Making Do with Less: Calibrating a True Travel Demand Model Without Traditional Survey Data

Steve Ruegg, Parsons Transportation Group

Abstract

For many small and medium-sized cities, funding a full Home-Interview survey, with costs as high as $100 per household, is not feasible. Traditionally, such a survey has provided the basic foundation for developing a truly useful travel demand model. Without it, model development has been handicapped by the lack of such behavioral, disaggregate data. This problem was faced recently in Jamestown, North Dakota, for which it was necessary to develop a travel demand model for small (population 16,600) city, with no recent origin- destination data.

The technique relied on the fact that a sufficiently robust set of simple traffic counts contain a great deal of travel behavior information implicitly. Using a set of consistent, concurrently-taken counts, external travel behavior from an older study, and a detailed zone system, a technique has been developed which can produce a fully-specified, classical travel demand model. Not only can this model be calibrated quite closely to existing counts, but it can be used as a forecast tool, requiring only socioeconomic and network data. In other words, the traffic count data was tied not only to a current origin-destination trip table, but to distribution parameters, time of day parameters, and trip generation rates at the zone level. The procedure involves an extended application of the "O-D from traffic count" technique (which is implemented through a macro in EMME/2), essentially working the four-step modeling process in reverse. Through iterations of the model, proper generation and distribution parameters can be developed, which, when finished, reliably reflect actual conditions.

This technique can be a very cost-effective way for small and medium-sized cities to obtain a travel demand modeling capability which is not simply a product of "borrowed" parameters from another area, but is indeed calibrated to local, observed conditions.

The transportation planning needs of small and medium-sized cities have become more sophisticated, as they attempt to address the needs of the community. The difficulty has been increased due to more stringent planning requirements imposed by the 1991 ISTEA (Intermodal Surface Transportation Efficiency Act) and by the lack of sufficient funding, which forces difficult choices, and requires clear justification of future plans. These demands often require transportation plans to be developed in a more analytical way, and a key element to these plans is often a travel demand forecasting model. While major metropolitan areas almost always have some type of model, smaller areas may have only old, or special-purpose models, or none at all. For these communities, the choices are:

- Proceed without a model, or use the existing model, which may not be adequate
- Develop a new travel demand model

The first choice, for reasons mentioned above, may lead to an inadequate or even misleading plan. The second option, however, can be expensive.

A truly useful model has traditionally required data collected from a comprehensive origin-destination survey, or set of surveys. Home interview surveys, which are typically used to collect this data, are expensive, ranging from $50 to $100 per household. For even a 1 percent sample, this can quickly become a major expense for even a small city. Keep in mind that, even for a small city, a minimum sample size is probably at least 1200-1600 households. Additional cost will be incurred processing the survey results, including geocoding, survey record cleaning, trip chaining, factoring and creating a calibration data base.

In addition to the cost, gaining approval from politically-appointed councils for survey funding is often difficult, since the task is several steps removed from showing direct benefits to the community.

For these reasons, travel demand forecast models have often been developed without contemporary origin- destination data, leaving the resulting models insufficient for their analytical tasks, or at least open to criticism and attacks on credibility of results.

To address this problem, a technique has been developed which seeks to strike a compromise between depending upon expensive and difficult to obtain O-D data, and model development without reference to local travel behavior data. This paper describes this technique, and the results of an application in Jamestown, ND.

Background

Barton-Aschman Associates was retained by the City of Jamestown, North Dakota to develop a Land Use and Transportation Plan for the City and surrounding area. The previous plan had been developed in 1970, and the lack of an updated plan was hindering the project development process for any major transportation infrastructure improvement. State and Federal funding for such improvements required an up-to-date plan for the area. As a part of this plan, a travel demand forecast model was prepared, and submitted to the North Dakota Department of Transportation planning office for review and approval. No previous travel demand model had been developed.

