FEASIBILITY OF A

FREIGHT MOVEMENT MODEL

PB99-123283

FINAL REPORT

by

Mark R. McCord
Yi-Ying Chung
Dario Hidalgo
Jun Zhang

Department of Civil and Environmental Engineering
The Ohio State University
2070 Neil Avenue
Columbus, OH 43210

December 1998

Prepared in Cooperation with
The Ohio Department of Transportation
and
The U.S. Department of Transportation,
Federal Highway Administration

State Job No. 14641(0)

FEASIBILITY OF A

FREIGHT MOVEMENT MODEL

EXECUTIVE SUMMARY

by

Mark R. McCord
Yi-Ying Chung
Dario Hidalgo
Jun Zhang

Department of Civil and Environmental Engineering
The Ohio State University
2070 Neil Avenue
Columbus, OH 43210

December 1998

Prepared in Cooperation with
The Ohio Department of Transportation
and
The U.S. Department of Transportation,
Federal Highway Administration

State Job No. 14641(0)

We investigated the feasibility and desirability of implementing mathematical freight forecasting models to be operated by the Ohio Department of Transportation. We interviewed potential users to elicit what they considered desirable model outputs and conducted a survey of freight forecasting efforts in other state DOT's. The many different outputs mentioned by potential Ohio users indicate a need for a general model. Despite many available freight models proposed in the literature, our survey results show that few states can be considered to be active in comprehensive freight forecasting. Still, several appear active in aspects of freight forecasting, primarily addressing truck trips and using a few publicly available databases, often supplemented with private data and special data collection efforts. More detailed investigations of statewide freight models being developed in three states showed that the models in these states are all being developed by consultants and designed to address statewide passenger, as well as freight forecasting. They all employ modular frameworks resembling the traditional 4-step urban transportation planning modeling system and presently emphasize assigning truck trips to the highway network, although the eventual intent is to allow multimodal freight assignment. We illustrate that simple models could presently be implemented with existing databases to serve either as components in such statewide models or as stand-alone models. However, we also show that different specifications can produce very different forecasts, and that it is not yet clear which alternatives would be most useful.

We expect that statewide freight modeling activities should increase in state DOT's in the near future, that many of the models will resemble the 4-step urban passenger forecasting system, and that implementations will concentrate on the highway network in the near future. We encourage ODOT to implement a statewide freight model using a similar framework if it commits to sustained development, research, and testing designed to regularly identify and implement improved model components and accelerate understanding the appropriate use of the model. Given the lack of experience with statewide freight models, we foresee that having consultants simply "develop and turn over" a model would fail without such a commitment, even if agency personnel are trained to run the model. To reduce costs, make expertise more readily available, and help sustain commitment to regularly implementing improved components, we propose that models be developed in collaboration with other states. Such collaboration would also allow a stronger influence to be exerted in the design of federal freight studies and data collection efforts. It could even be advantageous to formalize interstate collaboration by pooling funds to develop a regional model that could be scaled to appropriate resolutions for the participating states. Additionally, we encourage the formation of an advisory group consisting of experts who would ensure the relevance of model developments and increase the likelihood of practical use.

We also present three similar methods of updating truck origin-destination (OD) matrices from observed volumes and show that recently available network databases could be used with existing software to perform intermodal assignment. We suggest that ODOT estimate a statewide truck OD matrix from data recently collected in a roadside survey and routinely use some method to update the estimated matrix. However, since the quality of intermodal assignments is limited by a lack of acceptable logic and commodity OD matrices at this time, we propose that the use of intermodal databases and software should presently be limited to developing and testing intermodal assignment algorithms.

1. Report No.	2. Government Accession No.	3. Reci
FHWA/OH-98/016		PB99-123283

4. Title and Subtitle	5. Report Date
FEASIBILITY OF A FREIGHT MOVEMENT MODEL	September, 1998
	6. Performing Organization Code
	8. Performing Organization Report No.
7. Author(s) Mark McCord	10. Work Unit No. (TRAIS)
9. Performing Organization Name and Address The Ohio State University Department of Civil Engineering Columbus, OH 43210-1063	11. Contract or Grant No. State Job No. 14641(0)
	13. Type of Report and Period Covered Final Report
12. Sponsoring Agency Name and Address Ohio Department of Transportation 25 S. Front St. Columbus, OH 43215	14. Sponsoring Agency Code

15. Supplementary Notes

Prepared in cooperation with the U.S. Department of Transportation, Federal Highway Administration

16. Abstract

We investigate the feasibility and desirability of implementing freight forecasting models at the Ohio Department of Transportation. Interviews with potential users indicate a desire for a general model, while a survey of state DOT's leads us to believe that freight modeling for issues of interest to state DOT's is relatively new but should increase in the future. More detailed investigations of models being developed in three states indicate similar approaches to development.

We conclude that it would be feasible and desirable to develop an Ohio statewide freight model with a framework similar to those being developed in other states. However, many different specifications of the model components could be considered, and it is not yet clear which would be most useful. We, therefore, recommend implementing a statewide freight model at this time only if a commitment is also made to sustained development and testing of model components. We also recommend close collaboration with other states when developing the model, perhaps even to the point of pooling funds to develop a regional model that could be scaled to appropriate resolutions for participating states. In addition, we encourage the formation of an advisory group of freight experts to ensure model relevance and increase the likelihood of responsiveness. We also suggest that ODOT develop a statewide truck OD matrix from data recently collected in a roadside survey and regularly use a method to update the matrix from observed link volumes.

17. Key Words	18. Distribution Statement
Freight, Models, Forecasts, Surveys, Databases, Indicators	No Restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified		

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

FEASIBILITY OF A

FREIGHT MOVEMENT MODEL

FINAL REPORT

by

Mark R. McCord
Yi-Ying Chung
Dario Hidalgo
Jun Zhang

Department of Civil and Environmental Engineering
The Ohio State University
2070 Neil Avenue
Columbus, OH 43210

December 1998

Prepared in Cooperation with
The Ohio Department of Transportation
and
The U.S. Department of Transportation,
Federal Highway Administration

State Job No. 14641(0)

DISCLAIMER

Acknowledgments

This report was prepared under The Ohio Department of Transportation Agreement No. 8389, State Job No. 14641(0), for the research project entitled, "The Feasibility of a Freight Movement Model." We gratefully acknowledge the funding of The Ohio Department of Transportation and the assistance of Bill Edwards and Roger Green. We acknowledge the assistance of Mark Byram, Larry Sutherland, Pat Moore, Greg Giaimo, and Dave Moore of the Ohio Department of Transportation, and Lou Jannazo of the Ohio Rail Development Commission. We thank the interview participants, listed in Appendix 2.1, and questionnaire respondents, listed in Appendix 2.2. We especially thank Steve Smith (Indiana DOT), Bill Upton (Oregon DOT), and Richard Nellet (Michigan DOT) for insights into their statewide models. We acknowledge the early participation of Dr. A. Ziliaskopoulos and. Jennifer Donnaly. Finally, we are grateful to the comments of Mark Byram, Greg Giaimo, and Ansen Wu on an initial draft of this report. However, all views presented are those of the principal investigator of the project.

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF FIGURES

LIST OF TABLES

Section 1. Introduction

This report documents the results of our investigation into the feasibility of developing a freight movement model for use in public sector planning in Ohio.

1.1 Motivation for Freight Modeling

Interest in freight planning issues has been increasing due, in part, to ISTEA legislation and to the increased recognition of the economic importance of an efficient freight transportation system. Large-scale infrastructure and policy alternatives are costly and irreversible in the short term. Therefore, it is useful to be able to predict the resulting impacts of the various alternatives before deciding which alternative to implement.

Freight flows can be considered a relatively direct impact of implemented alternatives because of the economic benefits associated with freight transportation and handling in a region. For instance, building or expanding an intermodal facility would generate more freight flow in the surrounding region, increasing the economic activity in the region. Freight flows can also be considered inputs to other impacts, such as congestion, pavement deterioration, and accidents. For example, increased activity derived from the new intermodal facility could generate increased truck or at-grade rail traffic, which in turn could negatively affect highway network traffic unless further investments are made or policies changed. Also, some freight attracted to the intermodal facility might be diverted from facilities in other regions of the state, causing decreased economic activity in these regions.

Although such impacts might be qualitatively envisioned through more informal means, it would be preferable to support large-scale infrastructure and policy decisions with quantitative estimates of the potential flows. Such estimates could conceivably be produced from mathematical models based on explicit and understandable rules.

A similar motivation exists for predicting passenger movements. The field of passenger movement forecasting has matured to the point where there are well-established and widely-recognized theories, algorithms, data collection procedures, and institutional arrangements indicating who has what responsibility for the various components of the process. Like other states, Ohio has an operational, model-based system that is used to predict passenger flows. A similar model might be envisioned for freight flow modeling.

Yet, even the mature field of passenger flow forecasting is presently being reconsidered, and although forecasting freight movements appears to have many general similarities to forecasting passenger flows, there are important differences that complicate freight movement forecasting. For example:

- Freight transportation involves multiple commodities with distinct cost characteristics and time requirements.
- The unit of analysis may not be the same across transport modes.
- Freight shipments can be greatly consolidated, making independence assumptions among units of flow much less reasonable than in passenger flow modeling.
- Route and mode choice decisions are the result of decisions of both shippers and carriers.
- Freight cost functions are not always convex or even continuous; for instance, rail cost functions are concave (marginally decreasing with higher volumes) and discontinuous (jump to higher levels) when capacity is reached.
- Empty rolling stock is an important component of freight transportation, and special attention must be paid to backhauling operations.
- The confidentiality of some freight data bases can make model estimation difficult.

Such complexities make it unlikely that theoreticians and practitioners will soon agree on a comprehensive, operational freight movement model or set of models. Indeed, as we see in Section 3, there are many types of freight movement models being developed. Although most of these are developed by academics, there have been parallel modeling efforts undertaken at state departments of transportation. There have also been federally funded studies attempting to coalesce the various freight modeling efforts into practical guidance, but as of yet, we have seen no acknowledged acceptance of these guidelines by practitioners. Indeed, as we see in Section 2, responses to a survey of state departments of transportation indicate that states vary greatly in their experience with modeling.

1.2 Setting and Objectives

During the course of our work, some individuals proposed that the types of complexities mentioned above make it unlikely that freight movement models useful to practitioners could ever be developed. However, all models, including those dealing with passenger flows, are simplifications of reality; they will, therefore, never replicate existing conditions perfectly or produce error-free forecasts. Rather than evaluate the freight forecasting system by how closely it produces outputs exhibited in reality, the system can only be evaluated as to whether the benefits of using such a system outweigh its costs.

There have been no conclusive studies documenting that the benefits of modeling passenger flows outweigh the costs. Nevertheless, ongoing investments in passenger models and their widespread use demonstrate an important belief that the inaccuracies and uncertainties inherent with modeling the trip making process are overcome by the benefits of having these model outputs. Such revealed belief in the benefits of a flow forecasting system can only come with time, however, and we feel that it would be too difficult to determine whether a freight forecasting system would eventually demonstrate net benefits. A review of freight forecasting efforts and data availability (see Section 3) showed that specific freight models could be pursued. We, therefore, turned our attention to *assessing whether the time was appropriate to take steps* toward the implementation of a freight forecasting system and *identifying what these steps might be.*

We considered appropriate steps toward implementing a freight forecasting system that would be regularly used for analysis of freight issues to public agencies in Ohio. Discussions with individuals involved with freight transportation led us to conclude that the Ohio Department of Transportation (ODOT) would have primary responsibility for any regularly used freight system developed in the near future. As seen in Section 2, we investigated the desires of organizations other than ODOT, but we considered these desires so that ODOT could take appropriate steps toward a system that would be responsive to these needs.

There may be no more accurate means of understanding freight movements, and even of quantifying and forecasting the magnitudes of the movements, than that of using expert judgments. Indeed, if ODOT eventually commits to a formal freight forecasting system, and as such a commitment would move closer toward implementation, we recommend that appropriate *individuals be* formally *designated as an advisory group scheduled to discuss freight issues on a regular basis.* The benefits of such a group would lie in its ability to foresee structural changes that could not be forecast with a mathematical modeling system which assumes that past correlations hold in the future. Expert judgment would also be valuable in determining the values of parameters used in mathematical models, modifying such values estimated from other means, and checking the outputs of more traditional mathematical models for reasonableness.

Although we recognize the great benefits associated with using a designated group of experts to provide judgments on freight movements, we focused on steps toward implementing a mathematical forecasting system. Such a system can produce outputs and a formal structure that help clarify expert judgment. It is also more conducive to an ongoing, formal process that can be institutionally controlled by a lead organization. This type of process would also lead to systematic collection, processing, and interpretation of data that serve as inputs to the mathematical models, but that may also have secondary uses and indirect benefits. In addition, a mathematical modeling system gives the appearance of being less susceptible to individual biases than a system based solely on expert judgment. This perceived objectivity would help generate support for propositions based on the model results.

1.3 Design and Overview of Report

In summary, we focused our investigation on determining the desirability of taking steps toward implementing mathematical freight forecasting models eventually operated by ODOT that address planning issues of concern to Ohio and identifying what the next steps might be. To accomplish this, we investigated the perceived needs for a forecasting model in Ohio, freight forecasting efforts in other state departments of transportation, and the availability of models and freight databases. The results are presented in Sections 2 through 4, which serve as Part I of our report.

In Section 2 we report on the perceived needs for a forecasting model in Ohio and freight forecasting efforts in other states. We interviewed individuals in Ohio for the types of outputs they would desire from a freight forecasting model. The interviewees provided many, varied, and general responses; no single output was predominant. Table 2.2 summarizes the general types of outputs mentioned by the interviewees and the questions they said the outputs would help answer. We also present the results of our survey of freight forecasting efforts in other states in Section 2. Table 2.3 summarizes the responses to our survey. Similar to what was found in a survey conducted by Cambridge Systematics, Inc. *et al.*, (1997), few states consider themselves to be active in comprehensive freight forecasting efforts. However, we found that more than the few states mentioned in the Cambridge Systematics survey are involved with aspects of freight forecasting, and many more are involved with forecasting truck trips. Table 2.4 presents a more detailed, "conditional" breakdown of the responses. It shows that, other than for monitoring truck trips on the highway network, the same states seem to be involved with several aspects of freight forecasting.

In Section 3 we summarize our investigations of existing freight databases and models. We consider proposed models in the literature and modeling efforts underway in three states. The overall impression is that many freight models exist that could conceivably be employed in practice, but very few are. The efforts underway in the states investigated are similar. They are being developed by consultants and use a framework resembling the modular approach of the traditional 4-step urban transportation planning process. The intention is to develop intermodal and multimodal assignment models, but the present emphasis is on assigning truck trips to the highway network.

We draw conclusions on Part I in Section 4. We believe that there is presently little systematic freight modeling routinely conducted at state DOT's, but we expect activity to increase. We also believe that the activities in various states will probably resemble variations of the 4-step process and concentrate on the highway network in the near future. We propose that such a model could be developed in Ohio, but that it would be risky to simply develop an Ohio statewide freight model at the present time. We feel that statewide freight modeling is relatively new and expect that models produced in the near future will be somewhat inaccurate and unresponsive to users' needs, needs which we saw to be not well articulated in our survey. We, therefore, conclude that if ODOT is to pursue systematic freight modeling, it must be committed to sustained development that would allow improvements in model components to be implemented on a regular basis. There must also be an accompanying program of research, testing, and performance tracking of model components that would allow the improvements to be made. It would be advantageous to pursue this long-term development and refinement in formal collaboration with other states. Formal collaboration would reduce costs to the participating states, facilitate the sharing of expertise, and help sustain commitment to model development and refinement after initial efforts are concluded. It would also lead to a stronger influence on federal studies and data collection efforts.

In Part II, we consider a few simple, specific models or components that could either stand alone or be incorporated in more complex statewide models. In Section 5, we illustrate that indicator models could be used to forecast future commodity generation in Ohio. These forecasts could be useful in themselves or modified for use in a trip generation module of a larger model. In Section 6, we consider a model proposed to forecast freight usage at intermodal facilities. The model formulation is similar to that used for discrete choice analysis, which frequently forms the basis of components of the 4-step model. Like the indicator models of Section 5, the model proposed in Section 6 could presently be used in Ohio. However, we illustrate that different alternatives of the models in these two sections can produce very different forecasts, and it is not clear which alternatives would be most useful either as a stand-alone model or as a components in a statewide model.

In Section 7, we discuss three similar methods of updating truck origin-destination (OD) matrices from observed truck volumes. Since ODOT has recently completed a roadside survey from which a truck OD matrix could be estimated, and since truck volumes are routinely collected, it would be relatively inexpensive to use any of these procedures to maintain estimates of a statewide truck OD matrix. Simply determining an OD matrix that is consistent with observed traffic flows does not directly lead to forecasts of future conditions, but a good estimate of a current OD matrix could be used to calibrate or validate components of a larger-scale model. Moreover, accurate estimates of current OD patterns could assist experts when forecasting future patterns.

In Section 8, we show that recently available databases could be used with existing software to perform intermodal assignment. Although the quality of the assignments is presently limited by a lack of acceptable intermodal assignment logic, these databases could be used to develop, test, and experiment with intermodal assignment algorithms in the future.

In Section 9, we conclude by encouraging ODOT to pursue a statewide freight modeling if a commitment is made to sustained development and modification; the development efforts are similar to those that will be made in other states; and parallel efforts are made to investigate, test, and track the performance of alternative component formulations. We also recommend that ODOT try to formalize collaboration with other states by pooling funds to developed a regional model that could be scaled to the appropriate levels for the participating states. We further suggest that ODOT estimate a statewide truck OD matrix and use observed truck volumes on highway segments to update the matrix on a regular basis. Finally, if ODOT pursues systematic freight modeling, we encourage the formation of an advisory group consisting of experts who would ensure the relevance of model developments and increase the likelihood of practical use.

PART I: Needs, Surveys, and Reviews

Section 2. Background on Needs and Previous Efforts

In this section, we present what planners in Ohio expressed as desirable outputs from a freight forecasting model and responses to a survey of freight forecasting activities in other states. ODOT would want to ensure that any system developed would respond to the needs of those who would use its outputs. We, therefore, interviewed selected individuals in an open-ended, but structured manner to elicit what they thought would be useful outputs from a forecasting model and the types of questions these outputs would help answer. We see from the results in Section 2.1 that the individuals provided many and varied responses. Many of the responses are associated, as least indirectly, with forecasting freight on the highway network. Still, the responses were so different that they illustrate that there is presently no consensus as to what an Ohio freight forecasting model should produce.

In Section 2.2 we describe the survey of forecasting efforts in other states that we administered. The general impression obtained from the results is that no mature, widely accepted state DOT freight forecasting system exists; nor does it appear that one will likely emerge in the near future. However, it appears that there is interest and activity in freight forecasting and that this interest and activity may be increasing.

2.1 Local Interviews

Cambridge Systematics, *et al.*, (1997) surveyed state departments of transportation about needs they hoped freight forecasting models could fulfill. The needs most often cited by the 38 states responding to the survey are listed in Table 2.1.

One can see from this table that truck-related needs dominated the responses. Moreover, the first four needs, and to some extent the fifth need, listed in the table would all require some type of truck trip assignment model, i.e., a model that would forecast which highway segments would carry large volumes of truck traffic.

Table 2.1 Most often cited forecasting needs for a freight model

Forecasting Needs	Total No. of States
Highway needs analysis	36
Truck routes and restrictions	35
Highway planning	35
Truck size and weight regulations	34
Planning of truck/rail intermodal facilities	35 (*sic*)
Airport planning	31
Rail facility and access planning	31
Promotion of economic development	30

Source: Cambridge Systematics, Inc., *et al.*, 1997

The Cambridge Systematics survey of state needs is informative in at least two ways. It can help federal officials direct freight modeling research and development efforts that would be responsive to the needs of the states. It might also indicate where individuals undertaking various modeling efforts in the states might expect collaboration or shared expertise from other states. That is, it would be advantageous for a specific state to develop freight models of interest to other states, thereby raising the likelihood of being able to receive feedback from a greater pool of transportation professionals faced with similar circumstances and to share results, expertise, and experiences with these professionals.

Since our focus was on Ohio, we wanted to determine the perceived freight modeling needs of key individuals in Ohio. We, therefore, met with the individuals listed in Appendix 2.1 to discuss what they would want from a freight modeling system.

We allowed open-ended responses but tried to structure the discussions by asking a predetermined set of questions. Specifically, we first asked for specific outputs the individuals desired from freight forecasting models and for specific questions that these outputs would help answer. We emphasized that the individuals should not consider technical, economic, or institutional feasibility of producing these forecasts, but that they should think of what they would like to be able to know in terms of freight flow. We then described the types of outputs produced from each component of the traditional passenger flow 4-step process (Trip Generation-Trip Distribution-Modal Split-Traffic Assignment) and asked if they considered any of these to be valuable in freight context.

We found that our structured approach was useful in eliciting responses from the individuals, but not in eliciting specific model outputs desired. Rather, the responses to the questions on the types of outputs desired triggered many comments that demonstrated a general desire for freight modeling outputs. Except for the few individuals very familiar with transportation modeling, however, these questions did not elicit well-formulated specifications for model outputs. Our impression was that the individuals did not feel especially hampered by not having a specific forecast available but, when asked, could think of many general outputs they would consider useful. Perhaps due to this lack of well-formulated desires for specific outputs, reactions to the types of the outputs consistent with each component of the 4-step process did not help focus the needs any further, except for those few individuals who were very familiar with this process.

In Table 2.2 , we list the types of outputs mentioned in our interviews and the questions they might help answer. Since our questions encouraged open-ended responses, and these responses were not expressed as cogent, well-formulated model outputs, we had to edit and interpret answers to produce this table. However, we tried to limit our personal interpretations and avoided extending responses outside of what was explicitly stated. For example, some of the outputs listed in the table could be used in slightly different ways, but we do not list these different ways if they were not explicitly mentioned in the interviews. Similarly, one individual commented that knowing the time when intermodal flows would increase would be important for budgeting purposes. This timing issue would have similar importance for other model outputs, as well, but we only list it where it was explicitly stated.

Although we had hoped to use a list like that in Table 2.2 to help prioritize the types of freight forecasting models to be developed in Ohio, we strongly caution against doing so. Most of the questions listed in Table 2.2 could only be answered if freight movements on the highway network could be forecast, and developing truck models may, therefore, be a logical first step in developing a freight modeling system. Still, the conclusion that we wish to draw from this list is that there is no consensus of opinion on specific outputs desired from Ohio freight models. Moreover, when coupled with the general impression we received from these meetings, we do not believe that pursuing efforts to elicit freight forecasting needs of individuals in Ohio would be fruitful at this time. There has recently been much change in many state agencies, and we expect this to continue. Indeed, some of the individuals interviewed no longer hold the same positions they did when they met with us. Many of the needs alluded to seemed to stem from personal experiences, rather than needs associated with the position held by the individual. Therefore, this list would probably not be stable through time, and one would not wish to use it to fix long-term plans.

Table 2.2 Desirable freight modeling outputs stated in discussions and motivating questions

Modal diversion:
What is the impact, in terms of truck-induced highway congestion, revenues, ... of different rail investment scenarios?
What would be the diversion from truck to rail in response to increased highway costs?
How much freight expected at a new facility would be carried by modes other than trucks (to allow highway capacity planning around the facilities)?
How much air freight using other state airports would use Ohio airports in response to investments in facilities?
Hazardous material routes:
What hazardous material routes should be designated by policy makers?
How much hazardous material is/will be carried on specific highway segments?
Origin-destination data:
How many truck trips will be destined for specific communities?

Table 2.2 Desirable freight modeling outputs stated in discussions and motivating questions (continued)

	How much truck traffic could use local by-pass routes?
	In which corridors are there/will there be significant shipments of commodities that could be carried by both truck and rail (to determine "intercept" markets for intermodal transport)?
	Which corridors have significant truck traffic (to prioritize investments in highways corridors)?
	What OD table should be used in local assignment models?
	What OD table should be used in state-wide assignment models?
Truck assignment data:	
	What are pavement and geometric requirements on highway segments?
	How much traffic is expected at railroad-highway grade crossings?
	How many empty trucks are using a segment (to determine potential for backhaul traffic)?
	Which segments should receive snow and ice removal priority (to facilitate carrying high-value, time-sensitive goods)?
	What quantities of commodities that could be carried by rail are/will be carried on highway routes that parallel existing rail infrastructure?
	What will be future levels of truck-induced pollution under various highway investment scenarios?
	What will be future highway capacity needs?
	What will be truck volumes for alternative projects (to be used for project prioritization)?
Rail movements:	
	How much traffic is expected at railroad-highway grade crossings?
	What will be rail freight schedules (to allow coordination with intermodal movements)?
Commodity levels:	
	How much freight, by commodity type, is/will be traveling in Ohio?

Table 2.2 Desirable freight modeling outputs stated in discussions and motivating questions (continued)

Truck response to tolls:
How many trucks will divert from Ohio toll roads in response to increased tolls?
Intermodal data:
How much in- or outbound freight at intermodal facilities will be carried by truck (for highway capacity planning)?
How much intermodal freight will use a potential or existing facility?
When will critical levels of intermodal freight use a facility (to allow for capital budgeting)?
Aggregate regional data:
How much freight originates or terminates in a region (to help prioritize regional investments targeted toward job retention)?

2.2 DOT Surveys

As mentioned above, it would be advantageous if freight forecasting efforts in Ohio paralleled those of other states. This is even more true, since our interviews in Ohio revealed such varied desires on the types of outputs to be produced. We, therefore, investigated general efforts underway in other state departments of transportation. We did this by investigating available literature and by surveying the states.

Although some of the published literature mentioned models with state names attached, we discovered that these could not be taken as indicative of efforts ongoing in the states. After contacting some states, we found that some models associated with states in the literature were either no longer used by the states or never were. In addition to surveying states for their freight forecasting needs, Cambridge Systematics, *et al.*, (1997) also asked for the freight forecasting efforts undertaken by the states. They concluded that most states had "little or no experience in freight forecasting," mentioning Iowa, Oregon, and Washington as exceptions.

We did not wish to duplicate the Cambridge Systematics, Inc. (CSI) survey. However, based on the discussions with Ohio personnel discussed above and our review of freight models (see Section 3), we were beginning to believe that any freight model pursued by a DOT would have to be highly modular, with the possibility of developing components that may be useful in themselves, as well as when combined with other components in a more complex model. Therefore, we designed and administered our own survey of the states.

Based on our model review (see Section 3), the CSI survey of state desires, and our local interviews, we felt that the following components would be the most likely to be contained in freight models developed at ODOT in the near future: 1) a component that would correlate aggregate freight traffic with economic "indicators;" 2) a component that produces a freight-based origin-destination matrix; 3) forecasts of truck trips on highway segments. We wanted to keep the survey short to encourage response. Therefore, we limited our questions to these components. The survey and accompanying cover letter can be found in Appendix 2.2.

Through various means we found contacts that were likely to be associated with freight modeling at the state departments of transportation. We contacted these individuals by email or fax with a preliminary letter (see Appendix 2.2), asking to whom we should send the survey and whether they preferred to receive the survey by email or fax. We received 37 responses to this inquiry and sent out surveys to the names provided in these 37 responses. We received 23 (62.1%) completed surveys from the 37 we sent out. A list of responding agencies and the contact points for these agencies are provided in Appendix 2.2.

The responses are summarized in Table 2.3. Some of the responses confirm what was found in the CSI survey, but some give a different impression. Specifically, as in the CSI survey, we found few states actively involved with most aspects of freight modeling.

From Table 2.3, it is especially noteworthy that 16 of the 23 respondents said that they never track the correlation between freight data and indicator variables (Question 1), and that the other 6 said that they tracked this correlation "sometimes, but not regularly." In the same way, only eight of the 23 respondents have or produce freight OD tables (Question 5), and only one of these said that that they update these regularly (every five years).

Table 2.3 Responses to survey of freight modeling efforts in state agencies (23 states returned questionnaire)

Questions	Responses
Question 1. Track Correlation between Freight Data and Indicator Variables	
Regularly	1 (Illinois)
Sometimes	6 (Alaska, Indiana, Iowa, Kansas, Minnesota, and Oregon)
Never	16
Question 2. Type of Freight Data Tracked	Commodity Flows and Truck Weight, Tons Moved by Origins & Destinations by Commodity, Daily Traffic Counts, and Vehicle Miles Traveled.
Means to Obtain Freight Data	
Special Surveys	4 (Alaska, Illinois, Minnesota, and Iowa)
Private Data Supplier	2 (Minnesota and Iowa)
Public Data Source	7 (Alaska, Illinois, Indiana, Iowa, Kansas, Minnesota, and Oregon)
Other	2 (Illinois and Indiana)
Question 3. Type of Correlated Variables Tracked	Fuel Consumption, Economic Indicators, Agricultural Production, and Demographic Indicators (e.g., Population, Employment).
Means to Obtain Indicator Variables	
Special Surveys	1 (Kansas)
Private Data Supplier	2 (Iowa and Kansas)
Public Data Source	3 (Iowa, Kansas, and Oregon)
Other	2 (Illinois and Indiana)

Table 2.3 Responses to survey of freight modeling efforts in state agencies (23 states returned questionnaire) (continued)

Questions	Responses
Question 4. Explicit Analysis of Correlation among Freight Data and Other Variables	
Yes	3 (Illinois, Kansas, and Oregon)
No	4
Type(s) of Explicit Analysis	
Graphing Trends	2 (Indiana and Kansas)
Calculating Ratios	3 (Illinois, Indiana, and Kansas)
Regression Analysis	1 (Kansas)
Time Series Analysis	1 (Oregon)
Other	1 (Oregon)
Question 5. Has/Uses Freight O-D Table	
Yes	8 (Colorado, Illinois, Indiana, Iowa, Kansas, Maine, Minnesota, and Oregon)
No	15
If Yes, Updated How Often?	
Every Five Year	1 (Maine)
Irregularly	7
If Yes, O-D Table Kept by Commodity?	
Yes	7 (Colorado, Illinois, Indiana, Iowa, Kansas, Minnesota, and Oregon)
No	1
If Kept by Commodity, Type of Category	
STCC	5 (Colorado, Indiana, Iowa, Minnesota, and Oregon)
SITC	0
Commodity Names (e.g., Coal, Grain)	3 (Illinois, Indiana, and Kansas)
Other	0
Associated Commodity Units	
Tons/Year	4 (Colorado, Illinois, Iowa, Kansas)
Carload/Day	3 (Illinois, Indiana, and Oregon)
Means to Update the O-D Tables	
Special O-D Survey	3 (Illinois, Kansas, and Oregon)
Models	4 (Indiana, Kansas, Maine, and Oregon)
Other (e.g., Receive and/or Purchase Information; Public Data Sources)	4 (Colorado, Illinois, Iowa, and Minnesota)

Table 2.3 Responses to survey of freight modeling efforts in state agencies (23 states returned questionnaire) (continued)

Question 6. Monitor Truck Trips on Highway Network Regularly	
Yes	14 (Alaska, Arizona, Colorado, Connecticut, Illinois, Indiana, Iowa, Kansas, Kentucky, Maine, Minnesota, North Dakota, Oregon, and West Virginia)
No	9
Forecast Truck Trips	
Yes	13 (Arizona, Colorado, Connecticut, Illinois, Indiana, Iowa, Kansas, Kentucky, Maine, Minnesota, North Dakota, Oregon, and West Virginia)
No	1
Methods Used to Forecast	
Trend Projection	9 (Arizona, Colorado, Connecticut, Indiana, Iowa, Kentucky, Minnesota, North Dakota and West Virginia)
Correlation	2 (Iowa and Minnesota)
Trip Assignment	6 (Connecticut, Iowa, Kansas, Kentucky, Maine, and Oregon)
Other	3 (Illinois, Minnesota, and West Virginia)

The results differ from the CSI survey when considering the responses to Question 6, however. It was not surprising that more than half the respondents said that they "regularly monitor truck trips on highway links." (We were actually surprised that more than 14 of the 23 respondents did not respond affirmatively, since such activity is an important activity of most state DOT's. Therefore, there may have been some misunderstanding in this question, and one must be careful in drawing too strong a conclusion from the responses.) However, collecting these data does not necessarily mean that the states would also forecast truck trips. Yet, almost all (13 of 14) states monitoring truck trips also said that they forecast truck trips.

We were also interested in the data sources and methods used by those responding affirmatively to the questions. The states that track correlation between general freight data and economic indicators mentioned Commodity Flows, Truck Weight, Tons Moved by Origins & Destinations by Commodity, Daily Traffic Counts, and Vehicle Miles

Traveled as freight variables and Fuel Consumption, Economic Indicators, Agricultural Production, and Demographic Indicators as indicator variables. Although private data suppliers and special surveys were used, public data sources were almost always used. The public sources most often mentioned were the 1993 Commodity Flow Survey (Bureau of Transportation Statistics, Department of Transportation), BEA Regional Projections to 2040 Publication (Bureau of Economic Analysis), and Estimated Waterborne Commerce Statistics Publication (US Army Corps of Engineers). The two private sources listed on the surveys were Transearch (Reebie Associates) and Woods & Poole Economics, Inc. Only three of the seven states responded that they "explicitly analyze" the correlation (Question 4). Again, it is possible that different respondents interpreted this question to mean different things, but we intended it to mean -- and expect the respondents to have meant -- a systematic effort to assess the quality of the correlation, as opposed to simply calculating some arbitrary correlation to be used (see Section 5).

Eight of the 23 states said that they had or produced freight OD tables (Question 5). Some of these stated that these tables were kept by commodity, with most keeping these by Standard Transportation Commodity Code (STCC). A list of the STCC is provided in Appendix 2.2. Four of these eight said that they used models to produce the OD tables, with Commodity Flow, Network Flow Transshipment, and Economic Input-Output Models mentioned explicitly.

Trend projection was the most commonly stated means of forecasting truck trips on highway links (Question 6), but assignment techniques were also mentioned frequently.

We also looked at "conditional" responses to the survey. Specifically, in Table 2.4 we present the number of states responding affirmatively to aspects of freight modeling, conditional upon the number that responded affirmatively to other aspects. For example, the first row of numbers in the table shows that of the seven states which said that they conduct some type of analysis of economic indicators, (seven conduct some type of analysis of economic indicators), six estimate some type of freight origin-destination matrix, all seven monitor truck traffic on highway segments, and six forecast truck traffic on highway segments. As another example, the entries in the row labeled "Indicators, O-D matrices (6)," mean that of the six states that both analyze economic indicators and estimate freight origin-destination tables, (all six analyze economic indicators), (all six estimate freight origin-destination tables), all six monitor truck traffic on highway segments, and all six forecast truck traffic on highway segments.

The entries in the table indicate that the same six or seven states are involved with most of the freight modeling activity. Several other states are involved with monitoring and forecasting truck traffic on highway segments, but if a state is going to be involved with some other aspect, it seems that it will be involved with several aspects.

Table 2.4 Number of states conducting aspects of freight modeling, conditional on having conducted other aspects of freight modeling

	Indicators	O-D Matrices	Monitor Highway Links	Forecast Highway Links
Indicators (7)	7	6	7	6
O-D Matrices (8)	6	8	8	8
Monitor (14)	7	8	14	13
Forecast (13)	7	8	13	13
Indicators, O-D Matrices (6)	6	6	6	6
Indicators, Monitor (7)	7	6	7	6
Indicators, Forecast (6)	6	6	6	6
O-D Matrices, Monitor (8)	6	8	8	8
O-D Matrices, Forecast (8)	6	8	8	8
Monitor, Forecast (13)	7	8	13	13
Indicators, O-D, Monitor (6)	6	6	6	6
Indicators, O-D, Forecast (6)	6	6	6	6
Indicators, Monitor, Forecast (6)	6	6	6	6
O-D, Monitor, Forecast (6)	6	6	6	6
Indicators, O-D, Monitor, Forecast (6)	6	6	6	6

Section 3. Review of Models and Databases

When beginning this project, we had not seen reviews of models or available data for operational use. We, therefore, began searching the literature for proposed freight models and looking on the internet for various public and private data sources. During our searches, other helpful reviews appeared. Moreover, we determined that there were far too many *proposed* data sources and models to review during this project in any meaningful way. We, therefore, did not attempt to be comprehensive in our reviews. Rather, we focus in this section on mentioning other reviews and on summarizing the characteristics of selected databases that were mentioned in responses to our survey (see Section 2) and on selected ongoing efforts in other state Departments of Transportation.

3.1 Review of Databases

Data could be useful in developing freight models in several ways.

1) Data on the freight movement variable (dependent variable) could be extrapolated in time to form forecasts. These extrapolations could be formal, using time series or trend analysis techniques. Or, they could be more informal, based on holistic judgments after observing past and present data values.

2) Data on inputs (independent variables) to some existing freight movement model may be used to run the models under alternative infrastructure or policy scenarios, "predicting" the characteristics of freight movements under these alternative scenarios. If the data only exists for present and past time periods, freight movements that "could have existed" under the various scenarios could be predicted. Or, like the dependent variables, the historical data could be extrapolated into the future to predict the freight movements under the various alternatives under future conditions. On the other hand, if future values of the independent variable are presently forecast by other organizations, these values could be used in the analysis.

3) Paired data on dependent (output) and independent (input) variables could be used to investigate correlation between the variables and to test the degree of correlation achieved with various models. Investigating correlation could be relatively informal, e.g., plotting the paired data to see if they appear to be correlated. Or, the investigation could be more formal, e.g., using the paired data to calibrate parameters of a mathematical relation correlating the freight

> movement (dependent) variables with the (independent) variables thought to influence the freight movement. Tests of correlation or statistics summarizing the fit of the model to the data may be used to help determine if a given model should be pursued, to choose among competing models, or to quantify the amount of uncertainty in the forecasts.

Therefore, we thought it useful to investigate existing databases of variables that might be relevant to freight forecasting models. When beginning this project, we had not seen reviews of databases that could be used for freight forecasting. We, therefore, began searching the literature and the internet for various public and private data sources. However, during our searches, we found that good database reviews had previously been conducted.

Cambridge Systematics, Inc., *et al.* (1997) reviewed and summarized approximately 50 freight databases from various public and private data sources. Most of the data sources that they listed are also available in the Directory of Transportation Data Sources (BTS) of the U.S. Department of Transportation. The authors tabulated the selected data sources into different categories based on key characteristics, such as the modes, commodities and types of movements covered in the data sources.

Fang, *et al.* (1996) investigated a number of available data sources that could be useful in helping to forecast commodity flows between the U.S. and Mexico. The authors listed seven major sources, along with brief descriptions of other data sources reviewed by other study groups. They concluded that the accuracy of the predictive models depends greatly on the quality of the available data.

The authors of the *Quick Response Freight Manual* (Cambridge Systematics, Inc., *et al.*, 1996) listed state data centers from which users can obtain many freight-related data when forecasting or modeling freight movements. Two data centers were mentioned for Ohio:

The Ohio Department of Development: The Ohio Department of Development provides both historical and forecasting data on the general demographic statistics and regional economic growth.

The Ohio State University, School of Public Policy and Management: The School of Public Policy and Management at the Ohio State University published *Benchmark Ohio* (Shkurti and Bartle, 1991). This publication tabulates a series of statistical indicators in Ohio, such as population and economic variables, and provides a guide to public sources with information on the categories (e.g., education, taxation, etc.) contained in the publication.

In addition to the state data centers, the authors also listed a number of other public and commercial freight data sources. They categorized each source in terms of the perceived usefulness to a quick response freight modeling process.

We review the most often mentioned databases from our survey (see Section 2) and those databases that we discuss in other sections of this report. In addition to a brief summary of each database, examples to illustrate the layout of the databases are also provided. We group the databases according to whether they are available from a public agency or private organization.

Public Data Sources: In general, data from public sources are usually easier to obtain and less expensive than data from private sources. The majority of the public sources mentioned in our Section 2 survey and used by us in Part II are provided by the U.S. Department of Transportation (DOT) and related federal agencies. Moreover, many of the public databases, are available over the Internet or can be obtained from the agencies on other media (CD-ROM, Prints) for free.

Commodity Flow Survey 1993 Database: The *1993 Commodity Flow Survey* was conducted by the Census Bureau with major funding provided by the U.S. DOT. Flows of goods and materials from origins to destinations within the United States are presented in this database. Both detailed and aggregated data are prepared at different levels of aggregation by geographic area (e.g., origin-destination pairs at the County, National

Transportation Analysis Region (NTAR), or State level). Commodity information is categorized according to Standard Transportation Commodity Codes (STCC) up to 5-digits. Other information collected for the sampled shipments are mode of transport, weight, value of the shipment, and ton-miles. Table 3.1a provides an example for a summary table from the database. This table indicates the "total" commodities shipped into Ohio from neighboring states. An excerpt from the database, indicating shipment characteristics by commodity (indicated by STCC codes) and mode of transportation (indicated by different modes), can be found in Table 3.1b.

Table 3.1 Examples of 1993 Commodity Flow Survey Data

3.1a Example of 1993 Commodity Flow Survey Data: Shipments destined for Ohio from neighboring states

State of Origin	Value (million $)	Weight (thousand tons)	% value of state's shipments	% weight of state's shipments
Indiana	14,299	11,258	8.0	3.9
Kentucky	8,985	30,161	8.0	8.5
Michigan	20,735	26,873	8.1	8.3
Pennsylvania	11,288	15,705	4.5	3.8
West Virginia	4,136	27,520	11.8	11.7

Source: *1993 Commodity Flow Survey*, TC92-CF, 1996

3.1b Example of 1993 Commodity Flow Survey Data: 1993 Shipment characteristics by commodity and mode of transportation for Ohio

STCC code/Description/Modes	Value		Tons		Ton-miles		
	Number (million $)	%	Number (thousands)	%	Number (millions)	%	Average miles per shipment
All Commodities Total	325,626	100.0	469,652	100.0	89,974	100.0	362
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
STCC 14, Nonmetallic Minerals							
Total	1,225	100.0	128,639	100.0	6,357	100.0	41
Single Modes							
Parcel, U.S. Postal Service	(S)	(S)	---	---	---	---	(S)
Private truck	681	55.6	81,049	63.0	3,309	52.1	34
For-hire truck	505	41.2	40,494	31.5	2,161	34.0	55
Air	---	---	---	---	---	---	---
Rail	14	1.1	3,150	2.4	647	10.2	21.4
Inland water	(D)	(D)	(D)	(D)	(D)	(D)	(D)
Great Lakes	(D)	(D)	(D)	(D)	(D)	(D)	(D)

3.1b Example of 1993 Commodity Flow Survey Database: 1993 Shipment characteristics by commodity and mode of transportation for Ohio (continued)

STCC code/Description/Modes	Value		Tons		Ton-miles		
	Number (million $)	%	Number (thousands)	%	Number (millions)	%	Average miles per shipment
Deep sea water	---	---	---	---	---	---	---
Pipeline	---	---	---	---	---	---	---
Multiple Modes							
Private and For-hire trucks	---	---	(S)	---	---	---	(S)
Truck and air	---	---	---	---	---	---	---
Truck and rail	---	---	---	---	---	---	---
Truck and water	---	---	---	---	---	---	---
Truck and pipeline	---	---	---	---	---	---	---
Rail and water	---	---	---	---	---	---	---
Inland water and Great Lakes	---	---	---	---	---	---	---
Inland water and deep sea	---	---	---	---	---	---	---
Other Modes							
Other and unknown modes	19	1.5	2,635	2.0	168	2.6	84
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
STCC 39, Miscellaneous Products of Manufacture							
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

---Represents zero or less than 1 unit of measure
(D) Denotes figures withheld to avoid disclosing data for individual companies
(S) Data do not meet publication standards due to high sampling variability or other reasons
Source: *1993 Commodity Flow Survey*, TC92-CF-36, Table 6, 1996

Bureau of Economic Analysis Regional Accounts Databases: Several databases are prepared by the Bureau of Economic Analysis (BEA), which is a U.S. Department of Commerce agency. The *Gross State Product* database covers the years from 1977 to 1994 and is categorized by industry. Table 3.2a provides a sample entry of the database format.

The *State Personal Income* database includes tables for the years 1929 to 1996. These are prepared on an annual basis, while quarterly tables are available from 1969 to 1997. Table 3.2b demonstrates the format used for the income database.

The *Local Area Personal Income and Per Capita Personal Income* database contains average and total wages at the county and Metropolitan Statistical Area (MSA) levels on a per capita basis. This information is available from 1969 to 1996. Table 3.2c provides an excerpt from this database.

The *Projections* database currently includes data starting in 1969 and projected to the year 2045. Projections are made 50 years in the future, and the BEA updates its U.S. economic and population projections every five years. The primary database components are Gross State Product, Employment, and Income data. Illustrations excerpted from the *Projections* database are provided in Table 3.2d.

Table 3.2 Examples of Bureau of Economic Analysis Regional Accounts Data

3.2a Example of Bureau of Economic Analysis Regional Accounts Data: Gross State Product 1977 - 1994

Industry	1977 (million dollars)	...	1993 (million dollars)	1994 (million dollars)
Total Gross State Product	97,740	...	256,050	274,844
Farms	1,059	...	1,730	2,121
Mining	1,053	...	1,154	1,238
Nonmetallic minerals	143	...	300	304

Source: U.S. Department of Commerce, BEA, Regional Economic Analysis Division

3.2b Example of Bureau of Economic Analysis Regional Accounts Data: Per Capita Personal Income, by state and regional, 1996

Region	Income (dollars)	% of national average	Dollar difference from national average	U.S. Rank	% change 1995-1996
U. S.	24,426	100	0	...	4.6
Great Lakes					
Illinois	26,848	110	2,422	7	4.9
Michigan	24,954	102	519	16	3.7
Ohio	23,457	96	-969	21	4.0
Wisconsin	23,320	95	-1106	22	4.2
Indiana	22,601	93	-1825	28	4.1

Source: BEA, U.S. Department of Commerce

3.2c Example of Bureau of Economic Analysis Regional Accounts Data:
Ohio personal income and per capita personal income

Area Name	1994 personal income (million $)	1995 personal income (million $)	% change 1994-1995	1994 per capita income ($)	1995 per capita income ($)	Rank (within state)
Ohio	236,544	250,865	6.1	21,317	22,531	---
Metropolitan region	199,903	212,308	6.2	22,179	23,508	---
Non-metro. Region	36,641	38,557	5.2	17,587	18,338	---
Franklin Co.	23,874	25,410	6.4	23,787	25,193	5

Source: USDOC/BEA, Survey of Current Business, August 1997

3.2d Example of Bureau of Economic Analysis Regional Accounts Database:
Projections of Employment, Income, and GSP in Ohio (1969-2045)

	1969	...	1992	...	2045
Employment (thousand jobs)					
All-industry total	4,687.9	...	5,881.7	...	8,012.4
Farm	126.3	...	99.2	...	74.5
⋮	⋮		⋮		⋮
Income (million of $87)					
All-industry total	94,391.7	...	121,203.3	...	239,226.2
Farm	1,311.1	...	936.9	...	1,250.8
⋮	⋮		⋮		⋮
GSP (million of $87)					
All-industry total	(NA)	...	203,155.0	...	425,938.7
Farm	(NA)	...	2,040.4	...	3,565.1
⋮	⋮		⋮		⋮

Source: Bureau of Economic Analysis, U.S. Department of Commerce

Estimated Waterborne Commerce Statistics Publications: These publications contain statistics on the commercial movement of foreign and domestic cargo. Several publications and data are prepared by the United States Army Corps of Engineers (USACE). The *Internal U.S. Waterway Monthly Tonnage Indicators* database presents trends in commodity tonnage flows from January 1994 to February 1998. The "Internal" qualifier denotes that the commodities moved solely within the boundaries of the U.S. Four different tonnage indicators (total monthly indicator, coal indicator, petroleum and chemicals indicator, and food and farm products indicator) are included in this database and updated on a monthly basis. The data are presented in both tabular and graphical formats. Table 3.3a and Figure 3.1a illustrate the total monthly tonnage indicator in this database.

The *Internal U.S. Waterway Tonnage Comparisons* database presents short tons and percent change in short tons between adjacent years for both fiscal and calendar years. Table 3.3b gives an example of the database comparing calendar year 1997 and 1996.

The *Final Estimated Waterborne Commerce Statistics* database is prepared and published annually. The most recent report that we found is for calendar year 1996. In this database, national tonnage and tonnage on selected inland waterways are reported on a yearly basis from 1987 to 1996. Additional national summaries of the U.S. waterways and domestic and foreign harbors are also provided in the database. Data are presented in both tabular and graphical formats. Table 3.3c shows annual traffic on the Ohio River (Upbound) from 1987 to 1996, categorized by commodity. Upbound, downbound, and total commerce are all listed for the selected inland water. Table 3.3d and Figure 3.1b illustrate the total monthly tonnage indicators in this database.

