Commodity Flow Modeling

William R. Black
Indiana University

INTRODUCTION

The project reported on here was undertaken with the primary objective of creating a database of commodity flows into and out of the counties of Indiana and to allocate this commodity traffic to the transportation network of the state. Such a database was to form the foundation for a general transportation model for Indiana that would assist state decision-makers in evaluating the various alternatives for public investment in the transportation infrastructure of the state.

The project had two phases. The first phase was completed in 1993 [1]. It was recently described as "one of the most sophisticated modeling exercises done by or for state transportation departments" [2]. This paper covers the second phase of the project [3].

The organization of this paper follows a classical transportation planning process. It begins with an identification of the study area and the existing transport network, or networks in the present case. This is followed by a discussion of the commodities examined and the commodity traffic generation for the counties of Indiana. The models developed for estimating the production and attraction of these commodity flows are also presented and discussed. Distribution of the traffic from origins to destinations is accomplished using a gravity-type of distribution model. This is followed by the use of a modal split procedure, which allocates traffic to specific modes of transport on the basis of historical (1993) patterns of mode use. The final component of this process is an assignment of the distributed traffic to links of the appropriate modal network. This process was undertaken here for nineteen groups of commodities that include the entire range of manufactured goods, coal, farm products, non-metallic minerals, and waste products. There is also some analysis of mail flows, although much of this is at a rudimentary level of analysis due to a lack of data on these flows.

The end result of the above process is a database that can be used to estimate future flows on the Indiana's rail and highway networks. This estimation of future commodity traffic flows was accomplished with traffic projections for 2005 and 2015, however space limitations preclude a discussion of that portion of the study here.

Before proceeding a few words are in order on the geographic information system (GIS) used in the preparation of this report. The system used was TransCAD [4], a GIS system developed primarily for transportation applications. Most of the databases developed here were generated by this system. Although independent estimation procedures, such as multi-variate regression analysis and entropy-based gravity model algorithms, were used in several cases, TransCAD was used for network creation, updating, and assignment of traffic as well as for the maps included here.

This study is in many ways the state or multi-state version of the transportation plans undertaken thirty years ago for the nation's metropolitan areas. It includes 145 "traffic analysis zones", two modes, two networks (70,000 miles of highways and more than 100,000 miles of railway), and the equivalent of 21 trip purposes (commodities here). It is no exaggeration to say that this study was made possible primarily because of the availability of TransCad. It has made the most complex transport and mapping problems manageable.

STUDY AREA AND NETWORK DEFINITION

The primary study area for this research is the state of Indiana and its ninety-two counties. While the flows to, from, and between these counties are of interest, we can not look only at these; the analysis cannot be limited to intrastate flows. A significant amount of the commodity traffic in Indiana has neither an origin nor destination in the state, but instead represents goods or materials that are passing through the state. This overhead traffic may contribute little to the state's economy, but it adds to urban congestion, air pollution, wear and tear on the highways, and rail traffic. Therefore, what happens beyond the state's borders must also be examined here. As a result, this study includes, in addition to the 92 counties of Indiana, several major terminals outside the state. The latter group consists of 48 nodes representing the contiguous 47 states (excluding Indiana, since it is represented as a series of 92 areas) and the District of Columbia, as well as additional nodes for contiguous states: two for Ohio, and one in Illinois, Kentucky, and Michigan, for a total of 145 nodes. This study was concerned primarily with the highway and railway systems. Flows on the other networks are considered implicitly if motor carriers or rail are used in part of the movement.

THE HIGHWAY NETWORK

This study used the interstate highways of the 1992 digital highway network developed for the Federal Highway Administration (FHWA) by the Oak Ridge National Laboratory. Detection of boundary effects during flow assignments led the study staff to use a large circular ring of all digital highways centered on Indianapolis and extending out 200 miles. It was believed that the use of such a ring would eliminate the aforementioned boundary effects.

The Indiana Department of Transportation wanted this project to use a much more detailed state highway network than was available in the FHWA network. The detailed network of interest included the links in the State Roadway Inventory [5]. This database contains a wealth of information on links of the state’s highways. The resulting network consisted of 34,154 links, 31,557 nodes and 70,620 miles of highway. Of these amounts, 15,074 links, 14,330 nodes, and 11,319 miles of highway network were for these elements of Indiana’s State Roadway Inventory.

The Railway Network

The rail network developed by the Federal Railroad Administration (FRA) is based on topographic maps with a scale of 1:2,000,000. The accuracy of such a map is approximately 1,200 meters, about three quarters of a mile, in terms of root-mean square. There are other networks available at a scale of 1:100,000, but these did not contain some of the data necessary for analysis purposes in this study. For this reason the 1:2,000,000 network was used here. It consists of approximately 16,000 links and 11,500 nodes. Among the attribute data available in this network are a link identification code, origin and destination end nodes for each link, owner of the link, abandonment status, and traffic density. It was this latter variable that was critical to later analysis.

COMMODITIES EXAMINED AND DATA SOURCES

It is desirable to have the commodities in a flow study be as detailed as possible. There is a very good reason for this. If the commodity groups are too general it is possible to get nonsensical results during flow modeling. For this study, all two-digit categories of the Standard Transportation Commodity Code (STCC) were examined in terms of their significance to Indiana's economy. A set of 18 commodity groups was identified. An additional group of five commodities was included in what is called “STCC 50" here. These additional manufactured goods include commodities of little economic significance to Indiana (see Table 1).

