

Appendix T

SANDAG Travel Demand Model and Forecasting Documentation

Appendix T Contents

SANDAG Travel Demand Model Documentation and
Methodology

Series 13 Demographic and Economic Forecasting Model (DEFM)
Technical Documentation

Series 13 Subregional Allocation Technical Documentation

SANDAG Travel Demand Model Documentation

SANDAG Travel Demand Model Documentation and Methodology

This document describes the San Diego Association of Governments (SANDAG) Activity-Based Model (ABM) specification. This ABM will serve as the major travel forecasting tool in the San Diego region for decades to come. This model has been developed to ensure that the regional transportation planning process can rely on forecasting tools that will be adequate for new socioeconomic environments and emerging planning challenges. It is equally suitable for conventional highway projects, transit projects, and various policy studies such as highway pricing and HOV analysis.

The SANDAG model is based on the CT-RAMP (Coordinated Travel Regional Activity-Based Modeling Platform) family of Activity-Based Models. This model system is an advanced, but operational, AB model that fits the needs and planning processes of SANDAG. The CT-RAMP model, which is fully described in the following section, adheres to the following basic principles:

- The CT-RAMP design corresponds to the most advanced principles of modeling individual travel choices with maximum behavioral realism. In particular, it addresses both household-level and person-level travel choices including intra-household interactions between household members. This approach is fundamentally different from the more simplified AB models developed or being developed in such regions as San Francisco County, Sacramento and Denver, where all travel choices are modeled at the person level, independently of choices made by other household members.
- CT-RAMP is a proven design, intensively tested in practice in several regions. The New York model was developed in 2002, and was used in the New York region to analyze numerous projects. The Columbus, Ohio model (the first fully-fledged member of the CT-RAMP family) was developed in 2004 and has since been applied by the MORPC for various transit and highway projects. The Lake Tahoe model was created in 2006 largely by transferring main components of the Columbus model. The Atlanta, Georgia (ARC) model has been co-developed with the MTC Model. Future developments of CT-RAMP include models for the San Diego region (SANDAG) and Jerusalem, Israel (JTMT). In each case, the model system has been tailored to address the specific issues and markets that are particular to the region.
- Operates at a detailed temporal (half-hourly) level, and considers congestion and pricing effects on travel time-of-day and peak spreading of traffic volume.
- Reflects and responds to detailed demographic information, including household structure, aging, changes in wealth, and other key attributes¹.
- Is implemented in the PB Common Modeling Framework, an open-source library created specifically for implementing advanced models.
- Offers sensitivity to demographic and socio-economic changes observed or expected in the dynamic San Diego metropolitan region. This is ensured by the enhanced and flexible population synthesis procedures as well as by the fine level of model segmentation. In particular, the SANDAG ABM incorporates different household, family, and housing types including a detailed analysis of different household compositions in their relation to activity-travel patterns.

- *Accounts for the full set of travel modes.* Our experience with previously developed ABMs has shown that mode choice is one of the least transferable model components due to regional differences in mode choice and accessibility.
- *Integrates with other model components.* The CT-RAMP model is one component (person travel) and can easily integrate with other components such as the existing SANDAG truck model, a model of inter-border commuting (particularly important given the interaction of San Diego with Mexico and with Orange, Riverside, and Imperial County), models of non-resident visitor travel, airport travel, and special event travel. Furthermore, the model developed for SANDAG will be integrated with the Production, Exchange, and Consumption (PECAS) land-use model system.
- *Provides detailed inputs to traffic micro-simulation software.* The CT-RAMP model operates at a half-hour time scale, which can provide detailed inputs to traffic micro-simulation software for engineering-level analysis of corridor and intersection design.

