Final Report on Contract Number NCTIP97-21:

Identifying Number: NCTIP 1997-21 Extension

Project Title: Development of a Freight Forecasting Model to Forecast Truck Flow Between NJ Counties Themselves and Between NJ Counties and Other States

Principal Investigators:

Dr. Kenneth Lawrence

School of Management

New Jersey Institute of Technology
973-596-6241
Lawrence@Tesla.NJIT.Edu

Dr. Gary Kleinman

Silberman College of Business Admin.

Fairleigh Dickinson University

201-692-7258

Gklei49593@cs.com

Prepared by the principal investigators For the New Jersey Department of Transportation Division of Research and Technology and the National Center for Transportation and Industrial Productivity, at the New Jersey Institute of Technology.

Identifying Number: NCTIP 1997-21 Extension

Project Title: Development of a Freight Forecasting Model to Forecast Truck Flow Between NJ Counties Themselves and Between NJ Counties and Other States

Principal Investigators:

Dr. Kenneth Lawrence

School of Management

New Jersey Institute of Technology
973-596-6241
Lawrence@Tesla.NJIT.Edu

Dr. Gary Kleinman

Silberman College of Business Admin.

Fairleigh Dickinson University

201-692-7258

Gklei49593@cs.com

External Contact:John Powers
CN600

NJ Department of Transportation

1035 Parkway Avenue, Trenton, NJ

Project Abstract: First, to create a model capable of predicting commodity flow information via trucking in order to assist the appropriate agencies in understanding when and where new roads are needed. The prediction model will be based upon a gravity flow model. The gravity flow model argues that truck traffic between any two points is inversely proportional to the distance between the two points, and otherwise proportional to the population of the two areas and the economic health of the areas. A database of economic and trucking information will be developed for use in predicting truck-based freight flows between Bureau of Economic Affairs (BEA) regions, between NJ and other states, and between counties within the state of New Jersey. Database tools will be developed to make different views of this data accessible to the decision maker. Second, we will create a database, data management, and data forecast analysis system to assist the appropriate agencies in understanding when and where new roads are needed. This prediction model will be based upon the gravity flow model that was found to best predict historic freight flows.

Table of Contents

Introduction	4
Methodology	5
Data Sources and Descriptions	7
Statistical Testing	9
Variable Definitions	10
Results	11
Pivot Table and Chart Systems	13
The Forecasting Process	17
Bibliography	18
Table 1: Characteristics of Best Between State Freight Flow Prediction Model	20

The Development of a Model and Decision Support System to Use in Forecasting Truck Freight Flow in the Continental United States

Introduction

This research develops a regression-based model for forecasting truck borne freight in the continental United States. This model is capable of predicting freight commodity flow information via trucks to assist transportation planners who wish to understand when and where new road facilities are needed. Such an understanding is important because shipments by truck account for 53% of total tonnage shipped within the US and 72% of total shipments for value (Chin, Hopson & Hwang, 1998). The methods used here are can be generalized to other transportation modalities. When, as was done here, this model is allied with databases of forecast economic and population data, it can be used to forecast future truck freight flows.

This research begins with the use of a traditional gravity model to predict freight flow within the states of the continental United States. Such a model posits that freight volume between any two areas is a direct function of the attraction of each area and inversely proportional to the distance between the two areas. Obviously, the populations of the destination and origin states serve as one possible measure of their demand for, and ability to supply, goods and services. The greater the distance between the destination and origin states, however, the less likely that freight will move between them since shipment costs will be higher.

Population alone, however, has certain limitations as an indicator of the power of a region to draw freight flows from any other area since the purchasing power of the population may be low. In order to increase the model's predictive ability, we included several socio-economic variables. These include each region's total employment, earned income, and total personal income. Each of these is described more fully below.

METHODOLOGY

I. General Description of Data Sources

To accomplish our objectives, we created a database of economic and trucking shipment information. The economic information was obtained from the Bureau of Economic Affairs (BEA) web site (URL http://www.stat-usa.gov/BEN/ebb1). The latter contained information on state population, total employment, total wages paid, and total personal income. Of the data available, the database used to develop the prediction model only included data for

1993 since the commodity flow data existed only for 1993 as of the time the model was developed.