Also examined here are the flows of mail by the U.S. Postal Service and overnight express mail operations of companies such as FedEx. In these cases the analysis is of movements not covered by the other analyses, e.g, the movement of manuscripts, contracts, magazines, and the like.

Data Sources

The original design of this project anticipated using primarily models derived from the 1977 Census of Transportation to estimate the volume of goods produced or attracted by counties in Indiana and the states of the United States. This design was chosen because the most recent commodity flow data available was from that former census. There was no other census of such multimodal flow traffic available, although a commodity flow survey was underway at the time this project began. The data of that commodity survey was to become available in 1996, but this project was scheduled for completion much earlier than that. In effect, much of what is included here in the way of modeling was done on the 1977 data. This could not be redone with the 1993 data simply because much of this is not available even now. On the other hand all of the discussion of commodity flows into, through or within, and out of Indiana pertains to the 1993 data. In addition to using elements of the 1977 Census of Transportation [6], and portions of the 1993 Commodity Flow Survey [7], this study also made use of the various years of County Business Patterns [8] and the carload waybill sample on CD-ROM [9] in the modeling.

EXISTING FLOWS OF MANUFACTURED GOODS

Based on the 1993 Commodity Flow Survey, Indiana originated commodity flows valued at $178.7 billion. These flows weighed in excess of 285.8 million tons. The major commodities involved in these moves in terms of value were transportation equipment (19.2%), primary metal products (9.8%), food and kindred products (9.5%), electrical machinery (8.9%) and chemicals and allied products (6.4%). The major products in terms of weight were slightly different: petroleum and coal products (21.9%), non-metallic minerals (20.1%), farm products (14.0%), primary metal products (9.8%), stone, clay and glass products (7.7%), food and kindred products (7.4%), and chemicals and allied products (4.2%) (see Table 1).

Destinations of Indiana Shipments

Data are not published at this time on the destinations of Indiana's commodity shipments as such. Data are published on the destinations for "all shipments." For these data the major destinations in terms of value were Michigan, Illinois, Ohio, California, and Kentucky. The major destinations in terms of weight were Illinois, Michigan, Ohio, Kentucky, and Louisiana.
As one might expect, Indiana was the major destination of its own shipments in terms of value or weight. This is typical of most states.

Modal Choices

Mode choices for all shipments originating in Indiana were primarily truck, 77.3% in terms of value, but it is considerably higher than that for some of the manufactured goods examined here. Parcel and express mail account for 7.1% (based on value) of the shipments and these are most likely all manufactured goods. Rail moved only 6.9% of the traffic based on value and 15.2% based on weight. Air freight (excluding parcels) and truck-air moves accounted for only 1.9% of the value and .05% of the weight moved. These numbers are not consistent with other national figures in part due to the nature of the data collected. The figures are based on traffic originating in Indiana. Traffic passing through Indiana (or the U.S.) or traffic originating outside the country are not included here.

A discussion of the 19 sets of commodities and their importance to the Indiana can be found in the final report. For many small communities in the United States the only regular motor carrier traffic is the arrival of a U.S. mail contract motor carrier. These communities may also see the occasional overnight express delivery truck. If the items being sent are economic goods in the sense of a manufactured product or a retail item, these were included implicitly in the 1993 Commodity Flow Survey as such a good shipped by parcel and delivered by any of the aforementioned services or the U.S. mail. There are some exceptions to this statement; these are printed matter. Books, magazines, newspapers, and the like were not included in the survey. U.S. mail flows were estimated at 25.7 pounds per capita and shipments by express mail carriers were estimated at 6.8 pounds per capita, both on an annual basis.

Intermodal Traffic

There is considerable interest today in traffic shipments that are intermodal. For Indiana as a whole intermodal transport is practically insignificant according to the 1993 flow survey. In terms of tonnage, intermodal traffic represented less than a quarter of one percent of the total. This is looking at truck-rail, truck-air, and truck-water moves. If we throw in private truck-for hire trucks as intermodal, the percentage climbs to just under 1.25. Intermodal traffic volume represents about 3.6% of the total traffic originated in the U.S.

If we look at the value of goods the numbers increase as one would expect. In this case intermodal traffic has 3.2% of the total shipment value in Indiana. For the nation this statistic is 4.7% of the total shipment value. Nearly half of these percentage totals come from intermodal truck-air transport.

Summary

The various commodity groups that are included in this study of traffic flows (excluding the mail flows) represent 93.7% of the value of all products shipped from Indiana and 93.9% of the value of all products shipped in the United States according to the 1993 Commodity Flow Survey. In terms of weight the commodity groups represent 98.9% of the originating tons in Indiana and 96.0% of the originating tons nationally based on the same 1993 survey.

The 1993 Commodity Flow Survey is shipper based and (for obvious reasons) only shippers in the U.S. were surveyed. Therefore, we get exports from Indiana, but we get no imports since the shippers of these are in foreign countries. Some crude estimates were made to partially correct this problem.

