APPENDIX C: MODELING PROCESS
Modeling Process

Travel demand forecasting models (TDMs) are a major analysis tool for the development of long-range transportation plans. These mathematical models are designed to calculate the number of trips, connect their origins and destinations, forecast the mode of travel, and identify the roadways or transit routes most likely to be used in completing a trip. Models are used to determine where future transportation problems are likely to occur, as indicated by modeled roadway congestion. Once identified, the model can test the ability of roadway and transit system improvements to address those problems.

For the 2040 Transportation Plan, MACOG contracted with Resource Systems Group (RSG) to conduct a major update of the travel demand forecasting model. A hybrid model, the new design blends aspects of traditional four-step models and activity-based models. The model can be described as trip-based, as it produces aggregate trip table matrices of trips between origins and destinations rather than disaggregate records detailing individual travelers’ activities. However, it can also be described as tour-based since the travel patterns predicted can be mathematically proven to be consistent with tours and all travel is segmented within the model by types of tours, eliminating the non-home-based trips problematic in traditional four-step models.

Unlike traditional four-step models which are entirely aggregate and activity-based models which are entirely disaggregated, the hybrid model includes both aggregate and disaggregate component...
models. Despite the inclusion of disaggregate choice models, there are no random number draws or Monte Carlo simulations included in the TDM. As a result, MACOG’s model results are reproducible, unlike the results of activity-based or other simulation models. Any difference between two model runs is directly attributable to differences in their inputs as with traditional trip-based models. Whereas, in simulation models, multiple model runs are necessary when comparing alternatives to ensure that the difference between model runs results from differences in the alternative inputs rather than from differences in the random numbers drawn for each run.

Significant elements of the MACOG TDM are as follows:

HELPViz Land Use Model
HELPViz was developed by RSG as part of the Sustainable Evansville Area Coalition’s Regional Plan for Sustainable Development. Using the Land-Based Classification System’s activity-based codes, 2002 aerial photography and 2013 oblique photography was used to describe land use changes in the urbanized areas of the region over a 10-year period which was then used to adapt HELPViz to the area.

This land use model offers sensitivity to land use zoning, building codes and infrastructure facilities such as the transportation network, water and sewer utilities. HELPviz allocates the future population and employment regional totals to the TAZs based on build out capacities, the transportation network and infrastructure facilities. HELPviz uses Nested Logit model framework and uses information at both TAZ and parcel levels.

Population Synthesis
In recent years there has been a shift away from the application of demand models directly to entire traffic analysis zones in favor of representing individual households (and sometimes persons) and modeling travel behavior at their level. This shift is to avoid the aggregation bias that occurs when non-linear demand models are applied to aggregate or average characteristics rather than to populations with a range of attributes around the group averages. For example, a mode choice model may predict no significant transit mode share when applied to a zone with 100 households with an average of 2.2 cars per household. However, the same mode choice model, applied to the same households individually, may predict a significant number of transit trips if 5 of the households have no vehicles and 15 have only one vehicle.

The MACOG TDM generates a disaggregate synthetic population of households based on the demographic information associated with the traffic analysis zones. For each zone, individual households are created. Each household has a total number of persons, workers, students, and a binary variable indicating whether any of the household members is over the age of 65. Each household also has an income variable that indicates whether the household belongs to the lower (under $35,000/year), middle ($35,000 - $75,000/year) or upper (over $75,000/year) income category, each of which comprises approximately a third of the households in the region. The number of vehicles available to each household is modeled separately, after the population synthesis, based on these variables and other variables describing the zone in which the household is located.

The synthetic population is developed in two steps. First, a set of ordered nested logit models predict for each variable (such as household size, number of workers, etc) the number of households which have each level of that variable (one person, two persons, etc; zero workers, one worker, two workers, etc). Second, iterative proportional fitting is used to develop the synthetic population based on a seed population of households from the household travel survey and the marginal distributions for each variable provided by the logit models. Unlike the procedures used to develop synthetic populations in many activity-based models, this procedure is entirely deterministic and does not introduce randomness or simulation error into the model using any random draws. This is possible since the model is allowed to produce more or less individual households that exist in the real population, creating consistency instead by weighting those households so that their weighted sum is the total
actual number of households in each zone.