The *1996 Waterborne Tonnage for Principal U.S. Ports and all 50 States and U.S. Territories* database reports tonnage for domestic, foreign (imports and exports), and intra-State waterborne traffic at the selected U.S. ports. This information is presented in a tabular format. Table 3.3e and Table 3.3f illustrate the database format.

Table 3.3 Examples of Estimated Waterborne Commerce Statistics Publications Data

3.3a Example of Estimated Waterborne Commerce Statistics Publications Data: Total monthly tonnage indicator 1994-1998

Total Tons (millions)												
	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec
1994	39.9	41.0	49.7	51.1	52.4	50.0	50.1	53.0	51.9	55.6	54.0	51.4
1995	49.5	44.7	52.6	53.4	45.1	47.2	53.3	52.1	51.8	55.2	54.4	50.5
1996	42.7	41.1	52.8	51.5	49.5	50.0	52.3	48.8	47.7	52.8	53.0	44.6
1997	39.5	44.3	44.2	46.9	52.3	50.4	49.6	50.8	47.9	55.4	52.3	49.0
1998	41.8	43.8										

Source: Waterborne Commerce Statistics Center, 1998

3.3b Example of Estimated Waterborne Commerce Statistics Publications Data: Internal U.S. waterways short ton comparisons:

	CY 1996	CY 1997	Percent Change
National Domestic Total	1100.7	1111.7	1.0
National Lakewise Total	114.9	123.1	7.2
National Coastwise Total	267.4	265.8	-0.6
National Internal Total:			
All Internal Commodities	622.1	625.8	0.6
Food and Farm	89.3	85.2	-4.6
Coal	176.3	178.8	1.4
Chemicals	52.1	53.4	2.5
Petroleum	151.8	147.2	-3.0
Metal	29.2	29.9	2.6
Other	123.4	126.4	2.4
Waterways:			
Alabama-Coosa River	0.7	0.7	3.5
Allegheny River	3.3	4.0	18.4
Apalachicola River	0.6	0.5	-17.3
Atlantic Intracoastal	4.3	3.9	-9.0
Black Warrior River	24.9	26.0	4.6
Columbia River	18.3	19.2	5.0
Cumberland River	17.2	23.0	34.0
Gulf Intracoastal	118.0	120.4	2.0
Illinois Waterway	46.2	45.4	-1.9
Kanawha River	24.8	24.7	-0.5
McClellan-Kerr Wtwy	10.6	11.7	10.6
Mississippi River	319.6	317.6	-0.6
Missouri River	8.2	8.2	0.0
Monongahela River	36.6	37.3	2.0
Ohio River	237.7	238.2	0.2
Snake River	5.7	6.0	5.2
Tennessee River	45.5	48.7	7.0
Tennessee Tombigbee	8.0	8.5	6.8

Source: Waterborne Commerce Statistics Center, 1998

3.3c Example of Estimated Waterborne Commerce Statistics Publications Data: Commerce on Ohio River - Upbound, by commodity, 1987-1996 (Millions of tons)

Year	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
Total	88.34	85.52	88.98	101.40	93.48	96.15	107.65	110.87	107.61	109.45
Coal	50.65	40.21	43.63	52.94	48.82	49.56	53.13	55.47	53.09	54.40
Petro & Chem	19.93	21.51	21.68	19.81	17.73	18.51	19.64	20.53	18.87	17.92
Nonmetal	8.54	13.86	13.28	15.47	14.46	15.16	19.01	17.71	19.01	20.71
Other		10.23	10.40	13.17	12.47	12.92	15.87	17.16	16.65	16.41

Source: Waterborne Commerce Statistics Center, 1997

3.3d Example of Estimated Waterborne Commerce Statistics Publications Data: Commerce on U.S. total and selected waterways and internal traffic, 1987-1996 (millions of tons)

Year	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
Total Internal	569.83	588.12	606.01	622.60	600.39	621.04	607.25	618.41	620.32	621.88
Mississippi	293.23	298.76	298.87	306.19	301.67	315.71	298.26	314.58	323.02	318.46
Ohio	197.17	192.59	202.67	224.70	218.32	226.39	227.24	236.66	234.06	236.84
Tennessee	41.71	47.10	43.06	44.51	42.09	46.08	48.16	49.14	46.39	45.53
Illinois	41.41	40.97	39.67	43.30	43.11	42.67	45.64	50.88	47.43	46.24

Source: Waterborne Commerce Statistics Center, 1997

3.3e Example of Estimated Waterborne Commerce Statistics Publications Data: Tonnage for selected U.S. ports in 1996 ranked by total tons

Rank	Port_Name	Total	Foreign	Imports	Exports	Domestic
1	South Louisiana, LA	189,814,564	83,769,483	25,172,134	58,597,349	106,045,081
2	Houston, TX	148,182,876	87,058,288	58,041,465	29,016,823	61,124,588
3	New York, NY & NJ	131,601,244	56,485,614	48,472,360	8,013,254	75,115,630
⋮	⋮	⋮	⋮	⋮	⋮	⋮
40	Cleveland, OH	16,720,837	3,977,549	3,367,610	609,939	12,743,288
41	Lorain, OH	15,977,949	121,947	121,947	-	15,856,002
42	Portland, ME	15,242,802	13,369,237	13,289,315	79,922	1,873,565
⋮	⋮	⋮	⋮	⋮	⋮	⋮
149	Huron, OH	1,003,830	13,485	10,178	3,307	990,345
150	Redwood City, CA	985,392	513,392	227,175	286,217	472,000

Source: Waterborne Commerce Statistics Center, 1998

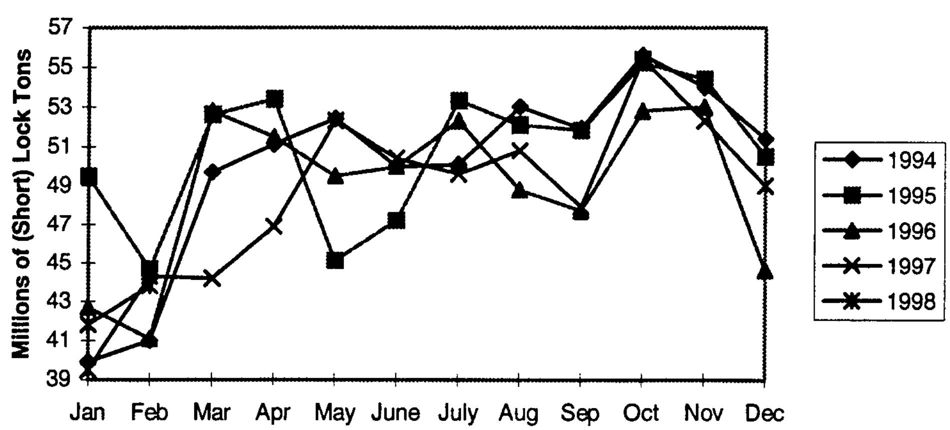
3.3f Example of Estimated Waterborne Commerce Statistics Publications Data: CY1996 Waterborne tonnage by state (in units of 1000 tons) sorted by tons

STATE		Shipping to		Receiving From		
	Totals	Domestic	Foreign	Domestic	Foreign	Intra-state
Louisiana	494,249	99,374	114,616	133,560	104,692	42,007
Texas	385,585	48,533	53,765	23,943	209,355	49,988
California	181,165	6,779	46,925	52,100	48,075	27,285
Ohio	123,459	24,247	13,535	61,193	5,762	18,722
Florida	117,430	13,718	18,337	58,352	23,056	3,968
⋮	⋮	⋮	⋮	⋮	⋮	⋮
Vermont	0	0	0	0	0	0
		Interstate	Exports	Interstate	Imports	Intra-state
Totals	2,284,063	767,715	450,794	767,715	732,592	332,962

Source: Waterborne Commerce Statistics Center, 1998

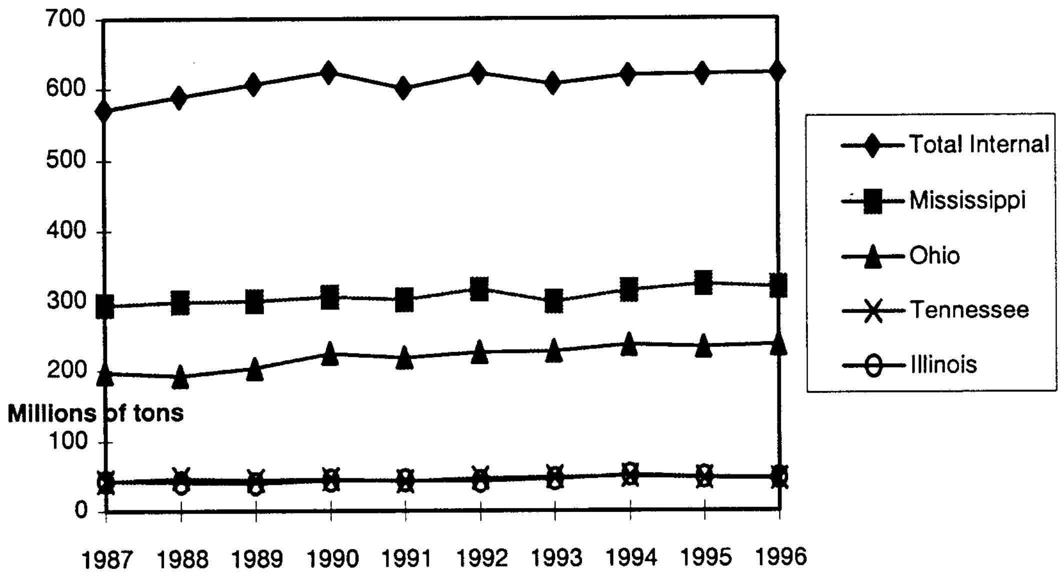
Figure 3.1 Examples of Estimated Waterborne Commerce Statistics Publication Data

3.1a Example of Estimated Waterborne Commerce Statistics Publications Data: Total monthly tonnage indicator 1994-1998



Source: Waterborne Commerce Statistic Center, 1998

3.1b Example of Estimated Waterborne Commerce Statistics Publications Data: Commerce on U.S. total and selected waterways and internal traffic, 1987-1996



Source: Waterborne Commerce Statistics Center, 1997

ICC Carload Waybill Sample Database: The *ICC Carload Waybill Sample Database* contains detailed information on Class I freight railroads. This database was prepared by Association of American Railroads, and the information is based on a one percent sample of rail waybill data from 1988 to 1992 supplied to the Interstate Commerce Commission (ICC) on actual rail shipments. The information includes commodities carried; railroad involved; origin, destination, and junction points; number of carloads; tons transported; and total revenues. For confidentiality reasons, the public-use version of the sample only contains movements reported at the BEA-to-BEA level and commodities reported at the 5-digit STCC code level. Table 3.4 provides an example of the public-use data. This example contains rail shipments from and to Ohio.

Table 3.4 Example of ICC Carload Waybill Sample Database: Rail shipments of the five largest commodities (ranked by weight) from and to Ohio (1992)

Commodity	Tonnage	Percent of state total
Originated within Ohio		
Metallic ores	11,975,461	19
Primary metal products	8,434,228	14
Coal	7,227,043	12
Farm products	7,118,235	12
Nonmetallic minerals	5,095,968	8
Terminated within Ohio		
Coal	33,620,329	41
Primary metal products	7,870,137	10
Chemicals	6,897,453	8
Nonmetallic minerals	5,765,024	7
Metallic ores	5,039,708	6

Source: Rail Waybill Data, compiled by DOT Federal Railroad Administration (Washington, DC:1996)

National Transportation Statistics, Annual Report: The *National Transportation Statistics, Annual Report* contains transportation statistics and mode profiles along with economic, safety, energy, and environmental information. Since data are aggregated at the national level, it would not be a very good source for state-wide freight modeling. Table 3.5 contains an example from the transportation and economy section.

Table 3.5 Example of National Transportation Statistics, Annual Report Data: Employment in for-hire transportation by mode: 1983-1995 (in percent)

Mode	1983	1995
Trucking	44.5	47.6
Air	16.6	20.1
Transit	9.3	10.8
Services	8.3	10.5
Railroad	13.7	6.1
Water	6.9	4.4
Pipeline	0.7	0.4

Source: Table 2-5, Transportation Statistics Annual Report, BTS, U.S. DOT, 1997.

1992 Census of Transportation: The Bureau of Census prepares the Census of Transportation, Communications, and Utilities database to provide periodic (every five years), detailed data on transportation, communication, and utility establishments and activities. This database is specialized to cover only those commodities categorized from Standard Industry Commodity Codes (SIC) 40 to SIC 49, except railroads (SIC 40), the U.S. Postal Service (SIC 43), and large certificated passenger air carriers (part of SIC 4512). Data are prepared and updated for the entire United States, each state, the District of Columbia and Selected Metropolitan Statistical Areas (MSAs) on a five year basis for years ending in "2" and "7." Table 3.6 demonstrates the layout of this database.

Table 3.6 Example of 1992 Census of Transportation Data: 1992 Summary Statistics for the United States and States

SIC code	Geographic area and kind of business	Establishment (number)	Revenue ($1,000)	Annual payroll ($1,000)	First-quarter payroll ($1,000)	Paid employees (number)
	United States					
	Total Transportation, communications, and utilities except U.S. Post Office	(NA)	869,251,440	178,424,510	(NA)	5,566,120
	Total transportation except U.S. Post Office	(NA)	327,623,049	92,211,414	(NA)	3,356,872
40	Railroad transportation	(NA)	28,348,895	8,752,862	(NA)	197,421
41	Passenger transportation	17,805	12,649,307	5,191,117	1,245,956	354,913
411	Local and suburban passenger transportation	8,275	5,968,003	2,623,812	612,903	153,278
4111	Local and suburban transit	1,135	1,363,966	837,711	198,901	37,653
⋮	⋮	⋮	⋮	⋮	⋮	⋮
42	Motor freight transportation and warehouse	110,908	143,794,366	39,895,651	9,196,480	1,580,095
421	Trucking and courier service, except air	101,169	135,436,985	37,760,025	8,691,103	1,484,655
⋮	⋮	⋮	⋮	⋮	⋮	⋮
4213	Trucking, except local	40,821	78,357,536	20,974,464	4,807,742	758,435
⋮	⋮	⋮	⋮	⋮	⋮	⋮
44	Water transportation	8,147	29,207,214	5,170,196	1,213,197	171,314
⋮	⋮	⋮	⋮	⋮	⋮	⋮
45	Air transportation	(NA)	82,670,356	24,530,166	(NA)	707,148
⋮	⋮	⋮	⋮	⋮	⋮	⋮
46	Pipelines, except natural gas	844	7,063,056	821,085	203,267	16,779
⋮	⋮	⋮	⋮	⋮	⋮	⋮
47	Transportation service	46,593	23,889,855	7,850,337	1,854,022	329,202
⋮	⋮	⋮	⋮	⋮	⋮	⋮
48	Communication	39,244	230,667,167	47,057,941	12,335,145	1294,236
⋮	⋮	⋮	⋮	⋮	⋮	⋮
49	Electric, gas, and sanitary service	20,049	310,961,224	39,155,155	9,720,355	915,012
⋮	⋮	⋮	⋮	⋮	⋮	⋮
	ALABAMA					
⋮	⋮	⋮	⋮	⋮	⋮	⋮
	OHIO					
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Source: Table 1, 1992 Census of Transportation, Communication, and Utilities, Bureau of Census, U.S. Department of Commerce, Washington, DC: 1995.

Private Data Sources: Private data suppliers request fees for their products. The private firms collect national data on the transportation system and related indicator variables primarily by conducting regional surveys and by compiling and repackaging of public data. Based on the information collected, the private firms also produce future projections.

Woods & Poole Economics, Inc.: Woods & Poole Economics, Inc. specializes in detailed long-range projections for regional and demographic data. The database contains more than 550 economic and demographic variables for every state, region, county, Metropolitan Statistical Area (MSA) and Designed Market Area (DMA) in the U.S. for every year from 1970 to 2020. Upon contact, Woods & Poole sent us an example of their database layout, which can be found in Appendix 3.1. The information contains both historical and projected data on different variables, such as population (grouped by different age cohorts and races) and total employment (categorized in different industries), and others, for Alameda County, CA. Additional information may be requested from Woods and Poole Economics, Inc. (Phone Number: 202-332-7111) for a fee.

TRANSEARCH: TRANSEARCH is available from Reebie Associates. According to the company's description, "*TRANSEARCH is a market-to-market, industry (commodity)- and mode-specific database of freight traffic activity throughout the country.*" The database is produced annually. Table 3.7 shows an illustration of the layout at the "Business Economic Area (BEA)" level. An illustration of the BEA map for the eastern part of U.S. is shown in Figure 3.2. The O-D information is shown by Origin BEA (followed by affiliated city/state) and Destination BEA (affiliated city/state). The commodity information is classified by STCC (up to 5-digits) codes and given in annual tons by mode. For example Table 3.7 presents commodities shipped from BEA region #18 (Philadelphia, PA) to BEA region #55 (Memphis, TN). The first line shows that there were 134 tons of Farm Products (STCC-01) shipped by only two modes (i.e., For Hire Truckload [101 tons] and Private Truckload [33 tons]) during the given year.

Information beyond this example may be requested from the company (Phone Number: 203-661-8661) for a fee.

Table 3.7 Example of TRANSEARCH database
(Source: TRANSEARCH, Reebie Associates, 1997)

[1]Origin: BEA 18 Philadelphia PA		Destination: BEA 55 Memphis TN							
Two-digit STCC Summary		Total	[2]Rail		Highway			Air	Water
			Carload	Intermodal	For Hire		Priv./Ex		
STCC	Commodity	Tons			TL	LTL			
01	Farm products	134	0	0	101	0.	33	0	0
11	Coal	[3]595	0	0	595	0	0	0	0
20	Food or kindred products	7413	1008	0	4664	1399	342	0	0
22	Textile mill products	1613	0	0	586	285	742	0	0
23	Apparel or related products	406	0	0	0	274	63	69	0
24	Lumber or wood products	4443	0	0	737	0	3706	0	0
25	Furniture or fixtures	127	0	0	0	127	0	0	0
26	Pulp, paper or allied products	21719	9334	0	7258	2201	2926	0	0
27	Printed matter	67	0	0	0	67	0	0	0
[4]28	Chemicals or allied products	36568	18630	0	11109	507	6322	0	0
29	Petroleum or coal products	10258	9560	0	533	165	0	0	0
30	Rubber or misc. Plastics	6249	0	0	2346	2890	1013	0	0
32	Clay, concrete, glass or stone	7285	722	0	2673	407	3483	0	0
33	Primary metal products	13308	4868	0	7594	586	260	0	0
34	Fabricated metal products	2145	159	0	610	1150	226	0	0
35	Machinery	5718	0	0	2627	2513	407	171	0
36	Electrical Equipment	2095	0	0	468	737	890	0	0
37	Transportation Equipment	9468	5747	0	3172	103	443	3	0
38	Instrum, photo eq., optical eq.	645	0	0	0	249	396	0	0
39	Misc. manufacturing products	539	0	0	454	29	56	0	0
40	Waste or scrap materials	722	722	0	0	0	0	0	0
	[5]Total	136724	50750	920	45646	17847	21318	243	0
	Percent of Total	100.0	37.1	0.7	33.4	13.1	15.6	0.2	0.0
Additional STCC Detail		Total	Rail		Highway			Air	Water
			Carload	Intermodal	For Hire		Priv./Ex		
STCC	Commodity	Tons			TL	LTL			
01195	Potatoes, other than sweet	81	0	0	61	0	20	0	0
01399	Misc. fresh vegetables	53	0	0	40	0	13	0	0
11112	Prepared anthracite	595	0	0	595	0	0	0	0
20000	Food or kindred products	1229	0	0	1190	0	39	0	0
20100	Meat or poultry, fresh or chill	673	0	0	32	641	0	0	0
20300	Canned or preserved food	515	0	0	251	264	0	0	0
20330	Canned fruits, vegetables, etc.	1237	0	0	1087	147	3	0	0
20334	Canned fruit juices	824	0	0	671	151	2	0	0
20359	Sauces or seasonings	22	0	0	21	0	1	0	0
20381	Frozen prepared food or soup	551	0	0	362	0	189	0	0
20520	Biscuits, crackers or pretzels	107	0	0	93	0	14	0	0
CONT									

[1] Specific traffic lane
[2] Seven modes of transport
[3] Annual tons
[4] Listing of commodities
[5] Totals of traffic lane

Figure 3.2 Illustration of Business Economic Areas (BEA) map of the eastern part of the U.S. (Source: Bureau of Economic Analysis, 1996)

U.S. Air Freight Origin Traffic Statistic: The *U.S. Air Freight Origin Traffic Statistic* database is prepared by The Colography Group, Inc.. According to a review from the *Quick Response Freight Manual* (Cambridge Systematics Inc., 1997), this database contains express and freight air traffic shipments, based on a survey of firms that generate air traffic shipments. The data are reported on a geographic basis at the 4-digit SIC level by industry type. The originating data are also categorized as either domestic or exports, depending on their destinations. This database is compiled based on the survey results from shippers, which included both express mail and small shipments (under 3,000 lbs.). Detailed information may obtain from the company (Phone Number: 770-565-0464).

3.2 Review of Models

Models in the Literature: Winston (1983) provided an extensive review of freight demand models. He classified freight demand models into aggregate and disaggregate models. In aggregate models, the basic unit of observation is the share of a specific freight mode at a relatively coarse geographic level, such as a region. In disaggregate models, the unit of observation is an individual shipment. Winston claimed that the aggregate models would be useful in analyzing freight flows for policy analysis or for practical prediction in the context of large-scale regional or national studies. He noted that while disaggregate models are generally more attractive from a theoretical perspective, their extensive data requirements represent a significant disadvantage. Winston also discussed the extent and nature of intermodal competition, the importance of service quality, and the effect of the regulatory environment in the context of freight demand modeling.

Fang, *et al.*, (1996) reviewed approximately 20 models dealing with choice of freight modes, distinguishing between aggregate econometric, disaggregate econometric, and network-based models. The authors further broke down the aggregate econometric models into regression and aggregate logit models, and the disaggregate econometric models into abstract mode, linear programming, microeconomic, and discrete choice models. They also specified the required inputs for the various models. They proposed that although aggregate models are usually used in specific studies, they are inferior to the disaggregate models on theoretical grounds.

Cambridge Systematics, Inc., *et al.* (1996) reviewed approximately 40 freight forecasting models. The authors classified the models into those employing a structural approach and those employing a direct approach. The structural approach forecasts freight demand by applying techniques similar to those used in the traditional Four-Step Urban Transportation Modeling System (trip generation, trip distribution, mode choice, and route assignment). The direct approach uses simpler techniques to estimate correlation between freight demand and other variables that are expected to influence freight demand. They did not draw any conclusions about the relative merits of the two approaches in their review.

Before discovering these model reviews, we began reviewing freight models found in the literature according to the structure described in Appendix 3.2. We felt that this structure that would be useful for determining the appropriateness of proposed models for operational use in public planning issues. Given the many models found in the literature, we determined that typing the models in this way was beyond the scope of this project. We, therefore, terminated this effort but illustrate it in Appendix 3.2 with freight modeling components discussed in the second part of this report.

DOT Models: We became aware of ongoing freight modeling efforts at various state DOT's and contacted three states -- Indiana, Michigan, and Oregon -- to investigate efforts there. We conducted phone conversations with representatives of Indiana (Smith, 1998) and Oregon (Upton, 1998) and left a phone message at Michigan, which resulted in an email response (Nellet, 1998). We also found further descriptions of the freight modeling efforts through web sites (Parsons Brinckerhoff Quade & Douglas, Inc., *et al.*, 1996 a and b) and mailed reports (Cambridge Systematics, Inc., 1998; Black, 1997).

The models being developed are sophisticated enough that we do not review or critique them in detail. Rather, we highlight the features that should be helpful in clarifying steps ODOT should take in determining its immediate modeling efforts.

The freight forecasting models being developed by Indiana, Michigan, and Oregon are components of statewide models being developed that also include passenger forecasting components. All the models are being developed by consultants, and although it appears that DOT employees are committed to development and implementation, the efforts presently appear to be geared toward having individuals outside the organizations work with DOT personnel to develop tools for use by the organization. In at least one of the states, the possibility has been raised that the models might be run on an occasional basis by individuals outside of the organization. At the time of this writing, the models are all fairly advanced in terms of initial development, but they have not yet been used in supporting policy or alternative analysis.

In spite of the many methods and variations presented in the literature and the difficulties mentioned in Section 1, all the models seem to be based on the traditional 4-step passenger transportation forecasting process. The models differ from each other in details, and there are also slight differences from what one might consider to be a direct transfer of the 4-step process to statewide freight modeling. We also noticed discussion of combining some of the four steps, just as there is discussion of doing so in passenger modeling. Still, the basic framework of generating freight productions and attractions by zones, developing origin-destination matrices, splitting origin-destination flows by mode, and assigning flows to the network seems to be the backbone of the freight models under development. And, although there is some discussion of feeding back results from later stages into earlier stages, the implication seems to be one of retaining a highly modular structure.

In all the models, the initial emphasis seems to be on forecasting truck traffic on the highway network, but the intention is to allow future development into truly multimodal and intermodal models. The basic approach seems to be one of generating commodity flows and distributing the flows generated to form commodity flow origin-destination (OD) tables. The OD tables are converted to modal OD commodity flow tables, and truck commodity flow units are converted to truck trips before assigning them to the highway network.

We observed a common optimism in the potential of freight modeling. This optimism is heavily based on anticipated improvements in quality and quantity of freight data. The consultants are relying on several of the data sources mentioned in Section 3.1 to forecast independent variables and to calibrate model components, with the 1993 Commodity Flow Survey receiving most mention. Still, data from surveys or other special studies are required. Moreover, delays in making elements of the Commodity Flow Survey available were mentioned several times as slowing progress in model development. Like the advances in data availability, all the efforts seem to rely heavily on newer generations of transportation GIS and modeling software. Specific packages are mentioned.

Again, our objective is not to critique or evaluate the models based on our admittedly brief review of material available to us. We expect more methodological documentation to be available in the future, but it is apparent that the consultants have not "written the final chapter" on statewide freight modeling. Nevertheless, it seems that they are producing state-of-the-art operational models, striking a reasonable compromise between applicability and a rigorous behavioral basis. As with all large-scale system models involving elements of human decision making, the consultants have had to blend theoretical principles, judgments, and, at times, almost arbitrary adjustments to make model outputs match flow observations, where the observations are themselves imperfect estimates obtained from limited data. Depending on one's background, an individual might argue that more emphasis should be given to one component or another of this blend, but we feel that the consultants, working with the representatives of the DOT's, are presently in the best position to make these decisions.

Discussion is given to validating the models by comparing predictions to observations. However, as mentioned in Section 1.2, one cannot expect model outputs to match observations completely, and, therefore, one must resort to the fundamental concept of determining whether the benefits of making simplifying assumptions outweigh the costs of doing so. As mentioned above, a lack of experience with applications makes it too early to evaluate the models on this criterion, but the DOT representatives contacted are presently optimistic about the potential of the models to assist the DOT's.

Many of the difficulties (with methodology, data, ...) were not discovered until development was underway. This is not surprising, but it is important to recognize that useful freight models will not be developed and implemented exactly as proposed. That is, the states must be committed to development over the long run, and a one-time project devoted to developing a freight model that would then be turned over to the state would probably not be sufficient. We return to this point in the following section. Similarly, the developers note that the present efforts must be thought of as first phases and that the models must therefore be flexible and designed in such a way that modules can be added and components changed to keep up with results obtained from research or other freight studies.

Section 4. PART I Conclusion

A review of freight models and databases shows that much work has been and continues to be devoted to developing freight forecasting models and to making freight data available for use in such models. Despite this work, our survey indicates that, other than monitoring truck traffic on highway networks, there is presently little systematic freight modeling conducted at state DOT's. Similarly, the results of our interviews with potential users of Ohio freight models exhibit no present consensus on the desired outputs of a freight model to be used for public planning issues.

Still, our survey of freight modeling activity at state DOT's shows more activity than mentioned in the Cambridge Systematics, *et al.*, (1997) survey. There may have been some misunderstanding of our questions, but responses implied that any misunderstanding would have underestimated forecasting activity. Moreover, our personal communications with DOT personnel developing freight models in Indiana, Michigan, and Oregon, and the documentation on these developments, lead us to believe that freight modeling activity to support public planning issues is increasing and will be more widespread in the future.

Similarly, even though there is no real consensus on desired freight model outputs in Ohio, or apparently in other states, many of the outputs and motivating questions of Table 2.2 were related in some way to freight forecasting on the highway network. This was also the case in the results of another independent survey (see Table 2.1). Additionally, the intent at the Indiana, Michigan, and Oregon DOT's appears to be one of developing intermodal and multimodal models, but the initial emphasis is on forecasting truck trips on the highway network. These models will probably produce origin-destination (OD) tables in units of commodity flows, but in the near future the OD tables will be converted to truck OD tables, which are then assigned to the network. Black (1997, p. 101) offers a good explanation of why assigning rail trips to a network is inherently more difficult than assigning truck trips. Such difficulties, along with the apparent broader and deeper interests in highway network issues seen in Section 2, probably explain much of the reason that these states are beginning with an emphasis on the highway network. Therefore, we expect to see an increase in truck trip forecasting to support the analysis of public planning issues in the near future.

The similarities we observed in the developing Indiana, Michigan, and Oregon models also lead us to believe that future activities in freight modeling will be similar to each other. If the Ohio Department of Transportation (ODOT) is going to begin freight modeling activities, it would make sense to *pursue methods that will be common to several states*. This does not mean that the methods could not eventually be modified to suit the needs and characteristics specific to Ohio. However, the benefits of being able to share experiences and expertise with colleagues faced with similar issues make commonality desirable, at least commonality in the general modeling framework, architecture, and primary data sources exploited. In addition, pursuing models common to other states would be helpful in estimating external flows (i.e., flows from, to, or through states other than Ohio) that would eventually travel in Ohio, an issue that seems to require a great deal of effort in the models presently under development.

It would also seem that the basic elements of the traditional 4-step passenger forecasting system -- generating flows originating or destined for a region or zone; distributing originating flows among destinations and destined flows among origins to form origin-destination matrices; splitting origin-destination flows among modes; and assigning the flows to mathematical representations of the physical transportation networks -- should still be relevant in any modeling system eventually implemented. Indeed, despite the many different freight models proposed in the literature and the difficulties mentioned in Section 1.1, the developing Indiana, Michigan, and Oregon models strongly resemble the 4-step process, with a highly modular sequencing of the steps, and where estimated dependent variables are fit against estimated independent variables to produce some type of best-fit relations at each step that are assumed to remain constant in the future.

There are several advantages to this approach. There is a long experience with using it for forecasting passenger flows on highway networks. Therefore, although relatively complex, the framework is generally familiar to modelers, administrators, and users. Not only is it familiar to these groups, but it has been producing results that are used in a routine manner in the passenger forecasting context. The intermediate outputs produced (productions, attractions, origin-destination flows, and origin-destination flows by mode), as well as the final assigned link volumes, have proven useful when estimating measures of interest to policy studies. Therefore, even though there is no consensus on the desired outputs of a freight model, one is tempted to think that something useful can be obtained from the outputs of one or more of the stages of a 4-step freight forecasting model. Moreover, progress in the states investigated has already demonstrated that existing software can produce outputs from these methods using available data supplemented with limited special studies. Finally, the 4-step model is attractive because experience is being gained from developing similar models in different states, and this experience should be transferable to development efforts of similar models in the future.

Despite its advantages, trying to implement this type of model at ODOT at the present time is risky. As mentioned previously, developing such large-scale, systems models requires a judicious blend of theoretical rigor and practical compromises. The typical approach also involves extensive model calibration and validation. ODOT does not presently have sufficient resources to commit its own personnel to such an effort. The blend of required knowledge and the objective of producing an operational tool would make such work suitable and attractive to only a few consulting firms, where the expertise is geographically distant from ODOT and is expected to be increasingly in demand from other states. It is doubtful that any new, large-scale model such as this will be adequate the first time around, that is, at the end of the initial contract.

The danger, then, is that a consulting arrangement will produce a model that initially is not frequently used and which is turned over to DOT employees who feel no sense of ownership in it. There are several individuals at ODOT who understand flow modeling issues sufficiently to realize that developing a workable and useful large-scale freight model is a long-term project. However, these individuals have so many daily and longer term commitments in the present environment that, unless they feel strong ownership in the product, it will be much easier for them to dismiss an initial effort as a failure than to try to convince administrators that additional contracts are needed to produce a useful product. The documentation associated with the Indiana, Michigan, and Oregon efforts demonstrates that these DOT's and their consultants realize this potential danger. However, no specific solution is offered to reduce the likelihood of its occurring.. In short, we feel that developing a large-scale freight model at this point is doomed to failure in the present environment unless ODOT is willing to *commit to long-term development*

We *propose* that *working with several states in a formal relationship* is one way to motivate ODOT (or any other state) to remain committed to the model development throughout and after initial efforts. Such a relationship could take the form of a working group with regularly scheduled meetings. Or, we feel it would be more beneficial to join other states in collaboratively funding model development by consultants. We envision pooling funds to develop either identical models in the participating states or, if the participating states were contiguous, a regional model that could be scaled down to sufficient resolution in each of the states. There may be certain unique details in the various states that would need to be addressed, but there would be common framework, architecture, hardware, and data requirements.

Working with colleagues in other states to supervise and direct the development of a common model would motivate DOT technical representatives to remain philosophically committed to the development process over the long run and allow one state to call on others to convince administrators to remain financially committed. In addition, such a process should facilitate discussion of issues and sharing of expertise, produce economies in the development process, and, if a regional model is produced, lead to easier and more accurate estimation of external flows. Finally, working with other states would focus interest in such a way that the group could exert stronger influence on federal studies. That is, a coalition of states would be in a better position than an individual state to define a research agenda or request data collection efforts.

Any large-scale model like one based on the 4-step process also poses the danger of turning into a black box that produces results which are consistent across the steps, but does so at the price of reducing the transparency of the relationships among its variables, inhibiting the ease of drawing insights from the numbers produced, and limiting the ability to modify the forecasts produced with expert opinion or common sense. Calibrating relations on historical data may produce good fits and even good results in validation studies, but the relations must hold into the future if the model is to produce good forecasts. In these models, there are many intertwined relations with no real behavioral justification. Moreover, even where there is a behavioral basis, there is no reason to believe that human behavior remains constant over time. Therefore, the flows forecast from the model cannot be expected to be realized in the future.

If model forecasts are know to be erroneous, it would make sense that they be modified by expert judgment when intended for use in supporting policy or alternative analysis. The many and dependent relationships in a large-scale model makes this modification more difficult. This is especially true in the anticipated freight models, where little experience has been gained. The models being developed in Indiana, Michigan, and Oregon appear to be state-of-the-art. However, we believe that a panel of true experts could produce more accurate forecasts than the first generations of these models. Such a panel would also have much more flexibility than a large-scale model in forecasting different freight measures. The hope is that the model forecasts could eventually assist the experts develop better forecasts than they could otherwise produce without the use of the models. We doubt that this hope will be realized without more experience gained from applications and from systematic model testing, tracking, and experimentation.

To expedite the development of this experience, we *propose* that *research and experimentation with alternative model components* be conducted *and* that the forecast *accuracy* of the components *be tracked* on an ongoing basis. Modular systems like those being developed have the advantage that their components can be readily modified without having to reinvent the entire framework. However, if developed under contract for immediate implementation without any accompanying experimentation with its components, there is a strong possibility that modifications will not be made and the entire model will be dismissed as irrelevant.

Black (1997) proposes that research is needed on the freight generation step, and the consultants developing the Oregon model (Parsons Brinckerhoff Quade & Douglas, Inc., *et al.,* 1996b) emphasize the need recognized in other publications to calibrate models over longer time periods than are traditionally considered. We agree. However, we also feel that every component of the freight forecasting system would benefit from systematic research and that it would be better to assess the fits of alternative specifications and functional forms of model components than to restrict the efforts to calibrating a component specified at the outset.

We again mention that there may no be no more accurate forecast than that produced by a panel of experts, and the real value of a forecast produced from a mathematical model might best be measured in its ability to modify these expert forecasts. Even if a model or model component is shown to produce good fits to existing data, it will be of little predictive use if it does not change "prior" expert forecasts, those that would have otherwise been produced without its availability. (These prior forecasts might not change either because of a lack of confidence that the good-fitting empirical correlations used in the model would hold in the future or because the model would produce forecasts that could have been obtained with simpler reasoning.) Therefore, rather than limit the research to assessing correlations between dependent and independent variables, alternative forms of the components of the models should also be considered in terms of an ability to modify expert forecasts of measures useful for policy analysis. This is an open field, since there is presently no set of agreed upon useful measures, but this openness does not mean that this aspect should be ignored. As a minimum, it would appear that simplicity in the components should be valued.

We also propose that forecast accuracy be systematically tracked. Validation studies, where calibrated outputs are compared to existing flows and then assumed to perform well enough if some deviation measure is low, are useful but not sufficient to assess predictive performance. Like the call to calibrate models over longer time periods, predictive ability should also be systematically assessed over longer time periods. This type of analysis should not only be considered as leading to assessments of whether a model is performing "well enough," but also of producing a feel for the level of uncertainty associated with the forecasts. This feeling is underestimated by traditional simulation experiments, where output variability is produced by simply varying inputs and/or model parameters. This type of work is sorely lacking even in the passenger forecasting field, and it is complicated in the freight modeling context because of data limitations. Still, it must be considered if one is to develop any appreciation on proper use of the outputs before users dismiss the entire model as being too inaccurate.

The proposed agenda for research and experimentation with model components is ambitious and likely beyond what ODOT (and most DOT's) sees as its present mission. However, we feel that it is important if an operational statewide freight model is to be accepted for use. Again, pooling resources with other states could help implement this work: Some of the economies gained by pooling resources could be put toward funding these efforts; different tasks could be funded by different states with group oversight of the entire program; and the group could collectively provide input to the federal research and development program.

PART II: Illustrative Models and Components

In this part, we consider a few simple, specific models or components that could either stand alone or be incorporated in more complex models. In Section 5, we consider simple indicator models that could be used to forecast future commodity generation in regions. These forecasts could be useful in themselves or modified for use in a trip generation module of a larger model. In Section 6, we consider a model proposed to forecast freight using intermodal freight facilities. The model formulation is similar to that used for discrete choice analysis, which is frequently the basis of trip distribution and mode choice models. The models considered in Sections 5 and 6 could be presently used in Ohio. However, we illustrate that different specifications are possible and that they can produce very different forecasts. Also, as mentioned in the previous section, we believe that simply finding which specification fits past data best is not sufficient for determining which will be most valuable for forecasting in practice.

In Section 7, we discuss different methods of updating truck origin-destination (OD) tables from observed truck volumes. We feel that the methods considered would lead to similar results. Simply determining an origin-destination table that is consistent with observed traffic flows does not directly lead to forecasts of future conditions. However, since truck volumes are routinely collected, using any of these procedures would be inexpensive to implement on a systematic basis, and a good estimate of a present OD matrix could be used to calibrate or validate model components. Moreover, accurate OD estimates could assist experts when forecasting future OD patterns.

In Section 8, we show that recently available databases could be used with existing software to perform intermodal assignment. Although the quality of the assignments is presently limited by a lack of acceptable intermodal assignment logic, these databases could be used to develop, test, and experiment with intermodal assignment algorithms in the future.

Section 5. Freight Indicators

Freight indicators are variables describing socio-economic activity expected to be correlated with freight movement activity. The assumption motivating the use of indicator methods is that it is easier to forecast the indicators and infer forecasts of freight movement variables than to forecast the freight movement variables directly. Therefore, it is necessary that data, and hopefully forecasts, are available for the freight indicator variables. As claimed in Cambridge Systematics, Inc. *et al.*, (1996), such methods can be used to establish rough, but quick and inexpensive forecasts of statewide, regional, or even local and facility-specific freight activity. They are also closely related to traditional trip generation methods. The major advantages of forecasting freight movements through forecasts of indicators are ease of use and low expense.

In this section, we show that indicators can be used with available databases to produce forecasts of commodity flows generated in Ohio at an aggregate level. As such, indicators methods could be readily implemented for use in Ohio. However, we also demonstrate that there are several alternatives -- approaches, functional forms, and independent variables -- that can be used. We feel that the regression methods described offer the most modeling flexibility with relatively little added complexity. We also determine goodness of fit statistics that can be used to judge the degree of correlation between the freight movement and indicator variables. However, as mentioned in the previous section, longer term analysis would be required to determine which indicators predict freight movements well over time, gain some sense of the variability inherent in such forecasts, and assess which indicators can be best used to help support policy or alternative analysis by themselves or in modules of more complex models.

5.1 Methods

Dependent Variable Growth Factor: Perhaps the simplest model for forecasting future freight movements is one which applies a "growth factor" to the variable of interest *Y*, indicating freight flows (Cambridge Systematics, Inc. *et al.*, 1996). Actually, the Dependent Variable Growth Factor method is not truly an indicator method, since there is no explicit relation between the freight movement variable Y and some economic indicator X. Still, it is related to the Independent Variable Growth Factor method presented below and is often mentioned in the context of simple models.

If Y_t is the level of the freight movement variable value in year t and Y_{to} is the level in year t_o, then the Dependent Variable Growth Factor method determines Y_t from Y_{to} as:

$$Y_t = Y_{to}(1+r)^{(t-to)}, \tag{5.1}$$

where r is an annual growth rate. This growth rate r could either be assumed or calculated from observations of the freight movement variable in two years, e.g., Y_{t1} in year t_1 and Y_{t2} in year t_2. Substituting Y_{t1}, Y_{t2}, t_1, and t_2 , $t_2 > t_1$, in (5.1), and rearranging:

$$r = -1 + (Y_{t2}/Y_{t1})^{(1/(t2-t1))} \tag{5.2}$$

As an example, the Bureau of Census (Bureau of Transportation Statistics, 1997C) showed that there were 140,040 thousand tons of Nonmetallic Minerals shipped in Ohio in 1989 and 128,639 thousand tons shipped in 1993. Using Equation (5.2):

$$r = -1 + (128{,}639/140{,}040)^{(1/(1993-1989))} = -0.021.$$

To forecast the Ohio shipments of Nonmetallic Minerals in the year 2000, for example, this value is substituted in Equation (5.1):

$$Y_{2000} = 128{,}639(1 - 0.021)^{(2000-1993)} = 111{,}000\ [10^3 \text{ tons}].$$

Another example of this method can be found in *Quick Response Freight Manual* (Cambridge Systematics Inc. *et al.*, 1996).

If observations of the freight movement variable are available for more than two years in a period considered representative of that being forecast, an econometric procedure could be used to fit a growth factor.

This growth factor model assumes that traffic "grows" by itself or compounds like an interest-bearing investment. This would seem difficult to accept, since one of the principles of transportation analysis is that transportation is a derived good, dependent on other factors. By ignoring any relation with an independent variable X, this model ignores this basic principle. Also, by not modeling any relation to an independent, "causal" variable, there is no room to forecast freight flows under different policy alternatives. The model also assumes that the growth calibrated in some time interval will hold in other time intervals. The only way that this could be argued as acceptable is if one could claim that "all relevant changes in factors in freight movement during the period of analysis will be the same" as during the calibration period.

In spite of its limiting assumptions, this growth factor method could form the basis of critical reasoning. Specifically, freight movements could be forecast for some future time using an assumed rate or one calibrated from past data. Then, one could use expert opinion to decide whether more, less, or about the same amount of growth is expected than that calculated. In this way, the calculated value would serve as a lower or upper bound, or a best guess estimate.

Independent Variable Growth Factor (Unit Elasticity): The previous growth factor method can be amended to allow a relation between some economic indicator variable X and the freight movement variable of interest Y (Cambridge Systematics, Inc. *et al.*, 1996). Again let Y_t be the level of the freight movement variable value in year t and Y_{t_o} be the level in year t_o, and let X_t and X_{t_o}, respectively, be the levels of the economic indicators in years t and t_o. This model assumes that the growth in Y from year t_o to year t will mirror that of X in the same period. Specifically, X is assumed to grow at a constant yearly rate between t_o and t:

$$X_t = X_{to}(1+r')^{(t-to)}, \tag{5.3}$$

where *r'* is the annual growth rate in the economic indicator *X*. Rearranging Equation (5.3), this rate is found as:

$$r' = -1 + (X_t/X_{to})^{(1/(t-to))}; \tag{5.4}$$

and used to forecast Y_t from Y_{to}:

$$Y_t = Y_{to}(1+r')^{(t-to)}. \tag{5.5}$$

As an example, consider a case where Ohio Employment in the Non-Metallic Minerals Mining sector is used as an indicator for Total Tons of Non-Metallic Minerals shipped in Ohio. From the Bureau of Economic Analysis (1996), Employment in Non-Metallic Minerals Mining was 4.8 [10^3] in 1993 and forecast to be 5.0 [10^3] in 2000. From the Bureau of Census (Bureau of Transportation Statistics, 1997C), there were 128,639 [10^3 tons] of Non-Metallic Minerals shipped in 1993. Using Equation (5.4):

$$r' = -1 + (5.0/4.8)^{1/(2000-1993)} = 0.0058,$$

and substituting this in Equation (5.5):

$$Y_{2000} = 128{,}639(1+0.0058)^{(2000-1993)} = 134{,}000\ [10^3 \text{ tons}].$$

Again, the growth rate in *X* could be assumed or decided upon from expert opinion, rather than using observations of *X* in different time periods. One would then use Equation (5.5) directly, and by comparing to Equation (5.1), the process is seen to be identical to that in which the growth factor is applied directly to the freight movement variable *Y*. Also, as before, if there are observations on the economic indicator in more than two years, some econometric technique could be used to fit a value of *r'* to the data.

This method can be simplified by noticing that substitution of Equation (5.4) in Equation (5.5) yields:

$$Y_t = Y_o (X_t/X_o). \tag{5.6}$$

Equation (5.6) can be used to find Y_t directly if Y_o, X_t, and X_o are given. So, in the above example, one could simply calculate:

$$Y_{2000} = 128{,}639(5.0/4.8) = 134{,}000\ [10^3 \text{ tons}].$$

The relation in Equation (5.6) can be rewritten as:

$$Y_t = (Y_o/X_o)\, X_t, \tag{5.7}$$

and as:

$$Y_t/Y_o = X_t/X_o. \tag{5.8}$$

The form in Equation (5.7) becomes important when discussing linear regression models below. The form in Equation (5.8) becomes significant when determining the elasticity implied by this growth factor method. The elasticity is the ratio of relative changes in the freight movement variable to the relative changes in indicator variables:

$$\text{elasticity} = [(Y_t-Y_{to})/Y_{to}] \,/\, [(X_t-X_{to})/X_{to}]. \tag{5.9}$$

Algebra on Equation (5.9) yields $elasticity = (Y_t/Y_{to} - 1) / (X_t/X_{to} - 1)$, and using Equation (5.8), yields:

$$\text{elasticity} = 1.0.$$

That is, the Independent Variable Growth Factor method implies a unit elasticity, which means that changes in *Y* correspond exactly to the changes in *X*. There is no room for economies of scale. Also, as in the preceding method, which applied a growth factor directly to the freight variable, the assumption in this model is that the growth in *X* during the period of analysis will be the same as that during the period when the growth factor is calibrated.

In spite of these strict assumptions, this growth factor model could be used to reason as before, i.e., by calculating a number with the method and using expert opinion to decide whether this is a high, low, or middle estimate. In this case, however, the reasoning would be on the anticipated growth in economic indicator *X*, and the levels of the freight flow *Y* would be deduced. Before doing this, however, one would want to ensure that there really is a strong unit elastic correlation between *Y* and *X*. This could be done by looking at several years of paired *X* and *Y* data to see either if Equation (5.6) holds or if the unique value of *r'* resulting from the econometric method fits the data well. Alternatively, two years of paired *X-Y* data could be used from several geographical areas. We discuss a related approach in the regression models below.

Constant Elasticity Method: The Independent Variable Growth Factor method allowed *Y* to be correlated with some economic indicator. The correlation was seen to be such that the elasticity--the ratio of relative changes in the freight variable to relative changes in the indicator variable--is unity. This unit elasticity could be relaxed slightly to say that the ratio of relative changes is constant but equal to some value *b*, which is not necessarily unity. That is:

$$\text{elasticity} = [(Y_t - Y_{to})/Y_{to}] \,/\, [(X_t - X_{to})/X_{to}] = b. \qquad (5.10)$$

Solving for Y_t:

$$Y_t = Y_{to}[1 + b(X_t - X_{to})/X_{to}]. \qquad (5.11).$$

So, for example, the Bureau of Economic Analysis (1996) data show that employment in the Ohio Non-Metallic Minerals sector was 5.1 [10^3] and 4.8 [10^3] in 1989 and 1993, respectively, and the Bureau of Census (Bureau of Transportation Statistics, 1997C) data show that there were 140,040 [10^3 tons] and 128,639 [10^3 tons] shipped in Ohio in these years. Using Equation (5.10), the elasticity *b* is estimated as:

$$b = [(128{,}639 - 140{,}040)/140{,}040] \,/\, [(4.8 - 5.1)/5.1] = 1.38.$$

To estimate Non-Metallic Minerals shipments in 2000, this elasticity, an employment forecast of 5.0 [10^3] in the Ohio Non-Metallic Minerals sector in 2000 (see Bureau of Economic Analysis, 1996), and some base year data would be used in Equation (5.11). Using the 1993 year as the base:

$$Y_{2000} = 128{,}639[1+1.38(5.0-4.8)/4.8] = 136{,}000\ [10^3\ \text{tons}].$$

Again, an elasticity b could alternatively be determined from expert opinion. Similarly, if there are paired observations on X and Y in more than two years, an econometric technique could again be used to fit an elasticity.