There are other components of the traffic stream that are not examined here. One of these is household moving vans. Twenty percent of the population in the U.S. moves each year and many of these use established moving companies. Data on these companies were once compiled by the Interstate Commerce Commission; it is not clear that data are compiled on individual moves, but it may be possible to estimate this from the decennial census.

A second component of the commercial vehicle traffic stream not examined here is service transport. Armored trucks moving bank receipts, tow trucks moving disabled vehicles, carpet cleaner vehicles, commercial laundry vehicles, construction vehicles, plumbers, lawn care vehicles, and many more are included in this service transport sector. It is a sector that has all but been ignored by transport planners, yet these are the commercial vehicles that make urban area arterials significantly more congested than rural arterials. If such vehicle moves are addressed at all in urban transportation studies, they are frequently handled by a growth factor (e.g., ten percent additional vehicles to account for trucks) or as a component of traffic counts (e.g., ten percent of the flow is trucks). This data shortcoming is far beyond the scope of this study to correct. It should be addressed first at the national level by a group such as the Transportation Research Board, which could undertake an examination of common practices by state transport planning agencies.

These are relatively minor components for a study of interstate and inter-county commercial transport flows involving Indiana. One can feel confident that the major non-local commercial vehicle flows are included here.

MODELING THE FLOWS

Let us now examine the methodologies used in estimating coal, non-metallic mineral, farm products, and manufacturing commodity flows for the state of Indiana. We will begin with an overview of the research design used here. This will be followed by specific discussions of the procedures used in the commodity production and attraction phases of the traffic generation analysis. The distribution of traffic, identifying the origins and destinations of commodity flow shipments, is discussed next. Once the origins and destinations of flows are known the next concern is estimating the modes taken by these flows; the method of splitting this distributed traffic between modes is discussed next. The final step in the analysis is to assign the traffic to the appropriate transport networks. In this case these are the highway and railroad networks discussed earlier in the paper.

Traffic Generation

As part of this commodity flow study it was necessary to identify the traffic originating and terminating within geographic areas across the United States. The approach used here involved determining the functional relationships that exists between production and attraction of commodity traffic and key variables capable of statistically explaining these flow variables. It is well established that the total flow of a commodity from a given place is statistically related to the total amount of the commodity produced there. Similarly, total flows to an area are related to measures of local markets. The objective then is to model these productions and attractions. In order to model productions and attractions, it is necessary once again to work with the flows of 1977. The flows existing at that time are statistically explained by using the levels of related variables at that time. The models derived can then be used with the level of the explanatory variables for 1993 to yield 1993 commodity productions and attractions. This research utilized multiple regression analysis to develop traffic production and attraction models for each of the nineteen commodity groups examined here.

The quantity of a manufactured commodity exported from an area is a function of the level of production of that commodity within the area or its supply. Unfortunately, commodity production data are also not available. Nevertheless, it has been demonstrated repeatedly that an excellent indicator of a sector's production is employment in the sector. Therefore, a key variable in the traffic production models developed is employment in the sector of interest or related sectors. Some of the commodity may never leave the production area since it is consumed locally. To incorporate this tendency, use is made of a population variable to represent this consumer market in several cases.

Flows of manufactured commodities into an area or the attractiveness of an area is a function of the demand for the product. For most manufactured goods there are two markets: the personal consumer market and the industrial market. With regard to the personal market, it is not meant that the manufacturing firms deal directly with consumers; they will most often go through a retailer or wholesaler. Nevertheless, the magnitude of this market is best reflected by the level of local population. The industrial market is often more difficult to identify. As an example, consider a commodity group such as food and kindred products. This group includes all the processed foods consumed by individuals as well as all the ingredients used in preparing other foods. As a result the level of manufacturing in these further stages of manufacturing also represent a market. Once again, employment is used as an indicator of this industrial market.

Returning to the problem at hand, the 1977 production and attraction levels formed the basis for models of the same based on 1977 population estimates derived by the U.S. Census and employment data derived from the 1977 *County Business Patterns* [10]. Models of non-manufactured goods (coal, non-metallic minerals, farm products, and waste) were not developed in the Phase 1. Models were developed for these sectors here using the 1993 CFS and census data.

The models derived along with an indicator of model accuracy appear as Table 2. While other variables important in explaining the levels of production and attraction will no doubt come to mind, there has been a conscious effort made here to keep the variable base limited and readily accessible. All of the models have used only variables on employment by sector, population, or some economic indicators. Forecasting the variables used into the future may be required and all of these have series available from the aforementioned *County Business Patterns*, from population forecasts, or from other government censuses. On a couple of occasions the variables used are a function of other variables estimated. For example, the level of lumber and wood product flows into an area is a function of the level of traffic production in that sector. Derivation of these models yielded a method of estimating traffic produced and attracted by sector for all states of the United States and counties of Indiana in 1993. Overall the models tend to be accurate based on the adjusted coefficients of determination presented.

The appearance of the 1993 commodity flow survey changed the need to use the models derived for estimating state level productions and attractions to some extent. No data were published on the activities at the county level and as a result the models were used to generate Indiana county level productions and attractions.