**Tour and Stop Generation**
The new TDM generates tours and stops rather than trips. The number of tours and stops of each type is estimated using multiple regression models applied to the disaggregated synthetic population of households. First, the number of tours, of each type, is estimated for each household. Then, for each stop type, the ratio of stops per tour is modeled and the total number of stops produced by multiplying this ratio by the number of tours.

In this framework, the modeled behavior is dominated by the tour generation equations, with the stop generation playing a secondary role (in some ways similar to, albeit simpler than, activity-based approaches which allow more tradeoffs). This is reflected in their goodness-of-fit which is quite good for the tour generation equations, but rather modest for stop generation since stop rates per tour are relatively constant. As mentioned previously, more elaborate model frameworks which allocate stops to tours may be developed at a later date, giving the model additional behavioral fidelity. However, the simple framework adopted here still offers improved sensitivity over traditional models.

Although cross-classification models were once viewed as an ad-

<table>
<thead>
<tr>
<th>Table C-1: Factors Affecting Household Tour and Stop Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Workers</strong></td>
</tr>
<tr>
<td><strong>Work Tours</strong></td>
</tr>
<tr>
<td>Work Stops</td>
</tr>
<tr>
<td>Other Stops</td>
</tr>
<tr>
<td><strong>School Tours</strong></td>
</tr>
<tr>
<td>School Stops</td>
</tr>
<tr>
<td>Other Stops</td>
</tr>
<tr>
<td><strong>Other Tours</strong></td>
</tr>
<tr>
<td>Short Maintenance Stops</td>
</tr>
<tr>
<td>Long Maintenance Stops</td>
</tr>
<tr>
<td>Discretionary Stops</td>
</tr>
</tbody>
</table>

**Key**
- Variable (column) increases tour/stop rate (row)
- Variable (column) decreases tour/stop rate (row)

Source: MACOG Travel Model: Model Development and Validation Report
vance over regression models for generating trips, this was due to their ability to reduce aggregation bias compared to regression models which were applied to zones as a whole. By applying regression models instead to a disaggregate population, aggregation bias is eliminated altogether in the approach adopted here. While cross-classification models are limited to two or three variables at most, regression models can include more variables, introducing sensitivity in resulting trip rates to factors like gas prices and accessibility variables, in addition to the basic demographic characteristics. Although interaction effects were widely tested, the only interaction effect that proved significant was the interaction of gas prices and household income; increasing gas prices decreased certain stop rates, but only for low income households.

The number of work tours was mostly a simple function of the number of workers. Vehicle ownership, the presence of seniors and household income offered some additional explanatory power. The presence of seniors in a household made work tours slightly less frequent, perhaps because senior workers are less likely to work full time.

The number of work stops is calculated for each household and allocated to income groups based on the household’s income. The number of work stops per work tour is relatively constant. However, the number of work stops per work tour is slightly higher for high income workers, probably reflecting greater frequency of eating out for lunch which results in two work stops (before and after lunch). Accessibility also makes work stops marginally more frequent because it implies that commute times are shorter, so it is easier to get back and forth between home and work, such as going home for lunch, returning to work after dinner, work activities on weekends, etc.

The number of other stops per work tour is significantly increased by the number of household students from workers stopping to drop off students on the way to work and decreases with the number of non-workers in the household who can drop off the students instead. Here also, we see income and vehicle ownership increasing other stops on work tours, again perhaps increased lunch stops out.

The number of (primary and secondary) school tours is largely a simple function of the number of students in a household. The number of school tours does increase with accessibility, like with work stops, because it is easier to get back and forth between home and school. Income also marginally increases the number of school tours with more students, perhaps indicating that higher income households are more likely to send their children to different schools or that their high school students drive separately and their primary school children are picked up/dropped off on another tour.