The formulation of elasticity in Equation (5.10) deals with finite, discrete changes in variables, $(Yt\text{-}Yto)$ and $(Xt\text{-}Xto)$. Infinitesimal changes dY and dX could be used instead to write:

$$\text{elasticity} = b = (dY/Y) / (dX/X) = (dY/dX)\ (X/Y). \qquad (5.12)$$

In this case, a model of the form:

$$Y = aX^b, \qquad (5.13)$$

would be consistent with Equation (5.12), since $(dY/dX)\ (X/Y) = (abX^{b-1}) \times (X/aX^b) = b$. Paired indicator X and freight movement Y values in two years would be sufficient to calculate values of the two parameters a and b of Equation (5.13). If paired observations are available for more than two years, or if 2-year data pairs are available from several geographic locations, an econometric technique could once again we used to fit these parameters. We discuss this in more detail when presenting the regression method below.

For example, consider the Bureau of Economic Analysis and Bureau of Census data cited above for Non-Metallic Minerals employment and shipments in 1989 and 1993. Substituting these data in Equation (5.13) to form two equations in a and b, and solving yields $a = 14{,}300$, $b = 1.4$. Substituting these values in Equation (5.13) with $X_{2000} = 5.0$ would yield:

$$Y_{2000} = 14{,}300(5.0^{1.4}) = 136{,}000\ [10^3 \text{ tons}].$$

Although this value is identical to that found above, there could be some difference because of rounding error and, especially, because of the difference of estimating "arc" elasticities based on discrete differences and "point" elasticities based on infinitesimally small changes.

As before, this method assumes that the elasticities calculated in one period hold during another period. And, as before, the model could be used as a basis for reasoning such as, "If the ratios of changes were the same, we would expect 136,000 [tons] (for example). We expect that certain increases in efficiency will make the changes of Y, relative to changes of X, increase more than previously. Therefore, we feel that there would be more than 136,00 [10^3 tons] produced in 2000."

Regression Methods: Although there is some merit to the above approaches, we feel that regression approaches to indicator variable modeling are more general. They, therefore, incorporate the advantages of the other approaches but allow more flexibility--essentially the use of more data to calibrate parameters. Perhaps, the simplest regression models to consider would be *linear* specifications:

$$Y_t = aX_t + k, \tag{5.14}$$

and *exponential* specifications:

$$Y_t = aX_t^b. \tag{5.15}$$

If the "intercept" k in the linear specification is set equal to zero, a comparison between Equations (5.7) and (5.14) shows that the linear specification corresponds to the Independent Variable Growth Factor method, with the parameter a equal to the ratio of freight movement variable Y to the indicator variable X in some base year. As mentioned above, this means that the elasticity of Y to X is equal to unity. Similarly, as discussed above, the exponential specification would correspond to a constant elasticity method with elasticity equal to the exponent *b*. Therefore, these specifications would imply assumptions similar to those discussed in the Unit and Constant Elasticity methods.

The parameters of the specifications would be fit to paired X-Y data in several years or in several regions using ordinary least squares techniques. The exponential form would first be transformed to a linear one by taking the logarithm of both sides to obtain:

$$\log Y_t = \log a + b \log X_t. \tag{5.16}$$

The advantage of using the regression techniques is that they allow more than the minimum number of data observations to be used to determine "best fit" parameters of the model. They also produce r^2 statistics that indicate how well the assumed specification fits the data, i.e., how well the underlying assumptions of the method are being satisfied. The regression methods produce other indicative statistics, such as F- and t-statistics, although their use in hypothesis testing should be considered with caution because of the additional assumptions required. Multivariate regression of the freight movement model Y against several indicator variables $X_1, X_2, \ldots, X_n$ could be considered. Indeed, this would provide a better estimation of the parameters when several indicators vary simultaneously in data sets. However, this would be more similar in spirit to trip generation models than to simple indicator models. Using a single indicator has the advantage of simplicity. Doing so can be more easily justified when considering different indicators for different commodities or types of movement, as is illustrated in the next section.

5.2 Numerical Illustrations

We provided simple numerical examples above when discussing the different methods. Here, we compare the methods further and illustrate how they may be used with available data bases. Specifically, we compare the Independent Variable Growth Factor (Unit Elasticity) method, the Constant Elasticity method, and Regression methods using both Linear and Exponential specifications. For simplicity, we use the estimate of the exponent in the Exponential Regression as the estimated elasticity in the Constant Elasticity Method. The rationale and difficulties with doing this were presented above.

Calibration: One would be interested in identifying indicators and forecasting freight movements for important commodities. We, therefore, first used the Commodity Flow Survey data (Bureau of Transportation Statistics, 1997C) to identify the most important commodities, in terms of tons shipped, originating in Ohio in 1993. In Table 5.1, we present the 13 most important commodities, along with the tonnage and percentages shipped by truck, rail, multimodal, and other or unknown mode.

Table 5.1 Shipments originating in Ohio by commodity and mode in 1993

	Total Tons (thousands)	Truck	Rail	Multimodal	Other/ Unknown
Non-Metallic Minerals	128,639	94.5%	2.4%	0.0%	2.0%
Petroleum/Coal Products	80,418	57.0%	3.6%	0.0%	26.1%
Primary Metal Products	40,819	86.6%	9.4%	0.6%	0.9%
Clay, Concrete, Glass, Stone	35,467	97.3%	0.0%	0.0%	0.9%
Food, Kindred Products	32,785	88.5%	9.8%	0.1%	1.5%
Farm Products	32,680	40.8%	44.7%	0.0%	13.0%
Coal	27,391	45.4%	n.a.	n.a.	43.0%
Chemicals, Allied Products	25,003	74.3%	8.8%	0.9%	0.2%
Waste, Scrap Materials	11,445	56.7%	41.0%	0.0%	0.0%
Lumber, Wood Products	11,021	38.6%	0.3%	0.2%	0.1%
Transportation Equipment	10,416	71.9%	9.3%	12.4%	6.0%
Fabricated Metal	8,574	89.3%	5.3%	0.1%	4.4%
Pulp, paper, Allied Products	8,570	95.4%	2.5%	0.2%	1.6%
Other	39,939	54.0%	5.9%	2.3%	37.8%
Total	493,167	78.5%	8.2%	0.6%	12.2%

Source: Table 5, CFS 1993, Report by State of Origin, Ohio (Bureau of Transportation Statistics, 1997C)

As indicators for the commodities in Table 5.1, we considered various data available from the Bureau of Economic Analysis (1996), as summarized in Table 5.2. Rather, than look at time series data to perform the regressions, we considered cross-section data. Specifically, we used the tons of the commodity shipped in each state in 1993 as the freight movement variable Y, paired with the value of the indicator variable

X in the same state in 1993. For the Linear Regression Specification, we assumed no intercept ($k = 0$) to allow comparison with the Unit Elasticity Method results.

Table 5.2 Potential indicators available from the Bureau of Economic Analysis

Indicator	Level
Gross State Product -GSP-by Industry	57 Industries (State), 14 Industry Groups (BEA Economic Areas, Metropolitan Statistical Areas)
Employment by Industry	57 Industries (State), 14 Industry Groups (BEA Economic Areas, Metropolitan Statistical Areas)
Earnings by Industry	57 Industries (State), 14 Industry Groups (BEA Economic Areas, Metropolitan Statistical Areas)
Personal Income	State, BEA Economic Areas, Metropolitan Statistical Areas
Population	State, BEA Economic Areas, Metropolitan Statistical Areas

Source: Bureau of Economic Analysis, U.S. Department of Commerce (1996)

As an example, we regressed the sum over all commodities shipped against the Total Industry Employment to obtain:

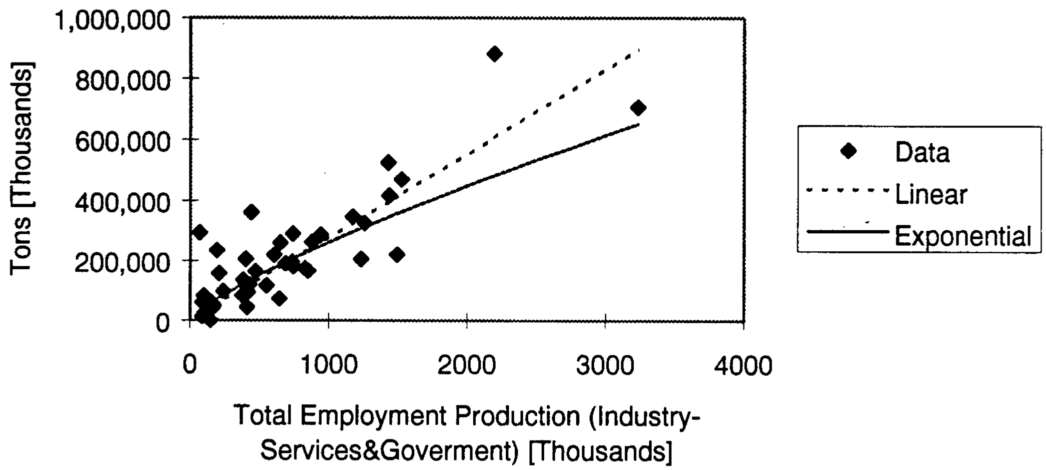
Linear Specification:	$Y = 57.45\ X$,	with $r^2 = 0.596$;
Exponential Specification:	$Y = 445.57X^{0.76}$,	with $r^2 = 0.583$;

and also the sum over all commodities shipped against the Total Production Employment (defined as the Total Employment minus that in the Services Industries and Government) to obtain:

Linear Specification:	$Y = 277.04\ X$,	with $r^2 = 0.703$;
Exponential Specification:	$Y = 1193.18\ X^{0.78}$,	with $r^2 = 0.607$.

Both indicators show that the tons shipped would increase with the employment variable, as expected. Therefore, one might consider the r^2 values to say that the Total Production Employment indicator performed better than the Total Industry Employment indicator. We represent the Linear and Exponential specification fits to the data in Figure 5.1.

Figure 5.1 Linear and exponential fits of Total Tons Shipped vs. Total Production Employment using 1993 cross-section data



We continued in this way with the 13 commodities listed in Table 5.1, performing regressions of the linear and exponential specifications using a few of the indicators listed in Table 5.2. In Table 5.3 we present the results of those producing the best results, in terms of expected signs of coefficient a and fit parameter r^2. (In the table, we present the estimate of *ln a,* the natural logarithm of the coefficient a for the exponential specification.) Note that, when looking across commodities, the estimated exponents b of the exponential specification were not very different from unity, indicating that a unit elasticity might be reasonably assumed in this data. Although we present the results for these "best" indicators in Table 5.3, this exercise was meant to be illustrative only. Further analysis is recommended before deciding on best indicators for practice.

Table 5.3 Regression results for selected indicators for commodities in Table 5.1

		Model 1- Linear y=ax		Model 2 - Exponential ln y = ln a + b ln x		
Commodity (y)	Indicator (x)	Estimated a (Standard Error a)	r^2	Estimated ln a (Standard Error ln a)	Estimated b (Standard Error b)	r^2
All Commodities	Production Employment[1]	57.451 (3.948)	0.596	7.084 (0.582)	0.7801 (0.0946)	0.607
Non-Metallic Minerals	Employment Non-Metallic Minerals Mining	16053.80 (1100.35)	0.646	9.351 (0.140)	1.1572 (0.137)	0.646
Petroleum and Coal Products	Employment Petroleum Products	11138.77 (444.03)	0.920	9.920 (0.0893)	0.6275 (0.0627)	0.730
Primary Metal Products	Employment Primary Metal Products	413.16 (12.241)	0.940	5.467 (0.1675)	1.1241 (0.0715)	0.849
Clay, Concrete, Glass and Stone Products	Employment Stone, Clay and Glass	1292.0 (66.63)	0.777	7.773 (0.1308)	0.78251 (0.05841)	0.803
Food and Kindred Products	Employment Food Manufacturing	493.999 (17.925)	0.879	6.137 (0.1382)	1.0165 (0.04256)	0.927
Farm Products	Employment Farming	199.894 (26.861)	0.246	3.715 (0.8410)	1.2982 (0.2046)	0.521
Coal	Employment Coal Mining	8229.99 (1797.5)	0.317	9.315 (0.2899)	0.7925 (0.1669)	0.634
Chemicals and Allied Products	Employment Chemical Products	424.25 (61.00)	0.320	6.924 (0.2343)	0.7149 (0.08237)	0.653
Waste, Scrap Materials	Production Employment[1]	3.9347 (0.2215)	0.823	-2.317 (0.7403)	1.5073 (0.1185)	0.818
Lumber and Wood Products	Employment Lumber and Wood Products	859.72 (72.122)	0.611	5.949 (0.2697)	1.1432 (0.10062)	0.755
Transportation Equipment	Employment Motor Veh.+ Other Transp. Equipment	48.392 (3.2616)	0.787	3.989 (0.3432)	0.88835 (0.10304)	0.668
Fabricated Metals	Employment Fabricated Metals	63.259 (2.9216)	0.859	4.061 (0.1714)	0.99288 (0.05739)	0.872
Pulp, Paper and Allied Products	Employment Paper Products	287.98 (17.672)	0.654	5.967 (0.1727)	0.9023 (0.06636)	0.826

1. Production Employment = Industry Employment - (Transportation + Wholesale and Retail + Financial, Insurance, Real Estate + Services + Government) Employment

Sources: CFS (BOC, 1996), SPEA (BEA, 1996), Own Calculations

Forecasting: We apply the four methods to future forecasts of the indicator variables for Ohio to illustrate how the indicators could be used in forecasting. We obtained these forecasts from the Bureau of Economic Analysis (1996) and show selected projections in Table 5.4.

Table 5.4 Forecasts of indicators

Indicators	1993	1998	2000	2005	2010	2015	2025	2045
Production Employment	1523.3	1558.9	1557.6	1568.3	1573.8	1569.6	1538.5	1617.3
Employment Non-Metallic Minerals Mining	4.8	4.9[1]	5.0	5.0	5.1	5.1	4.9	5.2
Employment Petroleum Products Manufacturing	7.8	7.7[1]	7.6	7.3	7.1	6.9	6.5	6.6
Employment Primary Metal Products Manufacturing	90.4	87.6[1]	86.5	83.5	80.9	78.5	74.2	75.5
Employment Stone, Clay and Glass Manufacturing	43.8	41.8[1]	41.0	39.2	37.6	36.3	34.0	34.1
Employment Food Manufacturing	63.3	63.3[1]	63.3	62.8	62.3	61.6	59.9	62.8
Employment Farming	99.5	97.4[1]	96.2	93.9	91.0	87.3	79.6	74.5
Employment Coal Mining	4.8	3.8[1]	3.4	2.9	2.5	2.3	1.9	1.4
Employment Chemical Products Manufacturing	67.8	68.0[1]	68.1	68.3	68.4	68.3	67.2	71.1
Employment Lumber and Wood Products Manufacturing	27.2	29.5[1]	30.4	32.0	33.2	33.9	34.3	36.7
Employment (Motor Veh.+Other Transp. Eq.) Manufacturing	130.5	131.6[1]	132.1	131.9	131.2	130.2	126.4	131.1
Employment Fabricated Metals Manufacturing	123.4	119.1[1]	117.4	113.5	110.0	106.9	101.1	102.4
Employment Paper Products Manufacturing	37.2	37.3[1]	37.4	37.4	37.4	37.2	36.5	38.4

1. Interpolated value not provided by BEA.
Source: SPEA 1993-2045 (BEA, 1996)

To demonstrate, consider applying the procedures to forecast Total Tons of Non-Metallic Minerals shipped in Ohio in 1998. From Table 5.3, the relevant indicator for this freight movement variable Y is Employment in the Non-Metallic Mineral Mining sector in 1998. We use 1993 as the base year t_o in the Independent Variable Growth Factor (Unit Elasticity) and Constant Elasticity methods, since freight movement data can be found in the Commodity Flow Survey for this year. Table 5.1 shows that $Y_{to} = 128{,}639$

[10^3 tons] of Nonmetallic Minerals shipped in Ohio in 1993. Table 5.4 shows that the Employment in the Non-Metallic Mineral Mining sector was 4.8 [10^3] in 1993 and forecast to be 4.9 [10^3] in 1998.

Using Equation (5.7), the 1998 forecast using the Independent Variable Growth Factor (Unit Elasticity) method would be:

$$Y_{1998} = (4.9/4.8)\ 128{,}639 = 131{,}000\ [10^3 \text{ tons}].$$

From Table 5.3, the estimated elasticity for this freight movement and indicator variable would be 1.1572, rounded to 1.16. Using Equation (5.11), the 1998 forecast using the Constant Elasticity Method would be:

$$Y_{1998} = 128{,}639[1 + 1.16(4.9\text{-}4.8)/4.8] = 132{,}000\ [10^3 \text{ tons}].$$

From Table 5.3, the estimated parameter a in the Linear Specification regression method would be 16,053.80, rounded to 16,100. Using Equation (5.14) with $k=0$, the 1998 forecast using this method would be:

$$Y_{1998} = 16{,}100\ (4.9) = 78{,}700\ [10^3 \text{ tons}].$$

From Table 5.3, the estimated parameters a and b in the Exponential Specification regression method would be $e^{9.35} = 11{,}500$ and 1.16, respectively. Using Equation (5.15), the 1998 forecast using this method would be:

$$Y_{1998} = 11{,}500\ (4.9)^{1.16} = 72{,}600\ [10^3 \text{ tons}].$$

The differences between the Elasticity Methods (Unit and Constant Elasticity) and Regression (Linear and Exponential) results are apparent. The differences stem from the fact that the ratio of the tons of Non-Metallic Minerals shipped to the Employment in the Non-Metallic Mineral Mining sector was higher in Ohio than the least squares ratio of the fifty states in 1993, 26,800 (=128,639/4.8) compared to 16,100. The Unit Elasticity (Independent Variable Growth Factor) and Constant Elasticity results are similar because the constant non-unit elasticity used in the latter method was close to unity. This also explains the closeness between the Linear and Exponential Specification regression results.

We forecast the 1998, 2000, 2005, 2010, 2015, and 2025 tons shipped in Ohio for the commodities listed in Table 5.1 with each of these four methods. We present the results in Table 5.5. We also total these forecast tonnage and list these in the table. From Table 5.1, one can deduce that these commodities represented approximately 92% of the total tons shipped in Ohio in 1993. The totals in Table 5.5 could be factored up by dividing by 0.92 if the contribution of these commodities was expected to stay the same in the future. However, we would propose a less drastic assumption, one that says that the sum of these commodities will continue to represent a "large proportion" of tons shipped in Ohio unless significant structural changes occur. In the absence of these changes, the calculated totals in Table 5.5 would represent a large percentage of shipments in Ohio in the corresponding year. Again, we emphasize that these numbers are intended only to be illustrative of the type of results that could be obtained with the use of indicators, and not to represent accurate forecasts.

Table 5.5 Forecasts of freight shipments originating in Ohio 1998-2025 [million tons] using various indicator methods

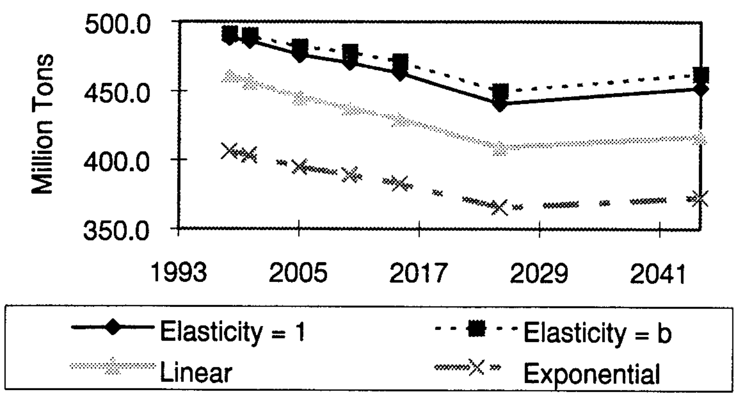
	1993		1998	2000	2005	2010	2015	2025
Non-Metallic Minerals	128.6	Elasticity=1	131.6	134.0	134.0	136.7	136.7	131.3
		Elasticity=b	132.0	134.8	134.8	137.9	137.9	131.7
		Linear	78.8	80.3	80.3	81.9	81.9	78.7
		Exponential	72.6	74.1	74.1	75.8	75.8	72.4
Petroleum/Coal Products	80.4	Elasticity=1	79.5	78.4	75.3	73.2	71.1	67.0
		Elasticity=b	79.8	79.1	77.2	75.9	74.6	72.0
		Linear	85.9	84.7	81.3	79.1	76.9	72.4
		Exponential	73.3	72.6	70.8	69.6	68.3	65.8
Primary Metal Products	40.8	Elasticity=1	39.6	39.1	37.7	36.5	35.4	33.5
		Elasticity=b	39.4	38.8	37.3	36.0	34.8	32.6
		Linear	36.2	35.7	34.5	33.4	32.4	30.7
		Exponential	36.1	35.6	34.2	33.0	31.9	30.0
Clay, Concrete, Glass, and Stone Products	35.5	Elasticity=1	33.9	33.2	31.7	30.4	29.4	27.5
		Elasticity=b	34.2	33.7	32.6	31.5	30.7	29.3
		Linear	54.0	53.0	50.6	48.6	46.9	43.9
		Exponential	44.1	43.4	41.9	40.6	39.5	37.5

Table 5.5 Forecasts of freight shipments originating in Ohio 1998-2025 [million tons] using various indicator methods (Continued)

	1993		1998	2000	2005	2010	2015	2025
Food, Kindred Products	32.8	Elasticity=1	32.8	32.8	32.5	32.3	31.9	31.0
		Elasticity=b	32.8	32.8	32.5	32.3	31.9	31.0
		Linear	31.3	31.3	31.0	30.8	30.4	29.6
		Exponential	31.4	31.4	31.1	30.9	30.5	29.7
Farm Products	32.7	Elasticity=1	32.0	31.6	30.8	29.9	28.7	26.1
		Elasticity=b	31.8	31.3	30.3	29.1	27.5	24.2
		Linear	19.5	19.2	18.8	18.2	17.5	15.9
		Exponential	15.7	15.4	14.9	14.3	13.6	12.1
Coal	27.4	Elasticity=1	21.7	19.4	16.5	14.3	13.1	10.8
		Elasticity=b	22.9	21.1	18.8	17.0	16.1	14.3
		Linear	31.4	28.0	23.9	20.6	18.9	15.6
		Exponential	32.0	29.3	25.8	22.9	21.5	18.5
Chemicals and Allied Products	25.0	Elasticity=1	25.1	25.1	25.2	25.2	25.2	24.8
		Elasticity=b	25.1	25.1	25.1	25.2	25.1	24.8
		Linear	28.9	28.9	29.0	29.0	29.0	28.5
		Exponential	20.8	20.8	20.8	20.9	20.8	20.6
Waste, Scrap Materials	11.4	Elasticity=1	11.7	11.7	11.8	11.8	11.8	11.6
		Elasticity=b	11.8	11.8	12.0	12.0	12.0	11.6
		Linear	6.1	6.1	6.2	6.2	6.2	6.1
		Exponential	6.4	6.4	6.5	6.5	6.5	6.3
Lumber, Wood Products	11.0	Elasticity=1	12.0	12.3	13.0	13.5	13.7	13.9
		Elasticity=b	12.1	12.5	13.2	13.8	14.1	14.3
		Linear	25.4	26.1	27.5	28.5	29.1	29.5
		Exponential	18.4	19.0	20.2	21.0	21.5	21.8
Transportation Equipment	10.4	Elasticity=1	10.5	10.5	10.5	10.5	10.4	10.1
		Elasticity=b	10.5	10.5	10.5	10.5	10.4	10.1
		Linear	6.4	6.4	6.4	6.3	6.3	6.1
		Exponential	4.1	4.1	4.1	4.1	4.1	4.0
Fabricated Metal Products	8.6	Elasticity=1	8.3	8.2	7.9	7.6	7.4	7.0
		Elasticity=b	8.3	8.2	7.9	7.6	7.4	7.0
		Linear	7.5	7.4	7.2	7.0	6.8	6.4
		Exponential	6.7	6.6	6.4	6.2	6.0	5.7
Pulp, Paper, Allied Products	8.6	Elasticity=1	8.6	8.6	8.6	8.6	8.6	8.4
		Elasticity=b	8.6	8.6	8.6	8.6	8.6	8.4
		Linear	10.7	10.8	10.8	10.8	10.7	10.5
		Exponential	10.2	10.2	10.2	10.2	10.2	10.0
Total	453.2	Elasticity=1	447.1	444.9	435.6	430.5	423.4	403.1
		Elasticity=b	449.3	448.3	440.8	437.3	431.1	411.5
		Linear	422.1	417.9	407.3	400.3	393.0	373.9
		Exponential	371.7	368.9	361.1	356.0	350.2	334.2

The differences in the freight movement forecasts produced by the different methods are apparent in Table 5.5. We illustrate these differences by graphing the forecast totals in Figure 5.2. This would indicate that the choice of indicator method used could make a difference in the forecast produced. We personally feel that the regression methods are preferable to the elasticity methods, but note that there is no reason to believe that the forecast produced by any method would be more accurate than that produced by any other method for a specific year. Therefore, if a forecast produced from the indicator methods would be useful in supporting policy or alternative analysis or as an input to a more complex model, we would presently suggest using all the methods to produce a range of forecasts that would serve as a lower bound indicator of the uncertainty in the expected freight movements.

Figure 5.2 Forecasts of freight shipments 1998-2045 [million tons] based on Table 5.1 commodities originating in Ohio using various indicator methods



5.3 Discussion

We demonstrated that indicator methods can presently be used with available public data sources to forecast generated freight at the state level. We have also been experimenting with indicators at the finer National Transportation Analysis Regions (NTAR) level (see Section 3) and obtaining similar results. We also showed that different indicator variables and methods could be considered. Although we believe that the regression methods should be preferred because they impose less restrictive assumptions than the other methods, this does not mean that these methods would produce more accurate forecasts for any given year. Similarly, the correlation statistics we calculate provide an indication of how well the calculated relationships fit freight variables to indicator variables, but they offer no indication of how well the forecasts will perform over time or how useful the forecasts will be in supporting analysis of public planning issues. An indicator method can be decided upon as a stand alone model or as a component of a more complex model if necessary, but longer term analysis would be required before one could use the results with any confidence.

Section 6. Intermodal Flows at New Facility: Example of Discrete Choice

Freight movement models could be used to help forecast usage at new freight terminals or transfer points. We illustrate variants of one such approach in this section. Although we do not expect new facility models to be used on an ongoing basis, we discuss this application in the event that ODOT would want to conduct or contract for such analysis of a new freight facility. We also note that like many models, alternatives for freight shipments -- the facility used, in this case -- are modeled as depending on certain level of service or performance characteristics. Indeed, we see that the heart of the methodology is similar to traditional modal choice models, where instead of facilities, the alternatives are the modes.

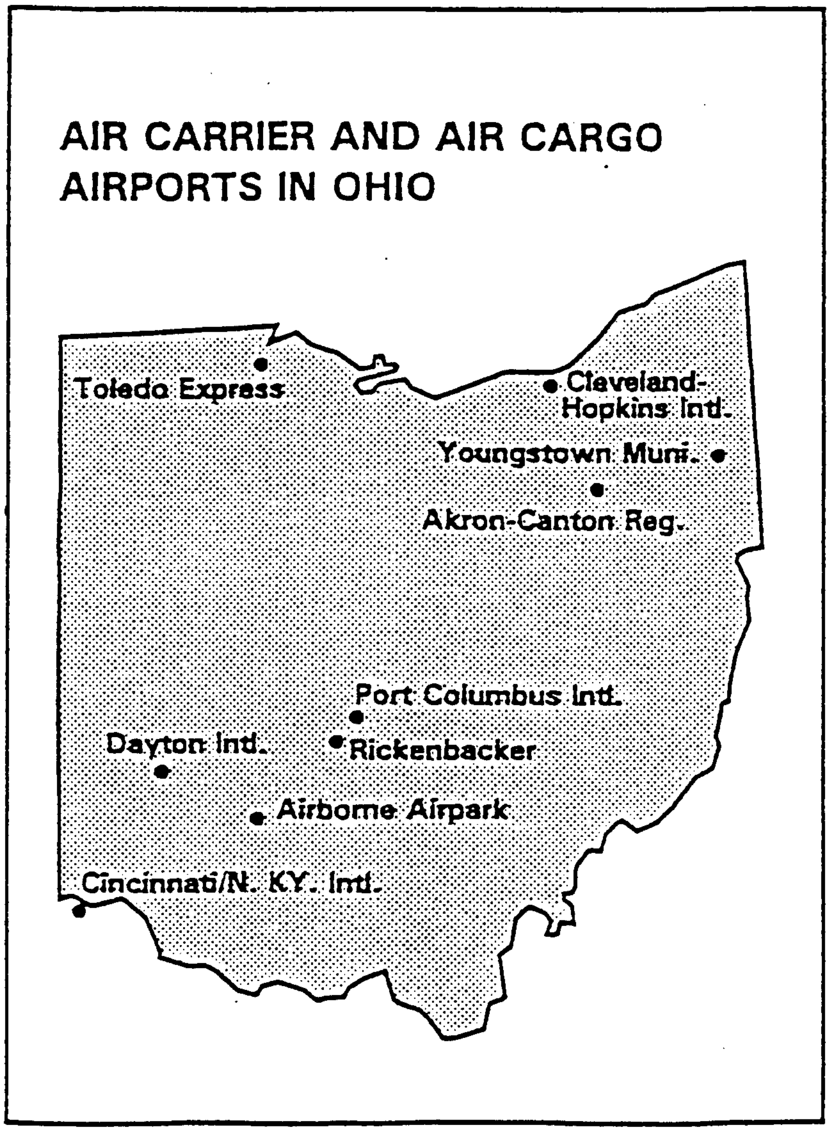
Again, we wish to illustrate that explicit freight movement modeling is possible, that models have been proposed and used for such analysis, and that variations of the models could easily be imagined. We illustrate this by discussing the application of what we call a New Facility Model *NFM* to the analysis of converting Rickenbacker military base, located in Central Ohio, to an intermodal, air-based facility. The *NFM* model has been proposed and used for new freight facility siting (Weinblatt and Edwards, 1997; Cambridge Systematics, Inc., *et al.*, 1997). Much of the data required would be obtained from private databases discussed in Section 3. Keeping with the illustrative nature of this report, we do not obtain these real data, but present the examples with realistic data. That is, the form of the data used in this analysis would be similar to that which could be obtained, but we assumed the actual numerical values of the data.

We also show that the different ways of combining independent variables proposed by the *NFM* developers can lead to very different forecasts. Different ways of fitting dependent variables to independent variables leads to forecasts that are very similar, but the differences due to the way in which the independent variables are combined illustrate how these similarities can be misleading in the absence of further investigation.

6.1 Setting of Illustrative Example

We consider converting Rickenbacker military base into an intermodal, air-based terminal that would compete for intermodal freight with existing facilities in Ohio. We consider the six facilities listed in the *Access Ohio* report (Ohio Department of Transportation, 1993). The locations of the facilities, and that of Rickenbacker, are shown in Figure 6.1.

Figure 6.1 Locations of Rickenbacker and Competing Ohio Air Freight Facilities (source: Exhibit 2-8 in Ohio Department of Transportation (1993))



Forecasting the amount of freight that would use Rickenbacker would be of interest when assessing the desirability of converting it to an intermodal facility and the desirability of alternative capacity and operational options of the converted facility. One component of freight usage would be traffic either arriving at the facility by ground and departing by air, or arriving by air and departing by ground. Similar to what is proposed in Cambridge Systematics, Inc., *et al.* (1997), we discuss one direction, that dealing with freight arriving by ground and departing by air, what we call ground-to-air freight. Separate models could be developed for the air-to-ground component, or this volume could be assumed to equal the ground-to-air component over an extended time period, such as a year (Cambridge Systematics, Inc., *et al.*, 1997). Also, similar to the methodology proposed in Cambridge Systematics, Inc., *et al.*, (1997), we only consider the freight diverted from existing facilities. Traffic generated by adding a new facility could eventually be added to the diverted traffic, but this would require a different methodology. However, the methodology presented here is not limited to existing freight patterns, but can just as easily be applied to forecasts of generated freight.

We base our analysis on that proposed in (Weinblatt and Edwards, 1997) and used in North Carolina (Cambridge Systematics, Inc., *et al.*, 1997). The methodology requires dividing the study area into geographic freight generating regions, calculating distances from these regions to the air facilities, and determining facility level-of-service variables. For demonstration purposes, we arbitrarily selected regions roughly corresponding to the National Transportation Analysis Regions (Bureau of Transportation Statistics, 1997b) in Ohio (see Figure 6.2) and calculated the distances from geographic centroids of these regions to the facilities (see Table 6.1). Freight generating regions could, of course, extend beyond the state borders. This would not complicate the type of analysis presented in this section.

Figure 6.2 Regions used in New Facility Model example applications

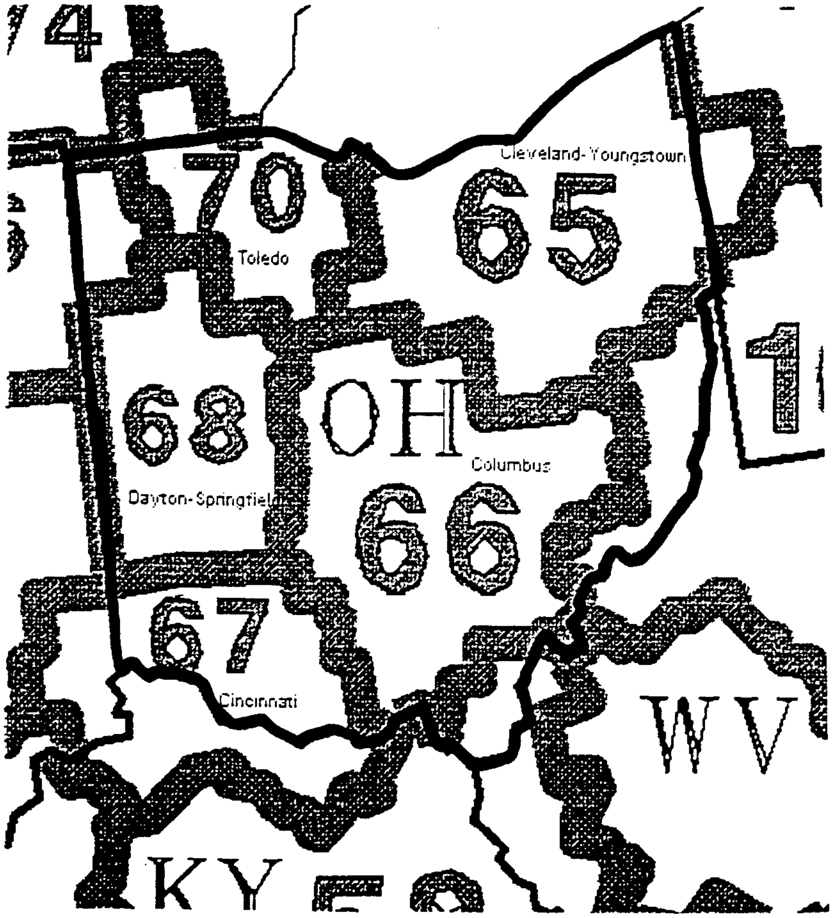


Table 6.1 Distances in miles from facilities of Figure 6.1 to regions of Figure 6.2

	Region I (Cleveland, NTAR65)	Region II (Columbus, NTAR66)	Region III (Cincinnati, NTAR67)	Region IV (Dayton, NTAR68)	Region V (Toledo, NTAR70)
Airborne Airpark	180	40	70	45	155
Dayton	215	80	50	10	145
Toledo	110	140	190	145	5
Cincinnati	240	105	5	50	190
Cleveland-Hopkins	10	140	240	215	110
Columbus	135	5	105	80	140
Rickenbacker	200	30	120	120	165

The facility level-of-service variables could take several forms. The North Carolina study used number of operations per year, available capacity, frequency of service, and cost of shipments. For demonstration purposes, we used frequency of service (measured by number of operations per year) and number of destinations served. We could not find values for these variables, even for the existing facilities, but note that the values used in the North Carolina study were obtained from the Colography Group, Inc., a private data supply firm reviewed in Section 3. Rather than purchase such data for an illustrative study, we assumed values of the data and present them in Table 6.2. We assumed values in such a way that the airports shown to handle more freight in 1991 (Ohio Department of Transportation, 1993) had larger numbers of operations and destinations served than those that handled less freight.

Table 6.2 Values of Level-of-Service variables assumed for the air facilities of Figure 6.1

	Frequency	No. of Destination
Airborne Airpark	49,000	100
Dayton	44,500	85
Toledo	42,000	63
Cincinnati	35,000	40
Cleveland-Hopkins	18,000	30
Columbus	800	13

The *Access Ohio* report (Ohio Department of Transportation, 1993) provides magnitudes and percentages of freight enplaned in 1991 at the existing Ohio airports but gives no information on the origins of the freight. To generate such values, we arbitrarily assumed that 20% of the ground-to-air freight would be generated in each of the five regions of Figure 6.1. These shares could be refined using public or private data bases, but the illustrative nature of this study did not warrant such refinement. We considered the unit of ground-to-air freight to be the ton, assumed the shares of ground-to-air tons using the existing facilities equaled the shares of total freight enplaned at the existing facilities in 1991, and arbitrarily assumed that the busiest facility (Airborne Airpark) would enplane an average of 957.5 tons per day (= 349387.5 tons per year).

With these conditions we calculated a total of 184732.8 (annual) tons of ground-to-air freight generated in each region. To determine the tons of ground-to-air freight generated in each region using each of the existing facilities, we used a multinomial logit model:

$$P(a'|r') = \frac{\exp[\beta_{FREQ}FREQ(a') + \beta_{NDES}NDES(a') + \beta_{DIST}DIST(r',a')]}{\sum_{a=1,6} \exp[\beta_{FREQ}FREQ(a) + \beta_{NDES}NDES(a) + \beta_{DIST}DIST(r',a)]},$$

$$a' = 1,2,\ldots,6;\ r' = 1,2,\ldots,5, \quad (6.1)$$

where $P(a'|r')$ is the probability that a ton of ground-to-air freight generated in region r' would use facility (airport) a', $FREQ(a)$ and $NDES(a)$ are, respectively, the number of yearly operations at and number of destinations served by facility a, $DIST(r',a)$ is the distance from the centroid of region r' to facility a, and β_{FREQ} ,β_{NDES} , and β_{DIST} are, respectively, coefficients of the *FREQ, NDES,* and *DIST* variables. We found the values of these coefficients by using the 1991 shares as estimates of probabilities that a unit of ground-to-air freight would be enplaned at the facility, and regressing these probabilities against the *FREQ, NDES,* and *DIST* variables in Equation (6.1). (Regressing shares against variables in a way consistent with the multinomial logit model is described below, where the emphasis is on methodology.) We found $\beta_{FREQ} = 3.41E\text{-}05$, $\beta_{NDES} = 1.28E\text{-}02$, $\beta_{DIST} = -1.75E\text{-}03$. This led to the $P(a'|r')$ values shown in Table 6.3.

Table 6.3 Probability of a unit of ground-to-air freight generated in region r' using facility *a'* assumed for study

	Region I (Cleveland, NTAR65)	Region II (Columbus, NTAR66)	Region III (Cincinnati, NTAR67)	Region IV (Dayton, NTAR68)	Region V (Toledo, NTAR70)
Airborne Airpark	0.365	0.397	0.382	0.382	0.351
Dayton	0.243	0.262	0.280	0.288	0.253
Toledo	0.203	0.164	0.152	0.158	0.224
Cincinnati	0.095	0.102	0.123	0.109	0.095
Cleveland-Hopkins	0.070	0.047	0.040	0.040	0.054
Columbus	0.025	0.027	0.023	0.023	0.023
Total	1.00	1.00	1.00	1.00	1.00

We assumed that the annual number of tons of ground-to-air freight using each facility in the base case (i.e., the scenario in which Rickenbacker did not exist as an air freight facility) was the expected (mean) number of tons using the facility, found by combining the number of tons assumed to be generated in each region above with the conditional probabilities of Table 6.3. We present these numbers in Table 6.4. All numbers would change with time and fluctuate from year to year, of course, but we assumed one year of data for the illustrations used in this section.

Table 6.4 Tons of ground-to-air freight generated in region *r'* using facility *a'* assumed for study

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	67,443.8	73,445.7	70,555.7	70,650.7	64,875.6	346,972	37.6%
Dayton	44,907.0	48,447.4	51,725.6	53,172.9	46,736.4	245,019	26.5%
Toledo	37,393.4	30,242.5	28,053.2	29,091.6	41,374.8	166,155	18.0%
Cincinnati	17,447.2	18,866.7	22,754.3	20,157.9	17,563.5	96,819	10.5%
Cleveland-Hopkins	12,880.2	8,744.7	7,432.0	7,442.1	9,955.5	46,454	5.0%
Columbus	4,631.2	4,955.8	4,212.0	4,217.6	4,227.0	22,243	2.4%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100%

In summary, for our illustrations we assume that the six facilities other than Rickenbacker shown in Figure 6.1 operate with frequencies given in Table 6.2 and serve the number of destinations shown there. Ground-to-air freight using these six facilities is generated from the five regions portrayed in Figure 6.1. The distances from these five regions to these six facilities are given in Table 6.1, and the quantities of freight generated in the regions and using the airports, measured in units of tons, are those in Table 6.4.

6.2 New Facility Model

We call the model proposed in Weinblatt and Edwards (1997) to predict freight flows to a new facility and applied in North Carolina (Cambridge Systematics, Inc., *et al.*, 1997) New Facility Model *NFM*. We summarize the model as consisting of the following steps:

1) Determine competing facilities;
2) Divide the study area into freight generating regions;
3) Forecast the annual freight volumes produced in (or attracted to) each region;
4) Assign a "proximity score" for each region-facility combination;
5) Assign a "Level-of-Service (LOS) score" for each facility;
6) Combine the proximity and LOS scores to determine an "overall score" for each region-facility combination;
7) Based on the overall scores of Step 6), assign the region freight forecasts of Step 3) to the facilities;
8) Sum the assignments of Step 7) across regions to obtain freight forecasts for each facility.

We illustrate this procedure by applying it to the Rickenbacker conversion example.

Steps 1) , 2) and 3) have been illustrated in Section 6.1. The resulting facilities and regions are shown Figures 6.1 and 6.2. The regional freight forecasts are found as the Regional Totals in Table 6.4.

The developers for *NFM* suggest that the proximity score of Step 4) be based on the "highway distance" from the approximate centroid of economic activity in the region to the facility. A default function is provided to map these distances into proximity scores and illustrated in Figure 6.3. The region-facility distances were given above in Table 6.1. Applying the function of Figure 6.3 to the distances of Table 6.1 leads to the proximity scores of Table 6.5. The authors propose that the proximity function could be adjusted to consider transport cost, time, and reliability, although no guidance or default function is provided. However, the functions could be calibrated in a manner similar to that shown below.

Figure 6.3 Suggested proximity score function for New Facility Model (taken from: Weinblatt and Edwards, 1997)

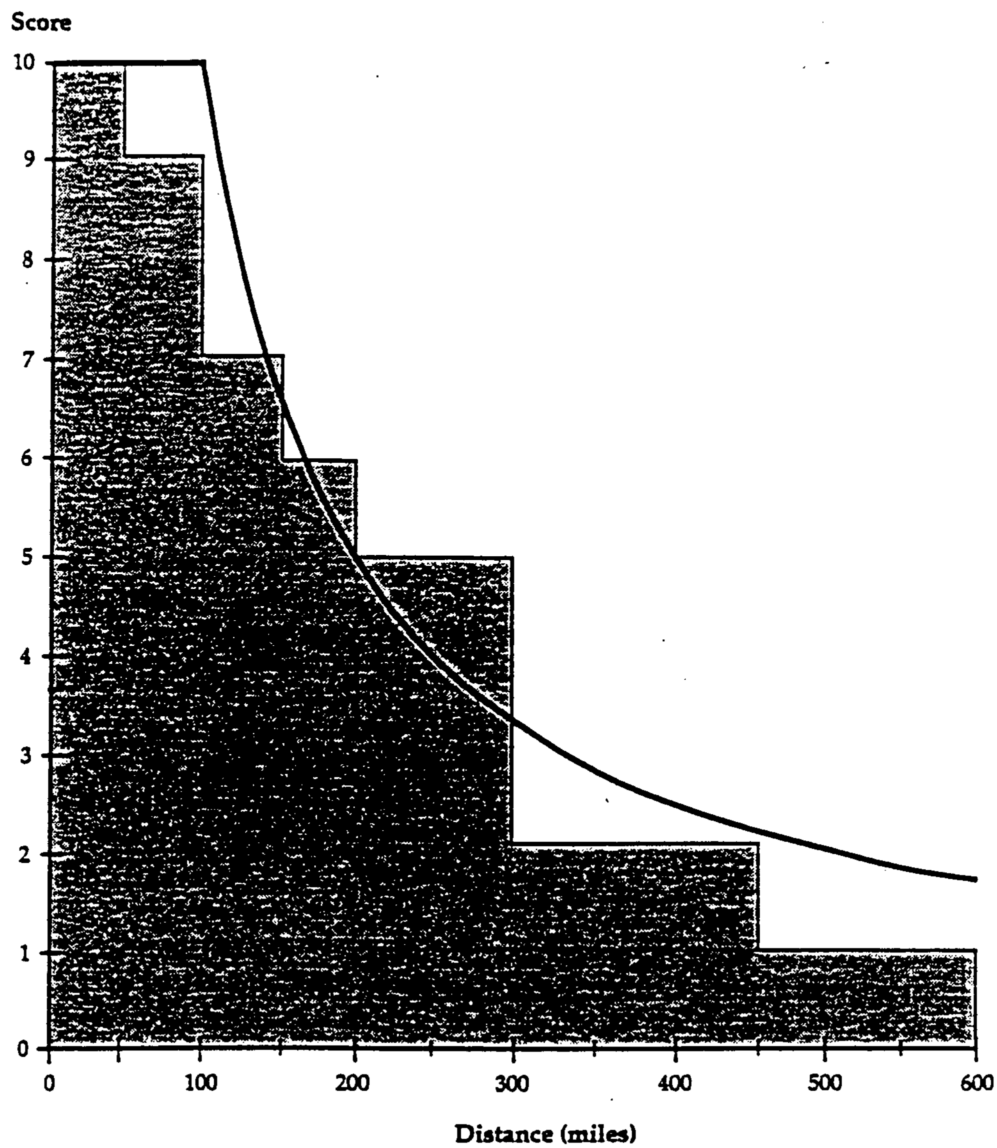


Table 6.5 Proximity scores for New Facility Model based on region-facility distances of Table 6.1 and proximity score function of Figure 6.3

	Region I (Cleveland, NTAR65)	Region II (Columbus, NTAR66)	Region III (Cincinnati, NTAR67)	Region IV (Dayton, NTAR68)	Region V (Toledo, NTAR70)
Airborne Airpark	5.4	10.0	10.0	10.0	6.5
Dayton	4.7	10.0	10.0	10.0	6.8
Toledo	9.0	6.9	5.2	6.8	10.0
Cincinnati	4.1	9.5	10.0	10.0	5.2
Cleveland-Hopkins	10.0	6.9	4.1	4.7	9.0
Columbus	7.2	10.0	9.5	10.0	6.9
Rickenbacker	5.0	10.0	7.0	7.0	6.0

To develop the LOS score in Step 5), the *NFM* developers suggest that the facilities be compared in terms of number of destinations served, frequency of service, and unit costs of carriers using the facilities. Other than suggesting that the LOS scores be determined on the basis of judgment, that the facility considered best on these criteria receive a score of 10, and that the scores of the other facilities be determined in relation to this score, no guidance is provided for determining these scores. In our illustrative example, we base LOS scores on destinations served and frequency of service (number of operations).