The 1993 Commodity Flow Survey

As noted previously a commodity flow survey was undertaken in 1993. It was a survey of approximately 200,000 firms in the United States. It was not expected the data would be available for use in this study and this is part of the reason why the alternative methods noted above were developed. As the progress of the study slowed it became apparent that at least some of the data from the survey might be available before the project was over. The United States summary volume appeared in November of 1996 and some state volumes (including Indiana's) have also been published. In January of 1997 a CD-ROM was released by the Bureau of Transport Statistics of the U.S. Department of Transportation that gave among other things data on the amount of commodities produced (in a traffic generation sense) for the nearly all of the industrial sectors of interest here. After considering the quality of the data being released it was clear that the 1993 data being released was of a much higher quality than the 1977 Census of Transportation data. This was due in large part to the fact that the 1977 commodity flow data was based on a sample of 20,000 shippers and the 1993 data was based on a survey of 200,000 firms. The latter study is also much more aware of the statistical nature of the data collected, e.g., coefficients of variation are presented for most data and "unstable" data (usually based on small numbers in the sample) are not published. It seemed logical to use production and attraction data from the 1993 flows if this was at all possible.

Traffic originations were simple, but traffic attractions presented more of a problem. Recall that all of the state volumes had not been released. It is not clear that this would solve the problem or not. The problem quite simply is that there do not appear to be any figures given for total attractions by state and industry. There are tables on the CD-ROM that yield flows from an origin state to destination states by commodity and this might appear to yield a route to the data of interest, i.e., one could add the flows of each commodity of interest from all states to the destination state and get total attractions. This would be quite possible if all the data appeared on the CD-ROM, but interstate flows are often very scarce and the data are withheld for proprietary reasons. These same proprietary concerns would not enter into consideration if the data included the total traffic attracted by industrial sector. In other words the Bureau of Transport Statistics has the data and could do a special aggregation of the data for state modeling purposes.

There was not sufficient time to pursue negotiations for the release of attraction data by the Bureau of

Transport Statistics. Instead two pieces of aggregate information were used along with the regression models previously noted to estimate the traffic attracted to destinations. In the first case the models were run to estimate the “volume” that would be attracted to each state. The sum of these estimates by commodity was equated with the total attractions by commodity for the country; the latter total attractions is one of the pieces of aggregate information that kept the system in line with actual data.

The second piece of information was the total traffic attracted by commodity for the state of Indiana. This piece of data was used in the same manner as the other national data were used. In effect, the aforementioned models were run to estimate attractions for Indiana counties and the total attraction for the state became the flow limit for allocating commodity traffic to destination areas.

The result of these various operations can be summarized succinctly as follows: the total flows produced by the states are equal to the total flow produced by the nation; the total flows attracted by the states are equal to the total flow attracted by the nation; these same statements also apply to the counties of Indiana and the sum of their productions and attractions are equal to these values for the state. These controls enhance the accuracy of the methods used here.

Traffic Distribution

The distribution of traffic in the Phase 1 report was accomplished by developing several statistical models for traffic generation, as described above, and using these models to identify the production and attraction vectors for input to a fully constrained gravity model. The latter model is also sometimes called an entropy model (see Wilson [10] and [11]). Such a model can, given productions and attractions and an average length of shipment, yield estimates of the flow between all origins and destinations. It does this in such a way that the estimated traffic volume from any origin is equal to the value inputted to the model, the estimated traffic volume to any destination is equal to the value inputted to the model, and, in general the average shipping distance of the estimated flows is equal to the value inputted to the model. In effect, the outcome is constrained to meet all of the initial input parameters of the problem. This tends to yield the most accurate results of any general flow model currently in use. One could alter the output of any model and get more accurate results.

But the basis for such alterations is not clear, nor is it usually defensible.

This model has the general form:

$$S_{jk} = A_j B_k O_j D_k \exp(-\beta c_{jk})$$

where

S_{jk} = the amount of a given commodity shipped from origin area j to destination area k;

O_j = the amount of a given commodity available for shipment at origin j;

D_k = the amount of a given commodity demanded by destination k;

c_{jk} = a measure of the cost or impedance of moving from j to k.

In addition,

$$A_j = [\Sigma\, B_k\, D_k\, exp\,(-\beta\, c_{jk})\,]^{-1}$$

$$B_k = [\Sigma\, A_j\, O_j\, exp\,(-\beta\, c_{jk})\,]^{-1}$$

and

The above formulation is rather straightforward. Some comments are in order on the impedance or cost of movement factor, $c_{j\,k}$. This factor is defined here as the distance between a location *j* and a location *k*. Obviously other functions could be used, but this seemed as good as any [12]. The purpose of the impedance factor is to exercise a negative influence on interaction at increasing distances, other things being equal. Some researchers believe that this function should be actual road distance between places. If it were this would certainly complicate subsequent analysis which in part tries to identify the need for additional roads.

The model should be viewed as capable of replicating major shipping patterns in the area of interest; this is the United States and Indiana in the present instance. In effect, if certain aspects of the flow are known this model should yield what could be called the most probable set of flows given this information. This is sometimes called the most probable macro-state. This yields the major patterns of flow, which should be the major concern of the analysis.

The project also used actual data for Indiana to refine the modeled estimates for the counties of the state, i.e., the sum of the county estimates for each category was set equal to the state total. As a result the following statements can be made with regard to commodity flows examined in this study:

1. Total flows from all states as used by the gravity model are equal to actual traffic productions by manufacturing category for those states.

