ESTIMATION AND PREDICTION OF INTERSTATE SECTORAL FREIGHT FLOWS FOR NEW JERSEY:
A REVIEW OF THE LITERATURE AND DATA SOURCES

by

Theologos Homer Bonitsis*
and
Kenneth D. Lawrence**

Co-principal Investigators

National Center for Transportation and Industrial Productivity
(NCTIP)

(A research report submitted under NCTIP 9-91992.)

*Associate Professor of Finance
**Professor of Management Science
School of Industrial Management
New Jersey Institute of Technology
University Heights
Newark, New Jersey 07102-1982
Tel. no.: (201) 596-3263
Tel. no.: (201) 596-6425
Fax no.: (201) 596-3074
May 28, 1996
ESTIMATION AND PREDICTION OF INTERSTATE SECTORAL FREIGHT FLOWS FOR NEW JERSEY:

A REVIEW OF THE LITERATURE AND DATA SOURCES

ABSTRACT

This report is divided into two parts. Part one is a succinct non-mathematical review of the contributions to the literature with regard to estimating and predicting traffic flows with an emphasis on the gravity model. Complementary to this, part two presents various data sources for freight flow estimation to and from the State of New Jersey.

PART ONE: Traffic Flow Estimation Techniques

Interest in investigating interstate traffic flows has a long history. Spiegelglas (18), for example, investigated state to state commodity flows within the U.S. as early as 1960. He noted at the time that more is probably known about international commodity flows than interregional state flows. This point is essentially valid today for interstate commodity flows, and particularly for intermodal flows where no database exists. Pfister (13), lifting a tool from international trade theory, suggests that the terms of trade is useful in the analysis of relative regional economic positions. By extension, this suggests that the terms of trade should also have explanatory power to explain regional commodity flows.

There are several empirical models that may be employed to estimate traffic flows. One method of estimation would be the formulation of an input-output model. Moses (12) in an early work discusses trading patterns between regions within the context of an input-output model. He
notes that this methodology embodies three premises: first, structural stability of the model's technical coefficients; second, uniform trade relations across sectors within a region; and third, stability of trading patterns. These caveats, given the current rapid pace of innovation in production and transport technologies, are very strenuous to accept today. Moreover, the construction of an input-output model is extremely cumbersome. It is clear, therefore, that there is no dearth of alternative approaches to address the issues of traffic flow estimation.

A widely used model to predict traffic flow is the gravity model, it is an extension of Isaac Newton's theory of gravity to an application of traffic flows. The gravity model essentially states that trip volume between areas is directly proportional to the relative attraction of each area and inversely proportional to some function of the spatial separation between areas. Consider, for example, the following standard gravity model:

$$T_{ij} = \frac{P_i \ A_j \ F_{ij} \ K_{ij}}{\sum_{j=1}^{n} A_j \ F_{ij} \ K_{ij}}$$

Where $T_{ij}$ is the total trips from zone $i$ to zone $j$, $P_i$ is the total trips produced by zone $i$; $A_j$ relates the total trips attracted to zone $j$; $F_{ij}$ is an empirical friction factor that represents the spatial separation effects of zone $i$ and zone $j$; and $K_{ij}$ is an adjustment factor based on socioeconomic characteristics of the origin and destination zones $i$ and $j$. Hence, five parameters are needed to estimate the model. One factor not extensively dealt with in the literature is a parameter that is an array of zone-to-zone socioeconomic adjustment factors that influence travel patterns but are not embodied elsewhere in the model, e.g., income and occupation.

Wilson (22) offers a theoretical basis for spatial distribution models so as to propose a methodology that leads to new models. He points out an important deficiency in the gravity model; namely, that if both origin and destination each double then there is a quadrupling of number of trips between zones. To correct this shortcoming Wilson offers a modified gravity model with two constraints that are solved iteratively, this satisfactorily addresses the problem. In
addition, the study's findings indicate the theoretical robustness of the gravity model over the intervening opportunities model. Similarly, Black (2) concludes that gravity models developed from transportation studies are robust for estimating interregional commodity flows. He also finds that the distance parameter's exponent is inversely related to the magnitude of the proportion of total shipments from the largest shipper and directly related to the magnitude of the proportion of intraregional total flow. Interestingly, Black finds only a weak relationship for a direct relationship between the exponent of the distance parameter and the economic value of the commodity flow. Related to this work is the study by Gordon (8), which investigates several hypotheses with regard to transport costs, commodity values, and distance deterrence. Interestingly, Gordon finds by employing a gravity model, that transport costs is a small ratio of total distance costs and that the steepness of the distance deterrence function is directly related to the weight-value ratio of a commodity; the later finding, contrary to other studies, is found to be statistically significant.