We assumed data values so that the busiest existing facility was the best on each of these measures, the next busiest existing facility was next best on each of the measures, and so on (see Table 6.2). We also assumed data for the Rickenbacker facility so that its numbers of destinations and operations were equal to those of the Cincinnati facility. Therefore, the ranking of facilities was obvious. Since no guidance is given for mapping the LOS variables into LOS scores, we simply assigned scores for the existing (without Rickenbacker) case so that the scores were uniformly distributed between 1 and 10. We then assigned the Rickenbacker facility a score equal to that of Cincinnati. These LOS scores are given in Table 6.6.

Table 6.6 New Facility Model Level-of Service scores based on facility measures of Table 6.2

Airports	LOS
Airborne Airpark	10.0
Dayton	8.2
Toledo	6.4
Cincinnati	4.6
Cleveland-Hopkins	2.8
Columbus	1.0
(Rickenbacker)	(4.6)

Developers of the model suggest that the proximity and LOS scores be combined either additively or multiplicatively in Step 6). That is, the overall score for a region-facility combination could either be the sum or the product of the region-facility proximity score determined in Step 4) and the LOS score of the facility determined in Step 5). The sets of overall scores formed by using the additive and multiplicative combinations with the LOS scores of Table 6.6 are presented in Table 6.7.

Table 6.7 New Facility Model overall scores formed by combining LOS scores of Table 6.6 with proximity scores of Table 6.5 additively and multiplicatively

6.7a Additive combination

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)
Airborne Airpark	15.4	20.0	20.0	20.0	16.5
Dayton	12.9	18.2	18.2	18.2	15.0
Toledo	15.4	13.3	11.6	13.2	16.4
Cincinnati	8.7	14.1	14.6	14.6	9.8
Cleveland-Hopkins	12.8	9.7	6.9	7.5	11.8
Columbus	8.2	11.0	10.5	11.0	7.9
(Rickenbacker)	(9.6)	(14.6)	(11.6)	(11.6)	(10.6)
Total	73.4 (83.0)	86.3 (100.9)	81.8 (93.4)	84.5 (96.1)	77.4 (88.0)

6.7b Multiplicative combination

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)
Airborne Airpark	54.00	100.00	100.00	100.00	65.00
Dayton	38.54	82.00	82.00	82.00	55.76
Toledo	57.60	44.16	33.28	43.52	64.00
Cincinnati	18.86	43.70	46.00	46.00	23.92
Cleveland-Hopkins	28.00	19.32	11.48	13.16	25.20
Columbus	7.20	10.00	9.50	10.00	6.90
(Rickenbacker)	(23.00)	(46.00)	(32.20)	(32.20)	(27.60)
Total	204.20 (227.20)	299.20 (345.18)	282.30 (314.46)	294.70 (326.88)	240.80 (268.38)

In Step 7), the *NFM* developers suggest that the freight from a region be assigned to a facility in direct proportion to the fraction of the total overall score for a facility that comes from the region. For example, consider the freight to be assigned based on using an additive combination of proximity and LOS scores. The overall scores for this scenario are given in Table 6.7a. In this scenario, the proportion of, for example, Region I freight assigned to the Airborne Airpark facility would be the proportion of the total freight from Region I (184732.8 tons according to Table 6.4) obtained by taking the ratio of the Region I-Airborne overall score (15.4) to the sum of the region-facility overall scores for the Airborne facility (73.4). That is, the forecast use of Region I freight using Airborne Airpark would be (15.4/73.4) x 184732.8 = 38758.7 [tons].

Summing across regions in Step 8) is straightforward. The forecast region-facility assignments and totals for subareas and regions for the four overall score scenarios considered are given in Table 6.8.

Table 6.8 Forecast freight in tons using facilities by region based on the four sets of overall scores of Table 6.6

6.8a Additive composition for overall score, without Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	38,758.7	42,811.8	45,166.9	43,723.7	39,381.0	209,842.1	22.7%
Dayton	32,466.7	38,958.7	41,101.9	39,788.6	35,800.9	188,116.8	20.4%
Toledo	38,758.7	28,469.8	26,196.8	28,857.7	39,142.4	161,425.3	17.5%
Cincinnati	21,896.1	30,182.3	32,971.9	31,918.3	23,389.9	140,358.6	15.2%
Cleveland-Hopkins	32,215.0	20,763.7	15,582.6	16,396.4	28,163.4	113,121.1	12.2%
Columbus	20,637.7	23,546.5	23,712.6	24,048.1	18,855.2	110,800.1	12.0%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664.0	100.0%

6.8b Additive composition for overall score, with Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	34,275.7	36,617.0	39,557.3	38,446.0	34,637.4	183,533.4	19.9%
Dayton	28,711.5	33,321.5	35,997.2	34,985.8	31,488.5	164,504.5	17.8%
Toledo	34,275.7	24,350.3	22,943.3	25,374.3	34,427.5	141,371.1	15.3%
Cincinnati	19,363.6	25,815.0	28,876.9	28,065.5	20,572.5	122,693.5	13.3%
Cleveland-Hopkins	28,488.9	17,759.2	13,647.3	14,417.2	24,771.0	99,083.7	10.7%
Columbus	18,250.7	20,139.4	20,767.6	21,145.3	16,584.0	96,886.9	10.5%
Rickenbacker	21,366.7	26,730.4	22,943.3	22,298.7	22,251.9	115,590.9	12.5%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664.0	100.0%

6.8c Multiplicative composition for overall score, without Rickenbacker

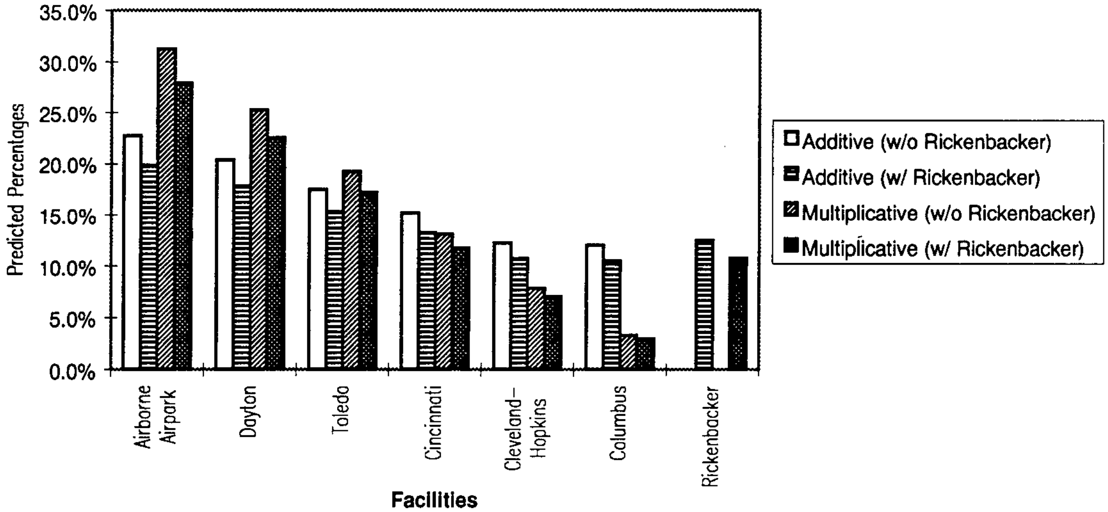
	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	48,852.0	61,746.4	65,447.7	62,689.3	49,869.7	288,605.1	31.2%
Dayton	34,865.8	50,632.0	53,667.1	51,405.2	42,780.6	233,350.8	25.3%
Toledo	52,108.8	27,267.2	21,781.0	27,282.4	49,102.5	177,541.8	19.2%
Cincinnati	17,062.0	26,983.2	30,106.0	28,837.1	18,352.1	121,340.3	13.1%
Cleveland-Hopkins	25,330.6	11,929.4	7,513.4	8,249.9	19,334.1	72,357.5	7.8%
Columbus	6,513.6	6,174.6	6,217.5	6,268.9	5,293.9	30,468.6	3.3%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664.0	100.0%

6.8d Multiplicative composition for overall score, with Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	43,906.6	53,517.8	58,746.0	56,514.0	44,741.2	257,425.5	27.9%
Dayton	31,336.3	43,884.6	48,171.8	46,341.4	38,381.0	208,115.1	22.5%
Toledo	46,833.7	23,633.5	19,550.7	24,594.9	44,052.8	158,665.5	17.2%
Cincinnati	15,334.8	23,387.3	27,023.2	25,996.4	16,464.7	108,206.4	11.7%
Cleveland-Hopkins	22,766.4	10,339.6	6,744.0	7,437.2	17,345.8	64,633.1	7.0%
Columbus	5,854.2	5,351.8	5,580.9	5,651.4	4,749.4	27,187.7	2.9%
Rickenbacker	18,700.9	24,618.2	18,916.2	18,197.5	18,997.8	99,430.6	10.8%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664.0	100.0%

In Figure 6.3 we plot the percentages of total generated freight using the various air facilities for the four different cases corresponding to Table 6.8--with and without Rickenbacker, each with the additive and multiplicative composition for the overall score. In general, the differences between the percentages predicted by the additive and multiplicative forms for a given scenario (either with or without Rickenbacker) are greater than the differences between the percentages predicted in the with and without Rickenbacker cases for a given functional form (either additive or multiplicative). For example, the difference between the percentages of freight predicted to use Airborne Airpark for the additive and multiplicative compositions in the without Rickenbacker case can be seen to be approximately 8% in Figure 6.3 (and calculated to be 8.5% from Table 6.8). For the additive composition, the difference between the percentages of freight predicted to use Airborne Airpark in the with and without Rickenbacker cases can be seen to be only approximately 3% in Figure 6.3 (and calculated to be 2.8% from Table 6.8). The exception is the difference in the Rickenbacker and, perhaps, the Toledo and Cincinnati facilities. At the Rickenbacker facility there is a great difference between the with and without scenarios (roughly, 10-12%) because the facility had no freight in the without scenario, while the difference between the additive and multiplicative forecasts when Rickenbacker is added is only approximately 2%. The differences in forecasts due to functional form and those stemming from adding Rickenbacker cases are approximately the same in the Toledo and Cincinnati cases.

Figure 6.4 Predicted percentages using the facilities under additive and multiplicative composition assumptions in the with and without Rickenbacker



Since neither the additive or multiplicative form is prescribed as more appropriate, the differences due to the different forms can be thought of as a lower bound on the accuracy of the models; that is, one should not be able to predict percentages to a level finer than these differences. Since the differences in the predicted percentages between the with and without Rickenbacker cases are less than or equal to these differences, this model would not be useful in predicting changes in freight usage at the facilities.

Both the additive and multiplicative compositions capture the general trends of Table 6.4 (see Table 6.9). This is not very surprising, since although the Table 6.4 data was generated using a model different from the NFM, it was generated using the same causal variables assumed in the NFM--frequency, number of destinations served, and distance. The forecasts obtained with the multiplicative formulation match the Table 6.4 data better than do the forecasts obtained with the additive formulation. However, once again, no reason is given to prefer the multiplicative over the additive formulation. Although the forecasts obtained with the additive composition do rank the facilities in the same order as does the generated data leading to Table 6.4, the differences in the magnitudes of the predicted percentages are great enough to indicate that such a model could only be used as a very rough estimate of the actual magnitude of freight using the facilities.

Table 6.9 Percentages of freight using the facilities in the without Rickenbacker scenario generated by region predicted in the New Facility Model under additive (from Table 6.8a) and multiplicative (from Table 6.8c) score compositions and base data

	Additive	Multiplicative	Base data
Airborne Airpark	22.7%	31.2%	37.6%
Dayton	20.4%	25.3%	26.5%
Toledo	17.5%	19.2%	18.0%
Cincinnati	15.2%	13.1%	10.5%
Cleveland-Hopkins	12.2%	7.8%	5.0%
Columbus	12.0%	3.3%	2.4%
Total	100.0%	100.0%	100.0%

6.3 Model Calibration

The new facility model *NFM* presented above is essentially a two-stage model, one that "generates" freight in regions, and one that assigns the freight generated in each region to the various facilities. The methodological details in the model lie in the latter stage. The suggested method is typical of many passenger trip distribution and modal choice models in that the assignments are based on measures that indicate the relative attractiveness of the alternative (the facility, in this case) and of the "trip maker"-alternative (region-facility, in this case) pair. In the example developed above, the assignment is based on the distance of the region centroid to the facility and on the number of destinations served and annual operations at the facility.

The empirical results of the preceding section illustrated that the *NFM*, as presented, could only be used to obtain a very rough estimate of freight using the various facilities. The advantage of this type of model is that it is easy to use, and for the assignment stage presents default, "quick response-like" parameter values. As with most forecasting models, one might want to use data to calibrate parameters of the model. This will ensure a better fit to the conditions used in the calibration. Of course, this improved fit to existing conditions will not necessarily translate to improved forecasts of altered or future conditions if no acceptable behavioral meaning can be ascribed to the functional form of the model. Still, we discuss calibrating the models to data because: *i)* it is commonly suggested for transportation models, even those with no behavioral interpretation; and *ii)* if done over many data sets, it can be used to identify robust correlations among independent and dependent variables. We discuss calibrating the *NFM* model, and an alternative logit formulation of the *NFM*.

Assuming that the proximity scores for region r_j to facility a_i and the LOS score for facility a_i--denoted $PROX_{i,j}$ and LOS_i, respectively--are combined additively in Step 6) of the *NFM*, the model can be written:

$$F(a_{i'}) = \sum_{j=1,J} F(a_{i'}, r_j) \quad , \qquad i' = 1, \ldots, I \qquad (6.2a)$$

$$F(a_{i'}, r_j) = F(r_j) \frac{(PROX_{i',j} + LOS_{i'})}{\sum_{i=1,I} (PROX_{i,j} + LOS_i)} \quad , \quad i' = 1, \ldots, I;\ j = 1, \ldots, J \qquad (6.2b)$$

where, $F(a_{i'})$ is the amount of freight using facility $a_{i'}$, $F(a_{i'},r_j)$ is the amount of freight using facility $a_{i'}$ generated in region r_j, $PROX_{i,j}$ is the proximity score of facility a_i to region r_j, LOS_i is the Level-of-Service score for facility a_i, there are I airports, and there are J regions. A similar model can be written for the multiplicative combination of proximity and LOS scores by changing the $PROX_{.,.} + LOS_.$ term to $PROX_{.,.} \times LOS_.$. For ease in exposition, we only discuss the additive combination here.

Given the form of equation (6.2b), the assumptions in the *NFM* model are related to the way in which the LOS and PROX scores are determined. The proximity score function of Figure 6.3, which is a function of distance *DIST*, is claimed to be based on experience, and the developers also state that judgment should be used in deciding upon

LOS scores from the LOS variables, $FREQ(a_i)$ and $NDES(a_i)$ in the example above. One could think of the *PROX.,.+ LOS.* expression (or the multiplicative equivalent) as a value function, *V(PROX, LOS) = PROX.,.+ LOS..* With this concept, one could investigate the form of *V(PROX, LOS) = V(DIST, FREQ, NDES)* using either structured value assessment techniques (e.g., Keeney and Raiffa, 1976) from experts if no historical data are available or econometric fitting techniques if such data are available.

To illustrate the type of fitting that could be done with historical data, consider what we call "Data Calibration Scenario 1." This scenario assumes that the amount of freight using the various facilities $F(a_i)$, $i = 1, ..., I$, and the amount of freight from the various regions $F(r_j)$, $j = 1, ..., J$, are known. For example, the amount of freight using the existing facilities was found from the *Access Ohio* Report (Ohio Department of Transportation, 1993), and the freight from the various regions could be found from historical data or estimated in the first stage of *NFM* using generation models, such as the indicator models considered in Section 5. Consider that one accepted the proximity score function of Figure 6.3, assumed that the highest ranked facility had a LOS score of 10, and wanted to fit the LOS scores for the other facilities. In this case, the set of *I* equations formed when substituting the equations of (6.2b) in those of (6.2a) would contain *I-1* unknowns (the LOS for all but the LOS =10 facility). One could conceivably use some fitting technique, but inspection of (6.2) shows that the equations would be nonlinear in the LOS unknowns, and some efficient solution techniques would need to be used.

We used trial and error to obtain what we thought were reasonable fits of the LOS variables in the "without Rickenbacker" case. We then set the Rickenbacker LOS equal to that of the Cincinnati facility, as before. The LOS values and the resulting forecasts for use of the various facilities in the "without" and "with" Rickenbacker cases are presented in Tables A6.1a-c in Appendix 6.

In "Data Calibration Scenario 2, " we still use the historical data on the freight using the facilities and the freight generated from each region, but relax the constraint that the LOS score of the best facility is equal to 10; rather, we fit all the LOS scores. This offers more flexibility and makes it easy to handle new facilities with better LOS scores than the previous best facility. As such, it seems preferable. In this case the I equations formed when substituting the equations of (6.2b) in those of (6.2a) would now contain I unknowns (the I LOS scores). Again, the equations would be nonlinear in the I unknowns and not easy to solve. Using the same data as before, we again found "good fitting" LOS values by trial and error. The LOS values and the resulting "forecasts" are given in Tables A6.2a-c in Appendix 6. Note that the "forecasts" in the "without" Rickenbacker case should replicate the original data of Table 6.4. We came close but did not replicate these data exactly because we did not find values that would solve the 6 nonlinear equations exactly.

Historical data on the amount of freight shipped from each region to each facility $F(a_i,r_j)$ could be used advantageously. This data would require special surveys which could be commissioned from private firms (see Section 3.2). With this data, the $I \times J$ equations in (6.2b) could be transformed into:

$$-F(r_j)PROX_{i',j} + F(a_{i'},r_j)\sum_{i=1,I} PROX_{i,j} = F(r_j)LOS_{i'} - F(a_{i'},r_j)\sum_{i=1,I} LOS_i \quad ,$$

$$i' = 1, \ldots, I; \quad j = 1, \ldots, J \qquad (6.3)$$

The left-hand side of (6.3) would be known, and the right-hand side would contain the unknown LOS values to be fit. The advantage over the preceding scenarios is that the $I \times J$ equations are now linear in the unknowns and standard linear regression packages could be used to solve for best fit LOS values.

In "Data Calibration Scenario 3," we fix the best LOS at 10 and fit the 30 ($I = 6$ facilities $\times$ $J = 5$ regions) equations to the 5 (= I-1) LOS values. In "Data Calibration Scenario 4," we do not constrain the best LOS to be 10 and regress the 30 equations to find estimates of the 6 unknown LOS values. Again, the LOS value of Rickenbacker is considered to be that of the Cincinnati facility. The results are presented in Tables A6.3 and A6.4 in Appendix 6. Notice that this technique can, and did, lead to negative LOS values. The negative LOS values could lead to predictions of negative flows from a region to a facility (see, e.g., Table A6.3b and c). Negative flows would not make sense, and some type of *ad hoc* adjustment could conceivably be made. Our interest was in showing concepts, and we see no need to propose such adjustments at this time. Therefore, we leave the negative LOS values and flow predictions. Although the *NFM* model was proposed as intuitive, without claim to theoretically appealing foundations, the potential for negative values indicates a theoretical problem with the model.

One could also consider explicit functional forms for the LOS and proximity scores, for example:

$$LOS_i = \beta_1 FREQ(a_i) + \beta_2 NDES(a_i), \quad i = 1, 2, \dots I, \tag{6.4}$$

$$PROX_{i,j} = \beta_3 DIST(a_i,r_j), \qquad i = 1, 2, \dots I; \quad j = 1, 2, \dots, J, \tag{6.5}$$

where, as above, $FREQ(a_i)$ and $NDES(a_i)$ are, respectively, the number of yearly operations at and number of destinations served by facility a_i, and $DIST(a_i,r_j)$ is the distance between facility a_i and region r_j. Alternatively, one could use $PROX_{i,j} = \beta_3 / DIST(a_i,r_j)$ to approximate the proximity function in Figure 6.3 better. Doing so would not complicate any of the following. Accepting the proximity function of Figure 6.3 would eliminate (6.5) and, therefore, the need to estimate β_3. One could also add constant terms to the LOS equation--e.g., $LOS_i = \beta_i + \beta_1 FREQ(a_i) + \beta_2 NDES(a_i)$--to reflect the unspecified influence of variables other than *FREQ* and *NDES* on the LOS score of the facility. For illustration purposes, we use (6.4) and (6.5).

Substituting the equations of (6.4) and (6.5) in those of (6.2b), one could then proceed as above. Specifically, if one only had estimates of the amount of freight produced in each region $F(r_j)$ and shipped to each facility $F(a_i)$, one would have to estimate the three β values from the $I = 6$ equations formed by substituting (6.2b) in (6.1). Unfortunately, the equations would be nonlinear in the β values, and solving for a good solutions would not be easy. As before, we used trial and error for this "Data Calibration Scenario 5" and found the β and LOS values given in Table A6.5a in Appendix 6. Here, β_{FREQ} = -1.3E-06, β_{NDES} = 0.004492, and β_{DIST} = -0.00035. The LOS and proximity scores for Rickenbacker would be found from (6.4) and (6.5) using the Rickenbacker *FREQ*, *NDES*, and *DIST* values with the estimated β's. The freight forecasts are given in Tables A6.5b and A6.5c in Appendix 6.

The functional form used in (6.2b) to predict freight assignments from a region to a facility based on the proximity and LOS scores could also be reconsidered. For example, the freight could be assigned using a multinomial logit formulation:

$$F(a_{i'}, r_j) = F(r_j)\{\frac{\exp \alpha(PROX_{i',j} + LOS_{i'})}{\sum_{i=1,I} \exp \alpha(PROX_{i,j} + LOS_i)}\},$$

$$i' = 1,\ldots, I; \quad j = 1,\ldots, J, \qquad (6.6)$$

where *exp(.)* is the inverse function of the natural logarithm, α is a scaling parameter, and all other variables are as before. Much of theoretical appeal for using a logit model disappears in light of the difficulties discussed in the Introduction of this report. Still, logit is a widely used model with properties just as appealing as those of the *NFM* formulation. Moreover, it is familiar to many transportation planners and engineers.

Fitting proximity and LOS scores to data would proceed as with the *NFM* formulation. Fitting the LOS scores in a way similar to either Data Calibration Scenario 1 or 2, where the $F(a_i,r_j)$ are unknown, would involve solving nonlinear equations in the unknowns. If the $F(a_i,r_j)$ were known, one could form $(I-1)\times J$ linear equations in the unknowns from (6.6) by noting that:

$$\ln\{\frac{[F(a_{i'},r_j)/F(r_j)]}{[F(a_{i''},r_j)/F(r_j)]}\}=\ln\{\frac{F(a_{i'},r_j)}{F(a_{i''},r_j)}\}=\alpha(PROX_{i',j}-PROX_{i'',j})+\alpha(LOS_{i'}-LOS_{i''}) \quad ,$$

$$\text{for } i' = 1, \ldots, I, \; i' \neq i''; \; j = 1, \ldots, J; \text{ and some } i'', \qquad (6.7)$$

In "Data Calibration Scenario 6", we fixed the Airborne LOS to 10, assumed the proximity score function given in Figure 6.3, and used the $(I\text{-}1)\times J = 5\times 5 = 25$ equations of (6.7) to find the five other LOS values and the value of the scaling parameter α presented in Table A6.6a of Appendix 6. As before, the LOS score for Rickenbacker is assumed to be equal to that of the Cincinnati facility. The resulting freight forecasts for Data Calibration Scenario 6 are given in Tables A6.6b and A6.6c in Appendix 6. Table A6.6a exhibits negative LOS scores. This is perfectly compatible with the intervally-scaled interpretation of the function used in logit models (see, e.g., McCord and Villoria, 1987; Ben-Akiva and Lerman, 1985), however, and negative signs will not produce negative freight flows.

Specifying PROX and LOS as functions of their independent variables *DIST*, *FREQ*, and *NDES*, as in (6.4) and (6.5), and fitting these specifications to data would also proceed as in the *NFM* formulation. Unknown $F(a_i,r_j)$ would lead to nonlinear equations in the three β values of (6.4) and (6.5). On the other hand, if the $F(a_i,r_j)$ were known, substitution of (6.4) and (6.5) in (6.7) would lead to $(I\text{-}1)\times J$ linear equations to be fit to the three β values. (The value of the scaling parameter α is absorbed in the β values in this case.) We present the results of this "Data Calibration Scenario 7" in Tables A6.7a-c of Appendix 6.

Table 6.10 Summary of Calibration Scenarios

	Data Calibration Scenario 1	Data Calibration Scenario 2	Data Calibration Scenario 3	Data Calibration Scenario 4	Data Calibration Scenario 5	Data Calibration Scenario 6	Data Calibration Scenario 7
Freight from Regions, $F(r_j)$	Known	Known	Known	Known	Known	Known	Known
Freight at Facilities, $F(a_i)$	Known	Known	Known	Known	Known	Known	Known
Freight from Region to Facility, $F(a_i, r_j)$	Unknown	Unknown	Known	Known	Unknown	Known	Known
Structure (LOS, PROX)	NFM	NFM	NFM	NFM	NFM	Logit	Logit
LOS Score	$LOS(a_i)$, best LOS = 10	$LOS(a_i)$, best LOS uncon-strained	$LOS(a_i)$, best LOS = 10	$LOS(a_i)$, best LOS uncon-strained	$\beta_1 FREQ + \beta_2 NDES$	$LOS(a_i)$, best LOS = 10	$\beta_1 FREQ + \beta_2 NDES$
Proximity Score	From Fig. 6.3	From Fig. 6.3	From Fig. 6.3	From Fig. 6.3	$\beta_3 DIST$	From Fig. 6.3	$\beta_3 DIST$
Results	Tables A6.1	Tables A6.2	Tables A6.3	Tables A6.4	Tables A6.5	Tables A6.6	Tables A6.7

We summarize the Calibration Scenarios in Table 6.10. In Table 6.11, we present the predicted percentages obtained under the various Calibration Scenarios (taken from the tables of Appendix 6), along with the predicted percentages in the uncalibrated models. As expected, Calibration Scenario 7 replicates the Table 6.4 data, since Calibration Scenario 7 assumes the same model used to generate the data. (Calibration Scenario 6 also replicates the percentages of Table 6.4, but upon inspection of the detailed data in Appendix 6, it does not replicate the exact facility-region volumes.) In Table 6.12a, we present the absolute values of the deviations from the Table 6.4 data of the without Rickenbacker forecasts obtained under the various Calibration Scenarios. We did not generate any data for the with Rickenbacker case, but if we followed the logit generation scheme used in the without Rickenbacker case, the data would have been identical to that obtained in the with Rickenbacker case under Calibration Scenario 7. Therefore, in Table 6.12b we present the absolute values of the deviations from with Rickenbacker-Calibration Scenario 7 percentages of the forecasts obtained under the various Calibration Scenarios.

Table 6.11 Summary of forecast freight percentages handled at facilities under different data calibration scenarios

6.11a Without Rickenbacker case

	Calibrated							Uncalibrated	
	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Additive	Multiplicative
Airborne Airpark	35.6%	37.1%	37.4%	37.6%	34.9%	37.6%	37.6%	22.7%	31.2%
Dayton	27.8%	27.5%	26.5%	26.5%	28.7%	26.5%	26.5%	20.4%	25.3%
Toledo	19.0%	18.1%	18.2%	18.0%	18.7%	18.0%	18.0%	17.5%	19.2%
Cincinnati	9.8%	9.7%	10.2%	10.4%	9.1%	10.5%	10.5%	15.2%	13.1%
Cleveland-Hopkins	5.5%	5.4%	5.4%	5.1%	6.2%	5.0%	5.0%	12.2%	7.8%
Columbus	2.3%	2.2%	2.2%	2.3%	2.4%	2.4%	2.4%	12.0%	3.3%
Sum	100%	100%	100%	100%	100%	100%	100%	100%	100%

6.11b With Rickenbacker case

	Calibrated							Uncalibrated	
	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Additive	Multiplicative
Airborne Airpark	32.7%	34.0%	34.4%	34.4%	32.1%	34.2%	34.1%	19.9%	27.9%
Dayton	25.5%	25.2%	24.4%	24.2%	26.3%	24.1%	24.1%	17.8%	22.5%
Toledo	17.4%	16.7%	16.8%	16.5%	17.2%	16.4%	16.3%	15.3%	17.2%
Cincinnati	9.0%	8.9%	9.3%	9.5%	8.4%	9.5%	9.5%	13.3%	11.7%
Cleveland-Hopkins	5.0%	4.9%	5.0%	4.7%	5.7%	4.6%	4.6%	10.7%	7.0%
Columbus	2.1%	2.1%	2.0%	2.1%	2.2%	2.2%	2.2%	10.5%	2.9%
Rickenbacker	8.3%	8.2%	8.1%	8.7%	8.1%	9.1%	9.3%	12.5%	10.8%
Sum	100%	100%	100%	100%	100%	100%	100%	100%	100%

Table 6.12 Absolute value of deviation of forecast freight percentages handled at facilities from simulated base data

6.12a Without Rickenbacker case

	Calibrated							Uncalibrated	
	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Additive	Multiplicative
Airborne Airpark	2.0%	0.5%	0.2%	0.0%	2.7%	0.0%	0.0%	14.9%	6.4%
Dayton	1.3%	1.0%	0.0%	0.0%	2.2%	0.0%	0.0%	6.1%	1.2%
Toledo	1.0%	0.1%	0.2%	0.0%	0.7%	0.0%	0.0%	0.5%	1.2%
Cincinnati	0.7%	0.8%	0.3%	0.1%	1.4%	0.0%	0.0%	4.7%	2.6%
Cleveland-Hopkins	0.5%	0.4%	0.4%	0.1%	1.2%	0.0%	0.0%	7.2%	2.8%
Columbus	0.1%	0.2%	0.2%	0.1%	0.0%	0.0%	0.0%	9.6%	0.9%
Sum	5.6%	3.0%	1.3%	0.3%	8.2%	0.0%	0.0%	43.0%	15.1%

6.12b With Rickenbacker case (Compared with Table A6.7c)

	Calibrated							Uncalibrated	
	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Additive	Multiplicative
Airborne Airpark	1.4%	0.1%	0.3%	0.3%	2.0%	0.1%	0.0%	14.2%	6.2%
Dayton	1.4%	1.1%	0.3%	0.1%	2.2%	0.0%	0.0%	6.3%	1.6%
Toledo	1.1%	0.4%	0.5%	0.2%	0.9%	0.1%	0.0%	1.0%	0.9%
Cincinnati	0.5%	0.6%	0.2%	0.0%	1.1%	0.0%	0.0%	3.8%	2.2%
Cleveland-Hopkins	0.4%	0.3%	0.4%	0.1%	1.1%	0.0%	0.0%	6.1%	2.4%
Columbus	0.1%	0.1%	0.2%	0.1%	0.0%	0.0%	0.0%	8.3%	0.7%
Rickenbacker	1.0%	1.1%	1.2%	0.6%	1.2%	0.2%	0.0%	3.2%	1.5%
Sum	5.9%	3.7%	3.1%	1.4%	8.5%	0.4%	0.0%	42.9%	15.5%

The absolute values of the differences from the Table 6.4 data presented in Table 6.12a are small for all Calibration Scenarios. Although this would be expected for the explicit least squares fits of Calibration Scenarios 3, 4, 6, and 7, we were able to find good "trial-and-error" fits in the nonexplicit Calibration Scenarios 1, 2 and 5, where we assumed that we did not have access to the more refined *F(a,r)*. Table 6.12b indicates that the fits remain good in the with Rickenbacker case, even when using the *NFM* model structure to forecast data generated from a different (logit) structure and when using the "trial-and-error" fits. However, unless behavior, and therefore freight flow, follows the assumed structure, this says nothing about the ability to predict flows under altered or future conditions with calibrated models. It does indicate, however, that choosing between the *NFM* or logit form and selecting the approach to data fitting may not be very important. Of course, this result would have to be replicated on a much larger scale and with more realistic data before it could be accepted for practical use.

6.4 Discussion

We presented the New Facility Model to illustrate a type of discrete choice model proposed for freight transportation. Variations of the model could be considered--additive and multiplicative combinations of independent variables; logit formulation versus that proposed in the model. The difficulties presented in Section 1.1 eliminate much of the behavioral justification traditionally proposed for using logit models. Therefore, one would be hard-pressed to justify one formulation over the other on theoretical grounds.

The numerical differences in the results produced when combining the independent variables either additively or multiplicatively were seen to be too great to consider using *NFM* for anything but a rough estimate of freight flows at a facility. Calibrating the additive version of the model and a logit variant of it in several ways led to fairly stable forecasts from the model. However, this stability does not mean that such forecasts would be accurate. Rather, it indicates that once calibrated on a common set of data, the forecasts will tend to mimic each other. In fact, the uncalibrated versions of the model were presented based on previous work by the model developers. We feel that the differences seen when transferring these versions to the problem presented here would be more representative of forecasting performance than the differences produced from different data fits of a single additive formulation. Once again, models of this type could presently be developed and parameters could be estimated, but more analysis would be required to have confidence in the accuracy and usefulness of the forecasts produced.

Section 7. Monitoring and Updating Truck Origin-Destination Tables

Origin-destination (OD) matrices, or trip tables, are at the center of many freight movement applications. Several of the expressed needs for freight movement models in Section 2 could be answered, at least partially, with a freight OD matrix. Moreover, OD matrices are used as inputs when assigning trips to links of a network, another important output of freight movement models. ODOT has recently undertaken a large roadside survey. The type of vehicle, trip origin, and trip destination are included in the collected information, and a truck OD matrix could be estimated from the data.

Although OD matrices are perceived as useful, obtaining the data to estimate these matrices is expensive and time consuming. Much of the information obtained from surveys like that recently conducted by ODOT is collected for purposes other than freight studies. Still, the millions of dollars of costs involved are mostly fixed and could not be avoided by collecting only freight information. As such, an OD matrix represents a substantial financial investment, and inexpensive methods to keep its entries current would be useful.

In this section, we illustrate inexpensive methods of updating or even estimating OD matrices. The methods use observed truck volumes on highway segments, data that are routinely obtained in ODOT's traffic monitoring programs. They also require some truck assignment logic. This logic is comprised of two elements: one representing the choice rule for selecting paths between origins and destinations, and another representing the performance of the paths in terms of the attributes influencing the choice of paths. Determining the relevant attributes of path selection for highway trips may be a subject for future research. We feel, however, that routing trucks on the minimum time path between its origin and destination would be sufficient at this point.

Since the effect of congestion on path times would probably only be of concern in urban areas at selected times of the day, one might not need to incorporate congestion (volume-delay) functions on most link segments. Incorporating these functions is straightforward with most assignment algorithms and software packages, however. Therefore, in this section, we consider the assignment logic to be one of selecting routes between origins and destinations that minimize path travel times, while the travel times respond to passenger-car equivalent volumes through the well-known Bureau of Public Roads (BPR) functions:

$$t_a(X_a) = t_{o,a}\,[1 + \alpha(X_a/C_a)\beta], \qquad (7.1)$$

where $t_a(X_a)$ gives the travel time on segment (arc) a as function of the volume X_a in passenger-car equivalents on this segment, $t_{o,a}$ is the free-flow travel time on segment a, C_a is a parameter based on the capacity of segment a, and α and β are parameters, which may or may not depend on the segment. In this section, we set α and β to their typically used values of 0.15 and 4.0, respectively.

7.1 Updating OD Tables with Observed Link Volumes

Using observed link volumes (ground counts) to update OD tables has been discussed in the urban transportation context. These techniques could be readily modified to update truck OD tables. We discuss two of these methods. We also discuss a technique we found that was developed especially for use in estimating statewide truck OD tables.

Method 1: Sheffi (1985) presents a general method to estimate an OD matrix directly from link flows. The general approach is one of finding an OD matrix that can reproduce the observed conditions when the network is at equilibrium. Specifically when the OD matrix is assigned to the network, the assigned flows should produce assigned OD travel times equal to the observed OD travel times. However, different OD matrices can produce the same set of arc flows and, therefore, the same set of travel times. Therefore, a second step consists of choosing, from all the OD matrices that produce the observed travel times, the matrix that is the closest to a "target matrix." A target matrix is usually obtained from a previous field survey.

The problem is solved by an iterative descent method using partial Lagrangians. In Appendix 7.1, we translate the general description of this method (Sheffi, 1985) into a set of algorithmic steps and illustrate these steps. To compare to methods represented below, we apply the method to a small example. The network used in this example is shown in Figure 7.1. In this figure, the numbers in the ovals identify the links of the network, and the numbers in the boxes identify the nodes of the network. The parameters of the BPR link performance functions and the observed arc flows $\{\tilde{x}_a\}$ are given in Table 7.1. The target OD matrix $\{\tilde{q}_{rs}\}$ is shown in Table 7.2. We use a matrix that has primarily zero entries to make the problem simple enough to solve with a spreadsheet. A computer code could be developed for more realistic problems. Extending the concepts is straightforward. There would simply be more computations, and an available shortest path algorithm would be required to find minimum time paths.

Figure 7.1 Example Network (to illustrate Method 1)

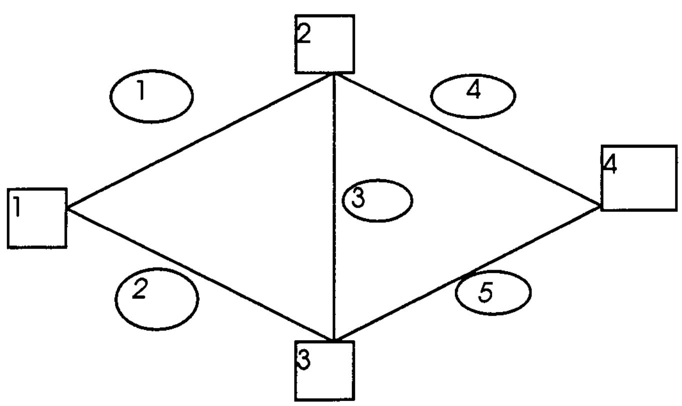


Table 7.1 Parameters for network of Figure 7.1

Link ID, a	BPR function parameters				Observed flow $\tilde{x}_a$
	Free flow travel time $t_{o,a}$	Capacity C_a	α	β	
1	5	4	0.15	4	2.23
2	15	4	0.15	4	2.77
3	10	4	0.15	4	3.10
4	20	6	0.15	4	9.18
5	5	6	0.15	4	13.81

Table 7.2 Target OD matrix for the example

From \ To	1	2	3	4
1	0	0	0	7
2	0	0	0	9
3	0	0	0	6
4	0	0	0	0

The final OD and link flow estimates from this method are presented in Tables 7.3 and 7.4. We also present the percent differences between the estimated (assigned) and the observed minimum OD path times and link flows. This simple example required 16 iterations in Part A (see Appendix A7.1) and an average of 25 iterations for each value of γ in Part B to reach convergence.

Table 7.3 Estimated OD matrix and travel times obtained by applying Method 1 to example problem compared to observed OD times

From\To	1	2	3	4	Estimated OD time	observed time	Difference(%) |Est-Obs|/Obs
1	0	0	0	6.59	41.65	41.57	0.2
2	0	0	0	9.40	36.46	36.49	0.1
3	0	0	0	6.96	25.75	26.06	2.6
4	0	0	0	0	N/A	N/A	N/A

Table 7.4 Estimated link flows obtained by applying Method 1 to example problem compared to observed link flows

Link ID	Estimated flow x_a^*	Observed flow $\tilde{x}_a$	Difference(%) |Est-Obs|/Obs
1	3.41	2.23	52.9
2	3.18	2.77	14.8
3	3.63	3.10	17.1
4	9.18	9.18	0
5	13.76	13.81	3.6

Method 2: Park and Smith (1996) developed a state-wide truck demand model for Wisconsin using a small OD travel survey data and extensive truck classification count data. Like Method 1, the general approach of what we shall call Method 2 is also one of finding an OD matrix that can reproduce the observed conditions when the network is at equilibrium. The target OD matrix in Method 2 is generated by a gravity model. The gravity model uses trip productions and attractions estimated from a small-scale OD survey, zonal populations and trip rates. After assigning the target OD matrix to the network, the estimated trip productions and attractions are adjusted by the ratio between assigned and observed flows on selected links. The new estimated productions and attractions are input to the gravity model to produce a new OD matrix, which is assigned to the network. The ratios of assigned and observed volumes on selected links are used to estimate new productions and attractions, and the procedure continues until the ratios of assigned and observed volumes are close to one, where closeness is defined according to some convergence criterion.

Again we describe the steps of Method 2 and illustrate it with a simple example in Appendix 7.2. We present the results of the same example here. The example can be solved using a spreadsheet. All steps could be programmed, however, and more realistic problems could be readily solved with the method.

The network used in this example is shown in Figure 7.2. Again, the numbers in the ovals identify the links of the network, and the numbers in the boxes identify the nodes of the network. The parameters of the BPR link performance functions and the observed "selected link" flows $\{\tilde{x}_r\}$ are given in Table 7.5. The small-scale OD survey matrix $\{\bar{q}_{rs}\}$ is shown in Table 7.6. Note that the survey matrix $\{\bar{q}_{rs}\}$ contains all study zones except zone 2a, corresponding to an assumption that no survey data is available for zone 2a. The distances between zone pairs are shown in Table 7.7.

The example was solved with a spreadsheet. A computer code could be developed for more realistic problems. Extending the concepts is straightforward. There simply would be more computations in the parameter estimation and traffic assignment tasks.

Table 7.5 Parameters for network of Figure 7.2

	BPR function parameters				Observed flow $\tilde{x}_r$
Link ID,a	Free flow travel time $t_{o,a}$	Capacity C_a	α	β	
1	5	4	0.15	4	54
2	15	4	0.15	4	38
3	10	4	0.15	4	55
4	20	6	0.15	4	/
5	5	6	0.15	4	62
6	13	4	0.15	4	57
7	10	6	0.15	4	42

Figure 7.2 Example Network (to illustrate Method 2)

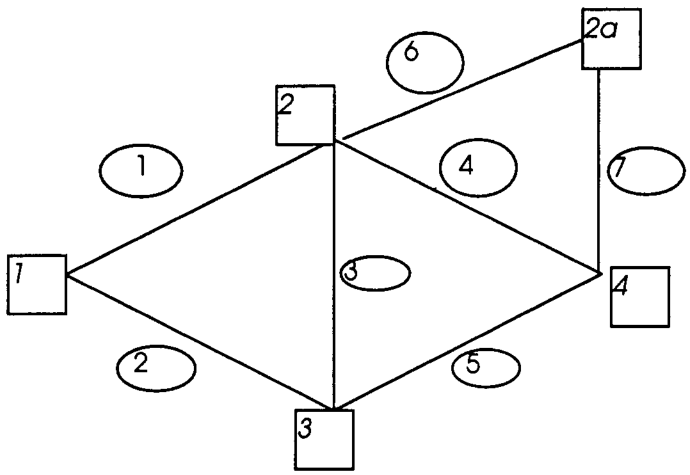


Table 7.6 Results $\{\bar{q}_{rs}\}$ of small-scale OD survey for the example of Method 2

From \ To	1	2	3	4
1	7	18	9	6
2	18	11	17	9
3	9	12	8	8
4	9	10	6	5

Table 7.7 Assumed distances between zone pairs $\{d_{ij}\}$

From \ To	1	2	2a	3	4
1	46	55	142	196	66
2	55	61	77	49	76
2a	142	77	54	89	106
3	196	49	89	55	65
4	66	76	106	65	50

The estimated OD table is shown in Table 7.8, and the estimated link flows are shown in Table 7.9. We also present the percent differences between the estimated (assigned) and the observed minimum link flows in Table 7.9. No differences in link travel times are shown, since we used an all-or-nothing traffic assignment algorithm (see Appendix 7.2), and link travel times were, therefore, always considered to be free flow travel times. In Table 7.9 we see that the percentage differences in link flows are very small (all but one are less than or equal to 0.05%).

Method 2 requires different input than Method 1. Method 1 requires a target matrix with every zone in the estimated target matrix. Method 2 assumes models for trip generation and trip distribution. A small-scale survey is used to calibrate a trip distribution model (a gravity model in this case), in order to produce a calibrated target OD from trip productions and attractions. As presented by Park and Smith (1996), Method 2 also assumes that trip productions and attractions for each zone are based on its zonal population and trip rates. Original estimates could be obtained from historical data. Also independent variables other than population could be used for the trip generation step. When calibrated, of course, Method 2 provides more information than Method 1, especially information on trip generation and distribution models.

Table 7.8 Estimated OD table of the problem illustrating Method 2

From \ To	1	2	2a	3	4	P_i
1	6.1	15.4	6.4	6.3	2.9	37.1
2	12.7	7.7	6.6	8.2	3.0	38.2
2a	19.5	24.5	13.4	14.5	6.1	78.0
3	23.6	37.3	17.9	16.8	8.2	103.8
4	5.2	6.5	3.6	3.9	1.6	20.8
A_j	67.1	91.4	47.9	49.7	21.8	/

Table 7.9 Estimated link flows of the problem illustrating Method 2

Link ID	Estimated link flow	Observed link flow	Difference (%) |Est-Obs| /Obs
1	54.02	54	0.03
2	37.99	38	0.03
3	54.00	55	1.82
4	0.000	N/A	N/A
5	62.03	62	0.05
6	56.99	57	0.02
7	41.99	42	0.02

Method 3: Nielsen (1994) presents a third method for estimating trip matrices. Using a target OD matrix, a network, and a set of traffic counts as input, his Single Path Matrix Estimation Method (SPME) estimates a new trip matrix to minimize the average deviation between counted and assigned traffic along the minimum time path between each zone pair. The estimated trip patterns reflect the route choice patterns given by a specified traffic assignment, and various traffic assignment algorithms can be used. If a traffic dependent assignment, such as user equilibrium is used, iterations of "inner loop" estimations are required for each "outer loop" iteration. In this way the algorithm is similar to Method 1 above, where there are main loops and traffic assignment iterations within each loop. We summarize the algorithm in Appendix 7.3.

Nielsen's algorithm (Nielsen, 1994) is implemented by Caliper (1996) in TransCAD 3.0®. We used TransCAD 3.0® to run the problems used to illustrate Methods 1 and 2 above. The results using the data used to illustrate Method 1 are presented in Tables 7.10 and 7.11. We also present the differences between the estimated and the observed minimum OD path times and link flows. Comparing with Method 1 (see Tables 7.3 and 7.4), Method 3 gives similar, but slightly poorer results. In Method 1, it took approximately 190 iterations (16 iterations in Part A plus an average of 25 iterations for each of the seven values of γ we tried in Part B) to reach convergence. Method 3 took only 50 iterations to converge to its solution. Since each iteration takes similar time for both methods (to run user equilibrium traffic assignment, find the shortest path, etc.), Method 3 converged more rapidly than Method 1 in this example, while yielding similar results.

Table 7.10 Estimated OD matrix and travel times obtained by applying Method 3 to example problem in Method 1(after 50 iterations).

From\To	1	2	3	4	Estimated OD time	Observed time	Difference (%) |Est-Obs| /Obs
1	0	0	0	6.48	39.37	41.57	2.9
2	0	0	0	9.47	34.12	36.49	6.5
3	0	0	0	6.25	23.45	26.06	10.0
4	0	0	0	0	N/A	N/A	N/A

Table 7.11 Estimated link flows obtained by applying Method 3 to example problem in Method 1 (after 50 iterations)

Link ID	Estimated flow x_a^*	Observed flow $\tilde{x}_a$	Difference (%) |Est-Obs| /Obs
1	3.04	2.23	36.3
2	3.44	2.77	24.2
3	3.91	3.10	26.1
4	8.84	9.18	11.3
5	13.36	13.81	3.3

The results using the data used to illustrate Method 2 are presented in Tables 7.12 and 7.13. Again, we show the differences between the estimated and the observed link flows. No differences are shown for link travel times, since only free flow travel times are used in Method 2. Method 3 converged after two iterations and gave a set of estimated link flows that are all within 5.2% of the observed ones. Method 2 yielded link flows closer to the observed flows for this example (see Table A7.2.5).

Table 7.12 Estimated OD matrix and travel times obtained by applying Method 3 to example problem in Method 2 (after 2 iterations).

From \ To	1	2	2a	3	4	Pi
1	6.1	17.4	9.9	9.2	8.9	51.5
2	17.3	15.6	18.9	14.9	11.9	78.6
2a	9.7	19.2	10.1	7.2	13.6	59.8
3	10.7	15.6	7.9	4.7	4.8	43.7
4	8.1	12.1	12.3	4.4	6.9	43.8
Aj	51.9	79.9	59.1	40.4	46.1	/

Table 7.13 Estimated link flows obtained by applying Method 3 to example problem in Method 2 (after 2 iterations)

Link ID	Estimated link flow	Observed link flow	Difference (%) |Est-Obs| /Obs
1	54.27	54	0.5
2	36.72	38	3.4
3	54.49	55	0.9
4	0.000	N/A	N/A
5	65.23	62	5.2
6	57.71	57	1.2
7	40.99	42	2.4

In order to compare the estimated OD matrices of Methods 2 and 3, we use the sum of square difference between the estimated OD matrix and the target matrix, i.e., $\sum_{rs}(\hat{q}_{rs} - q_{rs})^2$. The target matrix in Method 2 is the one calibrated from the gravity model in Table A7.2.3. The smaller this value is, the closer the estimated OD matrix is to the target Matrix. While the results of the link flows are similar, Method 3 yields an OD matrix closer to the target OD matrix. $\sum_{rs}(\hat{q}_{rs} - q_{rs})^2$ =165 for Method 3, while $\sum_{rs}(\hat{q}_{rs} - q_{rs})^2$ = 1629 for Method 2. It should be noted, however, there is no reason to want an updated OD matrix close to the target matrix. The target matrix is simply a seed to allow the estimation of a matrix that will update link flows.

