EPORT NO. DOT-TST-76T-38

MASTER

TRANSPORT OF SOLID COMMODITIES VIA FREIGHT PIPELINE

DEMAND ANALYSIS METHODOLOGY
VOLUME IV

July 1976
FIRST YEAR FINAL REPORT
UNDER CONTRACT: DOT-OS-50119

PREPARED FOR

U.S. DEPARTMENT OF TRANSPORTATION
Office of the Secretary
Office of University Research
Washington, D.C. 20590

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

Technical Report Documentation Page

1. Report No. DOT-TST-76T-38	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Transport of Solid Commodities Via Freight Pipeline - Demand Analysis Methodology, Volume IV		5. Report Date July 1976 6. Performing Organization Code
7. Author(s) W. B. Allen and T. Plaut		8. Performing Organization Report No.
9. Performing Organization Name and Address Department of Civil & Urban Engineering University of Pennsylvania Philadelphia, Pennsylvania 19174		10. Work Unit No. (TRAIS) 11. Contract or Grant No. DOT-OS-50119
12. Sponsoring Agency Name and Address Office of University Research Office of the Secretary U. S. Department of Transportation Washington, D. C. 20590		13. Type of Report and Period Covered First Year Final Report 14. Sponsoring Agency Code OST/TST-60

15. Supplementary Notes

OST Technical Monitor: David C. Ryan, Jr., TST-10

16. Abstract

This report describes findings of research performed during the first year of work under Contract DOT-OS-50119 for the Office of University Research, Office of the Secretary of Transportation. The application of freight pipeline for the movement of solid goods offers a new option in the field of transportation. Thus, the purpose of the first year of research was to evaluate the technical and economic feasibility of freight pipeline as an intercity transportation mode.

The report for the first year consists of the following five separate volumes:

VOLUME I - Cost and Level of Service Comparison
VOLUME II - Freight Pipeline Technology
VOLUME III - Cost Estimating Methodology
VOLUME IV - Demand Analysis Methodology
VOLUME V - Impact Assessment

The second year of research currently is being devoted to sharpening the concepts, broadening the areas of concern and applying the tools of analysis developed in the first year to a specific origin-destination transportation corridor.

17. Key Words Freight Pipeline Technology, Commodities, solid goods, Intercity Transportation, Impact Assessment, Cost Comparison, Demand Analysis Methodology	18. Distribution Statement Document is available to the U. S. Public through the National Technical Information Service, Springfield, Virginia 22161

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

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

This document is
PUBLICLY RELEASABLE
Barry Steel
Authorizing Official
Date: 7-20-07

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED
SHB

PREFACE

This report describes findings of research performed during the first year of work under contract DOT-OS-50119 for the Office of University Research, Office of the Secretary of Transportation. The application of freight pipeline for the movement of solid goods offers a new option in the field of transportation. Thus, the purpose of the first year of research was to evaluate the technical and economic feasibility of freight pipeline as an intercity transportation mode.

The report for the first year consists of the following five separate volumes:

I.	Cost and Level of Service Comparison	I. Zandi; B. Allen; E. Morlok, K. Gimm; T. Plaut; J. Warner
II.	Freight Pipeline Technology	I. Zandi and K.K. Gimm
III.	Cost Estimating Methodology	Section A: J. Warner and E. Morlok Section B: K. K. Gimm and I. Zandi
IV.	Demand Analysis Methodology	B. Allen and T. Plaut
V.	Impact Assessment	I. Zandi and K.K. Gimm

The second year of research currently is being devoted to sharpening the concepts, broadening the areas of concern and applying the tools of analysis developed in the first year to a specific origin-destination transportation corridor.

The authors wish to acknowledge gratefully the assistance given by Mr. David C. Ryan Jr. of the Office of R & D Policy, Office of the Secretary of Transportation. His numerous technical and editorial suggestions have been of great help to us.

Barry Silverman, Melissa Clark-Rhodes, and Janet Hines have also contributed to this document in various capacities.

Iraj Zandi
Principal Investigator

FIRST ANNUAL REPORT
TRANSPORT OF SOLID COMMODITIES
VIA FREIGHT PIPELINE

July 1976

Volume IV

Demand Analysis Methodology

The Demand for Freight Transportation
The Case of the Intercity Freight Pipeline

Executive Summary

In order to determine the feasibility of intercity freight pipelines, it was necessary to determine whether sufficient traffic flows currently exist between various origins and destinations to justify consideration of a mode whose operating characteristics became competive under conditions of high traffic volume.

An intercity origin destination freight flow matrix was developed for a large range of commodities from published sources. A physical screening was then applied to yield a flow matrix which consisted of only goods which could be physically moved by pipeline. Rather than consider all possible origins and destinations, a high freight traffic density corridor between Chicago and New York and another between St. Louis and New York were studied. These corridors, which represented 18 cities, had single direction flows of 16 million tons/year. If trans-shipment was allowed at each of the 18 cities, flows of up to 38 million tons/year were found in each direction. These figures did not include mineral or agricultural products.

After determining that such pipeline eligible freight traffic volumes existed, the next step was to determine the ability of freight pipeline to penetrate such markets. This entailed a modal split analysis. Since no markets presently exist in which freight pipeline is involved, it was not possible to empirically observe such modal competition and shippers' behavior therein. Thus an abstract demand-modal split model was formulated where shipper's reaction to an abstract set of pipeline performance characteristics, e.g., rate, time in transit, reliability, etc., was assumed to be the same as their reaction to the same set of abstract performance characteristics when exhibited by existing modes. Since shippers' behavior with respect to truck and rail performance characteristics was already observable, it was possible to determine how shippers choose among freight transportation modes using different abstract performance characteristics.

Modal split models were run on aggregate data from the 1967 Census of Transportation. Transportation rates and transportation times by both truck and rail were estimated and then in turn used in the modal split models. Such models (logit estimated) yielded reasonable classifications of existing traffic into truck or rail by 3 digit STCC commodity.

Modal split models were also run on disaggregate data specially collected for this study. Six major national firms were contacted and data (rate, time, reliability) on the chosen mode and on the non-chosen mode were collected for four of them. Modal split models (logit estimated) were then developed for three of the four firms. Such models performed quite well for one firm and with moderate success for the other two.

The freight pipeline service characteristics were then substituted into both the aggregate and disaggregate models (truck versus pipeline and then rail versus pipeline) and estimates of pipeline penetration into particular STCC commodity groups were made. In general, pipeline was estimated to be able to penetrate 20-50% of each of the STCC markets investigated *ceteris paribus*. For some commodities, however, the estimated penetration was 50% or more.

Based on these very preliminary results, it appears that freight pipeline has market penetration potential that is consistent with high volume participation in the intercity freight market. Needless to say, caveats are necessary at this point in time. The results of the second year of study should enable more definitive conclusions to be drawn.

FIRST ANNUAL REPORT
TRANSPORT OF SOLID COMMODITIES
VIA FREIGHT PIPELINE
July 1976
Volume IV

Demand Analysis Methodology

Table of Contents

FIRST ANNUAL REPORT
TRANSPORT OF SOLID COMMODITIES
VIA FREIGHT PIPELINE

July 1976

Volume IV

Demand Analysis Methodology

List of Tables

FIRST ANNUAL REPORT
TRANSPORT OF SOLID COMMODITIES
VIA FREIGHT PIPELINE
July 1976
Volume IV

Demand Analysis Methodology

List of Figures

FIRST ANNUAL REPORT
TRANSPORT OF SOLID COMMODITIES
VIA FREIGHT PIPELINE
July 1976
Volume IV

Demand Analysis Methodology

List of Maps

THE DEMAND FOR FREIGHT TRANSPORTATION: THE CASE OF THE INTERCITY FREIGHT PIPELINE

Introduction

While literature on freight transportation demand exists ranging from a macroeconomic point of view (1,2) to a microeconomic point of view (3,4,5), a literature search revealed that the presentation of theory was discussed more frequently than empirical studies (with the exception of Polenske, (6)). Recently, however, some freight demand and modal split models have been estimated (e.g. Benishay and Whitakker, (7), used Samuelson's (3) Model; Reebie, (8), estimated containerizable flows between 130 aggregate Office of Business Economics (OBE) regions; and Hartwig and

---

(1) For a gross regional product-econometric transportation demand model see Mathematica, Studies on the Demand for Freight Transportation, Princeton, NJ, 1967.

(2) For a discussion of a multi-regional input-output flow model see Lang, A.S., "Demand and Supply: The Technology of Transportation" in E. Williams, ed., The Future of American Transportation, Prentice Hall, Englewood Cliffs, NJ, 1971, pp. 41-57.

(3) Samuelson, P., "Spatial Price Equilibrium and Linear Programming", American Economic Review, Vol. XLII, 1952, pp. 165-177.

(4) Baumol, W. and H. Vinod, "An Inventory Theoretic Model of Freight Transport Demand", Management Science, Vol. 16, 1970, pp. 413-422.

(5) Allen, W.B. and L.N. Moses, "Overseas Trade: Competition Between Air and Sea" Transportation Research Forum Papers, Richard B. Cross and Co., Oxford, Indiana, 1968, pp. 235-248 & Allen, W.B., "The Demand for Freight Transportation: A Micro Approach", Transportation Research, Vol. 11, 1977, pp. 9-14.

(6) Polenske, K. et.al., State Estimates of the Gross National Product, D.C. Heath and Co., Lexington, Mass., 1972.

(7) Benishay, H. and G. Whitakker, Demand and Supply in Freight Transportation, The Transportation Center, Northwestern University, 1965.

(8) Reebie Associates, National Intermodal Network Feasibility Study, prepared for U.S. DOT, FRA, Washington, DC, 1976,

Linton (9), estimated a disaggregate modal split model with techniques used in the current urban passenger transportation models).

Freight demand theory work is still in its infancy, and that which has been done is limited to a handful of researchers. The bulk of the work which has been done is also relatively recent, and there is more work currently in progress than has probably been produced in the past. However, much of the current work is either theoretical in nature or based in the heavily empirical-oriented field of input-output. The latter deals in levels of geographical aggregation and production function aggregation which may be too large for the purpose of the research on freight pipeline (although some attempt to use some of this data shall be made and shall be developed in the second year report), i.e., it is not commodity or route specific enough to be of use in determining pipeline feasibility.

Thus, while the intention here was not to "reinvent the wheel", the existing literature has been surveyed and it was determined that much of it was not useful for the purposes of this freight pipeline research project. It must be remembered that pipeline transport of freight is an area where little market research has been done. Of course, no observations of the existing demand for freight transport via pipeline exist since not very many freight pipelines yet exist in the United States. (Petroleum and its products are excluded from the analysis since pipelines already move the product in great numbers; only one coal slurry pipeline exists in the U.S. and some chemical lines exist--see Zandi, (10), and Kim, (11). Thus, the task of estimating demand for freight pipeline was made more difficult because no past evidence was available.

Since a direct approach using past freight pipeline experience was not possible, an indirect approach (as outlined below) was undertaken. This approach suggested that even without past pipeline experience, prospective shipper's behavior toward pipeline could be inferred from their behavior regarding the modes with which they currently have a choice, i.e., the Quandt-Baumol (12) abstract mode approach.

---

(9) Hartwig, J. and W. Linton, "Disaggregate Mode Choice Models of Intercity Freight Movement", Unpublished Master's Thesis, Transportation Center, Northwestern University, 1974.

(10) Zandi, I., "Future of Pipeline--Beyond Liquid and Gas", Proceedings of the Fifteenth Annual Meeting, Transportation Research Forum, Richard B. Cross, Oxford, Indiana, 1974, pp. 187-193.

(11) Kim, K.S., "A Review of Practical Experience with Solid Pipelines", Proceedings of the Fifteenth Annual Meeting Transportation Research Forum, Richard B. Cross, Oxford, Indiana, 1974, pp. 371-380.

(12) Quandt, R. and W. Baumol, "The Demand for Abstract Transportation Modes: Theory and Measurement", Journal of Regional Science, Vol. 6, 1966, pp. 13-26.

It must be noted that whether a demand for a freight pipeline is obtained (traffic generation and new location decisions aside) depends (1) on whether a sufficient traffic volume exists and (2) on whether the pipeline can competitively attract a sufficient amount of that traffic. Thus, if an insufficient freight traffic volume exists to sustain a pipeline, (2) becomes superflous.

The Flow Data

The first task was to determine what freight traffic volumes now exist. Since a pipeline would flow from point to point, disaggregate flows on a geographical basis were desired, where possible. Due to size or special handling requirements it will be physically impossible for some items to go through a pipeline. Therefore, it was desirable to obtain data on freight traffic volumes as disaggregated as much as possible by commodity codes in order to eliminate STCC (Standard Transportation Commodity Code) codes from the freight traffic data which had physical limitations such as to preclude movement by pipeline.

Freight transportation volume data for the U.S. are not readily available in any generally consistent form. The Peat, Marwick, Mitchell data (gathered for the Reebie study (8) are presented only in terms of total flow of containerizable products. Apparently these data were once in a form disaggregated by commodity (at least on the DOT multiregional input-output level), and it seems that once a product was associated with a containerizable class, its STCC commodity code identification was dropped. Thus, commodity specific data were erased and only general aggregations of containerizable commodities were maintained.

The problems of finding consistent flow data for the U.S. were documented in Whitten (13) and were brought up to date in Reebie (8). This situation has not improved much since 1968, although the Peat, Marwick, Mitchell data could have been a basic improvement had the commodity code identification not been lost. Existing sources of data differ with respect to their coverage -- both in terms of commodity detail and geography. Some freight transportation modes are better covered than others. Some data give comprehensive information on origins but not on destinations, while other data give the reverse.

Data which are the most consistent for a freight transportation mode are provided by the 1% Rail Waybill currently administered by U.S. DOT. These data are

---

(13) Whitten, H., ed., Transport Flow Data: Proceedings of the National Transportation Flow Statistics Forum, The Transportation and Logistics Research Center, The American University, Washington, DC, 1968.

available on a five digit commodity level (which is quite disaggregate) and on a state-to-state basis. There are dangers in expanding the 1% waybill up to an estimate of complete enumeration since the waybill is not strictly a 1% sample. These and other problems in the waybill were described by Banks (14) in the Burden Study. Noted was that as the degree of disaggregation was increased for geographical and commodity data, the lower the confidence which could be placed on extrapolations of flows from the 1% sample. Nevertheless, with the exception of TOFC/COFC (Trailer on Flat Car/Container on Flat Car) movements (somewhat less than 5% of rail carloadings), the 1% waybill sample is inclusive of all commodities. This data source is an important one.

However, the authors must express caution in using the data. Using state-to-state data for iron ore in a study of commodity flows done for U.S. DOT (15), one found an estimated complete enumeration of a yearly flow of iron ore from Minnesota to Illinois which would exhaust the yearly capacity of the Illinois Steel works complex (exclusive of Lake movements of iron ore) several times over. Nor was this magnitude of commodity trans-shipped to waterway in Illinois. When a comparison was made of complete enumeration data derived from the 1% waybill sample for individual railroads with the quarterly commodity statistics (QCS) reports (true enumerations) of those railroads, errors of over 100% were found for some carriers.

Railroad TOFC/COFC movements on an origin-destination basis were not publicly available. While the FRA-Reebie study (8) apparently has such data, it is not clear that these are available for public consumption.

The best sources of multimodal freight traffic data were found to be the three Census of Transportation (1963, 1967, 1972) taken to date. This data includes all transportation modes except pipeline (and express, etc.). There was an unfortunate lack of commodity coverage in the Census. Only the manufacturing STCC's (20-39) were covered and then only if the movement was from a manufacturing point. Any movement of manufactured goods from a warehouse or a distribution point (a movement most likely by truck) was excluded. Also excluded were local movements (defined as originating and terminating within the same municipal boundaries or as terminating within 25 miles of the origin).

---

(14) Banks, R.L. and Associates, "An Estimation of the Distribution of the Rail Revenue Contribution by Commodity Groups and Type of Rail Car, 1969", Washington, DC, for the USDOT, October, 1972.

(15) U.S. DOT, "The Future Market for Rail Transport in the Northeast", Office of Policy and Plans Development, TPI-30, Washington, DC, 1973.

The latter exclusion was not disturbing for purposes of the freight pipeline study since the goal herein was to investigate an intercity freight pipeline. However, the lack of coverage of the raw material movements (STCC 1-18--primarily agricultural products and products of mines, STCC 19 is ordnance) and the lack of coverage of scrap and waste and of containerized movements (STCC's 40-46) was most important. Since it is possible to envision compatible movements of raw materials and manufactured products, e.g., a slurry of coal as the mode of conveyance for a capsule, it would have been desirable to have all possible commodity flow information. However, flows of unprocessed products were not available from the Census. The 1977 Census of Transportation will rectify this commodity coverage difficulty with respect to raw materials.

Public availability of tapes for the 1967 Census of Transportation enabled construction of an origin-destination (O-D) matrix with 25 origins and 59 destinations (the 1972 tapes allowed 27 origins). The 25 origins were dubbed production areas and were SMSA's (Standard Metropolitan Statistical Area) or combinations of SMSA's (e.g. the Philadelphia production area was the five county Philadelphia SMSA, the one county Trenton (NJ) SMSA, and the three county Wilmington (Del.) SMSA). Thus the origins and destinations were much broader in scope than the names attributed to them in row and column headings of matrices. The 59 destinations included the 25 origins plus 25 other production areas plus 9 regional catchalls. The destination list was thus exhaustive of the country; the origin list was not however--but did cover the major origins of traffic.

An origin-destination matrix could have been constructed by commodity (commodity detail was up to 5 digit STCC), by mode and by shipment size. However, the same level of detail was not available for each production area. For large and diverse production areas (e.g., Philadelphia, Chicago, Los Angeles) there was available considerable commodity coverage and quite a number of five digit commodities were listed. Small and less diverse production areas (e.g., Denver, Atlanta) had only a few two digit STCC's reported.

Thus, to obtain an origin destination matrix with a minimum number of zeros elements it was necessary to aggregate commodities. Precision was lost, of course, in the course of aggregation. On a disaggregate level many zeros would be reported and many of these would be in cells in which the true reading was not zero, but was included in a higher level of aggregation because it was small or to assure that inadvertent disclosure of privileged data was not made.

The Census data was found to present other difficulties which may cast doubt on its reliability. Comparing national complete enumerations of QCS rail data with the complete rail enumerations estimated from the Census on a commodity-by-commodity basis yielded Census estimates of from 25% to 400% of actual (QCS) totals. Other difficulties exist with respect to the Census of Transportation. These were outlined in Crecine, Moses, and Stucker (16).

Despite these difficulties, during this study, nine matrices were produced, each for 25 by 59 origin-destination pairs, and at the three digit STCC commodity code level of aggregation. These matrices reflected commodities which passed the physical screen on "pipelineability" (17), but also met certain prerequisites with respect to the number of production areas where observations existed and where percent of goods shipped was over 100 miles. The micro demand analysis (to be discussed below) concentrated on these commodities. These matrices shall be discussed below.

Some attempt could be made to fill in some of the data gaps from the Census of Transportation. That is the Rail 1% Waybill covered the raw material and scrap STCC's. Department of Agriculture (USDA) data shows origin by state to 41 destination cities (some of which--but not all-coincided with the Census' 25 production areas) for fresh fruits and vegetables. An investigation of the Census of Agriculture can frequently identify counties where products are grown in a state, and hence likely origins can be determined. Many of these products move by truck due to the agricultural exemption.

There was very little systematic information on grain traffic volumes except for the 1% Rail Waybill. Truck information was non-existent except for some local area studies (Beuthe, (18), Iowa State University, (19)). The Army Corps of Engineers does provide port information on a commodity basis; however, the original origin of a movement was not given, nor was the ultimate destination. The USDA did not track grain movements. As discussed below, some surmising was done with the data as available.

The Department of Interior, Bureau of Mines (20) produces origin destination

---

(16) Crecine, J., L. Moses, J. Stucker, "The Census of Transportation: An Evaluation", Transportation Research Forum Papers, Richard B. Cross, Oxford, Indiana, 1966, pp. 87-105.

(17) Gimm, K.K.,"Inter and Intra Urban Freight Transportation Via Pipeline", Ph.D Dissertation, in Civil and Urban Engineering, University of Pennsylvania, Philadelphia, Pennsylvania, 1976.

(18) Beuthe, M., "A Predictive Model of Regional Demands for Freight Transportation", Journal of Regional Science, Vol. 12, #1, 1972, pp. 85-94.

(19) Iowa State University, An Economic Analysis of Alternative Grain Transportation Systems: A Case Study, prepared for the Department of Transportation, Federal Railroad Administration, Washington, DC, November, 1973.

(20) U.S. Department of Interior, Bituminous Coal and Lignite Distribution, Calendar Year 1974, Bureau of Mines, Division of Fossil Fuels, Washington, DC, April 18, 1975.

information for bituminous coal and lignite movements. Origins are 23 defined areas and destinations and are all states (except for Louisiana for some reason). These movements were also given by mode.

Finally the U.S. Department of Commerce, the U.S. DOT, and the Army Corps of Engineers, (21) cooperated to produce a report on the internal origin of U.S. exports and internal destination of U.S. imports for the year 1970. Balfe et.al. (22) have cast some doubt as to the utility of this data.

First Cuts with the Flow Data

Because the flow data were in very rough form, a study approach was to examine the Reebie study (8) flow data. The Reebie freight traffic volume data were reported between 130 origins and destinations in 40 foot trailer equivalents. These origins-destinations were OBE (Office of Business Economics-U.S. Department of Commerce) regions or their aggregates and were basically exhaustive of the whole country. The name of a primary city designated a particular region.

