Application of demographic and econometric models to regional transportation planning*

by

C. CHARLES KIMM

Battelle Memorial Institute, Columbus, Ohio, U.S.A.

INTRODUCTION

One of the most pressing problems facing the transportation planner today is the ability to identify, design, and implement a reliable methodology to relate uncertain future socioeconomic conditions to the transportation system requirements needed to support the demand. To complicate the matter further, factors such as resource limitations, environmental degradation, and problems related to interregional integration of widely dispersed people and productive units, which are more recent issues facing the planning, must also be recognized as important factors affecting overall transportation needs.

The United States Transportation System is expected to face new challengers in coming years due to the scarcity of energy resources, increasing costs to construct new facilities and to maintain existing networks, and increasing Federal interest in using the Highway Trust Fund as a source to finance a wider variety of transportation improvement programs. At a time when all these developments are taking place, energy conservation measures are expected to cause a reduction in motor-fuel tax revenues of many states, since the major highway funding source is a gallonage tax on fuel sold to highway users.

In recognition of these problems and uncertain future transportation requirements, Battelle researchers have developed two modeling frameworks as an alternative attempt to project transportation demands, for both passengers and freight, by modes and origin-destination links. The first forecasting model is built around Battelle Demographic and Economic Modeling System (DEMOS) to project passenger travel and vehicle-miles of travel (VMT). The second forecasting model is based on Battelle's regional input/output (I/O) model to project freight transportation. The use of these models in a regional setting will be described in this paper.

DEMOS MODEL

DEMOS is a dynamic simulation model which is composed of various submodels. The model is a simulation as opposed to an optimizing model; it begins with a set of initial conditions, but does not provide optimal values as a result of any simultaneous solutions. It is dynamic as opposed to static: it generates values for many subsequent time periods rather than for a given point in time.

Through the use of DEMOS, subregional level forecasts for population, employment, vehicle ownership and trip generation were obtained. From the subregional level, projections were aggregated by geographic zones. The aggregation was done on the basis of similarity of socioeconomic characteristics such as large urban area influence zones, small urban area influence zones, and rural area influence zones.

The model is composed of three submodels – the demographic submodel, the economic submodel, and the feedback mechanism. Each segment incorporates real world relationships which are structured into a system operating within a simulated time frame. The ability of the variables to change and interact over time gives the model its dynamic nature.

For purpose of inclusion in the model, real world relationships have been identified, expressed in mathematical form, and estimated by means of regression analysis. Although not all relationships which exist in the real world are incorporated in the model, those felt to be most important in terms of their potential effects on the projected variables are included. Figure 1 summarizes the structure of the DEMOS model and indicates the interrelationships among its components.

Demographic Submodel

The demographic submodel is based upon the cohort-survival method of projecting population. In this method the three components of population change – births, deaths, and migration – are summed over time and added to the previous year's population to give a new age specific population for the subregional units.

Since the population changes vary by age group, the three components of population are age specific. The population is separated into 1-year age cohorts and to these cohorts are applied their respective rates. The net natural change in population plus the net migration are then summed to give the new population. This process is reiterated for every year simulated.

Subregion specific births are determined by birth rate equations which are adjusted for current fertility trends. Deaths are calculated by applying age specific mortality rates to the various population cohorts. Age and subregion specific migration is calculated through use of a migration equation in which migration is a function the difference between the local and national unemployment rates.

Changes in birth and migration rates over time are brought into play by means of the feedback mechanism which will be discussed later. These changes over time are changes that occur within the model, a result of its dynamic nature, and not changes related to the alternate assumptions concerning fertility and migration rates which are discussed at a latter point.

* *This paper is based on research work conducted by Battelle Columbus Laboratories for the Commonwealth of Kentucky in 1975.*

Economic Submodel

Employment projections by industry are the result of the economic submodel. The actual employment structure used is confined to those sectors shown in Table 1. As in the demographic submodel, real world relationships are simulated over time to generate employment projections within the various industries in the subregional economy. Basically, this submodel consists of various relationships among the subregional export industries, household serving industries, and business serving industries. Employment in each of these industrial sectors responds and changes for different reasons.