Similarly, there are numerous studies on calibrating the gravity model. Williams (9) conducts a comparative study of three calibration techniques for the standard doubly constrained gravity model with exponential costs functions to conclude that a modified Hyman method is most efficient, i.e., produce the most robust estimates. While a study by Sen and Soot (16) presents several methods for gravity model calibration, these are the Iterated Proportional Fitting Procedure (IPFP), linear least squares, and maximum likelihood. With regard to the last method, Sen (17) shows within a theoretical context that maximum likelihood estimates of gravity model parameters exist and are unique. Another study by Yun and Sen (19) shows that a Modified Scoring Procedure seems to be an efficient procedure for extracting maximum likelihood estimates of gravity model parameters.

A more sophisticated technique for calibration is suggested by Flowerdew and Aitkin (5), their method entails constructing a Poisson model to overcome the Problems associated with estimating a lognormal specified gravity model. In another paper Flowerdew (6) presents an iterative weighing method to correct for heteroscedasticity in estimating lognormal gravity models; albeit the poison method still yields more robust results.

Weber (21) concludes that the signs of both the derivatives and elasticities of gravity model flows are not independent of the calibration process employed; hence, the reconciliation of
origin and destination flows is an integral part of the gravity model of flows. In an earlier paper, Weber and Sen (20) present a novel approach to estimating incremental changes in gravity model generated forecasts resulting from incremental changes in input values. This is achieved by a linearization of the iterative Deming-Stephan-Furness (DSF) procedure. Goncalves and Ulysseaneto (7) integrate both the gravity model and the intervening-opportunities model so as to construct a "new gravity-opportunity model" for trip distribution; this model's attributes are its accuracy, operationality, and simplicity.

With regard to applications of the gravity model Christerson (3) finds that trade flows in world apparel are consistent with those predicted by a simple gravity model based on distance, size and cost parameters. This finding is particularly encouraging in that the gravity model is found to have properties consistent with sound theory; namely, simplicity, generality, and good predictive power. A study by Bailey and Munford (1) applies a more sophisticated gravity model that assumes a Poisson distribution for the flows under study with similar robust fitted results. In contrast, an earlier paper by Pitfield (14) applies a doubly constrained gravity model with a deterrence power function to predict U.K. rail freight for an array of goods, his findings do not generate qualitatively good predictions. Knudsen (10) employing only a Poisson regression analyzes interannual variation in U.S. rail freight flows, he concludes that flow variability dominate regional import-export variability. It is noteworthy to point out that the gravity model can be applied to any type of flow modelling problem. Plane (15), for example, uses a doubly constrained gravity model to investigate interstate population movements; Cromley, Hempel, and Hillyer (4) employs the model to examine the parameters of market attractiveness.

PART TWO: Data Sources for Freight Flow Estimation

I. AN INTRODUCTION TO AVAILABLE DATA RESOURCES FOR FREIGHT FLOW PREDICTIONS TO AND FROM NEW JERSEY
This research document is intended to delineate the data resources available in the analysis and prediction of freight flow to and from New Jersey. The basic structure of each of these data is given in the report. The data sets to be reviewed are as follows:

1. **BEA Regional Projections and Historical Data for Each State and Each Region**
2. **1992 Census of Transportation**
3. **1993-1994 Regional Truck Corridor Survey, Port Authority of New York and New Jersey and NJDOT**
4. **1993 Commodity Flow Data**
5. **Flow of Trucks into, through, and out of New Jersey**

**II. BUREAU OF ECONOMIC ANALYSIS DATA**

BEA Regional Projections and Historical Data for Each State and Each Region


Projections: 1995-2040

1. Total personal income
2. Total earnings
3. Total population
4. Per capita personal income
5. Total employment
6. Sources of personal income by 57 industries
7. Employment by place of work by 57 industries