The Reebie flow data were built up from: the 1% Rail Waybill, a special TOFC/COFC survey of the railroads; the Census of Transportation; the fresh fruits and vegetables data from the Department of Agriculture; Post Office Department flow information; and a special survey on the inland origin and destination of U.S. exports and imports (for 1970). This did not, however, represent the U.S. universe of freight shipments. Freight traffic flows were allocated to regions by each region's population or manufacturing employment. These freight traffic data were then normalized to the year 1971 since originally they were from various sources and based on different years.

The Reebie study (8) flow data were those which were dubbed as prime or suitable containerizable by previous Maritime Administration studies. Marginal and unsuitable flow data were excluded from the data base.

In consultation with the Pipeline Technology Group of this project, it was decided that after the exclusion of perishable products, 90 percent of those products

---

(21) Domestic and International Transportation of U.S. Foreign Trade: 1970, U.S. DOC, Bureau of the Census, U.S. DOT, Office of the Secretary, Department of the Army, Corps of Engineers, Washington, USGPO, 1972.

(22) Balfe, M., R. Heilmann, J. Johnson, and W. Wendling, "Limitations of Policy Formulation with Imperfect Information: A Case Study with Respect to the Great Lakes", Mimeo, University of Wisconsin, Milwaukee, Wisconsin Sea Grant Program, 1975.

which were prime or suitably containerizable were likewise prime and suitable for pipeline. The origin-destination flow data provided by the Reebie study (8) presented a percentage of perishable products thus making it possible to state freight traffic volumes in terms of non-perishable commodities. Ninety percent of the remaining total freight traffic volume was used as the rough approximation of the potential freight pipeline traffic.

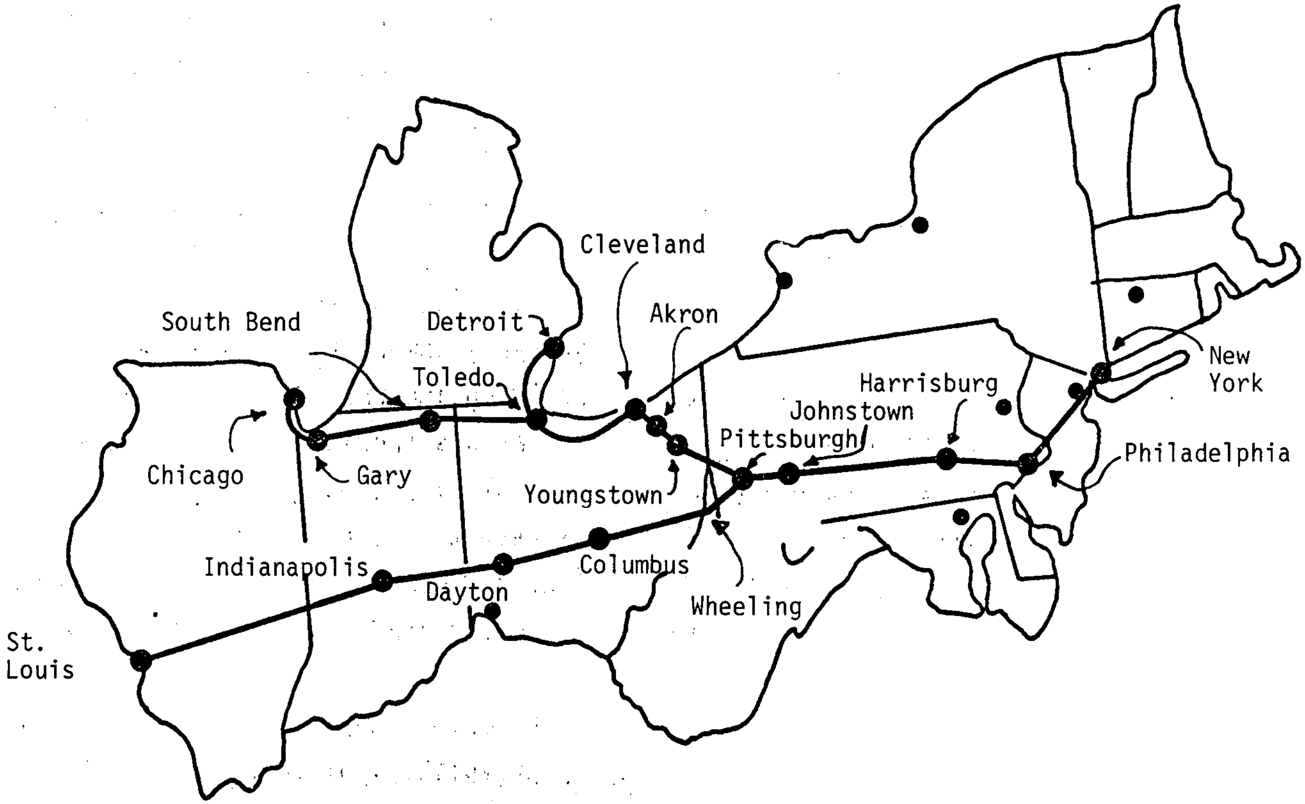
However, the Reebie study (8) data did not take into account flows of ores, coal, and minerals (STCC's 10, 11, and 14) which were considered to be prime candidates for pipelineable commodities. Consequently, it was necessary to determine these flows and put them into the Reebie Study format. Since most of the products which move any appreciable distance go by rail and/or barge and since reliable data only existed for rail (except for coal), a rail origin-destination matrix was constructed from the 1% Waybill sample. This was compared with the coal flow data derived from the Department of Interior as a check for consistency.

It was decided that the investigation of the Reebie study (8) data would be used to concentrate on the Northeast and Midwest states--basically the 17 state area from Maine to Maryland and West Virginia, then west along the Ohio River and up the Mississippi River to encompass Illinois. The analysis focused on 45 of the Reebie study's (8) 130 areas and thereby eliminated the difficulty of dealing with an unwieldy 130 by 130 matrix. This decision was also due to the fact that a priori evidence was available from previous personal work of the authors with the Census of Transportation which showed that certain high freight traffic density corridors exist.

Since the origin destination data was collected on the basis of production areas or OBE regions which bear the names of cities, a hypothetical pipeline running along two high freight traffic density corridors in the Northeast and Midwest states connecting 18 major cities was assumed for analysis. The first assumed pipeline routed from Chicago, through Gary, South Bend, Toledo, (with a spur to Detroit), Cleveland, Akron, Youngstown, Pittsburgh, Johnstown, Harrisburg, and Philadelphia to New York. The second assumed pipeline routed from St. Louis, through Indianapolis, Dayton, Columbus, Wheeling, and then on to Pittsburgh to join with the other line (See Map 1).

Data were analyzed in several ways. In one analysis data were examined for only the 18 cities actually on the assumed pipeline network. Thus an 18 by 18 flow matrix was made of freight traffic volumes both originating and terminating in the 18 areas. This analysis (initially) excluded minerals and ores. A flow matrix of containers, less perishable goods, was constructed (shown in Table I) and the entries multiplied by .9 (the estimate of the non-perishable containerizable goods that were assumed pipelineable). The container equivalents were then multiplied by 17.75 tons. This

Map 1. Hypothetical Pipeline Network



equivalence was derived from the average of the weights given in a flow of containers from New York given in the Reebie study (8) report.

The weight of many of the containers given in Reebie study report (8) was 22 tons/container although other, lower weights, of course, exist. A weighted average over all commodities could have been used, but time precluded its calculation. From the point of view of making a preliminary estimate, the 17.75 tons/ 40 foot container based on the New York sample was deemed adequate.

These flows are shown in Table I. (They must be multiplied by .9 x 17.75 to yield tonnages). Note that intra-area flows were excluded (the main diagonal-x), as were flows between areas virtually contiguous, e.g., Chicago-Gary, Akron-Youngstown, x. In some cases the Reebie study (8) data showed no flows existing - y, e.g., Youngstown to South Bend. In other cases, flows do occur, but it was assumed that they would not move by pipeline due to circuity, z, in the network presented herein, e.g., Chicago-St. Louis.

The sums of the rows and columns of Table I did not represent the total exports and imports of the respective cities, but only those exports (imports) to (from) the other 17 cities. Since 112 other possible exporters/importers exist, these flows could be small relative to the total activity of the area. However, since these areas were relatively close to one another and were some of the larger population centers, a gravity model hypothesis would lead to the conclusion that fairly high freight traffic volumes occur between these areas.

Table II shows the loading on the various links of the hypothetical pipeline network using the data from Table I (converted to tons). The westbound link with the highest freight traffic volume on line one was between Akron and Cleveland with an estimated flow of 9.8 million tons of potentially pipelineable products. The eastbound link with the highest freight traffic volume was between Harrisburg and Philadelphia where 16.2 million tons per year were moved. Line two had a much lower volume of freight if traffic were constrained to only the network cities. No eastbound line generated the freight traffic volume as the least link in Network One. The westbound links generated even less freight traffic volume.

The potential magnitude of these flows can be put into perspective. The capacity of a 60" diameter pipeline was estimated at 7 million tons in each direction per year. Route one exceeded that constraint on many links (see Table II). In addition, the Technology Group reported that for freight traffic flows (east and westbound) totalling 14 million tons per year, the cost per ton-mile approached 1.6¢. Such costs would make the pipeline line-haul cost competitive with those for rail movements. If the speed, reliability, loss and damage costs performed better than those for rail or truck, the freight pipeline could stand an excellent

Table I. Containerizable Flows from 18 Origins to 18 Destinations

Origin \ Des't	Chicago	Gary	South Bend	Toledo	Detroit	Cleveland	Akron	Youngstown	Pittsburg	Johnstown	Harrisburg	Philadelphia	New York	St. Louis	Indianapolis	Dayton	Columbus	Wheeling
Chicago	x	x	13268	13238	72186	24391	12084	9696	14208	y	5566	28578	104184	z	z	z	z	1381
Gary	x	x	7360	6120	39307	12959	7539	4210	17340	y	4054	9746	35909	z	z	z	z	x
South Bend	24234	18294	x	2644	y	4313	y	y	y	y	729	y	4499	z	z	z	z	x
Toledo	25134	6382	2650	x	x	35389	22320	19442	28069	1925	6717	11823	21689	z	z	z	z	14571
Detroit	68614	45524	17076	x	x	50936	22769	15366	18257	y	2309	24331	96875	z	z	z	z	7128
Cleveland	31833	15801	4066	16227	79982	x	44110	48419	24211	y	5872	28888	55675	z	z	z	z	z
Akron	17529	6298	1578	14369	43556	40529	x	x	10849	861	3627	17033	28623	z	z	z	z	z
Youngstown	15497	6587	y	6532	21206	32251	x	x	45928	y	4271	15558	20281	z	z	z	z	z
Pittsburg	19870	9410	y	5559	22607	27375	11960	28006	x	2348	12418	81772	113785	6724	y	7111	4018	x
Johnstown	4239	1059	y	y	1436	2101	y	2218	y	x	3182	8971	11661	654	860	y	y	y
Harrisburg	5232	2595	y	1360	4539	2108	6605	1674	27458	10084	x	66016	77323	1000	y	655	y	y
Phila.	34920	9476	y	5688	28859	30390	11482	10427	88458	8412	39779	x	64954	8774	5651	10197	8763	4027
New York	81356	21132	1880	6902	37744	19721	4483	3398	53989	3275	18653	117921	x	16637	4978	6051	7935	4146
St. Louis	z	z	z	z	z	z	z	z	4360	y	868	8145	16366	x	10017	3891	2313	y
Indianapolis	z	z	z	z	z	z	z	z	6607	y	1499	13865	17583	16521	x	6014	1016	1515
Dayton	z	z	z	z	z	z	z	z	9612	y	5418	16101	25283	5646	9547	x	21894	7332
Columbus	z	z	z	z	z	z	z	z	8618	y	6505	14454	20839	3037	y	9452	x	3791
Wheeling	4529	2712	y	4876	6334	z	z	z	x	y	1262	10692	18159	1414	y	2484	2506	x

x - no movement reported because of close proximity
y - no flows recorded
z - no movement reported because of circuity

* Conversion to tonnage by multiplying by .9 x 17.75

Source: Calculated by the authors from Reebie (8)

Table II

Tonnage Flows on Links of Hypothetical Network
(Assumes that only network cities can originate and terminate traffic)

Link	Westbound	Eastbound
Chicago-Gary	5,319,467	4,801,766
Gary-South Bend	7,640,156	7,110,856
South Bend-Toledo	7,396,888	6,975,979
Toledo-Cleveland	9,670,866	11,232,054
Cleveland-Akron	9,775,757	12,636,672
Akron-Youngstown	8,348,727	11,786,179
Youngstown-Pittsburg	7,757,971	11,555,755
Pittsburgh-Johnstown	9,477,265	14,808,585
Johnstown-Harrisburg	9,624,267	15,106,998
Harrisburg-Philadelphia	9,546,388	16,189,384
Philadelphia-New York	6,552,961	11,720,665
St. Louis-Indianapolis	942,365	734,211
Indianapolis-Dayton	1,037,081	1,342,571
Dayton-Columbus	1,358,210	2,552,438
Columbus-Wheeling	1,540,137	3,015,457
Wheeling-Pittsburgh	1,965,899	3,552,952

Source: Calculated from Table I and Reebie (8).

chance of attracting a significant share of the assumed market (see below).

However, it may be too confining to limit the pipeline's service to just the 18 areas located on the two assumed pipelines. Freight movements which either originate and/or terminate in areas not on the line may still use the pipeline for a portion of their haul, e.g., a movement from Milwaukee (or Seattle) to Philadelphia may come to Chicago to be trans-shipped onto the pipeline. Or a movement from Milwaukee to Baltimore could involve Milwaukee to Chicago access, line haul by pipeline from Chicago to Harrisburg, and egress Harrisburg to Baltimore.

Assuming that trans-shipment was physically possible (reserving judgement on the economics for a moment), two approaches were taken: one of the approaches allowed for long haul moves on the pipeline system with a majority of movements either originating or terminating on the system, while the other, less restrictive, approach allowed almost any east-west move, which did not involve absurd route circuity, to enter the system.

The first step was to set up a 130 by 130 inverted L-shaped matrix representing intra-regional shipments in a 45 by 45 northwestern corner of the matrix. Added eastward to the 45 by 45 submatrix was a 45 by 85 submatrix representing the exports by the 45 areas in the study region to the 85 areas outside of the study region. Added southward to the 45 by 45 submatrix was an 85 by 45 submatrix representing the imports to the 45 areas in the study region from the 85 areas outside of the study region. (This large matrix is unwieldy and is not reproduced here). The remaining 85 by 85 submatrix was not completed since it entailed shipments which in no way involved the study region or those movements flowing through the study region.

Using the more restrictive approach for trans-shipment and translating the container equivalents to tonnages (17.75 x .9) for each link, yields Table III. As can be seen, the traffic on the links jumped appreciably. The westbound link with the highest freight traffic volume was between Johnstown and Harrisburg with 19.3 million tons per year. The same link had the highest eastbound flow of 35 million tons per year. On route two, the St. Louis-Indianapolis link had 12.7 million tons per year westbound, and the Pittsburgh-Wheeling link had 12 million tons per year eastbound. Such freight traffic volumes were more than enough to absorb the capacity of the assumed pipeline system described above.

The last case was the less restrictive trans-shipment one (shown in Table IV). The maximum link westbound on route one was Gary-South Bend with 22.8 million tons per year; eastbound was Johnstown-Harrisburg with 38.3 million tons per year. On route two, St. Louis-Indianapolis had 16.7 million tons per year westbound and Wheeling-Pittsburgh had 13.3 million tons per year eastbound.

Table III

Tonnage Flows on Links of Hypothetical Network

(Assuming Restrictive Trans-Shipment)

Link	Westbound	Eastbound
Chicago-Gary	13,866,300	15,340,952
Gary-South Bend	17,244,165	18,266,837
South Bend-Toledo	17,257,962	21,358,207
Toledo-Cleveland	19,038,286	24,454,146
Cleveland-Akron	17,928,247	26,166,842
Akron-Youngstown	16,126,938	24,963,733
Youngstown-Pittsburg	15,479,679	24,797,928
Pittsburgh-Johnstown	19,121,068	34,737,669
Johnstown-Harrisburg	19,281,569	34,996,496
Harrisburg-Philadelphia	15,989,233	31,238,377
Philadelphia-New York	10,954,520	22,871,088
St. Louis-Indianapolis	12,745,334	7,320,096
Indianapolis-Dayton	11,632,851	8,735,450
Dayton-Columbus	8,319,109	10,832,647
Columbus-Wheeling	6,541,842	10,701,956
Wheeling-Pittsburgh	7,089,913	12,021,107

Source : Calculated from non-perishable origin-destination matrix and Reebie (8).

Table IV

Tonnage Flows on Links of Hypothetical Network (Assuming Less Restrictive Trans-shipment)

Link	Westbound	Eastbound
Chicago-Gary	18,902,961	15,710,070
Gary-South Bend	22,841,550	18,867,784
South Bend-Toledo	22,728,367	22,604,704
Toledo-Cleveland	20,445,316	27,339,215
Cleveland-Akron	19,400,231	26,922,747
Akron-Youngstown	17,645,713	26,829,437
Youngstown-Pittsburgh	17,104,720	26,458,050
Pittsburgh-Johnstown	21,366,466	38,725,380
Johnstown-Harrisburgh	22,120,326	38,252,712
Harrisburgh-Philadelphia	17,423,692	33,394,347
Philadelphia-New York	11,900,256	23,537,392
St. Louis-Indianapolis	16,746,400	8,412,371
Indianapolis-Dayton	12,975,805	9,385,041
Dayton-Columbus	13,917,116	11,663,698
Columbus-Wheeling	7,275,878	11,341,083
Wheeling-Pittsburgh	8,667,620	13,292,110

Source: Calculated from non-perishable origin-destination matrix and Reebie (8).

Although seasonality could be a problem when dealing with yearly flow data, e.g., if the eastbound flows occurred during one part of the year and the westbound flows at another as opposed to balanced flows in direction and time, the available data did not contain information on such seasonality. However, the three scenarios shown above did point out a problem faced by a system using containers/vehicles. The eastbound flows dominated the westbound flows on many links by a ratio which in some cases approached 2 to 1. This implied a potential empty backhaul problem for the container/vehicles. If the flows were balanced with respect to direction and with respect to timing such that equipment would be fully utilized, the hypothesized network could carry a given amount of traffic at a minimum cost.

The authors were somewhat skeptical of the above data base. Past experience with expanding the 1% rail waybill to represent complete enumeration had produced flows which were not feasible. The comparison of estimated rail complete enumeration with QCS data was also bothersome. Others have been skeptical about the validity of the import-export inland destinations-origins data (see Balfe et. al., (22)). It should be noted that real goods imports plus exports have ranged between 12 and 20 percent of the real goods, i.e., excluding services, Gross National Product in recent years. Thus the imports-exports portion of the data base was not nearly as important as the totally domestic portion of the data base. Any controversy regarding the accuracy of the domestic data. The Census of Transportation contained large sampling variability for many commodity codes. Thus, it was entirely possible that if all estimates of traffic for a given O-D pair were fortuitously on the high side, then the cumulative effects of such errors could be quite an overstatement. Of course, an understatement was also possible, or the errors could be countervailing ones. Based on Census of Transportation flows that were developed for this project, the above flows seemed high. However, it should be remembered that the flows of minerals and ores have not been added. This shall be done below.

The Reebie study (8) data were attractive because they represented containerizable freight, most of which was estimated to be pipelineable. However, they unfortunately suffered from a lack of commodity detail. At one time in its compilation, it was apparently on a four digit STCC basis. At another juncture, it was in a 76 sector input-output code. Flow data on such a comprehensive geographical basis as in the Reebie study (8) and also commodity disaggregated was unheard of up until now. Unfortunately, it appears that once a commodity or portion thereof was deemed to be prime or suitable with respect to containerizability, the total containerizable was noted and the commodity identification discarded. This resulting lack of commodity identification made the modal split task of this project more difficult as is explained below.

Nevertheless, this first cut at the flow data yielded some "ball park" figures based on the best available data suggesting that significant traffic flows which were pipelineable currently existed on an east to west axis from Chicago and St. Louis to New York. Other significant corridors could also exist, e.g., Cleveland, Erie, Buffalo, Rochester, Syracuse, Albany, Springfield, Boston, or an Eastern seaboard north-south corridor, Atlanta to Boston. This study has concentrated on the east-west corridor because it involved the largest flows.

The concept of trans-shipment loomed large in the above network analysis. This held true with regard to both the restrictive and less-restrictive estimates (involving cities not on the network) and also with regard to the network cities (since the cities are not points). The hypothesized network cannot directly serve all of the traffic postulated. Thus, the shipments would be brought to a pipeline facility. The ease and cost of this access and egress to and from the network will play a very important role in the competitiveness of this system vis-a-vis the existing modes.

The Reebie study (8) data specifically excluded STCC's 10 (ores), 11 (coal), and 14 (minerals) as being non-containerizable. However, such commodities were considered to be pipelineable and currently represent the major flows of non-petroleum movements by pipeline.

Two attempts were made to add such flows to the analysis. The first used the Department of the Interior's Bureau of Mines (20) data on bituminous coal and lignite distribution. The second used the state-to-state flows of STCC's 10, 11, and 14 from the 1972 1% rail waybill sample.

The coal data were for 1974. These data were used instead of 1971 data (Reebie's data were normalized to 1971) because coal flows have grown since 1971 and the intent was to determine the most up-to-date information on flows as possible.

Although 23 coal producing regions were identified by the Department of the Interior, due to aggregation and omission, it was possible to identify 20 origin areas. Destinations were given as states and only Louisiana data was omitted of the 48 contiguous states. A 20 by 47 origin-destination matrix for bituminous coal and lignite was then constructed.