Export Employment

Export employment is defined as that employment serving a demand external to the specific subregion. Generally, export employment is limited to the agriculture, mining, and manufacturing sectors. Since export sectors produce goods and services which are exported to markets outside the subregion, the employment required within a given export industry and the changes in that employment over time depend upon factors beyond the control of the subregion's economy. As a result, the manner in which export industries grow is assumed to be exogenous. Generally, that growth is based on national economic forecasts.

Business Serving Employment

Business serving employment in a subregion is a function of the total employment within the subregion. Since demand for business serving industries depends, by definition, upon other businesses, it follows that as total employment increases, business serving employment increases and vice versa. Given a change in subregion-based export industry employment, there will be a corresponding change in total employment and thus busi-

Table 1 – List of economic sectors considered for each subregional unit in the regional DEMOS model

Agricultural, Forestry, and Fishing
Mining
Construction
Furniture, Lumber, and Wood
Metals Industry
Machinery, Excluding Electrical
Electrical Machinery
Transportation Equipment
Other Durable Goods
Food and Kindred Products
Textile and Textile Products
Printing and Publishing
Chemicals
Other Nondurable Goods
Railroad
Trucking
Other Transportation
Communications
Utilities and Sanitary Service
Wholesale Trade
Food and Dairy Stores
Eating and Drinking Places
General Merchandising
Motor Vehicle Retailing
Other Retail Trade
Finance
Insurance and Real Estate
Business and Repair Services
Private Households
Other Personal Services
Entertainment
Hospitals
Other Health Services
Government Education
Private Education
Other Educational Services
Religions and Nonprofit Organizations
Professional Organizations
Public Administration