Model Features and SANDAG Planning Needs

The SANDAG CT-RAMP model has been tailored specifically to meet SANDAG planning applications, as outlined below. These planning applications consider current and future projects and policies and also take into account the special markets that exist in the San Diego Region. The model system addresses requirements of the metropolitan planning process, relevant federal requirements, and provides support to SANDAG member agencies and other stakeholders. The ABM structure fully complies with the following major planning applications:

- *RTP, TIP, and Air Quality Conformity Analysis.* The ABM will be carefully validated and calibrated to replicate observed traffic counts and other monitoring data sources with the necessary level of accuracy. The output of traffic assignment can be processed in a format required by the emission calculation software used by SANDAG, including either EMFAC or MOVES.
- *Corridor Studies, Development Impact Studies, and other planning studies.* The ABM will have more realistic travel patterns that will lend itself to a high level of credibility with respect to routine planning studies conducted by SANDAG staff and other model users.
- *FTA New Starts Analysis.* The ABM application software package includes an option that produces the model output in a format required by FTA for the New Starts process. This output can be used as a direct input to the FTA software Summit used for calculation and analysis of the User Benefits. In order to meet the FTA “fixed total demand” requirement for comparison across the Baseline and Build alternatives, the ABM includes a run option for the Build alternative with certain travel dimensions fixed from the Baseline run.
- *Different highway pricing and managed lanes studies.* One of the advantages of an ABM over a 4-step model is a significantly improved sensitivity to highway pricing. This includes various forms of toll roads, congestion pricing, dynamic real-time pricing, daily area pricing, license plate rationing and other innovative policies that cannot be effectively modeled with a simplified 4-step model. The explicit modeling of joint travel was specifically introduced to enhance modeling of HOV/HOT facilities.
- *Other transportation demand management measures.* There are many new policies aimed at reducing highway congestion in major metropolitan areas, including telecommuting and teleshopping, compressed work weeks, and flexible work hours. ABMs are specifically effective for modeling these types of policies since these models are based on an individual micro-simulation of daily activity-travel patterns.

- *Enhanced Environmental Justice analysis.* The model system features a full micro-simulation of the population, providing the ability to perform virtually unlimited market analysis. Environmental justice disparity analysis can be performed across highly disaggregated user groups, providing information for Title VI and other types of environmental justice studies.

General Model Design

The SANDAG ABM has its roots in a wide array of analytical developments. They include discrete choice forms (multinomial and nested logit), activity duration models, time-use models, models of individual micro-simulation with constraints, entropy-maximization models, etc. These advanced modeling tools are combined in the ABM design to ensure maximum behavioral realism, replication of the observed activity-travel patterns, and ensure model sensitivity to key projects and policies. The model is implemented in a micro-simulation framework. Micro-simulation methods capture aggregate behavior through the representation of the behavior of individual decision-makers. In travel demand modeling, these decision-makers are typically households and persons. The following section describes the basic conceptual framework at which the model operates.