The matrix values were then shrunk to account for just rail flows of coal (modal split information is given by the Bureau of Mines). It was decided that truck hauls of coal (0-50 miles) were much too short to be hauled by an intercity pipeline system and that movement by waterway was a competitor whose low price ≅ .3¢ per ton-mile would be difficult to reach.

The 20 coal mining areas were assigned cities on the network as origin (or feeder points). Although a nontrivial amount of coal is currently trucked to a rail or water head, the distance involved is short and a feeder type of arrangement (over long distances) from a hub city is not too likely (although the analogous principle of gathering and delivering lines in oil pipelines could be explored). Thus, because of limited pickup and delivery possibilities, it was decided to limit coal origins <u>and</u> destinations to the network area and destinations to network cities.

Specifically, coal originating in the Eastern Pennsylvania District (District #1) was assigned to Johnstown; coal originating in the Western Pennsylvania District (District #2 was assigned to Pittsburgh; coal originating in the Ohio District (District #4) was assigned to Columbus; coal originating in the Panhandle District (District #6) was assigned to Wheeling; coal originating in Illinois (District #10) was assigned to St. Louis; and coal originating in Indiana (District #11) was assigned to Indianapolis. Although some of the other 14 districts (especially Southern Districts #1 and #2) originated coal which terminated in the states served by the pipeline, it was decided to exclude those flows on the basis of a long feed required to reach a potentially originating network city.

Destination flows were based on states. To assign state flows to a particular city, two approaches were followed. The first identified the location of iron and steel workers in each state by SMSA (from the Census of Manufacturers). Then all coal shipments by rail which went to the "all others" category which Interior reports were assumed to be split on the basis of SMSA share of employment in iron and steel production.

The second approach (and the one which represented by far the greatest tonnage) was to assign the remaining rail flows of coal to electric utilities (the chief consumer, to coke and gas plants, and to retail dealers) to the SMSA's of each state on the basis of each SMSA's share of total SMSA population of the state.

From the two approaches, only the flows to network cities from network cities were noted. The flows are shown in Table V. The linkage flows contain a large number of zeros due to the limited number of origins of coal and the tendency for coal to be consumed fairly near where it is produced, e.g., most Indiana coal is consumed in Indiana. In addition, since most coal consumption occurs in the same state where it is produced (and for Illinois, Indiana, and Ohio, coal production is in the southern part of the state and many consumption areas are in the northern part of the state), many(north-south) flows are excluded from the east-west pipeline network assumed herein.

Unfortunately, such detailed flows were not available for STCC's 10, 14 and anthracite coal. However, making the same presumption as with coal concerning truck and water movement, if rail movements were available, an estimate of pipelineable STCC's 10, 40, and anthracite could be made. Unlike bituminous coal and lignite,

Table V

Flows of Bituminous Coal and Lignite By Rail (1974)
Network Cities Only (Tons)

	Eastbound	Westbound
New York-Philadelphia	7,650,443	0
Philadelphia-Harrisburg	20,732,038	0
Harrisburg-Johnstown	22,077,690	0
Johnstown-Pittsburgh	13,954,324	1,434,043
Pittsburgh-Youngstown	0	5,574,513
Youngstown-Akron	0	5,071,900
Akron-Cleveland	0	4,823,850
Cleveland-Toledo	0	3,922,315
Toledo-South Bend	0	598,860
South Bend-Gary	0	437,127
Gary-Chicago	0	48,904
Wheeling-Pittsburgh	8,769,112	1,505,510
Columbus-Wheeling	0	3,512,329
Dayton-Columbus	0	2,064,778
Indianapolis-Dayton	0	721,555
St. Louis-Indianapolis	0	79,510

Source: Calculated by the authors from U.S. Department of the Interior (20) Table I.

a complete enumeration of rail movements was not available. However, state-to-state movements of five digit STCC's were available from the 1972 1% waybill sample. Taking these flows and expanding them to represent complete enumeration yielded a very rough estimate (subject to the difficulties mentioned above) of the rail (hence pipelineable) flows of minerals, ores, and anthracite coal.

The assignment of origin and destination was done in a comparable method to the coal assignment. The Mineral Yearbook yielded potential countries of origin for many ores and minerals in the states. From the rail waybill sample, an estimate of the distance travelled by the shipments could be made - - call it X. With the likely origin (county of production or port), the state of destination, and the distance from origin and destination, likely destinations in the destination states could be made, i.e., what cities in the destination state were X miles from the likely origin?

Given the above methodology, assignment of origin and destination to the expanded flows was made. When both likely origin and likely destination were network cities (or close to a network city) in an east-west alignment, the flow was assigned to the pipeline network. Table VI shows the results. North-south movements and the general lack of mineral and ore production in the pipeline territory tended to keep the flows low.

Table VII shows the sum of Table V and VI. Table VIII gives the combination of Table II and Table VII. Thus it represents the restrictive network (or almost network) area flows. As can be seen, the addition of the STCC 10, 11, and 14 flows made the Wheeling-Pittsburgh and the Columbus-Wheeling links more significant (in terms of 14 million tons per year total flow). While the flows on the Dayton -Columbus link were large, the flows on the Dayton-Indianapolis link and especially the Indianapolis-St. Louis link remained low. Only when the less restrictive feeder relationship was allowed did these latter two links generate significant traffic volumes.

Nevertheless, this rough cut shows that significant levels of traffic of physically pipelineable goods could move in the area served by the hypothetical network.

The Census of Transportation Flows

To date there have been three Census of Transportation taken - - 1963, 1967, and 1972. There is approximately a three year gestation period before the information needed to construct flows is released to the public. At the time when this research began (June, 1975), the 1972 information was not yet released. Thus, the 1967 infor-

Table VI

Flows of Minerals, Ores, and Coal (Except Bituminous) Coal and Lignite) By Rail (1969) Network Cities Only

(Tons)

	Eastbound	Westbound
New York-Philadelphia	672,200	362,700
Philadelphia-Harrisburg	847,400	736,500
Harrisburg-Johnstown	218,500	1,333,100
Johnstown-Pittsburgh	215,500	1,270,000
Pittsburgh-Youngstown	206,900	726,200
Youngstown-Akron	169,200	561,000
Akron-Cleveland	169,200	591,800
Cleveland-Toledo	39,000	433,600
Toledo-South Bend	13,200	388,100
South Bend-Gary	13,200	148,700
Gary-Chicago	3,000	131,900
Pittsburgh-Wheeling	38,800	414,800
Wheeling-Columbus	54,800	548,000
Columbus-Dayton	51,700	346,500
Dayton-Indianapolis	51,700	346,500
Indianapolis-St. Louis	39,600	308,000

Source: Calculated by the authors from the 1972 1% Rail Waybill Sample - 5 digit STCC level
See text for description of STCC's considered.

Table VII

Flows of All Minerals, Ores, and Coal To and From Hypothetical Network Cities Only (Tons)

	Eastbound	Westbound
New York-Philadelphia	8,322,643	362,700
Philadelphia-Harrisburg	21,579,438	736,500
Harrisburg-Johnstown	22,296,190	1,333,100
Johnstown-Pittsburgh	14,169,824	2,704,043
Pittsburgh-Youngstown	206,900	6,300,713
Youngstown-Akron	169,200	5,632,900
Akron-Cleveland	169,200	5,415,650
Cleveland-Toledo	39,000	4,355,915
Toledo-South Bend	13,200	986,960
South Bend-Gary	13,200	585,827
Gary-Chicago	3,000	180,804
Wheeling-Pittsburgh	8,807,912	1,920,310
Columbus-Wheeling	54,800	4,060,329
Dayton-Columbus	51,700	2,411,278
Indianapolis-Dayton	51,700	1,068,055
St. Louis-Indianapolis	39,600	387,510

Source: Calculated from Tables V and VI

Note: 1974 and 1972 data were mixed because coal flows are known to have increased over time due to the energy crisis while 1972 data was the last year available for the other products.

Table VIII

Tonnage Flows on Links of All Pipelineable Commodities To and From Hypothetical Network Cities Only (Tons)

	Eastbound	Westbound
New York-Philadelphia	20,043,308	6,915,661
Philadelphia-Harrisburg	37,768,822	10,282,888
Harrisburg-Johnstown	37,403,188	10,957,367
Johnstown-Pittsburgh	28,978,409	12,181,308
Pittsburgh-Youngstown	11,762,655	14,058,684
Youngstown-Akron	11,955,379	13,981,627
Akron-Cleveland	12,805,872	15,191,407
Cleveland-Toledo	11,271,054	14,026,781
Toledo-South Bend	6,989,179	8,383,848
South Bend-Gary	7,124,056	8,225,983
Gary-Chicago	4,804,766	5,500,271
Wheeling-Pittsburgh	12,360,864	3,886,209
Columbus-Wheeling	3,070,257	5,600,463
Dayton-Columbus	2,604,138	3,769,488
Indianapolis-Dayton	1,394,271	2,105,136
St. Louis-Indianapolis	773,811	1,329,875

Source: Calculated from Tables II and VII.

mation was used. As discussed later in Table XVI, the flows of selected pipelineable commodities (nine) derived from the 1967 production area were utilized to estimate transport demand.

During the research year, the 1972 production area data were released. Flows from the published data for 1972 were utilized to estimate the macro demand model shown in equation 2. The flows of STCC 291 (petroleum products) were omitted from the calculations since they were products which have already been proven to be economically feasible to be moved by pipeline. Since the Census information excluded movements by pipeline from its coverage, the STCC 291 movements which are excluded from the flows represent flows not now moving by pipeline but which are amenable to this mode of transport.

The Census of Transportation was limited in its coverage. The 1972 Census had 27 origins and 28 destinations (the 27 origins and an "all other" category). The flows were based on a sampling procedure and were "expanded" to represent the complete enumeration. The flows only accounted for manufactured goods flowing from firms with greater than 20 employees. Any flow of manufactured products from warehouses and distribution centers were excluded from the analysis. Since such hauls were likely to be over short distances (although this was by no means necessary), their absence may not be crucial from the freight pipeline point of view. Since only a portion of the country was included in the 27 production areas, some movements which originated and/or terminated on the network cities were excluded from the data base. In toto, neglecting sampling variability problems, it appeared that the Census would understate manufacturing flows in the study area. In addition, as one disaggregated commodity-wise, some flows were not given due to disclosure problems (the Census will not release data in a form which will enable the operations of an individual firm to be distinguished).

Using the published 1972 data, a 27 by 28 origin-destination matrix for all commodities (except STCC 291) was constructed. Some of the production areas used in the Census were much larger or covered different jurisdictions than the area used in the Reebie study (8) analysis above, e.g., the Philadelphia production area included Wilmington; the Cleveland production area included Akron, Youngstown, and Erie; the Pittsburgh production area included Wheeling; the Detroit production area included Flint, Toledo, and Ann Arbor; the Cincinnati production area included Dayton and Springfield; the Chicago production area included Gary; among others. The ancillary area mentioned with a production area title above were separate areas in the Reebie analysis (8).

The total flows involving just the network cities are shown in Table IX. The main diagonals were excluded to allow for strictly longer hauls. Table X gives the

Table IX
Flows of Manufactured Commodities Between the Hypothetical Network Cities, 1972
(1000's of tons)

To From	New York*	Philadelphia	Harrisburg	Pittsburg	Cleveland	Detroit	Chicago	St. Louis	Indianapolis
New York	x	2565	360	272	344	360	560	129	133
Philadelphia	4851	x	797	592	404	469	404	132	132
Harrisburgh	693	784	x	107	157	190	132	33	17
Pittsburgh	839	732	303	x	1852	1514	1237	177	177
Cleveland	1537	824	262	1906	x	5169	1648	z	z
Detroit	1993	907	90	519	2447	x	2916	z	z
Chicago	1138	622	285	622	1529	3544	x	z	z
St. Louis	359	155	19	535	z	z	z	x	290
Indianapolis	224	127	37	22	z	z	z	231	x

*Includes North Jersey

Source: Calculated by the authors from U.S. Department of Commerce, 1972 Census of Transportation, Production Area Series.

Table X

Flows of Manufactured Products on the Hypothesized Pipeline Network Links, 1972
(1000's of tons)

	Westbound	Eastbound
New York-Philadelphia	4,723 (1067)	11,634 (2629)
Philadelphia-Harrisburg	5,088 (1150)	10,934 (2471)
Harrisburg-Pittsburgh	4,567 (1032)	10,453 (2362)
Pittsburgh-Cleveland	7,623 (1723)	10,705 (2419)
Cleveland-Detroit	11,683 (2640)	10,152 (2294)
Detroit-Chicago	6,897 (1559)	7,740 (1749)
Pittsburgh-Indianapolis	930 (210)	1,768 (400)
Indianapolis-St. Louis	702 (158)	1,358 (307)

Source: Calculated from Table IX.

flows of total manufactured goods (less STCC 291) on the various links of the hypothetical pipeline. Of course, not all of these flows were physically pipelineable.

Table XI shows flows involving the network cities and assuming that the network cities are also feeder cities for off network cities, e.g., a Milwaukee to Boston could go Chicago to New York via pipeline. Table XII shows the loads on the links of the hypothetical pipeline.

What percent of this traffic was physically capable of going by pipeline? If the data was disaggregated by commodity, a large number of zeros appear in the matrix. Thus, to get the maximum amount of flow, the flows undifferentiated by commodity have been presented. Two digit STCC matrices could have been constructed, but time and expense precluded doing so.

A clue existed as to the pipelineable products in a paper done by Wallin and Frost (23) for the Forest Service, U.S. Department of Agriculture. Wallin and Frost estimated the national percentage of the nation's manufactured traffic which was palletizable based on the Census of Transportation for 1967, (they were interested in wood consumption for pallets). Their estimate was 20%. The current basic dimensions of a pallet load are 40" by 48" with a diagonal of 62.5" which will fit into a 72 inch pipeline.

Although the palletizable flow by commodity by origin-destination for the 1972 Census was not known, a rough idea of the type of commodity which is moving can be found by looking at the industrial base of each production area in the pipeline region. Aggregating major flow items (albeit without regard to destination) for the whole region enabled a determination of the share of each commodity type in total flows. Wallin and Frost's percent palletizable was then applied to the flow figures given above. As previously, the idea was a ball park estimate of traffic flows. When this percentage (approximately 22.6% for the Middle Atlantic and East North Central regions) was applied to Tables X and XII, the results were the columns in parentheses in Tables X and XII.

As can be seen, although the volume of tonnage was cut substantially from the Reebie flows, there was still a large amount of traffic which was palletizable i.e., physically pipelineable. While the magnitude was nowhere near the Reebie data magnitude, it must be remembered that the data set used for this exercise was just a subset of the Reebie data.

Wallin and Frost gave an actual origin-destination matrix for 1967 palletizable flows. These are shown in Table XIII (for just the network cities) and in Table XIV

---

(23) Wallin, W. and R. Frost, Production Flow in a National Pallet Exchange Service, Northeastern Forest Experiment Station, Forest Products Marketing Laboratory, U.S. Department of Agriculture, Princeton, West Virginia, 1973.

Table XI

Flows of Manufactured Commodities Involving the Hypothesized Network Cities As Generators and Feeders, 1972 (1000's of Tons)

From \ To	New* York	Phila-** delphia	Harris- burg***	Pitts- burgh****	Cleve- land	De- troit	Chi-+ cago	Cinci-++ natti	Indiana- polis	St.+++ Louis
New York*	x	3013	481	382	569	563	1664	233	202	1497
Philadel- phia**	6632	x	797	675	518	614	863	241	132	1013
Harris- burg***	1048	784	x	296	321	266	751	215	35	797
Pitts- burgh****	1367	934	787	x	1852	1514	2377	587	243	1854
Cleveland	2137	974	674	1906	x	5169	2510	z	z	z
Detroit	2707	997	404	519	2447	x	4646	z	z	z
Chicago+	2690	1117	891	1370	1764	4700	x	z	z	z
Cincinatti++	599	204	277	292	z	z	z	x	321	1023
Indianapolis	321	142	477	141	z	z	z	328	x	75
St. Louis+++	3498	1181	343	1004	z	z	z	795	552	x

*	Includes N.Jersey, Hartford, Boston
**	Includes Allentown
***	Includes Baltimore, Syracuse
****	Includes Buffalo
+	Includes Milwaukee, Minneapolis, Seattle, San Francisco
++	Proxy for Dayton
+++	Includes Dallas, Houston, Denver, Los Angeles, Kansas City

Source: Calculated by the authors from U.S. Department of Commerce, 1972 Census of Transportation, Production Area Series.

Table XII

Flows of Manufactured Products on the Hypothetical Pipeline Network Links, 1972
Less Restrictive Unadjusted
(1000's of Tons)

	Westbound	Eastbound
New York-Philadelphia	8,604 (1945)	20,999 (4746)
Philadelphia-Harrisburg	10,444 (2360)	20,700 (4678)
Harrisburg-Pittsburgh	11,847 (2677)	22,721 (5135)
Pittsburgh-Cleveland	11,872 (2683)	16,386 (3703)
Cleveland-Detroit	16,291 (3682)	14,906 (3369)
Detroit-Chicago	12,811 (2895)	12,532 (2832)
Pittsburgh-Dayton	7,049 (1593)	8,479 (1916)
Dayton-Indianapolis	7,117 (1608)	8,230 (1860)
Indianapolis-St. Louis	6,259 (1415)	7,373 (1666)

Source: Calculated from Table XI.

Table XIII

Millions of Tons of Palletizable Product
Network Cities Only, 1967

D O	N.Y.	Philadelphia	Harrisburg	Pittsburgh	Cleveland	Detroit	Chicago	St. Louis
New York	x	.62	.06	.10	.07	.10	.12	.03
Philadelphia	1.31	x	.16	.13	.12	.11	.10	.01
Harrisburg	.13	.18	x	.04	.02	.01	.02	.00
Pittsburgh	.61	.30	.10	x	.64	.29	.27	.05
Cleveland	.31	.16	.04	.34	x	.76	.27	z
Detroit	.27	.29	.02	.26	.48	x	.61	z
Chicago	.25	.07	.03	.08	.23	.67	x	z
St. Louis	.10	.04	.01	.02	z	z	z	x

Loads on Links of Hypothetical Pipeline

	Westbound	Eastbound
New York-Philadelphia	1.10	2.98
Philadelphia-Harrisburg	1.11	2.71
Harrisburg-Pittsburgh	.98	2.50
Pittsburgh-Cleveland	1.96	2.12
Cleveland-Detroit	2.05	1.98
Detroit-Chicago	1.39	1.37
Pittsburgh-St. Louis	.09	.17

x = no movement reported due to close proximity
z = no movement reported due to circuity

Source: Calculated by the authors from Wallin and Frost, 1973,pp. 68-69.

Table XIV

Millions of Tons of Palletizable Product With the Hypothesized Network Cities as Feeders, 1967

D O	N.Y.	Philadelphia	Harrisburg	Pittsburgh	Cleveland	Detroit	Chicago	Cincinnati	St. Louis
New York	x	.78	.09	.14	.13	.15	.43	.05	.27
Philadelphia	1.84	x	.16	.14	.14	.12	.18	.03	.11
Harrisburg	.20	.18	x	.11	.06	.03	.19	.02	.11
Pittsburgh	.78	.34	.30	x	.64	.29	.54	.25	.25
Cleveland	.43	.20	.11	.34	x	.76	.38	z	z
Detroit	.37	.35	.12	.26	.48	x	.80	z	z
Chicago	.70	.33	.26	.22	.33	.92	x	z	z
Cincinnati	.19	.03	.05	.05	z	z	z	x	z
St. Louis	.70	.27	.11	.06	z	z	z	.26	x

Loads on Links of Hypothetical Pipeline

	Westbound	Eastbound
New York-Philadelphia	2.04	5.21
Philadelphia-Harrisburg	2.14	5.07
Harrisburg-Pittsburgh	2.41	5.64
Pittsburgh-Cleveland	2.90	3.69
Cleveland-Detroit	3.07	3.42
Detroit-Chicago	2.52	2.76
Pittsburgh-Cincinnati	1.09	1.46
Cincinnati-St. Louis	.82	1.40

x = no movement reported due to close proximity
z = no movement reported due to circuity

Source: Calculated by the authors from Wallin and Frost, (23), pp. 68, 69, 70.

(for the network cities as feeders). In general, the 1967 flows were greater than the 1972 flows shown in Tables X and XII. Since traffic flows have grown between 1967 and 1972 (Allen 24), these results indicated that disaggregation by production area and commodity type (1967) is likely to result in larger flows than regional average palletizable (1972).

The Census data indicated that pipelineable flows existed such that pipeline would incur costs per ton mile of 4-5¢ for the lowest traffic volumes and 1.5¢ for the highest traffic volumes if all such traffic were moved by pipeline. Since other modes currently offer some service at lower rates than those assumed for pipeline, whether pipeline can win traffic will be a function of pipeline's service vis-a-vis the other modes. Since the pipeline showed pronounced economies of scale, to the extent that the pipeline cannot capture some of the physically pipelineable products, pipeline cost will rise. This, in turn, will reduce traffic diverted to pipeline. Therefore an iterative process will be necessary to determine the final traffic volume that the pipeline can sustain. The process of splitting modes will be spelled out in a later section.

Macro Demand Models

Once an understanding existed with regard to the estimated size of the total traffic physically pipelineable freight traffic pie, a determination of the potential share of that pie which could be reasonably expected to be carried by pipeline was made. This entailed a modal split analysis. Such analysis was necessary since although the flows under observation were assumed pipelineable, they are also railroadable, truckable, containerizable, etc. To understand modal split behavior, it was helpful to understand something about demand behavior.