2. Total flows from Indiana and total flows to Indiana, by commodity, as generated by the model are equal to the actual flows as given in the commodity census.

3. The sum of the total flows as generated by the states for productions and attractions are equal to national totals for these.

The primary item being modeled here is tonnage of commodities shipped between origins and destinations in the United States for the year of 1993. As a result these flows are annual tonnages. In order to have something to compare average daily traffic with, it was necessary to reduce the flows to a daily basis. An examination of the *Highway Capacity Manual* [13] revealed that truck traffic generation was approximately equal from Monday through Friday, but on the week end the truck traffic generation rate appears to be about 44%. Adding five days and two days at .44 each yields 5.88 days per week or 306 days per year of trucking. Therefore, dividing the total annual flow by 306 yields a good estimate of commodity truck trips per weekday. Multiplying this weekday rate by .44 gives a reasonable estimate of the traffic on a Saturday or Sunday.

Modal Split

Once traffic is distributed between origins and destinations there remains a question of the modes selected for the movement of that traffic. The modal split computer model written for this purpose (NEWMODE) utilized 1993 data for nine single modes and eight multiple mode categories. While additional detail is always desirable it sometimes creates problems as well. The primary areas of interest in this document are highway and rail traffic. It would simplify the project if only truck and rail were given, but this would ignore several modal combinations that seemed to be rather common, e.g., truck and rail or truck and air. So that this traffic would not be lost all seventeen categories examined here has a set of distances (less than 50 miles, 50 to 99 miles, 100 to 249 miles, 250 to 499 miles, 500 to 749 miles, 750 to 999 miles, 1000 to 1499, 1500 to 1999, and 2000 or more miles) and for each distance group there are seventeen modal categories, plus some summary classes such as highway, or rail.

Commodity Density

To obtain density factors the tonnages of the 19 commodities coming into Indiana, and leaving Indiana by rail according to the expanded Waybill Sample were aggregated by commodity and rail carloads. Division of the former by the latter yields tons, by commodity, per carload, or commodity density. As one might expect these density factors differ based on whether they are in inbound or outbound and this may reflect the commodity’s stage in the manufacturing process. A weighted average of these inbound and outbound density factors was calculated for rail cars and these appear as Table 3.

The density factors for motor carrier traffic assume that a rail car can handle 100 tons and a truck can carry 40 tons or 80,000 pounds, or 40% of the same product. In other words it was assumed that the motor carriers could handle 40% of the density factor of a rail car by commodity, and this is the source of the motor carrier density factor in the table. One could argue that some states permit vehicles with weights in excess of 80,000 pounds, but there is also a considerable amount of highway traffic moving in vehicles smaller than this. As a result, assuming all motor carriers are 40 ton vehicles seems a reasonable standard for use here.

Mail arriving in your local community usually arrives by a contact mail carrier operating a tractor-trailer or semi. Included in the trailer are bags of mail that can weigh no more than 70 pounds. Most are filled to this limit. A typical trailer can hold an estimated 450 bags or 31,500 pounds of mail. This translates into 15.75 tons per motor carrier.

Some counties may receive considerably more trucks than estimated here. That does not mean that our estimates are off. We were looking only at non-commercial mail (letters, manuscripts, photographs from your children or parents, contracts and the like). Other mail that may include commercial materials, e.g., a shirt from a mail order house, or some fresh fruit from a popular West Coast mail order firm, are included as parcel moves in the commodity flows examined elsewhere in this report.

Traffic Assignment

Methods of assigning traffic to a network are numerous. The simplest of these methods is referred to as “all or nothing” assignment. In this procedure the methodology assigns traffic moving between some area *j* and some area *k* to the shortest path (route) between this origin and destination. All possible pairs of origins and destinations have their traffic assigned in exactly the same manner. There is no consideration given to the capacity of links in the paths selected or whether travel time on the links will be affected by congestion.

“All or nothing” traffic assignment has more than a few critics against its use in urban transportation planning, however the focus here is on regional or statewide transportation planning and most of the criticisms seem inappropriate in the regional context. For example, nearly everyone in Indianapolis that found they suddenly had to drive to Chicago would take Interstate 65 simply because of the shortest travel time of this route. In an urban context there might be some question about your route in getting to this interstate highway, and probably other assignment procedures would be appropriate for this portion of the trip. This does not matter in the present study because the flows being examined are inter-county and interstate trips.

The Highway Network and Cost of Movement

Traffic assignment techniques requires the construction of a network over which movement can take place. This network connects all origins to all destinations and includes the “cost” of movement over the links and in some cases the capacity of the links to hold traffic. Cost may be a misleading term because the measure used is rarely in dollars and cents. Instead studies over the years have used distance, travel time, or traffic flow functions related to distance or travel time. This project used travel time as its initial measure of travel cost. For large scale studies over an area the size of the United States travel time is rarely known. Instead it is approximated by the following:

Install Equation Editor and double-
click here to view equation.

Here the length is in miles and the speed is in miles per hour. This results in travel time being measured in hours or parts of hours.

One final point on the network is that whenever the cost of travel, or the way in which it is being measured, changes, or whenever the links in the network change for some reason, it is necessary to generate a new network for assignment purposes since the network expects to move flows between centroids, the network nodes must be consistent with the defined network.