Jamestown, North Dakota is a community of about 16,600 persons, with an additional 5,600 persons in the surrounding county (Stutsman County). The city lies just off of I-94, about halfway between Bismarck on the west, and Fargo on the east, approximately 95 miles from each city. North/South access is provided by US281. The economy is agriculturally-based, with some light industry, bulk food and cattle processing. Jamestown is the home of a buffalo museum, State Hospital, a school for children with physical disabilities, and a small 4-year liberal arts college--Jamestown College. Bisecting the community is the main line of the Burlington Northern Railway, which primarily hauls coal from mines to the west, to power plants in Minnesota and the port of Duluth/Superior. It is a heavily used line (about 27 trains/day), which passes through the heart of the downtown. One grade separated crossing exists, along with 6 at-grade intersections involving local streets. Major transportation issues center on reducing train/vehicle interaction, and reducing the North/South truck traffic through the downtown, which currently has no other route option.

Over the past 10 years, Jamestown has remained static in growth. For the purposes of the plan, however, significant growth was assumed, to a population of 20,600. In recent years, employment has increased, without causing a significant increase in population. Future growth in employment, it is believed, will finally begin to expand the population and household base.

In order to adequately address the transportation impacts of anticipated growth, and of proposed major transportation improvements, including a north/south bypass, a travel demand model was necessary. In addition, the North Dakota Department of Transportation required a model-based plan. However, as mentioned above, no previous model existed. In addition, only a 1973 external travel survey was available, providing only the most general information on trip origins and destinations. Resources simply did not exist for conducting a home interview survey. However, some information was available, including

- A very comprehensive, contemporary and consistently-obtained set of traffic counts for all major areas of the city.
- A very complete, and comprehensive inventory of households, and employment, with locational information. The households were classified by density, and employment was divided into retail and non-retail.
- Summary information regarding work trips for the county and city from the 1990 Census Transportation Planning Package (CTPP).
- Good information regarding roadway widths, speed limits, signal locations, and general roadway network information within the City.

The availability of these data, and the quality of the data, supported the approach of creating a relatively simple, traditional travel demand model for the City. The model would have the advantage of containing parameters based on local travel behavior, rather than relying upon "borrowed" model coefficients, as is often done. The model could be calibrated not just in the sense that it replicated traffic counts, but to some degree also reflected community trip generation, distribution and trip length characteristics.

The key to this was the excellent count information available. The North Dakota Department of Transportation (NDDOT) conducted a set of tube counts in the fall of 1994 at over 150 locations in the city. These were then seasonally adjusted, and published on a map. All counts were taken within a few weeks of one another. Peak hour turning movement counts were also taken at selected intersections. A set of counts such as this may be thought of as a "snapshot" of the community's vehicular travel behavior. While any individual count shows only volume on a street at one location, taken together, the counts are indicative of all travellers' trip-making behaviors. Implicit in the data set is the trip generation, and distribution characteristics needed for a traditional travel demand model. The information on the socioeconomic data (households and employment) was also available in detail. Combined with the other data sources, we began the model development process knowing much about the beginning and end of the model, and something about how to get between them in terms of traveller behavior--all with locally-based data.