The goal of this section is to analyze several macro demand flows. By macro is meant an aggregation over commodity type, over geographical area, and/or over shipper. This analysis contrasts with the micro demand analysis below which investigates transportation behavior of individual shippers, shipping specific commodities from specific origins to specific destinations. The intention of the micro and macro demand modal split sections of the report was to (1) model individual shipper's behavior, (2) compare and contrast the individual behavior with the behavior found in the macro analysis, (3) build a linkage between the micro and macro models, and (4) split the macro flows between pipeline and other modes.

---

(24) Allen, W.B., "Some Observations on Improving Railroad Productivity", Transportation Research Forum Papers, Richard B. Cross & Company, Oxford, Indiana, 1976.

The first attempt at a macro demand model involved an aggregation of all manufactured commodities in the production area series of the 1972 Census of Transportation. The model to be tested was a basic gravity model of the form

$$T_{ij} = K \frac{P_i^A P_j^B}{D_{ij}^C}$$

Where

T_{ij} = Number of tons flowing from origin i to destination j

P_i = Population of Region i (or manufacturing employment or value added)

P_j = Population of Region j (or manufacturing employment or value added)

D_{ij} = Distance from i to j (the airline distance from the cities which identify i and j)

K,A,B,C = Parameters estimated from the data

A basic discussion of the gravity model was found in Isard (25). In short, P_i is a push variable (generating flows from origin i), P_j is a pull variable (attracting flows to destination j), and D_{ij} is a proxy for transportation costs, which tend to inhibit movements from i and j.

Since population, manufacturing employment, and value added were all highly correlated (.92 level), they all basically gave the same results when estimated in equation (1). Equation (1), estimated in the log linear form of equation (2) i.e., taking the logarithm of both sides of equation (1), yielded equation (2).

(2) $\log T_{ij} = \log K + A \log P_i + B \log P_j - C \log D_{ij}$

with results of

(2') $\log T_{ij} = \underset{(.720)}{-2.099} + \underset{(.063)}{.799} \log P_i + \underset{(.060)}{.823} \log P_j - \underset{(.033)}{.919} \log D_{ij}$

where the values in parentheses are the standard errors.

As can be seen, the signs of A, B, and C were as expected and all coefficients were statistically significant at the 1% level. The coefficient of determination, i.e., R^2 was .655 meaning that 65% of the variability in the flow data was explained by the three gravity model variables. The F value was a significant 357. Thus the basic gravity model explained total flows quite well and yielded a rudimentary understanding of the demand for transportation. While no general rule exists with re-

---

(25) Isard, W., "Regional Commodity Balances and Interregional Commodity Flows", American Economic Review, Vol. 43, 1953, pp. 167-180.

gard to the magnitude of R^2 with the exception of 1.00 which signifies a perfect fit of the data and .00 which signifies no relationship of the dependent and independent variables, casual inspection of the literature utilizing such statistical estimation techniques will show that an R^2 of .65 is quite a good result.

These results basically confirmed the a priori theory that the demand for transportation traffic volume was a derived demand (since the numerator variables had positive signs) and was negatively influenced by transportation costs (since the distance variable had a negative sign).

The gravity model had been estimated (in the constrained sense below) on STCC 20 (food) flows by O'Sullivan and Ralston (26) and on total flows for 1967 by Black (27).

The O'Sullivan and Ralston model was run on a constrained form, i.e.,

$$(3) \quad T_{ij} = B_j O_i D_j d_{ij}^{-\beta}$$

$$\text{where} \quad B_j = \left[\sum_i O_i d_{ij}^{-\beta}\right]^{-1}$$

O_i = tons originating from origin i

D_j = tons terminating in destination j

d_{ij} = distance from i to j

β = parameter to be estimated

As shown by Wilson (28) and as in equation (3).

A doubly constrained program (Wilson, 28), has recently been made available to the authors

$$(4) \quad T_{ij} = A_i B_j O_i D_j d_{ij}^{-\beta}$$

$$\text{where} \quad A_i = \left[\sum_j D_j d_{ij}^{-\beta}\right]^{-1}$$

and the second year report will present its results on the 1972 data.

These gravity models enabled the authors to estimate traffic volumes between population centers in cases where actual flow information did not exist.

As Byler and O'Sullivan (29) show, gravity model parameters tend to be fairly

---

(26) O'Sullivan, P. and B. Ralston, "Forecasting Intercity Commodity Transportation in the USA. Regional Studies," Vol. 8, No. 2, 1974, pp. 191-195.

(27) Black, W., "Interregional Commodity Flows: Some Experiments with the Gravity Model", Journal of Regional Science, Vol. 12, #1, 1972, pp. 107-118.

(28) Wilson, A.G., Entropy in Urban and Regional Modeling, Pion Press, London, 1971.

(29) Byler, J. and P. O'Sullivan, "The Forecasting Ability and Temporal Stability of the Coefficients of Gravity Models Applied to Truck Traffic", Traffic Engineering Control, 1974.

stable over time. Dynamically stable gravity model parameters tend to suggest structual stability in the economy. Hence if traffic flows were to be projected into the future (which is potentially risky due to possible future structural changes in the economy), the gravity method could be a useful first approximation.

A Disaggregated Macro Approach

The next step was to take macro data and introduce transportation modes and attempt to explain the traffic volumes by modes between origin and destination. As explained above, the only information which contained numerous flows by commodity and also contained modal split information was the Census of Transportation (in public use computer tape form). Since it was expensive to use the tape, a research strategy had to be decided upon fairly early in the project. A large number of potential commodities existed that were physically pipelineable. If origin-destination matrices were constructed by mode for all such commodities, the computer budget for this project would have been overspent many times. In addition, if one attempted to achieve commodity detail, matrix cells with zero entries started to proliferate.

Thus, in consultation with the pipeline technology group, it was decided to investigate nine pipelineable commodities on a three and four digit STCC basis. While three digits was fairly aggregative, the trade-off in going to five (or in some cases even four) digits was an unacceptable number of zeros in the flow matrix. The nine three digit STCC's had a significant number of origins and destinations in the northeast and midwest region, a significant number of shipments greater than 10,000 pounds (to insure reasonable pipeline container loads) and a significant number of shipments greater than 75 (to insure a reasonable line haul distance). In addition, these commodities were selected to be similar with the micro demand/modal split interviewing with shippers which was undertaken and described below.

In this model, the objective was to explain the flows T_{ijh} (tonnage between i and j by mode h) on the basis of the assumed abstract transportation characteristics of each. There were inherent problems with this approach. As explained below, shippers exhibit idiosyncratic perceptions of the transportation modal services and also exhibit idiosyncratic corporate goals. Since the data utilized herein was more disaggregate than the data of the last section but still entailed different products and/or different shippers, the model developed and estimated herein implied that individual shippers could be assumed to have similar decision criteria for modal selection. While individual shippers do, in fact, have fairly systematic decision criteria, the modal split of an aggregation of such shippers as in the Census of

Transportation Data, with different, but systematic, criteria may not appear to be orderly.

Given the modal choice explanations for these commodities, the addition of the pipeline mode as an alternative to the shippers (who chose between truck and rail in the above data) was hypothesized. The modal choice model hypothesized choice based on modal characteristics (ideally shipper and shipment characteristics should also be considered -- see Allen and Moses, (5); Gilmour, (30) -- for a list over 30 modal choice variables, and Allen (5). Consultation with the pipeline technology group yielded values for comparable modal characteristics for pipeline (e.g., rate, transit time). The modal split analyis allowed the determination of how the traffic would split between the three modes using the methodology of Quandt and Baumol(31). The details of the models and the modal splitting are given below.

The Abstract Mode Model

The abstract mode model was first proposed by Quandt and Baumol and has been used to estimate demand for passenger transportation (31). Relatively little work has been devoted to applying this approach for estimating demand for freight transportation (32). However, as will be seen later, this approach was ideal for estimating the potential demand for a new mode of transportation (such as the freight pipeline). The abstract mode approach viewed the intermodal demand for various transportation modes not directly, but as being derived from a demand for the "attributes" of these

---

(30) Gilmour, P., "An Evaluation of the Marketing Strategy for Transportation Services", Mimeo, Department of Economics, Monash University, Melbourne, Australia, 1975.

(31) See R.E. Quandt and W.J. Baumol, "The Demand for Abstract Transport Modes: Theory and Measurement", Journal of Regional Science, Vol. 6 #2, 1966, Reuben Gronau and Roger E. Alcaly, "The Demand for Abstract Transport Modes: Some Misgivings", JRS, Vol. 9 #1, 1969, R.E. Quandt and W.J. Baumol, "The Demand for Abstract Transport Model: Some Hopes", JRS, Vol. 9 #1, 1969, E. Philip Howrey, "On the Choice of Forecasting Models for Air Travel", JRS, Vol 9 #2, 1969, and Kan Hua Young, "The Abstract Mode Approach to the Demand for Travel", Transportation Research, Vol. 3, 1969.

(32) See Brian C. Kullman, "A Model of Rail/Truck Competition in the Intercity Freight Market", Volume 15 in Studies in Railroad Operations and Economics, prepared for the Federal Railroad Administration, December, 1973, and James H. Herendeen, "Theoretical Development and Preliminary Testing of a Mathematical Model for Predicting Freight Modal Split", Pennsylvania Transportation and Traffic Safety Center, Pennsylvania State University, 1969.

modes (33). For example, a shipper's decision to utilize rail or truck did not depend on what these modes were called. Rather, his decision to ship by a particular mode was based on the levels of service (e.g., rates, transit times, reliabilities, probabilities of loss and damage, etc.) provided by that and other modes.

An abstract mode model of the intermodal demand for freight transportation can be formulated as

$$T^{\ell}_{ijk} = AX_i^{a_1} \; X_j^{a_2} \; (R_k/R_b)^{b_1} \; R_b^{b_2} \; (TT_k/TT_b)^{b_3} \; TT_b^{b_4} \; (Rel_k/Rel_b)^{b_5} \; Rel_b^{b_6} \qquad (5)$$

where T^{ℓ}_{ijk} was the flow of commodity ℓ by mode k from origin i to destination j; X_i and X_j were vectors of exogeneous variables representing production and consumption at origin i and destination j (i.e., population, manufacturing employment, value-added in manufacturing, etc.), R_k was the rate charged by mode k between i and j, R_b was the rate charged by the best (cheapest) mode between i and j, TT_k was the transit time provided by mode k between i and j, TT_b was the best (fastest) transit time between i and j, Rel_k was the transit time reliability provided by mode k between i and j, Rel_b was the transit time reliability between i and j of the best (most reliable) mode, and A, a_1, a_2, b_1, b_2, b_3, b_4, b_5, and b_6 were the parameters to be estimated. Equation (5) accounted for both freight generation and model split. The freight generation model could be written as:

$$T^{\ell}_{ij} = AX_i^{a_1} \; X_j^{a_2} \; R_b^{b_2} \; TT_b^{b_4} \; Rel_b^{b_6} \qquad (6)$$

where T^{ℓ}_{ij} was the total flow of commodity ℓ from origin i to destination j. The form of this model was very similar to that for gravity model formulation. The flow between origin i and destination j was a function of the relative attractiveness of the origins and destinations (X_i, X_j) and of the "friction" between them (R_b, TT_b, Rel_b). Thus, if a new mode was introduced that reduced the "friction" between i and j, say for example, by changing a lower rate, the total flow would be expected to increase. The estimation of freight generation and model split together (as in equation (5)) was extremely complex and will not be attempted here. Instead, we concentrated on just the modal split in the existing market. However, it must be remembered that the introduction of a new mode will lead not only to a new division of existing flows, but also may lead to an increase in total flows (by stimulating latent demand).

Here, we were interested in estimating the modal split between truck and rail.

---

(33) Lancaster, K., "A New Approach to Consumer Theory", Journal of Political Economy, 74, 1966, pp. 132-157.

It is important to note that this was essentially a binary mode choice. Consequently, in the spirit of the abstract mode model, the following equation was formulated:

$$(\text{Pertruc})^{\ell}_{ij} = (\text{Const.})\ (\text{TR/RR})^{\alpha}_{ij}\ \ (\text{RT/TT})^{\beta}_{ij}\ \ (\text{RRel/TRel})^{\gamma}_{ij} \qquad (7)$$

where $(\text{Pertruc})^{\ell}_{ij}$ was the percent of the flow of commodity ℓ going by truck between origin i and destination j, $(\text{TR/RR})_{ij}$ was the truck rate relative to the rail rate between i and j (the rail rate was almost always the least costly), $(\text{RT/TT})_{ij}$ was the rail time relative to the truck time between i and j (truck time was almost always the fastest), $(\text{RRel/TRel})_{ij}$ was the rail reliability relative to truck reliability between i and j (truck reliability was almost always the best), and (Const.), α, β, and γ were the parameters to be estimated. Various surveys have indicated that rates, transit time, and transit time reliability were the three most important factors in the modal decisions made by shippers (<u>34</u>). Thus, these variables were included in equation (7). Also the multiplicative form of this function implied that a shipper traded off relative rates, transit times, and reliabilities in making his modal selections. Thus; there appeared to be a behavioral basis for this multiplicative formulation.

Equation (7) can be estimated using ordinary least squares (OLS) assuming a multiplicative error and taking the logarithms of both sides, i.e.,

$$\begin{aligned} \text{Ln}(\text{Pertruc})^{\ell} &= \text{Ln}(\text{Const.}) + \alpha\text{Ln}(\text{TR/RR}) + \beta\text{Ln}(\text{RT/TT}) \qquad (8) \\ &+ \gamma\,\text{Ln}(\text{RRel/TRel}) + \mu \end{aligned}$$

where μ is a random error. It is well known that estimates of (Const.), α, β, γ, obtained by applying OLS to equation (8) will be asymtotically unbiased. However, these estimates will not be asymtotically efficient (least variance of all estimators). This was because the dependent variable (Pertruc) was bounded by 0 and 100, and thus one of the basic assumptions of least squares regression, homoscedasticity, was violated. It can also be shown in these estimates that the estimated variances of the estimated coefficients were biased and therefore the usual significance tests cannot be applied. However, this problem can easily be avoided by taking the logistic transformation of the dependent variable, obtaining

---

(<u>34</u>) For example, see the surveys identified by Kullman, B.C., "A Model of Rail/Truck Competition in the Intercity Freight Market", Vol. 15 in Studies in Railroad Operations and Economics, prepared for Federal Railroad Administration, December, 1975.

(Pertruc/100 - Pertruc), and using the logarithm of this variable as the dependent variable in equation (8). The dependent variable will now be unbounded, and the estimated coefficients will be both asymtotically consistent and unbiased. Once equation (7) had been estimated, the abstract service characteristics of the freight pipeline (rates, transit time, reliability) could be substituted in to predict the percentage of commodity ℓ that would go by this new mode. This will be illustrated in a later section of the report entitled Potential Demand for the Freight Pipeline.

The Data and Variables - Manufactured Commodities

Data for the dependent variable in equation (7) was obtained from the 1967 Census of Transportation--Commodity Transportation Survey public use tapes. The Census data showed flows from 25 production areas (one or a combination of SMSA's) to 59 destination areas (one or a combination of SMSA's and nine geographic regions) of manufactured commodities by Standard Transportation Commodity Code (on the 2, 3, 4, and 5 digit levels). Unfortunately, because of disclosure problems, data on the four and five digit levels were available on only a few origin-destination pairs. Thus, the greatest possible disaggregation by commodity was on the three (sometimes, four) digit STCC level. Pertruc was calculated by eliminating all less-than-truckload shipments (less than 10,000 pounds), and dividing truckload tonnage by total tonnage (truckload and rail carload) in a given origin-destination pair. This calculation was performed for nine three digit commodity groups which were deemed "pipelineable" (could physically move by a freight pipeline (35)). Among these commodity groups transportation by rail and truck was dominant-water, air, and other modes were extremely insignificant. The nine commodity groups chosen in this analysis by STCC were 208 (Beverages or Flavoring Extracts), 264 (Converted Paper or Paperboard Products), 265 (Containers or Boxes, Paperboard), 283 (Drugs and Medicines), 284 (Soaps and other Detergents), 285 (Paints and Allied Products), 307 Miscellaneous Plastic Products), 346 (Metal Stampings), and 364 (Electric Lighting or Wire Equipment).

For each origin-destination pair, it was necessary to calculate the truck rate and rail rate by commodity groups. Rail rates were calculated from nine regressions estimated on data from the 1969 Carload Waybill Statistics (36). Parameters for

---

(35) This was determined in consultation with the Technology Group.

(36) 1969 Carload Waybill Statistics, U.S. Department of Transportation, Federal Rail Administration, Office of Economics,

the following equation were estimated for each commodity group:

$$RR = A(SS)^{b_1}(Haul)^{b_2} \qquad (9)$$

where RR was the rail rate in cents per hundredweight, SS was the shipment size in tons per carload, and Haul was the shipment distance in miles. The multiplicative form of this function allowed both a quantity and distance taper in the rate structure. As shipment size increased, the rail rate was expected to decrease at a decreasing rate--b_1 was expected to be negative. As Haul increased, the rail rate was expected to increase, but at a decreasing rate--b_2 was expected to be positive. This equation was estimated on a sample of state-to-state movements by five digit STCC within each commodity group. The logarithm of both sides of equation (9) was taken, and A, b_1, and b_2 were estimated using ordinary least squares. The results of these regressions are presented in Table XV. It can be seen that the explanatory power of these regressions was fairly high, the estimated coefficients were usually significant, and b_1 and b_2 always had the expected signs.

Truck rates were calculated from the following equation which was estimated in three shipment size groups on data provided in Morton (37).

$$TR = A + b \ (1/Haul) \qquad (10)$$

where TR was the truck rate in cents per ton-mile, Haul was the length of haul in miles, and A and b were the parameters to be estimated. The form of this equation implied no distance taper in average rates (i.e., in cents/ton). However, as can be seen in Table XV, the three regressions fitted the data extremely well. The coefficient b was always positive and significant (at the 1% level). These regressions were estimated on data for the Northeast during 1967. Thus, it may not be valid to calculate truck rates for the entire country from these equations. Also, equation (10) was not estimated by commodity groups as were the rail rate regressions. However, Morton has shown that truck rates vary significantly with the commodity shipped(37). Still, the regressions had to be used because they were estimated on the best (and it seems only) available data concerning truck rates.

To calculate rail and truck rates (and rail and truck times) it was necessary to obtain data on distance between the 25 production areas and the 50 destination areas by rail and highway. The nine geographical regions were not included in this

---

(37) Morton, A.L., "Competition in the Intercity Freight Market", Office of Systems Requirements, Plans, and Information, Department of Transportation, Washington, D.C., 1971.

Table XV
Rate Regressions

Rail Rate Regressions

Commodity Group	A	b_1	b_2	F	R^2	N=
STCC 208 Beverages, etc.	2.654 (1.79)	-0.131 (-1.18)	0.567 (9.77)**	48.17**	.60	67
STCC 264 Converted Paper, etc.	13.343 (5.49)	-0.662 (-4.46)**	0.602 (11.41)**	66.01**	.63	.79
STCC 265 Containers, etc.	13.158 (10.28)**	-0.765 (-12.42)**	0.662) (16.31)**	159.00**	.86	53
STCC 283 Drugs, etc.	16.119 (6.801)**	-0.476 (-4.89)**	0.520 (13.87)**	106.03**	.77	65
STCC 284 Soap, etc.	1.583 (0.98)	-0.463 (-3.97)**	0.851 (15.72)**	123.56**	.85	46
STCC 285 Paint, etc.	1.972 (2.15)*	-0.133 (-0.188)	0.663 (16.78)**	141.59**	.84	56
STCC 307 Misc. Plastics	33.717 (14.21)**	-0.737 (15.61)**	0.525 (14.701)**	210.78**	.81	100
STCC 346 Metal Stampings	17.374 (11.56)**	-0.556 (-10.73)**	0.559 (14.26)**	122.62**	.81	62
STCC 364 Electric Lighting	17.392 (9.47)**	-0.751 (-9.89)**	0.659 (15.85)**	149.44**	.83	64

Truck Rate Regressions

SS Group	A	b	F	R^2	N=
10-12.5 tons	3.151**	939.218**	1656.70**	.99	22
12.5-15 tons	3.038**	762.683**	1163.33**	.98	21
15-20 tons	2.470**	684.973**	1565.65**	.99	21

Figures in parentheses are t-statistics
* Significant at the five percent level
** Significant at the one percent level

analysis. Both rail and highway distances were measured from the center of the largest city in each production or destination area. Rail short line miles were obtained from the 1967 Rand McNally Commercial Atlas because rail distance would not be expected to change greatly between 1967 and 1976. On the other hand, highway distances have changed during this period mainly because of the construction of the Interstate Highway System. Thus, highway short line miles were obtained from the 1970 Rand McNally Highway Atlas. The average circuity over short line miles in the United States has been estimated to be 16 percent by rail and 6 percent by truck (38). Thus, truck rates were calculated from equation (10) by multiplying truck short line miles by 1.06 and using the rate regression which includes the average truck shipment size (39) in that commodity group. Rail rates were calculated from the rate regressions reported in Table XV by multiplying rail short line miles by 1.16 and inserting for shipment size the average for that commodity group (40). Both rail and truck rates were converted to dollars per ton.