Discussion of Methods: In the examples all three methods produced solutions that reasonably replicated observed link volumes, indicating that they produced good OD matrices, as well. In general, accuracy in the solutions would depend on the accuracy of the model assumptions. All three methods are based on the assumption that truck volumes on highway segments can be modeled as the result of assigning a truck OD matrix to a network with some assumed traffic assignment logic (Sheffi, 1985). Method 1 assumes an "equilibrium" logic, and Method 3 is claimed to handle more general assignment algorithms, including the equilibrium assignment. Method 2 assumes an "all-of nothing" logic, which, unlike the equilibrium logic, ignores congestion effects. However, for intercity truck assignment, ignoring congestion is probably acceptable. Methods 1 and 3 lead to all-or-nothing assignments when flows are low enough that congestion is not an issue. Therefore, we believe that the three methods would be roughly equivalent on the basis of the acceptability of their assignment logic.

In Methods 1 and 3, the major assumptions are those of the traffic assignment logic, the logic that converts an OD matrix to link flows. Method 2 requires the additional assumptions associated with the trip generation and trip distribution models. If the only goal is to develop an OD matrix that replicates the observed truck volumes on highway segments, Methods 1 and 3, with their fewer assumptions, should perform better. Method 2 sacrifices accuracy in the updated OD matrix for the benefit of producing trip generation and distribution models. Of course, if the assumptions leading to the generation and distribution models are unacceptable, then the calibrated models would have no real forecasting power.

All three methods require a target OD matrix and observed truck volumes on highway segments as input. Preliminary analysis of Method 3 (Zhang, 1988) indicates that the outputs are much less sensitive to the target matrix inputs than to the observed link volume inputs. (We conducted the analysis on Method 3 because of software availability.) In their target matrices, Methods 1 and 3 require an estimate of OD flows for the entire matrix but do not place restrictions on how these are to be obtained. In

Method 2 a trip distribution model is estimated. (A traditional gravity model is proposed by the model developers.) This trip distribution model is then used with outputs of a trip generation model to estimate a complete OD matrix. The target OD matrices required in Methods 1 and 3 could be estimated in the same way. That is, none of the methods seems to offer an advantage in terms of required effort for the initial OD target matrix.

Methods 1 and 3 require volume counts on all segments that will be used by the OD flows, whereas the purported advantage of Method 2 is that it does not. However, we note that Method 1 (and, seemingly, Method 3) is guaranteed to estimate an OD matrix that approximates the observed volume counts when processed through the assignment algorithm assumed. As noted above, the assignment logic of Method 2 can be handled by Method 1. Therefore, if the volumes on links not considered when using Method 2 would impact the estimated OD matrix produced by Method 1, not considering these volumes in Method 2 would affect its matrix unless the effect is somehow compensated. The compensation comes from the trip generation and distribution models that lead to the estimated volumes on the unobserved links. That is, the fewer number of link observations come at the expense of requiring the additional trip generation and distribution models. The choice, then, is one between estimating link volumes from data or models. Since link volumes are routinely collected and available, the marginal cost of data collection is low, and the choice reduces to one based on perceived accuracy between volumes estimated from models or from collected data.

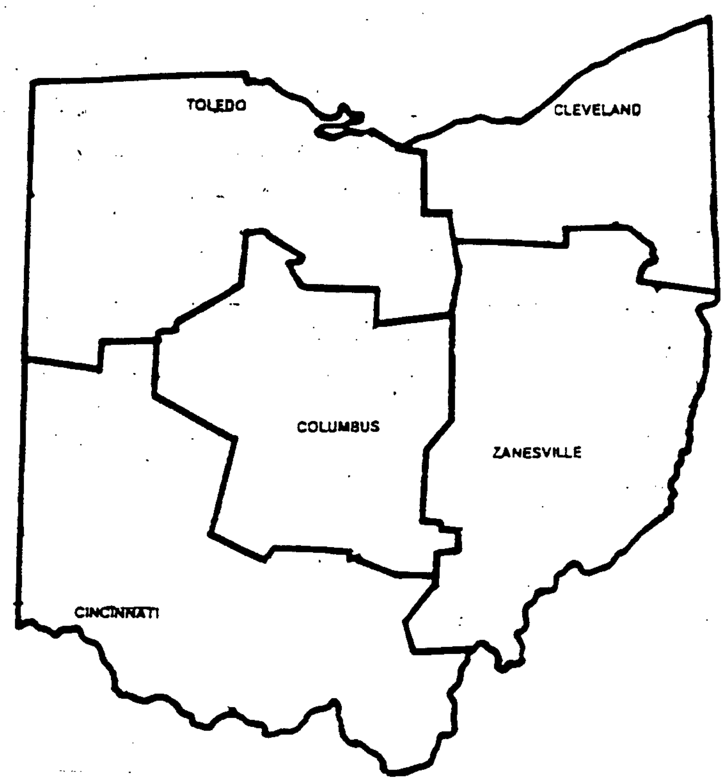
In the examples all three methods converged to solutions that closely approximated link volumes. Method 1 is special case of a more general mathematical programming problem which can be guaranteed to converge. Although they did not claim guaranteed convergence, Park and Smith (1996) found that Method 2 converged after three iterations for the Wisconsin network. (We required many more iterations in our simple example, since we used very strict convergence criteria for all methods. These criteria would be unrealistic to use for planning level applications.) Nielsen (1994) claims that Method 3 can be proven to converge, and the method has been claimed to work well in practice (Caliper, 1996).

7.2 Applying Method 3 to More Realistic Data

Method 3 is available in TransCAD 3.0®. We apply it to a more realistic network using two different origin-destination (OD) target matrices and the same set of observed link counts.

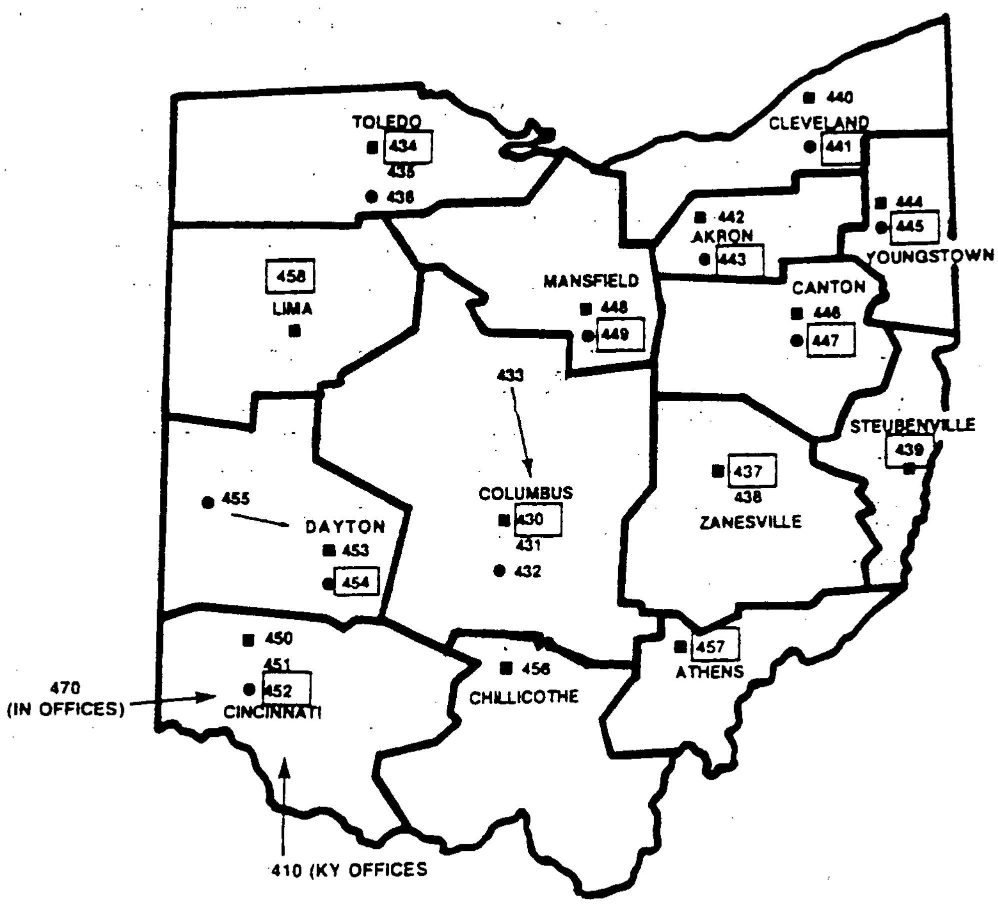
Origin-destination matrix: We used a subset of the data collected in ODOT's roadside survey to form the target matrices. We did not wish to commit resources to developing an accurate table, since our purpose is one of illustration. We developed a highly aggregated truck OD matrix representing five zones in Ohio (Figure 7.3), sampled the data sparingly, and did not explicitly account for external trips (trips with origin or destination outside Ohio) found in the data.

Figure 7.3 Illustration of five zones used in aggregated truck OD matrix



We obtained the raw data file from ODOT for six survey areas (Akron, Columbus, Dayton, Springfield, Toledo, and Newark) and extracted the truck data from the files. We identified the truck data based on the vehicle type field. After extracting the truck data, we sampled 10% of the truck trip entries. Because we were only interested in illustrating the OD updating methods and not in estimating an OD matrix, we did not sample randomly, but rather chose the first 10% of the entries in each file for which both the origin and destination zip codes were available from the data. Based on the origin and destination zip codes, we summarized each trip into a 14 x 14 trip table using the 3-digit zip code map defined by the U.S. Postal Service (1994). We show a portion of this map in Figure 7.4.

Figure 7.4 Portion of 3-digit zip code map (U.S. Postal Service, 1994) used for OD matrix estimation



We aggregated the 14 zones into the five regions of Figure 7.3, where Cincinnati, Columbus, Cleveland, Toledo, and Zanesville served as the major cities in the regions. We scaled up the entries by a factor of 10 to compensate for the 10% sample taken and again scaled up by a factor of 1.5 to produce the 24-hour matrix shown in Table 7.14. (The 1.5 factor was somewhat arbitrarily chosen in an attempt to account for the fact that the samples were taken in consecutive 12-hr periods expected to contain more than half the daily volumes.)

Table 7.14 Initial 5 x 5 24-hour truck origin-destination matrix used in analysis (no external trips included)

	Toledo	Cleveland	Cincinnati	Zanesville	Columbus
Toledo	3,602	105	225	105	195
Cleveland	75	3,563	0	75	950
Cincinnati	90	45	2,393	969	45
Zanesville	30	93	228	849	270
Columbus	90	210	60	660	4,274

The matrix in Table 7.14 table does not contain any "external" trips, trips with either an origin or destination outside Ohio. Since there is a large amount of external truck traffic traveling in Ohio, assigning Table 7.14 to the highway network would grossly underestimate the truck volumes on highway segments. Therefore, we believe that this matrix would be far from one that would be compatible with observed link counts.

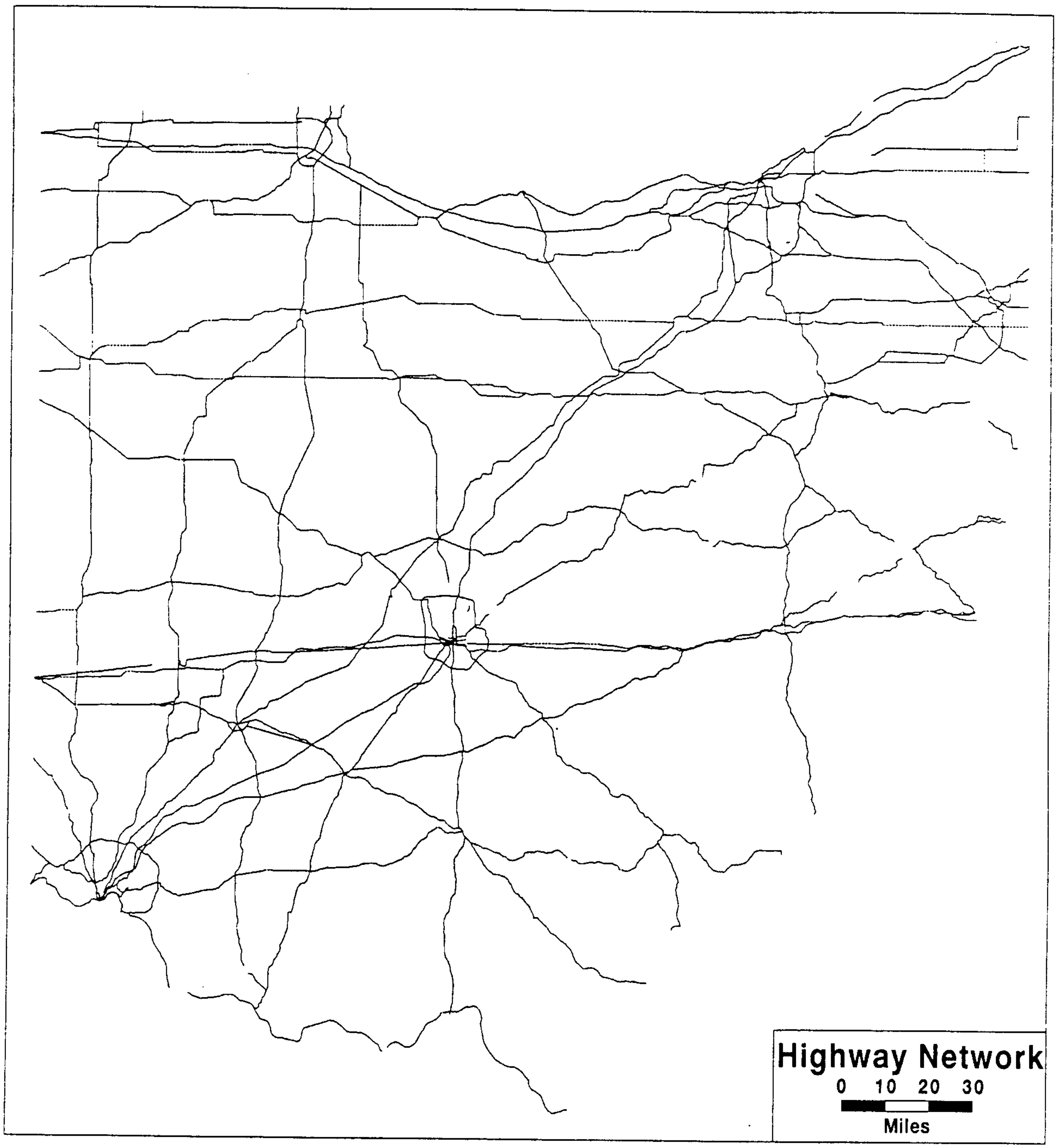
We created a second OD matrix to reflect external volumes more accurately, but in an admittedly artificial manner. Specifically, for each OD pair we found the path of minimum length between the origin and destination and used ODOT's state ADT map to identify the smallest truck count on any link of this path. This count was used as the number of truck trips made between this OD pair. For example, the shortest length path between Toledo and Columbus contains four segments, U20, U23, U36 and I71. The truck counts on them are 3500, 3170, 2500 and 7618, respectively. Since 2500 is the smallest among these four, we take it as the number of truck trips traveling between these two cities per day (1250 for each direction). The matrix developed in this way is presented in Table 7.15.

Table 7.15 Truck origin-destination matrix formed directly from link counts

	Toledo	Cleveland	Cincinnati	Zanesville	Columbus
Toledo	----	4,000	4,000	796	1,250
Cleveland	4,000	----	370	2,903	3,809
Cincinnati	4,000	370	----	1,152	3,809
Zanesville	796	2,903	1,152	----	4,000
Columbus	1,250	3,809	3,809	4,000	----

Network and link counts: We used the National Transportation Atlas Database (NTAD) obtained from the Bureau of Transportation Statistics to develop a computerized statewide highway network. For illustrative purposes, we only wanted a sparse network. Therefore, we extracted Ohio highway data from the NTAD, eliminated all links which were not designated as Interstate or US routes in the database, and converted the data into a geography file. The resulting geography file could be used by a standard transportation GIS software. We used TransCAD®, and we show the resulting network in Figure 7.5. The NTAD database also provided link attributes, such as number of lanes and length, which allowed us to determine link performance functions.

Figure 7.5 Highway network used for assigning truck trips



We obtained daily passenger car and truck count maps from ODOT's Bureau of Technical Services. Again, since we were only developing data for illustrative purposes, we did not wish to commit resources to entering the detailed volumes on each link of the Figure 7.5 network. Rather we identified those links with flows after running a user equilibrium traffic assignment using the truck OD matrix of Table 7.14 on the Figure 7.5 network. We entered the passenger car and truck volumes from the ODOT maps on these links. When we referred to the ODOT count map, we took the median of the flow range for one route within a county as the truck count for every link of this route. For example, the segment of U23 in Delaware county has a count range of 1500 to 5000 truck per day. Therefore, we considered all links of U23 in Delaware county (ID 85667, 85668, 85669) to have 3250 (i.e., the midpoint in the range) trucks per day. We did the same thing for passenger car counts.

We entered the passenger car volumes by link as a field in our database and the truck volumes as a second field. The passenger car volumes would be used as "preloaded" volumes below, and the truck volumes would be used as "observed" truck volumes below. We estimated free flow travel time on the arcs -- $t_{o,a}$ from Equation (7.1) -- based on distance and functional classification of the highway segment. In an *ad hoc* attempt to handle some peaking of the daily truck trips to be assigned, we estimated capacities C_a of the Equation (7.1) to correspond to a 10-hour capacity of the segment in terms of passenger car equivalents.

To be compatible with these functions, we converted the truck trip table to passenger equivalents by multiplying by a passenger car equivalent of 1.5. In practice, different passenger-car equivalents for different highway segments could be incorporated by dividing the segment-specific C_a parameter by the corresponding passenger-car equivalent for trucks on that segment. We used the passenger car segment volumes obtained from ODOT's Bureau of Technical Services and discussed above as "preloaded" volumes on the network. Although the software we used contains an option for assignment with preloaded traffic, other assignment codes could be easily modified to allow such assignment.

We assigned the truck trip tables of Tables 7.14 and 7.15 to this preloaded network and compared the results to the "observed" truck trips obtained from ODOT's Bureau of Technical Services discussed above. We used a user equilibrium traffic assignment algorithm (Sheffi, 1985), although we would expect most algorithms to lead to very similar results because of the minor congestion effect anticipated. After the assignment, each trip table produces a set of "assigned" link flows. In Figure 7.6a and 7.6b, we plot the distribution of the absolute values of difference between the "assigned" and the "observed" flows on each link, using the truck OD matrices from Table 7.14 and 7.15, respectively. Table 7.16 provides summary statistics of the results.

Figure 7.6a Distribution across links of absolute differences between "assigned" and "observed" flows using truck OD matrix of Table 7.14

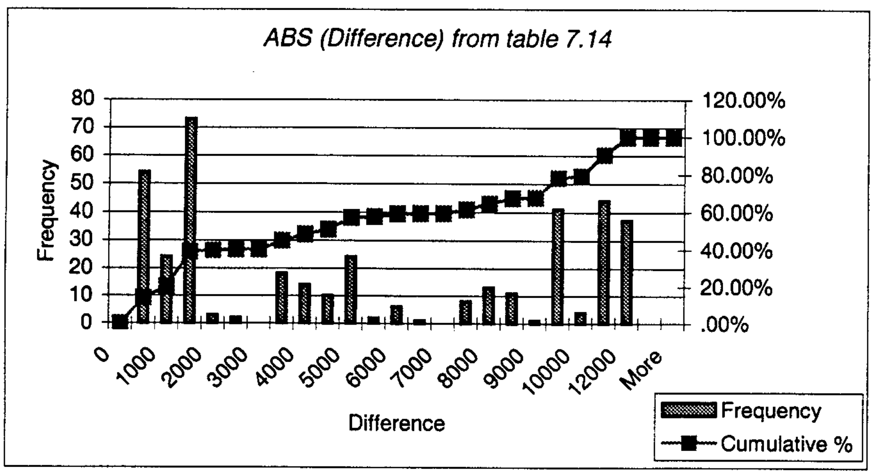


Figure 7.6b Distribution across links of absolute differences between "assigned" and "observed" flows using truck OD matrix of Table 7.15

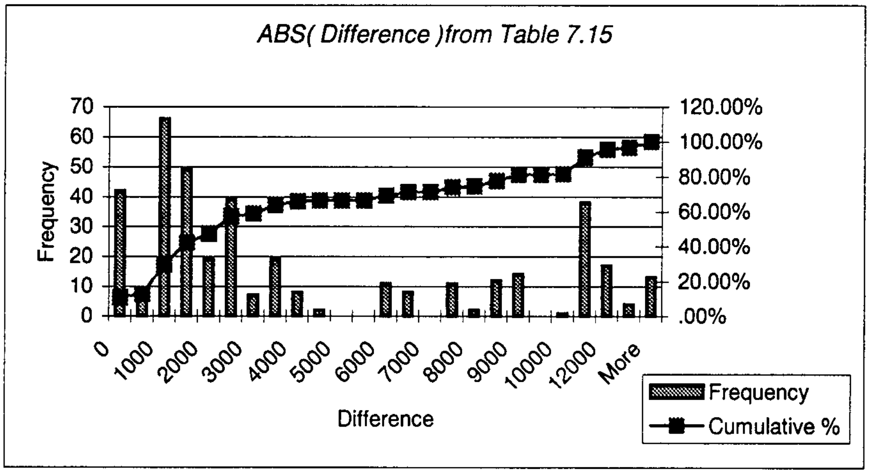


Table 7.16 Summary statistics of differences between “assigned” and “observed” flows using OD matrices from Tables 7.14 and 7.15 (390 observation)

ABS (Difference, from Table 7.14)		ABS (Difference, from Table 7.15)	
Mean	5162.6	Mean	4362.0
Median	4447.5	Median	2250.0
Standard Deviation	4188.3	Standard Deviation	4869.7
Skewness	0.2	Skewness	1.5
Minimum	97.5	Minimum	0
Maximum	11748	Maximum	24162

We see that the median difference in Figure 7.6a is 4500 vehicles, while it is only 2300 in Figure 7.6b. Moreover, 30% of the differences are greater than 9000 vehicles in Figure 7.6a, while fewer than 20% of the differences are greater than 9000 in Figure 7.6b. Table 7.16 also indicates that the differences between the “observed” and “assigned” flows are markedly greater when using the OD matrix of Table 7.14 than when using that of Table 7.15. That is, the absolute value of differences between the “assigned” and “observed” flows reflects the belief that Table 7.15 is a better OD matrix than Table 7.14.

Using the inferior Table 7.14 as the target matrix with Method 3 and the observed link counts produces Table 7.17 as the estimated OD matrix. Using Table 7.15, which we believed to be a much better matrix than that of Table 7.14 as the target matrix, produces Table7.18 as the estimated OD matrix.

Table 7.17 Method 3 estimation of new trip matrix using Table 7.14 as target matrix

	Toledo	Cleveland	Cincinnati	Zanesville	Columbus
Toledo	3602	1916	542	876	911
Cleveland	1916	3563	0	1215	3513
Cincinnati	507	134	2393	353	97
Zanesville	857	1249	353	849	473
Columbus	861	3456	149	473	4274

Table 7.18 Method 3 estimation of new trip matrix using Table 7.15 as target matrix

	Toledo	Cleveland	Cincinnati	Zanesville	Columbus
Toledo	0	1916	449	717	833
Cleveland	1916	0	65	1383	3277
Cincinnati	449	65	0	313	267
Zanesville	717	1383	313	0	499
Columbus	833	3277	267	499	0

We assigned the truck trip tables of Tables 7.17 and 7.18 to the preloaded network in Figure 7.5 and compared the results to the "observed" truck trips obtained from ODOT's Bureau of Technical Services discussed above. We used a user equilibrium traffic assignment, although we would except most algorithms to lead to very similar results because of the minor congestion effect anticipated. After the assignment, each trip table produces a set of "assigned" link flows. In Figure 7.7a and 7.7b, we plot the distribution of the absolute values of difference between the "assigned" and the "observed" flows on each link, using the truck OD matrices from Table 7.17 and 7.18, respectively. Table 7.19 provides summary statistics of the results.

Figure 7.7a Distribution across links of absolute differences between "assigned" and "observed" flows using truck OD matrix of Table 7.17

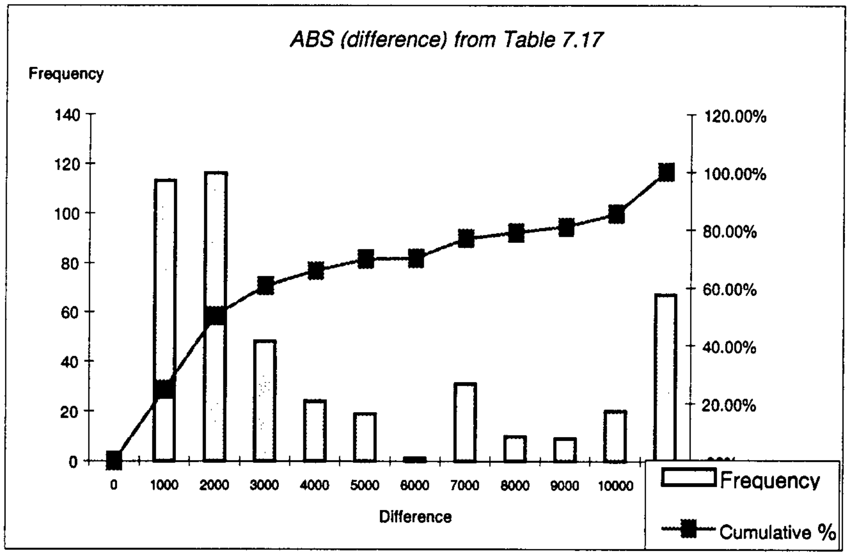


Figure 7.7b Distribution across links of absolute differences between "assigned" and "observed" flows using truck OD matrix of Table 7.18

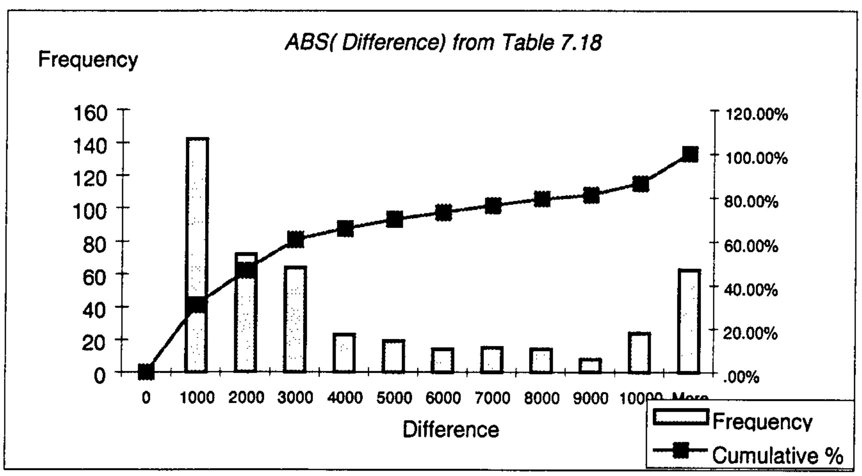


Table 7.19 Summary statistics of differences between "assigned" and "observed" flows using OD matrices from Tables 7.17 and 7.18 (390 observation)

ABS (Difference, from Table 7.17)		ABS (Difference, from Table 7.18)	
Mean	3351.1	Mean	3349.6
Median	1546.2	Median	2167.4
Standard Deviation	3513.1	Standard Deviation	3517.8
Skewness	1.1	Skewness	1.1
Minimum	73.1	Minimum	3.6
Maximum	13111.9	Maximum	12330.2

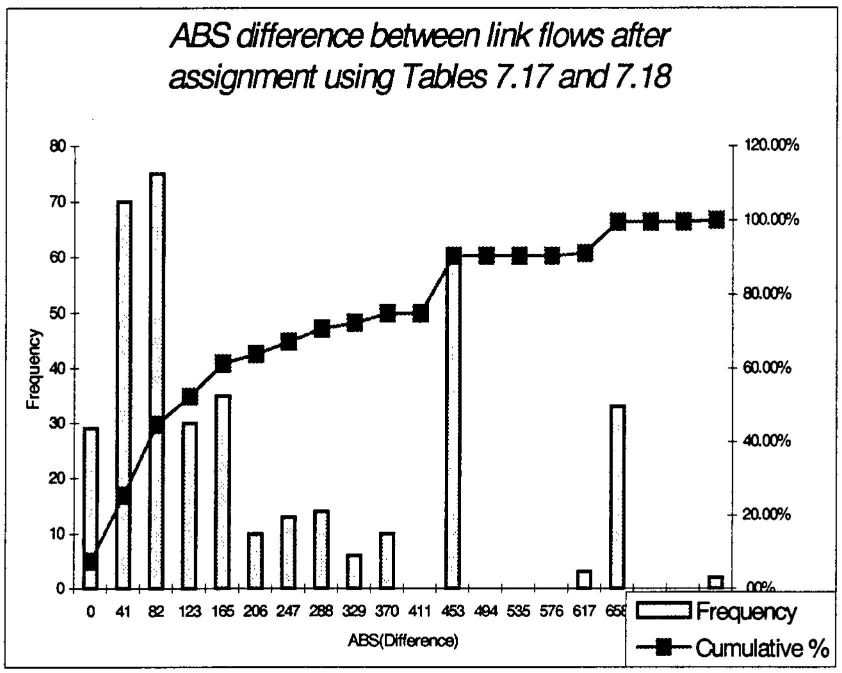
Comparing the distributions in Figure 7.7 and statistics in Table 7.19, we see that there is little difference in the distribution of assigned link flows, regardless of whether the OD matrix of Table 7.17 or that of Table 7.18 was used, even though the target matrix leading to Table 7.17 (Table 7.14). We also compared the assigned link flows between Tables 7.17 and 7.18. The results are shown in Figure 7.8 and Table 7.20. The mean of the absolute difference is around 200, which is much smaller than the means in Table 7.19. Although this is just one example, it indicates that Method 3 can produce OD matrices that lead to assigned link volumes which are fairly insensitive to the target matrix used.

Comparing the statistics in Table 7.16 and 7.19, we also observed marked decreases in means and standard deviations after updating the OD matrix with Method 3. For results from Table 7.17 and 7.18, the means decreased by 34% and 23%, respectively. The standard deviations decreased by 16% and 28%. That is, Method 3 not only seems to produce results that are insensitive to the target matrix, but the OD matrices assign link volumes that match observed link counts much better than the original target matrices do.

Table 7.20 Summary statistics of differences between "assigned" flows using OD matrices from Tables 7.17 and 7.18 (390 observation)

ABS (Difference, from Tables 7.17 and 7.18)	
Mean	204.5
Median	119.8
Standard Deviation	206.2
Skewness	0.9
Minimum	0.0
Maximum	781.7

Figure 7.8 Distribution across links of absolute differences between "assigned" flows using truck OD matrices of Tables 7.17 and 7.18



7.3 Discussion

We presented different methods of updating truck origin-destination (OD) matrices from observed truck volumes. The methods are based on determining an OD matrix that, when processed through a traffic assignment logic, replicates the observations. All the methods performed well on the simple examples examined. One of the methods is included in commercial software presently available and has been claimed to work well in practice. This same method has been used by the Indiana Department of Transportation in developing its statewide model (Smith, 1998).

ODOT has recently conducted a large roadside survey that contains data elements from which a statewide truck OD matrix could be estimated. Since truck volumes are routinely collected, using any of these procedures would be inexpensive to implement on a systematic basis. Therefore, it would be feasible and relatively inexpensive for ODOT to maintain a current estimate of a statewide OD matrix. Simply determining an origin-destination matrix that is consistent with observed traffic flows does not directly help forecast future conditions. However, such a matrix could be indirectly used for forecasting. An accurate estimate of a truck OD matrix could be used to calibrate or validate components of a larger-scale model that forecasts a truck OD matrix. Moreover, accurate estimates of present OD patterns could assist experts when forecasting future patterns.

Section 8. Intermodal Network

Assignment of freight onto networks is at the heart of many freight movement concerns. We discussed assigning truck traffic to a highway network in Section 7. Doing so required a truck origin-destination (OD) matrix. The truck OD pattern will be a function of how freight is distributed across the network encompassing complementary and competitive modes to transport the freight. There would be similar interest in assigning rail-, air-, and water-borne networks. Moreover, we saw in Section 2 that being able to analyze the potential to have freight switch modes in Ohio was expressly stated as a desire for a freight movement model.

Therefore, it would eventually be desirable for ODOT to possess an operational model capable of assigning freight onto an intermodal network. Network assignment models require a mathematical (computer coded) network, an OD table, and an assignment logic with parameters of the network-based impedance functions compatible with this logic. From Part I, it appears that some states are pursuing the development of commodity flow OD tables, and such tables could be developed for Ohio. Much of the freight transported in Ohio would be external freight, having at least one of its trip ends exterior to the state. Developing anything but a crude freight OD table useful in determining what type of exterior freight could be transported in Ohio would probably require a federal effort or, at least, parallel efforts in many states. Given the recent interest in freight issues, it is not unreasonable to imagine such efforts in the future. However, the quality of such products would need to be assessed before one could use them confidently.

Developing an acceptable intermodal assignment logic to distribute tons of commodities among modes and routes would appear to be more difficult. In reviewing models, we did not see anything that would appear to be operational and accurate enough for DOT use in the near term. Nevertheless, the wide dissemination of the Bureau of Transportation Statistics's (BTS) National Transportation Atlas Database (NTAD) (1997) and the increasing federal and state interest in freight movement models indicate a potential that some common efforts in intermodal assignment may be forthcoming in the future. In this section, we demonstrate that available software can be used with the NTAD to form a network for intermodal assignments with only minor modification. We do this by using standard software with a slightly modified version of the NTAD to show that the minimum impedance routes between a specified origin and destination, which are at the heart of network assignments, can encompass a different mode or a combination of modes, depending on the impedance of intermodal transfer points or the impedance characteristics of the modes.

The NTAD contains spatial information on transportation facilities, networks, intermodal terminals, and related attributes. Points, lines, and polygons are used to define the spatial features in the NTAD database. A point is given by an (x, y)-coordinate pair. Some examples of point data are airports, truck terminals, water ports, and highway/rail transfer terminals. Lines are series of connected points that represent the transportation networks. Highway, rail, and waterway links are examples of line data. Polygons are closed areas formed by joining lines. Polygons define specific regions such as a counties, states, National Transportation Analysis Regions (NTAR), or congressional districts. Some of the networks (e.g., highway and rail) also include related attributes for characterizing the nodes and links (e.g., length, speed limit, number of lanes, ...) of different networks. The most recent database that we obtained terminals (NTAD, Bureau of Transportation Statistics, 1997) does not provide explicit connections between modes and, and BTS will provide more information and accessibility on the intermodal connectivity and passenger and commodity flow information in future releases. However, the database was sufficient to conduct the following experiments.

We used the NTAD highway and rail networks in Ohio to demonstrate that minimum impedance paths could use database links depicting these two modes. Extensions to other modes would be similar. We used the TransCAD® software package to build the mathematical network, perform shortest path operations, and display results. To do this, we:

1) extracted the Ohio highway and rail data from the NTAD;
2) pasted the rail file into the highway file in Microsoft® Notepad;
3) converted this joint file to a geography file;
4) opened this geography file in TransCAD®;
5) used the list of intermodal points in the NTAD to identify the latitudes and longitudes of selected intermodal locations in Ohio;
6) found one highway and one rail node near each latitude-longitude pair identified in Step 5);
7) added an "intermodal" link between the highway and rail node pairs identified in Step 6);
8) opened the file in TransCAD®, which formed the "connected" intermodal network.

We show the original *NTAD* highway and rail networks in Figure 8.1, and the connected intermodal network in Figure 8.2. The circles on the intermodal network of Figure 8.2 represent the locations where we added intermodal links in the following analyses. These locations were found in the NTAD "Tofccofc.geo" file. We zoom in on the intermodal location in Columbus representing the Buckeye Yard in Figure 8.3.

Figure 8.1 Original *NTAD* Ohio rail and highway networks

Figure 8.1a Rail network

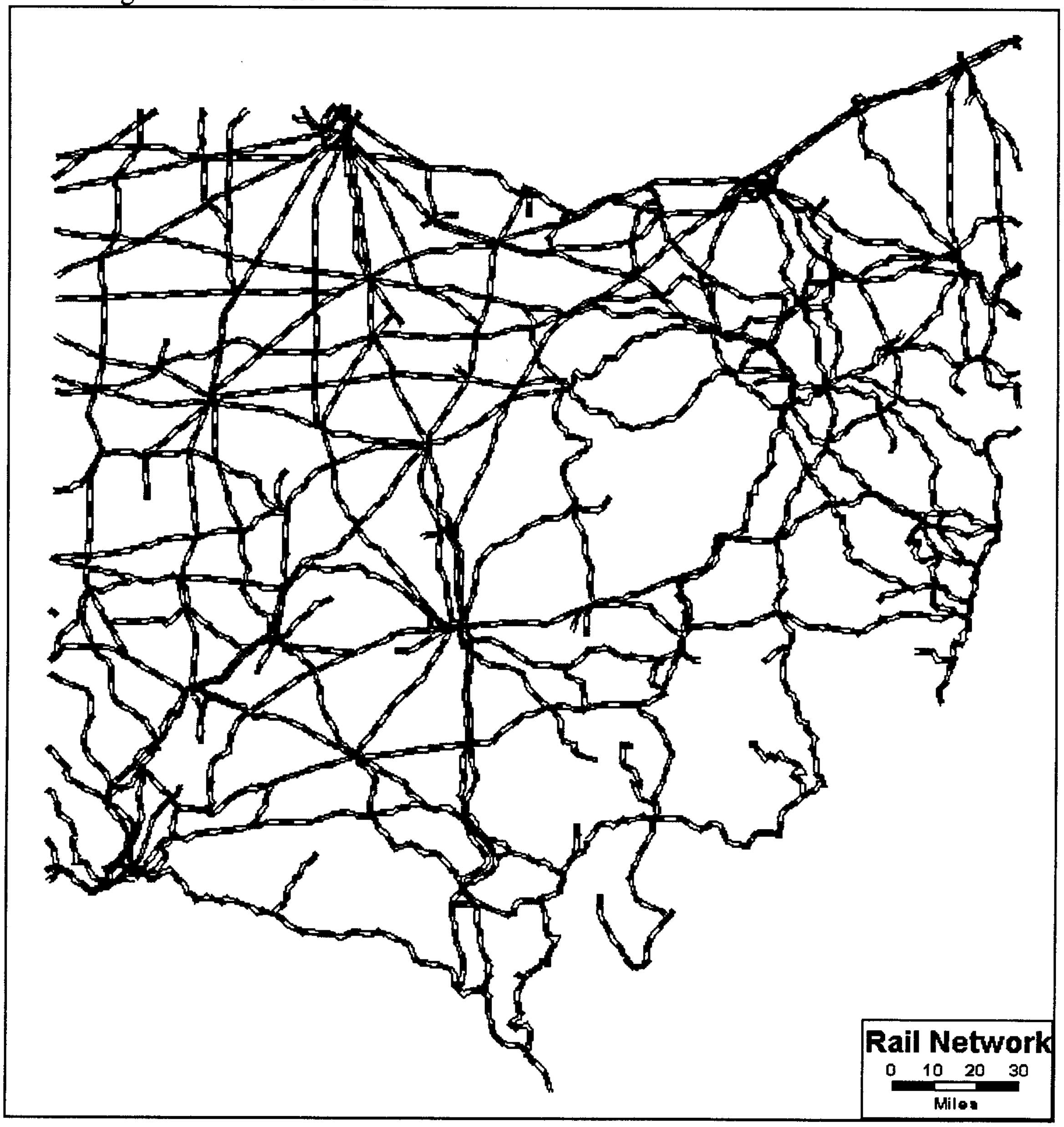


Figure 8.1b Highway network

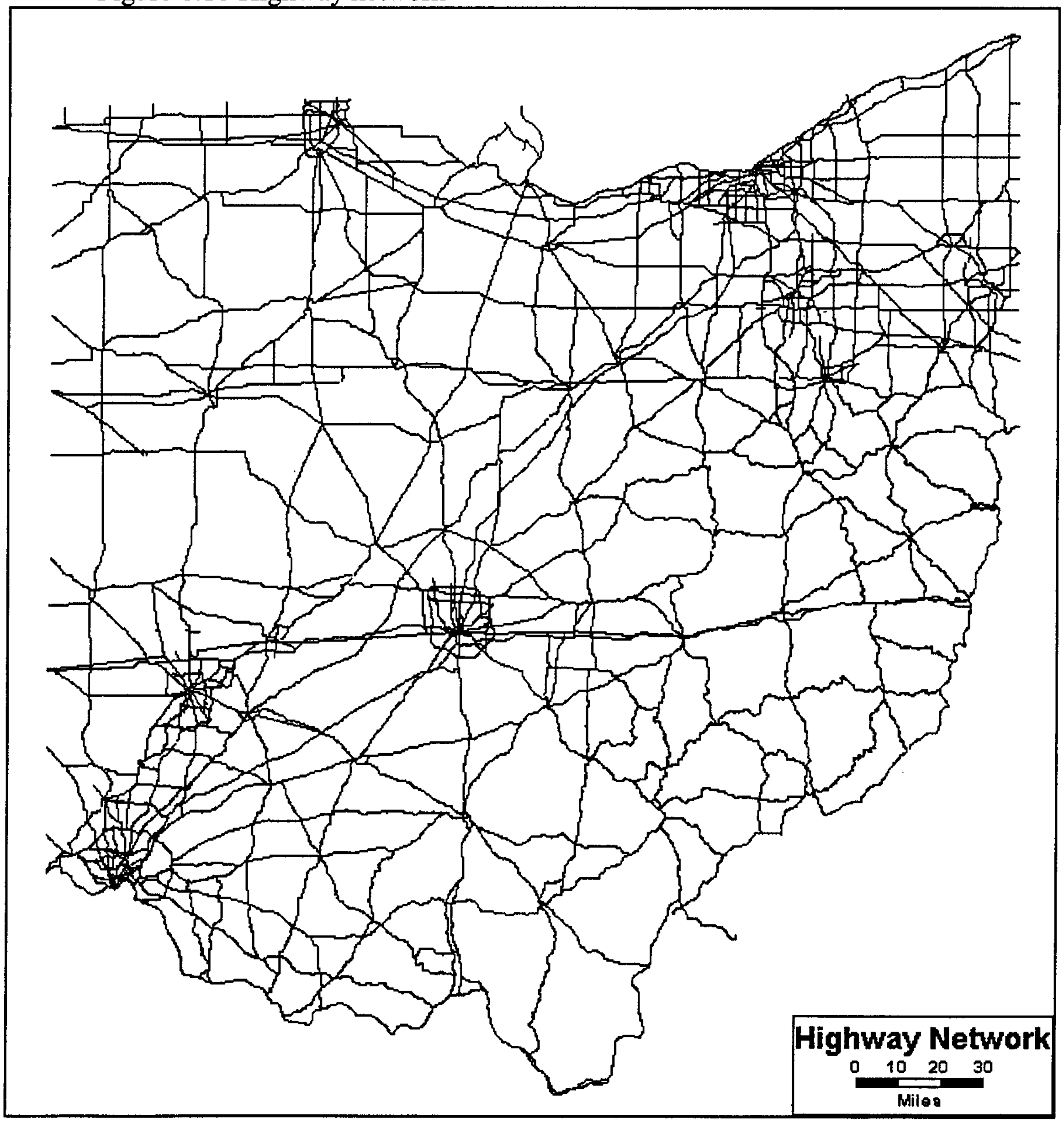


Figure 8.2 Ohio intermodal network formed by connecting the Figure 8.1 rail and highway networks at intermodal locations represented by symbols

