STEP2 for Clark County
Household Microsimulation for Transportation Policy Analysis

prepared for

Southern Nevada Regional Planning Coalition

prepared by

Caliper Corporation

November 2002

Table of Contents

Introduction ... 4
Outline of the Report ... 4
Background: Microsimulation, STEP, and the Objectives of STEP2 ... 5
Aggregate Model Applications versus Microsimulation ... 6
Existing Aggregate Transportation Modeling Capability of Clark County ... 6
Advantages of Modeling at the Level of the Individual ... 8
Overview of STEP2 for Clark County ... 10
Framework ... 12
Model Development ... 13
Simulating Individual Travel Patterns from a Travel Demand Model ... 13
Running the Model in TransCAD ... 14
Data Requirements ... 15
Outputs ... 17
Flexibility and Portability ... 17
Model Documentation ... 18
Generating the Synthetic Population ... 18
Population Synthesis ... 19
Aging the Population ... 21
Geographic Databases, Network Development and Network Skimming ... 24
Travel Demand Models ... 25
Long-term Lifestyle and Mobility Decisions ... 25
Residential Location ... 25
Labor Force Participation and Retirees ... 26
Work Location ... 27
Other Lifestyle and Mobility Decisions ... 29
Daily Trip Decisions ... 29
Work Participation and Tour ... 29
Non-Work Tour Trip Frequency ... 31
Time of Day ... 32
Destination/Mode ... 35
Output of Travel Demand Model: Individual Trips ... 36
Aggregation and Assignment ... 37
Aggregate Trip Matrices ... 37
Traffic Assignment ... 40
Future Directions ... 40
Conclusion ... 40
References ... 42

List of Figures

Figure 1: Representing Travel Patterns as Independent Trips 7
Figure 2: Complexity of Commute Patterns in Las Vegas... 9
Figure 3: Ways an Individual Might Respond to a Peak Period Toll 10
Figure 4: Model Framework for STEP2 11
Figure 5: Example of Simulating a Travel Outcome for an Individual 14
Figure 6: Custom TransCAD Model Toolbox for STEP2 15
Figure 7: Scenario and File Managers for STEP2 in TransCAD... 16
Figure 8: New TransCAD Procedures for Logit-based Microsimulation 17
Figure 9: New TransCAD Survey Processing Tools to Generate Inputs for STEP2................... 18
Figure 10: Inputs for Population Synthesis 19
Figure 11: PUMS Household and Person Data Records 20
Figure 12: Household and Person Files Generated by Population Synthesis 21
Figure 13: Workforce Participation and Retiree Model Estimation Results................................ 27
Figure 14: Distribution of Home-based Work Tour Type 31
Figure 15: Distribution of Non-work Tour Trips 32
Figure 16: Work Start and End Time Lookup Table 33
Figure 17: Before Work Stops and After Work Stops Timing Lookup Table............................. 34
Figure 18: Time of Day Distributions 34
Figure 19: Mode Choice Model Structure.. 36
Figure 20: Simulated Trips for Synthetic Persons 37
Figure 21: STEP2 Produces Trip Matrices by Socio-economic Group 38
Figure 22: Example Output of Trip Characteristics by Income Group.. 39
Figure 23: Example of Changes in Mode Share by Race, Base Year to Future Scenario 39

Introduction

This document describes the implementation of a household microsimulation travel demand model for Clark County, Nevada. The principal objective of this project is to develop and apply state-of-the-art microsimulation transportation planning techniques to capture the essence of environmental, land use, and transportation interrelationships. There are two key attributes that make this model fundamentally different than the aggregate, 4-step model currently used by RTC: (1) the model is applied at the level of the individual rather than based on zonal averages and (2) it includes a residential choice component to forecast the spatial distribution of households in the urban area.

The implementation builds upon the STEP policy analysis tool created by Greig Harvey in the 1970s and enhanced through the 80s and 90s (Harvey, 1978, and Harvey and Deakin, 1996). Harvey's work on STEP was based on two key philosophies. The first was the importance and value of applying travel demand models at the level of the individual decision-maker, rather than applying the models based on zonal averages. Application at the level of the individual both allows for the use of more behaviorally rich (i.e., realistic) travel demand models, and also allows for impact analysis to be performed for specific socio-economic groups. The second philosophy was to develop the tools in a way such that they would be accessible to transportation planning agencies without overly burdensome commitments of time and money. Therefore, STEP provided default models that could be calibrated to match local conditions, and the models are based on readily available data such as census data (including SF1, SF3, and PUMS), although it also makes use of household transportation surveys.

The philosophies of the original STEP model are also at the heart of this STEP2 project, in which the STEP modeling tool has been revived and enhanced for Clark County. STEP2 has gone beyond STEP in numerous ways, including running within a GIS environment, developing a user-friendly interface and scenario manager, incorporating realistic transportation networks and traffic assignment, integrating models that have been developed specifically for Clark County, and making use of the 1996 Las Vegas Valley household survey. The outputs of STEP2 for Clark County can be used to assess environmental, social, and economic impacts of various land use and transportation policies and plans. A valuable aspect of the output is the ability to produce impacts by socio-economic group, and therefore address issues of equity. Furthermore, the modeling analysis tool is easily expandable to include additional sensitivity to demographics and transportation policy variables by incorporating more detailed representation of travel and travel behavior.

Outline of the Report

This report starts out by presenting background on microsimulation, STEP, and the STEP2 project. Next the existing aggregate modeling capability used in Clark County is briefly described along with the advantages that can be gained from a microsimulation approach such as STEP2. Following this, an overview of STEP2 for Clark County is presented, including discussions of the model framework, model development, and basic information on running the model. Additional detail on the specific model components are presented in the following section on model documentation. Finally, future directions and conclusions are discussed.

Background: Microsimulation, STEP, and the Objectives of STEP2

The principal objective of this project is to develop and apply a state-of-the-art microsimulation transportation planning techniques to capture the essence of environmental, land use, and transportation interrelationships. This effort builds upon and enhances the STEP travel demand analysis package, which was originally developed in the 1970s as a sketch planning tool for the San Francisco area (Harvey, 1978). The key feature of STEP is that it is based on microsimulation, meaning that it uses the individual or household as the basic unit of analysis.

The essence of what makes STEP and STEP2 different than a traditional 4-step aggregate modeling implementation (e.g., the RTC Model in TransCAD) is that all of the processing is done at the level of the household and individuals in these households. That is, while an aggregate 4-step implementation is based on aggregate travel between zones, a microsimulation approach instead simulates a population of representative households and persons, and then makes forecasts by aggregating decisions made at the household level. There are numerous advantages to a microsimulation approach, including being able to tabulate impacts for subgroups of the population (for example, low income or elderly) and the capability of explicitly modeling realistic travel behavior patterns such as trip chaining.

The use of microsimulation for policy analysis was pioneered by the economist Guy Orcutt in the late 1950s (Orcutt et al., 1976). The driving force is that aggregate demand is made up of a large number of decisions made by individuals, and therefore it is necessary to do the behavioral modeling at the level of the individual. With this approach, one person (or household) is processed at a time, and then these individual decisions are summed up to produce summary statistics on the behavior (including the impacts of policies). It has long been recognized in transportation (since the 70s, at least) that there is great value in modeling transportation at the level of the individual. The basic argument is that people travel, not zones, and by averaging to the level of the zones, much information is lost. A driving force behind these ideas was Daniel McFadden, whose work on theory and methods for modeling choices at the level of the individual was awarded the Nobel Prize in Economics in 2000. He developed the widely used multinomial logit model (among many contributions), and his first application of it was to forecast ridership for BART in the San Francisco Bay Area. It is these logit-type models that form the building blocks for microsimulating travel demand.

While the theoretical advantages of microsimulation are well known, the technique has not been implemented to a wide extent as the vast majority of transportation planning agencies still relies on an aggregate, 4-step travel demand model. There are many reasons for this. One is that a microsimulation approach is significantly more computationally intensive. However, such computational limitations are being alleviated via dramatic increases in processing power and the use of multiprocessing techniques (for which microsimulation is a perfect application). A second reason for slow adaptation is the methodologies are more complex, and therefore require more expertise for development. They also tend to be more data intensive. The original STEP model made great strides in making microsimulation a more viable alternative for transportation planning agencies by creating default specifications that could be calibrated for different study areas. STEP2 aims to continue the progress in this direction. A third is that there is, as of yet, very little hard evidence of realized gains from modeling with microsimulation – a fact that will probably change if microsimulation tools are available to planners.

The objective of the STEP2 project is to build upon the original STEP work through development of a microsimulation planning tool for Clark County, Nevada. This project continues the work that Caliper (with the Environmental Defense) performed in the past to integrate STEP's home-based work models into TransCAD and apply the models to a case study in Baltimore (Slavin and Lam, 2001). This Baltimore project tested some key enhancements to STEP, including the use of a GIS platform and real transportation networks. STEP2 aims to further revive the STEP models and the general philosophy behind them, including modeling at the level of the household and developing tools in a way such that they are accessible to transportation planning agencies. In addition, the STEP2 project includes many enhancements to STEP, including:

- The development of flexible and generic microsimulation tools, data processing, and calibration capabilities so that the default specifications can be readily modified and enhanced.
- The use of TransCAD's GIS environment as a platform, so that model outputs can be analyzed visually using all of the capabilities of a powerful GIS.
- The use of real transportation networks and incorporation of traffic assignment so that changes in level of service resulting from changes in demand can be calculated internally. (In order to be able to be run quickly, the original STEP model did not have an internal transportation network representation or traffic assignment model.)
- The incorporation of tour- and activity-based modeling concepts by explicitly incorporating trip-chaining in the work tour.
- The capability of aging the population.
- And, specifically for Clark County,
  - The use of demand models developed for Las Vegas
  - Incorporation of travel behavior statistics derived from the 1996 Las Vegas household survey and recently released Census 2000 data.
  - The inclusion of a residential choice component to determine the spatial distribution of households.

Aggregate Model Applications versus Microsimulation

This section briefly discusses the key differences between aggregate modeling (as in a traditional 4-step transportation planning model) and a microsimulation (or disaggregate) approach. First the existing aggregate transportation modeling capability of Clark County are discussed, with an emphasis on aspects that help explain the advantages of microsimulation. Second, the advantages of microsimulation are presented.

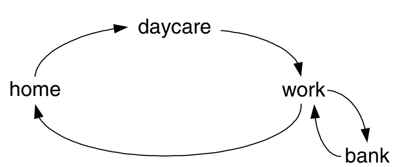
Existing Aggregate Transportation Modeling Capability of Clark County

The transportation tool that is currently available for Clark County is what is called an aggregate 4-step transportation modeling package (also known as the urban transportation modeling system or UTMS). Much of the model components were developed for RTC by Parsons Brinckerhoff Quade & Douglas in the 1990s, and these have been improved upon and ported to TransCAD by Caliper Corporation (see Caliper 2002 for more information). This type of aggregate model system has become standard practice for travel demand modeling, and some form of this model is employed by virtually all transportation planning agencies in the United States.

The basic approach of the aggregate model is to divide travel behavior into the four independent stages of trip generation (predicts the number of trips produced by and attracted to each zone), trip distribution (links productions to attractions to predict origin-destination flows), mode split (divides the flow between each origin-destination pair among different modes), and trip assignment (loads origin-destination flows on the network to estimate network link flows and level of service). The key to an aggregate application is that all data are aggregated to zones from the beginning of the model run and all models are applied to population averages.

The models predict number of trips going between zones by trip purpose (work, shop, etc.), time of day and mode, and then assign these trips to the transportation network to produce projected traffic flows on each link in the network. In such a model system, numerous things are done to simplify the problem. One critical simplification is that all travel patterns are reduced to a set of *independent* trips where a trip is defined as a journey from one location to another location. For example, say a particular person has the travel pattern shown on the left side of Figure 1. That is, she begins the day at home, drops her child off at daycare on the way to work, goes to the bank at lunch, and then returns from work at the end of the day. Such a travel pattern would be recorded in a trip-based system as 5 *independent* trips, as shown on the right side of Figure 1. The independence assumption is important, because it means that the trip-based model does not recognize that these trips are connected (or chained) in any way, and therefore cannot fully reflect the impact of trip chaining on travelers responses to transportation policies. It can also lead to confusing statistics, such as in the Phase 1 RTC Model implementation in which it appeared that not enough work trips were being generated. The issue was rather that the high amount of trip chaining on the work commute (stopping on the way to or from work) severely reduced the number of "home-based work" trips.

Travel Pattern Schematic



Representation as Independent Trips

1 "home-based other" trip	(home → daycare)
3 "non home-based" trips	(daycare → work, work → bank, bank → work)
1 "home-based work" trip	(work → home)

Figure 1: Representing Travel Patterns as Independent Trips

Another key aspect to such models is that they are applied based on average characteristics of the travelers at the zonal level. While this greatly simplifies the problem and allows straightforward model implementation, working with averages severely limits the ability of the model to capture the behavioral richness behind travel patterns (for example, trip-chaining behavior) and therefore its ability to reflect responses to transportation policy.

These two realities of independent trips and use of zonal averages are highlighted, because they represent the key differences between the STEP2 implementation and the current aggregate 4-step models. Each will be further elaborated below.

Despite these simplifications, the aggregate approach to transportation modeling has been widely implemented because it provides a practical way of reducing complex travel behavior into components that can be addressed with relatively simple techniques using reasonable data requirements. One primary reason for this is that there has not been an alternative that is available to municipalities that do not have large resources for model development. An important objective of STEP2 is to address this problem and make advanced techniques accessible for transportation policy analysis.

Advantages of Modeling at the Level of the Individual

STEP2 provides a substantially different modeling tool for Clark County than the aggregate 4-step model described above. While there are many similarities between the two models (including the basic concepts of the 4 stages as well as some overlap of models), the key difference is that the model is applied at the level of the household and not at the level of the zone. Thus, instead of applying the 4 stages at the zonal level, a population of individual households and individuals is simulated, and for each individual the number of trips by purpose and origin, destination, timing and mode are generated for each person. There are numerous advantages to working at the level of the individual.

First, note that even in the case of the aggregate 4-step model described above, most of the travel demand models were developed using data at the level of the individual. The specifications were determined and parameters or distributions of the models were estimated from survey data. There is a clear reason for this, which is that too much information is lost when data are averaged (or aggregated), and therefore it is not possible to estimate meaningful models from aggregate data. That is, the richness of the underlying population and behavior is necessary to develop the models. However, as long as these models are kept fairly simple, they can be applied on an aggregate zonal basis or to fairly course segments within each zone (for example, segmentation based on income and household size). This is the approach of an aggregate application. A microsimulation application, rather, focuses not only on retaining the behavioral richness in the model development process, but also in the model application process. In microsimulation, each model is applied individual by individual in the population, and each individual's travel patterns are simulated.