The number of school stops per school tour is essentially constant at just over one, although very slight increases result from higher income and accessibility. Other stops on school tours were also largely constant, but were somewhat more common for students from households with higher income. The increase related to higher income students may have more money to spend, hence may make more shopping stops, etc.

The number of other (non-work) tours made by a household is most influenced by the number of non-workers in the household: more non-workers generate more non-work tours. However, the non-work tours are also increased albeit less by workers and are more frequent for households with seniors and more vehicles. Non-work tours also decrease slightly as gas prices rise. The number of short (under 30 minutes) maintenance stops per other tour was largely constant, but somewhat higher for households with more people and income. The number of long (over 30 minutes) maintenance stops was also fairly constant and increased with the number of vehicles available; however, it also decreased with the number of students, who may curtail long shopping activities. The number of discretionary stops decreased slightly with the presence of seniors and increased with income and students with cars.

In the new hybrid tour-based framework, there are no attraction
generation models. Rather, attractions are modeled as part of the stop location choice models, instead of inputs to trip distribution. The model script does generate attractions, but only because TransCAD requires it. In fact, the actual attractions are part of the stop location choice models.

**Tour-Based Modal Choice**

In the new model, as in activity-based models, the mode of travel is developed in two stages: tour mode choice and trip mode choice. After tours are generated, they are assigned a primary mode by tour mode choice models. Then, after the spatial distribution of stops creates trips, individual trips are assigned a mode based on the primary mode of the tour in trip mode choice models.

The MACOG model makes use of four primary tour modes:

- Private Automobile
- Public Transit
- Walk / Bike
- School Bus

The primary mode for a tour is determined by a simple set of definitions or rules.

- Any tour containing a school bus trip is a school bus tour.
- Any other (non-school bus) tour containing a public transit trip is a public transit tour.
- Any other (non-transit) tour containing a private automobile trip is an automobile tour.
- Any other tour, which contains only walk or bike trips, is a non-motorized tour.

In this framework, the primary choice determining transit mode share is the tour mode choice. Trip mode choice ultimately reduces mostly to the determination of vehicle occupancy for automobile tours or the allocation of access modes for transit tours. Even in advanced activity-based models, fixed shares or other simple heuristics have been used for trip mode choice; whereas, tour mode choice models are more comparable to mode choice in traditional models.

The incorporation of behaviorally sensitive tour mode choice models in the TDM represents significant added value as compared to the previous model in which mode shares were fixed and totally insensitive to demographics, levels-of-service, or any other policy variables. The new model produces, in addition to automobile trips by occupancy class, the system-level transit ridership, the number of transit trips generated by each residence zone, and the total regional number of daily walk/bike trips. Moreover, the model architecture allows for the straightforward addition of future component models to produce transit and non-motorized trips at the route/street level. These component models and level of spatial fidelity could be developed in a future model upgrade.

The key difference between the tour mode choice models and those common in activity-based models is the way in which they measure the level-of-service provided by each competing mode and the related assumption of the hierarchy of travelers’ choices (i.e., whether travelers’ destination choices depend more on their mode choices or vice versa).

In activity-based models, as in traditional four-step models, tour mode choice is modeled after destination choice (or distribution) and can therefore use actual travel times between origins and destinations as level-of-service variables. This traditional model structure was first developed for very large metropolitan areas with significant choice rider markets and is more sensitive to changes in level-of-service provided by transit improvements and for testing their impacts on transit route ridership. However, it may be oversensitive to level-of-service variables and a source of optimism bias in transit forecasts, as this model structure is built on the assumption that travelers are more likely to change mode than destination. This may well be the case for affluent choice riders for their work.
## Table C-2: Factors Affecting Tour Mode Choice