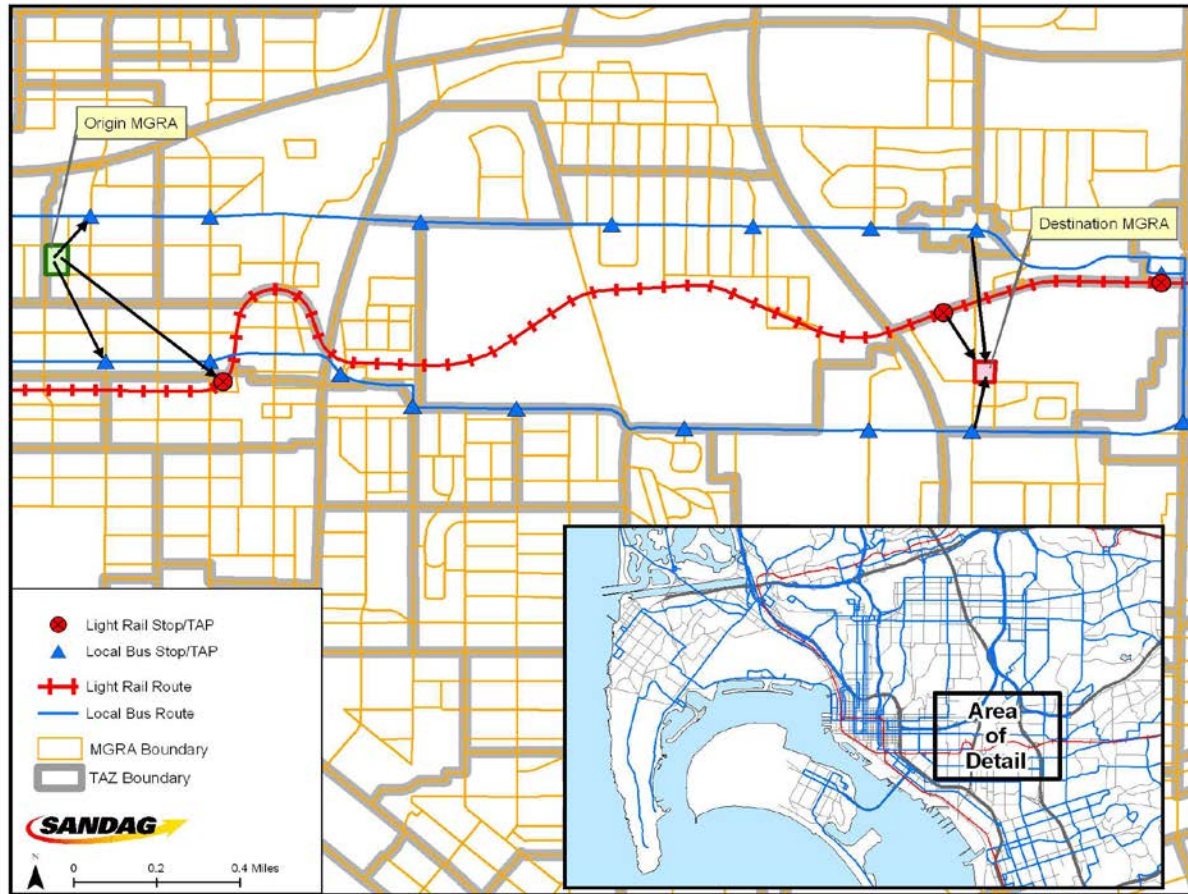
Treatment of space

Activity-based and tour-based models can exploit fine-scale spatial data, but the advantages of additional spatial detail must be balanced against the additional efforts required to develop zone and associated network information at this level of detail. The increase in model runtime and necessary computing power associated primarily with path-building and assignment to more zones must also be considered.

The use of a spatially disaggregate zone system helps ensure model sensitivity to phenomena that occur at a fine spatial scale. Use of large zones may produce aggregation biases, especially in destination choice, where the use of aggregate data can lead to illogical parameter estimates due to reduced variation in estimation data, and in mode choice, where modal access may be distorted. Smaller zones help minimize these effects, and can also support more detailed road network assignments. Strategies to address the modal access limitations of large zones through the use of transit sub-zonal procedures are discussed in the transit network section of this document.

The SANDAG CT-RAMP model will take advantage of SANDAG Master-Geographic Reference Area (MGRA) zone system, which is the most disaggregate zonal system currently in use in any travel demand model in the United States. Most large metropolitan area travel demand models consider between 1,500 and 4,000 zones. The SANDAG current MGRA system consists of 32,000 zones, which are roughly equivalent to Census block groups (see Figure T.2). To avoid computational burden, SANDAG relies on a 4,600 Transportation Analysis Zone (TAZ) system for highway skims and assignment, but performs transit calculations at the more detailed MGRA level. This is accomplished by generalizing transit stops into pseudo-TAZs called Transit Access Points (TAPs) and relying on TransCAD to generate TAP-TAP level-of-service matrices (also known as “skims”) such as in-vehicle time, first wait, transfer wait, and fare. All access and egress calculations, as well as paths following the Origin MGRA – Boarding TAP – Alighting TAP-Destination MGRA patterns, are computed within custom-built software. These calculations rely upon detailed geographic information regarding MGRA-TAP distances and accessibilities. A graphical depiction of the MGRA – TAP transit calculations is given in Figure T.1. It shows potential walk paths from an origin MGRA, through three potential boarding TAPs (two of which are local bus, and one of which is rail), with three potential alighting TAPs at the destination end.