The first advantage is that the microsimulation approach supports much more realistic and intricate behavioral models. It supports complexity on the travel patterns side, for example by being able to explicitly capture behavior such as trip chaining or the concept of a work tour (stopping on the way to or from work). It also is able to incorporate any number of socio-economic characteristics of the travelers that may be driving the behavior, for example not only income and household size as in a typical aggregate model, but also the role that the individual has in the household (single parent, female married parent, male married parent), age, occupation, and any other driving forces of behavior. Put simply, the richness of both the population and the travel patterns can be retained in the application process. Retaining the richness is important for policy analysis because:

1. *Real travel patterns are complex.* Figure 2 shows that only 49% of the workers in Las Vegas reported a "simple work commute", they did not make stops on the way to or from work or in the middle of the work day.

2. *Travel patterns vary based on demographics such as gender and household structure.* Figure 2 shows that the propensity of having a complex commute depends on these factors. A single adult with one or more children is least likely to have a simple commute, whereas a household with 2 or more adults and not children is least likely to have a complex commute.

3. *The response to policies depends on details of linkages among trips and demographics.* Figure 3 displays how a person's travel pattern may be impacted by a peak period toll. Before the toll is implemented, the person has a schedule of driving from home to work during the peak period, returning during the peak period and making a stop on the way home to shop. The toll may cause any number of adjustments, 3 of which are shown in the figure. The simplicity required in models used in aggregate applications cannot directly capture such responses. The motivation of tour-based and activity-based modeling is that by explicitly capturing the impacts of trip chaining and scheduling, the model will be able to more accurately reflect responses to transportation policies.

The second advantage is that when the models are applied throughout at an individual level, they can in the end be aggregated to any desired socio-economic (and trip type) groupings. For example, classifications of income groups by ethnicity by age provide rich information on who is impacted and how by various policy decisions.

A final advantage is that a microsimulation model can generate the data necessary for input into traffic simulation models, in which individual car movements by time period and, sometimes, personal characteristics, need to be generated.

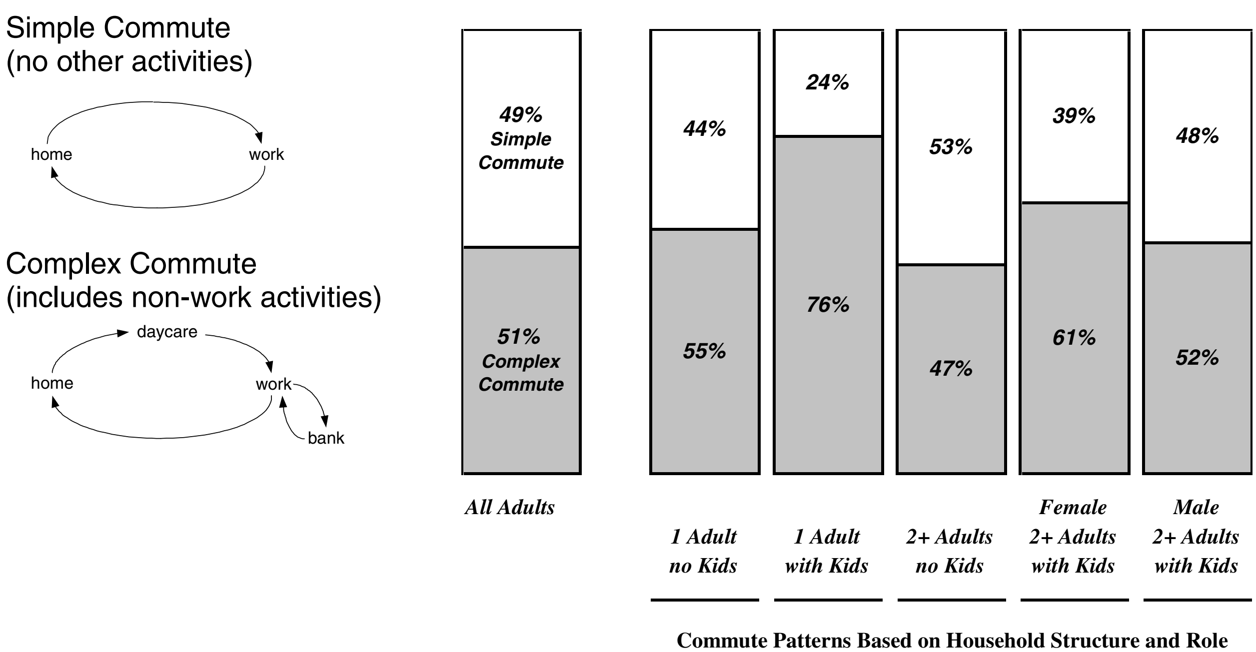


Figure 2: Complexity of Commute Patterns in Las Vegas
[from: 1996 Household Survey]

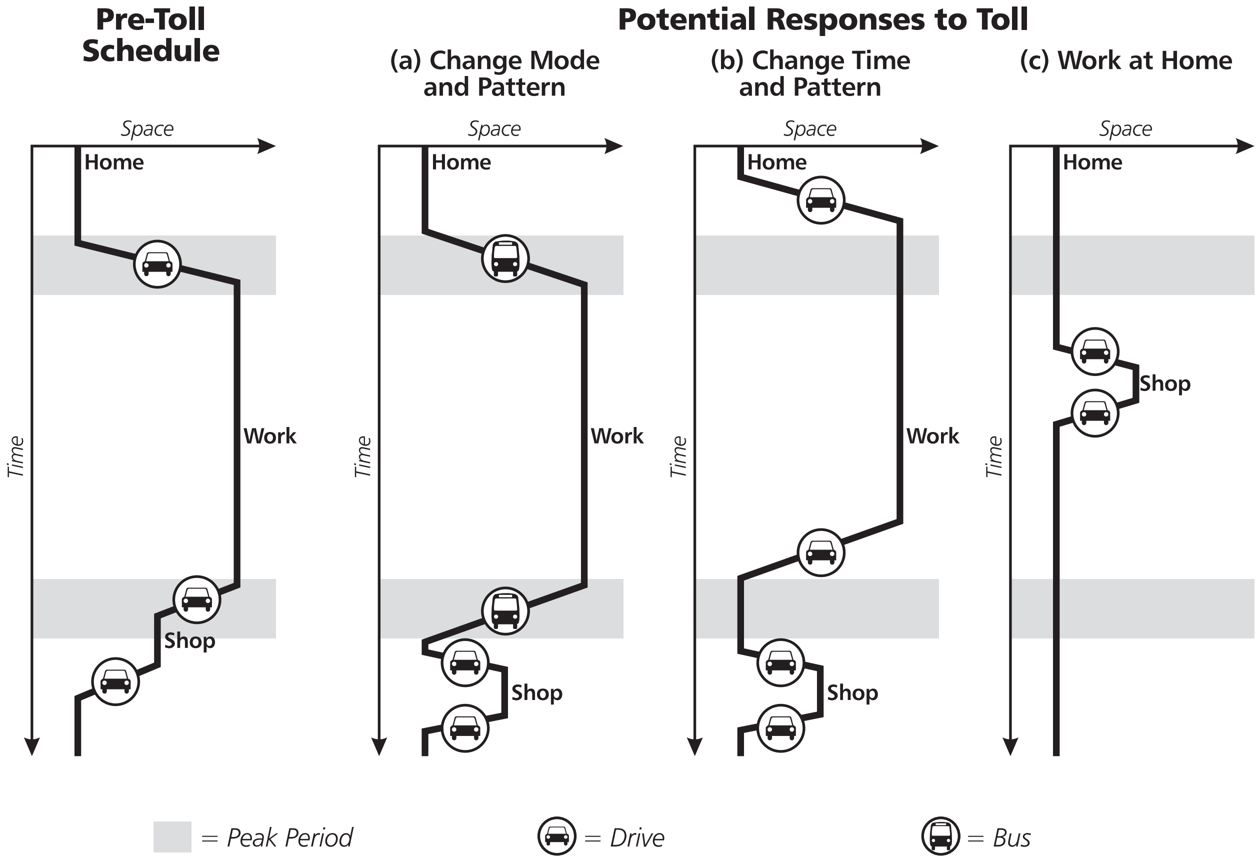


Figure 3: Ways an Individual Might Respond to a Peak Period Toll
[*adapted from Bowman, 1998*]

Overview of STEP2 for Clark County

The goal of the development effort was to develop a microsimulation-based travel demand model for Clark County. This means a set of integrated travel demand models that are applied at the individual level. The basic structure of the original STEP model is retained in STEP2, but many of the components have been modified and/or replaced with models developed directly for Las Vegas. Several new features were also incorporated that did not exist in STEP, including the use of a GIS platform, incorporating real transportation networks and traffic assignment, an aging process for aging the population to the study year, trip-chaining for the work tour, and new models for labor force participation and retirees. In addition, many of the STEP models, which were originally estimated using data from San Francisco, were replaced by models developed for Las Vegas from PUMS and the 1996 Las Vegas Household Survey. The sections below discuss the framework of the model, the model development process, and overview of how the model is run in TransCAD.

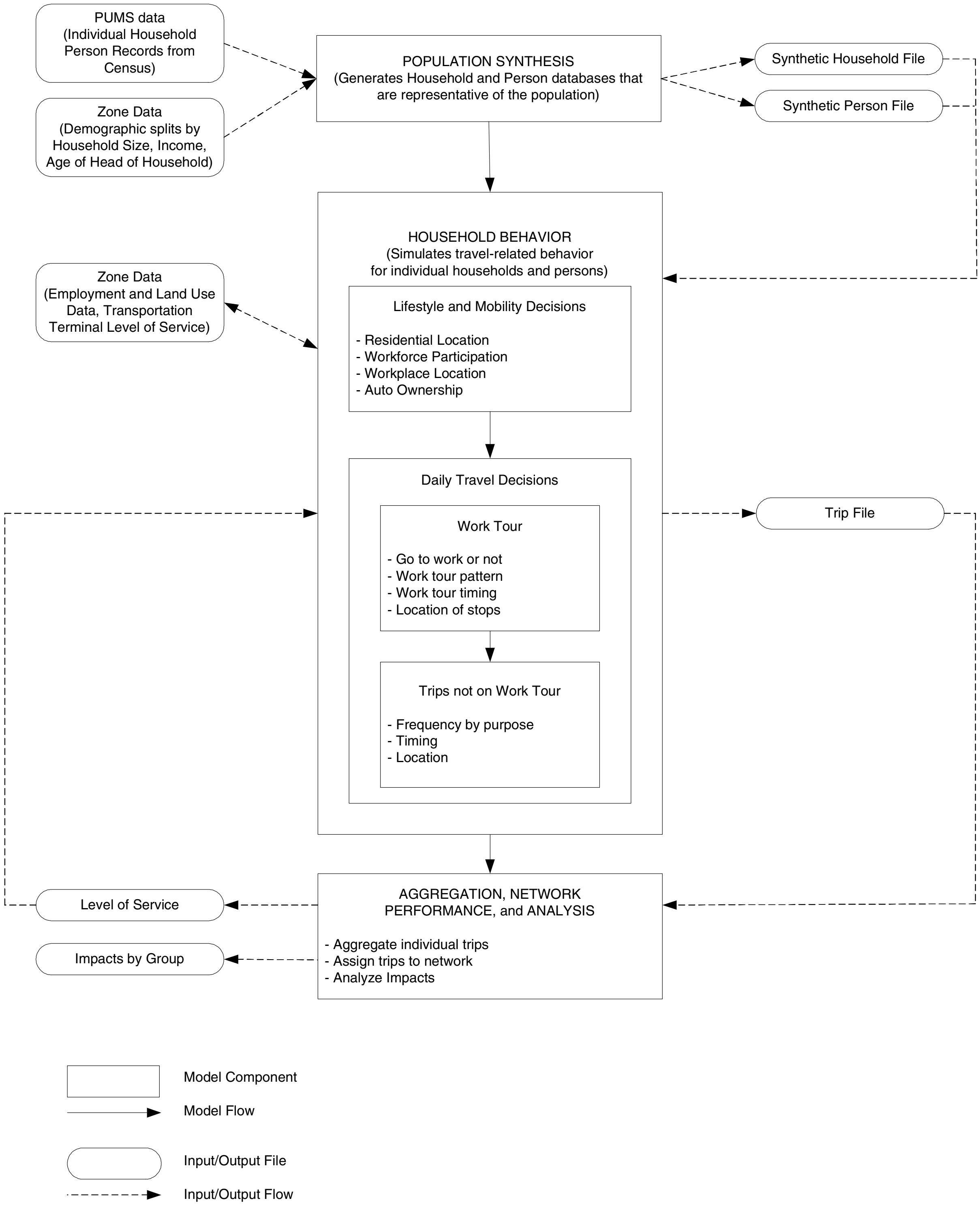


Figure 4: Model Framework for STEP2

Framework

The framework of the STEP2 model is provided in Figure 4. There are three major components to the model:

1. Population Synthesis
   Uses PUMS and aggregate zonal data to generate a representative population of specific (but synthetic) individuals and their personal and household characteristics.

2. Household Behavior
   Simulates travel-related behavior for each individual in the synthetic population. This includes both longer term decisions such as where people live and work as well as daily travel decisions of what trips to make on a given day.

3. Aggregation, Network Performance, and Analysis
   Aggregates the individual behavior to generate trip matrices by trip-type and or socio-economic characteristics and assigns these matrices to the transportation network to generate link flows and level of service.

Additional detail on the first and third stage is provided in the section below on model documentation. Most of the complexity of the model is in the second stage, and an overview is provided here. In this stage, specific travel patterns as well as other travel-related behavior are simulated for each individual in the synthetic population. For example, all of the following are simulated within STEP2 for each person:

- The TAZ in which the person's home is located
- Whether or not the person is a worker or retired
- For workers, the TAZ in which the person's work is located
- The person;s detailed travel patterns for a particular day:
  - For workers, whether the person goes to work
  - For those who go to work, whether they make a stop on the way to work, on the way home from work, or a mid-workday stop.
  - All other trips that are made in the day, including home-based shop trips, home-based other trips, and non-home-based trips
  - The location, timing, and mode of each of these trips.