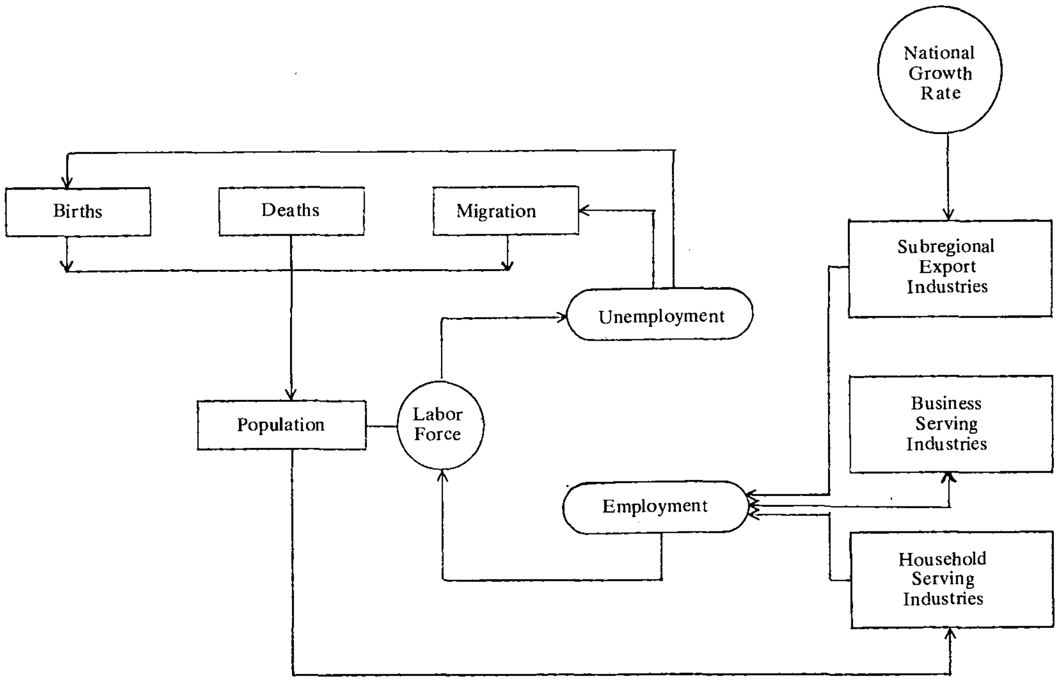

Figure 1 – Regional demographic and economic modeling system

ness serving employment within the subregion. The availability of labor in the subregion also affects the magnitude of the employment change in the business serving sector. For example, as employment in the export sector increases, the unemployment rate decreases making less labor available for any given job. Business within the subregion must now compete for the available labor supply and a specific business may not be able to hire all the labor it would like. This factor is represented within the model, and it dampens the relationship between business serving employment and total employment within the subregion.

Household Serving Employment

Household serving employment is a function of total population within the subregion. Household serving employment provides goods and services to the subregion population; as the population and their demand increase, so does employment in the household serving sector. Examples of household serving industries are hospitals, entertainment, eating and drinking places, and retail trade.

As before, the amount of labor desired by any particular household serving industry is adjusted to take account of the current labor availability in the subregion. Thus, desired change in employment may not be equal to actual change.

Feedback Sector

The feedback sector of the subregional DEMOS model relates the economic and demographic submodels of the system. The basic concept of the feedback sector is the relationship between unemployment and other variables within the model. The unemployment rate within a subregion serves as an indication of the subregion's economic vitality. Migration rates, labor force participation rates, and birth rates are all affected by the degree of unemployment within a given subregion.

Migration Rates

Migration, especially for the working ages, can largely be explained by economic opportunity. People gravitate toward areas having relatively low unemployment rates. This relationship between migration rates and unemployment has been incorporated into the feedback sector of the model. Basically, as the subregion unemployment rate becomes less than the national unemployment rate, there is net in-migration; and as the subregion unemployment rate becomes greater than the national unemployment rate, there is net out-migration. Out-migration reduces the total population in a subregion and thus reduces the household serving sector of the economy. Changes in total employment caused by changes in the household serving sector affect the business serving sector and out-migration may be further increased. A comparable process operates in a reverse manner to affect in-migration.

Birth Rates

Blirht Rates are also related to the level of unemployment within a given subregion. Studies have shown that national birth rates respond positively to business cycles within our economy. Within the regional DEMOS model it is assumed that this relationship also holds true for the subregion. Unemployment rates are used to represent the business cycle; and as the unemployment rate rises, the birth rate falls. Conversely, as the unemployment rate falls, the birth rate increases. One can sense this relationship intuitively by considering that when jobs become scarce, it is generally the young and minorities that are most severely affected. Since these groups are the most fertile groups in our society, and since family planning (including marriage) and personal income (including income derived from employment) are closely related, the relationship between unemployment rates and birth rates becomes apparent.

Labor Force Participation Rates

The labor force participation rate is an important link between the economic and demographic submodels. It is defined as the percentage of the population, age 16 and over, that is working or actively seeking employment. The labor force participation rate is computed for specific age groups within the model since actual rates vary by age group, geographic area, and over time. Past investigations have indicated that labor force participation rates are inversely related to the unemployment rates. As the unemployment rate increases and the number of jobs available becomes smaller, some workers become discouraged and withdraw from the labor force. Conversely, as the economic situation improves and the number of jobs increase, the number of persons seeking employment increases. As labor force participation rates change, they affect unemployment rates and other variables within the model, and once again the model's dynamitics come into play.

Registrations and Trips

In addition to the demographic and economic variables, other variables pertaining to vehicle registrations and trip generations were also projected. The actual variables forecast are:

- Automobile registration by subregion
  - Passenger cars, trucks, farm trucks, trailers, and other
- Trips generated by subregion.