Rail transit times were calculated from the following regression which was estimated by Martland (41):

$$\text{Days} = 1.2 + 0.0007\ \text{Miles} + 0.72\ \text{Hump} + 0.63\ \text{Flat} + 0.39\ \text{Inter} + 0.45\ \text{Unrel}$$

$$R^2 = .55, \quad F = 30.3 \tag{11}$$

(All coefficients significant at the five percent level, F significant at the one percent level) where Miles was the rail miles (again rail short line miles were multiplied by 1.16 to account for circuity), Hump was the number of hump yards (one every 250 miles, the national average was assumed), Flat was the number of flat yards (two assumed), Inter was the number of railroad interchanges (one assumed), and Unrel was a measure of transit time unreliability. Unrel equalled (100-Rel), where Rel was transit time reliability measured as the per-

---

(38) Bureau of Accounts, I.C.C., Rail Carload Cost Scales, 1973, Statement No, ICI-73, p. 131, and Bureau of Accounts, I.C.C., Cost of Transporting Freight by Class I and II Motor Common Carriers, 1971, Statement No. 2CI-71. p. 4.

(39) Average truck shipment sizes obtained from Freight Commodity Statistics of Motor Carriers of Property, 1967, I.C.C., Bureau of Accounts.

(40) Average rail shipment sizes obtained from the 1969 Carload Waybill Statistics, U.S. Department of Transportation, Federal Rail Administration, Office of Economics.

(41) Martland, C.D., Rail Trip Reliability: Evaluation of Performance Measures and Analysis of Trip Time Data, Studies in Railroad Operations and Economics, Vol. 2, prepared for the Federal Railroad Administration, DOT, June 1972, p. 87.

centage of shipments that arrive during the best three day period (80 percent, the national average was assumed). Truck transit times were calculated from the following equation:

$$\text{Days} = \text{Miles}/450 \qquad (12)$$

where Miles was the highway short line miles multiplied by 1.06 to account for circuity. The demoninator in the above equation was obtained by assuming a truck would travel a 45 mile per hour (the national average) (42) for ten hours per day.

As mentioned before, transit time reliability appeared to be one of the three most important factors in the shipper's modal choice. In fact, reliability may be the most important factor in this decision. In one of the shippers' surveys cited by Kullman (34), 73 percent of the shippers indicated dependability of delivery as very important in the selection of their transport mode. This was compared to 67 percent of the shippers who cited total transit time, 55 percent who cited freight rates, and 40 percent who cited loss and damage as very important in their modal selection. One possible measure of transit time reliability (as mentioned before) was the percentage of shipments that arrive during the best three day period. Nationally this figure appeared to be about 95 percent for truck and 80 percent for rail. Unfortunately there seemed to be no way to estimate this variable for each origin-destination pair. Thus, truck and rail reliability had to be excluded from the abstract mode model. Equation (8) was written as

$$\text{Ln}(\text{Pertruc})^{\ell} = \text{Ln}\ (\text{Const.})' + \alpha'\text{Ln}(TR/RR) + \beta'\text{Ln}(RT/TT) + \mu' \qquad (13)$$

Clearly equation (13) was mis-specified since the variable rail reliability relative to truck reliability was not included. If this excluded variable was correlated with the included variables, relative rates and relative transit times, the estimates of α' and β' will be biased. However, it seemed that this was probably not the case. It must be remembered that although equation (7) was based on individual(shipper) behavior, it was estimated on extremely aggregate data. Thus, although the range of modal reliabilities facing the shipper may be extremely large, the ratio of rail reliability to truck reliability may have very little

---

(42) Interstate Commerce Commission, Cost of Transporting Freight by Class I and II Motor Common Carriers of General Commodities, 1971, Statement 2C1-71, Washington, D.C., 1972, p. 51.

variance over the extremely aggregate origin-destination pairs. If this was the case, the constant term would pick up the reliability advantage of truck.

Empirical Results - Manufactured Commodities

Before presenting the results of the estimation of our abstract mode split model for the nine commodity groups, the implications of the aggregation problem in this case is briefly discussed. The dependent variable (Pertruc) represented an aggregation over many shipments of various commodities within each commodity group. On the other hand, rail rates, truck rates, rail transit time, and truck transit time all represented average values in that commodity group and for a particular origin-destination pair. It was very possible that the mix of commodities being shipped within each commodity group varied greatly over the origin-destination pairs. Since shippers of different commodities were expected to respond very differently to relative rates and relative transit times, this might lead to lower explanatory power for the mode split equation. Also, the Census production and destination areas resulted from aggregation over large land areas. However, within these areas shippers in the central city might be biased towards utilizing rail because of the availability of rail sidings, while in the suburbs shippers might be biased towards truck. Thus, a large variation in the spatial distribution of production and consumption among the origins and destinations might also lead to lower explanatory power in the mode split equations. In equation (13) the sign of α' was expected to be negative--the greater the ratio of truck to rail rates the smaller would be the percent shipped by truck. The sign of β' was expected to positive--the greater the ratio of rail to truck transit times the greater would be the percent shipped by truck. The results of the estimations of equation (13) using both the logarithms of Pertruc and of the logistic transformation (Pertruc/100-Pertruc) as the dependent variables are presented in Table XVI. The results of both versions appeared to be about the same, although the explanatory power of the equations using the logistic transformation as the dependent variable were generally slightly less. This was because the variance of (Pertruc/100-Pertruc) was, of course, much greater than that of Pertruc.

The estimated coefficients always had the expected signs. Also, in every equation except STCC 283 (Drugs) at least one of the coefficients was significant (at the one percent level). In four of the estimated equations, STCC's 208 (Beverages), 264 (Converted Paper Products), 285 (Paints) and 307 (Misc. Plastics), the coefficients associated with both relative rates and relative transit times were significant (at the one percent level). The calculated values of rail

Table XVI

Abstract Mode Results

Dependent Variable--Pertruc

Commodity	Const.	α	β	R^2	F	N=	Corr$^{1}_{12}$
STCC 208 Beverages, etc.	8.619 (3.30)**	-3.301 (-4.72)**	1.816 (9.94)**	.39	105.05**	339	-0.46
STCC 264 Converted paper, etc.	5.930 (2.45)*	-3.699 (-3.25)**	1.530 (6.15)**	.28	30.88**	164	-0.24
STCC 265 Containers, etc.	9.602 (5.29)**	-0.722 (-0.67)	1.004 (3.79)**	.12	8.78**	134	0.56
STCC 283 Drugs, etc.	35.945 (2.51)*	-2.704 (-1.74)	0.734 (1.51)	.21	7.36**	59	-0.64
STCC 284 Soap, etc.	3.171 (3.26)**	-1.551 (-1.59)	1.806 (4.63)**	.16	24.91**	264	-0.57
STCC 285 Paint, etc.	54.762 (9.82)**	-2.725 (-3.48)**	0.839 (5.11)**	.15	15.34**	181	0.27
STCC 307 Misc. Plastics	10.034 (4.56)**	-2.960 (-2.75)**	0.820 (2.78)**	.20	15.46**	127	-0.51
STCC 346 Metal Stampings	24.119 (3.74)**	-4.319 (-2.89)**	0.362 (0.89)	.13	7.54**	108	-0.48
STCC 364 Electric Lighting	0.436 (0.82)	-2.130 (-0.78)	2.052 (3.26)**	.21	6.01**	49	-0.13

Dependent Variable--(Pertruc/(100-Pertruc))

STCC208 Beverages, etc.	0.075 (1.97)*	-4.363 (-3.10)**	3.680 (10.01)**	.35	88.36**	339	-0.46
STCC 264 Converted Paper, etc.	0.125 (1.42)	-5.755 (-2.52)**	2.850 (5.70)**	.23	24.30**	164	-0.24
STCC 265 Containers, etc.	0.906 (1.09)	-0.960 (-0.42)	2.034 (3.62)**	.12	8.51**	134	0.56
STCC 283 Drugs, etc.	4.179 (0.50)	-5.075 (-1.68)	1.760 (1.86)	.24	8.70**	59	-0.64
STCC 284 Soap, etc.	0.054 (4.03)**	-1.163 (-0.58)	3.030 (3.80)**	.16	24.66**	264	-0.57
STCC 285 Paint, etc.	20.005 (3.33)**	-5.389 (-3.12)**	1.701 (4.68)**	.13	12.74**	181	0.27
STCC 307 Misc. Plastics	0.405 (0.89)	-6.538 (-3.03)**	1.930 (3.26)**	.24	20.01**	127	-0.51
STCC 346 Metal Stampings	2.321 (0.47)	-7.730 (-2.47)**	0.761 (0.90)	.10	5.90**	108	-0.48
STCC 364 Electric Lighting	0.002 (3.17)**	-3.242 (-0.61)	4.108 (3.33)**	.21	6.09**	49	-0.13

[1] Simple correlation between $L_n\,(^{RT}/_{TT})$ and $L_n\,(^{TR}/_{RR})$

**Significant at the one per cent level.

*Significant at the five percent level.

Figures in parentheses are t-statistics.

rates and transit times, and truck rates and transit times were essentially linear (or nearly linear) functions respectively of rail short line miles and truck short line miles. Since rail and truck short miles were highly correlated (the simple correlation between them was .977), multicollinearity problems in the estimation of the mode split models were anticipated. Although in some cases the logarithm of relative rates and logarithm of relative transit times were fairly highly correlated (See Table XVI), collinearity problems appeared to be severe only in the estimation of the Drugs (STCC 283) equation. The explanatory power of each of the equations estimated was not high, e.g., the R^2's were much below 1.00. This might in part reflect the problems in aggregation discussed before. Also, the Census of Transportation was subject to a large sampling error since the Census was based on a 1 in 100 sampling ratio. This would lead to lower explanatory power in the mode split equations, but would not bias the estimated coefficient if the random variation in the dependent variable was normally distributed (33).

From these estimated equations it was possible to calculate the marginal rate of substitution between relative rates and relative transit time--that is how much would (TR/RR) have to change given a change in (TT/RT) to stay on the same level of Pertruc. This concept is illustrated graphically in Fig. 1. (Where K_1; K_2 and K_3 are constants and $K_3>K_2>K_1$). Notice our equations were estimated on relative transit time in terms of (RT/TT) while here we are discussing the marginal rate of substitution between (TR/RR) and (TT/RT). However, it can be easily shown that the coefficient associated with (TT/RT), had it been included instead of (RT/TT) in equation (13), would be $-\beta$. The marginal rate of substitution was obtained by taking the total derivation of equation (13) and solving for d (TR/RR) / d (TT/RT). Thus,

$$\frac{d(TR/RR)}{d(TT/RT)} = \frac{\alpha}{\beta}(TR/RR)(RT/TT) = MRS \qquad (14)$$

It can be seen from this equation and from Figure 1 that the marginal rate of substitution was not constant, but was a changing function of relative rates and relative transit time. Table XVII presents the calculated marginal rates of substitution in eight of the commodity groups (all except STCC 283, Drugs). These marginal rates of substitution were calculated at the mean values of (TR/RR) and (RT/TT) and are shown for the regressions using Pertruc, and those using the logistic transformation (Pertruc/100-Pertruc) as the dependent variable. From the table it can be seen that the marginal rate of substitutions calculated from the logistic model were consistently larger in absolute value than those calculated from the equations using Pertruc as the dependent variable. Also, for those equations in which relative transit

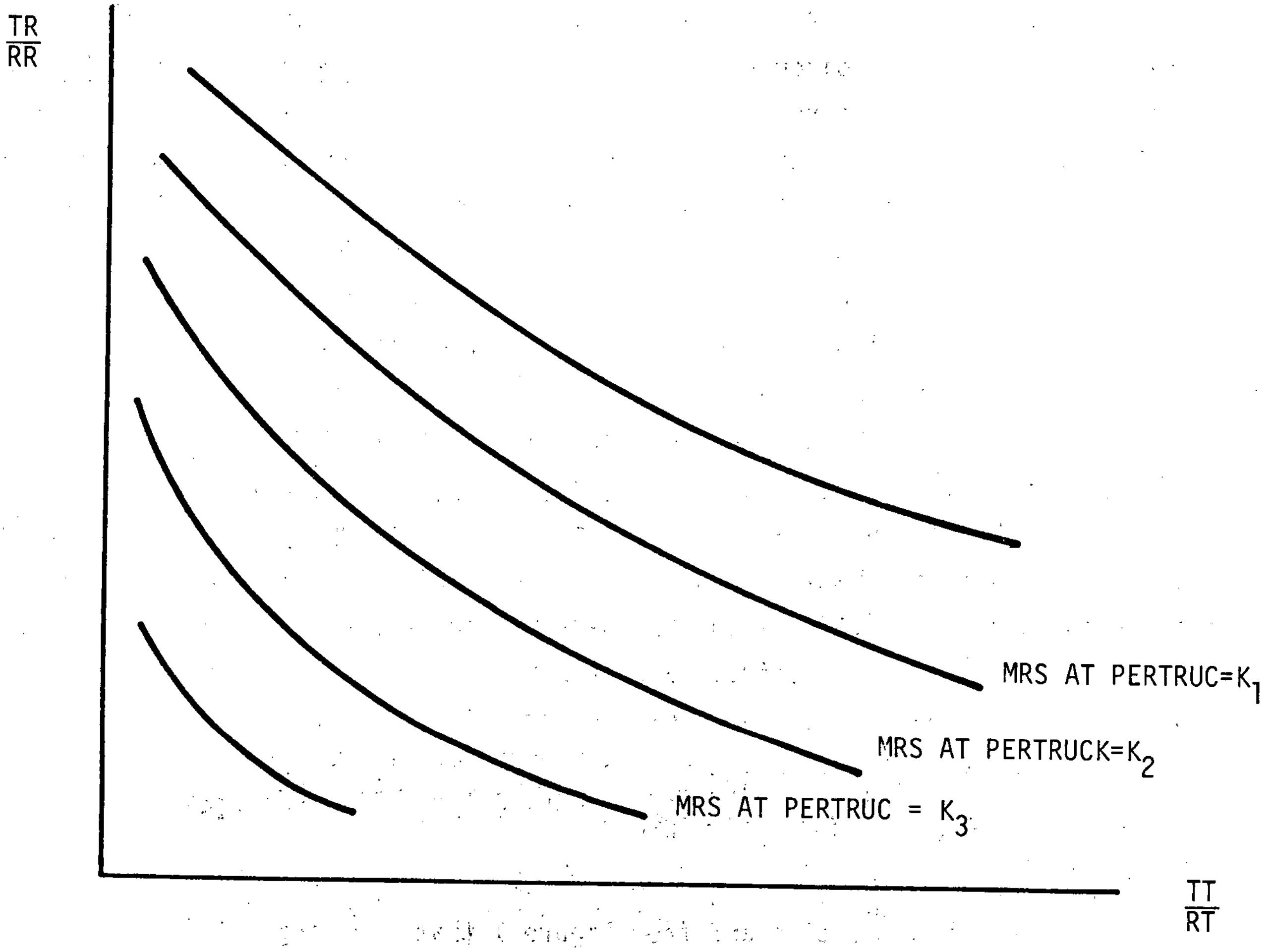


Figure 1

Marginal Rates of Substitution Between Relative Rates and Relative Transit Time

$(K_3>K_2>K_1)$

Table XVII
Marginal Rates of Substitution

Commodity Group	Dependent Variable Pertruc β/α	Dependent Variable Pertruc MRS	Dependent Variable Pertuc/(100-Pertruc) β/α	Dependent Variable Pertuc/(100-Pertruc) MRS
STCC208 Beverages, etc.	-0.550[1]	-6.187	-0.843[1]	-9.412
STCC264 Converted Paper	-0.414[1]	-4.135	-0.495[1]	-4.944
STCC 265 Containers, etc.	-1.391[2]	-14.788	-2.120[2]	-22.539
STCC 284 Soap, etc.	-1.164[2]	-10.297	-2.605[2]	-23.045
STCC 285 Paints, etc.	-0.308[1]	-3.251	-0.316[1]	-3.335
STCC 307 Misc. Plastics	-0.277[1]	-1.722	-0.295[1]	-1.834
STCC 346 Metal Stampings	-0.084[3]	-0.605	-0.098[3]	-0.706
STCC 364 Electric Lighting	-0.963[2]	-4.313	-1.267[2]	-5.676

[1] Both relative rates and relative transit time significant at the one per cent level.

[2] Relative transit time significant at the one per cent level.

[3] Relative rates significant at the one per cent level.

time was significant, the marginal rate of substitution between relative rates and relative time was very large (in absolute value). In other words, a very small change in relative transit time would require a very large change in relative rates to keep Pertruc at the same level. On the other hand, for that equation in which relative rates were significant (at the one percent level) and relative transit time was not, the marginal rate of substitution between relative rates and relative transit time was very small (in absolute value).

Empirical Results - Agricultural Commodities

Equation (13) was also estimated on data obtained from the 1974 Fresh Fruit and Vegetable Unload Statistics (43). Most of these fresh fruits and vegetables were extremely perishable and only three were deemed to be "pipelineable"--dry onions, potatoes and cabbage. These data show shipment of various agricultural commodities from states to 41 cities by truck and rail in rail carlot equivalents. Thus, in this case the dependent variable Pertruc was calculated by dividing rail carlot equivalents that went by truck by total carlots (by truck and rail). From the 1969 Census of Agriculture it was possible to identify the major counties in which these agricultural commodities were grown in the origin states. Highway short line miles were then measured from a city in or near these origin counties to the 41 destination cities. This information was obtained from the 1976 Rand McNally Road Atlas. Since highway and rail short line miles were assumed to be very similar, rail short line miles were not measured.

Rail rates were calculated from regressions estimated on data from the 1972 Carload Waybill Statistics allowing for circuity as previously described. These estimated regressions are (44):

Onions

$$RR = (0.383)\ (SS)^{-0.335}\ (miles)^{0.793**}\quad R^2 = 0.83\quad F = 138.58**\quad N = 60$$

Potatoes

$$RR = (0.331)\ (SS)^{0.017}\ (miles)^{0.759**}\quad R^2 = 0.85\quad F = 272.40**\quad N = 98$$

** Significant at the one percent level

Truck rates were calculated from the truck rate equations reported in Table XV again allowing for circuity. The calculated truck rates were then multiplied by 1.3

---

(43) U.S. Department of Agriculture, Agricultural Marketing Service, Fresh Fruit and Vegetable Unload Totals for 41 Cities, 1974.

(44) For reasons explained below, cabbage were dropped from the analysis. Thus a rate regression for cabbage is not given.

to bring them up to the 1972 level (45). One major problem with using these truck rate equations was that many agricultural commodities move by agricultural exempt motor carriers which tend to charge much lower rates than common carriers--the truck rate equations, however, were estimated on common carrier rates. Nevertheless, there seems to be little information on rates charged by agricultural exempt carriers. Thus, the ratio of truck rates to rail rates probably overstate the true rate ratio relationship. Also, no information on truck and rail rates was available for 1974. This ratio, however, probably had not changed much between 1972 and 1974. Rail and truck transit times were calculated (allowing for circuity) from equations (11) and (12) respectively.

Equation (13) was estimated for the three agricultural commodities using the logarithms of Pertruc and of the Logistic transformation (Pertruc/100-Pertruc) as the dependent variables. However, because of extreme multicollinerity between Ln(TR/RR) and Ln(RT/TT), the cabbage equations were not successful. The estimated equations for onions and potatoes were:

Onions

$$\text{Ln (Pertruc)} = 18.826 - 10.364\ \text{Ln(TR/RR)} + 3.013\ \text{Ln(RT/TT)}$$

$$R^2 = 0.42,\ F = 75.39,\ N = 213$$

$$\text{Ln (Pertruc/100-Pertruc)} = 36.622 - 24.867\ \text{Ln(TR/RR)} + 8.251\ \text{Ln(RT/TT)}$$

$$R^2 = 0.58,\ F = 146.90$$

Potatoes

$$\text{Ln (Pertruc)} = 10.306 - 9.173\ \text{Ln(TR/RR)} + 2.700\ \text{Ln(RT/TT)}$$

$$R^2 = 0.33,\ F = 54.00,\ N = 218$$

$$\text{Ln (Pertruc/100-Pertruc)} = 10.282 - 14.551\ \text{Ln(TR/RR)} + 5.133\ \text{Ln(RT/TT)}$$

$$R^2 = 0.30,\ F = 46.92$$

All coefficients and F statistics were significant at the one percent level.

Note that these estimated equations have much greater explanatory powers than the mode split regressions for manufactured commodities. This was probably partly due to the greater commodity disaggregation in the agricultural commodities.

As before it was possible to calculate from equation (14) the marginal rates of substitution between relative rates (TR/RR) and relative transit times (TT/RT). These calculated marginal rates of substitution are:

Onions

Dependent variable is Pertruc	MRS =-6.689
Dependent variable is (Pertruc/100-Pertruc)	MRS =-7.632

---

(45) From "A Brief Summary of MAC General Territory-Wide Changes for Class and Commodity Rates", Middle Atlantic Conference, Maryland.

Potatoes

Dependent variable Pertruc MRS = -3.099

Dependent variable is (Pertruc/100-Pertruc) MRS = -3.721

Again note that the marginal rates of substitution obtained from the models using the logistic transformation (Pertruc/100-Pertruc) as the dependent variable were slightly larger (in absolute value) than those using Pertruc. The marginal rates of substitution obtained from the onions and potatoes models indicated that relative transit time was a very important variable.

Conclusions

There appeared to be several problems in attempting to estimate abstract mode models for freight transportation from aggregate data. First, most aggregate market share data (for example the Census of Transportation) included under one commodity group several different types of commodities. Since modal decisions might be made very differently for various commodities, aggregate modal split models may not be expected to yield high R^2's and significant coefficient estimates. Also, aggregate data tended to utilize large land areas as origins and destinations. Partly for these reasons our mode split equations (especially for the manufactured commodities) did not have very high explanatory powers. Second, it was necessary to estimate values for truck rates, rail rates, truck transit times, rail transit times, etc., from several sources of secondary data. These estimates represented (at best) only broad averages for a given origin-destination pair. Because of peculiarities of each origin-destination pair, these estimates (at worst) may not have reflected the actual values at all. Finally, even though the abstract modal was based on individual (shipper) behavior, it was estimated on extremely aggregate data. These models were not truly behavioral. Thus, it might not be justified to use these models for predictive purposes. However, these disadvantages of the aggregate approach must be weighed against the major problem of the micro approach, generalizing from the micro to the aggregate level, which will be discussed in the next section.

