduce its bids to affect the solution and, finally, an estimate of the savings the shipper should expect.

There are two methodological challenges. First, it might be difficult to obtain the relevant information from the models. Because the models are integer programs, the dual information required to perform any type of sensitivity analysis is either not available or is biased due to variables being fixed. Some form of post-optimality logic will most likely be needed to calculate these dual values.

The second methodological challenge is to consider the effect of changing multiple bids at the same time. Traditional sensitivity analysis of mathematical programs only considers changes that are caused by modifying a single aspect of the model at a time. For the re-negotiation problem, however, we are interested in determining the effect that making a series of inter-related changes will have on the final solution.

In addition to these methodological challenges, the actual logistics of carrying out the re-negotiation need to be examined. Should carriers be presented with target bids for certain lanes or just be told to "improve their bids"? Should the information gained from the re-negotiation analysis be shared with the carriers, or kept confidential? These questions are as pressing and relevant as the methodological concerns.

Better Estimation of Load Matching Probabilities

While Chapter 5 suggested some methods to determine the probability of a reload using the historical data, this process could be refined further. Four possible extensions are envisioned.

First, the seasonality and distribution of the traffic flow could be better characterized. We assumed that the traffic on each lane followed a Poisson distribution which allowed us to make many simplifying assumptions. Other distributions should be examined for better fit. Additionally, we used the entire year for calculating the estimated flows. Seasonality effects should be included in the estimation.

Second, the approximate deadhead distance required in a zone could be estimated. Using information on the other shipping locations within an area, a lower bound on the expected empty miles required between loads can be developed. For example, if we assume that the origins and destinations within a zone are uniformly distributed across the zone, then, we can estimate the expected distance needed to travel between any two points. Larson and Odoni (1981, 135) show that this distance is approximately $c\sqrt{A}$, where c is a coefficient (between .52 and .60) describing the network and A is the area of the zone. Other results from geometric probability can be applied to better estimate deadheads and dwell times required between loads.

Third, it would be beneficial to combine these two ideas and develop an algorithm to determine which sets of lanes adjacent to a zone should be combined. The method we used in Chapter 5 was greedy in the sense that it added in all lanes to determine the reload potential at a zone. This method could be refined.

Finally, we have not considered the upper limit to the attractiveness of these increased follow-on probabilities to carriers. In order to secure a set of lanes to improve the probability of a reload, the carrier needs to share some of the savings due to economies of scope with the shipper. Thus, the carrier is attempting to increase its expected total profitability by increasing the reload probability while decreasing the lane contribution, that is, the carrier is trading off frequency of connection for payoff per connection. It would be interesting to determine what the value of this tradeoff is for different carriers and what types of carriers are the best candidates for using conditional bids.

Inclusion of Volume Dependent Rates

All of the carrier assignment models that we implemented in Chapter 6 concentrated on addressing economies of scope and therefore assumed that the rate was volume independent. In reality, carriers are sensitive to not only which lanes are assigned, but also the flow on the assigned network. Thus, the cost complementarities of the different lanes, rather than just the economies of scope, need to be considered.

It is not clear how carriers should submit this information. One way could be to provide a volume based discount which applies over a set of lanes, so that if the carrier is awarded, say, 500 loads per month out of Atlanta, then the price on each of these lanes is reduced by 5%. Another approach is for the carrier to submit a function or relationship relating the volume to the lane rates. While carriers are probably not able to provide this relationship itself, it could, perhaps be developed as an approximation of the carrier's costs. Allowing carriers to submit volume dependent rates, in any form, however, greatly complicates the assignment model.

Carrier-set assignment models, suggested and developed in Chapter 6, are a potential method of capturing volume dependence. By restricting the pricing of each carrier's assignment and volume allocation within a sub-problem, the master problem is greatly simplified. Unfortunately, this makes the sub-problem very difficult. The challenge to implementing the carrier-set models, then, is in finding a sub-problem that is simpler to solve than the proposed complete generation and clique methods which are not workable for a realistically sized network.