The equations used to forecast vehicle registrations are generally linear while the equation used to generate trips is multiplicative. The actual relationships used are as follows:

a. Passenger cars and truck registrations = f (total employment)
b. Farm truck registrations = f (time)
c. Trailer registrations = f (total population)
d. Other registrations = f (total employment)
e. Trips = f (total registration, total population).

Table 2 – Data inputs used in the regional DEMOS model

- Male population by 5-year age group
- Female population by 5-year age group
- Military workers
- Inmates of institutions
- Male labor force participation rate
- Female labor force participation rate
- Export serving employment
- Household serving employment
- Business serving employment
- Birth rate
- Annual migration rate by 5-year age group
- Labor force – civilian
- Esimated nonresident college students
- Motor vehicle registrations
- Census fertility series
- Export industry employment growth rate

DATA INPUTS TO THE REGIONAL DEMOS MODEL

Table 2 indicates the initial data inputs for the basic model. The DEMOS model is constructed so that it uses published secondary data inputs. These data are obtainable in Bureau of the Census of other generally available publications, and the data are standardized in the sense that they are available for every subregion or administrative unit in the region.

The use of various assumptions, such as different fertility trends or migration rates, allow us to simulate subregion specific conditions. It also, when future conditions are uncertain, allows us to simulate a range of possible conditions so that a range of possible outcomes can be generated.

The socioeconomic forecast generated by the Regional DEMOS are subregional unit specific. For each region's subregional unit DEMOS provides a dynamic simulation of future economic activity and generates forecasts of population, population distribution, migration patterns, employment by 39 economic sectors, labor force participation rates, and other variables for the projected years.

As shown in Figure 2, the outputs from DEMOS were used in a Gravity and Fratar Model to forecast passenger trips and vehicle-miles of travel. Since automobile travel has been and is expected to be the predominant mode of passenger transportation, greater emphasis is given for this mode of transportation. A Gravity model is used to forecast intraregional trips as well as extraregional trips using the region's highway in transit. The Gravity model develops an index of attraction between two points as a function of the population at each point and the distance between them. To operationalize the Gravity model, a vector of trips by zone is first generated. Since the trip behavior of people differ as a function of the area in which they live and work, a separate trip equation for