In theory, all of these decisions are interrelated. For example, where one lives depends on what types of trips and activities they like to make, and therefore how accessible these activities are. Another example is that non-work trips are dependent on the characteristics and timing of the work trips. However, incorporating all of the complex, multidimensional interdependencies makes the problem unmanageable. Therefore, the approach is to simplify the problem by representing it as a sequence of choices in a choice hierarchy (see Ben-Akiva and Lerman, 1985, for more discussion). The STEP2 framework is an example of such a choice hierarchy, in which the longer term decisions are shown on the top of the hierarchy (lifestyle and mobility decisions), and the shorter-term and more flexible daily travel decisions are shown below and conditional on the long-term decisions. Furthermore, within the daily decisions, the work trip and its characteristics are shown higher in the hierarchy, and the non-work trips are conditioned below.

The household behavior component is made up of a number of integrated behavioral models. Behavior that occurs lower in the hierarchy is conditional on decisions that are made above it in the hierarchy. Furthermore, decisions that are made higher in the hierarchy can be influenced by the potential of decisions made lower in the hierarchy. For example, the STEP2 implementation assumes that the choice of residential location is made above the choice of workplace location. This means that the choice cannot assume a fixed work location. However, the residential choice model does include a variable that represents accessibility to jobs, and the value of accessibility is derived from the workplace location model via what is called the logsum value (the mathematical form will be provided later). From the residential location model, a home TAZ is simulated for each person in the synthetic population. The work location model is then run conditional on the location of the home zone. These types of conditionality and feedback are introduced throughout the STEP2 model. Increasing these types of linkages lead to more behaviorally realistic models, but also increase the complexity of model estimation and application.

Model Development

The STEP2 model described by Figure 4 is comprised of a large number of behavioral models that are used to simulate each decision in the hierarchy. These models come in one of several forms including choice models (logit and nested logit), cross-classification lookup tables, and probability distribution lookup tables generated either from PUMS or the 1996 household survey. The STEP2 implementation makes heavy use of models developed for the original STEP program and models developed for Las Vegas by Parsons-Brinckerhoff for RTC's current aggregate 4-step travel demand model. In addition, new models were developed for STEP2 using both PUMS data and the 1996 household survey. The specifics of the models are described below under model documentation.

The model is based on the 1140 TAZ structure, and uses trip purposes similar to the RTC aggregate model, including home-based shop, home-based school, home-based other, and non-home-based. In addition, the concept of trip-chaining was explicitly incorporated for the work tour, and therefore includes trip purposes specific to the journey to work (jtw) including jtw-home-work, jtw-home-other, jtw-work-other, and journey away from work (not to home).

Simulating Individual Travel Patterns from a Travel Demand Model

In most cases, the outputs of a travel demand model are the probabilities with which each available outcome will occur. For example, for a work trip there would be the probability of going to work and the probability of not going to work. For the microsimulator, these probabilities must be used to generate a particular outcome for any given individual. The basic approach is to draw a realization from the probability distribution. The method of doing this in practice is fairly straightforward and is best described by example. Say the model is a destination choice model and there are 10 alternative destination TAZs. The destination choice model would predict, for each person, the probability with which each TAZ will be selected, an example is shown in the second column in the table below. A specific choice outcome is simulated for each person by drawing a realization based on the probability distribution generated by the model. This is done by calculating the cumulative distribution function (column 3 below), generating a random number between 0 and 1 (for each person and each choice situation). The bin of the cumulative distribution function into which this random number falls is the simulated chosen alternative. In this way, if an infinite number of realizations where drawn, the probability distribution would be replicated. This procedure is used for almost all of the models described below to translate the probability distribution into realized outcomes.

Outcome	Probability Distribution	Cumulative Probability Distribution
TAZ 1	0.0353	0.0353
TAZ 2	0.0906	0.1259
TAZ 3	0.0209	0.1468
TAZ 4	0.1120	0.2589
TAZ 5	0.1120	0.3708
TAZ 6	0.0842	0.4550
TAZ 7	0.1428	**0.5978**
TAZ 8	0.0795	0.6773
TAZ 9	0.1666	0.8439
TAZ 10	0.1561	1.0000

Say the random draw is 0.555, then the simulated outcome is TAZ 7.

Figure 5: Example of Simulating a Travel Outcome for an Individual

Note that the point of this simulation is not to be able to predict accurately nor exactly any particular individual's travel patterns. Rather, the objective is to simulate travel patterns that are representative of what people actually do. It is important to incorporate in the demand models the behavior that is driving these travel patterns in terms of the impact of individual characteristics (for example age, household structure, race, and income) and to incorporate transportation level of service and other policy variables and land use characteristics of interest. This is so that the models will provide information on the impact of policy changes on demand as well as who is impacted and how.

Running the Model in TransCAD

The STEP2 model runs as an Add-in to TransCAD 4.6, and is run via a custom interface shown in Figure 6. The specific functionality of this interface is described in the STEP2 User's Guide. Each model step can be run by clicking on the model buttons. TransCAD's powerful GIS capability can then be used to visualize the results and perform additional analysis.

The Add-in allows any number of scenarios to be stored. Upon installation, all of the necessary input files are provided for the base year and the base 2020 scenario for Clark County. Scenarios are defined by a scenario name, a set of input files, output files, and model parameters. There are special features in the Add-in to assist in setting up and modifying scenarios, which are shown in Figure 7. The Project Scenarios dialog box is invoked by clicking on the Setup button from the main dialog box, and is how the scenarios are managed. Scenarios can be added, deleted, sorted, described, and renamed. The Parameter Manager dialog box is launched by clicking the Contents button in the Project Scenarios dialog box, and this is where detailed information regarding the scenario is entered and modified. From the parameter manager dialog box input and output files can be opened or changed and model parameters can be viewed or changed.

Data Requirements

The specific inputs are described in the users guide, and include

- The Public Use Microsample (PUMS)
- TAZ zonal system, including information on
  - Number of households broken down by household size, income, and age of head of household
  - Employment by employment type
  - Designation (e.g., CBD or Strip)
  - Terminal transportation information (e.g., parking costs and walk access times)
- Transportation network files for auto and transit
- Behavioral model parameters and specifications (defaults are provided)

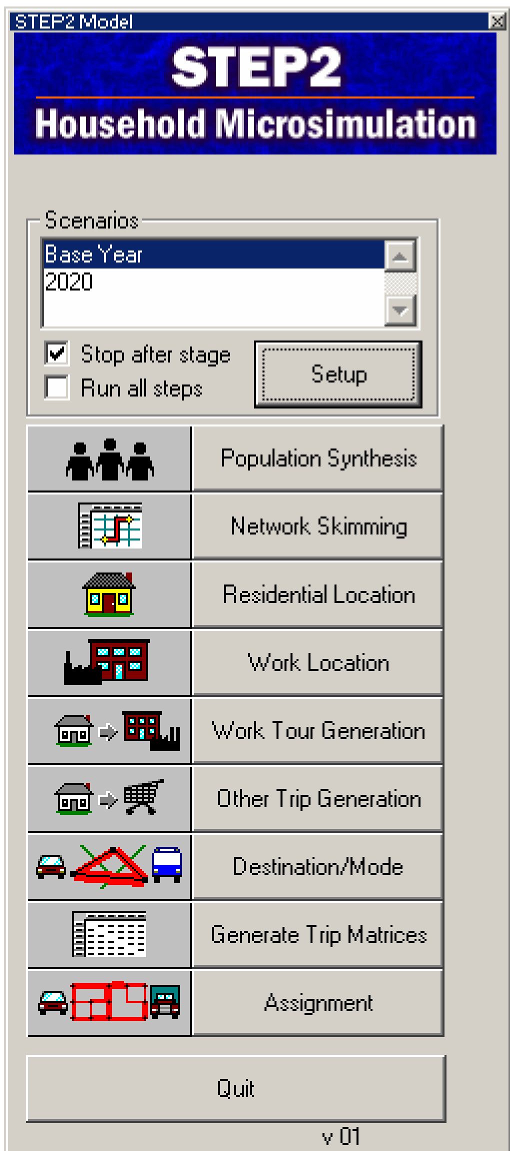
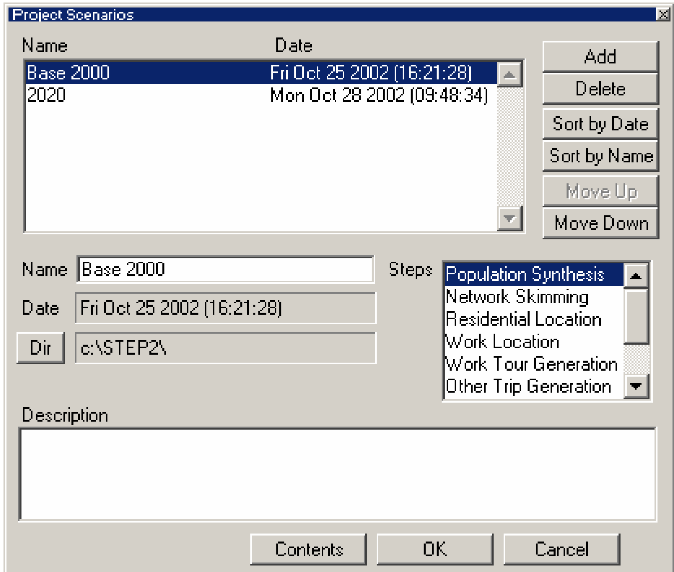


Figure 6: Custom TransCAD Model Toolbox for STEP2



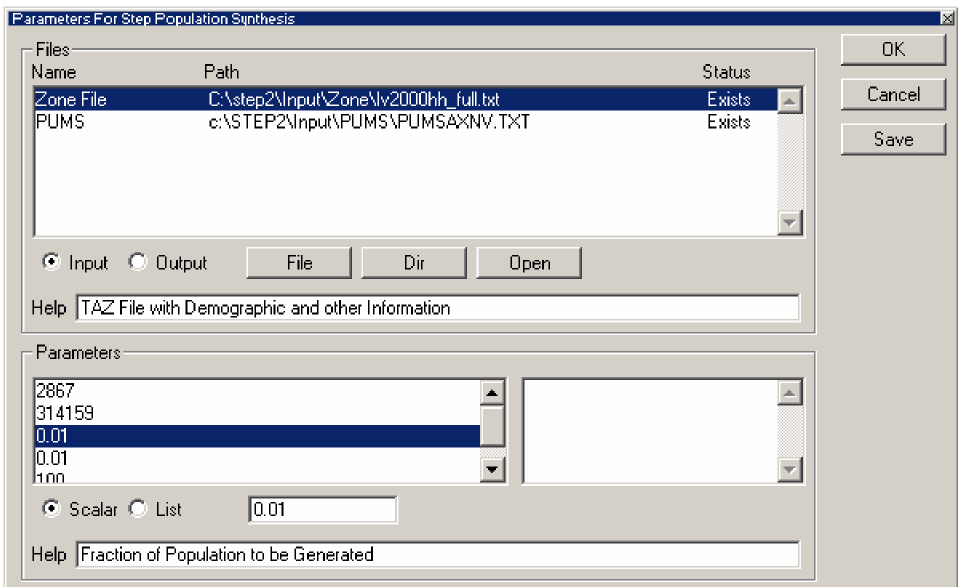


Figure 7: Scenario and File Managers for STEP2 in TransCAD

Outputs

The specific outputs are described in the user's guide, and include:

- Synthetic Household and Person databases that are representative of the population and include detailed information obtained from the PUMS databases (income, ages, occupations, race, etc.), and which has been aged from the census year to the study year.
- Simulated outputs from the Mobility and Lifestyle behavioral models on each household's TAZ of residence and each person's work status and TAZ of work location.
- Simulated trips for each person derived from the Daily Travel Decisions behavioral models, including each person's number of trips by trip purpose and each trip's origin, destination, mode and time of day.
- Aggregate forecasts of the spatial distribution of households by zone and by demographics.
- Aggregated origin-to-destination trips segmented by trip purpose, mode, time of day, income (low, moderate, middle, high), race (white, Hispanic, black, Asian, other), and many other aggregations are possible.
- Transportation network loadings and level of congestion.

Flexibility and Portability

One emphasis of this project is to develop STEP2 in a way such that it can be modified as desired for Clark County and also ported to different study areas. The scenario manager tools discussed above provides a means for making minor modifications to the parameters. In addition, TransCAD's batch mode and generic TransCAD procedures were used and or developed in many cases so that new specifications can be readily introduced. Procedures important to the implementation that that were developed are generic tools for applying nested logit (and multinomial logit) models and destination choice models at the individual level, and these interfaces are shown in Figure 8.

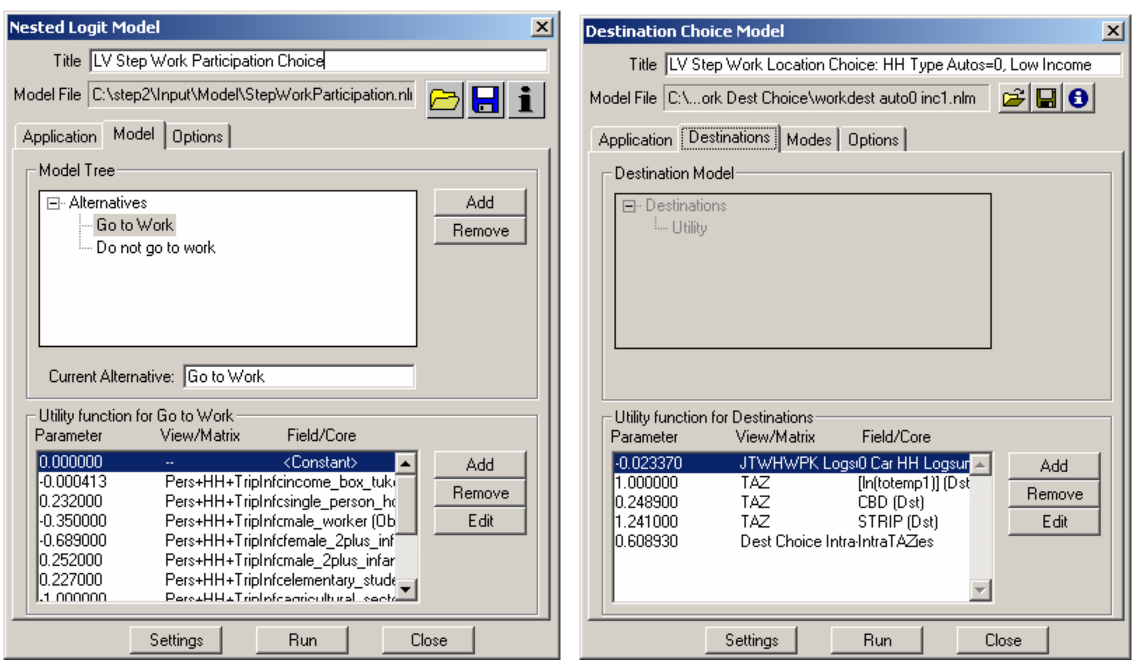


Figure 8: New TransCAD Procedures for Logit-based Microsimulation

In addition to creating generic procedures to setup and apply the behavioral models in STEP2, tools were also created to aid in processing household travel survey data to generate the input lookup tables necessary for STEP2. These include special tools to translate the activity schedule diaries reported in household surveys into trips by trip purpose and also into home-based tours (a round trip journey from home out to one or more activities and returning to home). These tools are used to generate the statistics on the types of work tours that people reported in the survey as well as the number and types of non-work-tour-based trips reported in the survey. The interfaces for these tools are shown in Figure 9.