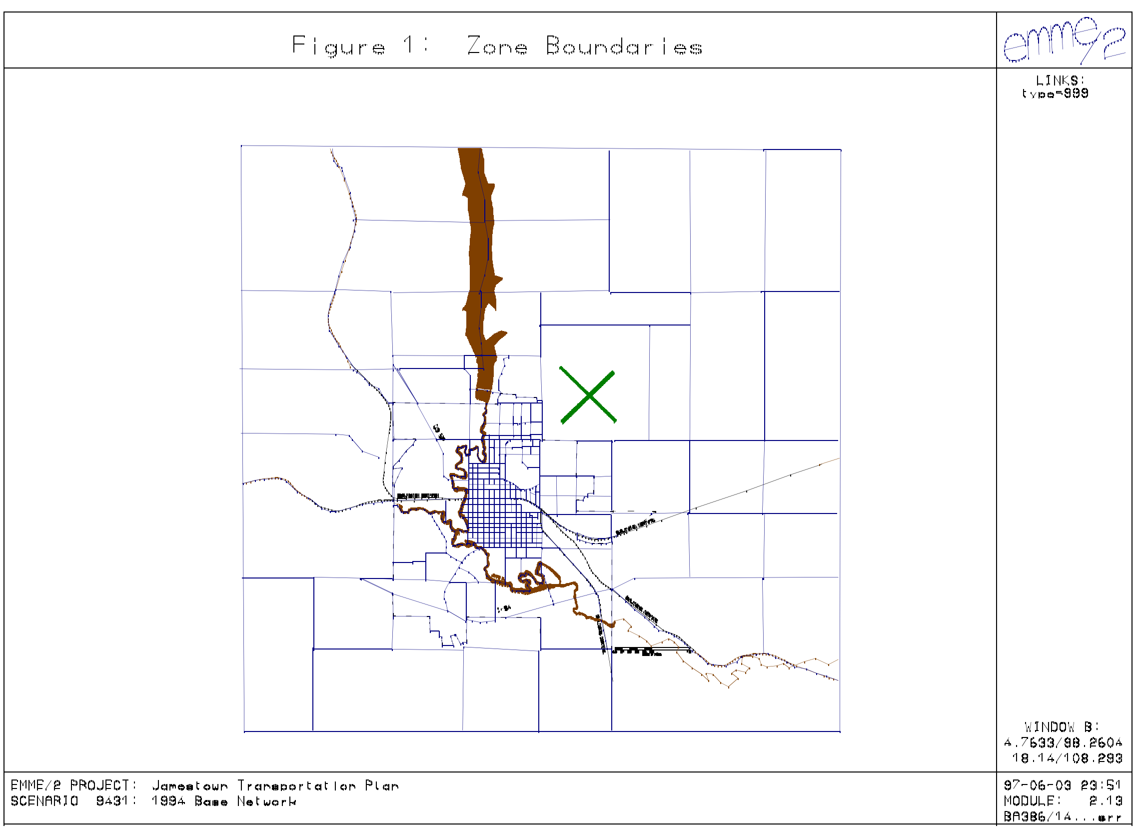
The real task, therefore, was how to extract the specific travel behavior information required for the model. This was done first by estimating a simple, generic model, using typical rates for trip generation, distribution, and other key parameters such as external trip percents. Executing this model provided an initial vehicle trip table, which can then be used in the second step. The initial trip table is adjusted to correspond with the set of observed counts in a process commonly known as O/D by traffic counts. Though not a unique solution, this adjusted trip table represents a most likely outcome, given the initial trip table. Next, the third step is to run the model "backwards" to obtain implied parameters regarding trip distribution (F and K factors) and trip generation rates. Finally, the model is run forward again, using the new parameters, to create a new seed matrix for adjustment. the third and fourth steps are repeated until no further improvement can be obtained. This is the overall process. The rest of this paper will discuss several key details of the process, the results for Jamestown, and conclusions regarding the process.

Initial Model Development

The first task was to develop an initial model, however approximate, which would ultimately supply the first trip table for the O/D by traffic count adjustment. The supply side of this model is represented by the network and zone system. A very detailed zone system was developed, which included 274 zones. In the core of the city, virtually every block was itself a zone, and almost all streets were explicitly coded. Eleven external zones were used for identifying major access locations. The modeled area consisted of roughly a square 8 miles by 8 miles in dimension. The detail of the network was necessary to support the number of counts. A detailed zone system would also enhance the trip table adjustment process, by providing the greatest level of specificity in terms of origin and destination locations. The zone system, and modeled area, are shown in Figure 1.

Once the network had been developed, and the socioeconomic data assigned to the proper zones, the demand side of the forecast model was developed, starting with the trip generation model.

Though many options exist for the functional form of the trip generation model, a linear model form was selected, based on experience in smaller cities, and in keeping with similar models developed in Grand Forks, Fargo and Bismarck. Three trip purposes were modeled, including Home-Based Work (HBW), Home-Based Other (HBO) and Non Home Based (NHB). The variables selected for the trip production model were single family detached households, households in groups of 2 to 9, and households in high-density situations (greater than 10). These groupings



probably serve, in part, as a surrogate for wealth in small cities such as Jamestown. The trip attraction equations used total employment for HBW attractions. Both HBO and NHB attraction equations used all three household types, plus retail and non-retail employment. The initial coefficients were selected based on standard practice and typical generation rates in other North Dakota cities. In all cases, the trip generation equations were assumed to estimate daily vehicle trips.