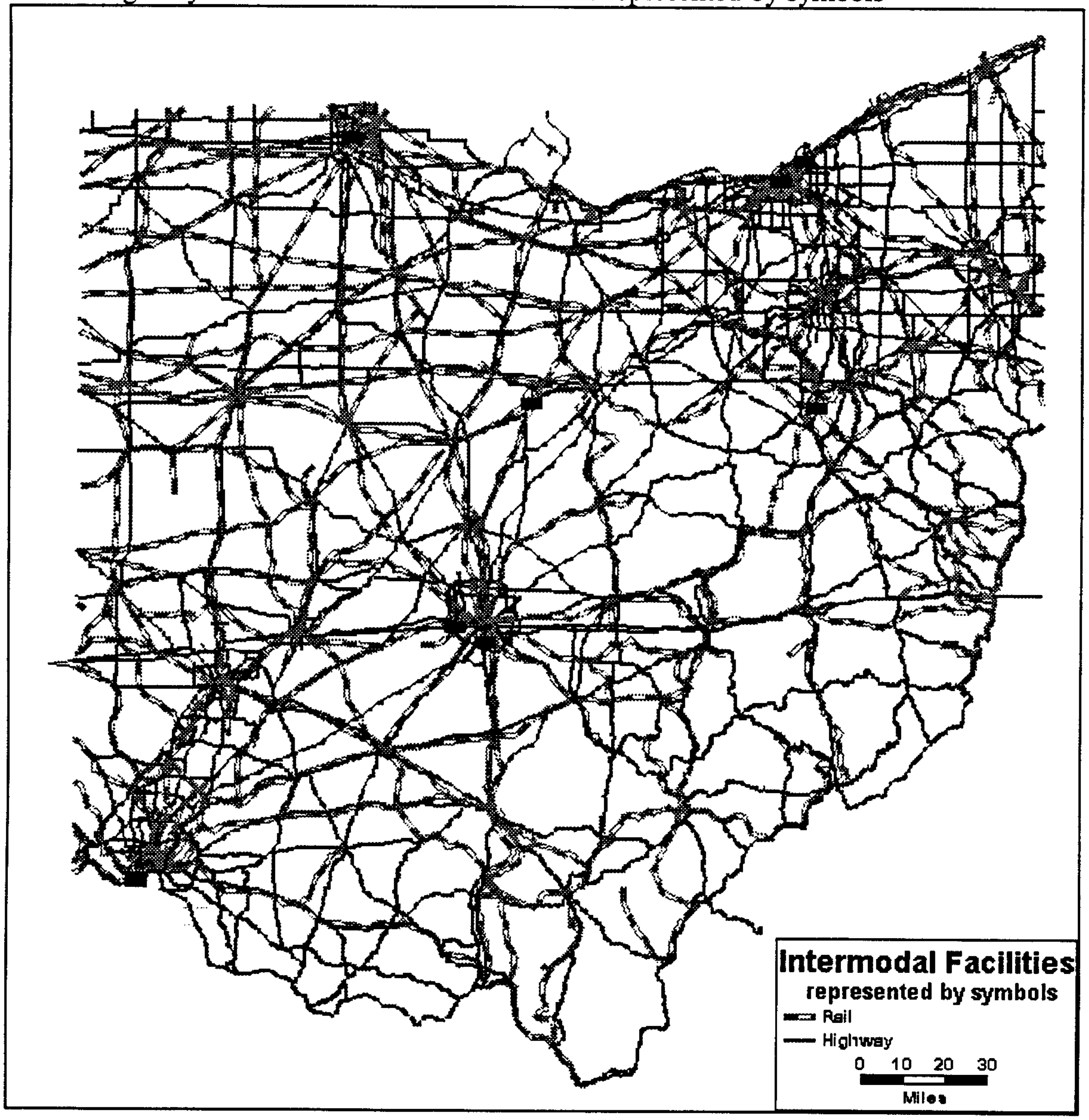
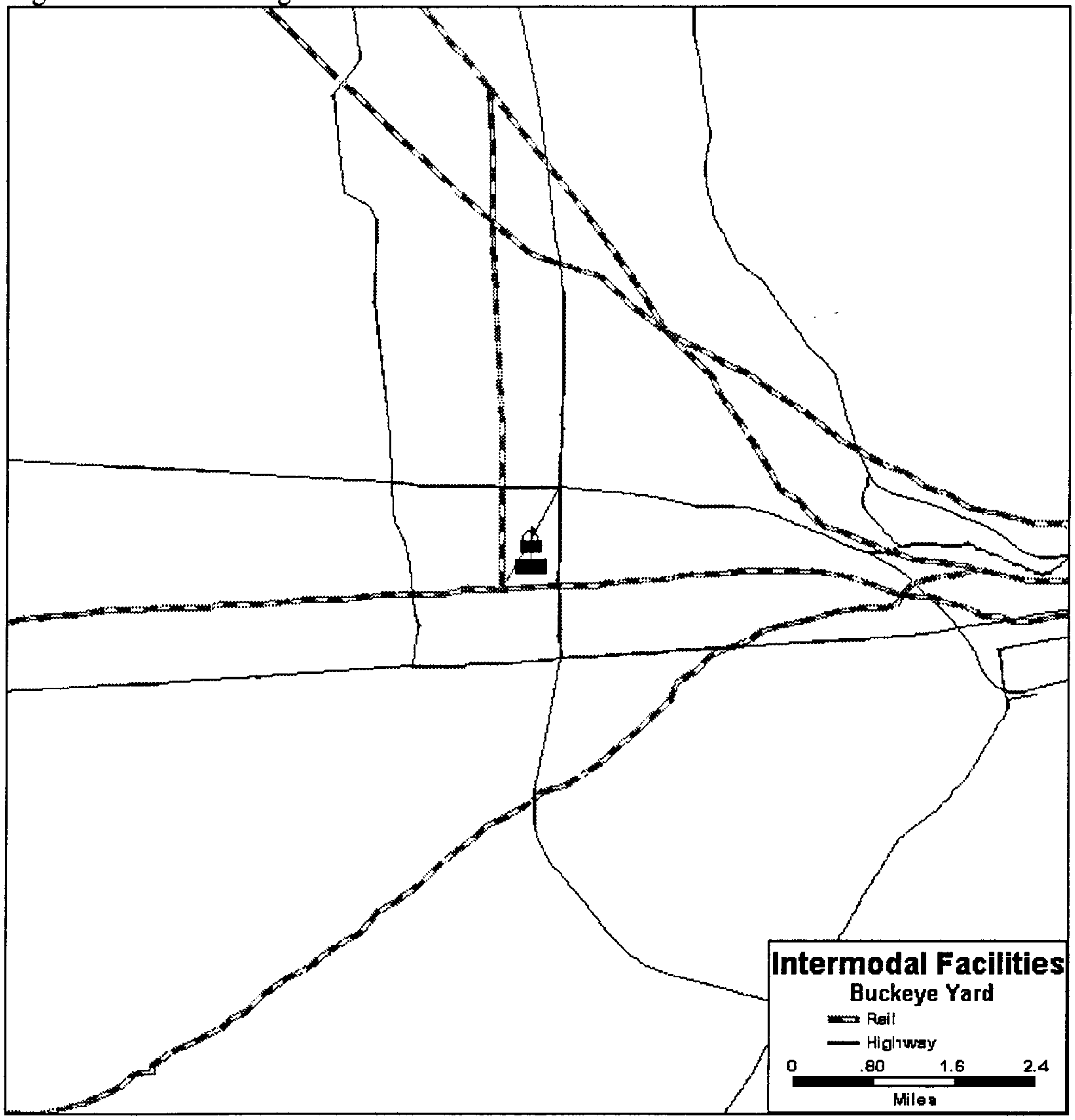
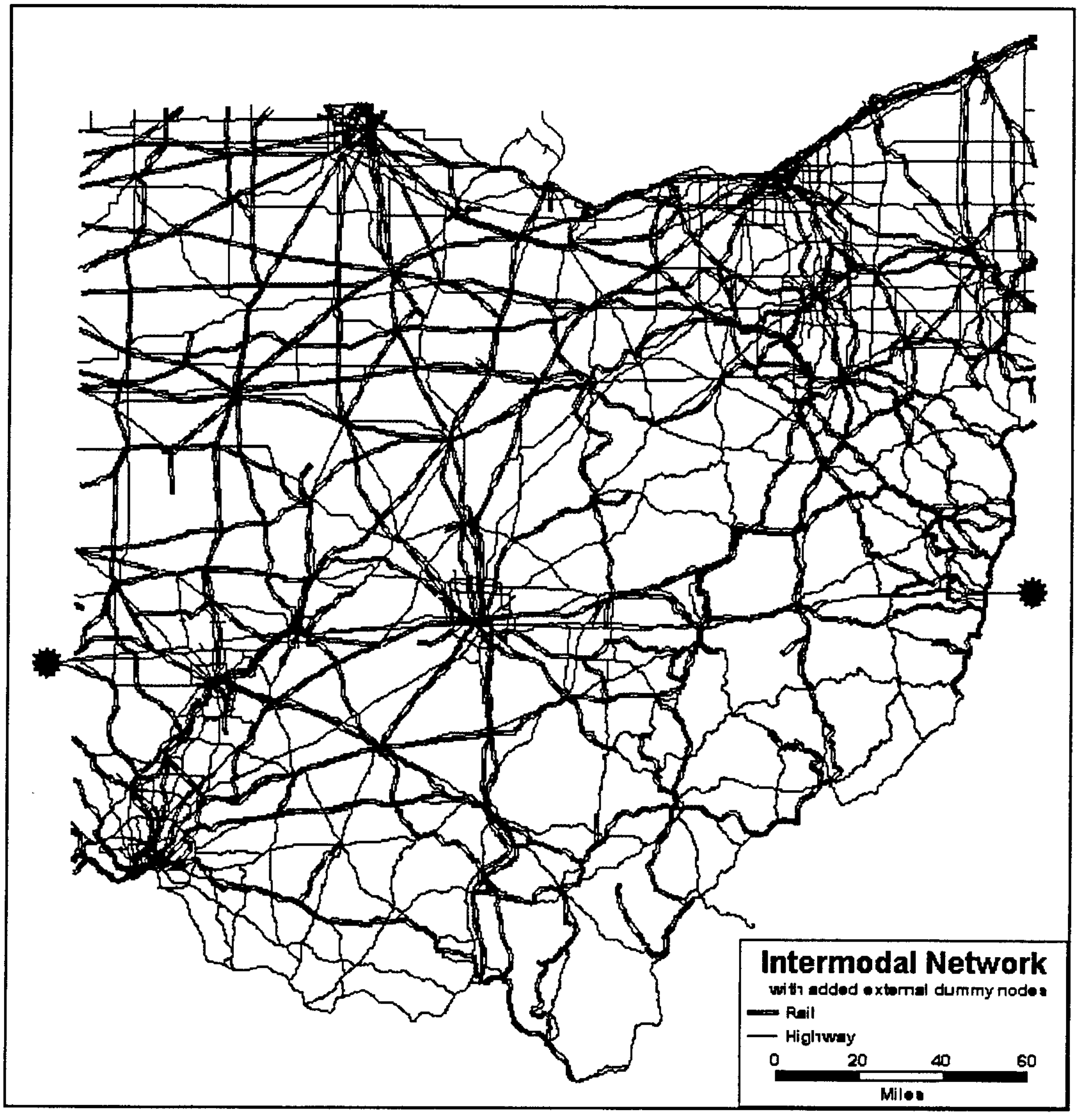


Figure 8.3 Zoom of Figure 8.2 intermodal network at Columbus intermodal location



We considered two cases of intermodal movements over our connected rail-highway network. The first case was intended to represent that a highway shipment crossing Ohio could switch to rail in Ohio by decreasing the impedance of an intermodal link. The second case was intended to represent modal choice for freight in Ohio according to impedance of the modes. We used the link distance as the impedance for both demonstrations. As mentioned above, determining appropriate measures of freight impedance is an open issue, but many measures would be related to distance, which we use as a first-cut measure of route impedance.

Figure 8.4 Ohio rail-highway network with added external dummy nodes and connector links



To illustrate the case in which exterior highway freight would switch to rail in Ohio, we added two dummy nodes, one to the east of the Ohio network and one to the west at approximately the latitude where I-70 enters and leaves the state. We connected each node to the intermodal network with two dummy links, one connecting the node to the nearest interstate highway link and one connecting it to the nearest rail link. This network is shown in Figure 8.4.

Using the lengths provided in the NTAD database as the impedance, we found the minimum impedance route from the western to the eastern node by running the shortest path routine in the TransCAD® software. This route, shown in Figure 8.5a, followed all highway links. We then decreased the impedance on the intermodal link shown in Figure 8.3 until we saw that the minimum impedance route, found in the same way, followed highway links up to the intermodal connector link, where it switched to a rail link, and followed rail links out of Ohio. In Figure 8.5b we show this intermodal route, and in Figure 8.5c a zoom of this route around the intermodal facility.

We needed to add a negative length to the intermodal link to force the shipment to change modes. Freight would not be expected to follow routes according to length only, however, and one should not draw conclusions from this unrealistic routing example. Still, in addition to indicating the potential of the NTAD database to serve as the underlying intermodal network for intermodal assignment, this would indicate the types of results that could be expected: If it truly did require an impossible negative impedance at an intermodal facility to fulfill an objective of getting a shipment off the highway and onto rail, then one could conclude that added investments in the intermodal facility would not be warranted.

Figure 8.5a Minimum impedance route from point west of Ohio to point east of Ohio
(Route follows all highway links)

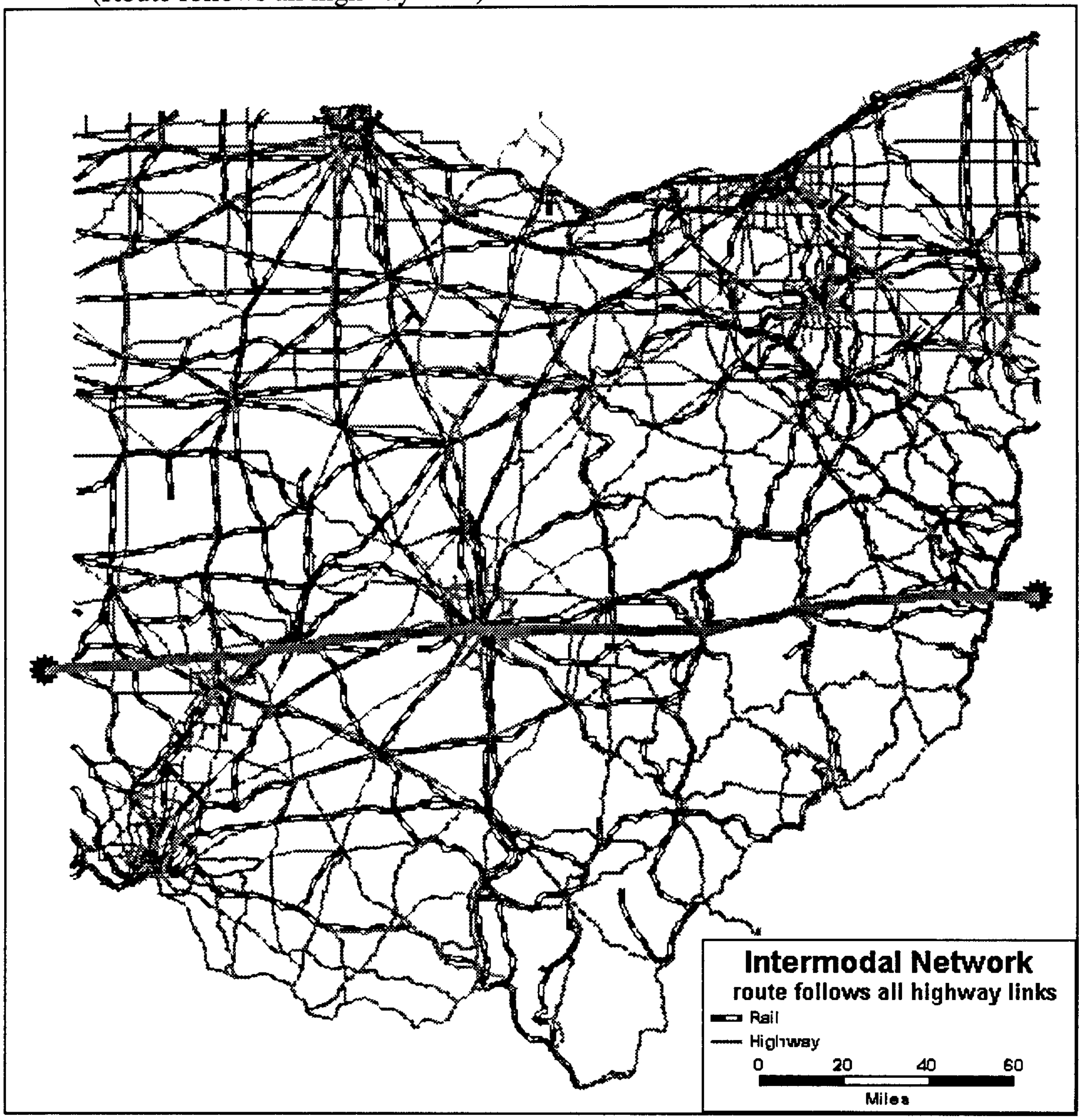


Figure 8.5b Minimum impedance route from point west of Ohio to point east of Ohio with decreased impedance on intermodal link (Route changes from highway to rail links)

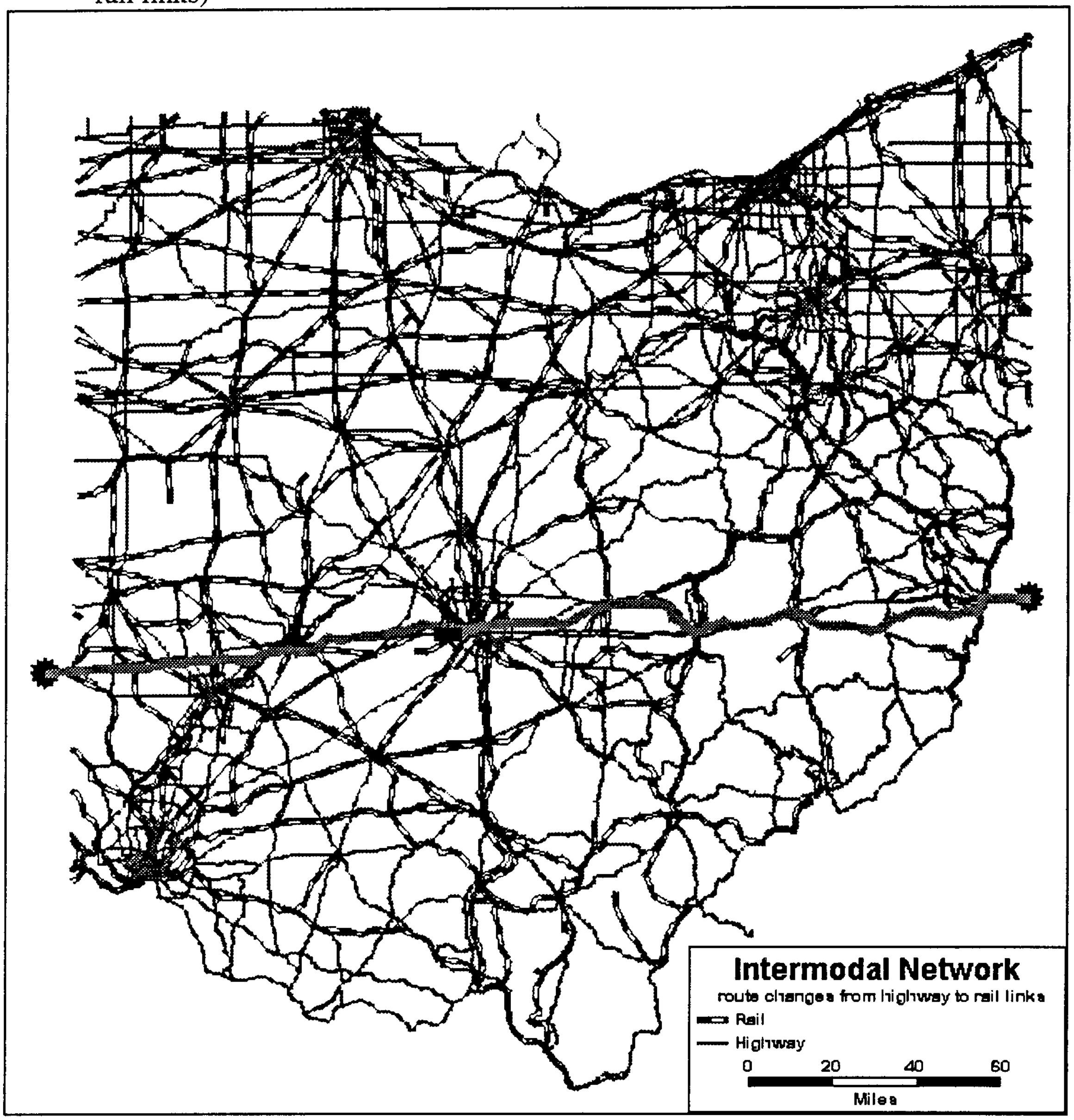
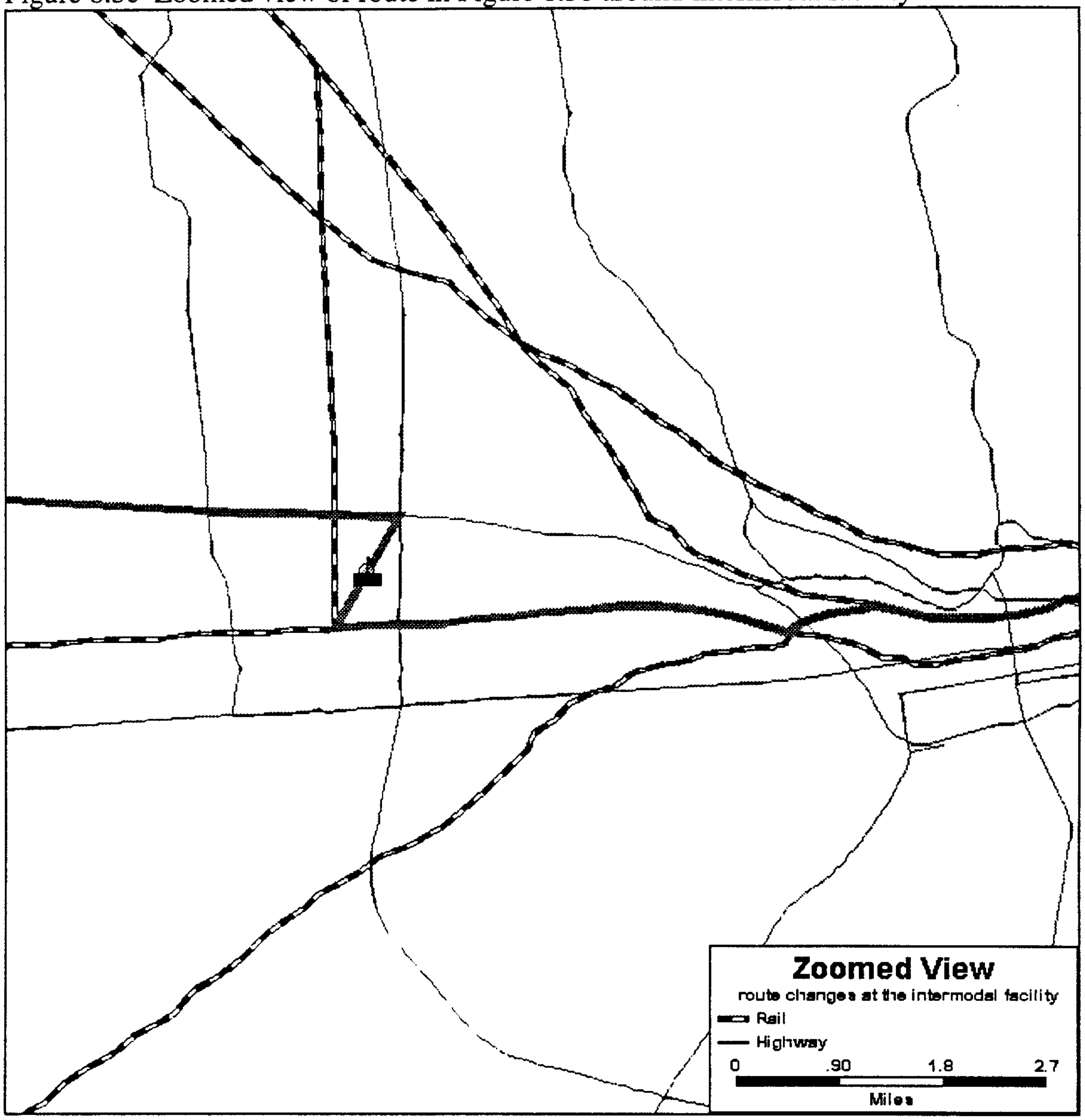


Figure 8.5c Zoomed view of route in Figure 8.5b around intermodal facility



To illustrate the effect of link impedance on the mode chosen through the network, we added a node in the intermodal links at the northeastern and southwestern Ohio locations depicted in Figure 8.2 and considered these as the origin and destination of the shipment. We found the minimum impedance route from the northeastern to the southwestern nodes again using the TransCAD®shortest path routine with the NTAD lengths as the impedance. This route, shown in Figure 8.6a, used all rail links. We then artificially increased the lengths of some of the rail links, and the route (Figure 8.6b) changed to one using all highway links.

The NTAD could be used, then, as a network that could eventually be used for intermodal assignment. Modal and intermodal routes depended on impedances of links representing intermodal transfers and mode-route selection in the manner expected. Adding the intermodal transfer links and changing impedances was straightforward. The major challenges associated with developing intermodal assignment capabilities would be in developing the freight OD matrix and, especially, an acceptable assignment logic and estimates of impedance parameters associated with this logic. The NTAD would appear to be a useful tool for developing and testing intermodal assignment algorithms.

Figure 8.6a Minimum impedance route using original impedance (Route follows all rail links)

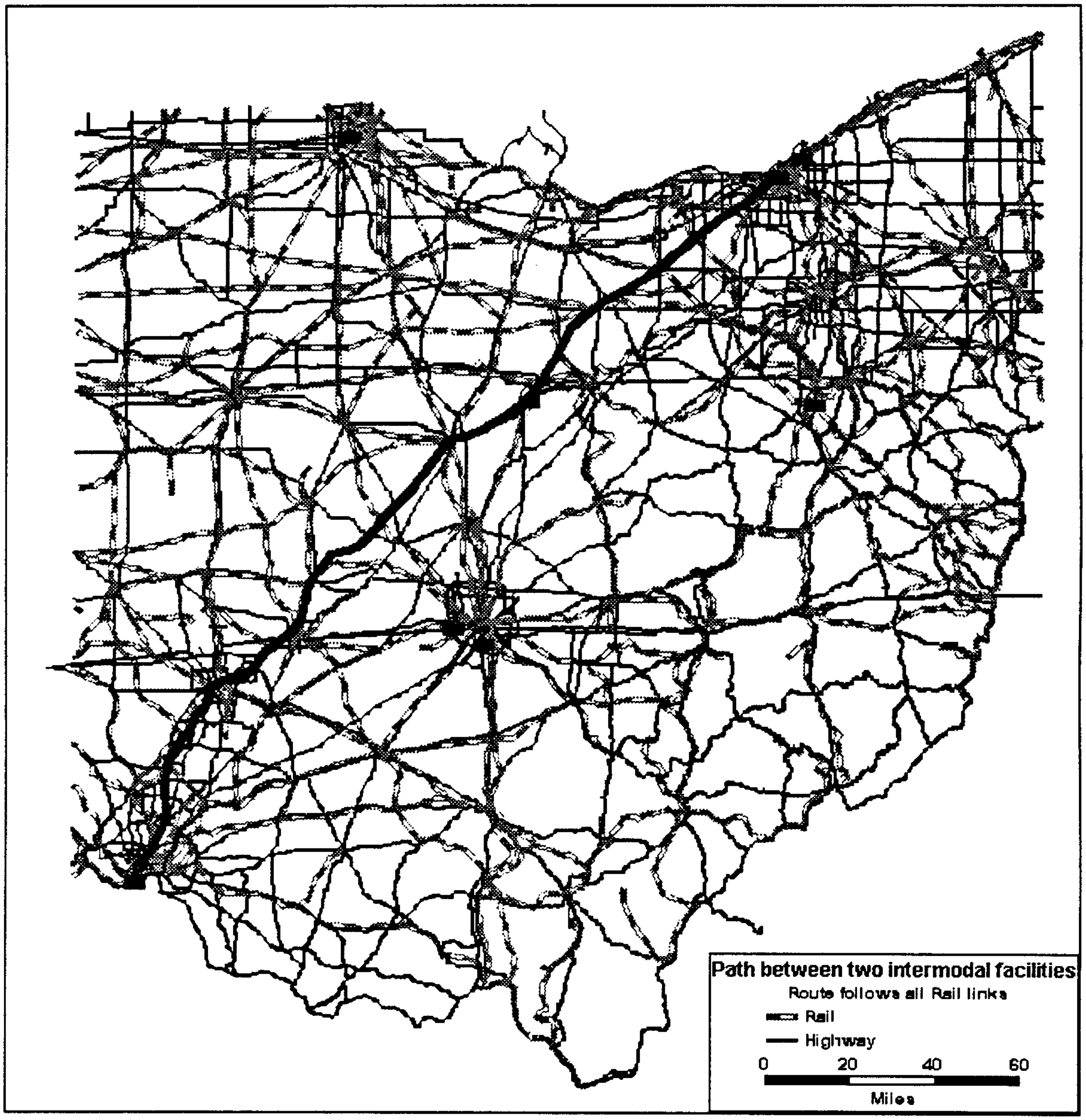
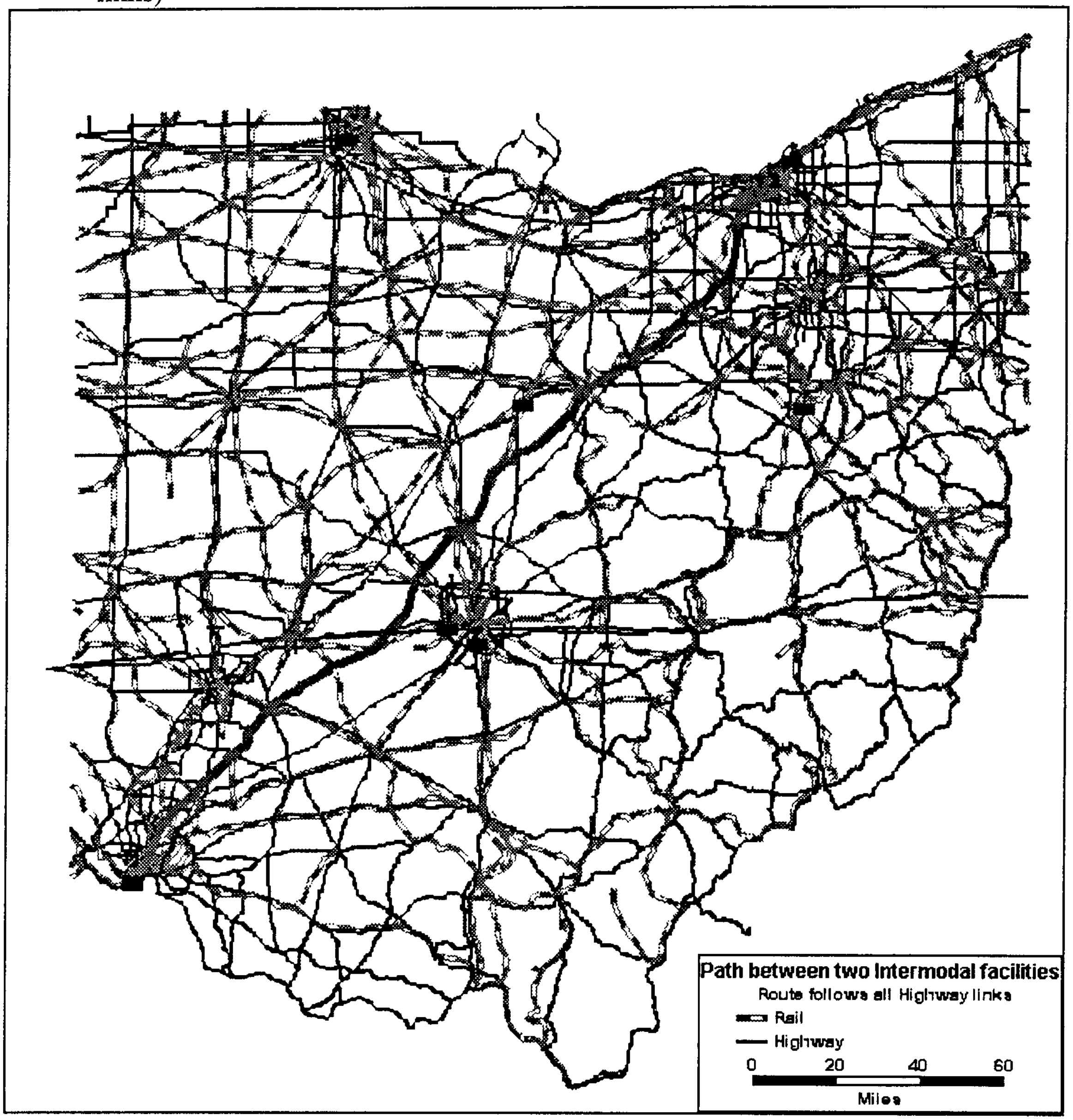


Figure 8.6b Minimum impedance route for same origin-destination pair of Figure 8.6a after increasing impedance on selected rail links (Route follows all highway links)



Section 9. Conclusions

Our review of freight modeling in the literature and in practice makes it clear that available software packages could be used with existing databases, supplemented by limited special data collection efforts, to develop working freight models in Ohio. These models could either be statewide, complex modeling systems, or more limited models. Our goal, then, was to assess whether the time is appropriate for ODOT to pursue freight model development and, if so, how to pursue it.

Model developments seem promising in the state DOT's we investigated. However, a lack of experience with applications in these states, coupled with a wide range of model outputs desired from potential users in Ohio, makes it difficult to foresee where Ohio freight models would see routine application. We feel that a panel of experts could forecast more accurately and more flexibly than mathematical freight models that would be developed in the near future, and that the experts could probably fit in the institutional framework surrounding policy or alternative analysis better. Freight modeling is new, and like most initial efforts in large-scale modeling, we expect products to be of limited initial value.

Despite our pessimism for the immediate value of an Ohio freight forecasting system, we are optimistic about the long-term potential of freight modeling. The importance of issues requiring freight forecasts will not go away. Moreover, it appears that freight forecasting activity will be increasing elsewhere. If those responsible for financing the models remain committed to development, initial results can be modified when it becomes apparent where improvements are possible. Although initial efforts would probably fall short, we expect sustained efforts to produce a useful product. Moreover, even if inaccurate, formal mathematical modeling systems appear objective and sophisticated, advantages that cannot be dismissed in the public arena. We, therefore, *encourage ODOT to pursue a statewide freight model if:*

i) a commitment is made to sustained, rather than one-time development efforts;

ii) the development efforts are similar to those that will be made in other states;

iii) parallel efforts are made to investigate, test, and track the performance of alternative formulations.

Whether out of need for the improved accuracy that could eventually result from a model or from the imposition of mandates, we believe that ODOT will eventually become involved with routine freight modeling, even if nothing is pursued at present. There are good reasons to join the effort now, however. Modeling capabilities will develop more quickly with more participants, and, therefore, ODOT would be "doing its part" in developing the field of freight forecasting for support of public policy issues. Of more direct benefit, getting involved early will allow ODOT personnel to more quickly gain experience and to more readily influence future model developments and federal research and data collection efforts.

Since there is no consensus in Ohio on desired freight model outputs, it makes sense to pursue a model that can produce fairly general outputs and that can be readily modified and expanded in light of future developments. The traditional 4-step urban demand modeling system has responded to these needs in the passenger forecasting context. This responsiveness and its familiarity to developers, administrators, and users probably explain why the statewide freight models being developed in the DOT's we investigated are based on this framework. We expect a modular, 4-step-like approach to form the basis of most, if not all, near-term efforts in other statewide freight models.

Since several states will be pursuing similar models, the opportunity exists for sharing expertise. This opportunity is a big reason for our optimism. More strongly, we *recommend against pursuing a unique model.* Given ODOT's present staffing situation, the large-scale, statewide freight models that presently seem most suitable would have to be developed under consulting contract and turned over to ODOT when finished. We expect ODOT personnel to be interested and involved in the development efforts, but we fear that they will be unable to commit the time necessary to develop a sense of ownership in the initial product. Without this sense of ownership, they would be insufficiently motivated to fight for further contracts when administrators become aware of the difficulties that will inevitably arise from the initial product. If ODOT develops a unique product, or proceeds without support of colleagues in other states, we foresee that development efforts would be halted after the initial contracts come to a close.

On the other hand, if ODOT could join with other states in sustained development, initial deficiencies would be overcome, and incremental progress should eventually lead to a valuable product. *We suggest* that collaboration with other states be formalized by *pooling funds to develop regional models that could be scaled to the appropriate levels for the participating states.* This should reduce development costs paid by any state, lead to easier and more accurate estimation of relevant external flows, and motivate individuals to discuss issues and share expertise. It would also focus interest in such a way that a formal group of states could exert stronger influence on federal studies and data collection efforts. Most importantly, perhaps, it would motivate technical DOT personnel to remain intellectually committed to the development process over the long run and help the states assist each other in trying to convince administrators to remain financially committed when potential benefits are questioned.

Research and experimentation with alternative model components and tracking performance on an ongoing basis *would be important* in hastening the usefulness of the model. Alternatives would exist for almost every component of a proposed model. Illustrations in Sections 5 and 6 indicate what is probably obvious: Different alternatives can produce different forecasts. Best-fit parameter estimates of certain specifications can be obtained with past data, and validation studies can be conducted on hold-out samples, but these do not indicate how well the model will forecast future flows or how useful the forecasts will be in updating expert forecasts. Moreover, a specification or estimate that appears to be inferior today may turn out to be superior under future conditions. To proceed toward implementation, initial choices must be made, and traditional tests of specification are good ways to make these choices. However, these initial choices should be reconsidered through time. The anticipated models should be designed flexibly enough to allow easy updating and substitution of improved modules, specifications, and parameter estimates as more freight forecasting experience is gained. However, if the model is simply turned over to the states after being developed with outside expertise, such updates and improvements will likely be made less frequently than they should. An ongoing research, development, and monitoring program that is explicitly targeted toward improving components would ensure that the model is designed with an open architecture and keep the states aware that substitutions and regular improvements are expected. Keeping in mind at the outset that the forecasts are intended to be modified by expert forecasts should also argue for developers making model components as transparent as possible.

Smaller-scale, stand-alone freight models are also feasible. If desired, the indicator methods of Section 5 and the method for forecasting freight at a new intermodal facility presented in Section 6 could be used in Ohio. These models are straightforward, and they probably would not be used frequently. Therefore, systematic development of these types of specific models seems unwarranted, except where they can lend insight on components of the anticipated statewide model. The smaller-scale models could be developed as needed for specific studies. However, in the absence of further investigation and development, we feel that a panel of experts could produce forecasts as useful as those produced from these models. Before investing in developing these models for a specific study, we would encourage that experts be polled to see if any of the possible model outputs could change their beliefs about the future. If not, development is clearly not warranted.

We also saw that the National Transportation Atlas Database could be used for intermodal assignment. However, we feel that legitimate intermodal assignment is presently limited by lack of acceptable logic. We do not necessarily suggest that ODOT pursue research in this area, but if a consortium of states wished to develop these capabilities, the supporting infrastructure seems available.

We do *suggest* that ODOT *estimate and use observed truck volumes* on highway segments *to* continually *update a* statewide *truck origin-destination matrix.* The data to estimate an origin-destination matrix are available from a recently conducted roadside survey. In Section 7 we saw that there are several methods that can use observed volumes to update target OD matrices. Since truck volumes are routinely collected, using any of these procedures would be inexpensive to implement on systematic basis. A good estimate of the present OD matrix would be useful in calibrating or validating components of developing models. It could also assist experts when forecasting future OD patterns.

If ODOT does pursue systematic freight modeling, we *encourage* the *formation of* an *advisory group of experts* in shipping, transportation, and freight logistics. The group would meet regularly to discuss freight issues with an objective of ensuring that model developments are relevant. Regular interaction between freight experts and modelers should markedly increase the likelihood that model outputs could be used with expert opinion to produce more valuable forecasts than could be produced by the model or the expert opinion alone. The experts would also be called upon to anticipate structural changes that models could not endogenously capture. And since experts can forecast more flexibly and probably more accurately than mathematical models, ready access to a group of experts would be valuable when forecasts are needed to support alternative or policy analysis. Such an advisory group would be even more useful if the experts could be drawn from several states and if they worked with technical personnel from many state DOT's in a formal manner. Once again, we feel that a pooled funding arrangement would facilitate organizing and profiting from such a group.

REFERENCES

Ben-Akiva, M. and S. R. Lerman, *Discrete Choice Analysis: Theory and Application to Travel Demand*, The MIT Press, Cambridge, Mass., 1985.

Black, W.R., *Transport Flows in the State of Indiana: Commodity Database Development and Traffic Assignment; Phase 2*, HPR-2085, in cooperation with the Indiana Department of Transportation and the Federal Highway Administration, Indiana Department of Transportation, Indianapolis, IN, July 1997.

Bureau of Census, *1992 Census of Transportation, Communication, and Utilities,* U.S. Department of Commerce, Washington DC, 1995.

Bureau of Economic Analysis (BEA), *BEA Regional Projections to 2045*, U.S. Department of Commerce, Washington DC., 1996.

Bureau of Transportation Statistics (BTS), *1993 Commodity Flow Survey*, Bureau of the Census, U.S. Department of Commerce, Washington DC, 1997a.

Bureau of Transportation Statistics (BTS), *1993 Commodity Flow Survey, Commodity Coding Manual,* U.S. Department of Transportation, Washington DC, 1997b.

Bureau of Transportation Statistics (BTS), *1993 Commodity Flow Survey, States and NTARs Tables 1-11, CD-ROM*, Bureau of the Census, U.S. Department of Commerce, Washington DC, 1997c.

Bureau of Transportation Statistics (BTS), *National Transportation Atlas Databases, CD-ROM*, U.S. Department of Transportation, Washington DC, 1997d.

Bureau of Transportation Statistics (BTS), *Directory of Transportation Data Sources*, U.S. Department of Transportation, Washington DC, 1996.

Bureau of Transportation Statistics (BTS), *Transportation Statistics Annual Report*, U.S. Department of Transportation, Washington DC, 1994.

Bureau of Transportation Statistics (BTS), *ICC Carload Waybill Sample Database: 1988-1992 CD-ROM*, U.S. Department of Transportation, Washington DC, 1996.

Caliper Corporation, *Travel Demand Modeling with TransCAD® 3.0*, Newton, Massachusetts, 1996.

Cambridge Systematics, Inc., Leeper, Cambridge & Campbell, Inc., Sydec, Inc., T. M. Corsi, and C. M. Grimm, *Forecasting Freight Transportation Demand, A Guidebook for Planners and Policy Analysts*, prepared for National Cooperative Highway Research Program Project 8-30, 1997.

Cambridge Systematics, Inc., COMSIS Corporation, University of Wisconsin-Milwaukee, *Quick Response Freight Manual,*; prepared for U.S. Department of Transportation (Federal Transit Administration, Federal Highway Administration, Office of the Secretary) and U.S. Environmental Protection Agency, September 1996.

Cambridge Systematics, Inc, *Major Corridor Investment-Benefit Analysis System: Model Documentation, Part I, Draft*, April 10, 1998.

Fang, Y. Y., R. Harrision, and H. S. Mahmassani, *Forecasting Freight Traffic between The U.S. and Mexico*, Research Report 1319-2, Center for Transportation Research, Austin, Texas, August 1996.

Keeney, R., and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, Wiley, New York, 1976.

McCord, M. R. and O. G. Villoria, "Nonlinear Utility in Time and Cost of Trips: Disaggregate Results from an Ordinal Methodology", *Transportation Research Record*, No. 1138, pp. 8-17, 1987.

Navigation Data Center, *Estimated Waterborne Commerce Statistics Publications CD-ROM*, U.S. Army Corps of Engineers, Alexandria, VA, 1997.

Nellet, R. Personal Communication (email description of MDOT statewide model), July 27, 1998.

Nielsen, O. A., "A New Method for Estimating Trip Matrices from Traffic Counts", Seventh International Conference on Travel Behavior, Chile, Preprints, 1994.

Ohio Department of Transportation, *Access Ohio: Ohio Multi-Modal State Transportation Plan to the Year 2020: Macro Phase*, Columbus, OH, October 1993.

Park, Man-Bae and R. L. Smith, Jr., "Development of a Statewide Truck Travel Demand Model with Limited O-D Survey Data", Transportation Research Board, 76th Annual Meeting, Preprints, Washington DC, 1996.

Parsons Brinckerhoff Quade & Douglas, Inc., Urban Analytics, Inc., ECONorthwest, and KJS Associates, Inc., *Consultant Recommendations for the Development of Phase II Databases, Models, and Forecasting Methods: Transportation and Land Use Model Integration Program, Phase I, Task 1.6.* Preliminary Draft for Review prepared for Transportation Development Branch, Oregon Department of Transportation, September 13, 1996a. (http://www.odot.state.or.us/tdb/planning/modeling/papers.html)

Parsons Brinckerhoff Quade & Douglas, Inc., Urban Analytics, Inc., ECONorthwest, and KJS Associates, Inc., *Consultant Scope of Work for the Transportation and Land Use Model Integration Program.* Internal Working Paper prepared for Transportation Development Branch, Oregon Department of Transportation, May 9, 1996b. (http://www.odot.state.or.us/tdb/planning/modeling/papers.html)

Reebie Associates, *TRANSEARCH Database*, Connecticut, 1997.

Sheffi, Y., *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1985.

Shkurti, W. J. and J. R. Bartle, *Benchmark Ohio*, School of Public policy and Management, Ohio State University and Ohio State University press, Columbus, OH, 1991.

Smith, S., Personal Communication (Phone conversation), July 23, 1998.

Transportation Management Group, Inc., Leeper, Cambridge & Campbell, Inc., and COMSIS Corporation, *North Carolina Air Cargo System Plan and a Global Air Cargo Industrial Complex*, February 1992.

Upton, W., Personal communication (Phone conversation), July 24, 1998.

U.S. Postal Service, *National Five Digit Zip Code & Post Office Directory*, 1994.

Weinblatt, H. and Edwards R., "Three Procedures for Forecasting Freight Transportation Demand", Transportation Research Board, 76th Annual Meeting, Washington, DC, Preprints, January 1997.

Winston, C., "The Demand for Freight Transportation: Models and Applications", *Transportation Research* Vol. 17A, No. 6, pp. 419-427, 1983.

Woods and Poole Economics, Inc., *1998 State Profile Series Database*, Washington, DC. 1998.

Zhang, J., "Sensitivity of Origin-Destination Matrix Estimation from Ground Counts", *M.S. Thesis,* Civil Engineering, The Ohio State University, 1998.

Appendix 2.1

Individuals Interviewed for Needs of Section 2.1

John R. Platt, Chief of Staff, Ohio Department of Transportation

Mark Byram, Chuck Gebhardt, and Greg Giaimo, Bureau of Technical Services, Ohio Department of Transportation

Gary Coburn and Larry Sutherland, Bureau of Planning, Ohio Department of Transportation

Lou Jannazo, Chief Planner, Ohio Rail Development Commission

Elena Constantine, Mid-Ohio Regional Planning Commission

Appendix 2.2

Questionnaire and Respondents to Survey of State DOT's

2.2a Preliminary Fax/Email Letter

The Ohio State University

Department of Civil & Environmental Engineering and Geodetic Science

470 Hitchcock Hall - 2070 Neil Avenue - Columbus, OH 43210 - USA

Date:

To: Alabama Public Service Commission

Fax No.: (334) 242-5218 **Phone No.**: (334) 242-5980

Message:

Dear Sir or Madam:

We are conducting a research project for the Ohio Department of Transportation to help them investigate the feasibility of developing a freight forecasting system. We have developed a short questionnaire to survey the other State DOTs. This questionnaire is regarding to some freight activities (Origin-Destination data/survey). In order to send this survey questionnaire to the right person, we need the email/snail mail addresses or fax number of the appreciate people. Please give us these information via email (chung.77@osu.edu), or fax (614-292-3780). Should you have any questions or comments, feel free to contact me (chung.77@osu.edu; 614-688-3761) or my advisor – Dr. Mark McCord (mccord.2@osu.edu; 614-292-2388).

Thank you for any information you can provide to us.

Sincerely,

Yi-Ying Chung
Graduate Research Associate

From: Dr. Mark R. McCord/Yi-Ying Chung
Phone No.: 614-292-2388 Fax No.: 614-292-3780
Number of pages (including cover page): 1

2.2b Cover Letter

Dear Sir or Madam:

I am conducting a research project for the Ohio Department of Transportation to help them investigate the feasibility of developing a freight forecasting system. One step is to investigate what other state DOTs have been doing in this area. Therefore, we have developed a short survey. The Ohio Department of Transportation has looked at this survey and given me permission to send it out. Completing this survey should take only a few minutes once it is in the hands of the individual familiar with any freight analysis conducted by your agency or who knows that none is conducted. Please email the completed survey to my assistant (chung.77@osu.edu) or fax (614-292-3780) it to me. Should you have any questions or comments, feel free to contact me (mccord.2@osu.edu; 614-292-2388).

Thank you for your anticipated cooperation.

Sincerely,

Dr. Mark R. McCord
Associate Professor
Civil Engineering
City & Regional Planning

Encl.

2.2c Survey Questionnaire

1. Does your organization track correlation between general freight data and other variables?
____ Regularly
____ Sometimes, but not regularly
____ Never (go to Question 5)

2. What type of <u>freight data</u> does your organization track? ______________________
____

Where do you obtain these data?
____ Special surveys
____ Private data supply company (Please specify)
____
____
____ Public data source (Please specify)
____
____
____ Other (Please specify)____
____

3. What type of <u>"correlated" variables</u> does your organization track? ______________
____

Where do you obtain these data?
____ Special surveys
____ Private data supply company (Please specify)
____
____
____ Public data source (Please specify)
____
____
____ Other (Please specify)____
____

4. Does your organization explicitly analyze the correlation between freight data and other variables?
____ Yes
____ No (go to Question 5)

What type of explicit analysis is performed to identify the correlation?
____ Graphing trends
____ Calculating ratios
____ Regression analysis
____ Advanced time series analysis
____ Other (Please specify) ______________________________________
____

5. Does your organization have/use any freight origin-destination (O-D) tables?
____ Yes
____ No (go to Question 6)

How frequently do you update these tables? (e.g., annually, every 5 years, irregularly, ...)

____

Do you keep O-D tables by commodity? ____ Yes; ____ No

If yes, by what kind of Commodity Category?

____ Standard Transportation Commodity Code (STCC)
____ Standard International Trade Classification (SITC)
____ General names of commodity (Example, please) ____________________
____
____ Other (Example, please) ____
____

What are the units associated with the commodity?
(e.g., ton/year; carloads/day; dollar value of commodity/quarter)

____

How does your organization update these commodity O-D tables?