References

- A.T. Kearney. 1983. *Private Fleet Benchmarks of Quality and Productivity*. Arlington, VA: The National Private Truck Council.
- Abshire, Roger Dale and Shane R. Premeaux. 1991. Motor Carrier Selection Criteria: Perceptual Differences Between Shippers and Carriers. *Transportation Journal*. v31 n3 (Spring): 31-35.
- Ahuja, Ravindra K., Thomas L. Magnanti, and James B. Orlin. 1993. *Network Flows: Theory, Algorithms, and Applications*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Alex Brown & Sons. 1995a. *Fifth Annual Alex. Brown & Sons Transportation Seminar*. Alex. Brown & Sons, Inc. (May).
- Alex Brown & Sons. 1995b. *Truckload Transportation Report*. (September).
- American Trucking Association. 1994. *American Trucking Trends, 1994-1995*. Washington, DC: American Trucking Association.
- Association of American Railroads. 1988. *Rail Stations by Standard Point Location Codes (SPLC)*. Transportation Codes Section, Economics and Finance Department.
- Baker, Gwendolyn H. 1984. The Carrier Elimination Decision: Implications for Motor Carrier Marketing. *Transportation Journal*. v24 n1 (Fall): 20-29.
- Banks, Jeffrey S., John O. Ledyard, and David P. Porter. 1989. Allocating Uncertain and Unresponsible Resources: An Experimental Approach. *RAND Journal of Economics*. v20 n1 (Spring): 1-25.
- Bardi, Edward J., Prabir K Bagchi, and T.S. Raghunathan. 1989. Motor Carrier Selection in a Deregulated Environment. *Transportation Journal*. v29 n1: 4-11.
- Barnhart, Cynthia, Christopher Hane, Ellis L. Johnson, and G. Sigismondi. 1994. A Column Generation and Partitioning Approach for Multi-Commodity Network Problems. *Telecommunications Systems*. To appear.
- Barnhart, Cynthia, Ellis Johnson, George Nemhauser, G. Sigismondi, and Pamela Vance, "Formulating a Mixed Integer Programming Problem to Improve Solvability," *Operations Research*, Vol. 41, No. 6, pp. 1013-1019.
- Barr, James L. and Timothy L. Shaftel. 1976. Solution Properties of Deterministic Auctions. *Journal of Financial and Quantitative Analysis*. (June): 287-311.
- Beier, Frederick J. 1989. Transportation Contracts and the Experience Effect: A Framework for Future Research. *Journal of Business Logistics*. v10 n2: 73-89.
- Beilock, Richard and Timothy Shell. 1992. Brokerage and the Potential for Electronic Marketing of Produce Transportation. *Transportation Journal*. v31 n4: 60-71.
- Ben Akiva, Moshe and Steven R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, MA: MIT Press.