Once the trip generation equations were defined, external trips were estimated by taking a percentage of each zone's productions and attractions. At the external stations, existing counts were used as a starting point. They were divided by purpose, and split into productions and attractions corresponding to control totals. These totals were derived by examining current count volumes, and the results of a 1970 survey, which was used to estimate through trip percentages. Standard values for characteristics of external trips, from NCHRP 1871 was also used. Table 1 summarizes some of the key parameters used in the initial model, which were held constant throughout this process.

Table 1: Key model parameters held constant in calibration process

Parameter		Value
Work Trips Per Employee		1.7
External Trip Purpose Shares	HBW	15%
	HBO	70%
	NHB	15%
Auto Occupancy	HBW	1.20 Persons/Vehicle
	HBO	1.32 Person/Vehicle
	NHB	1.24 Person/Vehicle
Exponent for Friction Factors	HBW	1.99
	HBO	2.40
	NHB	2.35

Finally, a set of special generators were identified. These included the Airport, Shopping Centers, the State Hospital, and Jamestown College. Trip generation for special generators was based on standard rates from the ITE Trip Generation Manual, and were not adjusted during the calibration process.

Trip distribution was next. For external trips, a fratar process was used to distribute trips to external stations. The initial distribution was based upon previous O-D survey percentages. Internal trip distribution was by type, and used a gravity model formulation. Friction factors were calculated by the following formula:

$$F_{ij} = C \times \frac{1}{t^n}$$

Where F_{ij} is the friction factor between zones i and j. C is a normalizing factor, t is the travel time, and n is a parameter which varies by trip purpose (the latter is listed in Table 1).

The trips were summed, and transformed into origin/destination format from production/attraction format. The trip table was assigned for reference purposes. However, the primary use of the trip table was to use as a seed matrix for the O/D by traffic count adjustment.

Adjustment of O/D trips

In this step, the information implicitly contained in the observed set of traffic counts is used to update the initial trip table. A total of 568 one-way count locations were used as a basis for the comparison. These were compiled from the counts taken as a part of this study (mid block and intersection) and historic count data. They were expressed as average weekday traffic volumes, adjusted for seasonal and day-of-week variations by the North Dakota DOT.

The particular application used for this study was developed by Mr. Heinz Spiess, of INRO Consultants, and is documented in a May 1990 CRT publication[1]. The process is based on minimizing an objective function. This function is a measure of the difference between observed and estimated volumes on the links. A gradient method is used to find this minimum, which in turn corresponds to a new trip table. Note that the nature of the problem means that there are a very large, if not infinite, number of “best” solutions, mathematically. One common problem in O/D trip table estimation is that of degeneracy--that is, as the search for a minimum objective function progresses, the new trip table becomes increasingly unlike the original matrix. In modeling terms, this is undesirable, since much valuable information may be lost if the final, adjusted trip table is quite different from the original. An advantage of the gradient method, used with optimum step lengths, is that the final trip table will not be radically different from the original. This will ensure that measures of travel behavior such as average trip time, vehicle-miles and vehicle-hours of travel, and trip generation rates, will not be radically changed because of the O/D adjustment process. In addition, zonal interchanges with no trips remained at zero in this process, and an additional constraint was placed on the process which limited any zone-to-zone trip adjustment to between 50 percent and 150 percent of the initial value. In this way, the total number of trips did not grow excessively, which was otherwise the case.

It is important to note that for the O/D adjustment to work properly, and produce reasonable results, several items required special attention, including:

- The counts must correctly entered, and consistent among themselves. Erroneous or unreasonable counts will distort the O/D trip table adjustments. Inconsistent counts will also distort the trip table adjustments, and lead to mismatches (observed vs. estimated) even among valid counts.
- The network must be accurately represented. Obviously, the observed counts reflect travel behavior on the actual street and highway network. To the extent that the modeled network does not reflect the travel times actually experienced, at least in relative terms, the O/D adjustment will be unable to properly develop a realistic trip table.
- The location of the count links should ideally be placed to capture all trips in the network. Counts near major generators may lead to distortions. Also, any count which might have high intra-zonal trips should be avoided, since these trips will not be adjusted or even accounted for in the assignment process.
- The requirement to minimize intra-zonal trips in the previous items leads to and advantage in creating a very detailed, small-zone network, which we did in the Jamestown model.