Target Flows

The flows used by the traffic assignment procedures were the flows for all goods examined in this study. In other words, it included shipping the total highway tonnage of all 15 manufacturing groups and the four resource based commodities and two types of mail included in this study. This represented the sum of the 21 gravity model distributions by O-D pair. The reason for evaluating the traffic assignment routine using total flow was that this was the only variable that came close to existing data on actual flows. Existing traffic count data are actually expanded numbers of commercial vehicles per day on Indiana’s highways over the period from 1991 to 1994. This became the target to which the traffic assignment had to demonstrate a relationship.

It goes without saying that commercial vehicle count data is not the best measure of manufactured or primary commodity traffic on the highways. After all commercial traffic includes the movement of delivery and large service vehicles, as well as empty trucks. We have not examined the empty tractor trailer combinations here. The major implication of this is that the target flows were larger than the flows to be assigned. In addition, this study looked at inter-county flows, not intra-county flows, and as a result the target flows were generally larger within an origin or destination county.

The Highway Assignments

Initial applications of the ‘all or nothing’ traffic assignment routine to the travel time network resulted in too many vehicles being assigned to the Interstate links and virtually nothing being assigned to other links. This was not that surprising, but it clearly required some adjustments be made.

Redefining Highway Travel Cost

To handle some of the other problems a new mathematical function was defined for speed. It was:

$$New\ Speed = Old\ Speed + (2 * \sqrt{(65 - Old\ Spe}$$

This results in the following changes in speeds:

Old speed	New Speed
65.00	65.00
60.00	64.47
55.00	61.32
50.00	57.75
45.00	53.94
40.00	50.00
35.00	45.95

and travel time was redefined as

$$Travel\ time = Length/NewSpeed$$

This new travel time function had the effect of making 55 or 60 mph routes that were slightly shorter than an Interstate link competitive with the latter. Therefore in cases where loading points were near two highways matching the above situation, traffic would be assigned in part to each route. A new travel time network was created and this was used in another run of the ‘all or nothing’ traffic assignment routine.

Assignment Results

This adjusted travel time network resulted in a more reasonable allocation of traffic across the highways of Indiana suggesting that perhaps individuals do not see a great deal of difference between these higher speeds, which is exactly what the new traffic function implies.

Evaluation of traffic assignment results is not an easy matter. A researcher's first impulse is to simply undertake a correlation and regression analysis of the statistical relationship between the assigned flows and the target flows from the total commercial vehicle road counts. In general this is not done in transportation planning studies at the scale of this study. Instead planners look at the distribution of trips and retain the assignment if it is “close” to the observed distribution. Nevertheless, a statistical analysis was undertaken here.

A sample of 40 locations in rural areas of the state, but including all types of highways, was drawn. The overall model "explains" about 48% of the variation in total commercial traffic using the flows assigned here to the 40 rural locations. This is a very significant result. Nevertheless, a higher level of explained variation would have been desirable for the overall relationship examined here.
A further evaluation of the traffic assignments was undertaken using the assignments of all goods (not just manufactured goods) to the Indiana highway network. The overall relationship dropped slightly yielding an adjusted coefficient of determination of .435. This gives an F statistic of 31.07, which is significant at the .001 level.

The Railway Network and the Cost of Movement

Although highway traffic assignment is controlled by travel time and the user’s desire to minimize this, or a cost version of the travel time, railway operations are not so preoccupied with this. It is certainly true that railroads want to move from origin to destination quickly, but speed is usually measured in days for railroads as compared to hours for highways. However, the problem is actually more complex than this.

Although there is some desire on the part of rail carriers to minimize the length of haul, they have a tendency to use mainline trackage even though secondary lines may be more direct. The question was how to represent this tendency with the rail data available for the digital network. Track condition plays a part in such decisions, but this is a very dynamic variable that would change more frequently than the database available. It seemed a new measure of spatial separation was necessary. The new measure of spatial separation would still incorporate an attempt to minimize shipping distance, but it would also pick those routes that the railroads tend to use. Short line or regional railroads that originate or terminate traffic are not important in this methodology, since the origin and destination of shipments must be reached. In other words these moves can be replicated by any methodology regardless of the cost attached to it simply because the end nodes of these moves are used as input to the methods.

The measure finally adopted had the form

$$I = (L\ (1/(D+1)))$$

where

I = the index of spatial separation;

L = the length of the line segment of the network; and,

D = the traffic density of the line in millions of gross ton-miles per year.

This measure diminishes the length of line segments by dividing the segment by its traffic density, i.e., by gross ton-miles per year. Typical traffic density values vary from 0 to about six million gross ton-miles per mile of line. When used on lines with high traffic density these routes “become shorter” and are always selected. Lines of low traffic density, do not become “longer” since their traffic density always has a unit value added to it. Lines of 0 traffic density would become lines of 0 length, if it were not for this correction factor.

Target Flows

In the railway case there are no target values that are route segment specific. Data that are made available in the public use carload waybill sample are too gross to be used for this purpose. Very detailed information that would allow comparisons are available in a complicated fashion for flows involving the movement to, from, or through the state of Indiana, but translating compiled data into this format is difficult. As a result one must visually examine the flows to see if they are consistent with expectations.