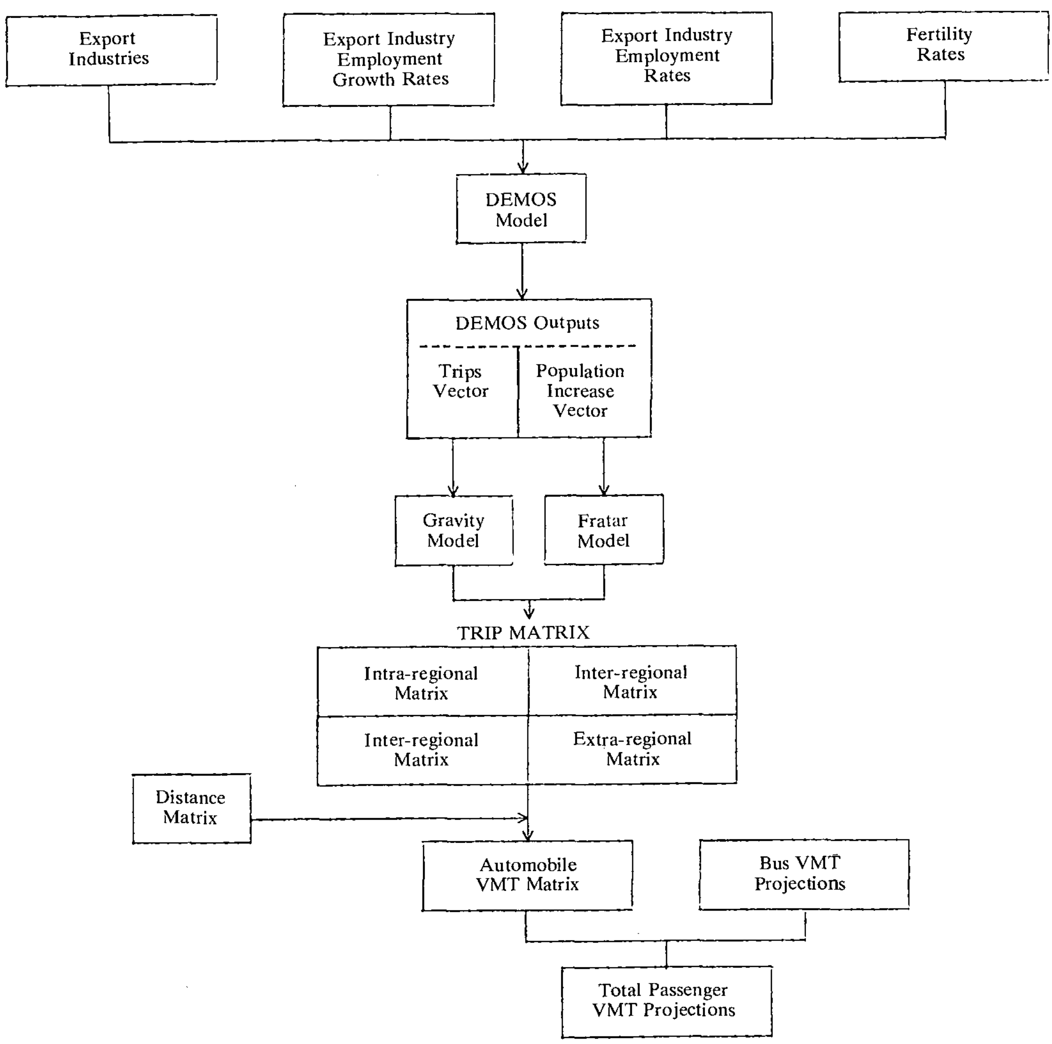

Figure 2 – Methodology to forecast automobile passenger demand

large urban, small urban, and rural influence zones is estimated. Based on the Household Travel Survey, a stepwise linear analysis was performed for a linear and logarithmic equation. In all three areas the linear equation was selected on the basis of the coefficients and determination (R^2) and F statistics. After the trips per zone are generated by using the appropriate equation, the Gravity model was used to assign trips to destinations.

The form of the equation is as follows:

$$T_{ij} = \left(\frac{\frac{P_i P_j}{D_{ij}}}{\sum_{i=1}^{n} \frac{P_i P_j}{D_{ij}}} \right) T_i \quad (i \neq j)$$

$$T_{ij} = \left(\frac{2 \frac{P_i P_j}{D_{ij}}}{\sum_{i=1}^{n} \frac{P_i P_j}{D_{ij}}} \right) T_i \quad (i = j)$$

where

T_{ij} = number of trips per day from i to j
P_i = population of origin zone i
P_j = population of destination zone j
D_{ij} = distance between i and j
T_i = trips originating in zone i.

When i = j (intra-zonal trips) a weighting factor of 2 is used to account for greater interaction within the zone.

The Fratar model is a proportional method of estimating link flows between nodes by distributing trips in proportion to the attractiveness of each zone. "Attractiveness" in this study being represented by the population of each zone. It is an interative method whereby observed zonal interchange was estimated for traffic between each pair of zones. For example, trips between zone i and zone j are estimated by distributing the total flow as originating either from i or from j, in proportion to the respective populations of the zones. The arithmetic average of the two estimates is taken as the estimate from the first iteration, and the computed number of trip-ends for each zone is obtained by summing the estimated zonal transfers to each zone. So then, the ratio of the originally forecasted trip-ends to the computed number of trip-ends represents the growth factor for the second iteration. This process is repeated with new growth factor until the ratio of forecasted to computed trip-ends for all zones are sufficiently close to one. The algebraic formulation of the Fratar model is as follows.