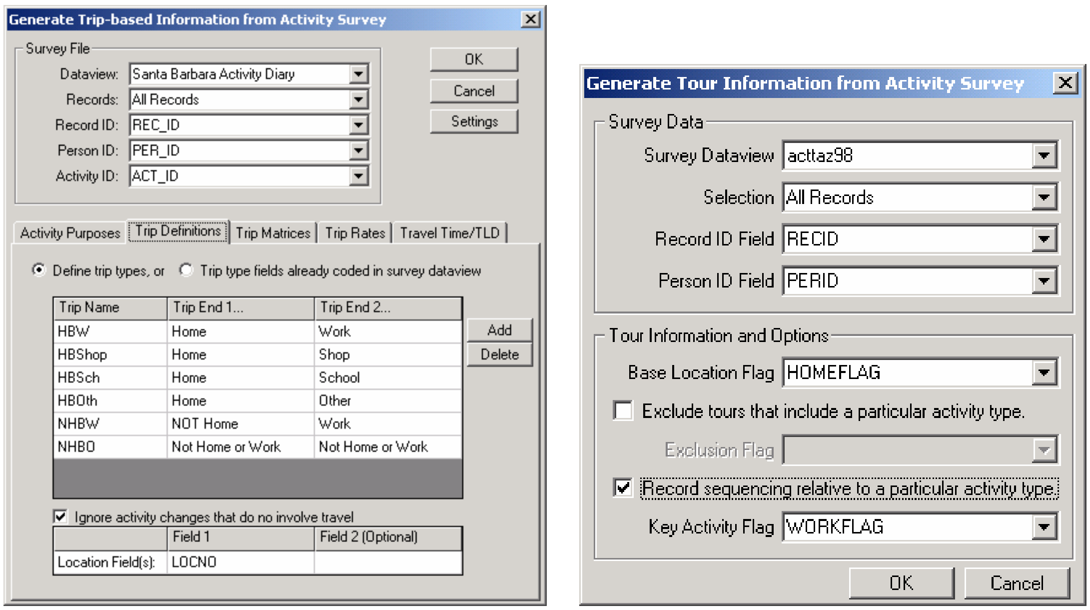


Figure 9: New TransCAD Survey Processing Tools to Generate Inputs for STEP2

Model Documentation

This document provides overall descriptions of each of the model components in the STEP2 program. The STEP2 User's Guide provides further information on the details of the input and output files and the specific procedures. Many of the models were borrowed from the original STEP and models developed by Parsons Brinkerhoff for the current aggregate model used in Las Vegas. Additional information on these models can be obtained from Harvey and Deakin (1996) and Parsons Brinckerhoff Quade & Douglas (2001). The 1996 Household survey is also used in many of the models, and more information on the survey can be obtained from Applied Management & Planning Group (1998).

Generating the Synthetic Population

Implementing a household microsimulation model requires that the model has a list of individual households and persons on which to apply the model. For each household, all of the explanatory variables that are used in the model (such as income, household size, gender and age of each person in the household) must be known. Since such a list of households is not available, simulation techniques are used to generate a synthetic population of households. The objective of the simulation is to generate a fictitious population such that it closely replicates key demographics of the real population. There are two steps to this process: synthesizing the population from PUMS to match aggregate zonal statistics, and aging the PUMS population from the year of data collection to the study year.

Population Synthesis

The technique employed here is similar to what was used in the original STEP program. It is a fairly standard technique used in transportation (see, for example, Dugway et al., 1976, for an early reference). The idea is to make use of Census data and/or aggregate demographic forecasts to generate a synthetic population consisting of individual household records and person records. Two pieces of census data are used to generate a base year population. First, SF3 data are used to obtain aggregate statistics of the population by TAZ that are to be matched by the synthetic population. For example the number of households by each income group may be tabulated for each TAZ, as well as the number of households by each household size. The PUMS data (a 5% sample of census household records) are then used to generate specific household records for each TAZ and they are generated in a way that matches the aggregate SF3 data compiled for each zone. In the STEP2 implementation, the characteristics that are matched are household size (4 categories), age of head of household (4 categories), and income (4 categories). SF3 provides the number of households for each zone and the proportion of households within each TAZ that are within each of these categories, an example of which is shown in Figure 10. The objective in population synthesis is to generate households that match these statistics.

		Proportion of TAZ Households in Various Demographic Groups											
	Number of	Income				Size of Household				Age of Head of Household			
TAZ	Households	Low	Moderate	Middle	High	1	2	3	4 plus	< 24	24-43	44-63	> 63
1108	925	29%	25%	24%	22%	17%	34%	18%	30%	2%	43%	36%	19%
357	764	34%	26%	24%	17%	29%	18%	13%	40%	13%	48%	23%	16%
112	662	23%	23%	25%	29%	8%	32%	23%	37%	2%	57%	32%	9%
707	297	24%	23%	25%	28%	26%	39%	18%	17%	3%	22%	52%	23%
1138	276	32%	25%	24%	19%	16%	18%	19%	47%	9%	56%	30%	5%
557	196	25%	24%	25%	27%	22%	44%	15%	18%	3%	29%	35%	33%
920	236	11%	19%	26%	43%	25%	39%	16%	20%	3%	31%	45%	21%
305	159	31%	25%	24%	20%	12%	31%	21%	36%	5%	49%	38%	8%
60	146	24%	23%	25%	28%	11%	34%	22%	34%	1%	55%	38%	6%
990	130	24%	23%	25%	28%	18%	36%	17%	29%	1%	46%	36%	17%
1122	76	27%	24%	24%	24%	15%	38%	18%	29%	1%	40%	47%	11%
⋮													

Figure 10: Inputs for Population Synthesis

An example of the SF3 inputs is provided in Figure 10. The first step of population synthesis is to take these marginal distributions and generate the joint distribution, which is the proportion of households in each TAZ that are within each permutation of the combination of the three demographic groups. For example, the proportion of households for each TAZ that are {low income, and have only 1 household member, and that person is under 24} through to the proportion of households the TAZ that are {high income, and have 4 or more people in the household, and the head of household is older than 63}. This procedure of generating the joint distribution from the marginal distribution is performed by a procedure called Iterative Proportional Fitting, or IPF (Deming and Stephan, 1940). It is a method that starts with a set of seed values for the joint distribution and then iteratively adjusts the values of the joint distribution to successively match each set of marginals (i.e., first adjust to income, then adjust to age, then adjust to household size, then adjust to income, etc.). The procedure eventually converges on a joint distribution that matches all marginal distributions.

Once this joint distribution is known, the PUMS household records come in. PUMS provides complete household records from the Census for 5% of the population. Data are included at both the household level and person level as shown in Figure 11.

PUMS Household Data

Household ID	Number of Persons	Home PUMA	Income	Autos ...
1001101	3	3200202	19687	2
1001102	1	3200202	37475	1
1001103	2	3200202	7050	0
1001104	3	3200204	125500	3
.				
.				
.				

PUMS Person Data

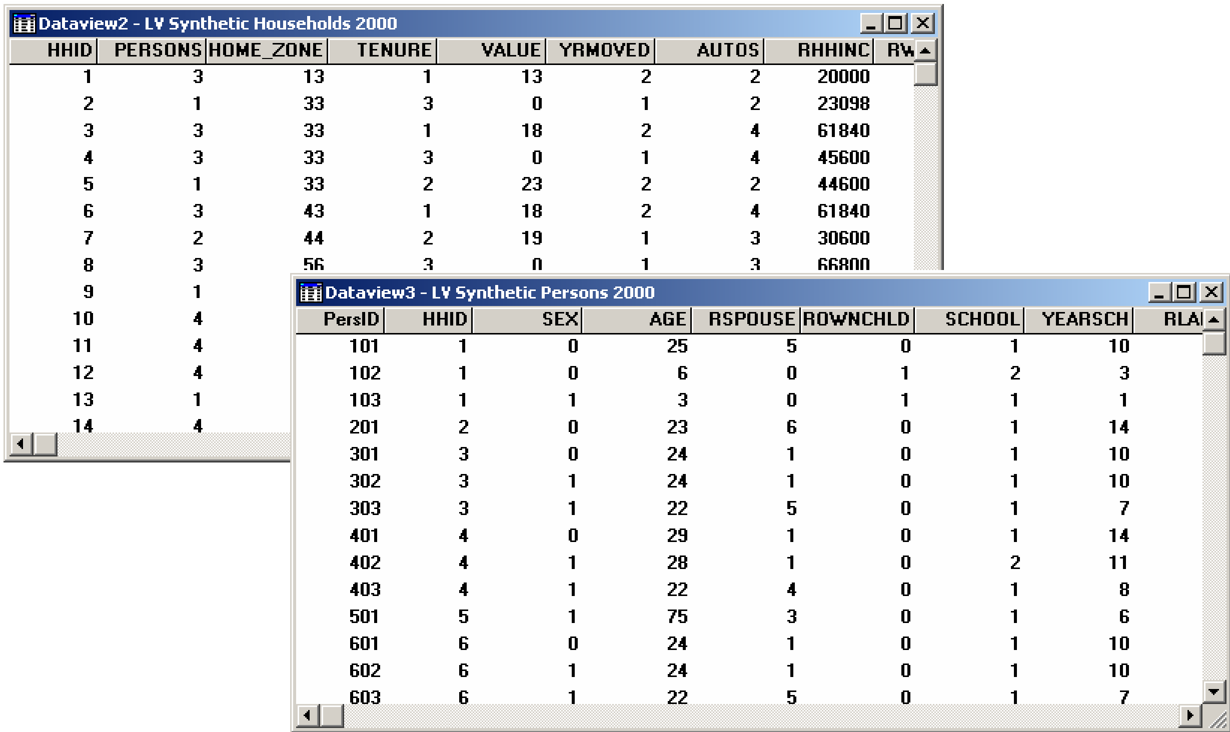
Household ID	Person Number	Relationship (1=householder)	Gender	Age ...
1001101	1	1	0	32
1001101	2	1	1	30
1001101	3	2	0	2
1001102	1	1	1	53
1001103	1	1	0	20
1001103	2	9	0	19
1001104	1	1	0	62
1001104	2	6	1	13
1001104	3	7	1	40
.				
.				
.				

Figure 11: PUMS Household and Person Data Records

The PUMS household records are used to generate actual households for each of the TAZs. The household records will not only include information on income, household size, and age of household, but also include all of the data that are included in the PUMS record (for example, occupation, gender, auto ownership, etc.). This is important as many of these other factors are used as explanatory variables in the travel demand models. Households are drawn for each TAZ according to the joint distribution of demographic factors for that TAZ, that is, so that the households in each TAZ match the input aggregate statistics for that TAZ. The result is a fictitious population of households and persons as shown in Figure 12.

For computational reasons, typically only a proportion of the population is synthesized, for example the resulting synthesized population may be a 10% sample of the entire population. Weights are calculated to expand the data to the full population. The models are applied to the 10% synthesized sample, and the resulting trips are expanded to generate travel patterns for the entire population.

Using the Census data, it is quite straightforward to generate a realistic synthetic population for census years. In order to create a synthetic population for a forecast year, the marginal distributions in need to be provided for the forecast year. This is done using output from the Clark County RTC Small Area Allocation Model, which provides forecasts for the number of households for each zone and average income per zone. Lookup distribution curves (from the PB model) are used to convert the average income in the zone to proportions of households in the zone that are in each of the 4 income categories. The distribution of household size and age of head of household are assumed to remain consistent with the census year.



Dataview2 - LV Synthetic Households 2000

HHID	PERSONS	HOME_ZONE	TENURE	VALUE	YRMOVED	AUTOS	RHHINC	RW
1	3	13	1	13	2	2	20000	
2	1	33	3	0	1	2	23098	
3	3	33	1	18	2	4	61840	
4	3	33	3	0	1	4	45600	
5	1	33	2	23	2	2	44600	
6	3	43	1	18	2	4	61840	
7	2	44	2	19	1	3	30600	
8	3	56	3	0	1	3	66800	
9	1							
10	4							
11	4							
12	4							
13	1							
14	4							

Dataview3 - LV Synthetic Persons 2000

PersID	HHID	SEX	AGE	RSPOUSE	ROWNCHLD	SCHOOL	YEARSCH	RLA
101	1	0	25	5	0	1	10	
102	1	0	6	0	1	2	3	
103	1	1	3	0	1	1	1	
201	2	0	23	6	0	1	14	
301	3	0	24	1	0	1	10	
302	3	1	24	1	0	1	10	
303	3	1	22	5	0	1	7	
401	4	0	29	1	0	1	14	
402	4	1	28	1	0	2	11	
403	4	1	22	4	0	1	8	
501	5	1	75	3	0	1	6	
601	6	0	24	1	0	1	10	
602	6	1	24	1	0	1	10	
603	6	1	22	5	0	1	7	

Figure 12: Household and Person Files Generated by Population Synthesis

Aging the Population

There is also an option in STEP2 to age the PUMS data to any future years. This ensures that the PUMS data year conforms to the rest of the data for any particular planning year. For each year that the data are brought forward, the following aging procedures are performed:

1. Age the population.
   1.1. Increase the age of each person in the household, up to a maximum of 90 years.
   1.2. Increase the education attainment of children between the ages of 6 and 17 by one and assign 6 year olds not attending school to public school, since the institution type is not used by the STEP model.
   1.3. Increase the personal income and wages using the personal income and wage growth rates. Tabulate the household income by summing the personal incomes of its members
2. Randomly eliminate some of the population, using the mortality rate by age and gender and a local scaling factor based on PUMA of residence.
3. Introduce population growth using the birth rate by age of parent and a local scaling factor based on PUMA of residence. The gender of each child is determined randomly using the proportion of male babies born. The racial and sample weight characteristics of the newborn follow the mother's characteristics.
4. Adjust households based on divorce and young adults moving out.
   4.1. Young adults who are children/stepchildren are assumed to leave home at age 22. They form a new household in the same PUMA with the same household weight. The presence of an automobile in the new household is determined randomly with probability equal to the ratio of automobiles to people in the original household. The initial industry is randomly chosen using the probabilities computed above. Next the personal income is

randomly computed using a normal distribution and aged to the correct year. The wage is assumed to be the computed fraction of income. The original household income and automobiles are reduced appropriately, except a single automobile will not be removed.