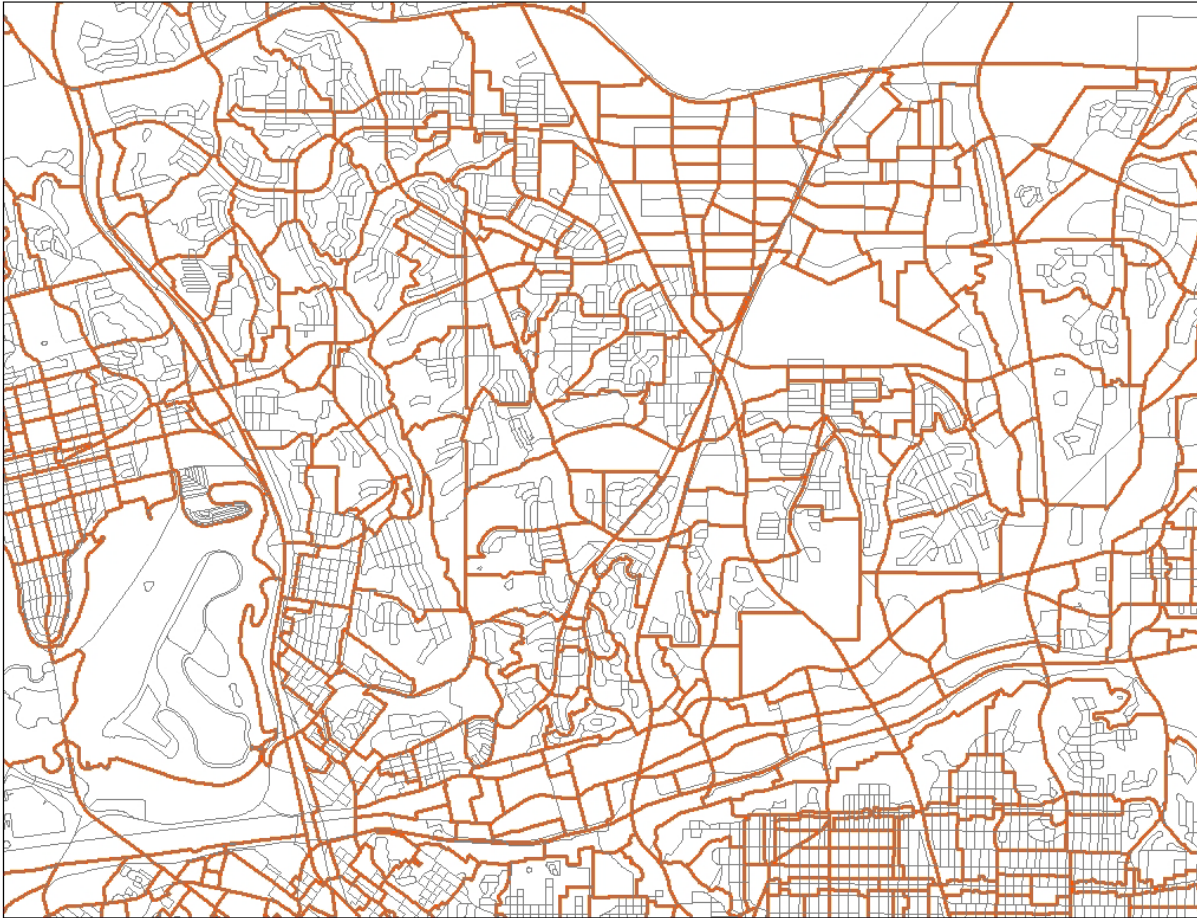
Figure T.1
Example MGRA – TAP Transit Accessibility



All activity locations are tracked at the MGRA level. There are model systems in use or under development which allocate activities to a unit smaller than the MGRA, such as a parcel. However, these model systems assume that the closest transit stop to the parcel is consistent with the zone-zone impedances calculated by the commercial transport software (TransCAD). In transit-rich environments this may not be the case and such assumptions can cloud User Benefit calculations required by FTA New Starts. The MGRA geography offers both the advantage of fine spatial resolution, and consistency with network levels-of-service, that makes it ideal for tracking activity locations.