- Bienstock, Carol. 1994. *The Effect of Outsourcing and Situational Characteristics on Physical Distribution Transportation Efficiency*. Unpublished Ph.D. Dissertation, Pamplin College of Business, Virginia Polytechnic Institute and State University.
- Bradley, Stephen P., Arnaldo C. Hax, and Thomas L. Magnanti. 1977. *Applied Mathematical Programming*. Reading, MA: Addison-Wesley Publishing Company.
- Brewer, Paul J. and Charles R. Plott. 1995. A Binary Conflict Ascending Price (BICAP) Mechanism for the Decentralized Allocation of the Right to Use Railroad Tracks. Social Science Working Paper 887. California Institute of Technology. (Revised February 1995).
- Brown, Terrence and Janet Greenlee. 1995. Private Trucking After Deregulation: Managers' Perceptions. *Transportation Journal*. v34 n1 (Fall): 5-14.
- Brown, Terrence. 1984. Freight Brokers and General Commodity Trucking. *Transportation Journal*. v24 n2 (Winter): 4-14.
- Brown, Terrence. 1990a. Property Brokers: A Pilot Study of Carriers' Perspective. *Transportation Journal*. v29 n2 (Winter): 32-39.
- Brown, Terrence. 1990b. Size and Operating Characteristics of Property Brokers. *Transportation Journal*. v29 n4 (Summer): 52-57.
- Brown, Terrence. 1991. Property Brokers: A Pilot Study of Shipper Perspectives. *Transportation Journal*. v30 n1 (Fall): 45-50.
- Brown, Terrence. 1992. *Transportation Brokers: History, Regulation, and Operations*. Alexandria, VA: Transportation Brokers Conference of America.
- Bykowsky, Mark M., Robert J. Cull, and John Ledyard. 1995. Mutually Destructive Bidding: The FCC Auction Design Problem. Social Science Working Paper 916. California Institute of Technology. (January).
- Capen, E.C., R.B. Clapp, and W. M. Campbell. 1971. Competitive Bidding in High Risk Situations. *Journal of Petroleum Technology*. v23 (June): 641-653.
- Caves, Douglas and Laurits R. Christiansen, "The Importance of Economies of Scale, Capacity Utilization, and Density in Explaining Interindustry Differences in Productivity Growth," *The Logistics and Transportation Review*, Vol. 24. No. 1 (1988), pp. 3-32.
- Cavinato, Joseph L. 1984. Transportation Contracts: Pointers and Pitfalls for Buyers. *Journal of Purchasing and Materials Management*. v20 n4: 9-13.
- CH Robinson. 1995. *Managing It the Right Way, All Ways*. C.H. Robinson Company Marketing Literature.
- Chiang, W. S. Judy and Ann F. Friedlaender. 1984. Output Aggregation, Network Effects, and the Measurement of Trucking Technology. *Review of Economics and Statistics*. v64: 267-276.
- Chow, Garland. 1978. *The Economics of the Motor Freight Industries*. Indianapolis, IN: Indiana University Press.
- Corsi, Thomas M. and Joseph R. Stowers. 1991. Effects of a Deregulated Environment on Motor Carriers: A Systematic, Multi-Segment Analysis. *Transportation Journal*. v30 n3 (Spring): 4-28.
- CPLEX Optimization, Inc. 1994. *Using the CPLEX Callable Library*.
- Cripps, Martin and Norman Ireland. 1994. The Design of Auctions and Tenders with Quality Thresholds: The Symmetric Case. *The Economic Journal*. v104 (March): 316-326.

- Crum, Michael R. and Benjamin J. Allen. 1990. Shipper EDI, Carrier Reduction, and Contracting Strategies: Impacts on the Motor Carrier Industry. *Transportation Journal*. v29 n4 (Summer): 8-32.
- Crum, Michael R. and Benjamin J. Allen. 1991. The Changing Nature of the Motor Carrier-Shipper Relationship: Implications for the Trucking Industry. *Transportation Journal*. v31 n2 (Winter): 41-54.
- Daganzo, Carlos, F. 1991. *Logistics Systems Analysis*. Lecture Notes in Economics and Mathematical Systems Series, vol. 361. New York: Springer-Verlag.
- Daughety, Andrew F., Forrest D. Nelson, and William R. Vigdor. 1985. An Economic Analysis of the Cost and Production Structure of the Trucking Industry. In *Analytical Studies in Transport Economics*. ed. Andrew F. Daughety. Cambridge: Cambridge University Press. Chapter 3.
- Distribution. 1983. Private Fleets as Brokers. *Distribution*. (May): 75.
- Distribution. 1993a. A Window on TL Logistics. *Distribution*. (February): 54-59.
- Dixon, Lance and Anne Miller Porter. 1994. *JIT II: Revolution in Buying and Selling*. Newton, MA: Cahners Publishing Company.
- Economides, Nicholas. 1994. The Economics of Networks. *Plenary Session Address, E.A.R.I.E. Conference, Chania, Greece*. (September).
- Engelbrecht-Wiggans, Richard. 1980. State of the Art: Auctions and Bidding Models: A Survey. *Management Science*. v26 n2 (February):119-142.
- Federal Communications Commission. 1993. *In the Matter of Implementation of Section 309(j) of the Communications Act - Competitive Bidding: Notice of Proposed Rule Making*. PP Docket No. 93-253. Federal Communications Commission. (12 October).
- Federal Communications Commission. 1994a. *In the Matter of Implementation of Section 309(j) of the Communications Act - Competitive Bidding: Second Report and Order*. PP Docket No. 94-61. Federal Communications Commission. (20 April).
- Federal Communications Commission. 1994b. *In the Matter of Implementation of Section 309(j) of the Communications Act - Competitive Bidding: Third Report and Order*. PP Docket No. 94-98. Federal Communications Commission. (10 May).
- Feldman, Robert A. and Rajnish Mehra. 1993. Auctions: Theory and Applications. *International Monetary Fund Staff Papers*. v40 n3 (September): 485-511.
- Foster, Jerry R. and Sandra Strasser. 1991. Carrier/Modal Selection Factors: The Shipper/Carrier Paradox. *Journal of the Transportation Research Forum*. v31 n1: 206-212.
- Friedlaender, Ann F. and Richard H. Spady. 1980. *Freight Transport Regulation: Equity, Efficiency, and Competition in the Rail and Trucking Industries*. Cambridge, MA: The MIT Press.
- Gentry, Julie, J. 1991. *Purchasing's Involvement in Transportation Decision Making*. Center for Advanced Purchasing Studies/National Association of Purchasing Management.
- Gibson, Brian J., Harry L. Sink, and Ray A. Mundy. 1993. Shipper-Carrier Relationships and Carrier Selection Criteria. *Logistics and Transportation Review*. v29 n4 :371-390.
- Gibson, Brian J., Ray A. Mundy, and Harry L. Sink. 1995. Supplier Certification: Application to the Purchase of Industrial Transportation Services. *Logistics and Transportation Review*. v31 n1: 63-75.