The Excel forecasting system, however, was developed using BEA forecast data on state total personal and earned income, and forecasts of the employment for the years 2000 to 2015. Population projections by county were obtained from the neighboring states' data centers. Creation of county income and employment will be described below. Neither the U. S. Bureau of Economic Affairs nor the U.S. Bureau of the Census provides economic or population projections by county.

The U.S. Department of Transportation's Commodity Flow Survey provides 1993 data on inter-state and intra-state commodity flows by truck. This data consists of weight of shipment, value of shipment, and ton-miles. We used the portion of data that was broken down by state of origin and destination. During the model's development, this was the latest data available.

Next, we describe the data used. These descriptions are pertinent to both the 1993 economic and population data and also the BEA's projections of economic data for New Jersey and its neighboring states for the years 2000 to 2015.

II. Specific Variable Descriptions

A. BUREAU OF ECONOMIC ANALYSIS (BEA)DATA

1. ***Total Employment.*** Employment includes each job that an employed person holds, in any employment setting.

2. ***Population.*** Population is defined by the BEA as the resident population as of July 1 for calendar year 1993.

3. ***Earnings.*** Earnings are defined as the sum of private and government wage and salary disbursements, other labor income, farm proprietors’ and non-farm proprietors' income. This is presented in constant 1987 dollars.

4. ***Total Personal Income.*** Personal income is defined as the sum of all income received by all persons, less personal contributions for social insurance. Personal income is presented in constant 1987 dollars.

B. THE COMMODITY FLOW SURVEY DATA

The 1993 Commodity Flow Survey (CFS) provides data on the movement of goods by mode of transportation. This section summarizes pertinent information on the CFS data.

The CFS provides data for each of the 48 contiguous, continental states. This data was collected from establishments in mining, manufacturing, wholesale trade, and selected retail and service industries. Certain other auxiliary establishments (e.g., warehouses) of multi-unit and retail companies were also covered.

The Bureau of the Census, which actually conducted the survey, generated its sampling frame from the Standard Statistical Establishment List (SSEL) of separate business locations with paid employees. The selected firms were required, for inclusion, to have had a non-zero payroll in at least one quarter of 1991. The total number of firms sampled came to some 250,000. Each contacted company was asked to record information on shipments that they made within a specified two-week period.

We used the information on tons and ton-miles of freight shipped solely by truck between any state of origin and destination.

1. **WEIGHT OF SHIPMENT (OR TONS)** is defined as the total weight of the entire shipment.

2. **TON-MILES**. Ton-miles equal the weight for a shipment multiplied by the mileage that shipment traveled. Mileage was calculated as "the distance between the shipment origin and destination ZIP codes (p. VIII, 1993 Commodity Flow Survey)." The actual distance calculation followed an algorithm developed by the Center for Transportation Analysis.

3. **DISTANCE BETWEEN ORIGIN AND DESTINATION**. The average distance between origin and destination was calculated by dividing the ton-miles variable by the weight of the total shipments that took this route.

Altogether, the database constructed for this study consisted of 2,304 observations (48 states of origin by 48 states of destination).

III. Statistical Testing

Eight potential models were tested using the standard regression techniques in SAS (Statistical Analysis System). Regression analysis provides a systematic method for building equations that summarize the relationships between the variables.

Seven models were subsets of the overall model structure given below. The overall, eighth, model was also run. We normalized the data by taking its natural log in order to ameliorate the effects of skewed data on the regression analysis outcomes.

VARIABLE DEFINITION

Where the dependent or criterion variable is:

Tonnage of freight between the origin state and the destination state.

Where the independent or predictor variables are:

Populations of the origin and destination states;

Distance between origin and destination state;

Personal incomes of the origin and destination states;

Wages of the origin and destination states;

Total employment of the origin and destination states.

Specifically, we analyzed eight different combinations of the variables in order to find the most descriptive model. Each model contained the basic gravity model. The latter consisted of the population of the origin and destination states and the average number of miles that each shipment traveled from state A to state B. First, we tested the predictive power of the gravity model itself. Then we added pairs (or sets) of the economic variables to the basic gravity model. In one extended model, both origin total personal income and destination total personal income were added. In another, total earned wages for the origin and destination states were added. In a third, total employment for the origin and destination states were added. In subsequent model analyses, we added the sets of economic variables two at a time into the same regression model. Finally, we ran a regression consisting of the basic gravity model and *all* three sets of economic variables.