Figure T.2

Treatment of Space – TAZs and MGRAs



Decision-making units

Decision-makers in the model system include both persons and households. These decision-makers are created (synthesized) for each simulation year based on tables of households and persons from census data and forecasted TAZ-level distributions of households and persons by key socio-economic categories. These decision-makers are used in the subsequent discrete-choice models to select a single alternative from a list of available alternatives according to a probability distribution. The probability distribution is generated from a logit model which takes into account the attributes of the decision-maker and the attributes of the various alternatives. The decision-making unit is an important element of model estimation and implementation, and is explicitly identified for each model specified in the following sections.

Person type segmentation

The SANDAG ABM system is implemented in a micro-simulation framework. A key advantage of using the micro-simulation approach is that there are essentially no computational constraints on the number of explanatory variables that can be included in a model specification. However, even with this flexibility, the model system will include some segmentation of decision-makers. Segmentation is a useful tool to both structure models such that each person type segment could have their own model for certain choices, and to characterize person roles within a household. Segments can be created for persons as well as households.

A total of eight segments of person types (shown in Table T.1) are used for the SANDAG model system. The person types are mutually exclusive with respect to age, work status, and school status.

**Table T.1
Person Types**

Number	Person type	Age	Work Status	School Status
1	Full-time worker ²	18+	Full-time	None
2	Part-time worker	18+	Part-time	None
3	College student	18+	Any	College +
4	Non-working adult	18 – 64	Unemployed	None
5	Non-working senior	65+	Unemployed	None
6	Driving age student	16-17	Any	Pre-college
7	Non-driving student	6 – 15	None	Pre-college
8	Pre-school	0-5	None	None

Further, workers are stratified by their occupation to take full advantage of information provided by the PECAS land-use model. The categories are given in Table T.2. These are used to segment destination choice size terms for work location choice, based on the occupation of the worker.

**Table T.2
Occupation Types**

Number	Description
1	Management Business Science and Arts
2	Services
3	Sales and Office
4	Natural Resources Construction and Maintenance
5	Production Transportation and Material Moving
6	Military

Activity type segmentation

The 2006 SANDAG home-interview survey included 24 different activity codes. Modeling all 24 activity types would add significant complexity to estimating and implementing the model system, so these detailed activity types are grouped into more aggregate activity types, based on the similarity of the activities. The activity types are used in most model system components, from developing daily activity patterns to predicting tour and trip destinations and modes by purpose.

The proposed set of activity types is shown in Table T.3. The activity types are grouped according to whether the activity is mandatory, maintenance, or discretionary. Eligibility requirements are assigned to determine which person types can be used for generating each activity type. The classification scheme of each activity type reflects the relative importance or natural hierarchy of the activity, where work and school activities are typically the most inflexible in terms of generation, scheduling and location, whereas discretionary activities are typically the most flexible on each of

these dimensions. When generating and scheduling activities, this hierarchy is not rigid and is informed by both activity-type and activity-duration.

Each out-of-home location that a person travels to in the simulation is assigned one of these activity types.