The following sectors are considered as they are of particular interest to transportation planners:

1. Railroads and related services
2. Local and suburban and transit and interurban highway passenger transportation

The Bureau of Economic Analysis, through the use of its regional input-output system provides a systematic analysis of the regional inter-industry relationship. Through the use of regional input-output multipliers, which account for the inter-industry relationships, an industrial distribution of inputs purchases and outputs sold is produced for each industry for each region.
BEA projects a set of regional relationships every five years for both the fifty, states and the BEA economic regions (336 economic regions). Earnings and employment are projected for fifty-seven industries for each of the fifty states. The industries are categorized according to the 1972 Standard Industrial Classification (SIC) codes. They reflect approximately the 2-digit industrial detail with this classification.

**Industries Projected for States and the Nation**

**A. Farm**

**B. Agricultural Services, Forestry, Fisheries, and Other**

**C. Mining**

1. Coal mining
2. Oil and gas extraction
3. Metal mining
4. Nonmetallic minerals, except fuels

**D. Construction**

**E. Manufacturing**

1. Nondurable goods
   a. Food and kindred products
   b. Tobacco manufacturers
   c. Textile mill products
   d. Apparel and other textile products
   e. Paper and allied products
   f. Printing and publishing
g. Chemicals and allied products  
h. Petroleum and coal products  
i. Rubber and miscellaneous plastic products  
j. Leather and leather products  

2. Durable goods  
a. Lumber and wood products  
b. Furniture and fixtures  
c. Stone, clay, and glass products  
d. Primary metal industries  
e. Fabricated metal products  
f. Machinery, except electrical  
g. Electric and electronic equipment  
h. Transportation equipment, excluding motor vehicles  
i. Motor vehicles and equipment  
j. Instruments and related products  
k. Miscellaneous manufacturing industries  

F. Transportation and Public Utilities  
1. Railroad transportation  
2. Trucking and warehousing  
3. Local land interurban passenger transit  
4. Transportation by air  
5. Pipelines, except natural gas  
6. Transportation services  
7. Water transportation  
8. Communication  
9. Electric, gas, and sanitary services  

G. Wholesale Trade  

H. Retail Trade  

1. Finance, insurance, and real estate  
   1. Banking and credit agencies  
   2. Holding companies and investment services  
   3. Insurance  
   4. Real estates  

J. Services  

1. Hotels and other lodging places  
2. Personal services  
3. Business and miscellaneous repair services  
4. Auto repair, services, and garages  
5. Amusement and recreation services and motion pictures  
6. Private household
7. Health services
8. Legal services
9. Educational services
10. Social services and membership organizations
11. Miscellaneous professional services

K. Government and government enterprises
   1. Federal, civilian
   2. Federal, military
   3. State and local

Earnings and employment are projected for fourteen industries for Metropolitan Statistical Areas (MSAs) and BEA economic areas.

B. BEA Data

The Bureau of Economic Analysis (BEA) developed a system of regional modeling in the 1970s. This modeling is based on the concepts and analysis of input-output analysis. An input-output table shows for each industry the industrial distribution of inputs purchased and outputs sold. The BEA input-output model shows the input and output structure of more than 500 industries within the United States. The BEA, by using the county-wage and salary, adjust the national input-output to show a regional (state) industrial structure and trading patterns. This data is available in CD-ROM format.

The following variables are produced as outputs and forecasts developed from the regional (state) input-output model.

1. \( p_i \): population centroid of each state (largest city in state) population of state
2. \( em_i^k \): employment level in the \( i^{th} \) state of the \( k^{th} \) industry
3. \( ea_i^k \): earnings level in the \( i^{th} \) state of the \( k^{th} \) industry
4. \( op_i^k \): output level in dollars in the state of the \( k^{th} \) industry
5. \( pe_i^k \): personal income level in the \( i^{th} \) state for the \( k^{th} \) industry
6. \( ep_i \): employment to population ratio in the \( i^{th} \) state
7. \( ppy_i \): percent of the population under 18 years of age in the \( i^{th} \) state
8. \( ppo_i \): percent of the population over 65 years of age in the \( i^{th} \) state
9. \( ppi_i \): per capita income in the \( i^{th} \) state
III. CENSUS OF TRANSPORTATION (1992)

The Truck Inventory and Use Survey provides data on the physical and operational characteristics of the U.S. truck population. This study was based on a probability sample of private and commercial trucks registered in each of the fifty states in 1992. The size of the sample was approximately 150,000 trucks out of a universe of over 60 million trucks. The major facets of the study included the following characteristics of trucks, 1987-1992. This data is available in printed format.