Flows Assigned

The commodities assigned to the rail network are a product of this project and programs developed by it. More specifically, the traffic assigned is the product of NEWMODE, the computer program that splits commodity specific traffic between modes based on the length of haul. The basis for these splits is data published for the United States in the 1993 Commodity Flow Survey. In the case of some manufactured goods this is a very small amount of traffic since railroads during the latter part of the 20th century have loss significant market share for many manufactured goods to motor carriers. The total rail flows assigned are consistent with expectations.

Sources of Errors in the Assignments

It is reasonable to offer some explanation for the errors observed in the assignment process, but in order to do this it is also necessary to examine possible sources of error throughout the planning and analysis effort undertaken here. Some of these reasons have been stated previously, but they are restated here simply as a catalog of items that should be considered in future applications.

1. The network used in this study is an abbreviated representation of the state highway network. In other words there are places in Indiana where a substantial amount of travel occurs on local or county roads.

2. A second network related problem resulted in some highways of the state simply ending at the state border.

3. Still another source of error related to the network is the placement of external nodes

4. A related point is that beyond the circular highway network that surrounds Indiana all of the highways are part of the Interstate Highway System.

5. The models used for Indiana traffic generation are not perfect, and this could also result in some error in the estimates of production and attraction.

6. Manufactured goods shipped to the U.S. from a foreign country are missing in the data used here, to some extent.

7. The distribution modeling did not perfectly replicate the average shipping distance. The major

reason for this is the average shipping distances between states in the western U.S. It is not possible to constrain flows if distances between places are very large, relative to observed averages. Nevertheless, this does not appear to be a major source of error here.

The use of “all or nothing” assignment procedures is probably the best choice for a study of this type. However, it does seem that highway traffic flows at the scale examined here are in need of a new traffic assignment method. This method would look at the three or four best (e.g., lowest cost) routes that could be taken between an origin and a destination and assign probabilities to these. Trucks would be assigned in a Monte Carlo fashion with assignments proportional to their probabilities. For example, truckers passing through Chicago en route to Pittsburgh may consider Interstates 80, 65-70, or even U.S. 30. The proposed method would consider all of these as possible choices. This would overcome some of the shortcomings of the single least cost route of "all or nothing" assignments being selected all the time.

TABLE 1 Value and Tonnage of Major Commodity Groups Originating in Indiana

STCC Code	Value (millions)	Tons (thousands)	Commodity Group
01	$ 5,794	39,902	Farm Products
11	281	10,759	Coal
14	463	57,341	Non-metallic Minerals
20	16,958	21,039	Food and Kindred Products
22	275	93	Basic Textiles
23	7,795	553	Apparel
24	3,235	4,131	Lumber and Woods Products
25	3,120	734	Furniture and Fixtures
26	3,194	2,814	Pulp and Paper Products
28	11,474	11,957	Chemicals and Allied Products
29	9,008	62,500	Petroleum and Coal Products
32	2,748	21,972	Stone, Clay and Glass Products
33	17,485	27,881	Primary Metal Products
34	10,363	4,572	Fabricated Metal Products
35	9,504	1,023	Machinery, except Electrical

36	15,914	1,909	Electrical Machinery
37	34,401	6,731	Transportation Equipment
40	703	4,474	Waste and Scrap Material
50*	14,811	2,421	Other Manufactured Products

* Category 50 here includes STCC 21 (Tobacco Products), STCC 27 (Printed Matter), STCC 30 (Rubber and Miscellaneous Plastic Products), STCC 31 (Leather and Leather Products), STCC 38 (Instruments, including Medical and Photographic, as well as Watches and Clocks), and STCC 39 (Miscellaneous Products of Manufacturing). However, no data are included here for STCC 27 due to sampling and definitional problems regarding shipments in the *1993 Commodity Flow Survey*.

TABLE 2 Models of Production and Attraction

Model Number	Model	Adjusted R^2
(1)	PROD01 = 1445 -.523 AGSER + .0048 CASH	.562
(2)	ATTR01 = .819 PROD01	.660
(3)	PROD11 = 7.6 COAL	.650
(4)	ATTR11 = 3.1 COAL + 5.3 MIN	.657
(5)	PROD14 = .078 MAN	.658
(6)	ATTR14 = .997 PROD14	.977
(7)	PROD20 = .282 FOOD	.940
(8)	ATTR20 = .832 POP + .162 FOOD	.965
(9)	PROD22 = .016 TEX	.931
(10)	ATTR22 = .003 APP + .0001 ALL	.743
(11)	PROD23 = .004 APP	.919
(12)	ATTR23 = .002 APP + .011 POP	.926
(13)	PROD24 = .668 LUM	.808
(14)	ATTR24 = .728 PROD24	.805
(15)	PROD25 = .017 FURN	.906