- Griffin, Gene C. and Julene M. Rodriguez. 1992. *Creating a Competitive Advantage Through Partnershipping with Owner-Operators*. The Upper Great Plains Transportation Institute. UGPTI No. 91. Fargo, ND: North Dakota State University.
- Grimm, Curtis M., Thomas M. Corsi, and Judith L. Jarrell. 1989. U.S. Motor Carrier Cost Structure Under Deregulation. *The Logistics and Transportation Review*. v25: 231-249.
- Hall, Randolph, W. 1987. Consolidation Strategy: Inventory, Vehicles, and Terminals. *Journal of Business Logistics*. v8 n2: 57-73.
- Harper, Lucinda. 1994. Trucks Keep Inventories Rolling Past Warehouses to Production Lines. *Wall Street Journal*. 7 February, A9A.
- Harris, Milton and Artur Raviv. 1981. Allocation Mechanisms and the Design of Auctions. *Econometrica*. v49 n6 (November): 1477-1499.
- Hausch, Donald B. 1986. Multi-Object Auctions: Sequential vs. Simultaneous Sales. *Management Science*. v32 n12 (December): 1599-1612.
- Hoffman, Kurt. 1993b. Annual Logistics Review: TL & Specialized Carriers. *Distribution*. (July): 24-36.
- Hoffman, Kurt. 1994. Annual Logistics Review: TL & Specialized Carriers. *Distribution*. (July): 22-25.
- Hoffman, Kurt. 1995. Annual Logistics Review: TL & Specialized Carriers. *Distribution*. (July): 20-28.
- Interstate Commerce Commission. 1977. *Empty/Loaded Truck Miles on Interstate Highways During 1976*. Bureau of Economics, Interstate Commerce Commission.
- Jara Díaz, Sergio R. 1981. *Transportation Cost Functions: A Multiproduct Approach*. Unpublished Ph.D. dissertation. Massachusetts Institute of Technology. Cambridge, MA.
- Jara Díaz, Sergio R. 1982a. The Estimation of Transport Cost Functions: A Methodological Review. *Transport Reviews*. v2 n3: 257-278.
- Jara Díaz, Sergio R. 1982b. Transportation Product, Transportation Function and Cost Functions. *Transportation Science*. v16 n4 (November): 522-539.
- Jara Díaz, Sergio R. 1983. Freight Transportation Multioutput Analysis. *Transportation Research*. v17A n6: 429-438.
- Jara Díaz, Sergio R. 1988. Multioutput Analysis of Trucking Operations Using Spatially Disaggregated Flows. *Transportation Research*. v22B n3: 159-171.
- Jara Díaz, Sergio R. and Cristian E. Cortes. 1995. Calculation of Scale Economies from Transport Cost Functions. Presented at 7th World Conference on Transport Research. Sydney, Australia (July).
- Keaton, Mark H. 1983. Are There Economies of Traffic Density in the Less-Than-Truckload Motor Carrier Industry An Operations Planning Analysis. *Transportation Research*. v27A n5: 343-358.
- Keeney, Regina. 1995. FCC Grants Ten Regional Narrowband PCS Licenses. Federal Communications Commission Press Release. (23 January).
- Kelley, Al and Ira Pohl. 1990. *A Book on C*. Redwood City, CA: Benjamin/Cummings Publishing Company, Inc. 2nd Edition.