A. Major Use
   1. Agriculture
   2. Forestry and lumbering
   3. Mining and quarrying
   4. Construction
   5. Manufacturing
   6. Wholesale and retail trade
   7. For hire transportation
   8. Utilities and service
   9. Personal transportation
  10. Other

B. Body Type
   1. Pickup panel or minivan
   2. Platform and cattlerack
   3. Van
   4. Public utility
   5. Dump
   6. Tank or liquid or dry bulb
   7. Other

C. Vehicle Size
   1. Light
   2. Medium
   3. Light-heavy
   4. Heavy-heavy

D. Annual Miles
   1. Less than 5000 miles
   2. 5000-9999
   3. 10000 to 19999
   4. 20000 to 29999
   5. 30000 or more
E. Year Model

1. 1-2 years old
2. 3-4 years old
3. Over 4 years old

F. Vehicle Acquisition

1. Purchase new
2. Purchased used
3. Leased from someone

G. Truck Type

1. Single-unit Trucks
2. Combination

H. Range of Operation

1. Local
2. Short-range
3. Long-range
4. Off the road

I. Fuel Type

1. Gasoline
2. Diesel
3. Not reported

J. Primary Products Carried

33 types

K. Hazardous Material Carried
24 classifications

L. Truck Fleet Size

M. Miles Per Gallon

N. Equipment Type

O. Maintenance

P. Engine Type and Size

Q. Cab Type

R. Lease Characteristic

S. Primary Operator Classification

T. Model Year

U. Average Weight

V. Weeks Operated

IV. 1993-1994 REGIONAL TRUCK CORRIDOR SURVEY, PORT AUTHORITY OF NEW YORK AND NEW JERSEY AND N.J.D.O.T

Introduction
In the period December 1993 through May 1994, surveys were conducted of truck drivers. This survey was conducted at nine of the toll facilities located along the outer area of the Port authority of New York and New Jersey 29-county planning region and long the long Island Expressway east-and west-bound service roads. The surveys were conducted for a 24 hour period in each direction that tolls were collected. In addition to the truck driver surveys, truck and van counts were conducted at five locations in the New York/New Jersey region. The objective of the corridor study is to better understand truck movement in and through the region and to quantify the origination and destination patterns of trucks using the region's highway network. This data is available on diskettes.

Survey
The 1993 - 1994 Truck Corridor study is a follow-up to the Phase 1 survey conducted in November 1992. The interview survey of questions the truck drivers about their origin and destination, the commodities they were carrying, and about their trucks' characteristics. The commodities were coded using the 4-digit standard transportation costs for food and waste products and two digit SIC codes for all other loads.

The survey was conducted at the following locations:
1. Newburgh-Beacon Bridge
2. Woodbury Ticket Barrier (north- and south-bound)
3. Exit 6 of the New Jersey Turnpike
4. New Rochelle Toll Barrier 1-95
5. Exit 7A of the New Jersey Turnpike
6. Yonkers Toll Barrier of I-287 (north- and south-bound)
7. Entrance to the Pennsylvania Turnpike, Interchange 3
8. Harrison (westbound and eastbound)
9. Delaware Memorial Bridge
10. Long Island Expressway Exists 40-42

Output of the Study
1. Distribution of trucks by number of axles at each study point
2. Distribution of trucks (total, small, large):
   a. straight
   b. utility
   c. container
d. flatbed
e. autocarrier
f. tanker
g. double trailer
h. no trailer
i. all other at each study point
3. Commodity carried at each study point
4. Local distribution at each study point:
   a. full
   b. partial
   c. empty
d. doesn't know
5. Origin of trip at each study point
6. Facility type pickup location at each study point
7. Distribution by destination at each study point
8. Facility type delivery location at each study point
9. Facility type pair (origin and destination) at each study point
10. Commodity facility pickup pair toll direction at each study point
11. Trailer width (102 inches, 96 inches or less, doesn't know at each study
12. Origin/destination pair toll direction at each study point

V. 1993 COMMODITY FLOW DATA

The 1993 Commodity Flow Survey provides data on the movement of goods by mode of transportation. This sample survey produced measures of the movement of goods by major type of commodity shipped and mode of transportation used. The survey covers establishments in such industrial groups as mining, manufacturing, wholesale trade and selected retail and service industries. The study excludes such establishments as farms, governments, transportation, households, etc. The industries covered are defined in the standard industrial classifications (SICs). This data will be available on diskette and CD-ROM.