(16)	ATTR25 = .033 POP + .002 FURN	.960
(17)	PROD26 = .103 PULP + .056 LUM	.886
(18)	ATTR26 = .085 PULP + .259 POP	.953
(19)	PROD28 = .150 CHEM + 1.164 PET	.758
(20)	ATTR28 = .077 CHEM + .455 PET +.683 POP	.851
(21)	PROD29 = 6.857 PET	.945
(22)	ATTR29 = 4.007 PET + 1.881 POP	.938
(23)	PROD32 = 2.882 POP	.851
(24)	ATTR32 = 2.914 POP	.871
(25)	PROD33 = .085 MET	.982
(26)	ATTR33 = .093 MET + .061 FAB	.923
(27)	PROD34 = .013 MET + .034 FAB	.927
(28)	ATTR34 = .035 FAB	.861
(29)	PROD35 = .013 MAC	.883
(30)	ATTR35 = .010 MAC	.878
(31)	PROD36 = .004 MET + .004 FAB + .003 ELEC	.826
(32)	ATTR36 = .005 FAB + .034 POP	.915
(33)	PROD37 = .040 TRAN	.753
(34)	ATTR37 = .027 TRAN	.837
(35)	PROD40 = .00048 POP	.704
(36)	ATTR40 = .0067 MAN	.791
(37)	PROD50 = 1.097 ATTR50	.858
(38)	ATTR50 = .245 POP	.857

Notes: Most of the explanatory variables above are employment in specific STCC (SIC) industrial classes according to County Business Patterns. Exceptions are the PROD and ATTR variables which represent tons of product shipped or received by STCC, e.g., PROD01 is the tons of farm products shipped and ATTR01 is the tons of farm products received. Other variables are defined as follows: AGSER = employment in SIC 07; ALL = total employment; APP = employment in SIC 23; CASH = gross cash receipts (in $1,000s) from farming; CHEM = employment in SIC 28; COAL = employment in SIC 11; ELEC = employment in SIC 36; FAB = employment in SIC 34; FOOD = employment in SIC 20; FURN = employment in SIC 25; LUM = employment in SIC 24; MAC = employment in SIC 35; MAN = total employment in Manufacturing, SIC 2 and SIC 3; MET = employment in SIC 33; MIN = employment in SIC 14; PET = employment in SIC 29; POP = total population; PULP = employment in SIC 26; TEX = employment in SIC 22; TRAN = employment in SIC 37.

TABLE 3 Traffic Density Factors for Rail Cars and Motor Carriers by Commodity

Commodity STCC	Import rail traffic	Export rail traffic	Weighted rail density (tons)	Weighted truck density (tons)
01	94.90	96.20	96.13	38.44
11	100.60	99.10	100.42	40.17
14	97.10	97.40	97.20	38.88
20	77.35	80.36	79.52	31.81
22	25.00	15.00	18.33	7.33
23	-------	-------	*10.00	*4.00
24	73.88	55.50	72.27	28.91
25	-------	15.00	15.00	6.00
26	64.82	50.64	62.10	24.84
28	85.11	90.11	87.58	35.03
29	63.20	77.16	65.90	26.36
32	86.70	77.10	81.15	32.46
33	87.48	85.21	85.82	34.33
34	28.40	16.16	19.76	7.90
35	68.75	21.70	28.42	11.37
36	18.80	16.25	16.69	6.68
37	19.93	23.40	22.50	9.00
40	75.40	82.60	78.47	31.39
**50	92.85	14.88	86.56	34.62

* Estimated values

** There is no STCC 50. It is used here to represent STCC 21, 27, 30, 31, 38 and 39.

REFERENCES

1. Black, W.R. and Palmer, J.A. (1993) *Transport Flows in the State of Indiana: Commodity Database Development and Traffic Assignment, Phase 1.* Bloomington, IN: Transportation Research Center, Indiana University.

2. Transportation Research Board (1997) *A Guidebook for Forecasting Freight Transportation Demand, NCHRP Report 388*, Washington, D.C.: National Academy Press.

3. Black, W.R. (1997) *Transport Flows in the State of Indiana: Commodity Database Development and Traffic Assignment, Phase 2.* Bloomington, IN: Transportation Research Center, Indiana University.

4. Caliper Corporation (1996) *TransCAD: Transportation GIS Software*, User's Guide, Version 3.0 for Windows, Newton, Mass.: Caliper Corporation.

5. Indiana Department of Transportation, *INDOT Road Inventory Computer Print-Out Supplement, A Users Guide*, Indianapolis: Indiana Department of Transportation, Roadway Inventory Section, Division of Program Development, undated.

6. Bureau of the Census (1982), *1977 Census of Transportation*, Washington, DC: U.S. Department of Commerce.

7. Bureau of the Census (1996), *1992 Census of Transportation, Communications, and Utilities, "1993 Commodity Flow Survey,"* Washington, DC: U.S. Department of Commerce.

8. Bureau of the Census (1979), *County Business Patterns*, Washington, DC: U.S. Department of Commerce.

9. Bureau of Transport Statistics, *Railroad Carload Waybill Sample*, various years, Washington, DC: U.S. Department of Transportation, CD-ROM.

10. Wilson, A.G. (1967), A Statistical Theory of Spatial Distribution Models. *Transportation Research 1 (3)*, pp. 252-270.

11. Wilson, A.G. (1970), *Entropy in Urban and Regional Modelling*, London: Pion.

12. Black, W.R. and Larson, R.C. (1972), A Comparative Evaluation of Alternate Friction Factors in the Gravity Model. *The Professional Geographer 24 (4)*, pp. 335-337.

13. Transportation Research Board (1992), *Highway Capacity Manual, 2nd Edition, Revised, Special Report 209*, Washington, D.C.: National Research Council.