- Kiesling, Max K. 1995. *A Comparison of Freight Distribution Costs for Combination and Dedicated Carriers in the Air Express Industry*. Unpublished Ph.D. dissertation. University of California at Berkeley.
- Koopmans, Tjalling C. and Martin Beckman. 1957. Assignment Problems and the Location of Economic Activities. *Econometrica*. v25: 53-76.
- Kreps, David M. 1990. *A Course in Microeconomics Theory*. Princeton, NJ: Princeton Press.
- La Londe, Bernard J. and Martha C. Cooper. 1989. *Partnerships in Providing Customer Service: A Third-Party Perspective*. Oak Brook, IL: Council of Logistics Management.
- La Londe, Bernard, James M. Masters, Arnold B. Maltz, and Lisa R. Williams. 1991. *The Evolution, Status, and Future of the Corporate Transportation Function*. American Society of Transportation and Logistics.
- Lane, Lee. 1987. Innovation in Trucking: Advanced Truckload Firms. *Transportation Research Record*. Washington, DC: Transportation Research Board.
- Larson, Richard C. and Amadeo Odoni. 1981. *Urban Operations Research*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Ledyard, John O. 1993. The Design of Coordination Mechanisms and Organizational Computing. *Journal of Organizational Computing*. v3 n1: 121-134.
- Ledyard, John O., Charles Noussair, and David Porter. 1995. The Allocation of a Shared Resource within an Organization. Social Science Working Paper 917. California Institute of Technology. (January).
- Leonard, Herman B. 1983. Elicitation of Honest Preferences for the Assignment of Individuals to Positions. *Journal of Political Economy*. v91 n3: 461-479.
- Lieb, Robert C. 1994. *Transportation*. Houston, TX: Dame Publications, Inc. 4th Edition.
- Lieb, Robert C. and Robert A. Miller. 1988. JIT and Corporate Transportation Requirements. *Transportation Journal*. v27 n3 (Spring): 5-10.
- Maltz, Arnold. 1993. Private Fleet Use: A Transaction Cost Model. *Transportation Journal*. v32 n3 (Spring): 46-53.
- Masten, Scott E., James W. Meehan, Jr., and Edward A. Snyder. 1991. The Costs of Organization. *The Journal of Law, Economics, & Organization*. v7 n1: 1-26
- McAfee, R. Preston and John McMillan. 1987. Auctions and Bidding. *Journal of Economic Literature*. v25 (June): 699-738.
- McAfee, R. Preston and John McMillan. 1995. Analyzing the Airwaves Auction. Unpublished manuscript.
- McAfee, R. Preston. 1993. *Auction Design for Personal Communications Services: Reply Comments*. Unpublished comments for Federal Communications Commission.
- McCabe, Kevin A. Stephen J. Rassenti, and Vernon L. Smith. 1991. Smart Computer-Assisted Markets. *Science*. v254: 534-538.
- McGinnis, Michael A. 1989. A Comparative Evaluation of Freight Transportation Choice Models. *Transportation Journal*. v29 n2 (Winter): 36-46.
- McMillan, John. 1994. Selling Spectrum Rights. *Journal of Economic Perspectives*. v8 n3 (Summer): 145-162.
- Melbin, Jodi. 1995. The Converting of Private Fleets: Fact or Fiction. *Distribution*. (July): 41.