4.2. Divorce is determined based on the age and gender of the head of household for married couples, scaling using a local factor based on PUMA of residence. The income and automobiles are split, with the special case of one-car households, each getting an automobile after the divorce. The children are randomly assigned to a parent based on a custody probability based on gender. Change the marital status of divorcees to single.

5. Merge some households by marrying single men and women. A single man who will marry is found by using the marriage rate by age along with a local scaling factor for the PUMA of residence. Next, a set of prospective brides is determined by taking all single women whose age is within three years of each eligible single man initially of the same race. For each man to be married, the woman is chosen randomly from among the set of brides. If the set is empty, candidates 4 and 5 years apart are allowed. If the set is still empty, 6 years apart of any race is allowed. The household is created by merging the two former households, with a maximum of 7 cars.
6. Migration is handled by PUMA
    6.1. For inflow, the existing households are sampled for duplication.
    6.2. For outflow, the existing households are sampled for deletion.

Obviously, this procedure provides only crude estimates of the population and its characteristics. However, it is far better than what is usually done in transportation forecasting, which is to ignore the fact that the population has aged.

Demographic Rates Used for the Aging Process
Net migration rate by PUMA for 1990 to 1998
Derived from *County Population Estimates and Demographic Components of Population Change: Annual Time Series, July 1, 1990 to July 1, 1999,*
http://www.census.gov/population/www/estimates/co_99_8.html
Net migration is computed by summing the net international migration, the net federal movement, and the net domestic movement. The 1998 value is used through 2004. Rates for the whole county are used for sub-county PUMAs.

Birth rate by age of woman (0 to 90+) for 1990 to 1997
Table 4, Births: Final Data for 1997, *National Vital Statistics Reports*, Volume 47, Number 18, April 1999, http://www.cdc.gov/nchs/data/nvs47_18.pdf
The 1997 value is used through 2004.

Birth rate scaling from the national average by PUMA for 1990 to 1998
Derived from *County Population Estimates and Demographic Components of Population Change: Annual Time Series, July 1, 1990 to July 1, 1999,*
http://www.census.gov/population/www/estimates/co_99_8.html
The 1998 value is used throught 2004. Rates for the whole county are used for sub-county PUMAs.

Proportion of births which are male for 1990 to 1999
National Population Estimates by Nativity,
http://www.census.gov/population/www/estimates/us_nativity.html
The 1999 value is used through 2004.

Death rate by age (0 to 90+) and gender for 1997
Tables 2 and 3, United States Life Tables 1997, *National Vital Statistics Reports*, Volume 47, Number28, December 99, http://www.cdc.gov/nchs/data/nvs47_28.pdf
The 1997 value is used for all years.

Death rate scaling from the national average by PUMA for 1990 to 1998
Derived from *County Population Estimates and Demographic Components of Population Change: Annual Time Series, July 1, 1990 to July 1, 1999,* http://www.census.gov/population/www/estimates/co_99_8.html
The 1998 value is used throught 2004. Rates for the whole county are used for sub-county PUMAs.

Marriage rate by age (0 to 90+) and gender for 1990
Table 6, Advance Report of Final Marriage Statistics, 1989 and 1990, *Monthly Vital Statistics Report,* Volume 43, Number 12 Supplement, July 1995,
http://www.cdc.gov/nchs/data/mv43_12s.pdf
The 1990 value is used for all years.

Marriage rate scaling from national average by PUMA for 1988
Table VS, *USA Counties 1998,* http://www.census.gov/statab/www/county.html
The 1988 value is used for all years. Rates for the whole county are used for sub-county PUMAs.

Divorce rate by age (0 to 90+) and gender for 1990
Table 5, Advance report of Final Divorce Statistics, 1989 and 1990, *Monthly Vital Statistics Report,* Volume 43, Number 9 Supplement, April 1995,
http://www.cdc.gov/nchs/data/mvs43_9s.pdf
The 1990 value is used for all years.

Divorce rate scaling from national average by PUMA for 1988
Table VS, *USA Counties 1998,* http://www.census.gov/statab/www/county.html
The 1988 value is used for all years. Rates for the whole county are used for sub-county PUMAs.

Proportion of divorces which the father receives custody for 1989
Table 18, Advance report of Final Divorce Statistics, 1989 and 1990, *Monthly Vital Statistics Report,* Volume 43, Number 9 Supplement, April 1995,
http://www.cdc.gov/nchs/data/mvs43_9s.pdf
The 1990 value is used for all years.

Personal Income Growth rate by PUMA for 1990 to 1996

Table CA1-3, Local area personal income, http://www.bea.doc.gov/bea/regional/reis/ca1_3.htm
The 1996 value is used through 2004. Rates for the whole county are used for sub-county PUMAs.

Wage rate growth by PUMA for 1990 to 1996
Table CA34, Local area personal income,
http://www.bea.doc.gov/bea/regional/reis/ca34/index.htm
The 1996 value is used through 2004. Rates for the whole county are used for sub-county PUMAs.

Geographic Databases, Network Development and Network Skimming

Important inputs to the travel behavior models and network assignment models are the transportation networks and level of service matrices (containing travel times and costs between zones). The original STEP models did not have an internal transportation network representation or traffic assignment model, so changes in level of service resulting from changes in demand had to be generated through some other transportation modeling process. At the time, this was done for computation reasons, primarily so that STEP could maintain its quick response capability. With increases in computational power and inexpensive data storage, such limitations are no longer necessary, and the Caliper STEP2 implementation has included internal representation of transportation networks and calculation of transportation level of service characteristics.

The geographic databases and transportation networks and processing used for STEP2 were created for the Phase 1 RTC Model for TransCAD. The *RTC Phase 1 Model Documentation* describes the transportation networks, network creation, and the process of generating zone to zone travel time matrices in detail, and a brief summary is provided here. The highway networks and transportation analysis zones are based on the 1140 zone structure. The highway networks in the region were vastly improved as part of the Phase 1 RTC Model project, including "conflating" the pre-existing stick network to reflect true geography and link distances. The data at the zonal level (and included in the TAZ) geography were obtained from Census SF1, Census SF3, and output from the Clark County Regional Transportation Commission Small Area Allocation Model. Transit route systems are also necessary in order to generate the transit level of service variables.

As described above, an important improvement to the original STEP program is that STEP2 runs in TransCADs GIS environment, which provides the advantage that model outputs can be analyzed visually using all of the capabilities of a powerful GIS. Therefore, geographic files of both the TAZs and transportation networks are, of course, included.

Travel Demand Models

As described above, the travel demand models can be separated into the more long-term lifestyle and mobility decisions, which include household and job location decisions, and more short term daily travel decisions, which include whether to go to work on a particular day and where and when to make other trips on a particular day.

Long-term Lifestyle and Mobility Decisions

Residential Location

A key input to STEP2 are the land use variables produced by the *Clark County Regional Transportation Commission Small Area Allocation Model (SAAM)*. This model produces by TAZ forecasts of population and housing units (by type and occupancy) and employment by industry. The population synthesizer generates households residing in each TAZ based on the total number of households forecast by the SAAM. While this is already a forecast of the spatial patterns of households, STEP2 provides additional capabilities for fine tuning this forecast by incorporating more detailed representation of accessibility and level of service. This model can be used to adjust the SAAM forecast, but perhaps is most useful to determine residential land use shifts under different future scenarios.

The residential choice model included in STEP2 is similar to the specification in the original STEP model, but has several modifications. The utility derived from living in each zone is as follows:

$$
\begin{aligned}
U_i^s = {} & \beta_1 (\ Job\ Accessibility_i^s\ ) \\
& + \beta_2 (\ Average\ home\ ownership\ cost\ in\ relation\ to\ household\ income_i^s\ \text{-}\ home\ owners\ only\ ) \\
& + \beta_3 (\ Average\ rental\ cost\ in\ relation\ to\ household\ income_i^s\ \text{-}\ renters\ only\ ) \\
& + \beta_4 (Re\ sidential\ density\ in\ housing\ units\ per\ acre_i\ ) \\
& + \beta_5 (\ Jobs\ per\ employed\ resident_i\ ) \\
& + \beta_6 (\ Manufacturing\ jobs\ per\ acre_i\ ) \\
& + \beta_6 (\ Violent\ crimes\ per\ capita_i\ ) \\
& + ln(\ Number\ of\ housing\ units\ (owner\ occupied\ or\ rental,\ depending\ on\ household)_i\ )
\end{aligned}
$$

The utility is a function of accessibility, prices (in relation to household income), local environment and job availability, and safety. The residential choice model is not conditional on job location, and the job accessibility variable is calculated as the logsum of the work location model for the head of household. The utility specification is a function of socio-economic characteristics (denoted by the *s* superscript) including income and owner or renter status of the household. The log of the number of housing units (often called the size variable) is necessary in a residential choice model to account for the varying number of units across zones. The higher the number of units in any zone, the higher the probability that a household will live in that zone.

In addition, there is a major twist to the residential choice model. The issue is that it is not desirable to move all of the households after the population synthesis stage. Population synthesis was performed by matching projected marginal distributions of income, household size, and age of head of household for each zone. It is assumed in STEP2 that these are based on reasonable forecasts, and the desire of the residential choice model internal to STEP2 is more to fine tune the spatial distribution and to allow for adjustments to projected scenarios. Therefore, the model needs to be adjusted so that the probability of a household being moved to a different zone is decreased. This is done in STEP2 by using the utility as specified above for all zones *except* for the zone in which the household currently resides. For this zone, we have additional and critical information related to the household's behavior, which is that the household already "selected" this zone from all available zones. Therefore, according to the theory, this zone has the maximum utility for this household. The true value of this maximum utility cannot be known, because there is an unknown additive error in all utility equations that cannot be directly measured. However, it is possible to quantify the *expected* maximum utility of any given choice situation, which is equal to the "logsum" or the log of the denominator of the choice probability ($ln \sum_j exp(U_j)$). Therefore, the utility for the zone to which the household was originally assigned during population synthesis is given this value of the logsum (or the expected maximum utility). This in effect increases the utility of the current zone, and therefore increases the probability with which the household remains in the originally assigned zone.

With the utilities for all zones except the household's current residential zone taking the utility as specified above, and the utility of the household's current residential zone having a utility set equal to the logsum, the probability of a household residing in a particular zone is equal to the logit probability:

$$P_n(i) = \frac{exp(U_i^s)}{\sum_{all\ TAZ\ t} exp(U_t^s)}$$ where person *n* is in segment *s*.

From the resulting probability vector, a specific residential zone is simulated for each household.

Labor Force Participation and Retirees

Models to predict workforce participation were developed by Caliper using PUMS data and 1996 Las Vegas Household Survey. These models predict the probability that an adult is a worker, and, for non-workers, whether the person is retired or not. Since these models are not documented elsewhere, a higher level of detail is provided here. There are two models. First, the workforce participation model is applied to determine whether an individual is a worker or not. Whether or not a person is a worker is influenced by the person's gender and household structure (married or not, children or not and what age) and the person's age and race. This model was estimated using 1990 PUMS data primarily because PUMS includes information on race, which is thought to be a significant factor in labor force participation. Figure 11 provides the estimation results. The model is a binary logit model, and the parameters provide the specification of the utility of being a worker (versus not being a worker), so the utility of working is specified to be:

Workforce Participation Model ***Dependent variable: worker (versus non-worker)***	Data: 1990 PUMS	
Explanatory Variables	Parameter	T-stat
Constant	-13.4	-12.2
Gender and Household Structure:		
Single Female with Children under 18	-0.540	-4.4
Single Female with Children under 6	-0.811	-5.3
Single Female without Children under 18	-0.432	-7.5
Single Male with Children under 18	-0.563	-6.1
Single Male with Children under 6	-0.104	-0.9
Single Male without Children under 18	-0.413	-6.9
Married Female with Children under 18	-1.08	-14.0
Married Female with Children under 6	-0.688	-8.5
Married Female without Children under 18	-1.01	-18.2
Married Male with Children under 18	0.567	5.7
Married Male with Children under 6	0.109	0.8
Married Male without Children under 18	0.000	---
Age (Piecewise Linear)		
Years 16 to 18	0.803	12.8
Years 18 to 25	0.177	13.3
Years 25 to 35	-0.0144	-1.8
Years 35 to 45	-0.0141	-1.8
Years 45 to 55	-0.0562	-6.9
Years 55 to 65	-0.194	-22.7
Years 65 to 75	-0.136	-10.5
Years beyond 75	-0.167	-4.8
Race		
Black	-0.470	-8.9
Asian	0.107	1.3
Hispanic	0.149	1.7
Other	-0.190	-2.4
White	0.000	---
Number of observations *	26421	
Log-likelihood (beta)	-12942	
Log-likelihood (zero)	-18314	
Number of parameters	24	
Model fit (rho-bar squared)	0.292	

* PUMS 16 years of age and older

Retiree Model ***Dependent Variable: Retired (versus not retired)***	Data: 1996 HH Survey	
Explanatory Variables	Parameter	T-stat
Constant	-6.92	-2.0
Gender and Household Structure:		
Single Female with Children under 18	-2.06	-4.1
Single Female without Children under 18	-1.40	-4.0
Single Male with Children under 18	-1.84	-2.2
Single Male without Children under 18	-0.970	-2.7
Married Female with Children under 18	-2.85	-4.8
Married Female without Children under 18	-2.21	-6.9
Married Male with Children under 18	-1.24	-2.4
Married Male without Children under 18	0.000	---
Age (Piecewise Linear)		
Years 16 to 35	0.124	1.1
Years 35 to 45	0.202	2.1
Years 45 to 55	0.164	3.3
Years 55 to 65	0.220	5.8
Years 65 to 75	0.109	2.6
Years beyond 75	-0.052	-1.2
(Race not available in HH Survey)		
Number of observations *	1483	
Log-likelihood (beta)	-373	
Log-likelihood (zero)	-1028	
Number of parameters	14	
Model fit (rho-bar squared)	0.624	

* LV HH Survey 16 years of age and older Non-workers
Retiree status not available in PUMS

Figure 13: Workforce Participation and Retiree Model Estimation Results

$$
\begin{aligned}
U_n = & -13.4 \\
& -0.540(\textit{ Single Female with Children under 18 dummy }) \\
& -0.811(\textit{ Single Female with Children under 6 dummy }) \\
& \textit{and so on, as specified in the estimation results.}
\end{aligned}
$$

The probability of an individual being a worker is then:

$$P_n(\,worker\,) = \frac{1}{1 + exp(-U_n)}$$

This probability is used to simulate whether a person is a worker or not. For non-workers, an additional model is applied to determine whether the non-worker is retired or not. Retirees are an important demographic group and may have significantly different travel patterns than other non-workers. Since PUMS does not provide information on retirement status, the 1996 household survey was used for this model. The specification of the utility and estimation statistics for the model are shown in Figure 13. Retirement status is shown to be a function of gender and household structure as well as age. Applying this model to each non-worker in the synthetic population produces a probability that the person is retired or not, and from this probability an actual retiree status (retired or not) is simulated for the person.