____ Special O-D surveys
____ Models and other data (Please specify; e.g., trip generation data and gravity models)
____
____
____ Other ____

6. Does your organization regularly monitor truck trips on highway links?
____ Yes
____ No (go to Question 7)

Does your organization also forecast truck trips on the links?
____ Yes
____ No (go to Question 7)

What forecasting models do you perform to forecast the link flows?
____ Trend projection
____ Correlation with other indicators
____ Trip Assignment
____ Other (Please specify)____

7a. Agency (State DOT) of individual answering this questionnaire:
____

7b. Section/Bureau/Office of individual answering this questionnaire:
____

8. Who could be contacted if we have any further questions?
   Name: ______________________________________
   Phone No.:____________________________________
   Email @: ____________________________________
   Bureau/Office: _______________________________

2.2d Sampled and Responding States

States Contacted	* = States to which questionnaire was sent	** = States returning questionnaire	Contact
Alabama (AL)	*		
Alaska (AK)	*	**	Jeff Ottesen Chief of Statewide Planning jeff_ottesen@dot.state.ak.us (907) 465-6971
Arizona (AZ)	*	**	Louis Tognacci Intermodal Division Ltognacci@dot.state.az.us (602) 255-8137
Arkansas (AR)			
California (CA)			
Colorado (CO)	*	**	Dave L. Ruble Jr. Intermodal Branch (303) 757-9819
Connecticut (CT)	*	**	Joseph Spragg Bureau of Policy and Planning (860) 594-2022 (860) 594-2056 (FAX)
Delaware (DE)	*		
Florida (FL)			
Georgia (GA)	*		
Hawaii (HI)	*	**	Gordon Lum Oahu MPO (808) 587-2015 (808) 587-2018 (FAX)
Idaho (ID)	*		
Illinois (IL)	*	**	James Johnson Office of Planning & Programming (312) 793-5744 (312) 793-5966 (FAX)
Indiana (IN)	*	**	Steve Smith Division of Planning & Programming planners376@aol.com (317) 232-5646 (317) 232-1499 (FAX)
Iowa (IA)	*	**	Craig O'riley Planning & Programming Division coriley@iadot.email.com (515) 239-1520
Kansas (KS)	*	**	Rick Miller Statewide Planning Unit rick@dtthpo.wpo.state.ks.us (913) 296-7441

States Contacted	* = States to which questionnaire was sent	** = States returning questionnaire	Contact
Kentucky (KY)	*	**	Rob Bostrom Division of Transportation Planning rbostrom@mail.kytc.state.ky.us (502) 564-7183
Louisiana (LA)	*	**	James B. Norman Weights & Standards jnorman@dotdmail.dotd.state.la.us (504) 377-7131
Maine (ME)	*	**	Edward W. Hanscom Planning Division www.ed.hanscom@state.me.us (207) 287-3131 (207) 287-3292 (FAX)
Maryland (MD)	*		
Massachusetts (MA)			
Michigan (MI)			
Minnesota (MN)	*	**	C. Snaft Freight T.R.I.M. Division (612) 276-1666
Mississippi (MS)			
Missouri (MO)	*		
Montana (MT)	*	**	Bill Cloud Data and Statistics Bureau (406) 444-6114 (406) 444-7671 (FAX)
Nebraska (NE)	*	**	Rick Ernstmeyer Transportation Planning Division dor5005@vmhost.cdp.state.ne.us (402)479-4520 (402) 479-3884 (FAX)
Nevada (NV)	*		
New Hampshire (NH)	*	**	Stephen W. Gary Highway Maintenance (603) 271-2693
New Jersey (NJ)	*		
New Mexico (NM)			
New York (NY)	*		
North Carolina (NC)			
North Dakota (ND)	*	**	Jeff Patten Planning Division (701) 326-4197 (701) 3281404 (FAX)
Ohio (OH)	*		

States Contacted	* = States to which questionnaire was sent	** = States returning questionnaire	Contact
Oklahoma (OK)	*	**	Sam Shehab Planning Division (405) 521-6433 (405) 521-6917 (FAX)
Oregon (OR)	*	**	Bill Upton Planning Section william.j.upton@odot.state.or.us (503) 986-4106
Pennsylvania (PA)	*		
Puerto Rico	*		
Rhode Island (RI)			
South Carolina (SC)	*		
South Dakota (SD)			
Tennessee (TN)	*	**	Awin H. Pearson Office of Public Transportation (615) 741-3227 (615) 741-3169 (FAX)
Texas (TX)	*	**	Joe Barnard Motor Carrier Division (512) 465-3044
Utah (UT)			
Vermont (VT)	*	**	Karen Songhurst ksonghurst@aot.state.vt.us
Virginia (VA)	*		
Washington (WA)			
West Virginia (WV)	*	**	Jerry L. Legg Transportation Planning Division jllegg@mail.dot.state.wv.us (304)558-2864
Wisconsin (WI)			
Wyoming (WY)			

2.2e List of Standard Transportation Commodity Code (STCC)

STCC Code	Commodity Description
01	Farm products
08	Forest products
09	Fresh fish or other marine products
10	Metallic ores
11	Coal
13	Crude petroleum, natural gas or gasoline
14	Nonmetallic minerals
19	Ordnance or accessories
20	Food and kindred products
21	Tobacco products, excluding insecticides
22	Textile mill products
23	Apparel or other finished textile products or knit apparel
24	Lumber or wood products, excluding furniture
25	Furniture or fixtures
26	Pulp, paper or allied products
27	Printed matter
28	Chemicals or allied products
29	Petroleum or coal products
30	Rubber or miscellaneous plastic products
31	Leather or leather products
32	Clay, concrete, glass or stone products
33	Primary metal products, including galvanized
34	Fabricated metal products
35	Machinery, excluding electrical
36	Electrical machinery, equipment or supplies
37	Transportation equipment
38	Instruments, photographic goods, optical goods, watches, or clocks
39	Miscellaneous products of manufacturing
40	Waste or scrap materials not identified by producing industry
41	Miscellaneous freight shipments
42	Containers, carriers or devices, shipping, returned empty
48	Waste hazardous materials or waste hazardous substances
--	Commodity unknown

Source: *Commodity Coding Manual*, 1993 Commodity Flow Survey, Bureau of the Census

Appendix 3.1

Example of Woods & Poole Economics, Inc. Database: State Profile Series

ALAMEDA, CA

Unit of Geography: County
1990 Land Area: 737.5 Square Miles
FIPS CODE: 06001

Part of the Far West Region
Part of the San Francisco-Oakland-San Jose, CA Economic Area
Part of the San Francisco-Oakland-San Jose DMA
Part of the San Francisco-Oakland-San Jose, CA CMSA
Part of the Oakland, CA PMSA

	ALAMEDA, CA	1970	1980	1990	1995	1997	1998	1999	2000	2005	2010	2015	2020
1	Total Population (Thousands)	1073.96	1108.79	1278.24	1321.98	1337.16	1346.04	1354.81	1363.43	1405.90	1450.37	1498.15	1546.84
2	Under 5 Years	83.80	73.27	100.71	103.28	98.18	96.73	95.56	94.80	93.87	96.46	100.28	103.05
3	5 To 9 Years	97.35	70.49	87.68	94.01	99.34	100.63	100.75	99.35	92.32	91.82	94.56	98.36
4	10 To 14 Years	98.04	79.26	77.42	84.21	85.75	87.27	89.67	92.69	99.06	92.52	92.55	95.62
5	15 To 19 Years	94.23	95.65	81.68	80.38	84.42	86.35	87.73	88.90	98.36	105.98	99.03	99.01
6	20 To 24 Years	112.04	110.83	105.40	94.07	89.59	90.75	93.02	95.53	105.72	116.66	124.97	116.74
7	25 To 29 Years	86.26	110.08	122.70	108.57	104.24	101.36	97.89	94.42	96.73	106.06	115.87	123.23
8	30 To 34 Years	65.83	104.72	126.74	125.56	118.64	115.08	112.02	110.21	96.66	98.58	107.12	116.18
9	35 To 39 Years	58.06	77.64	114.76	124.07	125.52	125.39	124.75	122.87	108.73	95.67	97.30	105.30
10	40 To 44 Years	62.13	59.61	102.66	112.35	118.05	120.55	122.27	123.74	122.66	108.78	95.79	97.34
11	45 To 49 Years	66.47	53.15	73.77	89.91	94.75	96.57	98.94	101.49	112.32	111.79	99.88	88.51
12	50 To 54 Years	58.79	55.94	56.21	64.73	71.63	74.53	78.02	81.71	92.86	103.32	103.64	93.54
13	55 To 59 Years	50.65	56.99	47.35	49.37	51.80	54.03	55.76	57.62	73.26	83.92	93.92	94.97
14	60 To 64 Years	41.56	46.52	46.03	42.91	42.90	43.55	44.31	44.75	52.18	66.65	76.89	86.49
15	65 To 69 Years	32.47	39.01	44.55	43.91	43.15	42.34	41.77	41.32	42.79	49.68	63.28	72.91
16	70 To 74 Years	25.80	28.95	34.26	39.22	39.13	38.90	38.56	38.48	36.41	37.76	43.83	55.75
17	75 To 79 Years	19.05	21.08	26.32	29.58	32.00	32.97	33.67	34.29	33.88	32.38	33.86	39.43
18	80 To 84 Years	12.82	13.72	16.30	19.23	20.29	20.64	21.12	21.63	25.27	25.24	24.30	25.70
19	85 Years and Over	8.61	11.88	13.69	16.61	19.79	18.40	19.00	19.63	22.81	27.09	31.08	34.67
20	White Population	857.40	802.02	834.18	828.98	827.05	825.91	824.56	823.08	814.85	807.77	803.79	801.66
21	Black Population	162.19	205.77	233.89	241.42	242.69	243.83	244.86	245.79	248.97	250.76	251.52	251.90
22	Other Population	54.37	101.00	210.17	251.58	267.43	276.31	285.39	294.55	342.08	391.85	442.84	493.27
23	Hispanic Population, Any Race	135.02	126.50	182.98	219.61	230.85	234.99	239.16	143.42	264.04	286.89	313.10	343.20
24	Population 0-19 Years (Thousands)	373.42	318.66	347.50	361.89	367.69	370.98	373.72	375.74	383.61	386.78	386.41	396.05
25	Population Age 20-64 Years	601.80	675.48	795.61	811.54	817.12	821.80	826.98	832.34	861.14	891.43	915.39	922.32
26	Population Age 65 Years and Over	98.75	114.65	135.12	148.56	152.35	153.26	154.11	155.35	161.15	172.16	196.35	228.47
27	Male Population (Thousands)	526.16	539.18	628.73	650.04	654.87	658.93	662.94	666.93	686.47	707.64	730.23	752.83
28	Female Population	547.80	569.61	649.51	671.94	682.29	687.12	691.87	696.49	719.43	742.73	767.92	794.01
29	Population Age 16 Years and Over	776.05	868.38	997.64	1024.36	1037.19	1044.75	1052.27	1059.49	1100.55	1150.63	1192.32	1231.11
30	Median Age of Population (Years)	27.81	30.67	32.48	33.90	34.55	34.80	35.03	35.24	35.91	35.91	35.71	35.97
31	Total Employment (Thousands)	499.91	601.72	761.59	791.10	819.98	829.34	838.70	848.07	895.56	944.60	995.48	1048.78
32	Farm Employment	3.76	2.14	1.44	1.43	1.37	1.35	1.32	1.30	1.20	1.13	1.09	1.07
33	Agricultural Services, Other	1.85	3.92	5.11e	6.16	6.54	6.58	6.62	6.65	6.87	7.12	7.42	7.77
34	Mining	1.24	0.96	0.84e	0.82	0.81	0.81	0.81	0.81	0.80	0.81	0.81	0.83
35	Construction	21.87	29.17	38.40	35.87	37.67	37.93	31.18	38.42	39.59	40.72	41.87	43.08
36	Manufacturing	85.90	87.19	84.93	87.02	87.42	87.79	88.14	88.47	89.89	91.00	91.85	92.51
37	Transport, Comm. & Public Utility	33.16	34.21	44.48	45.15	46.70	47.11	47.52	47.91	49.75	51.39	52.87	54.19
38	Wholesale Trade	24.39	31.18	43.19	48.33	49.37	49.98	50.57	51.16	53.99	56.69	59.32	61.96
39	Retail Trade	73.02	97.43	121.89	124.98	128.29	129.17	130.06	130.94	135.40	140.05	144.96	150.27
40	Finance, Ins. & Real Estate	34.60	46.54	54.08	49.31	50.03	50.40	50.76	51.13	52.95	54.80	56.68	58.63
41	Services	92.42	144.38	222.45	258.04	278.53	284.74	291.01	297.34	330.02	364.49	400.68	438.66
42	Federal Civilian Govt.	27.84	21.95	23.25	19.77	19.04	18.92	18.80	18.67	18.03	17.39	16.77	16.19
43	Federal Military Govt.	24.91	12.12	19.23	14.23	13.48	13.48	13.48	13.47	13.46	13.45	13.44	13.43
44	State and Local Govt.	74.94	90.52	102.32	100.01	100.73	101.09	101.45	101.80	103.61	105.56	107.72	110.16
45	Total Earnings (Millions 1992 $)	14309.32	17299.97	22943.85	248912.2	26612.12	27167.40	27728.00	28294.52	31255.02	34355.68	37633.53	41135.27
46	Farm Earnings	73.90	54.40	26.16	12.59	19.88	19.90	19.93	19.98	20.39	21.16	22.33	23.98
47	Agricultural Services, Other	37.24	57.00	101.12e	106.17	112.02	113.60	115.20	116.82	125.42	135.03	145.98	158.61
48	Mining	58.51	83.34	28.50e	22.27	27.77	28.00	28.19	28.36	29.08	29.92	31.05	32.54
49	Construction	940.09	1235.12	1681.64	1394.01	1468.96	1488.15	1507.06	1525.75	161.6.54	1704.43	1790.59	1876.80
50	Manufacturing	3072.49	3524.74	3361.53	3781.28	4192.42	4254.74	4315.98	4376.20	4662.50	4923.85	5157.79	5363.43
51	Transport, Comm. & Public Utility	1179.60	1454.92	1864.76	1853.11	1916.39	1950.54	1984.48	2018.23	2184.48	2346.49	2503.18	2654.16
52	Wholesale Trade	864.36	1162.32	1636.51	1887.57	1987.62	2025.54	2063.30	2100.90	2286.75	2470.20	2652.37	2835.28
53	Retail Trade	1643.78	2037.29	2426.82	2452.76	2529.60	2560.28	2590.94	2621.62	2776.41	2935.23	3099.15	3270.40
54	Finance, Ins. & Real Estate	624.10	836.30	1008.24	1136.90	1265.48	1295.02	1324.74	1354.63	1506.54	1661.61	1818.02	1974.16
55	Services	2070.68	3285.39	6079.97	7601.86	8380.00	8678.22	8983.47	9295.99	10972.39	12847.63	14927.26	17219.62
56	Federal Civilian Govt.	968.88	792.58	843.47	832.29	834.65	835.01	835.20	835.26	833.93	830.97	827.30	823.87
57	Federal Military Govt.	490.49	225.93	418.00	329.17	326.52	329.31	332.11	334.92	349.13	363.61	378.33	393.28
58	State and Local Govt.	2285.20	2550.64	3467.13	3486.81	3550.81	3589.10	3627.42	3665.87	3861.46	4065.56	4280.19	4509.14
59	Personal Income (Millions 1992 $)	17659.78	22631.96	31187.94	33170.65	35251.84	35966.26	36688.74	37419.85	41214.71	45261.21	19572.24	54171.10
60	Wage and Salaries	12417.92	14450.22	18795.46	20034.88	21536.42	22044.35	22477.36	22955.91	25406.16	28005.00	30767.59	33708.98
61	Other Labor Income	671.84	1395.14	2049.46	2530.85	2579.52	2628.03	2676.86	2726.07	2974.50	3232.83	3502.27	3783.83
62	Proprietors Income	1219.56	1454.61	2098.93	2331.07	2496.18	2535.02	2573.78	2612.54	2844.36	3097.84	3363.67	3642.47
63	Dividends, Interest & Rent	2483.58	3639.22	5361.51	5304.99	5592.84	5703.41	5814.83	5927.04	6498.93	7085.18	7680.26	8278.02
64	Transfer Pmts. To Persons	2057.57	3223.64	4170.13	5134.18	5364.75	5479.45	5597.77	5719.85	6391.32	7178.90	8105.52	9199.30
65	Less Social Ins. Contributions	681.62	903.49	1533.18	1717.72	1832.87	1886.99	1942.69	2000.08	2315.08	2684.80	3120.80	3638.28
66	Residence Adjustment	-509.07	-627.39	245.64	-447.59	-485.00	-497.01	-509.16	-521.48	-585.48	-653.75	-726.28	-803.22
67	Income Per Capita (1992 $)	16444	20411	24399	25092	26363	26720	27080	27445	29316	31207	33089	35021
68	Income Per Capita (Current $)	4849	11939	22669	27071	29693	30758	31892	33100	40421	50426	63487	79805
69	W&P Wealth Index (U.S. = 100)	114.04	115.12	144.55	113.45	114.85	114.95	115.04	115.12	115.41	115.58	115.67	115.72
70	Number of Households (Thousands)	366.81	428.82	481.23	492.46	503.57	508.02	512.40	516.71	537.71	557.73	575.67	589.37
71	Persons Per Household (People)	2.82	2.52	2.58	2.60	2.59	2.58	2.58	2.57	2.55	2.54	2.54	2.56
72	Retail Sales Per Household (1992 $)	19415	20079	21096	21706	21729	21665	21660	216.96	22220	23087	24231	25644
73	Mean Household Income (1992 $)	46679	51703	63382	65743	68552	69342	70142	70951	75120	79573	84484	90202
74	Mean Household Inc. (Current $)	13766	30241	58888	70930	77210	79821	82604	85569	103577	128580	162099	205551
75	Household with Money Inc. (Thousand)	366.81	428.82	481.23	492.46	503.57	508.02	512.40	516.71	537.71	557.73	575.67	589.37
76	Less than $10,000 (1990 $)	n.a.	n.a.	56.18	52.76	51.45	51.29	50.50	49.69	44.88	39.78	35.33	31.48
77	$10,000 to $19,999	n.a.	n.a.	64.02	60.12	58.63	58.44	57.55	56.62	51.14	45.34	40.26	35.87
78	$20,000 to $29,999	n.a.	n.a.	68.85	64.66	63.05	62.85	61.89	60.89	55.00	48.76	43.29	38.85
79	$30,000 to $39,999	n.a.	n.a.	64.83	61.33	59.81	59.62	58.71	57.76	52.17	46.25	41.07	36.59
80	$40,000 to $49,999	n.a.	n.a.	56.85	63.06	66.08	66.94	67.83	68.44	65.67	58.21	51.69	46.06
81	$50,000 to $59,999	n.a.	n.a.	46.36	51.81	55.62	56.80	58.71	60.72	73.11	80.61	76.21	67.91
82	$60,000 to $74,999	n.a.	n.a.	49.56	55.38	59.46	60.72	62.76	64.91	78.15	95.32	114.91	128.67
83	$75,000 to $99,999	n.a.	n.a.	42.26	47.22	50.70	51.77	53.52	55.35	66.63	81.28	97.98	115.71
84	$100,000 to $124,999	n.a.	n.a.	16.55	18.49	19.85	20.27	20.95	21.67	26.09	31.82	38.36	45.30
85	$125,000 to $149,999	n.a.	n.a.	6.45	7.21	7.74	7.90	8.17	8.45	10.17	12.40	14.95	17.66
86	$150,000 or More	n.a.	n.a.	9.33	10.42	11.19	11.43	11.81	12.22	14.71	17.94	21.63	25.54
87	Total Retail Sales (Million s1992 $)	7121.68	8610.34	10152.01	10689.19	10942.00	11006.26	11098.51	11210.67	11948.00	12876.30	13948.94	15113.71
88	Building Materials, Hardware	215.61	341.17	466.02	543.73	559.89	559.81	562.53	567.00	601.52	646.25	695.43	748.74
89	General Merchandise	933.92	961.34	1130.96	1196.11	1225.91	1240.73	1258.62	1278.76	1399.54	1539.82	1689.02	1852.61
90	Food Stores	1707.17	1911.50	2093.08	1984.30	1981.15	1975.68	1974.77	1977.05	2016.50	2082.18	2154.49	2225.81
91	Automobile Dealers	1259.45	1519.85	2040.31	2222.58	2331.06	2346.50	2367.01	2391.27	2548.72	2755.14	3021.18	3296.13
92	Gasoline Service Stations	516.04	736.26	632.07	624.73	651.07	658.41	666.07	673.96	716.74	767.14	825.21	894.76
93	Apparel and Accessories	458.69	349.56	476.64	521.47	523.35	526.29	530.39	535.34	568.25	611.14	661.43	721.30
94	Furniture, Home Furnishings	332.58	430.33	661.12	838.86	859.05	865.52	873.87	883.63	946.00	1022.92	1119.77	1222.62
95	Eating and Drinking Places	627.78	885.59	1143.58	1222.86	1225.92	1244.46	1266.08	1289.94	1431.16	1599.96	1790.16	2009.82
96	Drug Stores	367.54	381.34	508.68	519.73	543.89	552.91	562.69	572.95	626.75	680.44	730.04	777.85
97	Miscellaneous Retail Stores	702.92	1093.40	999.56	1014.83	1040.73	1035.94	1036.48	1040.77	1092.82	1171.31	1262.22	1364.07

Note Historical employment, earnings, and income data, 1969-1995, and population data, 1969-1996, is from U.S. Dept. of Commerce (USDoC); retail sales data is historical for 1972, 1977, 1982, 1987, and 1992 from USDoC; household data is historical for 1970, 1980, 1985 and 1990 from USDoC; Hispanic data is historical for 1970, 1980, and 1990-1996 from USDoC; all other years of data, 1969-1995, for retail sales, households, and Hispanic population are estimated by Woods & Poole (W&P), "e" indicates withheld employment and earnings data estimated by W&P; all data, 1996-2020 (1997-2020 for population), is projected by W&P; residential population as of July 1; Hispanic population is persons of Spanish origin regardless of race - White, Black, and Other sum to total population; other is the sum of Asian Americans, Native Americans, and Pacific Islanders; employment in number of jobs includes proprietors and part-time jobs; earnings by industry are earnings of workers - the sum of wages & salaries, proprietors income, and other labor income, historical data on households by money income bracket is for 1990 only (from U.S. Dept. of Commerce) and is not consistent with mean household total personal income data; the projections, 1991-2020, of households by money income bracket (in 1990 $) are based on 1990 data only; please read "Technical Description of the 1998 Regional Projections and Database" (Chapter 2 of this report) for an explanation of data sources, data definitions, and projection methods. H ISTORICAL DATA IS SUBJECT TO REVISION; PROJECTIONS ARE UNCERTAIN AND FUTURE DATA MAY DIFFER SUBDTANTIALLY FROM THESE PROJECTIONS. WOODS & POOLE DDOES NOT GUARANTEE THE ACCURACY OF THE PROJECTION OR HISTORICAL DATA CONTAINED IN THIS TABLE. THIS DATA IS PROVIDED SUBJECT TO ALL TERMS AND CONDITIONS OF THE WOODS & POOLE ECONOMICS, INC. END USER LICENSE AGREEMENT AND IS NOT AUTHORIZED FOR USE IN LEGAL OR FINANCIAL TRANSACTIONS.

ALAMEDA, CA | Comparative Data Table

Unit of Geography: County
FIPS Code: 06001

Selected Calculations
Rank of ALAMEDA, CA Among all 58 Counties in CA – In Parentheses

	1970-1995	1980-1995	1995-2020
Population Growth Rate	+0.83% (54)	+1.18% (45)	+0.63% (44)
Employment Growth Rate	+1.85% (47)	+1.84% (37)	+1.13% (36)
Population Change (Thousands)	+248.02 (12)	+213.19 (9)	+224.85 (11)
Employment Change (Thousand)	+291.19 (8)	+189.39 (8)	+257.67 (8)

	1970	1980	1990	2000	2010	2020
Percent of Population Age 0-19	34.77% (43)	28.74% (41)	21.19% (45)	27.56% (46)	26.67% (43)	25.60% (40
Percent of Population Age 65 and Over	9.19% (35)	10.34% (34)	10.57% (39)	11.39% (35)	11.87% (34)	14.77% (32)
Percent of Population White	79.84% (56)	72.33% (57)	65.26% (57)	60.37% (57)	55.69% (57)	51.83% (57)
Percent of Population Black	15.10% (1)	18.56% (1)	18.30% (1)	18.03% (1)	17.29% (1)	16.29% (1)
Percent of Population Hispanic (Any Race)	12.57% (21)	11.41% (27)	14.31% (27)	17.85% (28)	19.78% (29)	22.19% (29)

	1970	1980	1990	2000	2010	2020
Percent of Jobs in Manufacturing	17.18% (13)	14.49% (10)	11.15% (16)	10.43% (16)	9.63% (17)	8.82% (18)
Percent of Jobs in Services	18.49% (27)	23.99% (21)	29.21% (17)	35.06% (15)	38.59% (15)	41.83% (12)
Percent of Jobs in Framing	0.75% (55)	0.36% (55)	0.19% (56)	0.15% (55)	0.12% (55)	0.10% (55)
Percent of Jobs in Government	25.54% (18)	20.71% (21)	19.01% (22)	15.79% (25)	14.44% (33)	13.33% (35)

	1970	1980	1990	2000	2010	2020
Population Rank	4	5	6	7	7	7
Income per Capita Rank	9	11	8	7	7	7
Mean Household Income Rank	15	19	12	12	12	11
Retail Sales per Household Rank	32	27	24	24	24	24

Note: Average annual rate of growth in percent; historical data, 1970-1995, from U.S. Dept. of Commerce; projected data, 1996-2020, from Woods & Poole Economics, Inc.; retail sales, household, population by age, and Hispanic population data are estimated; government is Federal, military, and state and local. please read "Technical Description of the 1998 Regional Projections and Database" (Chapter 2 of this report) for an explanation of data sources, data definitions, and projection methods. H ISTORICAL DATA IS SUBJECT TO REVISION; PROJECTIONS ARE UNCERTAIN AND FUTURE DATA MAY DIFFER SUBDTANTIALLY FROM THESE PROJECTIONS. WOODS & POOLE DDOES NOT GUARANTEE THE ACCURACY OF THE PROJECTION OR HISTORICAL DATA CONTAINED IN THIS TABLE. THIS DATA IS PROVIDED SUBJECT TO ALL TERMS AND CONDITIONS OF THE WOODS & POOLE ECONOMICS, INC. END USER LICENSE AGREEMENT AND IS NOT AUTHORIZED FOR USE IN LEGAL OR FINANCIAL TRANSACTIONS.

Labor Force and Unemployment

	1994	1995	1996	8/97	9/97	10/97
Civilian Labor Force (Thousands)	683.33	682.17	684.76	701.01	697.83	701.63
Employed	641.51	642.88	650.76	667.81	665.03	670.88
Unemployed	41.81	39.30	34.00	33.21	32.80	30.75
Unemployment Rate	6.1%	5.8%	5.0%	4.7%	4.7%	4.4%

Note: Historical Labor force, employment and unemployment data is from Bureau of Labor Statistics; monthly data is not seasonally adjusted and is subject to revision; employment by place of residence, not by place of work; employment data excludes proprietors and government workers.

Private Non-Farm Establishments by Size

	1994 Total	1994 1 to 49 Employees	1994 50 or more Employees	1995 Total	1995 1 to 49 Employees	1995 50 or more Employees
Total Number of Business Establishments	33301	31366	1935	33461	31427	2034
Agricultural Services, Forestry, and Fishing	399	392	7	413	404	9
Mining	19	14	5	19	14	5
Construction	2509	2412	97	2463	2365	98
Manufacturing	2541	2167	374	2576	2169	407
Transportation and Public Utilities	1337	1191	146	1342	1193	149
Wholesale Trade	3272	3074	198	3313	3097	216
Retail Trade	7003	6594	409	6927	6521	406
Finance, Insurance, and Real Estate	3067	2973	94	2966	2880	86
Service	12930	12325	605	13110	12453	657
Unclassified Establishments	224	224	0	332	331	1

Note: Historical data from U.S. Dept. of Commerce; data excludes proprietors and government workers; industry classifications based on 1987 SIC definitions; unclassified establishments are businesses that cannot by classified in any industry group because of insufficient kind-of-business information; statewide establishments are assigned proportionally to counties within state by Woods and Poole

Composition of Hispanic and Other Population by Race

	1990	1996
Percent of Hispanic Population, Race White	85.78%	86.18%
Percent of Hispanic Population, Race Black	5.59%	6.00%
Percent of Hispanic Population, Race Native American	2.12%	1.91%
Percent of Hispanic Population, Race Asian and Pacific Islander	6.51%	5.91%
Percent of Other Population, Race Asian and Pacific Islander	94.94%	95.78%
Percent of Other Population, Race Native American	5.06%	4.22%

Note: Hispanic population can be of any race - the percentages in this table indicate the proportion of the Hispanic population by race; other population is the sum of Native Americans, Asians and Pacific Islanders - the percentages in this table indicate the proportion of other population by its components; percentages for Other include Hispanic Asians/Pacific Islanders and Hispanic Native Americans. historical data from 1990 Census modified age, race, and sex data and 1996 Census Bureau estimates.

Educational Attainment

	1970	1980	1990
Percent of Population Age 25+, Not Completing High School	37.0%	24.0%	18.6%
Percent of Population Age 25+, Completing High School Only	48.4%	53.7%	52.6%
Percent of Population Age 25+, Completing 4 Years of College or More	14.6%	22.3%	28.8%

Note: Educational attainment as percent of population age 25 and over is from 1970, 1980, and 1990 Census of Population; data is based on self-reporting by Census respondents.

Appendix 3.2

Illustration of Freight Model Classification Framework

We began reviewing freight models found in the literature according to the structure found in Table A3. We tried to develop a structure that would be useful for determining the appropriateness of a model for operational use in public planning issues. We note that we began developing our approach before finding the reviews discussed in Section 3. Therefore, we limit the entries in Table A3 to those models discussed in Part II of the report.

The columns of the table represent different aspects that we felt would be important when evaluating a model for use in an operational system. The first column indicates the specific model analyzed. The second column represents whether the model is descriptive or normative. The distinction is based on the anticipated use of the freight movement model. We use "descriptive" to mean a model that describes or forecasts some aspect of freight movement under different "what if" scenarios. That is, a descriptive model generates forecasts of movements or behavior, i.e., describes what the movements or behavior will be. On the other hand, a "normative" model suggests to decision makers what to do based on some norms or objectives. For example, an objective may be selecting a route for transporting freight between cities that minimizes the cost of shipments. The output of the model would be the route that should be followed to achieve this cost minimizing objective. This suggested route may or may not be followed, and, therefore, the model does not attempt to describe what the freight flows will be. It is possible, however, that the modeler would intend to use this cost minimizing objective to predict where shipments will travel. If knowing the route that minimizes cost is intended to be used to tell a decision maker how a shipment *should be* routed, the model would be classified as normative; if it is intended to be used by planners to predict or describe how a shipment *will be* routed, it would be classified as descriptive. Since we are primarily interested in models that will predict freight movements under different conditions, we mostly sought descriptive models in our literature search and limit the entries of Table A3 to descriptive models.

Table A3 Model classification table with selected entries

Name	D/N	Modes	Input		Output		Methodology	Comments
			General	Specific	General	Specific		
Weinblatt and Edwards 1997	D	Intermodal facility	Regional volumes Facility information Distances	Base-year volumes generated in or destined for each region Operational information (e.g., number of destinations served, capacity, cost, frequency of service) of facilities Highway distances between facilities and regional centroids	Freight flows between each region and facility	Freight volumes for existing and new facilities	Variation of gravity model	Applied in analysis of new air cargo facility considered in North Carolina
Sheffi 1985	D	Conceivably any mode	Network configuration Link performance functions Link counts Target OD matrix	Network topology Functions relating segment impedance (e.g., time) to segment volumes Observed volumes for all links. "Best-guess" OD flows for same specified time period	Updated OD matrix Segment volumes	OD matrix that maximizes the match with the target OD matrix and the correspondence between assigned and observed segment volumes Assigned segment volumes	Mathematical programming model incorporating traffic assignment	Developed in context of passenger vehicles, but applicable to any mode General concepts provided, but not specific steps
Park and Smith Jr. 1996	D	Truck	Trip generation rates Network configuration Zonal distances Segment travel times Link counts OD data	Generation rate based on population Network Topology Zone to zone distance matrix Free flow segment travel times Observed flows on selected links OD estimates for selected zonal pairs	OD matrix Trip productions and attractions Segment volumes	OD matrix expanded to encompass all zone pairs Trip productions and attractions Assigned volumes on all segments	Based on gravity model and all-or-nothing traffic assignment	Small scale survey is needed Wisconsin DOT supported research study, but not currently used operationally
Nielsen (TransCAD® 3.0) 1994	D	Conceivably any mode	Network configuration Link performance functions Link counts Target OD matrix	Network topology Functions relating segment impedance (e.g., time) to segment volumes Observed volumes for all links. "Best-guess" OD flows for same specified time period	Updated OD matrix Segment volumes	OD matrix that maximizes the match with the target OD matrix and the correspondence between assigned and observed segment volumes Assigned segment volumes	Commercial software package	Available in a GIS software package

The modes involved in the freight movement model are shown in the third column. Most of the freight movement models we saw dealt with truck trips. Furthermore, statistics indicate that truck trips account for approximately 75% of freight movements produced in the United States. (Bureau of Transportation Statistics, 1994). However, most of the models considered in Part II of this report and, therefore, summarized in Table A3 are more general.

The fourth and fifth columns indicate the inputs to the model. We address the inputs at two levels: general (fourth column) and specific (fifth column). The general inputs give an idea of the type of data required. Further specification would be required, however, until the data could be quantified in a form that could be used by the model. This specification is provided in the fifth column. For example, an origin-destination table would describe input at a general level. However, one would need to know the freight category (e.g., commodity group), units of measurement (e.g., vehicle loads, tons, dollar value, ...), and period of measurement (e.g., per day, per quarter, per year, ...) before being able to estimate input for the model.

The sixth and seventh columns indicate the model outputs. Again, we consider the outputs at general (sixth column) and specific (seventh column) levels, where the distinction is similar to that made between general and specific inputs. For example, link flows would be a general output, whereas trucks of five or more axles per day on a highway link would add the specification necessary to understand how the model output could eventually be used.

The eighth column provides a general description of the methodology underlying the transformation of inputs to outputs. The methods are almost all based on either simulation, econometric fits to data, or mathematical programming (optimization) techniques.

The final column allows for miscellaneous comments.

To demonstrate how to interpret this model review table, consider the first entry in Table A3. The first column provides the reference for this model. It can be found in our list of references under Weinblatt and Edwards (1997). The "descriptive" annotation in the second column indicates that the model is intended to predict freight movements that will occur--freight volumes from various regions to specific intermodal facilities, in this case--as opposed to being intended to dictate the movements that should occur to satisfy some objective.

According to column three, this model considers freight transferred at an intermodal facility. Inputs of the model are shown in the next two columns. The general inputs are essentially regional freight volumes, intermodal facility information, and distances between regions and facilities. More specifically, base-year annual volumes generated in or destined for each region considered, level-of-service characteristics of the intermodal facilities (numbers of destinations served and capacity are used as examples), and highway distances between regional centroids and the facilities are used as inputs. The sixth and seventh columns contain the general and specific outputs generated. In this case, the general and specific outputs are essentially the same, namely, the region-facility volumes for a given year. The methodology uses a variation of the gravity model, as seen in the eighth column. In the final column we note that this model was applied to predict freight that would use a new air-based intermodal facility in North Carolina.

Appendix 6

Tables for Calculations of Section 6

Table A6.1a Estimated LOS scores in Data Fit Scenario 1

Airports	LOS
Airborne Airpark	10.0
Dayton	7.9
Toledo	5.5
Cincinnati	3.0
Cleveland-Hopkins	1.7
Columbus	0.6
(Rickenbacker)	(3.0)

Table A6.1b Freight forecasts at facilities in Data Fit Scenario 1, without Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	57,248.62	70,192.57	73,813.40	70,944.66	57,518.84	329,718.10	35.6%
Dayton	39,363.72	55,452.13	58,312.59	56,046.28	47,537.10	256,711.80	27.8%
Toledo	52,477.90	26,638.08	21,110.63	26,533.30	48,669.78	175,429.70	19.0%
Cincinnati	13,039.96	20,004.88	22,144.02	21,283.40	13,804.52	90,276.78	9.8%
Cleveland-Hopkins	18,022.71	8,233.59	5,144.79	5,668.48	13,539.05	50,608.62	5.5%
Columbus	4,579.89	4,211.55	4,207.36	4,256.68	3,663.51	20,918.99	2.3%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Table A6.1c Freight forecasts at facilities in Data Fit Scenario 1, with Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	52,711.08	63,010.03	68,099.24	65,650.09	52,953.04	302,423.50	32.7%
Dayton	36,243.75	49,777.92	53,798.40	51,863.57	43,763.65	235,447.30	25.5%
Toledo	48,318.49	23,912.31	19,476.38	24,553.14	44,806.42	161,066.70	17.4%
Cincinnati	12,006.41	17,957.86	20,429.77	19,695.03	12,708.73	82,797.80	9.0%
Cleveland-Hopkins	16,594.23	7,391.08	4,746.52	5,245.44	12,464.33	46,441.59	5.0%
Columbus	4,216.89	3,780.60	3,881.66	3,939.01	3,372.70	19,190.85	2.1%
Rickenbacker	14,641.97	18,903.01	14,300.84	13,786.52	14,663.92	76,296.25	8.3%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Table A6.2a Estimated LOS scores in Data Fit Scenario 2

Airports	LOS
Airborne Airpark	10.5
Dayton	7.9
Toledo	5.3
Cincinnati	3.0
Cleveland-Hopkins	1.7
Columbus	0.6
(Rickenbacker)	(3.0)

Table A6.2b Freight forecasts at facilities in Data Fit Scenario 2 , without Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	59,802.17	72,702.19	76,296.83	73,464.92	60,035.30	342,301.40	37.1%
Dayton	39,161.46	54,699.74	57,404.28	55,273.61	47,254.16	253,793.30	27.5%
Toledo	50,309.76	25,321.13	20,026.10	25,215.96	46,620.82	167,493.80	18.1%
Cincinnati	12,972.96	19,733.45	21,799.10	20,989.98	13,722.35	89,217.84	9.7%
Cleveland-Hopkins	17,930.10	8,121.87	5,064.66	5,590.33	13,458.46	50,165.43	5.4%
Columbus	4,556.36	4,154.41	4,141.83	4,198.00	3,641.70	20,692.29	2.2%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Table A6.2c Freight forecasts at facilities in Data Fit Scenario 2 , with Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	55,084.67	65,353.58	70,475.40	68,052.29	55,295.88	314,261.80	34.0%
Dayton	36,072.20	49,170.79	53,024.35	51,201.25	43,523.73	232,992.30	25.2%
Toledo	46,341.07	22,761.72	18,498.11	23,358.14	42,940.39	153,899.40	16.7%
Cincinnati	11,949.58	17,738.83	20,135.83	19,443.51	12,639.06	81,906.81	8.9%
Cleveland-Hopkins	16,515.69	7,300.93	4,678.22	5,178.46	12,396.00	46,069.29	4.9%
Columbus	4,196.93	3,734.49	3,325.81	3,888.70	3,354.21	19,000.14	2.1%
Rickenbacker	14,572.66	18,672.45	14,095.08	13,610.46	14,583.53	75,534.18	8.2%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Table A6.3a Estimated LOS scores in Data Fit Scenario 3

Airports	LOS
Airborne Airpark	10.0
Dayton	4.80
Toledo	1.17
Cincinnati	-2.56
Cleveland-Hopkins	-4.49
Columbus	-7.51
(Rickenbacker)	(-2.56)

Table A6.3b Freight forecasts at facilities in Data Fit Scenario 3, without Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	68,048.41	67,535.61	73,588.78	69,833.31	66,542.35	345,548.50	37.4%
Dayton	41,992.35	49,987.38	54,467.72	51,688.06	46,794.47	244,930.00	26.5%
Toledo	44,943.67	27,254.60	23,442.36	27,832.69	45,051.91	168,525.20	18.2%
Cincinnati	6,784.34	23,419.19	27,357.95	25,961.79	10,628.06	94,151.33	10.2%
Cleveland-Hopkins	24,344.01	8,135.61	-1,437.63	730.74	18,185.34	49,958.06	5.4%
Columbus	-1,379.97	8,400.42	7,313.62	8,686.22	-2,469.33	20,550.96	2.2%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Table A6.3c Freight forecasts at facilities in Data Fit Scenario 3, with Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	64,302.60	59,454.92	67,615.50	64,431.77	61,900.05	317,704.80	34.4%
Dayton	39,680.83	44,006.35	50,046.52	47,690.04	43,529.87	224,953.60	24.4%
Toledo	42,469.69	23,993.56	21,539.52	25,679.86	41,908.88	155,591.50	16.8%
Cincinnati	6,410.89	20,617.06	25,137.28	23,953.67	9,886.60	86,005.50	9.3%
Cleveland-Hopkins	23,003.96	7,162.18	-1,320.94	674.21	16,916.65	46,436.06	5.0%
Columbus	-1,304.01	7,395.30	6,719.97	8,014.35	-2,297.05	18,528.55	2.0%
Rickenbacker	10,168.83	22,103.44	14,994.95	14,288.91	12,887.82	74,443.94	8.1%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Table A6.4a Estimated LOS scores in Data Fit Scenario 4

Airports	LOS
Airborne Airpark	23.31
Dayton	14.09
Toledo	7.46
Cincinnati	1.09
Cleveland-Hopkins	-2.75
Columbus	-6.69
(Rickenbacker)	(1.09)

Table A6.4b Freight forecasts at facilities in Data Fit Scenario 4, without Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	68,955.73	68,513.09	72,126.67	69,914.18	68,058.34	347,568.00	37.6%
Dayton	45,134.40	49,552.70	52,166.26	50,566.06	47,697.46	245,116.90	26.5%
Toledo	39,538.43	29,540.78	27,418.03	29,935.01	39,866.90	166,299.10	18.0%
Cincinnati	12,468.15	21,783.70	24,015.23	23,278.56	14,363.04	95,908.69	10.4%
Cleveland-Hopkins	17,413.89	8,537.08	2,924.82	4,094.37	14,270.17	47,240.32	5.1%
Columbus	1,222.19	6,805.45	6,081.80	6,944.62	476.90	21,530.96	2.3%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Table A6.4c Freight forecasts at facilities in Data Fit Scenario 4, with Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	63,895.62	60,982.55	65,878.86	64,028.16	62,574.51	317,359.70	34.4%
Dayton	41,822.34	44,106.18	47,647.47	46,308.94	43,854.22	223,739.10	24.2%
Toledo	36,637.02	26,293.84	25,043.00	27,414.80	36,654.61	152,043.30	16.5%
Cincinnati	11,553.22	19,389.37	21,934.96	21,318.76	13,205.74	87,402.04	9.5%
Cleveland-Hopkins	16,136.02	7,598.73	2,671.47	3,749.66	13,120.34	43,276.23	4.7%
Columbus	1,132.51	6,057.44	5,554.97	6,359.96	438.48	19,543.35	2.1%
Rickenbacker	13,556.08	20,304.69	16,002.06	15,552.53	14,884.91	80,300.27	8.7%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Table A6.5a Estimated LOS scores in Data Fit Scenario 5

β's	Value	Airports	LOS
β_{FREQ}	-1.3E-06	Airborne Airpark	0.3855
β_{NDES}	0.004492	Dayton	0.3240
β_{DIST}	-0.00035	Toledo	0.2284
		Cincinnati	0.1342
		Cleveland-Hopkins	0.1114
		Columbus	0.0574
		(Rickenbacker)	(0.1342)

Table A6.5b Freight forecasts at facilities in Data Fit Scenario 5, without Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	64,111.44	64,605.75	66,043.82	65,051.59	62,440.81	322,253.42	34.9%
Dayton	49,444.34	51,470.70	56,067.73	56,381.57	51,502.12	264,866.45	28.7%
Toledo	37,750.41	31,197.88	29,618.37	31,253.97	42,722.90	172,543.52	18.7%
Cincinnati	9,975.54	16,943.58	24,227.65	20,527.98	12,757.72	84,432.47	9.1%
Cleveland-Hopkins	21,442.05	10,844.72	5,005.43	6,352.98	13,734.15	57,379.32	6.2%
Columbus	2,009.02	9,670.17	3,769.80	5,164.72	1,575.11	22,188.82	2.4%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Table A6.5c Freight forecasts at facilities in Data Fit Scenario 5, with Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	59,969.61	57,868.12	60,519.10	59,801.64	57,923.43	296,081.90	32.1%
Dayton	46,250.05	46,102.90	51,377.53	51,831.32	47,776.12	243,337.92	26.3%
Toledo	35,311.59	27,944.31	27,140.72	28,731.63	39,632.04	158,760.29	17.2%
Cincinnati	9,331.09	15,176.56	22,200.95	18,871.28	11,834.74	77,414.62	8.4%
Cleveland-Hopkins	20,056.81	9,713.74	4,586.71	5,840.26	12,740.53	52,938.06	5.7%
Columbus	1,879.23	8,661.68	3,454.45	4,747.90	1,461.16	20,204.43	2.2%
Rickenbacker	11,934.42	19,265.49	15,453.33	14,908.76	13,364.79	74,926.79	8.1%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Table A6.6a Estimated LOS scores in Data Fit Scenario 6

α = 0.053714	Airports	LOS
	Airborne Airpark	10.0
	Dayton	3.58
	Toledo	-3.11
	Cincinnati	-13.22
	Cleveland-Hopkins	-26.39
	Columbus	-41.51
	(Rickenbacker)	-13.22

Table A6.6b Freight forecasts at facilities in Data Fit Scenario 6, without Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	65,957.70	71,337.53	72,665.54	71,558.40	65,721.18	347,240.30	37.6%
Dayton	45,004.24	50,540.02	51,480.87	50,696.50	47,317.42	245,039.10	26.5%
Toledo	39,569.32	29,861.77	27,763.16	29,793.76	39,216.21	166,204.20	18.0%
Cincinnati	17,673.84	19,954.72	20,879.49	20,561.37	17,610.46	96,679.87	10.5%
Cleveland-Hopkins	11,960.24	8,554.00	7,496.56	7,624.14	10,646.13	46,281.08	5.0%
Columbus	4,567.47	4,484.75	4,447.18	4,498.63	4,221.40	22,219.43	2.4%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Table A6.6c Freight forecasts at facilities in Data Fit Scenario 6, with Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	59,939.14	64,212.53	66,288.34	65,365.77	59,772.88	315,578.70	34.2%
Dayton	40,897.66	45,492.22	46,962.86	46,309.25	43,034.81	222,696.80	24.1%
Toledo	35,958.67	26,879.26	25,326.64	27,215.42	35,666.83	151,046.80	16.4%
Cincinnati	16,061.12	17,961.69	19,047.08	18,781.99	16,016.57	87,868.46	9.5%
Cleveland-Hopkins	10,868.89	7,699.65	6,838.66	6,964.35	9,682.57	42,054.11	4.6%
Columbus	4,150.69	4,036.82	4,056.89	4,109.33	3,839.33	20,193.06	2.2%
Rickenbacker	16,856.63	18,450.63	16,212.33	15,986.69	16,719.82	84,226.10	9.1%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Table A6.7a Estimated β's and LOS scores in Data Fit Scenario 7

β's	Value	Airports	LOS
β_{FREQ}	3.41E-05	Airborne Airpark	2.951
β_{NDES}	0.0128	Dayton	2.605
β_{DIST}	-0.00175	Toledo	2.239
		Cincinnati	1.706
		Cleveland-Hopkins	0.998
		Columbus	0.193
		(Rickenbacker)	1.706

Table A6.7b Freight forecasts at facilities in Data Fit Scenario 7, without Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	67,443.76	73,445.73	70,555.71	70,650.72	64,875.64	346,971.60	37.6%
Dayton	44,907.02	48,477.36	51,725.59	53,172.87	46,736.35	245,019.20	26.5%
Toledo	37,393.40	30,242.46	28,053.20	29,091.56	41,374.81	166,155.40	18.0%
Cincinnati	17,477.15	18,866.67	22,754.27	20,157.94	17,563.49	96,819.51	10.5%
Cleveland-Hopkins	12,880.24	8,744.69	7,432.06	7,442.06	9,955.47	46,454.52	5.0%
Columbus	4,631.24	4,955.89	4,211.98	4,217.65	4,227.04	22,243.80	2.4%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Table A6.7c Freight forecasts at facilities in Data Fit Scenario 7, with Rickenbacker

	Region I (Cleveland)	Region II (Columbus)	Region III (Cincinnati)	Region IV (Dayton)	Region V (Toledo)	Total	Proportion
Airborne Airpark	61,230.82	65,784.86	64,099.58	64,430.67	59,013.96	314,559.90	34.1%
Dayton	40,770.17	43,420.85	46,992.49	48,491.56	42,513.60	222,188.70	24.1%
Toledo	33,948.70	27,087.97	25,486.22	26,530.36	37,636.49	150,689.70	16.3%
Cincinnati	15,867.15	16,898.75	20,672.16	18,383.25	15,976.58	87,797.89	9.5%
Cleveland-Hopkins	11,693.71	7,832.56	6,751.99	6,786.87	9,055.97	42,121.10	4.6%
Columbus	4,204.61	4,438.96	3,826.57	3,846.33	3,845.12	20,161.58	2.2%
Rickenbacker	17,017.64	19,268.85	16,903.79	16,263.77	16,691.07	86,145.12	9.3%
Total	184,732.8	184,732.8	184,732.8	184,732.8	184,732.8	923,664	100.0%

Appendix 7.1

Description and Illustration of Method 1 for Updating OD Matrices from Ground Counts

We translate the general description of the method presented by Sheffi (1985) into a set of algorithmic steps, and illustrate this method with a small example that can be solved using a spreadsheet. All steps could be programmed, however, and real problems could be readily solved with the method.