The Micro Demand Modal Split Model

The demand analysis took still another approach at demand/modal split. Much analysis has been done of late on behavioral modelling of transportation choice. The bulk of this research has been developed in the context of urban passenger demand/modal split (see Watson (46), Charles River Associates (47), Domencich and McFadden (48), Lisco (49), Warner (50), Ben Akiva (51), to name but a few of a vastly growing literature). Only a few examples existed on the freight side (Kullman, (34), Hartwig and Linton (9), Miklius and Casavant(52), Watson (53), Reebie (8). Some research on behavioral modelling of mode choice is in progress concurrently with this project, i.e., at MIT (54) by Paul Roberts and Alan Stenger at Pennsylvania State University. Despite the attempts, the freight side of demand analysis was virtually unexplored.

The objective of this section was to develop a disaggregate model of shipper modal split. This model was based on observations of actual modal choices for individual shipments under the transport characteristics, i.e., rate, time, reliability, etc., for both the chosen and alternative modes from a specific origin to a specific destination. The commodity ambiguity and the geographical ambiguity, the effect of shipment size, the use of average times and rates all disappeared in this analysis. The need for all aggregates was eliminated. All averages were replaced by the actual values of service attributes which confronted the decision maker when the decision was made.

---

(46) Watson, P., The Value of Time: Behavioral Models of Modal Choice, D.C. Heath and Co., Lexington, Mass., 1974.

(47) Charles River Associates, Competition Between Rail and Truck in Intercity Freight Transportation, NTIS, Springfield, VA, 1969.

(48) Domencich, T. and D. McFadden, Urban Travel Demand: A Behavioral Analysis, American Elsevier, N.Y., 1975.

(49) Lisco, T., "The Value of Commuters' Travel Time--A Study in Urban Transportation", unpublished Ph.D. dissertation, Department of Economics, University of Chicago, 1967.

(50) Warner, S., Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice, Evanston, IL, Northwestern University Press, 1962.

(51) Ben Akiva, M., "Structure of Passenger Demand Models", Highway Research Record, #526, 1974.

(52) Miklius, W. and K. Casavant, "Estimation of Demand for Transportation of Agricultural Commodities", Mimeo, Department of Agricultural Economics, Washington State University, Pullman, Circa 1974.

(53) Watson, P., Urban Goods Movement: A Disaggregate Approach, D.C. Heath Co., Lexington, Mass., 1975.

(54) MIT, Freight Demand Modelling: A Policy Sensitive Approach, Center for Transportation Studies, CTS Report #75-6, April, 1975.

By observing modal choice when the shipper was confronted by various "bundles" of modal service characteristics <u>vis-à-vis</u> the characteristics of the goods he was shipping (and his own objectives and characteristics of the origin and destination, etc.), shipper trade-offs between the various independent (explanatory) variables could be determined. If the model explained the actual choice well, the addition of the pipeline characteristics might enable the determination to be made whether a shipment in question was likely to move by pipeline (the likelihood will be expressed in probability terms). Performing the analysis for all shipments would determine the likely percentage of that shipper's shipments which were likely (economic) candidates for freight pipeline.

This method of determining the modal split of a new mode can be contrasted with the Reebie study (<u>8</u>) approach where a group of shippers were confronted with the following type of question: Given a matrix with column headings of specific transportation service level changes and with row headings of specific transportation rate level changes, e.g., suppose entry i. j read a transport rate increase of 5% accompanied by a service increase of 10%, if event i, j occurs, would you change your shipment? Using shipper responses the matrix cells were then filled out with percentage diversion under the circumstances indicated by the row and column headings.

Two difficulties existed with such a format. The first was that since the shippers when interviewed were not making any real decisions, one had no idea whether they would really shift as they said they would. Secondly, the changes were stated in percentage terms and immediately brought to mind - percentage of what level? Constant percentage changes usually imply different behavioral responses depending on the base from which one starts. This is especially true if one views the cumulative normal "S" shaped modal choice curve which predominates in the urban passenger modal split literature. It is hard to imagine that the i, j situations postulated to the shippers in the Reebie study (<u>8</u>) analysis meant the same thing to each one.

To circumvent these deficiencies, the approach taken herein inferred the shipper's behavior from how he actually behaved (in a real life i, j situation) as opposed to how he said he would behave. Secondly, although the shipping behavior of an individual shipper was imposed on other shippers in the same commodity group (time and cost precluded such a modelling effort for all shippers), the aggregation was not carried across all shippers of all goods as it was in the Reebie study(<u>8</u>). It was felt, therefore, that since an individual firm's behavior was being observed from their actions in the marketplace, that an understanding of the shipper's response to a new batch of characteristics, i.e., pipeline, could be discerned from his response to existing characteristic batches. This was the abstract mode approach which suggested that choices were made based upon modal performance and not by what a mode is called.

In order to implement this part of the research, it was decided to survey major industries which shipped pipelineable products for which sufficient macro data was available. The goal was a sample of 500 individual shipment per firm with a sample mode split indicative of the firm's overall mode split. Information was collected on origin, destination, type of commodity, size of shipment, value of commodity, density of commodity, freight rate of chosen mode (truck or rail), freight rate of rejected mode (rail or truck), time in transit by chosen and rejected mode, and reliability by the chosen and rejected mode. While loss and damage rates were felt to be important a priori, no shipper was found that could give anything else but an overall average for rail and truck. Since each observation of modal choice would always have the same constant value for rail loss and damage, e.g., x, and the same constant value for truck loss and damage, e.g., y, the loss and damage figures had no discriminatory power (in addition to being collinear). Thus loss and damage was not present in the analysis herein.

Below, modal split models will be presented for three firms. The firms must remain nameless because the data involved in the analysis was regarded as proprietary. Suffice it to say that these were large, national firms. They had sophisticated traffic and distribution departments. A conscious effort was made to solicit information from more sophisticated firms with the general notion that they were leaders in their field and that their techniques of today would be the techniques of the "followers" tomorrow. In addition, because of the concentration of economic power, a relatively small number of big shippers tender a very large percent of the nation's total shipments. Thus the firms included in this study (and firms of the type that were included in this study) shipped a non-trivial portion of the nation's freight.

Contact was made with six firms at the level of Vice President Transportation or Traffic Manager. It was an original intention to gather information from a firm in every two digit STCC level which was considered pipelineable and which generated significant traffic flow. In some STCC's it was hoped that at least two firms could be surveyed to check the hypothesis implicit in the Reebie criticism above that within STCC behavior was more stable than between STCC behavior.

However, such a plan proved to be overly ambitious. In all cases the information which was sought was regarded as propriety. In all cases a meeting was required with a Vice President of Transportation or Traffic Manager so that the prospective contributor could assess our intentions and that assurances with regard to the security of the information could be made. In essence, the "product" which the research team was "selling" had to be "sold to management". This was a time consuming process, as was the data collection which followed (and the tabulation, reconfirmation, education with respect to the data, etc.). Thus it became clear that the objective of 20 or so micro demand models could not be accomplished.

Of the six firms contacted, data was collected from four. A fifth firm was willing to cooperate but couldn't allow the research team to process the data and could not, at that time, provide manpower to process the data for the team. Only one firm was unable to participate due to management policy.

Not only did each data set take a long time to generate on the part of the research team, but, in addition, each set entailed many man-hours of time from the cooperating firms. The process of identifying the actual values on shipments which did not occur (the non-chosen mode) was non-trivial. Much time was spent with high ranking people in the traffic departments learning their coding systems, discussing their evaluations of the modes, learning about company policy with respect to transportation, and other matters which influenced how the modal split model for that firm was constructed and how the results were interpreted. Much time was also spent with other employees of the cooperating firms in "straightening out" data problems. All in all, a great debt of gratitude must be expressed to the participating firms.

While the data gathering and processing was time consuming, it is important to point out that the data was available to perform the desired research. The firms were cooperative and receptive. The almost uniform willingness to cooperate and spend considerable amounts of time with researchers was most gratifying.

Part of the learning process which was obtained by talking to the shippers was the idiosyncratic nature of the firms talked to. What is important to firm X might not be important to firm Y. In addition, many items which influence modal choice were highly idiosyncratic and difficult, if not impossible to quantify, e.g., advertising, public relations, salesmanship of a mode; equipment availability, special loading or unloading requirements at the origin and/or destination. To the extent possible, any modal choices which might not truly involve choice, e.g., special loading facilities at a factory might preclude a particular mode, and such an item if not known to the research team might result in the team's choice of another mode in that situation, were eliminated from the data base which was gathered (excepting in the case of Firm C). Thus there were no physical reasons why either mode could not be chosen in the data set. In addition, emergency rush orders, untypical seasonal demands, and any behavior regarded as atypical by the shipper was not included in the data base.

Nevertheless, the idiosyncrasies prevailed. Although the characteristics faced by each firm were the same, the reactions to these characteristics would be different and other idiosyncrasies exist as well. If each shipper had these personal differences, why proceed with the modelling effort or why not build the idiosyncrasies into the individual model?

To the extent that the individual model built in shipper X's idiosyncrasies, the better the model would be expected to classify shipper X's modal choices, i.e., predicting an actual rail shipment as rail and an actual truck shipment as truck. However, the model would then lose generality. It was the view of the research team that all the idiosyncrasies that existed could never be captured in the model. Certain items, e.g., rates, transit time, etc., are faced by every shipper, however. If a shipper's modal behavior could be replicated reasonably well based on a model built from publicly observable data in a model form which was a priori acceptable, an important step would have been made.

Since the goal was to take the individual micro models and split the macro data and to apply the micro model to other secondary source data, in order to meet this objective, the model had to be simple and contain variables for the rest of the universe (the non-surveyed firms) which were easily accessible (which idiosyncrasies were not). Thus the test of the models developed herein will be their workability and not their completeness, i.e., did the models replicate behavior even though not every possible explanatory variable was included in the model.

The theory of demand and the results of three tests of the micro model are given below.

Disaggregate Models of Modal Choice

The Demand for Transportation

The demand for transportation is usually viewed as being derived from the demand and supply conditions in various regions. This could be easily illustrated through a simple two region trade model. In Figure 2, region A and region B were initially each in isolation -- thus, the equilibrium price and

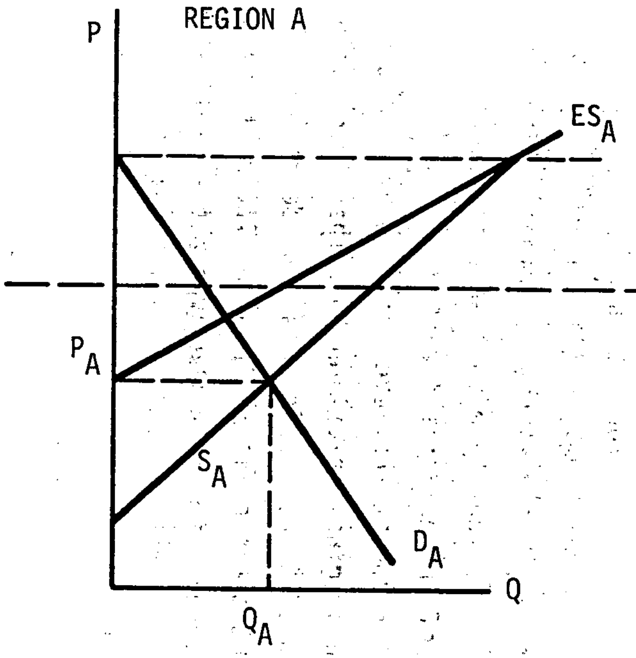


Figure 2A

Supply and Demand Equilibrium in Region A and the Derivation of Region A's Excess Supply Curve

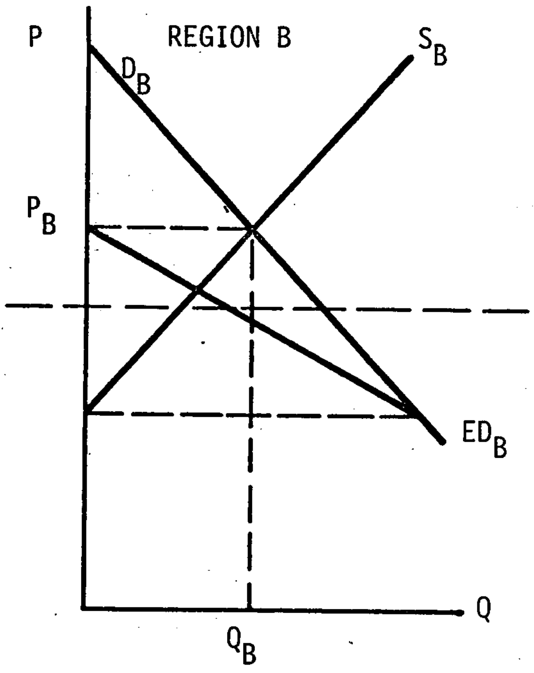


Figure 2B

Supply and Demand Equilibrium in Region B and the Derivation of Region B's Excess Demand Curve

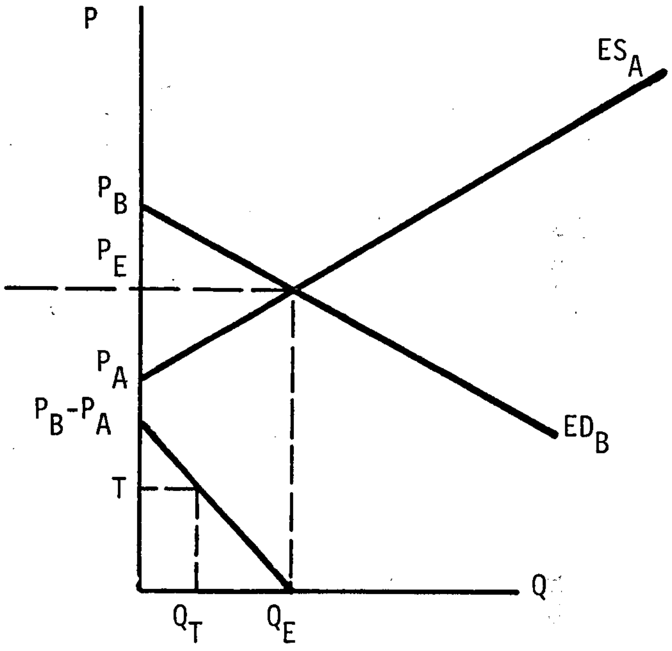


Figure 2C

The Derived Demand for Transportation Curve Between Region A and Region B derived From Region A's Excess Demand Curve from his response to existing characteristic batches. This was the abstract mode approach which suggested that choices were made based upon modal performance and not by what a mode is called.

In order to implement this part of the research, it was decided to survey major industries which shipped pipelineable products for which sufficient macro data was available. The goal was a sample of 500 individual shipments per firm with a sample mode split indicative of the firm's overall mode split. Information was collected on origin, destination, type of commodity, size of shipment, value of commodity, density of commodity, freight rate of chosen mode (truck or rail), freight rate of rejected mode (rail or truck), time in transit by chosen and rejected mode, and reliability by the chosen and rejected mode. While loss and damage rates were felt to be important a priori, no shipper was found that could give anything else but an overall average for rail and truck. Since each observation of modal choice would always have the same constant value for rail loss and damage, e.g., x, and the same constant value for truck loss and damage, e.g., y, the loss and damage figures had no discriminatory power (in addition to being collinear). Thus loss and damage was not present in the analysis herein.

Below, modal split models will be presented for three firms. The firms must remain nameless because the data involved in the analysis was regarded as proprietary. Suffice it to say that these were large, national firms. They had sophisticated traffic and distribution departments. A conscious effort was made to solicit information from more sophisticated firms with the general notion that they were leaders in their field and that their techniques of today would be the techniques of the "followers" tomorrow. In addition, because of the concentration of economic power, a relatively small number of big shippers tender a very large percent of the nation's total shipments. Thus the firms included in this study (and firms of the type that were included in this study) shipped a non-trivial portion of the nation's freight.

Contact was made with six firms at the level of Vice President Transportation or Traffic Manager. It was an original intention to gather information from a firm in every two digit STCC level which was considered pipelineable and which generated significant traffic flow. In some STCC's it was hoped that at least two firms could be surveyed to check the hypothesis implicit in the Reebie criticism above that within STCC behavior was more stable than between STCC behavior.

However, such a plan proved to be overly ambitious. In all cases the information which was sought was regarded as proprietary. In all cases a meeting output in each region was determined where demand equals supply (P_a and Q_a in region A and region B).

Obviously, if any trade was to be generated, it would be from region A to region B.(from the lower price region to the higher price one). Region A's excess supply curve (ES_A) was obtained through the horizontal subtraction of demand from supply. Similarly, region B's excess demand curve (ED_B) was obtained through the horizontal subtraction of supply from demand. These two curves have been transcribed to Figure 2c. If transportation from region A to region B was perfectly costless, equilibrium trade and price in both regions would be determined by the intersection of the excess demand and excess supply curves (Q_e and P_e). Note the equilibrium price is below the price in region B, but above that in region A. However, in reality transportation is never free. Thus, it was necessary to derive the demand for transportation -- this was simply the vertical subtraction of excess supply from excess demand (curve D_T is obtained). If the transportation rate was T the equilibrium trade would obviously be Q_T.

Figure 3 is a blow-up of the demand for transportation curve which has been derived.

Equilibrium trade was not only determined by the transportation rate, but also by other costs of shipping commodities. These associated costs might include time costs, costs because of uncertain transit time, loss and damage costs, etc. If these associated costs (which were drawn horizontally in Figures 3 and 4 for convenience's sake) could be measured in dollars, they would be added to the transportation rate (A + T) to determine the equilibrium quantity of trade (Q_T' in Figure 3).

In Figure 4, the transportation rate was no longer fixed. Rather, there existed two modes of transport each of which were willing to carry different quantities of the commodity at various rates (S_1 and S_2). Each mode also had an associated cost curve (A_1 and A_2). To obtain the total cost of transportation for each mode (TC_1 + TC_2) the supply curve and associated cost curve were vertically summed. The total cost of transportation (TC) was obtained through the horizontal summation of TC_1 and TC_2. The equilibrium level of trade was determined where the demand for transportation (D_T) equaled the "supply" (TC). At equilibrium,mode 1 carried Q_1 at the rate R_1, mode 2 carried Q_2 at a different rate R_2(55). From this simple analysis, it was possible to draw two important

---

(55) Stucker, J., An Econometric Model of the Demand for Transportation, Ph.D. dissertation, Northwestern University, 1969, also see Samuelson, P., Spatial Price Equilibrium and Linear Programming", American Economic Review, June 1952.

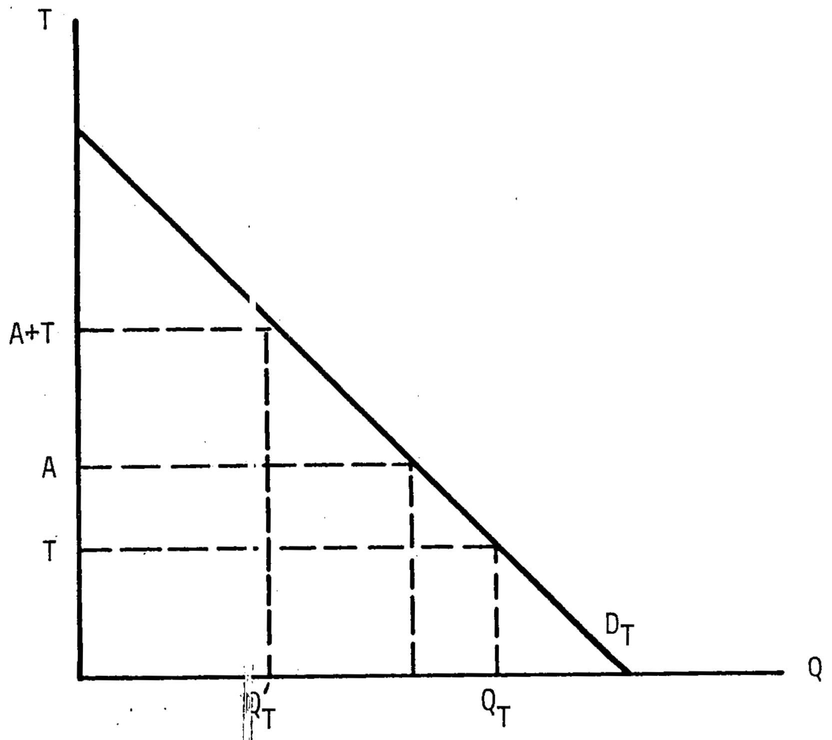


Figure 3. The Demand for Transportation Function and Equilibrium when Service Elements (A) Are Introduced in Addition to Transport Rate (T).