Work Location

A work destination choice model is used to determine the work TAZ for each worker in the synthetic population. The Parsons Brinckerhoff Las Vegas destination choice models for the home-based work trip purpose were used in STEP2. While the model is the same, the application is very different. In an aggregate model, the destination choice model is applied on a zone-to-zone basis. In the microsimulation framework, the model is applied at the level of the individual worker, and a specific work TAZ for each worker is determined. The assumption is that each worker will choose the work TAZ that maximizes his or her utility. The utility (according to the PB model) of making a home-based work trip from origin TAZ *o* to destination TAZ *d* for population segment *s* is defined as:

$$
\begin{aligned}
U_{od}^{s} = {} & \beta_1^s (\mathit{transportation\ level\ of\ service}_{od}^{s}) \\
& + \beta_2^s (\mathit{CBD\ dummy}_d) \\
& + \beta_3^s (\mathit{STRIP\ dummy}_d) \\
& + \beta_4^s (\mathit{INTRATAZ\ dummy}_{od}) \\
& + \ln(\mathit{HBW\ attractions}_d^s)
\end{aligned}
$$

The probability of person *n* working in TAZ *d* is then:

$$P_n(d) = \frac{\exp\left(U^{s}_{\text{home zone of } n,d}\right)}{\sum_{all\,TAZ\,t} \exp(U^{s}_{\text{home zone of } n,t})} \qquad \text{where person } n \text{ is in segment } s.$$

A different set of parameters β are used for each population segment *s*. The population is segmented based on household income and auto ownership. The level of service that is used in the model is the log of the denominator of a mode choice model (sometimes called the logsum or expected maximum utility), which also varies by population segment. Using the logsum makes the work location choice decision sensitive to both auto level of service and transit travel times and costs. Furthermore, the logsum does so in a way such that the transit level of service is significant in the model only for trips and market segments in which transit competes with auto. The CBD, STRIP, and INTRATAZ dummies have positive parameters β and act to increase the number of workers in the central business district and strip TAZs and also workers who work in the same TAZ in which they live. The HBW attractions term is an output from the PB trip attraction model and represents the number of workers in each TAZ by population segment. This term (often called the size variable) is necessary in a destination choice model to account for the varying number of jobs across zones. The higher the number of jobs in any zone, the higher the probability that a worker will work in that zone. The parameters and other details of the specification are described in the *RTC Phase 2 Model Documentation*.

In STEP2, for each worker, the probability with which the worker will work in each TAZ is calculated using the above equation. Then an actual TAZ is simulated for the person by drawing a TAZ from the probability distribution as described above.

Other Lifestyle and Mobility Decisions

STEP2 currently includes models described above as the key lifestyle and mobility determinants. All other lifestyle decisions are assumed to be consistent with the household record drawn from the PUMS data set during population synthesis. However, many of these variables could also be modeled directly. At the top of the list of variables relevant for transportation analysis are models of auto ownership. Both the Parsons Brinckerhoff and original STEP models include auto ownership models that can be built in to STEP2.

Daily Trip Decisions

The specific travel patterns that an individual is projected to make during a day are separated into several components. The first two steps of work participation, work tour, and nonwork trip frequency are the disaggregate (or microsimulation) equivalent to the trip generation step in the aggregate models. However, rather than working with average trip making behavior of groups of people (as in a standard cross-classification trip rate look up table), actual realizations of trips must be simulated. For example, rather than assigning an average of 2.268 home-based work trips to all 3 person households, a microsimulation approach models whether or not each particular worker in the synthetic population goes to work or not. Not only do integer values of work trips have to be assigned to each worker, but integer values of all of the trip types (home-based shop, home-based school, etc.) need to be assigned to each person. A valuable aspect of the disaggregate approach is that concepts such as trip chaining can be explicitly included in the model, as is done here with the work tour model. Thus, it is not only modeled that a particular worker goes to work, but also whether he or she makes a stop on the way to or from work. In this initial version of STEP2, the trip generation is done via lookup tables generated from the Las Vegas 1996 household survey. Future versions will introduce discrete choice models that are more of a function of socioeconomic characteristics in order to more accurately capture behavior.

Along with knowing the number of trips by trip type made by each person in the synthetic population, the time of day, destination, and mode must also be modeled. Time of day is done through lookup tables using time of day histograms generated from the survey. Destination and mode are determined from a destination choice model.

Travel is simulated for all persons in the synthetic dataset who are 16 years of age or older.

Work Participation and Tour

This model step involves two components. The first is whether a worker goes to work or not on the simulated day. For this model, the original STEP work participation model is used. It is a binary choice model (go to work or do not go to work), where the “go to work” utility is specified as:

$$\begin{aligned}U_n = {} & \beta_1 (\ Box - Tukey\ transform\ of\ household\ income_n\ ) \\ & + \beta_2 (\ Live\ Alone\ dummy_n\ ) \\ & + \beta_3 (\ Male\ dummy_n\ ) \\ & + \beta_4 (\ Female\ with >= 2\ infants\ dummy_n\ ) \\ & + \beta_5 (\ Male\ with >= 2\ infants\ dummy_n\ ) \\ & + \beta_6 (\ Elementary\ School\ Students\ in\ HH\ dummy_n\ ) \\ & + \beta_7 (\ Employed\ in\ Agriculture\ Sector\ dummy_n\ ) \\ & + \beta_8 (\ Employed\ in\ Construction\ Sector\ dummy_n\ ) \\ & + \beta_9 (\ Employed\ in\ Transportation\ Sector\ dummy_n\ ) \\ & + \beta_{10} (\ Employed\ in\ Service\ Sector\ dummy_n\ ) \\ & + \beta_{11} (\ Peak\ Highway\ Travel\ Time\ to\ Work_n\ /\ Free\ Flow\ Highway\ Travel\ Time\ to\ Work_n\ )\end{aligned}$$

The probability of any individual worker *n* going to work is then:

$$P_n(\ go\ to\ work\ ) = \frac{1}{1 + exp\left(-U_n\right)}.$$

This probability is used to simulate whether or not the person goes to work or not. A random number is drawn between 0 and 1. If the random number is less than the probability, then the person goes to work, and otherwise the person does not go to work.

For those workers who go to work, the second component is to determine what type of work tour they have. A work tour is defined as a round trip journey from home out and about to one or more activities (one of which is work) and returns to home. As discussed above (Figure 2), the majority of the people in the Las Vegas area do not have a simple home to work to home commute, because they stop either on the way to or from work and make mid-workday trips. STEP2 captures this trip-chaining behavior by assigning particular work tour patterns to each worker. The 1996 household survey was used to generate the distribution of work tour types, and the distribution is shown in Figure 14. Home-based work tour patterns were restricted to 8 patterns, which result from the combinations of stopping before work or not, stopping after work or not, and making a mid-workday trip. This representation is a simplification in many ways, including that people may make multiple stops on the way to or from work or in the middle of the workday; a worker may have two home-based work tours (for example, if he or she goes home for lunch), workers may visit more than one work location, and workers may have incomplete work tours on any given day (where they either do not start the day at home or finish the day at home). This specification can be enhanced in future versions of STEP2. In addition, it would be desirable to incorporate a behaviorally rich specification, such as the labor force participation model and work participation model described above, which reflects the characteristic of the person as well as transportation level of service.

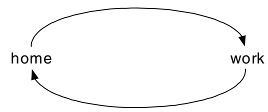
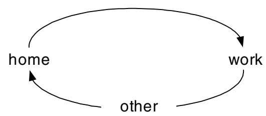
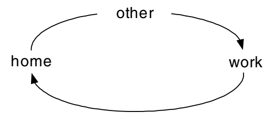
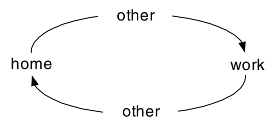
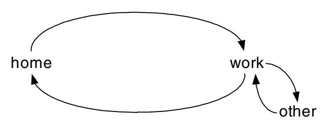
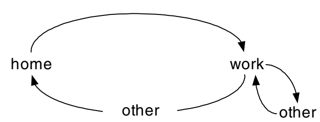
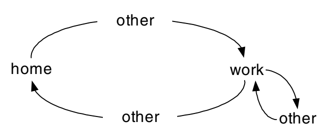
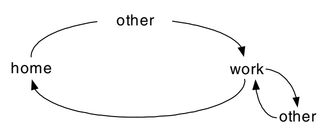
Work Tour Pattern	Probability Distribution of Work Tours
	59.8%
	16.2%
	7.3%
	6.2%
	5.4%
	2.5%
	1.4%
	1.2%

Figure 14: Distribution of Home-based Work Tour Type
[from: Work Tour Pattern Lookup.bin]

Non-Work Tour Trip Frequency

Along with simulating the work tour as described above, it is also necessary to simulate trips made outside of the work tour. School trips are dealt with independently from home-based shop, home-based other, and non-home-based trips.

The PUMS household record provides information on whether a person is a student or not. The 1996 household survey indicates that 86.1% of high school students go to school on any weekday, and 42.5% of post-high school students go to work on any school day. Therefore, students in the synthetic dataset who are 18 years old or younger are assigned with an 86.1%

probability 2 home-based school trips (home to school and school to home). Students in the synthetic dataset who are over 18 are assigned with a 42.5% probability 2 home-based college trips (home to college and college to home).

The home-based shop, home-based other, and non-home based trips (not part of the work tour) are simulated in STEP2 via another lookup table that represents various combinations of such trips that were reported in the 1996 household survey. What is simulated for each person is the number of home-based shop trips, number of home-based other trips, and number of non-home-based trips, where these numbers exclude any trips made on the work tour. For these non-work tour trips, trip chaining is not explicitly captured, although this could be incorporated in a future version of STEP2. A different distribution is used for people who make a work trip and people who do not. A portion of the lookup table is provided in Figure 15.

Potential Combinations of Non-work Tour Trips			Probability Distribution	
Number of Home-based Shop Trips	Number of Home-based Other Trips	Number of Non-home-based Trips	Persons without a Work Tour	Persons with a Work Tour
0	2	0	19.4%	48.3%
2	0	0	7.7%	15.4%
0	2	1	6.9%	6.8%
0	4	0	7.6%	4.8%
1	1	1	5.5%	4.2%
0	2	2	4.3%	4.0%
0	1	0	2.9%	3.3%
0	4	1	2.7%	1.6%
2	2	0	3.9%	1.5%
1	1	2	3.1%	1.4%
1	1	3	1.0%	1.1%
1	3	1	2.7%	1.0%
	.		.	.
	.		.	.
	.		.	.

Figure 15: Distribution of Non-work Tour Trips
[from: Non Work Tour Trip Rate Lookup.bin]

Time of Day

The time of day procedure models the hour in which each trip that the person makes commences. The general approach used is to generate trip time histograms (probability distributions) from the 1996 household survey, and then simulate realized trip times for each person from these probability distributions. This is similar to the aggregate implementation in the Phase 1/2 RTC Model, and some of the same lookup tables are used. The key difference here is that a specific time (hour) is simulated for each trip. Note that this implementation does not take into account the impact of congestion on travel time choice, but such an implementation will be introduced in future versions of STEP2 using a behavioral time of day choice model that is a function of level of service and characteristics of the trip (e.g., purpose) and traveler (e.g., income).

For work tours, first the start time and end time of the work day are determined. This is done via a lookup table based on distributions reported in the survey, and a start and end time is simulated for each worker with a work trip. A portion of the lookup table is shown in Figure 16. The most common work day reported in the survey is to begin the trip to work between 7 and 8 AM and leave work between 5 and 6 PM. The start time is applied to the worker's first trip on the work tour that has work as its destination. The end time is applied to the worker's last trip on the work tour that has work as its origin.

If there is a stop on the way to work, a lookup table is used to determine how many hours before the work trip started does the work journey start (that is, the first trip in the work tour). Similarly, if there is a stop on the way home from work, a lookup table is used to determine how many hours after the person left work does the trip home start (that is, the last trip in the work tour). The lookup table is shown in Figure 17, and shows that the majority of stops on the way to work are quick (e.g., drop offs or quick errands), whereas stops on the way home from work tend to have a longer duration.

Start Hour of First Trip with Work Destination	Start Hour of Last Trip with Work Origin	Probability Distribution
7 AM	5 PM	9.5%
6 AM	3 PM	5.8%
7 AM	4 PM	5.8%
6 AM	4 PM	5.4%
8 AM	5 PM	4.6%
5 AM	2 PM	3.6%
6 AM	5 PM	3.1%
7 AM	3 PM	2.7%
7 AM	6 PM	2.6%
5 AM	4 PM	2.4%
5 AM	3 PM	2.0%
6 AM	6 PM	2.0%
8 AM	4 PM	1.9%
9 AM	6 PM	1.8%
8 AM	6 PM	1.7%
5 AM	5 PM	1.5%
9 AM	5 PM	1.2%
6 AM	2 PM	1.2%
7 AM	2 PM	1.1%
9 AM	7 PM	1.0%
2 PM	11 PM	0.9%
3 PM	12 PM	0.9%
4 AM	2 PM	0.9%
6 AM	1 PM	0.9%
.		
.		
.		