In the actual application of this gradient approach, the process was executed through the use of a macro applied within the EMME/2 (c) transportation planning software package. The nature of the optimization problem means that the process is iterative, as the minimum of the objective function is searched. After examining the convergence characteristics of the process, ten iterations were selected. Further iterations showed very little improvement in observed vs. estimated comparisons.

---

1. Spiess, Heinz. “A gradient approach for the O-D matrix adjustment problem.” Publication #693, Center for Research on Transportation, Université de Montréal, May 1990.

Reversing the Model:

Once the trip table had been adjusted, considerable improvement was evident in the fit between observed and modeled link volumes. The next step was to develop a model which obtained this target trip table. In order to do this, the model was reversed, in the following way

Step 1: The resulting daily trip table first must be split into the HBW, HBO, and NHB trip purposes. This was done by factoring the total daily trip table by the previous splits. Note that this implicitly assumed that the O/D adjustments applied equally to all purposes, at least in terms of distributional factors.

Step 2: The purpose trip tables were then multiplied by purposes-specific occupancy factors, which provides person-trip tables. Though still in O/D format, these trip tables may then be used to adjust the trip distribution for internal trips. Internal/External and through trips were separated into separate matrices, and were used as seed matrices for the frataring process in the next iteration.

Step 3: Calculate new K-factors by the following formulas:

$$\text{Kfact Adj} = \frac{\text{Obs Trips}}{\text{Friction Factor}}$$

$$\text{Kfact Adj'} = \frac{\text{Kfact Adj}}{\left(\frac{\sum \text{Kfact Adj}}{n}\right)}$$

where:

Kfact Adj = Adjustment Factor to apply to previous iteration Kfactors (by zonal interchange)
Obs Trips = Observed Trips by purpose
Friction Factor = Friction Factors, based on travel time, and specific to purpose

Next, the Kfactor adjustment factor is “normalized” by dividing each cell value by the unweighted average, as shown in equation 3, where:

Kfact Adj’ = “normalized” Kfactor Adjustment Factor
n = Number of cell values (274x274 = 75,976)

Finally, the new Kfactors are estimated by multiplying the adjustment factor by the previous K- factors (initial Kfactors were set to 1.0 for all interchanges). This is shown in equation 4.

$$Kfact_{i+1} = Kfact_i \times KfactAdj$$

where:

$Kfact_{i+1}$ = New Kfactor, for iteration $i+1$
$Kfact_i$ = Previous iteration Kfactor (iteration i)

As an option, the Kfactors may be aggregated to district interchanges. This will tend to even out extreme Kfactor values.

Step 4: The observed trip tables, by purpose, were summed by row and column, which represented the total person-trips in and out of each zone on an origin/destination basis. These totals were then split into productions and attractions, using the P&A proportions from the previous model iteration trip generation results. This approach assumes that any changes resulting from the O/D adjustment apply to both production and attraction land uses equally. The result of this step was a set of observed productions and attractions for each zone, in terms of person-trips.

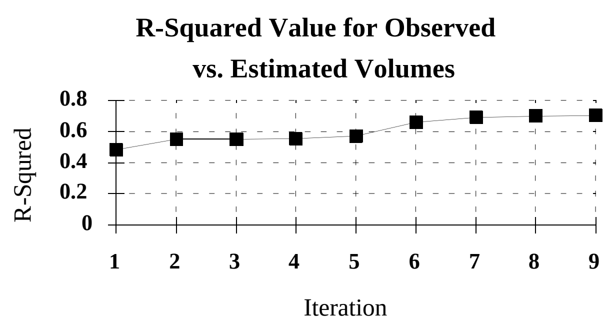
Step 5: The trip generation equations were re-estimated, using the new "observed" productions and attraction totals. This was done using a standard linear regression technique.