<table>
<thead>
<tr>
<th></th>
<th>LOS</th>
<th>Cost</th>
<th>Demographics</th>
<th>Built Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Access</td>
<td>Gas Price</td>
<td>Bus Fare</td>
<td>Workers</td>
</tr>
<tr>
<td><strong>Access by Mode</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Auto</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Transit</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Non-Motorized</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>School Tours</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Auto</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transit</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Motorized</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>School Bus</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Other Tours</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Auto</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Transit</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Non-Motorized</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Key**

- **Direct Increase**
- **Indirect Increase**
- **Indirect Decrease**
- **Direct Decrease**

*Source: MACOG Travel Model: Model Development and Validation Report*
commute in large cities. However, there are many situations where it is more reasonable to assume that travelers are more likely to change destinations than mode.

Local household survey data offer some support of this general assumption for the MACOG region that travelers are more likely to change destination than mode of travel. In general, this assumption seems more appropriate in markets similar to MACOG with few choice riders, where mode choice is generally a foregone conclusion on which destination choice is conditioned. For example, either the traveler has access to a car and does not even think of riding transit or they do not have access to a car and rely on transit, choosing their destinations, possibly even workplace, based on where the transit system can get them. “Reverse hierarchy” models such as those developed for the new MACOG model, which represent destination (or stop location) choice conditional on mode choice, still take the level-of-service provided by competing modes into account and allow for changes in ridership based on improvements to transit or highway modes. However, these models measure the level-of-service provided by each mode not directly by the travel times between origins and destinations but indirectly by the accessibility to various types of destination provided by each mode to a residence zone.

**Departure Time Choice**

The new regional travel model includes departure time choice models which distribute trips throughout the day. The models are capable not only of producing the traditional AM, PM and off peak trip tables for standard assignments, but also can produce trip tables for any or all 15-minute periods from 6 am to 9 pm. These 15-minute trip tables should be of significant value for traffic micro-simulations and could be used in the future in conjunction with a dynamic network assignment.

In addition to adding temporal resolution, the departure time choice models add sensitivity to new variables, most notably travel times and accessibility. The new models will reflect shifts in travelers’ departure times in order to avoid longer travel times. This effect, commonly referred to as peak-spreading as travelers leave earlier or later to avoid peak traffic, was modest, but already statistically significant in the household survey data. The effect was evident for all tour types but was most significant for Other Tours, which, in general, presumably have more flexibility in the timing of their activities than the other tour types.

The models also incorporate accessibility variables which allow departure times to vary geographically in the model, e.g., lower accessibility, rural travelers might generally leave for work earlier (since they have further to go to get to work).

Home-based and non-home-based trips for each tour type are represented by different models, since the first and last trips of a tour have different temporal distributions compared with mid-tour non-home-based trips. This segmentation is particularly important for midday/lunch traffic which is associated primarily with shorter, mid-tour non-home-based trips, as opposed to the AM and PM peaks which are more associated with longer home-based trips.

**University Student Travel Models**

**Michiana Area College Travel Study**

The university student travel models are supported by the Michiana Area College Travel Study. The College Travel Study closely paralleled the Michiana Area Household Travel Study in questionnaire structure and content. Six colleges agreed to participate in the study: Bethel College, Goshen College, Holy Cross College, Ivy Tech Community College, the University of Notre Dame, and Indiana University – South Bend.

Before administering the College Travel Study, the survey was soft-launched to 25 students from Goshen College. Goshen College was gracious to agree to soft-launch the survey as a way to test the data and ensure that the survey questions were clear and relevant to students taking the survey. After the soft-launch was completed, the data was reviewed. The College Travel Study was then administered with each participating college sending out an invitation...
email. Survey administration began on Wednesday September 18, and closed on October 14. This survey administration timing was specifically selected to ensure that the survey started after classes were in session (and the add/drop period had passed) and the survey was completed prior to the October break period. A total of 672 students completed the survey.