Figure 16: Work Start and End Time Lookup Table
[from: Work Time Start End Lookup.bin]

Hour Shift *	Probability Distribution	
	Before Work Stop	After Work Stop
0	54.23%	28.79%
1	27.45%	35.25%
2	9.97%	13.77%
3	3.17%	9.05%
4	2.24%	4.54%
5	2.94%	3.37%
6	0.00%	1.07%
7	0.00%	1.53%
8	0.00%	2.62%

* Absolute Value Of (Start hour of trip with work end
- Start hour of trip without work end)

Figure 17: Before Work Stops and After Work Stops Timing Lookup Table
[from: Work Before After Time Shift Lookup.bin]

Trip Hour	*Time of Day Probability Distributions based on Trip Type and Direction of Trip*										
	Work-based Midday		Home-based School		Home-based College		Home-based Shop		Home-based Other		Non-home-based
	From Work	To Work	From Home	To Home	From Home	To Home	From Home	To Home	From Home	To Home	
3	0.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.5%	0.0%	0.5%	0.2%	0.1%
4	0.0%	0.0%	1.0%	0.0%	2.4%	0.0%	1.3%	0.0%	0.0%	1.2%	0.0%
5	1.4%	0.0%	4.7%	0.0%	5.5%	0.0%	0.6%	0.1%	0.5%	3.6%	0.5%
6	0.4%	0.0%	25.8%	0.4%	20.8%	0.0%	2.2%	0.2%	1.1%	6.3%	1.5%
7	3.4%	0.0%	35.9%	0.2%	28.7%	0.0%	5.4%	0.8%	2.4%	9.1%	3.4%
8	5.8%	1.3%	25.1%	0.9%	15.1%	0.0%	7.6%	2.1%	5.1%	10.9%	5.1%
9	8.2%	4.0%	0.3%	0.6%	6.1%	5.2%	11.9%	3.8%	2.8%	7.5%	7.2%
10	8.9%	7.7%	0.9%	1.3%	5.5%	0.0%	9.0%	6.7%	5.0%	5.2%	10.0%
11	33.1%	7.8%	1.6%	4.8%	3.1%	7.1%	4.6%	9.8%	6.0%	4.3%	10.7%
12	16.2%	40.1%	1.7%	2.3%	2.4%	14.5%	6.6%	8.0%	5.2%	5.1%	9.5%
13	5.7%	18.4%	0.4%	11.3%	0.0%	17.9%	5.1%	9.3%	5.0%	3.7%	8.6%
14	5.5%	5.8%	0.2%	32.4%	0.0%	10.2%	7.4%	8.8%	7.2%	6.2%	9.4%
15	4.0%	8.8%	0.9%	26.4%	0.0%	13.0%	6.4%	9.7%	7.9%	5.1%	8.0%
16	5.4%	3.5%	0.5%	6.1%	2.4%	5.7%	5.0%	5.6%	6.7%	7.0%	6.7%
17	0.8%	1.4%	0.2%	9.6%	5.5%	3.7%	6.0%	8.5%	7.5%	7.5%	5.2%
18	0.0%	0.0%	0.3%	1.9%	2.4%	10.7%	6.9%	5.5%	8.6%	9.3%	3.5%
19	0.7%	0.5%	0.5%	0.6%	0.0%	3.4%	7.8%	7.4%	7.0%	3.4%	4.6%
20	0.0%	0.0%	0.0%	0.8%	0.0%	5.7%	3.6%	6.7%	8.8%	2.1%	2.9%
21	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	1.3%	4.5%	5.9%	1.1%	1.9%
22	0.0%	0.0%	0.0%	0.2%	0.0%	2.9%	0.8%	2.4%	4.8%	0.9%	0.6%
23	0.0%	0.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.1%	0.2%	0.6%
24	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
25	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
26	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Figure 18: Time of Day Distributions
[from: All Other Trip Time Distributions.bin]

Trip time histograms for mid-workday trips away from work are used to determine the trip time of trips occurring in the middle of the workday. The distribution, shown in Figure 18, is obtained from the Parsons Brinckerhoff model and is also used in the aggregate RTC TransCAD model. For each person, the full probability distribution is truncated to ensure that the mid-workday activities occur between the arrival and departure of work. First the time for the trip away from work is determined by drawing from the truncated distribution, then the return trip is determined conditioned to occur after the trip away from work and before leaving work.

The time of day of all trips that are not part of the work tour are determined from the probability distributions in Figure 18. In the current version, these trips are not conditioned on the timing of the work tour, although this could be implemented. For home-based school and home-based college, the return trip is conditioned to happen after the trip to school/college. For home-based shop, home-based other, and non-home-based, the time is drawn without restriction from the distributions in Figure 18.

Destination/Mode

The destination and mode choice models from the Parsons Brinckerhoff model were implemented in the Clark County STEP2. The destination choice models are very similar to the work destination choice model that is described above. The main difference from the work destination choice is that the destination choice models for the other purposes are not segmented by auto ownership and income.

The utility (according to the PB model) of making a trip from origin TAZ *o* to destination TAZ *d* for trip purpose *p* is defined as:

$$
\begin{aligned}
U_{od}^{p} = {} & \beta_1^p(transportation\ level\ of\ service_{od}^p) \\
& + \beta_2^p(CBD\ dummy_d) \\
& + \beta_3^p(STRIP\ dummy_d) \\
& + \beta_4^p(INTRACBD\ dummy_{od}) \\
& + \beta_5^p(INTRASTRIP\ dummy_{od}) \\
& + \beta_6^p(INTRATAZ\ dummy_{od}) \\
& + \ln(attractions_d^p)
\end{aligned}
$$

The probability of person *n* traveling from TAZ *o* to TAZ *d* for trip purpose *p* is then:

$$
P_n(d) = \frac{\exp\left(U^{p}_{\text{origin zone of trip},d}\right)}{\sum_{all\ TAZ\ t} \exp(U^{p}_{\text{origin zone of trip},t})}.
$$

A different set of parameters β are used for each trip purpose. The level of service is obtained from the mode choice model (the logsum), and is applied using either peak period travel conditions (6-9 AM and 2-6 PM) or off-peak travel conditions depending on the time of the particular trip. The attractions for each purpose are calculated using Parsons Brinckerhoff's trip attraction model.

For each individual, home-based shop and home-based other trips are paired into *from home* and *to home* trips, and only one shop/other destination TAZ is simulated for each pair of trips. If there is an odd number of home-based shop trips or home-based other trips, then an unpaired trip is generated with equal probability of the home-end being the origin or the destination. For non-home-based trips, the origin end is taken to be the simulated TAZ of the non-home end of one of the individual’s home-based shop or home-based other trips. The destination end of each non-home based trip is simulated using a destination choice model.

The mode choice models are used to determine which mode was used for each trip. A nested logit model is used, and the tree structure is shown in Figure 19. The utility of each mode is a function of the level of service (broken down into different travel times including in-vehicle and wait time), TAZ characteristics such as CBD and Strip dummy variables, and the auto ownership in the household. Different parameters are used for each trip purpose. As with the destination choice, there is some dependency of mode choice across trips made by an individual. Any pair of home-based-shop or home-based-other use the same mode, and all trip on the journey to and from work use the same mode (although a mid-day stop does not have to).

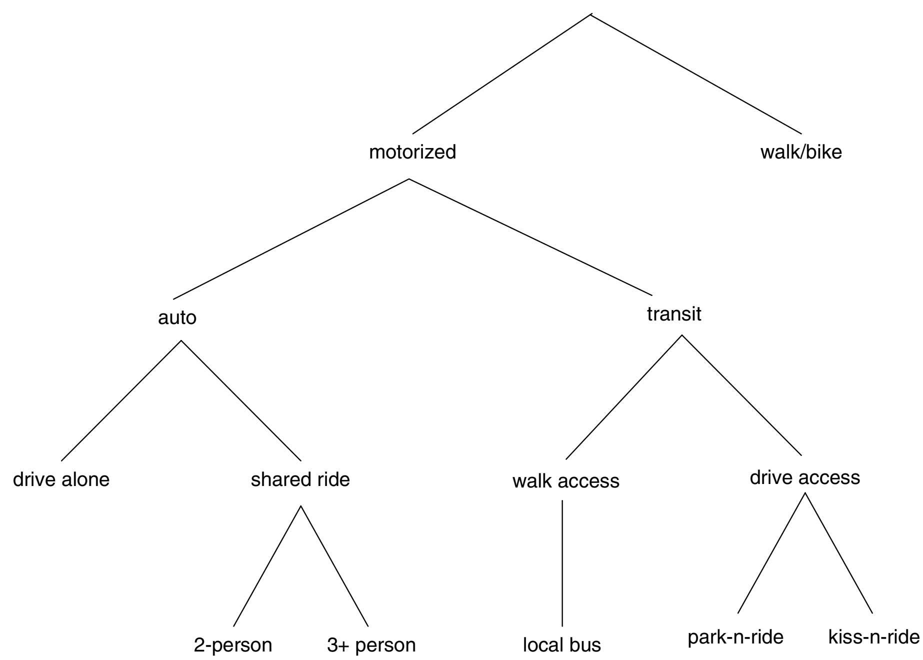


Figure 19: Mode Choice Model Structure

Output of Travel Demand Model: Individual Trips

At the end of the travel demand modeling steps described above, what has been generated is a list of simulated trips made by each person in the synthetic population. A sample of this table is shown in Figure 20. This is a strikingly different output from an aggregate application, in which only matrices of trips by limited socio-economic characteristics are generated. One benefit of producing specific trips assigned to individuals is that these trips can now be aggregated based on any trip or socio-economic characteristics. These trips can also be fed into a microscopic traffic simulator such as Caliper’s TransModeler program. One thing that may jump out from this table is that an individual’s travel patterns do not always make logical sense. This is a result of the various simplifications that are introduced in the modeling process. While many conditionalities are built in to the current STEP2 (for example, the home-based work tour has logically chained and timed trips), many simplifications remain. As additional complexities are introduced in the model (that is, removing existing simplifications), these travel schedules will become closer to reality.

Dataview1 - Trip_table

PersID	HHID	Purpose	[Orig Type]	[Dest Type]	[Orig TAZ]	[Dest TAZ]	Time
101	1	jtwHW	home	work	68	105	3
101	1	jtwHW	work	home	105	68	15
101	1	hboth	home	other	68	68	17
101	1	hboth	other	home	68	68	19
201	2	hboth	home	other	68	418	18
201	2	hboth	other	home	418	68	6
203	2	hbshop	home	shop	68	828	9
203	2	hbshop	shop	home	828	68	13
203	2	hboth	home	other	68	68	12
203	2	hboth	other	home	68	68	21
204	2	hboth	home	other	68	521	13
204	2	hboth	other	home	521	68	19
301	3	jtwHW	home	work	68	24	8
301	3	jtwHW	work	home	24	68	14
301	3	hbshop	home	shop	68	1129	9
301	3	hbshop	shop	home	1129	68	19
302	3	jtwHW	home	work	68	731	6
302	3	jtwWO	work	other	731	68	19
302	3	jtwHO	other	home	68	68	27
302	3	hbcol	home	college	68	240	17
302	3	hbcol	college	home	240	68	20
302	3	hbshop	home	shop	68	828	6
302	3	hbshop	shop	home	828	68	7
303	3	hbshop	shop	home	828	68	14
303	3	hboth	home	other	68	288	12
303	3	hboth	other	home	288	68	17
303	3	hboth	home	other	68	68	15

Figure 20: Simulated Trips for Synthetic Persons
[from: Trip_Table.bin]

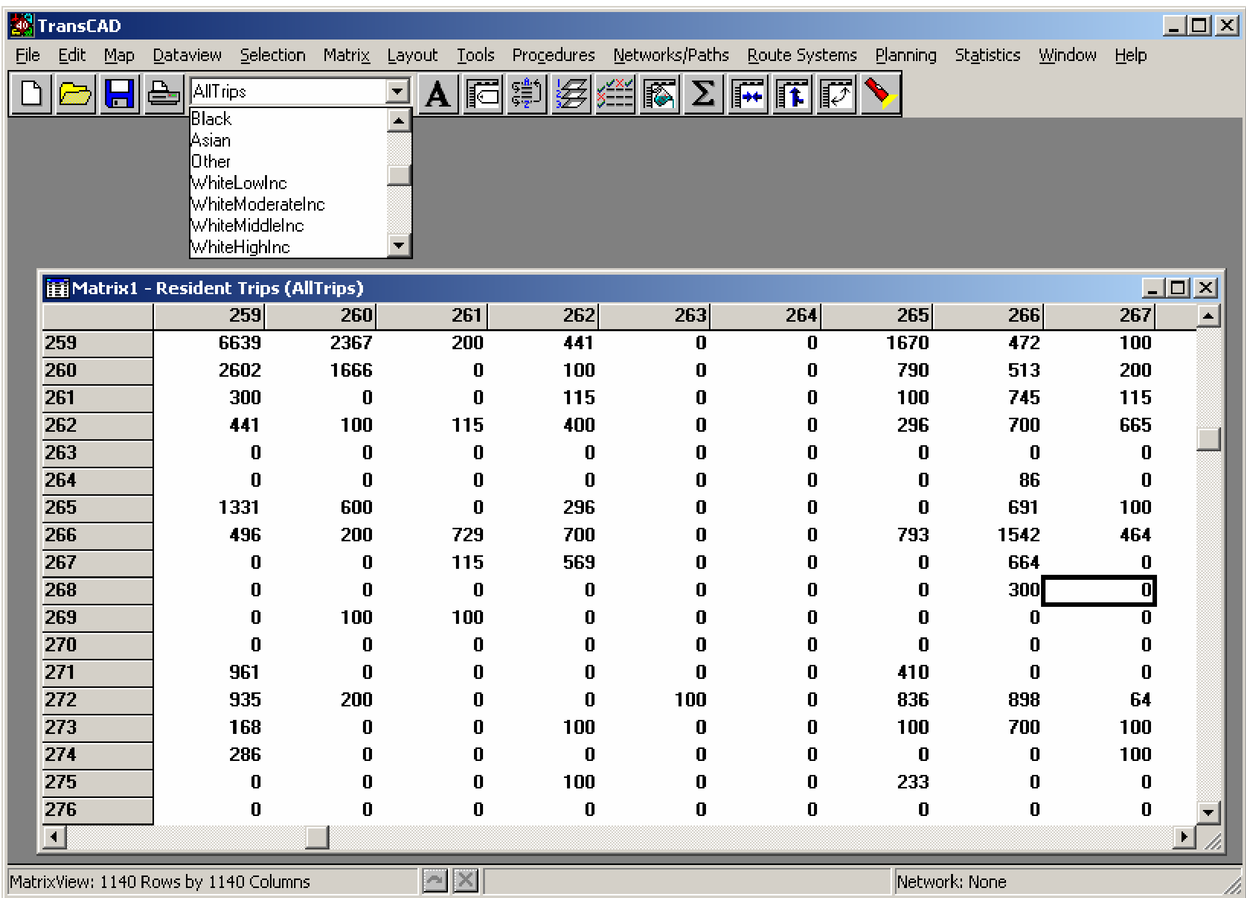
Aggregation and Assignment

The microsimulation travel demand model is performed by predicting travel made by individuals and focusing on behavior and decisions made at the level of the individual. Clearly these individual travel patterns are not of interest when making policy decisions. Rather, information at the aggregate level is necessary, including how these individual decisions aggregate to determine flows and level of service on the network, as well as summary statistics on travel characteristics on various segments of the population including how different groups are impacted by various policy decisions. Since a fraction of the population is simulated, this fraction needs to be expanded to the entire population.