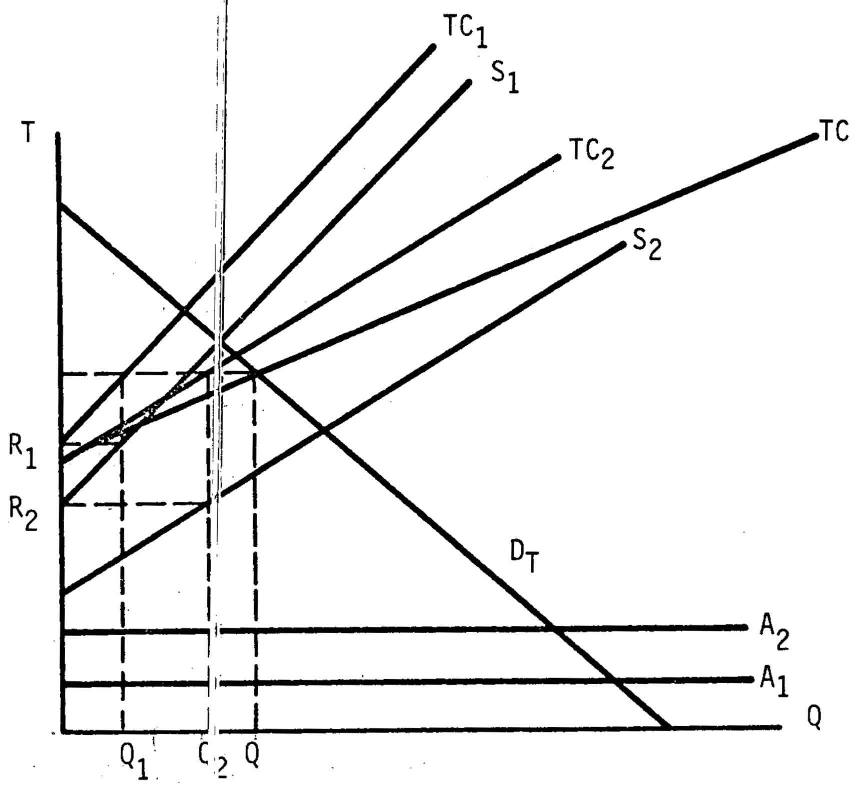


Figure 4. Transportation Equilibrium and Modal Split when Modal Service Elements (A_1 and A_2) Other than Rate Are Introduced in a Case of an Increasing Transportation Cost Function

conclusions. First, the equilibrium amount of trade and the modal split depended on the supply curve and the level of service (transit time, transit time reliability, etc.) provided by each mode. For example, if mode 1 improved its level of service (and thus reduced its associated costs), not only would the quantity carried by mode 1 be increased and that by mode 2 reduced, but also total amount of trade would be increased (a new equilibrium would result). Second, if we knew the the costs and levels of service provided by various modes (shifts in the supply curve), it was possible to identify the intermodal demand for transportation. During the previous discussion, the following demand hierarchy (by a shipper) had been implicitly assumed:

1. Choice of location and the level of activity,
2. Choice of market locations and volumes of trade, and
3. Choice of transport mode. (56)

These choices ranged from the very long run (1) to the short run (3). During this paper, we shall concentrate on the short run, i.e., decision three. However, it must be remembered that over the long run the costs and levels of service provided by various modes will not only affect the modal split, but will also affect the directions and total volume of trade flows and ultimately the location and level of economic activity.

A Model of Modal Choice

In this section we shall give an intuitive motivation for the logit model of mode choice. For a more rigorous development of this model see McFadden(57). Suppose a shipper had a choice of utilizing two alternative modes (truck or rail) each with a vector of "attributes" (shipment cost, transit time, reliability, etc.). Assume the probability the shipper chose truck for a particular shipment can be expressed as a linear function of the differences between truck and rail shipment costs, transit times, and transit time reliabilities. Thus:

$$P_T = a(TC - RC) + b(TT - RT) + c(TR - RR) \qquad (15)$$

---

(56) Terziev, Marc, Ben-Akiva, Moshe, and Roberts, Paul, "Freight Demand Modelling--A Policy Sensitive Approach", presented to the 47th National Operations Research Society of America Meeting, 1975.

(57) McFadden, Daniel, "Conditional Logit Analysis and Qualitative Choice Behavior", in P. Zarembka, ed. Frontiers in Econometrics, Academic Press, N.Y.C. 1973, and by the same author, A Disaggregated Behavioral Model of Urban Travel Demand, Charles River Associates, prepared for the Federal Highway Administration, U.S. Department of Transportation, 1973.

where (TC - RC) was the truck shipment cost minus the rail shipment cost, (TT - RT) was the truck transit time minus the rail transit time, (TR - RR) was the truck reliability minus the rail reliability, P_T was the probability of choosing truck and a, b, and c were the parameters of the equation.

A major problem with this formulation was that estimated values of P_T could fall outside the range of 0 to 1. To constrain this equation so that the probability of choosing truck fell between 0 and 1, the logistic transformation of P_T is taken, thus obtaining:

$$\mathrm{Ln}\ (P_T/(1 - P_T)) = a(TC - RC) + b(TT - RT) + c(TR - RR) \qquad (16)$$

where $(1 - P_T)$ was the probability of choosing rail (P_R).
From equation (16) the probability of choosing truck reduced to:

$$P_T = \frac{\exp\ (a(TC - RC) + b(TT - RT) + c(TR - RR))}{1 + \exp\ (a(TC - RC) + b(TT - RT) + c(TR - RR))} \qquad (17)$$

Multiplying the top and bottom of equation (17) by exp(axRC + bxRT + cxRR) obtained:

$$P_T = \frac{\exp\ (axTC + bxTT + cxTR)}{\exp\ (axRC + bxRT + cxRR) + \exp\ (axTC + bxTT + cxTR)} \qquad (18)$$

This was the logit model -- since its form was intrinsically non-linear, the parameters (a, b, and c) could be estimated using the method of maximum likelihood.

The Data and the Variables

The data used for the empirical analysis consisted of observations of individual shipments by truck and rail obtained from three(national) large shippers. These firms were very cooperative in allowing us to use this information. In return we agreed not to reveal their names. Thus, for the purposes of this report, we shall refer to them as Firms A, B, and C. For all three shippers the shipment observations were over a number of origin-destination pairs. When gathering the data from waybills, we attempted to select origin-destination pairs where the shippers were engaging in active mode splitting (between truck and rail). For each observation we collected (or calculated) the truck and rail shipment costs (in dollars per shipment), and the truck and rail transit times (in days). Also, from Firm A we were able to obtain the truck and rail transit time reliabilities(in the percentage of shipments arriving during the best three day period). Unfortunately, this information was not available for Firms B and C. In a previous analysis of this type, Hartwig and Linton (9) were able to collect information on the actual transit time for each shipment and then construct an index of reliability (standard deviation of the transit time) for a mode over a given origin-destination pair. However, information on actual transit times although available for most shipments by Firm C was not available for the other two firms. Therefore, the transit time figures used for Firms A and B, and the reliability figures used for Firm A were the expected values for a given origin-destination pair obtained through discussions with the shippers. This method seemed to be justified. If a shipper was reasonably informed about a mode's level of service (which ours seemed to be), his perceptions of transit time and transit time reliability over a given origin-destination pair would reflect the actual population values. Also, not enough information on transit times was available from Firm C to construct an index of modal reliability for each origin-destination pair.

For a shipment that was carried by truck it was necessary to calculate what the shipment cost would have been if it was carried by rail (and vice versa). Since these calculations could introduce a bias in the analysis, the method utilized will be detailed here. It was assumed that if a truck shipment went by rail, it would be combined with other truck shipments to form one rail shipment. The size of such a rail shipment was obtained through discussions with the shipper. For example, for Firm A this figure was 30,000 pounds. A 30,000 pound rail shipment almost always "went as 36,000 pounds" (the rate was charged as if the shipment was 36,000 lbs.). Thus, the calculated rail cost for Firm A was:

$$RC = SS \times \frac{36{,}000}{30{,}000} \times \text{ rail rate (on a 36,000 lbs. shipment)} \qquad (19)$$

where SS was the truck shipment size in pounds. For a shipment that was carried by rail, it was assumed that it would be broken down into a number of truck shipments. The number of truckload shipments required was obtained through discussions with the shipper. Since truck shipments were almost always carried at a truckload rate, the usual calculated truck shipment cost was:

$$TC = \text{Truckload Rate} \times \text{\# of Trucks required} \qquad (20)$$

Now we shall turn to the empirical results for Firms A, B, and C.

Empirical Results -- Firm A

The maximum likelihood estimate of equation (16) for Firm A was:

$$
\begin{array}{llll}
\mathrm{Ln}(P_T/1 - P_T) = & -0.0199\ (TC - RC) & - 1.0585\ (TT - RT) & \\
t\text{ - stat.} & (-9.94) & (-6.41) & \\
 & + 0.0011\ (TR - RR) & \lambda = 237.7 & \\
 & (0.09) & \text{Pseudo } R^2 = 0.49 &
\end{array}
$$

The likelihood ratio (λ) was used as a test of the significance of the estimated equation and was calculated as:

$$\lambda = -2\,(\log L\,(\theta = 0) - \log L\,(\theta = \hat{\theta})) \qquad (21)$$

where L $(\theta = 0)$ was the likelihood of the function with all coefficents set equal to zero and L $(\theta = \hat{\theta})$ was the likelihood of the functions at the estimated coefficients. The likelihood ratio was assymptotically chi-squared distributed with as many degress of freedom as the number of parameters estimated in the equation. Using this test, the estimated equation was highly significant (at the one percent level). The pseudo R^2 gave an intuitive feeling for the fit of the estimated equation to the data. It was calculated as:

$$\text{Pseudo } R^2 = 1 - \frac{\log L(\theta = \hat{\theta})}{\log L(\theta = 0)} \qquad (22)$$

Another test of the estimated equation was its success in classifying observations into the right group (truck or rail). With a cut-off probability of 0.5, this model misclassified only 22 observations (one truck and twenty-one rail) out of 350 observations. The performance of an estimated equation could also be analyzed by comparing the totals of the estimated probabilities for each group to the actual values in each group. In this case, the estimated total probabilities were exactly the same as the actual values (221 truck and 129 rail observations). From the pseudo R^2, its success in classifying observations, and the estimated total probabilities, it appeared the estimated equation fitted the data very well.

The signs of all the coefficients in the estimated equation were in the expected direction. The coefficient associated with cost (TC - RC) was expected to be negative -- larger the truck cost (or smaller the rail cost), the smaller was the probability the shipment would go by truck. The coefficient associated with transit time was also expected to be negative -- the larger the truck transit time (or smaller the rail transit time), the smaller was the probability the shipment would go by truck. Finally, the coefficient associated with transit time reliability was expected to be positive -- the larger the truck reliability (or smaller the rail reliability), the larger was the probability the shipment would go by truck. In the case of the logit model, the t-test cannot be used directly to test the significance of the estimated coefficients (the sampling distribution had a variance greater than the t). Therefore, since the tail of this distribution was "fatter" than the t distribution, the t-test could be used to test if a coefficient was not significant, but not if it was significant. Thus, the best method to test the significance of adding explanatory variables to an equation was to look at the increase in the likelihood ratio as these variables were added. Using this method, both cost (TC - RC) and transit time (TT - RT) were significant in the estimated equation (at the one percent level). However, the addition of transit time reliability (TR - RR) did not significantly increase the likelihood ratio of the estimated equation.

An estimate of the value of time by the shipper could be obtained by dividing the time coefficient by the cost coefficient. The time coefficient (b) equaled $\partial \ln (P_T/1 - P_T))/\partial$ (TT - RT) and the cost coefficient (a) equaled $\partial \ln(P_T/(1-P_T))/\partial$ (TC - RC). Therefore, b/a equaled ∂(TC - RC)/ ∂(TT - RT) which was the value of time. From the estimated equation this figure was approximately $53 per day for the average shipment.

Table XVIII presents the elasticities of the probability of choosing truck or rail with respect to the various independent variables $((\partial P_T/\bar{P}_T)/(\partial X/\bar{X})$ and $(\partial P_R/\bar{P}_R)/(\partial X/\bar{X})$, where $\bar{P}_T$, $\bar{P}_R$ and $\bar{X}$ were the values of these variables at their means). From these figures it was possible to determine the relative importance of the independent variables to the shipper in making his modal choice. The most important variables (with elasticities greater than one in absolute value) were truck cost, rail cost, and rail transit time. Truck cost was extremely important in the shipper's modal decision -- a one percent increase in the truck cost would increase the probability of choosing rail by 5.14 percent and decrease the probability of choosing truck by 1.64 percent. Both rail cost and relatively long rail transit time worked against the probability of choosing rail. Relatively long rail transit time was also an important variable working in favor of choosing truck. Truck transit time, truck transit time reliability, and rail transit time reliability did not seem to be very important variables in the shipper's modal decisions.

Empirical Results -- Firm B

Table XIX presents the results of five estimated equations for Firm B. In general the results for this firm were far less successful than those for Firm A.

Table XVIII. Elasticities, Firm A

Variable	Mode Truck	 Rail
Cost		
TC	-1.64	5.14
RC	0.88	-2.76
Transit Time		
TT	-0.18	0.56
RT	1.21	-3.78
Transit Time Reliability		
TR	0.03	-0.08
RR	-0.02	0.05

Table XIX. Estimated Equations, Firm B

Variable		1 Both Origins	2 Both Origins	3 Origin 1	4 Origin 2	5 Origin 2
Cost	TC - RC	-0.0160 (-4.70)[1]	-0.0240 (-7.64)	-0.0368 (-5.06)	-0.0071 (-2.33)	-0.0196 (-5.99)
Transit Time	TT - RT	0.1927 (3.30)		-0.8840 (-6.71)	0.2601 (4.33)	
Origin City Dummy	OC1	3.4950 (10.28)	3.1912 (9.24)			
Likelihood Ratio (λ)		240.1	228.6	119.5	110.0	91.03
Pseudo R^2		.509	.485	.556	.429	.355
Misclassified		39	36	14	27	25
Truck		23	21	2	27	24
Rail		16	15	12	0	1
Total Estimated Probability						
Truck		157.2	169.8	127.8	27.7	42.7
Rail		182.8	170.2	27.2	157.3	142.3

1) t - statistics are in parentheses

Shipments made by Firm B were from two origins -- one was the Middle Atlantic region (origin 1) and one was the Midwest (origin 2). Equations 1 and 2 (see columns in Table XIX) included as variables: cost (TC - RC), transit time (TT - RT) and an origin city dummy (0 if the shipment was made from the Midwest and 1 if it was made from the Middle Atlantic origin). In both these equations the coefficient associated with cost had the expected sign and was significant (at the one percent level). Unfortunately, in equation 1 the coefficient associated with transit time was also significant (at the one percent level), but its sign was in the wrong direction. In equation 2 the transit time variable has been deleted. According to the likelihood ratio test, both equations 1 and 2 were significant (at the one percent level). Also, the explanatory power of equation 2 was only slightly less than equation 1. Therefore, although the addition of transit time to the equation significantly increases the likelihood ratio, it appeared this variable was not very important to the shipper in making his modal choice (more about this when the computed elasticities from the equations are discussed). In both equations 1 and 2 the coefficient associated with the origin city dummy was positive and significant (at the one percent level). This implied shipments from the Middle Atlantic origin have a much higher probability of going by truck. Equation 1 misclassified 39 observations (23 truck and 16 rail) out of a total of 340 (159 truck and 181 rail). Also, the total estimated probabilities were very close to the actual values. Equation 2 misclassified 36 observations (21 truck and 15 rail). However, the total estimated probabilities from this equation were not very close to the actual values.

It appeared from the inclusion of origin city dummy variable that shipment decisions for the two origins might be made differently. Therefore, the logit model was estimated for each origin. The problem with this approach was that there was very little mode splitting from each origin -- from origin 1 only 23 out of a total of 155 shipments were made by rail, while from origin 2 only 27 out of a total of 185 shipments were made by truck. Despite this the results for origin 1 (equation 3) appeared to be somewhat successful. The coefficients associated with both cost and transit time had the expected signs and were significant (at the one percent level). The likelihood ratio was significant (at the one percent level) and the pseudo R^2 was fairly high. The equation misclassified 14 observations (2 truck and 12 rail) -- approximately one-half of the rail shipments were misclassified. The total estimated probabilities for this equation were fairly close to the actual values. The results for origin two (equation 4 and 5) were not so successful. Although the coefficient associated with cost had the expected sign and was significant (at the one percent level) in both equations, the coefficient associated with transit time in equation 4 had the wrong sign and was significant (at the one percent level). In equation 5 the transit time variable was deleted and the explanatory power of this equation was not much worse than equation 4. According to the likelihood ratio test, both equations were significant (at the one percent level). However, equation 4 misclassified all of the truck shipments, while equation 5 misclassified 24 of them. The total estimated probabilities for equation 4 were fairly close to the actual values (27 truck and 158 rail), while for equation 5 the total estimated probabilities were not very close to the actual values.

Table XX presents the elasticities computed at the mean of the probability of choosing truck or rail with respect to the various explanatory variables for equations 1 (both origins), 3 (origin 1), and 4(origin 2). From equation 1 it appeared both truck cost and rail cost were very important in the shipper's modal decisions - the elasticities associated with these variables were very large in absolute value. On the other hand, truck transit time and rail transit time did not seem to be very important. Although the signs of these elasticities were not as expected, their magnitudes were very small (in absolute value). From equation 3 it again appeared that truck cost and rail cost were very important in the shipper's modal choice. Also, relatively long rail•transit time seemed to be an important variable working against the possibility of choosing rail. Truck cost and rail cost were also important variables in equation 4. Also, although the truck transit time and rail transit time elasticities did not have the expected signs, the size of these elasticities were much smaller than those associated with truck and rail cost (in absolute value). Thus, it appeared that overall truck and rail transit time were not very important in this shipper's modal decisions. This conclusion was verified through discussions with the shipper. Firm B had a practice of calculating truck and rail cost indices for broad geographic regions, and then shipping by the cheapest mode to each region.

Empirical Results - Firm C

Table XXI presents the estimated equations for Firm C. Here, certain destination cities specifically requested that shipments to them be made by rail boxcar. This was accounted for by including a destination city dummy variable in the estimated equations, and assigning a value of one to those destination cities that requested rail deliveries. It can be seen that in both equations 1 and 2 (of Table XXI) the probability of shipping by truck was significantly lower (at the one percent level) to these destination cities. In equation 1 the coefficient associated with cost (TC - RC) had the expected sign and was significant (at the one percent level). However, the coefficient associated with transit time

Table XX. Elasticities, Firm B

	1- Both Origins		3- Origin 1		4- Origin 2	
Variable	Truck	Rail	Truck	Rail	Truck	Rail
Cost						
TC	-4.71	3.08	-1.15	9.87	-3.97	0.60
RC	4.01	-2.62	1.00	-8.53	3.36	-0.48
Transit Time						
TT	0.06	-0.04	-0.05	0.40	0.11	-0.02
RT	-0.58	0.38	0.43	-3.65	-1.20	0.17
Origin City						
OC	0.96	-0.63				

Table XXI. Estimated Equations, Firm C

Variable	1	2
Cost TC - RC	-0.0235 (-7.92)[1]	-0.0209 (-8.10)
Transit TT - RT Time	0.2075 (4.23)	
Dest. City Dummy	-2.9187 (-5.89)	-3.2755 (-6.87)
Likelihood Ratio (λ)	225.2	201.0
Pseudo R^2	.615	.549
Misclassified	24	29
Truck	14	14
Rail	10	15
Total Estimated Probability		
Truck	100.5	115.2
Rail	163.5	148.8

1) t - statistics are in parentheses

(TT - RT) did not have the expected sign and was significant (at the one percent level). In equation 2 the transit time variable had been deleted. The explanatory power of this equation was only slightly less than equation 1. According to the likelihood ratio test, both estimated equations were significant (at the one percent level). Also, the pseudo R^2 associated with both equations was fairly high. Equation 1 misclassified only 24 observations -- 14 truck out of 104 and 10 rail out of 160. Equation 2 misclassified only 29 (14 truck and 15 rail). Finally, the total estimated probabilities for both equations 1 and 2 appeared to be fairly close to the actual values (although equation 2 did somewhat worse).

Table XXII presents the elasticities computed at the means from equation 1. Although the elasticities associated with truck and rail transit times did not have the expected signs, truck cost and rail costs were by far the most important variables. A one percent increase in truck costs would decrease the probability of choosing truck by 15.6 percent and increase the probability of choosing rail by 3.35 percent. Conversely, a one percent increase in rail costs would increase the probability of choosing truck by 15.9 percent and decrease the probability of choosing rail by 3.41 percent. It appeared that mode choices made by this shipper were determined mainly by truck and rail shipment costs.

Conclusions

The empirical results from this and other studies (52, 53) had shown that logit analysis can be successfully applied to model the determinants of freight modal choice from disaggregate data. However, a major problem with this approach was how did one generalize from the disaggregate to the aggregate level? In other words, we were not really interested in how much Firms A, B, or C ship by rail or truck or even by pipeline. Rather, we would like to know how many tons of the commodity type shipped by these firms would go by pipeline. There was no easy answer to this problem. For the purposes of this report, we shall assume that these firms were representative of their respective industries. Since our firms were very large and constitute a significant part of the output in each of their industries, this might be a justifiable assumption. Researchers are just beginning to study how to make predictions from a disaggregate model of mode choice (58). In a later section of this report (Potential Demand for the Freight Pipeline), we shall illustrate how to use one of these dis-

---

(58) See Westin, Richard, "Predictions from Binary Choice Models", Journal of Econometrics, 1974, and Talvitie, Antti, "Aggregate Travel Demand Analysis with Disaggregate or Aggregate Travel Demand Models", Transportation Research Forum. 1973.

Table XXII. Elasticities, Firm C

Variable	Truck	Rail
Cost		
TC	-15.6	3.35
RC	15.9	-3.41
Transit Time		
TT	0.85	-0.18
RT	-1.58	0.34
Destination City Dummy	-0.81	0.17

aggregate models to predict the probability certain shipments will go by pipeline.

Pipeline Modal Split Analysis

Given the above macro and micro modal split models combined with a vector of expected pipeline performance with respect to rate, time, and reliability, evoking the theory of abstract modes (Quandt and Baumol), 12) yielded estimates of potential pipeline penetration into the existing truck and rail markets.