At the end of step 5, a new set of trip generation equations, new K-factors, and new seed matrices for Externally-based travel were ready. The model was then applied "forwards" to estimate a new daily vehicle trip table. At this point, the O/D by traffic count adjustment was applied and the process of re-estimating the model was started again. This process was continued until no further improvement could be obtained in terms of the observed vs. estimated count comparison.

Results

The success of the procedure may be measured by several factors. Most importantly, we should expect that the process would improve the model performance in terms of it's ability to replicate observed volumes. The results showed this to be the case. Initially, the modeled volumes compared with the observed counts with an r-squared value of 0.553. This eventually improved to an r-squared of 0.716, a 30 percent increase in r-squared. Figure 2 shows the progression of r-squared values by iteration. A total of 12 iterations were used before it was determined that no further improvement could be made. Eventually, the model was able to meet the NDDOT requirements for model performance criteria.

Figure 2



The final comparison plot showing observed vs. estimated link volumes, and the NDDOT criteria, are shown in Figure 3. The criteria permitted a percent deviation from observed, based on the magnitude of the observed volume. Ninety percent of the count locations must meet this criteria. In this case, those count locations not meeting the criteria were not located systematically, and were not key segments. The model also resulted in reasonable trip generation equations, as shown below:

HBWP = 3.272*SFDU + 2.591*MFDU + 2.976*HDDU
HBOP = 9.728*SFDU + 7.708*MFDU + 9.209*HDDU
NHBP = 2.052*SFDU + 6.216*MFDU + 1.035*HDDU

HBWA = 1.7*EMP
HBOA = 4.451*SFDU + 2.788*MFDU + 3.037*HDDU + 25.958*RET + 0.604*NRET
NHBA = 1.039*SFDU + 0.732*MFDU + 0.835*HDDU + 2.678*RET + 2.364*NRET

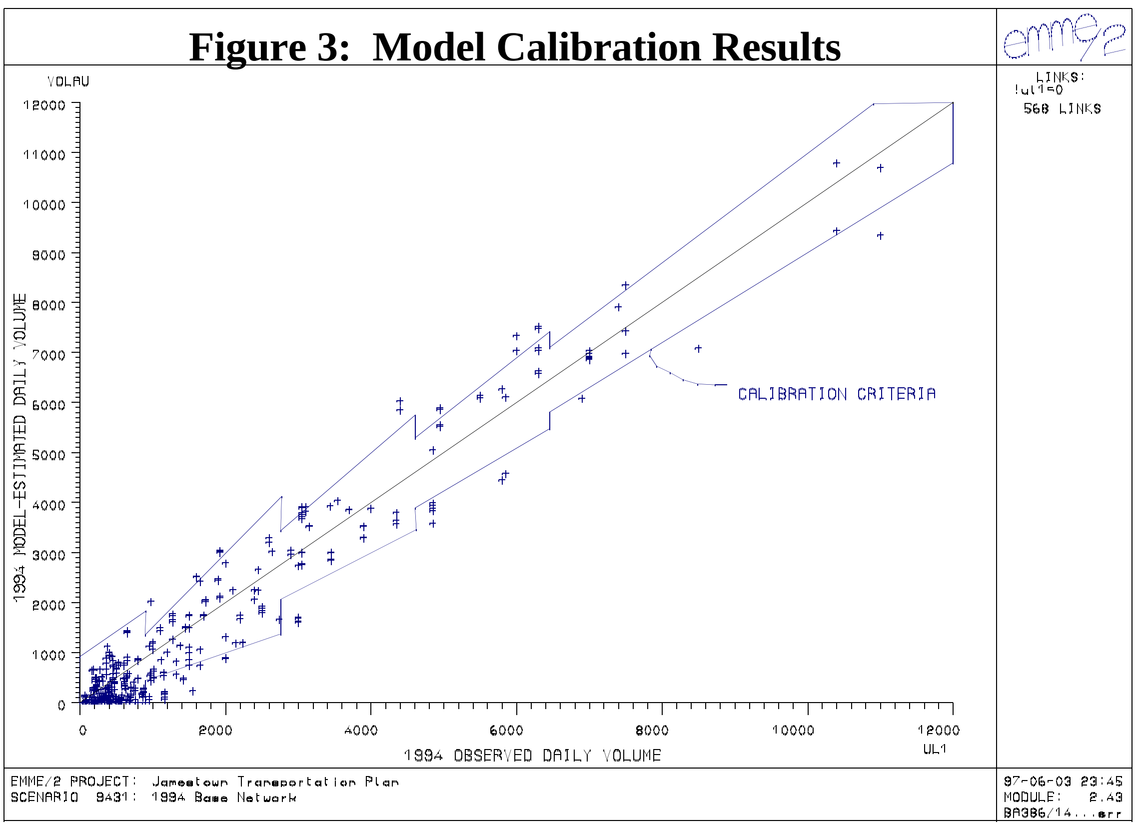
where:

HBWP = Home-Based Work Person-Trip Productions
HBOP = Home-Based Other Person-Trip Productions
NHBP = Non-Home Based Person-Trip Productions
HBWA = Home-Based Work Person-Trip Attractions
HBOA = Home-Based Other Person-Trip Attractions
NHBA = Non-Home Based Person-Trip Attractions
SFDU = Number of Single-Family Detached Homes
MFDU = Number of Households in groups of 2-10
HDDU = Number of Households in groups of 10 or more
EMP = Number of Employees
RET = Number of Retail Employees
NRET = Number of Non-Retail Employees

The relatively higher HBW and HBO production rates for HDDU may have resulted from the relatively few (18) zones in which these occurred, which provided a small observed data set for calibration.

The trip generation equations imply about 15 one-way trips per person per household. With an average occupancy of 2.3 persons per household, the model assumes about 6.5 one-way trips per

Figure 3: Model Calibration Results



person per day.

Screenline comparisons were very good. Three screenlines were used, following major east-west and North/south barriers within the city. No screenline deviated more than 1 percent, and the average deviation from observed counts was 0.3 percent low. Screenline comparisons were refined by the application of special K-factor adjustments.

Areas for Improvement

While the procedure resulted in a workable model, the process revealed several areas which could be improved. These included:

1. Estimation of new trip generation parameters. An alternative method might have been to add a unique zonal constant to each zone's trip generation equations. This would have preserved the original parameters yet more explicitly represented the variation in zonal trip characteristics. This would require, however, more confidence in the original parameters than was available for Jamestown. This type of methodology was applied in subsequent model estimation projects with some success.

2. Use of additional O-D information. Though not available for Jamestown, some additional origin-destination information is often available in the form of the Census CTPP data, previous model output or previous survey information. Since the trip table adjustment does not produce a unique result, it is important to use an initial matrix which is as accurate as possible.

3. Identification of special generators. In Jamestown, a small set of special generators were preselected. However, the calibration process may reveal other zones which do not fit the standard trip generation equations. The land use in these zones should be inspected to identify unusual land uses, which could be treated as special generators.

4. Automation of iterative calibration process. The steps of trip table adjustment, re-estimation of distribution parameters, re-estimation of trip generation parameters, model execution, and count comparison are essentially straightforward computational steps. These can and should be automated, to speed up the calibration process. Reporting these results in summary fashion allows the modeler to review progress and check for errors.

Conclusions

The process described here proved successful, in that it produced a reliable, conventional travel demand model for a community at a relatively low cost, and without the need for detailed origin-destination survey data. It was greatly aided by the OD trip table adjustment routine, an application of the EMME/2 software. The process does require a detailed zone and network system, with accurate travel times. Land use assumptions must be similarly detailed, with proper identification of special generators so that the basis for trip generation can be accurate. While some model assumptions, such as the trip generation variables, functional form, and friction factors, must be set exogenously, most of the actual model parameters were estimated based on the detailed traffic count database. This model estimation process may be particularly useful for communities with not history of travel demand models, and which lack resources to develop a detailed origin-destination trip database.