A7.1.1 Method Description

The program can be formulated as the solution to the following minimization program (Sheffi, 1985):

$$\min_{q} Z(\mathbf{x},\mathbf{q}) = \sum_{rs} (q_{rs} - \hat{q}_{rs})^2 \qquad \text{(A7.1.1)}$$

subject to

$$\sum_{k} f_k^{rs} = q_{rs} \qquad \forall\, r,s \qquad \text{(A7.1.2)}$$

$$f_k^{rs} \geq 0 \qquad \forall\, k,r,s \qquad \text{(A7.1.3)}$$

$$\sum_{a} \int_0^{x_a} t_a(\omega)d\omega - \sum_{rs} \tilde{u}_{rs} q_{rs} \leq \eta^* \qquad \text{(A7.1.4)}$$

where q_{rs} is the total flow from origin r to destination s; $\hat{q}_{rs}$ is the total flow from origin r to destination s in the target OD matrix; f_k^{rs} is the flow on path k from origin r to destination s; x_a is the flow on link a; $t_a(w)$ is the link performance function on arc a, which determines the time to traverse the link when the flow is w; $\tilde{u}_{rs}$ is the "observed" shortest travel time between origin r and destination s; η^* is the optimal (minimum) value of $\sum_{a} \int_0^{x_a} t_a(\omega)d\omega - \sum_{rs} \tilde{u}_{rs} q_{rs}$. By optimal, we mean the lowest possible value that satisfies the general flow constraints (see Part A below).

This program can be solved by decomposing it into two parts (Sheffi, 1985). In the first part, which we call Part A, a minimization program is solved to find η^*, the right hand side of constraint (A7.1.4). In what we call Part B, the value of η^* is used in the minimization program above. The specific steps corresponding to Part A and Part B can be stated as follows.

Part A. The minimization program to find η^* is:

$$\text{Min } Z(x,q)=\sum_a \int_0^{x_a} t_a(\omega)d\omega - \sum_{rs} \tilde{u}_{rs} q_{rs} \quad \text{(A7.1.5)}$$

subject to

$$\sum_k f_k^{rs} = q_{rs} \qquad \forall\, r,s \quad \text{(A7.1.6)}$$

$$f_k^{rs} \geq 0 \qquad \forall\, k,r,s \quad \text{(A7.1.7)}$$

The notation is that given above. In addition to the network structure and link performance functions, the only input to this program is the set of observed link flows $\{\tilde{x}_a\}$. The solution to this program produces a set of "assigned" link and OD flows denoted as $\boldsymbol{x}^*$ and $\boldsymbol{q}^*$ respectively. We then set $\eta^*=\sum_a \int_0^{x_a^*} t_a(\omega)d\omega - \sum_{rs} \tilde{u}_{rs} q_{rs}^*$ and calculate its value.

To solve the part A minimization program, one can use the following algorithm.

Part A algorithm

Initialization:

Step 1. Set an upper bound OD flow $\bar{q}_{rs}$ for each OD pair. This upper bound should be set high enough so that it would never be binding (i.e., the upper bound should be greater than the possible maximum OD flow for each OD pair).

Step 2. Arbitrarily generate an OD matrix $\{q_{rs}^1\}$ as the initial OD matrix to be assigned to the network.

Step 3. Use the observed link flow $\tilde{x} = \{\tilde{x}_a\}$ and link performance functions (e.g. BPR function given in equation (7.1)) to calculate "observed link travel times", $\tilde{t} = \{\tilde{t}_a\}$.

Step 4. Based on "observed link times" $\tilde{t}$ find the minimum "observed OD travel time" $\tilde{u}_{rs}$ between each origin-destination pair rs.

Step 5. Find the minimum "free flow travel path" $u_{rs,0}$ between each origin-destination pair rs. This is the path found using free-flow link travel time to,a.

Step 6. Set iteration number n to 1. Assign the set of OD flows of $\{q_{rs}^n\}$ on the paths found in Step 5 and obtain a set of assigned link flows $\{x_a^n\}$ where n denotes the iteration number.

Calculate link travel times:

Step 7. Use the link performance functions and the assigned flows $\{x_a^n\}$ to determine the assigned link travel times $\{t_a^n\}$.

Calculate auxiliary OD and link flows

Step 8. Based on the assigned link travel times $\{t_a^n\}$ of Step 7, find the minimum assigned travel time path m and corresponding time $u_{rs,m}^n$ for each OD pair rs.

Step 9. Compare the assigned OD travel time $u_{rs,m}^n$ from Step 8 with the "observed" travel time $\tilde{u}_{rs}$ from Step 4 for each OD pair. If the assigned shortest path time $u_{rs,m}^n$ is smaller than the observed time $\tilde{u}_{rs}$ for pair rs, set the corresponding auxiliary path flow equal to the OD flow upper bound set in step 1, and set the auxiliary path flows for other paths equal to 0, i.e. $g_m^{rs^n} = \bar{q}_{rs}$, $g_k^{rs^n} = 0 \ \forall \ k \neq m$. If the assigned shortest path time $u_{rs,m}^n$ is greater than the observed time $\tilde{u}_{rs}$, set all auxiliary path flows to zero, i.e., $g_k^{rs^n} = 0 \ \forall \ k$.

Step 10. Calculate the auxiliary OD flows υ_{rs}^n from origin r to destination s at iteration n and auxiliary link flows y_a^n for arc a at iteration n as:

$$\upsilon_{rs}^n = \sum_k g_k^{rs^n} \qquad \forall \text{ r, s}$$

$$y_a^n = \sum_k \sum_{rs} g_k^{rs^n} \delta_{a \cdot k}^{rs} \qquad \forall \text{ a} \qquad \text{(A7.1.8)}$$

where $\delta_{a,k}^{rs}$ is a indicator variable--- $\delta_{a,k}^{rs}=1$ if arc a is on path k between r and s; $\delta_{a,k}^{rs}=0$ otherwise.

Step 11. Find the value of α ($0 \le \alpha \le 1$) that minimizes

$$Z(\alpha)=\sum_a \int_0^{x_a^n+\alpha(y_a^n-x_a^n)} t_a(\omega)d\omega - \sum_{rs} \tilde{u}_{rs}(q_{rs}^n+\alpha(\upsilon_{rs}^n-q_{rs}^n)) \tag{A7.1.9}$$

where $\{x_a^n\}$ was found in Step 6, $\{q_{rs}^n\}$ was chosen in Step 2, and $\{\upsilon_{rs}^n\}$and $\{y_a^n\}$ were found in Step 10. Denote the value of α that minimizes (A7.1.9) by α^n.

Update flows

Step 12. Use the value of α^n found in Step 11 to update the link and OD flows:

$$x_a^{n+1}=x_a^n+\alpha^n(y_a^n-x_a^n), \qquad \forall a$$

$$q_{rs}^{n+1}=q_{rs}^n+\alpha^n(v_{rs}^n-q_{rs}^n), \qquad \forall rs \tag{A7.1.10}$$

Convergence test

Step 13. Check convergence.

If $Z(\mathbf{x}^n, \mathbf{q}^n)-Z(\mathbf{x}^{n+1}, \mathbf{q}^{n+1}) \le k$, set $x_a^*=x_a^{n+1}\ \forall a$; $q_{rs}^*=q_{rs}^{n+1}\ \forall$ rs, and stop,

if $Z(\mathbf{x}^n, \mathbf{q}^n)-Z(\mathbf{x}^{n+1}, \mathbf{q}^{n+1}) > k$, set n=n+1, and go to step 7;

where $Z(\mathbf{x}^n, \mathbf{q}^n)$ is the objective function in equation (A7.1.5), and k is a dimensionless convergence criterion set beforehand.

Once this algorithm terminates by passing the criterion test of Step 13, the set of flows $\{x^*_a\}$ and $\{q^*_{rs}\}$ are substituted in the right-hand side of equation (A7.1.5) to obtain $Z(\boldsymbol{x}^*,\boldsymbol{q}^*)$. The value of $Z(\boldsymbol{x}^*,\boldsymbol{q}^*)$ is denoted as η^*. Note that if the BPR function (equation (7.1)) is used for the link performance functions $t_a(\omega)$, solving the integral on the right-hand side of (A7.1.5) and, therefore, evaluating Z is straightforward.

Part B. By introducing a dual variable γ associated with constraint (A7.1.4), the original program (A7.1.1)-(A7.1.4) can be transformed into the partial Lagrangian problem:

$$\min L(\boldsymbol{x}, \boldsymbol{q}, \gamma)=\sum_{rs}(q_{rs}-\hat{q}_{rs})^2+\gamma[\sum_a \int_0^{x_a} t_a(\omega)d\omega - \sum_{rs}\tilde{u}_{rs}q_{rs}-\eta^*] \tag{A7.1.11}$$

subject to

$$\sum_k f_k^{rs} = q_{rs} \qquad \forall\, r,s \qquad \text{(A7.1.12)}$$

$$f_k^{rs} \geq 0 \qquad \forall\, k,r,s \qquad \text{(A7.1.13)}$$

$$\gamma \geq 0 \qquad \text{(A7.1.14)}$$

where γ is the dual variable associated with the (single) constraint (A7.1.4), and the other terms are defined above.

To solve the Part B minimization program, one can use the following algorithm. In addition to the network structure and link performance functions, the inputs to this program are the set of observed link flows $\{\tilde{x}_a\}$, the target matrix $\{\hat{q}_{rs}\}$ and the value of η^*. The solution to this program produces a set of estimated link and OD flows.

Part B algorithm

Initialization

Step 1. Set an upper bound of total possible flow $\bar{q}_{rs}$ for each OD pair. This upper bound should be set high enough so that it would never be binding (i.e., the upper bound should be greater than the possible maximum OD flow for each OD pair). Set the dual variable γ in equation (A7.1.11) to some positive value, $\gamma > 0$

Step 2. Use the target matrix $\{\hat{q}_{rs}\}$ as the initial OD matrix $\{qrs^1\}$ to be assigned to the network.

Step 3. Use the observed link flows $\{\tilde{x}_a\}$and link performance functions (e.g. BPR function given in equation (7.1)) to calculate "observed link travel times" $\{\tilde{t}_a\}$.

Step 4. Based on observed link times $\tilde{t}$, find the minimum observed travel time $\tilde{u}_{rs}$ between each origin-destination pair *rs*. These are the same observed times in the Part A algorithm.

Step 5. Find the shortest time path from every origin to every destination using free flow travel time to,a. This is the same free-flow path found in the Part A algorithm.

Step 6. Set the iteration number n to 1. Assign the OD flows $\{q^n_{rs}\}$ on the paths found in Step 5 and obtain the set of assigned link flows $\{x^n_a\}$, where n denotes the iteration number.

Calculate link travel times:

Step 7. Use the link performance functions and the assigned flows $\{x^n_a\}$ to acquire assigned link travel times $\{t^n_a\}$.

Calculate auxiliary OD and link flows

Step 8. Based on the assigned link travel times $\{t^n_a\}$ of Step 7, find the minimum assigned travel time path m and corresponding time on the path $u^n_{rs,m}$ for each OD pair.

Step 9. Compare the assigned OD travel time $u^n_{rs,m}$ from Step 8 with the observed travel time $\tilde{u}_{rs}$ from Step 4 for each OD pair. If the assigned shortest path time $u^n_{rs,m}$ is smaller than the observed time $\tilde{u}_{rs}$ for pair rs, set the corresponding auxiliary path flow equal to the OD flow upper bound set in Step 1, and set the auxiliary path flows for other paths equal to 0, i.e., $g^{rs^n}_m = \bar{q}_{rs}$, $g^{rs^n}_k = 0$ $\forall$ $k \neq m$. If the assigned shortest path time $u^n_{rs,m}$ is greater than the observed time $\tilde{u}_{rs}$, set all auxiliary path flows to zero, i.e., $g^{rs^n}_k = 0$ $\forall$ k.

Step 10. Calculate the auxiliary OD flows v^n_{rs} from origin r to destination s at iteration n and auxiliary link flows y^n_a for arc a at iteration n using equation (A7.1.8).

Step 11. Find the value of α $(0 \leq \alpha \leq 1)$ which minimizes

$$L(\alpha) = \sum_{rs} \left[q^n_{rs} + \alpha(v^n_{rs} - q^n_{rs}) - \hat{q}_{rs}\right]^2 +$$

$$\gamma[\sum_a \int_0^{x^n_a + \alpha(y^n_a - x^n_a)} t_a(\omega)d\omega - \sum_{rs} \tilde{u}_{rs}(q^n_{rs} + \alpha(v^n_{rs} - q^n_{rs})) - \eta^*] \qquad \text{(A7.1.15)}$$

where $\{x_a^n\}$ was found in Step 6, $\{q_{rs}^n\}$ was chosen in Step 2, $\{\upsilon_{rs}^n\}$ and $\{y_a^n\}$ were found in Step 10, $\{\hat{q}_{rs}\}$, $\{\tilde{u}_{rs}\}$ and the link performance function $\{t_a\}$ were given as input, γ was set in Step 1, and we found η^* in Part A. Denote the value of α that minimizes (A7.1.15) by α^n.

Update flows

Step 12. Use the value of α^n found in Step 11 to update the link and OD flows using equation (A7.1.10)

$$x_a^{n+1} = x_a^n + \alpha^n (y_a^n - x_a^n), \qquad \forall a$$

$$q_{rs}^{n+1} = q_{rs}^n + \alpha^n (v_{rs}^n - q_{rs}^n), \qquad \forall rs$$

Convergence test

Step 13. Check convergence.

If $L(\mathbf{x}^n, \mathbf{q}^n)-L(\mathbf{x}^{n+1}, \mathbf{q}^{n+1}) \le k$, set $x_a^* = x_a^{n+1}\ \forall a$; $q_{rs}^* = q_{rs}^{n+1}\ \forall$ rs, and stop;

If $L(\mathbf{x}^n, \mathbf{q}^n)-L(\mathbf{x}^{n+1}, \mathbf{q}^{n+1}) > k$, set n=n+1, and go to Step 7,

where $L(\mathbf{x}^n, \mathbf{q}^n)$ is the objective function in equation (A7.1.11), and k is a dimensionless convergence criterion set beforehand.

Find the value of γ that minimizes Z in equation (A7.1.11) subject to constraints (A7.1.12)-(A7.1.14)

Step 14. Given the set of flows $\{x^*_a(\gamma)\}$ and $\{q^*_{rs}(\gamma)\}$ from Step 13, calculate

$$dL(\gamma)/d\gamma = \sum_a \int_0^{x_a^*(\gamma)} t_a(\omega)d\omega - \sum_{rs} \tilde{u}_{rs} q_{rs}^*(\gamma) - \eta^* \qquad \text{(A7.1.16)}$$

Step 15. If $dL(\gamma)/d\gamma>0$, increase the value of γ and go to Step 1; if $dL(\gamma)/d\gamma<0$, decrease the value of γ, and go to Step 1 Part B. If $dL(\gamma)/d\gamma=0$(or close enough to zero according to some convergence criterion), then stop.

When $dL(\gamma)/d\gamma=0$, the set of flows $\{x^{n+1}{}_a\}$ and $\{q^{n+1}{}_{rs}\}$ constitute the solution.

A7.1.2 Example

To illustrate the approach we consider the same simple network shown in Figure 7.1.

Part A

Step 1. We choose upper bounds of 15 for all OD pairs, including those with zero OD flow in the target matrix, i.e. $\bar{q}_{rs}$=15, ∀ rs. We specify the convergence criterion k=0.01.

Step 2. We arbitrarily generate an initial OD matrix with OD flows q_{14}=3, q_{24}=4, q_{34}=1 and q_{rs}=0 ∀ other rs.

Step 3. Using the BPR function (7.1), the observed link flows and the link performance function parameters of Table 7.1, the observed link travel times can be found to be: $\tilde{t}_1$=5.07, $\tilde{t}_2$=15.52, $\tilde{t}_3$=10.52, $\tilde{t}_4$=36.49 and $\tilde{t}_5$=26.06.

Step 4. Based on the observed link travel times found in Step 3, the minimum observed time path can be found by inspection. Between OD pair 1–2, this path consists of arcs 1 and 4; between OD pair 2-4, it consists of arc 4; between OD pair 3-4, it consists of arc 5. The times on these paths are the "observed OD travel time" $\tilde{u}_{rs}$ shown in Table A7.11.

Table A7.1.1 Observed shortest OD travel time $\tilde{u}_{rs}$

OD (rs)	From	to	Time $\tilde{u}_{rs}$
14	1	4	41.57
24	2	4	36.49
34	3	4	26.06

Step 5. The free-flow link travel times are given in Table 7.1. The shortest free-flow time paths can be found by inspection. Between OD pair 1-5, this path uses arcs 2 and 5; between OD pair 2-4 the path uses arcs 3 and 5; and between OD pair 3-4, it uses arc 5.

Step 6. Assigning the initial OD flows from Step 2 to the shortest free-flow time paths results in link flows $x_1^{1}=0$, $x_2^{1}=3$, $x_3^{1}=4$, $x_4^{1}=8$ and $x_5^{1}=8$.

Step 7. The link performance functions and the assigned flows $\{x_a^n\}$ from Step 6 can be used to calculate the assigned link travel times: $t_1^1=5$, $t_2^1=15.71$, $t_3^1=11.5$, $t_4^1=20$ and $t_5^1=7.37$.

Step 8. Again the shortest travel time path for each OD pair after the assignment of Step 7 can be found by inspection. The paths and the path times are presented in columns two and three of Table A7.1.2.

Step 9,10. Here we must compare minimum assigned OD path travel times from Step 8 with the minimum observed OD travel times from Step 4, and assign auxiliary flows according to equation (A7.1.8). For OD pair 1-4, Table A7.1.2 shows the minimum assigned path time is 23.08. This time is smaller than the corresponding observed OD travel time of 41.57(see Table A7.1.1). Therefore, the corresponding auxiliary path flow $g_m^{14^1}$ is set to the OD flow upper bound $\bar{q}_{rs}=15$. We calculate auxiliary path flows for the other OD pairs in a similar manner (see Table A7.1.2). After all the auxiliary path flows are determined, we can calculate the auxiliary OD and link flows using equation (7.1.8). The results are shown in Table A7.1.2.

Table A7.1.2 Auxiliary path, link and OD flows for iteration 1 in Part A

OD pair (rs)	Links of the minimum time path after assignment at n=1	Assigned path time $u_{rs,m}^1$		Observed path time $\tilde{u}_{rs}$	Auxiliary flows		
					Path	OD	Link
1-4	2, 5	23.08	<	41.57	$g_m^{14^1}=15$	$v_{14}^1=15$	$y_2^1=15$
2-4	3, 5	18.87	<	36.49	$g_m^{24^1}=15$	$v_{24}^1=15$	$y_3^1=15$
3-4	5	7.37	<	26.06	$g_m^{34^1}=15$ and other $g_k^{rs^1}=0$	$v_{34}^1=15$	$y_5^1=15+15+15=45$ and $y_1^1=y_4^1=0$

Step 11. To find the move size α to solve equation (A7.1.9), we set

$$dZ(\alpha)/d\alpha = \sum_a (y_a^n - x_a^n) * t_a(x_a^n + \alpha(y_a^n - x_a^n)) - \sum_{rs} \tilde{u}_{rs}(v_{rs}^n - q_{rs}^n) = 0$$

The bi-section method (Sheffi, 1985) can be used to solve for α. In this case, 11 iterations of the bi-section method yields α^1=0.1439.

Step 12. Equation (A7.1.10) is used to update link and OD flows.

For example, with α^1=0.1439 from Step 10, $x_2^{\ 1}$=3 from Step 6, q_{24}=4 from Step 2, y_2^1=15 and v_{24}^1=15 from Step 10, we can calculate

$$x_2^2 = x_2^1 + \alpha^1(y_2^1 - x_2^1) = 3 + 0.1439 * (15 - 3) = 4.7268$$
$$q_{24}^2 = q_{24}^1 + \alpha^1(v_{24}^1 - q_{24}^1) = 4 + 0.1439 * (15 - 4) = 5.5829$$

Other link and OD flows are found in a similar manner.

Step 13. To check convergence, we use the difference in Z values at successive iterations.

At n=1, $Z(\mathbf{x}^1, \mathbf{q}^1) = \sum_{a=1}^{5}(t_{0,a}x_a^1 + \frac{\alpha \cdot C_a}{\beta+1}(\frac{x_a^1}{C_a})^{\beta+1}) - \sum_{rs}\tilde{u}_{rs}q_{rs}^1$ = -166.3142. Since there is no previous iteration to compare with, we set n=2 and go to Step 7 (using link flows x_a^2 from Step 11 to calculate link travel times). Continuing in this fashion, at iteration n=16, we found $Z(\mathbf{x}^{15}, \mathbf{q}^{15})$-$Z(\mathbf{x}^{16}, \mathbf{q}^{16})$= -355.8108-(-355.8194)= 0.0086. This satisfied our convergence test, since 0.0086 < 0.01. Therefore, we end Part A with η^* = -355.819. Although a set of assigned OD flows $\{q_{rs}^{16}\}$ and link flows $\{x_a^{16}\}$ are produced for this final iteration, they are only used to calculate η^*. They are not used in Part B.

Part B

Step 1. We again arbitrarily choose upper bounds of 15 for all OD pairs, including those with zero OD flow in the target matrix (i.e., $\bar{q}_{rs}$=15 for all rs) and arbitrarily set γ=4.

Step 2. We use the target matrix in Table 7.2 as the initial $\{qrs^1\}$.

Step 3. The observed link travel times are the same as those calculated in Step 3, Part A.

Step 4. The observed OD travel times $\tilde{u}_{rs}$ are the same as those calculated in Step 4, Part A and shown in Table A7.1.1.

Step 5. The minimum free-flow time paths for OD pairs are the same as those found in Step 5, Part A.

Step 6. Assigning the initial OD flows form Step 2 to the shortest free-flow time paths results in link flows $x_1{}^1$=0, $x_2{}^1$=7, $x_3{}^1$=9, $x_4{}^1$=0 and $x_5{}^1$=22.

Step 7. The link performance functions and the assigned flows $\{x_a^n\}$ from Step 6 can be used to calculate the assigned link travel times: $t_1{}^1$=5, $t_2{}^1$=36.10, $t_3{}^1$=48.44, $t_4{}^1$=20 and $t_5{}^1$=140.56.

Step 8. Again by inspection, we find the minimum travel time path for each OD pair after the assignment. These paths are presented in Table A7.1.3.

Step 9,10. Here we again compare the minimum assigned OD travel times in Step 8 with the observed minimum OD travel time from Step 4 and assign auxiliary flows according to equation (A7.1.8). For OD pair 3-4, Table A7.1.3 shows the minimum assigned path time is 68.44. This time is greater than the corresponding observed OD travel time of 26.06(see Table A7.1.1). Therefore, the corresponding auxiliary path flow $g_m^{34^1}$ is set to 0. We calculate auxiliary path flows for the other OD pair in similar manner (see Table A7.1.3). After all the auxiliary path flows are determined, we can calculate the auxiliary OD and link flows using equation A7.1.8. The results are shown in Table A7.1.3.

Table A7.1.3 Auxiliary path, link and OD flows for iteration 1 in Part B

OD pair (rs)	Arcs of the minimum time path after assignment At n=1	Assigned path time $u^1_{rs,m}$		Observed path time $\tilde{u}_{rs}$	Auxiliary flows		
					Path	OD	Link
1-4	1, 4	25	<	41.57	$g_m^{14^1}$=15	υ^1_{14}=15	y^1_1=15
2-4	4	20	<	36.49	$g_m^{24^1}$=15	υ^1_{24}=15	y^1_4=15+15=30
3-4	3, 4	68.44	>	26.06	$g_m^{34^1}$=0 and other $g_k^{rs^1}$=0	υ^1_{34}=0	$y^1_2 = y^1_3 = y^1_5$=0

Step 11. To find the move size α to minimize equation (A7.1.15), given dL(α)/dα=

$$\sum_{rs} 2(\upsilon^n_{rs} - q^n_{rs})[q^n_{rs} + \alpha(\upsilon^n_{rs} - q^n_{rs}) - \hat{q}_{rs}] + \gamma\{\sum_a (y^n_a - x^n_a) * t_a[x^n_a + \alpha(y^n_a - x^n_a)] - \sum_{rs} \tilde{u}_{rs}(\upsilon^n_{rs} - q^n_{rs})\}$$

we use the bi-section method (Sheffi, 1985) repeatedly for various levels of α.

In this case, performing seven iterations of the bi-section method yields α=0.331

Step 12. Equation (A7.1.10) is used to update link and OD flows.

For example, with α=0.331 from Step 11, $x_2{}^1$=7 from Step 6, q_{24}=9 from Step 2, y^1_2=0 and υ^1_{24}=15 from Step 10, we can calculate

$$x^2_2 = x^1_2 + \alpha^1(y^1_2 - x^1_2) = 7 + 0.331*(0-7) = 4.9668$$
$$q^2_{24} = q^1_{24} + \alpha^1(v^1_{24} - q^1_{24}) = 9 + 0.331*(15-9) = 10.9867$$

Step 13. To check convergence, we use the difference in L value at successive iterations.

At n=1,

$$L(\mathbf{x}^1, \mathbf{q}^1, \gamma) = \sum_{rs} (q^1_{rs} - \hat{q}_{rs})^2 + 4[\sum_{a=1}^{5} (t_{0,a}x^1_a + \frac{\alpha \cdot C_a}{\beta+1}(\frac{x^1_a}{C_a})^{\beta+1}) - \sum_{rs} \tilde{u}_{rs} q^1_{rs} \text{-(-355.819)}]$$

= 2320.97. Since there is no previous iteration to compare with, we set n=2 and go to Step 7 (using link flows x^2_a from Step 12 to calculate link travel times).

At iteration n=25, we found L($\mathbf{x}^{24}$, $\mathbf{q}^{24}$)-L($\mathbf{x}^{25}$, $\mathbf{q}^{25}$)= 1.2562-1.2533=0.0029. This satisfies our convergence test, since 0.0029 < 0.005. So for γ=4, we stop here with a set of flows $\{x^*_a\}$ and $\{q^*_{rs}\}$.

Step 14. At γ=4, $dL(\gamma)/d\gamma = 0.034 > k$. Therefore, we increase γ and redo all the steps in Part B above.

At γ=4.3, we stop after 26 iterations and get $dL(\gamma)/d\gamma = 0.027 > k$. Therefore, we increase γ and redo Part B. At γ=6, we stop after 25 iterations and get $dL(\gamma)/d\gamma = -0.046 < -k$. Therefore, we decrease γ and redo Part B again. We continue in this way until we try $\gamma = 5$. For $\gamma = 5$, we stop after 2 iterations and get $dL(\gamma)/d\gamma = 0.0025$. This satisfies our convergence test, since $0.0025 < 0.005$. Therefore, we end Part B with $\gamma = 5$ and the corresponding set of flows $\{x^*_a\}$ and $\{q^*_{rs}\}$ obtained at iteration 26 as the final estimates of link and OD flows.

Appendix 7.2

Description and Illustration of Method 2 for Updating OD Matrices from Ground Counts

A7.2.1 Method description

Method 2 of Section 7.1 can be summarized by the following algorithm. In addition to network structure and link performance functions, the inputs to this program include limited OD data $\{\bar{q}_{rs}\}$ obtained from a small-scale survey and a set of "selected link" counts $\{\tilde{x}_r\}$, where r indicates a "selected link".

Trip generation and friction factor estimation

Step 0. Set the convergence level k and the iteration number n=1.

Step 1. Obtain zonal population and trip rates for each study zone, calculate zonal trip productions and attractions.

$$P_i^1 = A_i^1 = r_i * POP_i \quad \text{(A7.2.1)}$$

where P_i and A_i are, respectively, the number of trips produced and attracted by zone i during a given time period (Park and Smith (1996) used one day as their time period); r_i is the trip rate for zone i; POP_i is the population of zone i; and one denotes the iteration number.

Step 2. Obtain zone-to-zone travel time of each zone pair in the small-scale survey, and derive trip length frequency (TLF) distribution curves of three trip types--Internal-Internal (I-I), Internal-External (I-E), and External-External (E-E).

Step 3. Use the information from Steps 1 and 2 to develop friction factor curves of the gravity model for each of the three trip types, namely, I-I, I-E, E-E. Alternatively, specifying a gravity model with "friction factors" that are functions of travel times is acceptable.

Step 4. Use the friction factors, trip productions and trip attractions with the gravity model to estimate an OD table $\{q_{rs}^n\}$, where n denotes the iteration number.

Traffic assignment

Step 5. Assign the OD table obtained from Step 4 to the highway network using an all-or-nothing traffic assignment based on free-flow travel times. The assignment yields assigned link volumes $\{x_a^n\}$ for each link a.

Step 6. Calculate the root mean square error RMSE between the assigned volumes $\{x_r^n\}$ and the observed volumes $\{\tilde{x}_r\}$ corresponding to links r where observed volumes are available:

$$\mathrm{RMSE}^n = \sqrt{\frac{\sum_r (\tilde{x}_r - x_r^n)^2}{N}} \tag{A7.2.2}$$

where N is the number of "selected links" (i.e., links where observed volumes are available), and the other notation has been described above.

Step 7. Check convergence.

If $\mathrm{RMSE}^n \leq k$, set $q^* = q_{rs}^n$ and stop;

If $\mathrm{RMSE}^n > k$, then go to Step 8,

where k is the convergence criterion set in Step 0

.

<u>Selected link adjustment</u>

Step 8. For each selected link r, calculate the link adjustment ratio:

$$a_r^n = \tilde{x}_r / \sum_{ij} t_{ij}^{r,n} \tag{A7.2.3}$$

where a_r^n is the ratio between observed traffic volume and assigned volume for "selected link" r at iteration n; $\tilde{x}_r$ is the observed volume (vehicle classification count data) for "selected link" r; $t_{ij}^{r,n}$ is the number of trips from zone i to zone j using "selected link" r after the traffic assignment at iteration n, $\sum_{ij} t_{ij}^{r,n} = x_r^n$.

Step 9. Calculate zonal adjustment factors for trip productions and attractions:

$$R_i^n = \sum_r [a_r^n t_{i\bullet}^{r,n} / \sum_r t_{i\bullet}^{r,n}] \tag{A7.2.4}$$

$$S_j^n = \sum_r [a_r^n t_{\bullet j}^{r,n} / \sum_r t_{\bullet j}^{r,n}] \tag{A7.2.5}$$

where $R^n{}_i$ is the adjustment factor for trip productions in zone i at iteration n; $S^n{}_j$ is the adjustment factor for trip attractions in zone j at iteration n; $t_{i\bullet}^{r,n} (= \sum_j t_{ij}^{r,n})$ is the sum of all the flows on link r that come from zone i at iteration n;

$t_{*j}^{r,n} (= \sum_i t_{ij}^{r,n})$ is the sum of all the flows on link r that go to zone j at iteration n;

and a_r^n for every selected link r at iteration n was determined in Step 8.

Step 10. Estimate new productions and attractions

$$P_i^{n+1} = R_i^n * P_i^n \quad (A7.2.6)$$

$$A_j^{n+1} = S_j^n * A_j^n \quad (A7.2.7)$$

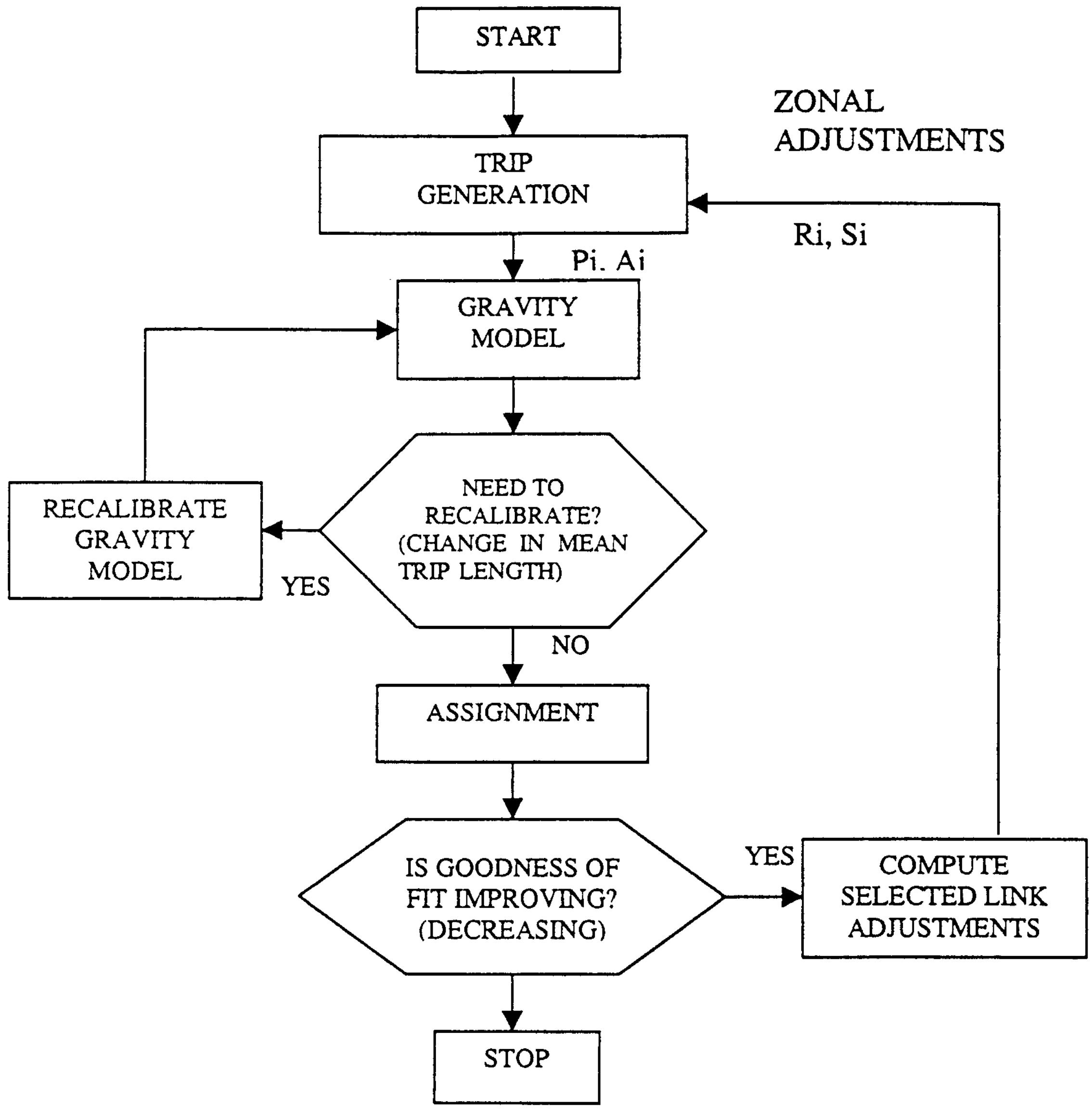
Step 11. Recalibrate the gravity model and estimate a new OD matrix $\{q_{rs}^{n+1}\}$ using the new productions and attractions found in Step 10.

Check gravity model stability

Step 12. Calculate the mean trip lengths for the three trip types (I-I, I-E, E-E, see Step 2) from $\{q_{rs}^{n+1}\}$ found in Step 11. If they are no longer "reasonably close" to the mean trip lengths from the small-scale survey (e.g., within 10%), then go to Step 3 and re-derive friction factor curves for the three trip types; If the mean trip lengths have not changed much, set n=n+1 and go to Step 5.

Method 2 terminates in Step 7 with output $\{q_{rs}^n\}$ and $\{x_a^n\}$. These represent the estimated OD matrix and the assigned link volumes, respectively. The algorithm is illustrated in Figure A7.2.1.

Figure A.7.2.1 Flow chart for algorithm (taken from Park and Smith, 1994)



A7.2.2 Example

To illustrate the method we consider the same simple network shown in Figure 7.2.

Trip generation and distribution estimation

Step 0. We arbitrarily set the convergence level k=0.05.

Step 1. In real problems the trip generation rates and zonal populations could be obtained from statistical data. For this hypothetical problem, we arbitrarily chose a set of trip productions and attractions. $P^1_1=55$, $P^1_2=70$, $P^1_{2a}=53$, $P^1_3=45$, $P^1_4=40$, $A^1_1=46$, $A^1_2=88$, $A^1_{2a}=50$, $A^1_3=32$, $A^1_4=47$

Step 2. Since developing the friction factor curves is relatively complicated and not the major concern of our study, we used a gravity model:

$$T_{ij} = \frac{P_i A_j F_{ij}}{\sum_j A_j F_{ij}} \qquad \text{(A7.2.8)}$$

where T_{ij} is the number of trips made from zone i to zone j; P_i is the total number of trips produced in zone i; A_j is the total number of trips attracted to zone j; F_{ij} is the friction factor between i and j, and friction factor F_{ij} is specified as a function of distance:

$$F_{ij} = e^{-\beta d_{ij}} \qquad \text{(A7.2.9)}$$

where d_{ij} is the distance between i and j, and β is a parameter.

Step 3. The unknowns in equation (A7.2.8) must be calculated before we can actually use this equation to estimate a trip matrix. The calibration strategy is to find values of F_{ij} that replicate the small-scale survey data in Table 7. 6 as closely as possible. F_{ij} is estimated using an iterative procedure. The results are shown in Table A7.2.1.

Table A7.2.1 Friction factor table

F_{ij}	1	2	2a	3	4
1	0.6593	1.3240		0.9138	1.0083
2	1.3240	0.6353		1.1495	1.0026
2a					
3	0.9138	1.1495		0.8757	1.0143
4	1.0083	1.0026		1.0143	0.9643

Step 3a. Equation (A7.2.9) can be transformed into

$$\ln F_{ij} = -\beta * d_{ij} \qquad \text{(A7.2.10)}$$

Using equation (A7.2.10) we regressed F_{ij} from Step 3 (Table A7.2.1) against d_{ij} of Table 7.7 (without 2a), to obtain an estimate of $\beta = 2.68E\text{-}5$

Step 3b. We then used the estimated value of β found in Step 3a, the distances between 2a and other zones, and equation (A7.2.9) to obtain the estimated friction factors associated with zone 2a. Combined with the other friction factors in Table A7.2.1, this leads to Table A7.2.2.

Table A7.2.2 Estimated friction factor matrix for the example of Method 2

F_{ij}	1	2	2a	3	4
1	0.6593	1.3240	1.00381	0.9138	1.0083
2	1.3240	0.6353	1.00207	1.1495	1.0026
2a	1.00381	1.00207	1.00145	1.00239	1.00285
3	0.9138	1.1495	1.00239	0.8757	1.0143
4	1.0083	1.0026	1.00285	1.0143	0.9643

Step 4. We used the gravity model of equation (A7.2.8) to estimate the OD matrix $\{q^1_{rs}\}$. The friction factors were obtained in Step 3b (Table A7.2.2), and the trip productions and trip attractions were given in Step 1. The estimated OD matrix is shown in Table A7.2.3.

Table A7.2.3 Estimated OD table for the example of Method 2 at iteration 1

From \ To	1	2	2a	3	4
1	6.1	23.4	10.1	5.9	9.5
2	17.0	15.6	14.0	10.3	13.2
2a	9.3	17.7	10.1	6.4	9.5
3	7.0	16.9	8.4	4.7	8.0
4	7.1	13.4	7.6	4.9	6.9

Traffic assignment

Step 5. We assigned the estimated OD matrix $\{q_{rs}^1\}$ in Table A7.2.3 to the network in Figure 7.2 using an all-or-nothing assignment based on free-flow travel time. This led to a set of assigned link volumes: x_1^1=59.8, x_2^1=29.5, x_3^1=53.8, x_4^1=0.0, x_5^1=70.9, x_6^1=51.1, x_7^1=31.9. The links constituting the shortest path between each OD pair are shown in Table A7.2.4.

Table A7.2.4 Links on minimum free-flow travel time paths for OD pairs

OD pair	Link ID						
	1	2	3	4	5	6	7
1-2	×						
1-2a	×					×	
1-3		×					
1-4		×			×		
2-2a						×	
2-3			×				
2-4			×		×		
2a-3					×		×
2a-4							×
3-4					×		

Step 6. Using equation (A7.2.2) to calculate the root mean square error between the assigned volumes $\{x_a^1\}$ from Step 5 and the observed volumes $\tilde{x}_r$ given in Table 7.5, we found RMSR^1= 7.34

Step 7. Since $RMSE^1 = 7.34 > 0.05$, we do not consider the algorithm to have converged, and we continue to Step 8.

Selected link adjustment

Step 8. As an example of how to use equation (A7.2.3) to calculate the link adjustment ratio, consider link 1(r =1). On this link, $\sum_{ij} t_{ij}^{1,1} = x_1^1 = 23.4+17+10.1+9.3 = 59.8$ from Step 5, and $\tilde{x}_1 = 54$. Therefore, using equation (A7.2.3), we obtained $a_1^1 = \tilde{x}_1 / \sum_{ij} t_{ij}^{r,1} = 54/59.8 = 0.903$. In a similar manner, we found: $a_2^1 = 1.288$, $a_3^1 = 1.023$, $a_5^1 = 0.874$, $a_6^1 = 1.116$, $a_7^1 = 1.315$.

Step 9. As an example of how to use equation (A7.2.4) to determine the adjustment factors for trip productions, consider zone one:

$$R_1^1 = \sum_r [a_r^1 t_{1*}^{r,1} / \sum_r t_{1*}^{r,1}] = \frac{a_1^1(q_{12}^1 + q_{12a}^1) + a_2^1(q_{13}^1 + q_{14}^1) + a_5^1 q_{14}^1 + a_6^1 q_{12a}^1}{q_{12}^1 + q_{12a}^1 + q_{13}^1 + q_{14}^1 + q_{14}^1 + q_{12a}^1}$$

$$= \frac{0.903 * (23.4 + 10.1) + 1.288 * (5.9 + 9.5) + 0.874 * 9.5 + 1.116 * 10.1}{23.4 + 10.1 + 5.9 + 9.5 + 9.5 + 10.1}$$

$$= 1.017$$

In a similar manner we found $R_2^1 = 0.983$, $R_{2a}^1 = 1.110$, $R_3^1 = 1.061$, $R_4^1 = 1.029$

As an example of how to use equation (A7.2.5) to get adjustment factors for trip attractions, consider zone one:

$$S_1^1 = \sum_r [a_r^1 t_{*1}^{r,1} / \sum_r t_{*1}^{r,1}] = \frac{a_1^1(q_{21}^1 + q_{2a1}^1) + a_2^1(q_{31}^1 + q_{41}^1) + a_5^1 q_{41}^1 + a_6^1 q_{2a1}^1}{q_{21}^1 + q_{2a1}^1 + q_{31}^1 + q_{41}^1 + q_{41}^1 + q_{2a1}^1}$$

$$= \frac{0.903 * (17.0 + 9.3) + 1.288 * (7.0 + 7.1) + 0.874 * 7.1 + 1.116 * 9.3}{17.0 + 9.3 + 7.0 + 7.1 + 7.1 + 9.3}$$

$$= 1.030$$

Similarly we found $S_2^1 = 0.986$, $S_{2a}^1 = 1.099$, $S_3^1 = 1.074$, $S_4^1 = 1.034$.

Step 10. As an example of how to use equations (A7.2.6) and (A7.2.7) to get new trip productions and attractions, consider zone one again:

$P_1^2 = R_1^1 * P_1^1 = 1.017*55 = 55.9$

$A_1^2 = S_1^1 * A_1^1 = 1.030*46 = 47.4$

Similarly, we found P^2_2=68.8, P^2_{2a}=58.8, P^2_3=47.8, P^2_4=41.1 and A^2_2=86.7, A^2_{2a}=55.0, A^2_3=34.4, A^2_4=48.6.

Step 11. Using the productions and attractions from Step 10, the friction factors in Table A7.2.2, and the gravity model of equation (A7.2.8), we estimate a new OD matrix $\{q_{rs}^2\}$ shown in Table A7.2.5.

Table A7.2.5 OD trip estimation q_{rs}^2 after link flow adjustment

From\ To	1	2	2a	3	4
1	6.2	22.8	11.0	6.2	9.7
2	16.5	14.5	14.5	10.4	12.8
2a	10.3	18.7	11.9	7.4	10.5
3	7.5	17.2	9.5	5.2	8.5
4	7.2	13.2	8.3	5.3	7.1

Step 12. If the friction factors were obtained from friction factor curves, here one would need to calculate the mean trip lengths for the three trip types (I-I, I-E, E-E, see Step 2) from $\{q_{rs}^{n+1}\}$ found in Step 11. If the calculated trip lengths were no longer "reasonably close" to the mean trip lengths obtained from the small-scale survey (e.g., within 10%), then one would go to Step 3 and re-derive friction factor curves for the three trip types. If the mean trip lengths had not changed much, then one would set n=n+1 and go to Step 5. Since we use "friction factors" that are functions of travel times, there is no need to check the mean trip lengths and re-derive friction factor curves. So we directly set n=2, and go to Step 5.

We repeated the steps above and checked convergence each time we reached Step 7: i.e., we decided whether the RMSE was still greater than 0.05. At Step 7 of iteration n=179, RMSE = 0.036 < 0.05. Therefore we stopped and obtained an estimated OD matrix from Step 11 of iteration n=178.

Appendix 7.3

Description and Illustration of Method 3 for Updating OD Matrices from Ground Counts

A7.3.1 Method Description

Given a network, a target matrix $\{\hat{q}_{rs}\}$ and a set of observed link counts $\{\tilde{x}_a\}$, the objective of the SPME method is to estimate a new trip matrix $\{q_{rs}\}$ that, when assigned to the network, minimizes the difference between the assigned link flows $\{x_a\}$ and the observed link flows $\{\tilde{x}_a\}$. As with the other algorithms, $\{x_a\}$ is obtained by running a traffic assignment model for $\{q_{rs}\}$. As we mentioned when introducing Method 1, different OD matrices can produce the same set of link flows. To find the one that is nearest to the target matrix, Nielsen uses the target matrix as a "seed" with which to start.

The key to Nielsen's approach is one of replicating "observed flows" by multiplying all OD flows using a particular link by the ratio between the "observed" and "assigned" flows on this link. This modified OD flow $q^n_{(E)rs,a}$ is called the "expected traffic" of OD pair *rs* using link a at iteration *n*. It is estimated by multiplying the assigned OD flow q^{n-1}_{rs} on link a at the previous iteration *n-1* by a ratio (which is the ratio between the assigned and observed link flow). Thus, the sum over all OD pairs of "expected traffic" using a link will be the observed flow on this link. Since we have different "expected traffic" for one OD pair with respect to the different links to which the OD flow is assigned, we take the average of them as the estimated one.

In addition to the network structure and link performance functions, the inputs of this algorithm are the set of observed link flows $\{\tilde{x}_a\}$ and a target matrix $\{\hat{q}_{rs}\}$. The solution to this program produces a set of estimated link and OD flows.

The procedure can be summarized as below:

Step 1. Set iteration number n = 1, initial OD matrix $q^0_{rs} = \hat{q}_{rs}$ and the convergence level to k.

Step 2. Assign q_{rs}^{n-1} to the network, using some traffic assignment algorithm — for example, an all-or-nothing algorithm — to obtain a set of assigned link flows $\{x_a^n\}$.

Step 3. For each OD pair rs, use a shortest path algorithm to find the shortest travel time path p from r to s through the network using link times estimated from the assigned flow $\{x_a^n\}$.

Step 4. Calculate the "expected traffic" $q_{(E)rs,a}^n$ for each OD pair rs using link a.

$$q_{(E)rs,a}^n = \frac{\tilde{x}_a}{x_a^n} * q_{rs}^{n-1} \tag{A7.3.1}$$

Step 5. Calculate the arithmetic mean of all the "expected traffics" for OD pair rs with respect to the links on its shortest time path p as the estimated number of trips q_{rs}^n between r and s:

$$q_{rs}^n = \frac{1}{N_{a \in (\tau,p)}} \sum_{a \in (\tau,p)} q_{(E)rs,a}^n \tag{A7.3.2}$$

where $q_{(E)rs,a}^n$ is the "expected traffic" between the zone pair r and s on link a found in Step 4; τ is the set of links with observed traffic along the shortest time path p; N is the number of links along the shortest path with observed traffic.

Step 6. Using the same traffic assignment algorithm used in Step 2, assign q_{rs}^n to the network to obtain a set of new assigned link flows $\{x_a^{n+1}\}$.

Step 7. For each OD pair rs, use a shortest path algorithm to find the shortest travel time path p from r to s through the network using link times estimated from the assigned flow $\{x_a^{n+1}\}$.

Step 8. Check convergence by calculating, for example:

$$\max_{a \in (\tau,p)} \left| x_a^n - x_a^{n+1} \right| . \tag{A7.3.3}$$

where x_a^n is found in Step 2; x_a^{n+1} is found in Step 6; a and τ are defined in Step 5; and the shortest travel time path p is found in Step 7.

If $\max_{a \in (\tau,p)} \left| x_a^n - x_a^{n+1} \right| < k$, then $\{ q_{rs}^n \}$ is the estimated matrix, and stop.

If $\max_{a \in (\tau,p)} \left| x_a^n - x_a^{n+1} \right| > k$, set n=n+1, and go to Step 4.