RESULTS

Based on results, we concluded that the extended gravity model, which included total personal income, and total salaries and wages, but not total employment, produced the best model. The best model's characteristics are shown in Table 1.

Practical Application of the Forecasting Model

The research underlying this paper has established a regression-based forecasting model. Having developed what we believe to be the best extant model for predicting freight flow between states, we then sought to develop a useful application of this technique. Specifically, we used the forecasting model developed above to forecast truck freight flow between New Jersey and the other 47 contiguous, continental states, between counties within New Jersey, and between New Jersey counties and non-New Jersey counties within 100 miles of the borders of New Jersey. These forecasts were embedded in an Excel spreadsheet, and were manipulated using pivot tables.

In order to do this, we developed a database of forecasts of population, personal income, wages, and total employment of the 48 states of interest. To do this, we used BEA forecasts of state economic data, and Bureau of the Census forecasts of state population. Projected data existed for the years 2000, 2005, 2010, and 2015. Historical data existed for the earlier years. Using accepted statistical techniques, we interpolated the variable values for the intervening years. Having completed the state database of projected data, we then used the best regression model found above to forecast freight flow for each inter-state linkage. Each forecast was then turned into a percentage of the total inter-state forecast. These percentages were then multiplied by the total interstate freight flow forecasts shown in the American Trucking Association (1999) study in order to get specific interstate freight flow projections. We used the same average distances between each pair of states as was used in developing the original model described above.

Similar procedures were used in constructing the inter-county databases and freight flow forecasts. In this case, the county-level economic data was estimated using the BEA’s forecast of state level economic information multiplied by each county's share of the projected state population, when no state-level forecasts of such data were available. Forecasts of county population were supplied by the data centers of the neighboring states. Forecasts of future state population were taken from the Bureau of the Census. Inter-county distance information was derived as the distance between the zip code of the county seat of the origin county and the destination county.

Based on the foregoing data, we used the best model developed during the first two years of the NCTIP97-21 contract to forecast freight flow a) between New Jersey counties alone, b) between New Jersey counties and counties within contiguous states (i.e., New York, Pennsylvania, and Delaware) if those counties were within 100 miles of New Jersey's borders, and c) to and from New Jersey and the other 47 contiguous states of the continental United States.

PIVOT TABLES AND CHART SYSTEMS

Pivot tables were constructed in each of the five freight forecast files that allow the user to select origin of shipments (whether state or county), destination of shipments (whether state or county), and any of five states of nature. The basic or baseline state of nature reflects the straight application of the prediction model to the forecast economic and population data for the years 2000 to 2015. Four alternate states of nature are presented. One of these four presents forecast freight flow based upon data that assumes that the population and economic data is ten percent higher than the baseline. The second of these four presents forecast freight flow based upon data that assumes that population and economic data is five percent higher than the baseline. The third of these four presents forecast freight flow based upon data that assumes that population and economic data is five percent lower than the baseline. The last of these four presents forecast freight flow based upon data that assumes that population and economic data is ten percent lower than the baseline.

The pivot tables embedded in the five forecast freight flow files (1. To NJ from the Other 47 states 062900 639 PM.xls; 2. From NJ State to Other States.xls; 3. Intra-New Jersey County to County Freight 062800 1134 AM.xls; 4. From NJ Counties to Other States 062900 9-15 pm.xls; and 5. To NJ Counties from Out-of-State 062900 749 PM.xls) are preset to allow the user to select origin county/state; destination county/state; year of forecast between 2000 to 2015; and whether the data used is inflated by 10%, 5%, baseline (0% inflation), or has been deflated by 5% or 10%

from the baseline economic and population data forecasts. While the number of selections within the pivot tables may differ due to the need to choose, say, both a state and a county, the somewhat common interface between the files should enable the user to easily master creation of customized tables and charts that meet the decision needs of the moment.

The user can, for example, customize the pivot table by choosing to see consecutive years forecasts of freight flow for the years 2000 to 2015, assuming no changes from the baseline economic or population data. Similarly, the user may choose to conduct a sensitivity analysis of sorts by looking at, say, a specific year (e.g., 2005), and look at how freight flow will change if the baseline data is presented as is, is increased by 10%, 5% or decreased by 10% or 5%. Other permutations of the data are also possible. For example, the user can look at freight flow from, say, Essex County, NJ, to the Bronx, NY, and within the same pivot table, to Bucks County, PA. This, of course, can be done for the same year (say 2005) or using baseline, inflated or deflated data. The manual accompanying the forecast freight flow files describes how to use the pivot tables and the accompanying charts.