The number of trips between zones i and j computed at zone i is:

$$t_{ij}(i)' = \left(\sum_{j \neq i} t_{ij} \right) F_i' \left[\frac{t_{ij} F_j}{\sum_{j \neq i} t_{ij} F_j} \right].$$

The number of trips between zones i and j is the same basic equation with j substitute for i at appropriate places:

$$t_{ij}(j)' = \left(\sum_{i \neq j} t_{ij} \right) F_j \left[\frac{t_{ij} F_i}{\sum_{i \neq j} t_{ij} F_i} \right].$$

The first approximation to t_{ij}' is the arithmetic mean of $t_{ij}(i)'$ and $t_{ij}(j)$ as follows:

$$t_{ij}' = \frac{t_{ij}(i)' + t_{ij}(j)'}{2}$$

The second approximation proceeds in a similar fashion, with:

$$t_{ij}(i)' = \left(\sum_{i \neq j} t_{ij}' \right) F_i' \left[\frac{t_{ij}' F_j'}{\sum_{i \neq j} t_{ij}' F_j'} \right].$$

And it continues in this manner while the desired accuracy is obtained between two iterations.

Using Gravity and Fratar models, a matrix of 0-D trips was created for the Regional DEMOS model. When this matrix is multiplied by a matrix of highway distances, the VMT matrix, by automobile, was generated.

Bus travel was found not to be linked to demographic or economic factors. A separate forecast of bus VMT was therefore compiled, based on historical growth trends.

INPUT/OUTPUT MODEL

A regional input/output model was developed to serve as a basis for freight transportation forecasting. The input/output model is a representation of region's economy in terms of inter-industry consumption and output. The Battelle model, currently operational at a 127-industry sector detail:

- Can be regionalized into any subregional administrative unit
- Reflects present or future technologies, i.e., is *not* confined, as are most other I/O models, to the technological situation that existed 8 or more years ago
- Produce an I/O table that allows distinctions to be made between those inputs into the region's industries which are produced within the region and those that must be imported; thus, it permits identifying purely regional impacts.

The overall purpose for deriving this particular set of tables was to provide accurate estimates of freight transportation needs by mode required to produce and deliver regions industrial output by sector for each of the projected years.

Preparation of the regional tables involves three basic steps:

1. Deriving the base year "unbalanced" table from the base year Battelle national table
2. "Balancing" the base table, i.e., distinguishing between the share of regional output retained in study region to satisfy total regional requirements (intermediate plus final demand) and the share which must be imported
3. Simulating the effects of estimated growth rates for each of a region's 127 producing sectors over the projected years.

There are several assumptions that lie behind the derivation of the regional input/output table from the national table, the most important of which are:

1. That region's share of national output for any given industry is proportional to its share of national employment in that industry.
2. That the technology employed by each industry in the region is that of the national average for that industry.
3. That regional productivity is the same as the national productivity.
4. That a sector's output is homogeneous in the sense that industry i's output can be used by any forwardly linked industry j.

Because of these four assumptions, the resulting I/O tables must be treated as being normative rather than precisely definitive. Nevertheless, it is felt that this type of table provides a good assessment of the amounts and levels of economic activity taking place within the region.

In actually deriving the tables, the region's share of total national output was obtained for each of the account sectors. These outputs were derived directly from the base year estimates of national total output, assuming that the region's total output has the same relationship to national total output as the relationship of the region's employment to national employment. That is, for each sector classification, the ratio of the employment in region to the total national employment was calculated. This ratio was then multiplied by total output for the national sector to provide the estimate of total regional output in that sector.

After estimating total output, similar estimates were made for final demands originating in the region. The personal consumption component of final demand was assumed to have the same distribution as personal consumption in the national table. The estimates of personal consumption expenditures in the region were obtained by multiplying each entry in the U.S. PCE (Personal Consumption Expenditures) table column by the ratio of personal income in the region to personal income in the nation.

Estimates of regional demand for gross private fixed capital formation were similarly obtained using the regional/national ratio for income arising in contract construction. The assumption implicit here is that all other plant and equipment expenditures are proportional to construction.