Travel Market Segmentation
University student travel is modeled in three distinct market segments:

- Full-time On-Campus Students
- Full-time Off-Campus Students
- Part-time Students

Each of these three market segments have distinct travel characteristics and is represented by its own set of travel models. Even the structure of these models differs between market segments, reflecting the fact that their travel decisions are different. Full-time students are assumed to be excluded from the Census household population. Their travel is represented by a fully distinct segment of non-household travel with its own fully separate models. In contrast, part-time students are modeled as members of a permanent household (which would have been counted in the Census) and is therefore represented in the households in the TAZ layer and synthetic population. Their university travel is generally considered just one, albeit special, segment of their overall household travel with trips to and from campus represented as special stops on work and other tours.

Full-Time On-Campus Student Travel
Full-time on-campus university student travel is represented with a hybrid trip-based model more similar to a traditional trip-based model. The demand model has four steps: trip generation, destination choice, mode choice and time-of-day split. The travel is decomposed into two segments or trip types:

- CB – Campus-Based Trips
- NC – Non-Campus Based Trips

These are analogous and nearly synonymous with home-based and non-home-based trips in traditional passenger models.

Full-Time Off-Campus Student Travel
Full-time off-campus university student travel is represented with a hybrid tour-based model, similar to household trips in the MACOG model. The demand model has five steps: trip generation, stop location choice, stop sequence choice, mode choice and time-of-day split. The travel is decomposed into one tour type and two stop types:

- UCT – Full-time, Off-Campus University Student Tours (Campus-Based)
- OCH – Student home (off-campus)
- OCO – Other off-campus location

These are analogous with the tour and stop types in the MACOG household travel models.

Part-Time Student Travel
Part-time university student travel is incorporated in the household travel models. Special university stops are included on work and other tours to represent part-time students’ travel. Most of the part-time student’s travel is therefore modeled together with other non-university travel generated by the same household.

Truck Model
Based on the method recommended in the Quick Response Freight Manual II, a commercial vehicle model was developed for predicting trips for four-tire commercial vehicles, single unit (SU) trucks, and multiple unit (MU) trucks. The model uses a four-step process. These steps are trip generation, distribution, choice of time of day and trip assignment. In addition, the special trip generators of inter-region and inter-modal trucks were added in the model to bet-
ter replicate the current inter-region and inter-modal truck movements.

The inputs to trip generation are the number of employees and the number of households by Traffic Analysis Zone (TAZ). These rates were obtained by adjusting the original generation rates in the Quick Response Freight Manual. To replicate the current truck traffic condition in the study area, the rates for four-tire commercial vehicles were further adjusted by a factor of 0.10.

The external-internal (EI) and internal-external (IE) truck trips were classified as a distinct type of trip in order to better replicate the in-balance direction truck flows at different time periods. Before the trip distribution, the trip origins and destinations were balanced for all TAZs and external stations for the following types of trips:

- EI-IE SU truck trips of all TAZs and external stations
- EI-IE MU truck trips of all TAZs and external stations
- Internal-to-Internal (II) SU truck trips of all TAZs
- Internal-to-Internal (II) MU truck trips of all TAZs
- Internal-to-Internal (II) 4-tire commercial vehicle trips of all TAZs

For four-tire commercial vehicles, it is assumed that the normal EI-IE trip attractions are proportional to the trip destinations. At the beginning, destinations are used as the normal EI-IE trip attractions and the balancing process scales to the total adjusted attractions.

For single-unit and multi-unit trucks, a destination choice model was applied separately to internal & external trips. The destinations chosen in these models (the sum over all origins) are scaled to the total number of trips produced in generation. This vector is then used as both the productions and attractions for a doubly-constrained gravity model to distribute the truck trips.

The time-of-day assignments were implemented in order to obtain better model results. To facilitate this, the trip tables from trip distribution must be factored to reflect morning peak, midday, and off-peak periods prior to trip assignment. The hourly time-of-day factors were derived from classification traffic counts provided by MACOG and applied to the MACOG Regional Travel Model.