Aggregate Trip Matrices

As described above, one of the key advantages to a microsimulation approach is that the entire richness of the population is preserved throughout the process, which permits tabulation for any subgroup (as long as there are sufficient observations to be statistically significant). STEP2 generates a host of different trip matrices, each containing the number of trips between each zonal pair by a particular segment of the population and/or type of travel. In terms of types of travel, a total trip matrix (all trips) is generated, as well as trip matrix of work tour trips, a trip matrix of non-work tour trips, and trip matrices for the AM Peak and PM Peak hours. In addition, trip matrices are generated by income group (low, moderate, middle, and high) and race (white, hispanic, black, asian, and other), and combinations of these traits (low income white, low income black, etc.).

Trip matrices can be generated not only by socio-economic group, but also by trip characteristics, such as journey to work or trips by various modes. With such information, summary charts explaining overall impacts by socio-economic group can be generated. Examples of such outputs are shown in Figure 22 and Figure 23. Figure 22 displays for each income group the number of work trips made by mode. Figure 23 is an example of scenario analysis results, in which the change in mode share by race between the base year and a future scenario is displayed.



	259	260	261	262	263	264	265	266	267
259	6639	2367	200	441	0	0	1670	472	100
260	2602	1666	0	100	0	0	790	513	200
261	300	0	0	115	0	0	100	745	115
262	441	100	115	400	0	0	296	700	665
263	0	0	0	0	0	0	0	0	0
264	0	0	0	0	0	0	0	86	0
265	1331	600	0	296	0	0	0	691	100
266	496	200	729	700	0	0	793	1542	464
267	0	0	115	569	0	0	0	664	0
268	0	0	0	0	0	0	0	300	0
269	0	100	100	0	0	0	0	0	0
270	0	0	0	0	0	0	0	0	0
271	961	0	0	0	0	0	410	0	0
272	935	200	0	0	100	0	836	898	64
273	168	0	0	100	0	0	100	700	100
274	286	0	0	0	0	0	0	0	100
275	0	0	0	100	0	0	233	0	0
276	0	0	0	0	0	0	0	0	0

Figure 21: STEP2 Produces Trip Matrices by Socio-economic Group

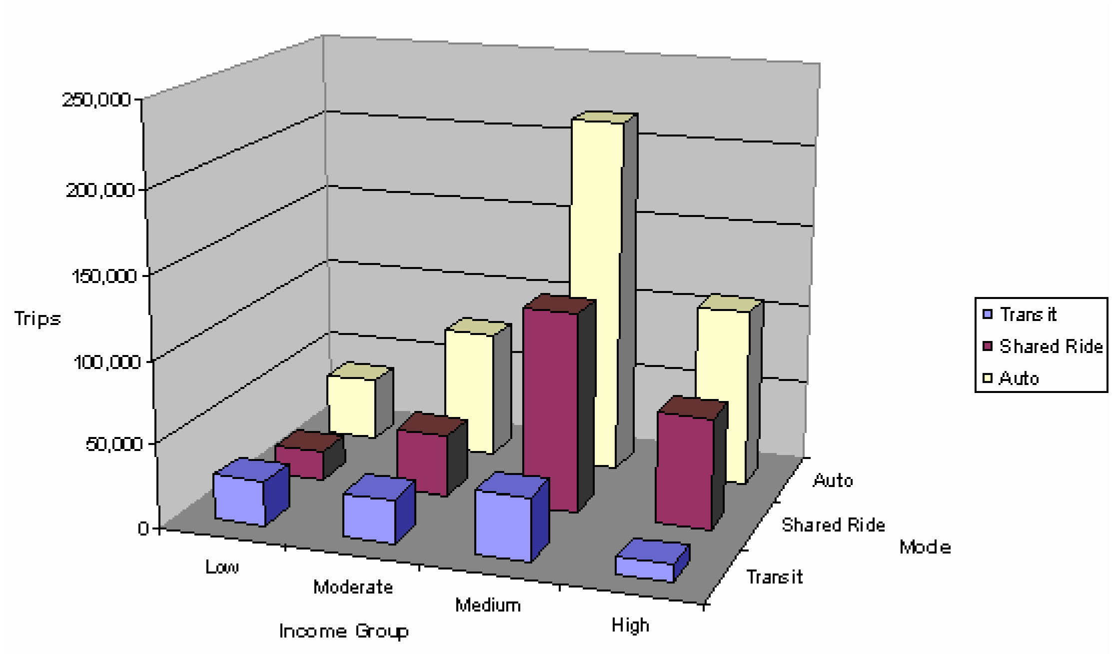


Figure 22: Example Output of Trip Characteristics by Income Group

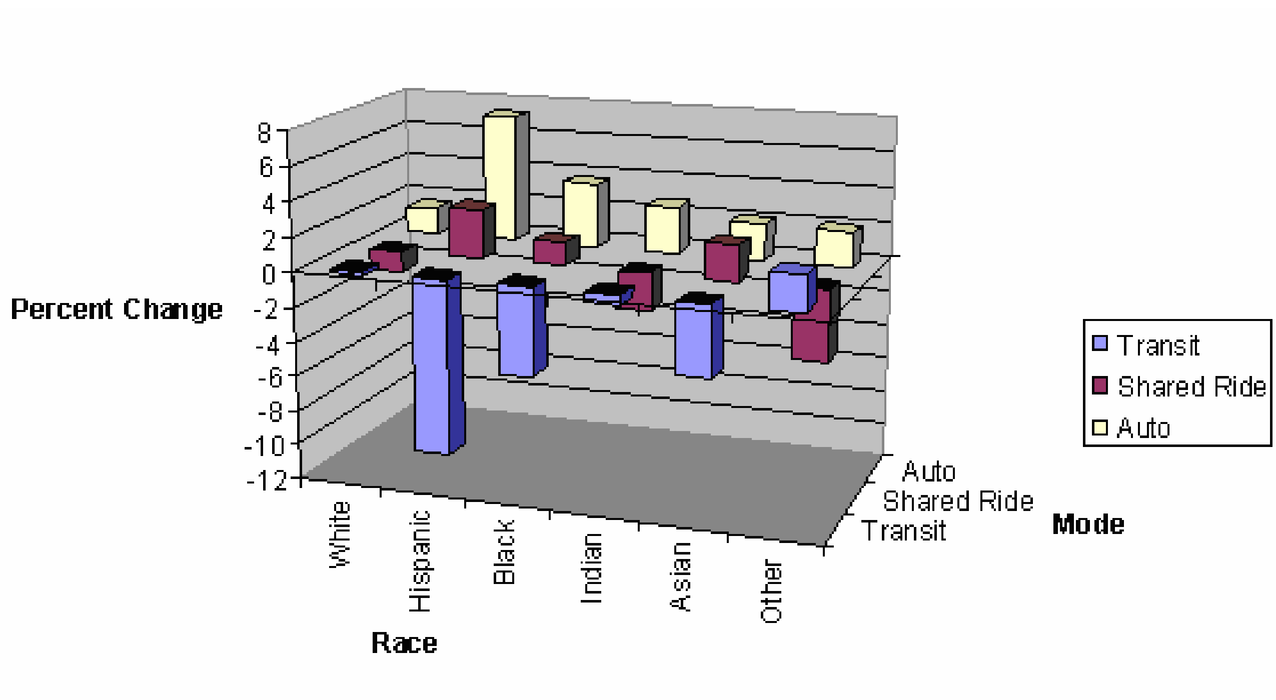


Figure 23: Example of Changes in Mode Share by Race, Base Year to Future Scenario

The other critical aspect of aggregating trip matrices is to generate input trip matrices for a traffic assignment procedure. In addition to developing matrices from the travel predicted in the microsimulation travel demand model, additional trips need to be included such as taxi trips, truck trips, external trips and visitor trips. For the Clark County implementation of STEP2, these matrices are imported from the RTC Phase 1 Model.

Traffic Assignment

Traffic Assignment is run for 3 key hours during the day, the AM Peak Hour (7 AM – 8 AM), the PM Peak Hour (4 PM – 5 PM) and the Peak Off-Peak Hour (11 AM – 12 PM). A User Equilibrium assignment based on the RTC Phase 1 Model implementation is used.

Future Directions

With microsimulation, the sky is the limit. These are very expandable methods that can be used to develop behavioral models based on a rich description of demographic and travel characteristics. In terms of the level of richness incorporated in the models, there is a wide range of detail that can be captured, from the most basic independent trip assumption (similar to an aggregate application) to a day activity schedule model that aims to model realistically all of the movements a person makes during the day. The future direction of STEP2 is to improve the models by increasing the level of detail in the choice hierarchy and improving the feedback and linkages among the model components. This can be done by making more use of nested logit model structures, in which conditionality is passed down the tree and logsums (the feedback) are passed up the tree. Some specific improvements to be made are to introduce behavioral models for trip generation that are a function of level of service and demographics, forecast other important socio-economic variables such as vehicle ownership, and create specialized travel demand models for retirees and other important demographic groups. As more detail is introduced in the models, computational speed becomes a major issue, and efforts need to be made to streamline the procedures and make use of techniques such as multiprocessing. Other important directions are to continue the development of tools to ease calibration of the model, data processing, and modification of the demand models. Finally, a valuable output of microsimulation models is the list of specific trips made by specific individuals, which can be used as the input to a microscopic traffic simulator and used to perform detailed traffic simulation to, for example, environmental hotspots on the transportation network.

Conclusion

This report described the STEP2 Household Microsimulator developed for Clark County, a state-of-the-art microsimulation transportation planning program. There are two key attributes that make this model fundamentally different than the aggregate, 4-step model currently used by RTC: (1) the model is applied at the level of the individual rather than based on zonal averages and (2) it includes a residential choice component to forecast the spatial distribution of households in the urban area. This effort builds upon and enhances the STEP travel demand analysis package, which was originally developed in the 1970s as a sketch planning tool for the San Francisco area (Harvey, 1978) and partially integrated into TransCAD for a project for Baltimore. STEP2 aims to further revive the STEP models and the general philosophy behind them, including modeling at the level of the household and developing tools in a way such that they are accessible to transportation planning agencies. STEP2 has gone beyond STEP in numerous ways, including running within a GIS environment, providing a user-friendly interface and scenario manager, incorporating realistic transportation networks and traffic assignment, explicitly incorporating the concept of trip-chaining, integrating models that have been developed specifically for Clark County, and making use of the 1996 Las Vegas Valley household survey. As part of the project, flexible and generic microsimulation tools, data processing, and calibration capabilities were developed so that the default specifications can be readily modified and enhanced.

The STEP2 model has three major components:

(1) A population synthesizer, which uses PUMS and aggregate zonal data to generate a representative population of specific (but synthetic) individuals and their personal and household characteristics, and can also age the population to the study year. The output is a synthetic household and person database that is representative of the population and includes detailed information from the PUMS databases (income, ages, occupations, race, etc.)

(2) Travel demand models to simulate travel-related behavior for each individual in the synthetic population. Travel demand is simulated through a choice hierarchy of decisions. At the top level are lifestyle and mobility decisions including where to live and whether to work and where to work. At the lower level are daily trip decisions, including whether to go to work, whether to makes stops on the way to or from work, what non-work trips to make, and when, where, and by what mode to make these trips. The travel demand component is a compilation of choice models (logit and nested logit) and cross-classification and probability distribution lookup tables. The STEP2 implementation makes heavy use of models developed for the original STEP program and models developed for Las Vegas by Parsons-Brinckerhoff for RTCs current aggregate 4-step travel demand model. In addition, new models were developed for STEP2 using both PUMS data and the 1996 household survey. There is conditionality and feedback built in throughout the model structure. The output of the travel demand component is a database of simulated trips for each person, including the number of trips by trip purpose and each trip’s origin, destination, mode, and time of day. Additionally, the residential choice component projects shifts in the distribution of households by zone and demographics in response to future scenarios.

(3) Routines to aggregate individual trips into trip matrices to assign on the network. The outputs are transportation level of service characteristics (including link flows and congestion) and aggregate summaries of the travel patterns and impacts made on different socio-economic groups (defined by income and race or any other characteristic).

The inputs required to STEP2 include the Public Use Microsample (PUMS), a TAZ zonal system including aggregate household demographics and employment information, and transportation network files. Valuable aspects to the output from STEP2 are the ability to produce impacts by socio-economic group (and therefore address issues of equity) and the creation of individual person/vehicle movements that can be fed into a microscopic traffic simulator. Furthermore, the modeling analysis tool is easily expandable to include additional sensitivity to demographics and transportation policy variables by incorporating more detailed representation of travel and travel behavior. The future directions of STEP2 are to improve the feedback and linkages among the travel demand models to generate

References

Applied Management & Planning Group in association with Parsons Brinckerhoff Quade & Douglas (1998) *Las Vegas Resort Corridor Major Investment Study – 1996 Las Vegas Valley Household Travel Survey*, prepared for Regional Transportation Commission, Las Vegas, Nevada.

Ben-Akiva, M. and S.R. Lerman (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*, The MIT Press, Cambridge, Massachusetts.

Bowman, J.L. (1998) *The Day Activity Schedule Approach to Travel Demand Analysis*, Ph.D. Dissertation, Massachusetts Institute of Technology.

Caliper Corporation (2002) *User's Guide for the RTC Las Vegas Phase 1 Model in TransCAD 4.0*.

Deming, W.E. and F.F. Stephan (1940) On a Least Squares djustment of a Sampled Frequency Table when the Expected Marginal Tables ar eKnonw, *Annals of Mathematical Statistics 11,* 427-444.

Dugway G., W. Jung, and D. McFadden (1976) *SYNSAM: A Methodology for Synthesizing Household Transportation Survey Data*. Working Paper No. 7618, University of California ant Berkeley.

Harvey, G. (1978) *STEP: Short-range Transportation Evaluation Program*, Metropolitan Transportation Commission, Oakland, California.

Harvey, G. and E. Deakin (1996) *Description of the STEP Analysis Package*, working paper.

Orcutt, G., S. Caldwell, and R. Wertheimer (1976) *Policy Exploration through Microanalytic Simulation,* Urban Institute Press, Washington, D.C.

Parsons Brinckerhoff Quade & Douglas (2001) *Las Vegas Travel Demand Model Methodology Report, Volume 1: Chapters 1-17*, prepared for Regional Transportation Commission of Clark County, Nevada.

Slavin, H. and J. Lam (2001) *STEP Microsimulation for Transportation and Equity Analysis*, presented at the Transportation Research Board Annual Meeting, Washington, D.C.