Similarly, the ratio of incomes paid by the Federal Government in the region to incomes paid by the Federal Government at the national level was used to estimate the component of final demand attributable to Federal activities in the region. This procedure is repeated to obtain final demand resulting from activities at the regional and subregional levels.

Estimates of gross imports (gross imports are defined here as foreign imports) were calculated in two ways. For about two-thirds of the sectors, gross imports were estimated by using the regional/national ratio used to estimate total output. For the remaining sectors, the ratio between regional total intermediate output and national total intermediate output was applied. The second procedure eliminates the possibility for significant amounts of foreign imports of more primary oriented products to enter a regional economy which has no apparent use for them, e.g., importing iron ore into a region which has no primary iron and steel processing.

The two remaining components of final demand, gross exports and inventory change, were estimated by assuming the regional/national ratio for each sector to be the same one used to estimate the regional total output. In this connection, attention is called to the fact that only the region's share of the nation's exports to other countries is under consideration here.

Intermediate (intra-industry) requirements were calculated by multiplying the national direct technical coefficients by the vector of estimated regional total output for the applicable industry. The result of this process is an *unbalanced* regional table which indicates regional final demand, total output, and intermediate requirements. Total requirements may equal, exceed, or fall short of actual regional output. The residual between regional requirements and output is then entered as interregional imports or interregional exports, depending upon whether the residual is negative or positive.

The next step was to determine the percentage of regional output retained in the region in order to satisfy total regional requirements (intermediate plus final demand). This determination was made in one of two ways, depending upon available data for each industry. The first involved personal knowledge or the use of actual survey information. In this case, people knowledgeable of the region's economy or representatives of the actual manufacturing concerns were questioned to determine the percentage of local output retained within the region.

The second procedure involved the use of a form of location quotient which compared the ratio of regional demand and supply to the ratio of national demand and supply for the same industry.

Once a determination of the region's output retained to satisfy regional demand was known, that amount was separated into the portion going to final demand and the portion going to intermediate demand. As personal knowledge did not dictate otherwise, this separation was consistent with the share between intermediate and final demand requirements of total requirements in the unbalanced tabel.

The second step involved reducing the unbalanced regional table to a balanced regional table. This step was accomplished by assuming that the national direct coefficients were reduced in region in proportion to the ratio of regional output retained to satisfy region's intermediate requirements to total regional intermediate requirements or:

$$r_{ij} = (a_{ij}) \left[\frac{t_{ri}}{T_{ri}}\right]$$

where

r_{ij} = regional direct technical coefficients from industry i to industry j

a_{ij} = national direct technical coefficients from industry i to industry j

t_{ri} = total region's output of industry i retained to satisfy regional intermediate demand

T_{ri} = total regional intermediate demand for industry i's output.

Given the set of regional direct technical coefficients, the regional table was balanced and inverted, and the total direct and indirect impacts resulting from the projected industrial growth rates were calculated.

For example, to simulate the effects of increasing the output of a sector, the regional exports final demand subvector was increased by the amount of the projected increase in output. Additionally, another final demand subvector, personal consumption expenditures, was changed as the increased output of the sector generated additional income which generated additional final demand. The inverted coefficients were then multiplied by the new final demand, and new values of total output, employment, income, etc., were generated. The impact was then calculated as the difference between the two sets of results.

In matrix notation the process is as follows:

$$X = (I - A)^{-1} Y$$

where

X = total output for the baseline conditions –1

(I – A) = inverted table of region's direct technical coefficients

Y = region's final demand for the baseline conditions.

$$X^* = (I - A)^{-1} Y^*$$

where

X* = the vector of total region's output after the output of a particular sector has been increased

Y* = new vector of final demand.

$$I = X^* - X$$

where

I = total direct and indirect impact generated by the direct change in output.