- Milgrom, Paul and Robert Wilson. 1993. *Replies to Comments on PCS Auction Design*. Unpublished comments for Federal Communications Commission.
- Milgrom, Paul R. 1995. Auctioning the Radio Spectrum. *Putting Auction Theory to Work*, Draft manuscript of Chapter 1.
- Milgrom, Paul R. and Robert J. Weber. 1982. A Theory of Auctions and Competitive Bidding. *Econometrica*. v50 n5 (September): 1089-1122.
- Milgrom, Paul. 1989. Auctions and Bidding: A Primer. *Journal of Economic Perspectives*. v3 n3 (Summer): 3-22.
- Military Traffic Management Command. 1994. *Guaranteed Traffic Solicitation for TL Carriers*. Department of Defense, Military Traffic Management Command. (29 November).
- Miller, Edward. 1973. Effect of Regulation on Truck Utilization. *Transportation Journal*. v13 n1 (Fall): 5-14.
- MIT Dictionary of Modern Economics*. 1992. 4th edition. Ed. David W. Pearce. Cambridge, MA: MIT Press..
- Moore, E. William, Janice M. Warmke, and Lonny R. Gorban. 1991. The Indispensable Role of Management Science in Centralizing Freight Operations at Reynolds Metals Company. *Interfaces*. v21 n1: 107-129.
- Moskal, B.S. 1989 Enough Premium Carriers in the 90's? Don't Bet On It. *Industry Week*. (August).
- Murty, Katta G. 1983. *Linear Programming*. New York, NY: John Wiley & Sons.
- Nalebuff, Barry J. and Jeremy I. Bulow. 1993. *Response to PCS Auction Design Proposals*. Unpublished comments for Federal Communications Commission.
- Nicholson, Walter. 1992. *Microeconomic Theory: Basic Principles and Extensions*. Orlando, FL: The Dryden Press. 5th Edition.
- Olson, Mark and David Porter. 1994. An Experimental Examination into the Design of Decentralized Methods to Solve the Assignment Problem with and without Money. *Economic Theory*. v4: 11-40.
- Owner Operator Independent Drivers Association. 1995. Cost of Operations Survey for 1992, 1993, and 1994. Unpublished documents and correspondence with Rick Craig.
- Palay, Thomas M. 1984. Comparative Institutional Economics: The Governance of Rail Freight Contracting. *The Journal of Legal Studies*. v13: 265-287.
- Phillips, Laurence T. 1991. Contractual Relationships in the Deregulated Transportation Marketplace. *Journal of Law and Economics*. v34 (October): 535-564.
- Pirrong, Stephen C. 1993. Contracting Practices in Bulk Shipping Markets: A Transactions Cost Explanation. *Journal of Law and Economics*. v34 (October): 937-976.
- Plott, Charles R. and David P. Porter. 1990. An Experiment with Space Station Pricing Policies. Social Science Working Paper 704. (March).
- Porter, Michael E. 1980. *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. New York, NY: Free Press, Inc.
- Powell, Warren B., Yosef Sheffi, and Sebastien Thiriez. 1984. The Dynamic Vehicle Allocation Problem with Uncertain Demands. *Proceedings of the Ninth International Symposium on Transportation and Traffic Theory*. eds. Dormuller and Hamerslag. The Netherlands: VNU Science Press: 357-374.