The basic idea behind the applications of the abstract mode theory shown below was the following: the existing models explained the existing modal choice between truck and rail by our three firms in the case of the micro models or the market split between truck and rail in the case of the macro models. For a given shipment in the micro case, the modal split model would designate either truck or rail (on a probability basis) as the carrier of that shipment. The vector of pipeline characteristics was now substituted for truck and the equation was calculated as a probability of going by rail or pipeline. Likewise the vector of pipeline characteristics was substituted for rail and the equation was calculated as a probability of going by truck or pipeline.

If the probabilities showed truck greater than rail, pipeline greater than rail, and pipeline greater than truck or rail greater than truck, pipeline greater than rail, and pipeline greater than truck, then pipeline unambiguously was assigned the traffic. As shown in the next section in a selected (chosen for variety purposes) sample of 20 observations from Firm A, under one assumption five and under a more liberal assumption eight observations could unambiguously be assigned to pipeline.

A similar type of analysis was performed with the macro models. Pipeline market shares were thus estimatable. The preliminary results which could be concluded from these equations were that pipeline appears capable of penetration of somewhere in the vicinity of 20-50% of the market sampled herein (although in one case almost total penetration resulted). It must be stressed that these results were quite preliminary and vary with the commodity investigated. Nevertheless, the acceptability to shippers could be shown by the use of the abstract mode methodology.

Since the origin-destination flows were not available by commodity nor did we have models for all pipelineable commodities, it was only by crude extrapolation that we could estimate the total flow potential of the pipeline. However, if our 20-50% penetration of the market held up, under the liberal matrix assumption, flows of between 10 and 25 million tons could be expected on the most dense links.

Potential Demand for the Freight Pipeline

In this section we shall demonstrate how the estimated micro and macro models would be utilized to estimate the potential demand that might exist for a freight pipeline system. It must be emphasized that at this point it was impossible to give a definitive answer to the question: how much tonnage would move by a freight pipeline? However, it was possible to begin to get a feeling for the relative competitiveness of a freight pipeline system vis-a-vis existing modes of freight transportation (truck and rail).

First, we shall demonstrate the use of one of the disaggregate models of modal choice. Here, we shall use the logit model estimated for Firm A. For purposes of comparison, a sample of twenty shipment observations representing a wide range of origin-destination pairs and shipment sizes was chosen. Of the twenty shipments, ten went by truck and ten by rail. For each shipment it was necessary to estimate the pipeline shipment cost, the pipeline transit time, and the pipeline transit time reliability. Pipeline shipment costs were calculated using figures obtained from the pipeline cost model. The commodity shipped by Firm A was of very low density (about seven lbs/cu ft)--pipeline costs were adjusted upwards to reflect this. Also, the pipeline costs were calculated assuming one hundred percent truck access at the origin and destination. Finally, these pipeline costs were multiplied by 1.10 to obtain the pipeline shipment costs to the shipping firm. This, of course, assumed pipeline rates would be ten percent above the costs to the "pipeline company." Pipeline transit times were calculated assuming twenty-one hours of terminal time at the origin and destination (the average of 18 and 24 hours, see above section) and a pipeline speed of 23.3 miles per hour. Both pipeline shipment costs and transit time were calculated based on a pipeline circuity of ten percent over straight line miles. Pipeline transit time reliability was assumed to be constant at 90 percent (90 percent of the shipments would arrive during the best three day period) over all origin-destination pairs.

For each sample shipment, three modal comparisons were made--truck versus rail, pipeline versus truck, and pipeline versus rail. For truck versus rail, equation 16 was used. The probability that the shipper would choose pipeline over truck was calculated for each shipment by substituting (PC-TC), (PT-TT), and (PR-TR) into the estimated logit model, where PC, PT, and PR were the pipeline shipment cost, pipeline transit time reliability respectively. Finally, the probability that the shippers would choose pipeline over rail was calculated by substituting into the estimated logit model (PC-RC), (PT-RT), and (PR-RR). The results of the calculations are presented in Tables XXIII and XXIV. The figures in Table XXIII were obtained using

Table XXIII. Modal Comparisons, Firm A

(Pipeline annual volume--Five million tons)

Truck vs. Rail		Pipeline vs. Truck		Pipeline vs. Rail		Shipment
Prob. Truck	Mode Selected	Prob. Pipe	Mode Selected	Prob. Pipe	Mode Selected	will go by:
0.93	Truck	0.04	Truck	0.39	Rail	Truck
0.74	Truck	0.16	Truck	0.35	Rail	Truck
0.99	Truck	0.21	Truck	0.97	Pipeline	Truck
0.99	Truck	0.08	Truck	0.97	Pipeline	Truck
0.17	Rail	0.60	Pipeline	0.22	Rail	Rail
0.03	Rail	0.26	Truck	0.01	Rail	Rail
0.05	Rail	0.80	Pipeline	0.12	Rail	Rail
0.09	Rail	0.99	Pipeline	0.48	Rail	Rail
0.46	Rail	0.03	Truck	0.01	Rail	Rail
0.09	Rail	0.99	Pipeline	0.98	Pipeline	Pipeline
0.17	Rail	0.84	Pipeline	0.49	Rail	Rail
0.01	Rail	0.99	Pipeline	0.93	Pipeline	Pipeline
0.82	Truck	0.21	Truck	0.54	Pipeline	Truck
0.51	Truck	0.79	Pipeline	0.78	Pipeline	Pipeline
0.04	Rail	0.99	Pipeline	0.55	Pipeline	Pipeline
0.38	Rail	0.40	Truck	0.30	Rail	Rail
0.98	Truck	0.39	Truck	0.97	Pipeline	Truck
0.77	Truck	0.14	Truck	0.34	Rail	Truck
0.78	Truck	0.29	Truck	0.58	Pipeline	Truck
0.66	Truck	0.99	Pipeline	0.96	Pipeline	Pipeline

Table XXIV. Modal Comparisons, Firm A

(Pipeline Annual Volume-- Ten Million Tons)

Truck vs. Rail		Pipeline vs. Truck		Pipeline vs. Rail		Shipment
Prob. Truck	Mode Selected	Prob. Pipe	Mode Selected	Prob. Pipe	Mode Selected	will go by:
0.93	Truck	0.07	Truck	0.51	Pipeline	Truck
0.74	Truck	0.21	Truck	0.43	Rail	Truck
0.99	Truck	0.25	Truck	0.97	Pipeline	Truck
0.99	Truck	0.11	Truck	0.98	Pipeline	Truck
0.17	Rail	0.81	Pipeline	0.44	Rail	Rail
0.03	Rail	0.70	Pipeline	0.01	Rail	Rail
0.05	Rail	0.99	Pipeline	0.26	Rail	Rail
0.09	Rail	0.99	Pipeline	0.59	Pipeline	Pipeline
0.46	Rail	0.10	Truck	0.01	Rail	Rail
0.09	Rail	0.99	Pipeline	0.97	Pipeline	Pipeline
0.17	Rail	0.99	Pipeline	0.63	Pipeline	Pipeline
0.01	Rail	0.99	Pipeline	0.98	Pipeline	Pipeline
0.82	Truck	0.27	Truck	0.63	Pipeline	Truck
0.51	Truck	0.82	Pipeline	0.87	Pipeline	Pipeline
0.04	Rail	0.99	Pipeline	0.81	Pipeline	Pipeline
0.38	Rail	0.55	Pipeline	0.42	Rail	Rail
0.98	Truck	0.51	Pipeline	0.98	Pipeline	Pipeline
0.77	Truck	0.23	Truck	0.51	Pipeline	Truck
0.78	Truck	0.37	Truck	0.67	Pipeline	Truck
0.66	Truck	0.99	Pipeline	0.98	Pipeline	Pipeline

pipeline costs based on a pipeline annual volume of five million tons, while those in Table XXIV were based on an annual volume of ten million tons. The mode ultimately chosen for each shipment was determined through a series of binary comparisons. For example, for the first sample shipment, truck would win over rail and truck would win over pipeline (using a cut-off probability of 0.50.). Thus, truck would be chosen.

Using this method based on an annual volume of five million tons, truck would be chosen 8 times, rail 7 times, and pipeline 5 times. It must be emphasized that our sample of shipment observations did not represent a random sample of all the shipments made by Firm A. Therefore, it was not possible to say 25 percent of all shipments made by Firm A would go by pipeline. However, our records did indicate that pipeline would probably capture all shipments made by Firm A over about 300 miles. If the annual pipeline volume was ten million tons, truck would be chosen 7 times, rail 5 times, and pipeline 8 times. These results indicated that at a fairly high volume pipeline would also be somewhat competitive at distances 100 to 300 miles. Since the longest shipment by Firm A was 550 miles, it was not possible to make any statements concerning the competitiveness at greater distances.

Here we shall demonstrate how the aggregate mode split models could be utilized to estimate the potential demand for the freight pipeline. For this purpose we used the estimated equations for three of the commodity groups: STCC 208 (Beverages, etc), STCC 264 (Converted Paper Products), and STCC 307 (Misc. Plastic Products). These equations were used to compare truck versus rail, pipeline versus truck, and pipeline versus rail at four distances -- 100, 300, 500, and 700 miles. To do so it was necessary to calculate truck, rail, and pipeline rates and truck, rail, and pipeline transit times. Truck and rail rates were calculated from the rate regressions presented in Table XV. Pipeline rates were obtained by multiplying pipeline costs derived from the pipeline cost model (assuming 100 percent truck access) by 1.10. As explained before this assumed pipeline rates would be ten percent above costs. Pipeline rates were adjusted to reflect the density of the commodity shipped. Also, truck and rail rates were multiplied by 1.4 and 1.3, respectively, to bring them up to 1973 levels. Rail and truck transit times were calculated from equations (11) and (12). Pipeline transit times were obtained by assuming 21 hours of terminal time and a pipeline speed of 23.3 miles per hour. These rates and transit times were calculated assuming circuities of 20 percent for truck, 25 percent for rail, and 10 percent for pipeline over straight line miles.

For the three commodity groups we first compared truck versus rail and calculated Pertruc (the percent of tonnage going by truck) by substituting (TR/RR) and (RT/TT) into the estimated mode split models; where TR, RR, TT, and RT were truck rate, rail rate, truck transit time, and rail transit time, respectively. Then we compared pipeline versus truck. We calculated Perpipe (the percent of tonnage going by pipeline) by substituting (TR/PR) and (PT/TT) into these equations and subtracting the result from 100. Finally, we compared pipeline versus rail and calculated Perpipe by substituting (PR/RR) and (RT/PT) into the mode split equations. Here, PR was the pipeline rate and PT was the pipeline transit time. A problem with this approach was how to handle the constant term in the mode split equations. The constant represented a consistent bias in favor of truck over rail. This might reflect in part the reliability advantage of truck compared to rail. As was assumed, there would be a constant bias in favor of pipeline when compared to rail (pipeline would be more reliable than rail). The percentage of the tonnage that would go by pipeline at each distance was calculated using the following equation:

$$\text{Percent Pipeline} = \left\{ \begin{array}{l} (\text{Pertruck})\text{X Perpipe (Pipe vs. Truck)} + \\ \quad (100 - \text{Pertruck}) \text{ X Perpipe (Pipe vs. Rail)} \end{array} \right\} /100 \qquad (23)$$

This assumed the percentage that would go by pipeline would reflect the estimated truck/rail split and the percent pipeline would take from truck and from rail. The results of these calculations for the three commodity groups and for two annual pipeline volumes (5 and 10 million tons) are presented in Tables XXV, XXVI, and XXVII.

A word of caution should be inserted before attempting to interpret these results. It must be remembered that the estimated mode split equations had fairly low explanatory powers. They are used here only to get a feeling for the relative competitiveness of pipeline versus truck and rail in terms of rates and transit times. Generally, pipeline rates fell in between truck and rail (cheaper than truck and more expensive than rail). Pipeline transit times also fell in between truck and rail (faster than rail and slower than truck). The results indicated that a freight pipeline system can be very competitive with both truck and rail. It appeared pipeline might be most competitive at distances between 300 and 500 miles. Pipeline rates relative to rail rates begin to increase between 500 and 700 miles and thus, pipeline becomes less competitive. Also, at higher annual volumes pipeline rates were lower and, therefore, pipeline was more competitive at all distances.

The results from both the micro and macro models indicated that pipeline could be competitive with truck and rail over a fairly wide range of commodities. However, it was not easy to estimate the potential demand for a new mode of freight transportation and much more research needs to be devoted to this subject.

Table XXV. Modal Comparisons, STCC 264 (Converted Paper Products)

Miles	Pipeline Annual Volume--5 Million Tons				Pipeline Annual Volume--10 Million Tons			
	Truck vs.	Pipeline vs.		Percent by	Truck vs.	Pipeline vs.		Percent by
	Rail	Truck	Rail		Rail	Truck	Rail	
	Pertruck	Perpipe	Perpipe	Pipeline	Pertruck	Perpipe	Perpipe	Pipeline
100	50.93	56.98	7.15	32.53	50.93	69.67	10.11	40.44
300	23.69	93.20	20.91	38.04	23.69	95.77	34.41	48.96
500	12.90	96.24	19.85	29.70	12.90	97.73	34.20	42.35
700	8.02	97.19	17.25	23.65	8.02	98.33	28.39	33.98

Table XXVI. Modal Comparisons, STCC 208 (Beverages, etc.)

	Pipeline Annual Volume--5 Million Tons				Pipeline Annual Volume--10 Million Tons			
	Truck vs.	Pipeline vs.			Truck vs.	Pipeline vs.		
Miles	Rail	Truck	Rail	Percent by	Rail	Truck	Rail	Percent by
	Pertruck	Perpipe	Perpipe	Pipeline	Pertruck	Perpipe	Perpipe	Pipeline
100	100.00	23.67	12.41	23.67	100.00	44.13	17.20	44.13
300	35.20	88.75	26.55	48.54	35.20	92.69	41.68	59.64
500	17.68	93.65	22.65	35.20	17.68	95.93	36.38	46.91
700	9.80	95.32	18.01	25.59	9.80	97.05	29.08	35.74

Table XXVII. Modal Comparisons, STCC 307 (Misc. Plastics)

	Pipeline Annual Volume--5 Million Tons				Pipeline Annual Volume--10 Million Tons			
	Truck vs. Rail	Pipeline vs. Truck	Pipeline vs. Rail	Percent by	Truck vs. Rail	Pipeline vs. Truck	Pipeline vs. Rail	Percent by
Miles	Pertruck	Perpipe	Perpipe	Pipeline	Pertruck	Perpipe	Perpipe	Pipeline
100	87.28	76.40	36.57	71.33	87.28	81.99	48.18	77.69
300	56.85	92.03	70.65	82.81	56.85	94.60	100.00	96.93
500	36.57	94.40	66.45	76.67	36.57	96.32	100.00	98.65
700	26.01	95.51	56.57	66.70	26.01	96.99	85.89	88.70

we wish to investigate in the second year of research, it can be concluded that the market which was economically feasible for freight pipeline to penetrate would be of a significant magnitude, e.g., 10-25 million tons/year in the corridor of the hypothetical network. However, such conclusions await the results of further research. If such results were substantiated in the second year research, the economic feasibility of a solid freight pipeline would appear to be assured.

It should be mentioned that the results herein must be conditioned by the availability of adequate data needed to carry out the research. The current status of flow data between origin-destination by commodity was woefully inadequate. To some extent this will be rectified by the 1977 Census of Transportation, which will increase its commodity coverage from manufactured products to include raw materials. However, movements of commodities from warehouses and distribution points still will remain uncovered. Although data on the national level was comprehensive, when one attempted to investigate specific commodity flows on specific origin-destination links, the sampling variability of the Census flows became very high. The feasibility of expanding the sample size should be explored. The Census would seem to be the likely vehicle to obtain such flow data. Open meetings should be held sponsored by DOT, and the Bureau of the Census, where the DOT and other users could express the research interests that could be fulfilled by a modified and improved Census. A meeting much like the one suggested above was held in November, 1976 in California and in March, 1977 in Washington. However, the former was a part of a much larger meeting dealing with Census data. Such a meeting should be the first of many.

Since the Census did collect information on mode split of the shipments surveyed, it would seem feasible for the Census to also collect information on the performance characteristics of the chosen mode and the best alternative mode for a sample of the observations. Modal split modellers would benefit from a common data base (for example the results of various model types could be directly compared against one another if a common data base was used). In addition, the substantial cost of developing a primary source data bank by many different researchers could be avoided. The FHWA is currently sponsoring a similar data set construction for urban passenger modal split modelling (current FHWA contract with Charles River Associates). The task suggested herein should be easier than FHWA's because Census already has the contact with the shippers.

Research such as that described within this volume would also benefit if a trucking analogue to the 1% railway bill sample were published. Trucking rate bureaus already collect such information so the problem is not one of collection but rather one of getting the information released and co-ordination. DOT coverage or legis-

Conclusions

The analysis done herein suggested that sufficient flows of physically pipelineable commodities existed on a hypothetical network connecting Chicago to New York and St. Louis to New York. When traffic was assumed to be limited in origin and destination to the 18 cities on the network, single direction tonnages of up to 16 million tons/year were found. If trans-shipment was allowed and hence other cities in proximity to the network could ship and/or receive via the network and some connecting mode, upwards of 38 million tons/year in a single direction could be generated.

Whether such tonnages which could physically move by pipeline were likely to do so depended on the service characteristics (rates, time in transit, reliability, loss and damage, etc.) offered by pipeline vis-a-vis those same characteristics offered by the competing modes.

The demand and modal split analysis performed showed that for selected STCC commodities that were considered physically pipelineable and for which Census of Transportation data was available, the existing modal split between truck and rail could be reasonably replicated given estimates of modal rates and transit times. Assuming the abstract mode approach of Quandt and Baumol (12), pipeline penetration into the market was estimated. In general, the STCC estimations found that between 20-50% of the commodity shipments could be penetrated by pipeline depending on the annual volume assumed for the pipeline and distance shipped. However, in one market, STCC 307, almost total penetration seemed possible.

The above analysis, called the aggregate analysis, was complemented with a disaggregated analysis on specially collected modal split data from three major national shippers. The disaggregate analysis showed that individual shipper's modal split decisions between truck and rail could be predicted with 90% accuracy using modal rates, time and reliability. The pipeline penetration analysis was performed on a sample of data of one of our three firms. This study showed pipeline penetration of between 25% - 40% of the shipments analyzed. More generally the pipeline would appear to penetrate all shipments of this firm greater than 300 miles.

To date the market penetration potential of pipeline for several of the STCC's which are physically pipelineable had been analyzed on both an aggregate and disaggregate basis. This was by no means a complete analysis of all commodities that made up our traffic volume matrix. Nor were the estimates of modal split free from criticism (since the R^2's were not always of high value - due, we hypothesized, to the lack of data availability on modal characteristics other than rate and time). Nevertheless, if the results thus far obtained hold up for the other STCC's which lation should be explored as a vehicle to obtain such information. Such a sample (or complete enumeration - it's been suggested by some that complete enumeration of rail movements would be no more costly than the 1% sample since most railroads computerize all movements and must abstract from the total to obtain the 1%) would enable the construction of motor carrier flows, estimation of truck rates, etc.

The collection of the above mentioned data is well within the physical capability of the existing or potential collectors. The economic feasibility of collecting such information should be explored with a DOT-TRB (Transportation Research Board) sponsored conference (much like the TRB conference held in March, 1977) as the best vehicle for obtaining information related to such feasibility.

Without such information, we will have little reliable information on current traffic flows and little information on the causation of freight modal split. Without knowing existing flows and the cause of existing modal splits, transportation infra-structure planning and the desirability of changes in the transportation institutions, e.g., regulation, will be difficult to carry out and assess.

*U.S. GOVERNMENT PRINTING OFFICE : 1977 0-732-014-162

REQUEST FOR FEEDBACK TO
The DOT Program Of University Research

DOT-TST-76T-38

YES	NO	
☐	☐	Did you find the report useful for your particular needs? If so, how?
☐	☐	Did you find the research to be of high quality?
☐	☐	Were the results of the research communicated effectively by this report?
☐	☐	Do you think this report will be valuable to workers in the field of transportation represented by the subject area of the research?
☐	☐	Are there one or more areas of the report which need strengthening? Which areas?
☐	☐	Would you be interested in receiving further reports in this area of research? If so, fill out form on other side.

Please furnish in the space below any comments you may have concerning the report. We are particularly interested in further elaboration of the above questions.

Cut Out Along This Line

COMMENTS

Thank you for your cooperation. No postage necessary if mailed in the U.S.A.

RESEARCH FEEDBACK

Your comments, please . . .

This booklet was published by the DOT Program of University Research and is intended to serve as a reference source for transportation analysts, planners, and operators. Your comments on the other side of this form will be reviewed by the persons responsible for writing and publishing this material. Feedback is extremely important in improving the quality of research results, the transfer of research information, and the communication link between the researcher and the user.

FOLD ON TWO LINES, STAPLE AND MAIL.

Fold Fold

DEPARTMENT OF TRANSPORTATION
OFFICE OF THE SECRETARY
Washington, D.C. 20590

Official Business

PENALTY FOR PRIVATE USE, $300

POSTAGE AND FEES PAID
DEPARTMENT OF
TRANSPORTATION
DOT 518

U.S. MAIL

Office of University Research
Office of the Secretary (TST-60)
U.S. Department of Transportation
400 Seventh Street, S.W.
Washington, D.C. 20590

Fold Fold

IF YOU WISH TO BE ADDED TO THE MAIL LIST FOR FUTURE REPORTS, PLEASE FILL OUT THIS FORM.

Name ______________________________ Title ______________
Use Block Letters or Type

Department/Office/Room ______________________________

Organization ______________________________

Street Address ______________________________

City ____________________ State __________ Zip __________