To a large degree, projected national growth rates for each of the 127 account sectors were applied to regional industries. However, where possible, more precise regional sector growth rates have been substituted for the expected national growth rates.

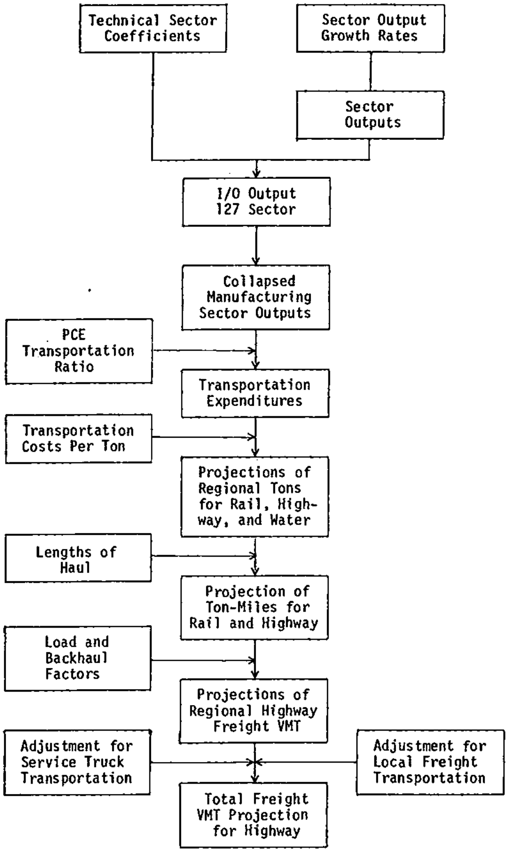

Figure 3 – Methodology to forecast freight transportation demand

As shown in Figure 3, the freight demand analysis for all sectors was based on conversion of the input/output table's sector output (in dollars) to a unit of transportation demand. In order to do this, it was necessary to first get a measure of the regional freight accounted for by each dollar of industrial output. This was done through use of the U.S. Department of Commerce, Bureau of Economic Analysis Personal Consumption Expenditure (PCE) "Poridge tables". These tables indicate transportation and other add-ons to producer prices. To convert transportation costs by sector to traffic volume expressed in tons, rate estimates derived from Interstate Commerce Commission and a series of interviews with the region's shippers representing the major industries were used. This application resulted in volume estimates by rail, truck, and water in the Input/Output model. Because ton-miles are not a meaningful indication of volume of inland waterway transportation, especially if the region's navigable waterways lie on regional boundaries, further freight forecasts in ton-miles were confined to rail and truck transportation. Rail and highway ton-miles were computed by multiplying the projected volumes by estimated average distance traveled within the region. Distance data were derived from the Censuses of Transportation and Manufacturing and the above-mentioned shipper interviews. In the case of truck transportation, it was necessary to go a step beyond ton and ton-miles forecasts. VMT projections in truck transportation were needed to add on to the VMT forecasts of passenger travel. To calculate vehicle-miles, the ton-miles were divided by the average load factor with allowance made for empty backhauls.

The input/output model's 127 industrial sector detail can be collapsed, based on the similarity of physical property of commodities and transportation characteristics, or any sector can be studied at greater depth exogenous to input/output model framework.

CONCLUSIONS

The DEMOS and Input/Output models provide an invaluable opportunity to transportation agencies, at all levels to evaluate the national, regional, and subregional based impacts of alternative growth levels in a wide variety of industries. The capabilities of these models have further been enriched to project freight vehicle-miles of travel and freight transportation demand as a function of complex socioeconomic interaction. In addition, the Gravity and Fratar models provide an added capability to project passenger origin-destination trips and vehicle-miles of travel. The transportation demand projections can be used to assess future system requirements and region's revenues as a source for funding transportation projects.

This comprehensive framework is expected to offer a unique opportunity to nations as an efficient tool in regional transportation planning. It will provide an input to facilitate the planning coordination of all modes in both public and private sectors.