- Powell, Warren B., Yosef Sheffi, Kenneth S. Nickerson, and Susan Atherton. 1988. Maximizing Profits for North American Van Lines' Truckload Division: A New Framework for Pricing and Operations. *Interfaces*. V18 n1 (January-February): 21-41.
- Powell, Warren, B. 1987. An Operational Planning Model for the Dynamic Vehicle Allocation Problem with Uncertain Demands. *Transportation Research*. v213 n3: 217-232.
- Private Fleet Directory*, 1995-1996 Edition. National Private Truck Council.
- Raiffa, Howard. 1982. *The Art and Science of Negotiation*. Cambridge, MA: The Belknap Press.
- Rakowski, James P. 1990. Marketing Myopia Meets Marketing Economies: Unexpected Results of U.S. Airline and Trucking Deregulation. *Transportation Quarterly*. v44 n4: 499-516.
- Rakowski, James P., R. Neil Southern, and Judith L. Jarrell. 1993. The Changing Structure of the U.S. Trucking Industry: Implications for Logistics Managers. *Journal of Business Logistics*. v14 n1: 111-129.
- Rassenti, S.J., V.L Smith, and R.L. Bulfin. 1982. A Combinatorial Auction Mechanism for Airport Time Slot Allocation. *The Bell Journal of Economics*. v13: 402-417.
- Richardson, Helen. 1993. Contracts Build Relationships. *Transportation & Distribution*. v34 n11: 53.
- Rinehart, Lloyd M. 1989. Organizational and Personal Factors Influencing the Negotiation of Motor Carrier Contracts: A Survey of Shippers and Motor Carriers. *Transportation Journal*. v29 n2: 4-14.
- Rothkopf, Michael H. and Ronald M. Harstad. 1994. Modeling Competitive Bidding: A Critical Essay. *Management Science*. v40 n3 (March): 364-383.
- Rothkopf, Michael H., Aleksandar Pekec, and Ronald M. Harstad. 1995. Computationally Manageable Combinatorial Auctions. *RUTCOR Research Report*. RRR 13-95. (April).
- Sarin, Rakesh and Martin Weber. 1993. Effects of Ambiguity in Market Experiments. *Management Science*. v39 n5 (May): 602-615.
- Schulz, John. 1995a. Evolution in Truckload Service. *Traffic World*. (14 August): 12-14.
- Schulz, John. 1995b. Revamped Carolina. *Traffic World*. (8 May): 54.
- Schulz, John. 1995c. Schneider National Making Major Inroads In Private Fleet Conversion Market. *Traffic World*. (15 May): 26-28.
- Schulz, John. 1995d. Big Catch: How Ryder Logistics Landed a \$56 million deal with Whirlpool. *Traffic World*. (13 March): 38-42.
- Schulz, John. 1996. WPE Carves Out Lucrative Time-Definite Regional Niche. *Traffic World*. (26 February): 18.
- Sheffi, Yosef. 1985. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Englewood Cliffs, NJ: Prentice-Hall Inc.
- Sheffi, Yosef. 1986. Carrier/Shipper Interactions in the Transportation Market: An Analytical Framework. *Journal of Business Logistics*. v7 n1: 1-27.
- Shubik, Martin. 1984. *A Game-Theoretic Approach to Political Economy*. Game Theory in the Social Sciences Series, vol. 2. Cambridge, MA: MIT Press.
- Simon, Herbert A. 1961. *Administrative Behavior*. 2nd Ed. New York, NY: MacMillan. Original Publication: 1947.

- Standard and Poor's Corporation. (1994) *Standard and Poor's Industry Surveys, Railroads and Trucking*. Section 2. (8 December).
- Standard and Poor's Corporation. (1995) *Standard and Poor's Industry Surveys, Railroads and Trucking*. (30 November).
- Stephenson, Frederick J. and Theodore P. Stank. 1994. Truckload Motor Carrier Profitability Strategies. *Transportation Journal*. v34 n2 (Winter): 5-17.
- Talley, Wayne Kenneth. 1988. *Transport Carrier Costing*. New York: Gordon and Breach.
- Vickrey, William. 1961. Counter-Speculation, Auctions and Competitive Sealed Tenders. *Journal of Finance*. v16: 8-37.
- Weber, Robert J. 1993. *Reply to Comments on FCC 93-455: Notice of Proposed Rule Making*. Unpublished comments for Federal Communications Commission.
- Williamson, Oliver E. 1985. *Economic Institutions of Capitalism*. New York, NY: The Free Press.