<table>
<thead>
<tr>
<th>SIC Code</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>10, ex. 108</td>
<td>Metal mining (excluding metal mining services)</td>
</tr>
<tr>
<td>12, ex. 124</td>
<td>Coal Mining (excluding coal mining services)</td>
</tr>
<tr>
<td>14, ex. 148</td>
<td>Mining and quarrying of nonmetallic minerals, except fuels (excluding nonmetallic minerals services)</td>
</tr>
<tr>
<td>20</td>
<td>Food and kindred products</td>
</tr>
<tr>
<td>21</td>
<td>Tobacco products</td>
</tr>
<tr>
<td>22</td>
<td>Textile mill products</td>
</tr>
<tr>
<td>23</td>
<td>Apparel and other finished products made from fabrics and similar materials</td>
</tr>
</tbody>
</table>
24 Lumber and wood products, except furniture
25 Furniture and fixtures
26 Paper and allied products
27, ex. 279 Printing, publishing and allied industries (excluding service industries for the printing trade)
28 Chemicals and allied products
29 Petroleum refining and related industries
30 Rubber and miscellaneous plastics products
31 Leather and leather products
32 Stone, clay, glass and concrete products
33 Primary metal industries
34 Fabricated metal products, except machinery and transportation equipment
35 Industrial and commercial machinery and computer equipment
36 Electronic and other electrical equipment and components, except computer equipment
37 Transportation equipment
38 Measuring, analyzing and controlling instruments; photographic, medical and optical goods; watches and clocks
39 Miscellaneous manufacturing industries
50 Wholesale trade --durable goods
51 Wholesale trade --nondurable goods
596 Catalog and mail-order houses
782 Motion picture and video tape distribution

Modes of Transportation

1. Parcel, US Postal Service or courier
2. Private Truck
3. For Hire Truck
4. Railroad
5. Inland Water or Great Lakes
6. Deep Sea Water
7. Pipeline
8. Air
9. Other Mode
10. Mode Unknown

Shipment

1. Ton Miles
2. Tons Shipped
3. Total Modal Activity
4. Value of Shipments
Flows between 89 BEA regions, as well as fifty states and Distinct of Columbia

1. $f_{ij}^k$: flow of freight by truck of the $k^{th}$ industry between the $i^{th}$ state and the $j^{th}$ state.

2. $d_{ij}$: distance between the population centroid of the $i^{th}$ state and the population centroid of the $j^{th}$ state.

VI. FLOW OF TRUCKS INTO, THROUGH AND OUT OF NEW JERSEY

In order to view the flow of fright carried by trucks into through and out of New Jersey, various agencies which manage the major highway systems and bridges were contacted. These groups provide monthly data on the flow of trucks over their bridges and highways. These flow of trucks were given on a monthly basis for the following entities:

1. Delaware River Joint Bridge Commission
   a. 7 toll bridges
   b. 13 non-toll bridges

2. New Jersey Turnpike Authority

3. Port of New York Authority

4. Delaware River Port Authority
5. New Jersey Department of Transportation

flow of trucks by size by month 1993

In combination with the various market research studies, statistical relationships between the flow of trucks and the freight carried can be devised. These dates are available on printouts from the various organizations. Moreover, List and Tumquist (11) suggest a method for estimating multi-class trip matrices from data bases that are fragmented and/or incomplete, but that are readily available; their method employs a routing algorithm to construct link use coefficients for each origin-destination pair. This approach, therefore, allows for the use of all available data independent of form and source and hence may be useful in the next phase of this research endeavor.

References for Modelling


Final 96M.REP


**References for Data Sources**


6. Correspondence with various bridge and highway authorities, 1993.