We have also included three Excel files containing the baseline forecast data. Filters were created within these files to enable the user to quickly search for forecast economic and population data for the 48 continental states, or the 21 NJ counties, or counties within states contiguous to New Jersey that are within 100 miles of the borders of New Jersey. The use of these filters is described in the manual that accompanies the files.

The combination of pivot tables and filters will allow the decision-maker to have a good idea as to how inter-county or inter-state freight flow will look in later years. This information, in effect a decision-support system for transportation planners, allows for the efficient production of freight flow forecasts between states in a given period with a given forecast scenario.

THE FORECASTING PROCESS

This study has demonstrated that it is possible to develop a highly reliable model for predicting the flow of freight between any two of the forty-eight contiguous, continental U.S. states. Being able to predict the flow of freight between states and regions is important for businesses, industries, and consumers in the U.S. Such predictions are useful as indicators of where to construct transportation facilities by indicating where truck freight flow is likely to be. To have an effective basis upon which to decide how to expend the required massive amounts of public funds, it is necessary to have an accurate method of forecasting the volume of traffic freight that moves between the states, between counties within New Jersey, and between New Jersey counties and those in neighboring states. The consequences of these expenditures are vital for the economic and social life of many communities. Highway facilities are an important means of promoting economic growth.

BIBLIOGRAPHY

Bonitsis, H. and K. Lawrence. 1996. Traffic Flow Estimation Techniques for Forecasting Interstate Freight Flows. *Proceedings of the North-East Business and Economics Association*. September, 1996, 51-53.

Chin, S., J. Hopson, H. Hwang. 1998. Estimating State-Level Truck Activities in America. *Journal of Transportation Statistics*, 1, Winter 1998, 63-74.

Flowerdew, R. 1982. Fitting the Lognormal Gravity Model to Heteroscedastic Data. *Geographical Analysis*, July, 1982, 14(3), 263-267.

Flowerdew, R. and M. Aitkin. 1982. A Method of Fitting the Gravity Model Based on the Poisson Distribution. *Journal of Regional Science*, 1982, 22(2), 191-202.

Gordon, I. R. 1973. Distance Deterrence and Commodity Values. *Environment and Planning A*, 1973, 10, 889-900.

Moses, L. N. 1955. The Stability of Interregional Trading Patterns and Input-Output Analysis, *The American Economic Review*, XLV(5), 803-826.

Pfister, R. L. 1961. The Terms of Trade as a Tool of Regional Analysis. *Journal of Regional Science*, 1961, 3(2), 57-65.

Spiegelglas, Stephen. 1960. Some Aspects of State

Commodity Flows in the United States. *Journal of Regional Science*, 26(2), 71-80.

United States Department of Commerce Bureau of Economic Affairs. 1995. *Bureau of Economic Affairs Regional Projections To 2045*. 1995, 1. Washington, DC.

U.S. Department of Transportation Bureau of Transportation Statistics. 1996. *1992 Census of Transportation, Communications, and Utilities: 1993 Commodity Flow Survey.* Washington, DC.

U.S. Department of Transportation Bureau of Transportation Statistics. 1996. 1*993 Commodity Flow Survey*. CD-CFS-93-2. Washington, DC.

Weber, J. 1987. Elasticities of Constrained Gravity Models. *Journal of Regional Science*, 1987, 27(4), 621-640.

Weber, J. and A. K. Sen, 1985. On the Sensitivity of Gravity Model Forecasts, *Journal of Regional Science*, 1985, 25(3), 317-330.

TABLE 1: CHARACTERISTICS OF BEST BETWEEN-STATE FREIGHT FLOW PREDICTION MODEL	
DEPENDENT VARIABLE: WEIGHT	**NUMERICAL VALUE OF MODEL STATISTIC**
MODEL STATISTIC	
MODEL F	1385.67
MODEL P	.0001
DF	1908
ADJUSTED R^2	.836
C(P)	8.000

Download Attachment Files (win zip file)