Table T.3
Activity Types

Type	Purpose	Description	Classification	Eligibility
1	Work	Working at regular workplace or work-related activities outside the home	Mandatory	Workers and students
2	University	College +	Mandatory	Age 18+
3	High School	Grades 9-12	Mandatory	Age 14-17
4	Grade School	Grades K-8	Mandatory	Age 5-13
5	Escorting	Pick-up/drop-off passengers (auto trips only)	Maintenance	Age 16+
6	Shopping	Shopping away from home	Maintenance	5+ (if joint travel, all persons)
7	Other Maintenance	Personal business/services, and medical appointments	Maintenance	5+ (if joint travel, all persons)
8	Social/Recreational	Recreation, visiting friends/family	Discretionary	5+ (if joint travel, all persons)
9	Eat Out	Eating outside of home	Discretionary	5+ (if joint travel, all persons)
10	Other Discretionary	Volunteer work, religious activities	Discretionary	5+ (if joint travel, all persons)

Treatment of time

The model system functions at a temporal resolution of one-half hour. These one-half hour increments begin with 3 A.M. and end with 3 A.M. the next day, though the hours between 1 A.M. and 5 A.M. will be aggregated to reduce computational burden. Temporal integrity is ensured so that no activities are scheduled with conflicting time windows, with the exception of short activities/tours that are completed within a one-half hour increment. For example, a person may have a very short tour that begins and ends within the 8:00 a.m. to 8:30 a.m. period, as well as a second longer tour that begins within this time period, but ends later in the day.

Time periods are typically defined by their midpoint in the scheduling software. For example, in a model system using one-half hour temporal resolution, the 9:00 a.m. time period would capture activities of travel between 8:45 a.m. and 9:15 a.m. If there is a desire to break time periods at “round” half-hourly intervals, either the estimation data must be processed to reflect the aggregation of activity and travel data into these discrete half-hourly bins, or a more detailed temporal resolution must be used, such as half-hours (which could then potentially be aggregated to “round” half-hours).

A critical aspect of the model system is the relationship between the temporal resolution used for scheduling activities, and the temporal resolution of the network simulation periods. Although each activity generated by the model system is identified with a start time and end time in one-half hour increments, level-of-service matrices are only created for five aggregate time periods: (1) early A.M.; (2) A.M.; (3) Midday; (4) P.M.; and (5) Evening. The trips occurring in each time period reference the appropriate transport network depending on their trip mode and the mid-point trip time. The definition of time periods for level-of-service matrices is given in Table T.4,

Table T.4**Time Periods for Level-of-Service Skims and Assignment**

Number	Description	Begin Time	End Time
1	Early	3:00 A.M.	5:59 A.M.
2	A.M. Peak	6:00 A.M.	8:59 A.M.
3	Midday	9:00 A.M.	3:29 P.M.
4	P.M. Peak	3:30 P.M.	6:59 P.M.
5	Evening	7:00 P.M.	2:59 A.M.

Trip modes

Table T.5 lists the trip modes defined in the SANDAG models. There are 26 modes available to residents, including auto by occupancy and toll/non-toll choice, walk and bike non-motorized modes, and walk and drive access to five different transit line-haul modes. Note that the pay modes are those that involve paying a choice or “value” toll. Tolls on bridges are counted as a travel cost, but the mode is considered “free.”

Table T.5**Trip Modes For Assignment**

Number	Mode
1	Auto SOV (Non-Toll)
2	Auto SOV (Toll)
3	Auto 2 Person (Non-Toll, Non-HOV)
4	Auto 2 Person (Non-Toll, HOV)
5	Auto 2 Person (Toll, HOV)
6	Auto 3+ Person (Non-Toll, Non-HOV)
7	Auto 3+ Person (Non-Toll, HOV)
8	Auto 3+ Person (Toll, HOV)
9	Walk-Local Bus
10	Walk-Express Bus
11	Walk-Bus Rapid Transit
12	Walk-Light Rail
13	Walk-Heavy Rail
14	PNR-Local Bus
15	PNR-Express Bus
16	PNR-Bus Rapid Transit
17	PNR-Light Rail
18	PNR-Heavy Rail
19	KNR-Local Bus
20	KNR-Express Bus
21	KNR-Bus Rapid Transit
22	KNR-Light Rail
23	KNR-Heavy Rail
24	Walk
25	Bike
26	School Bus (only available for school purpose)

Basic design of the SANDAG CT-RAMP implementation

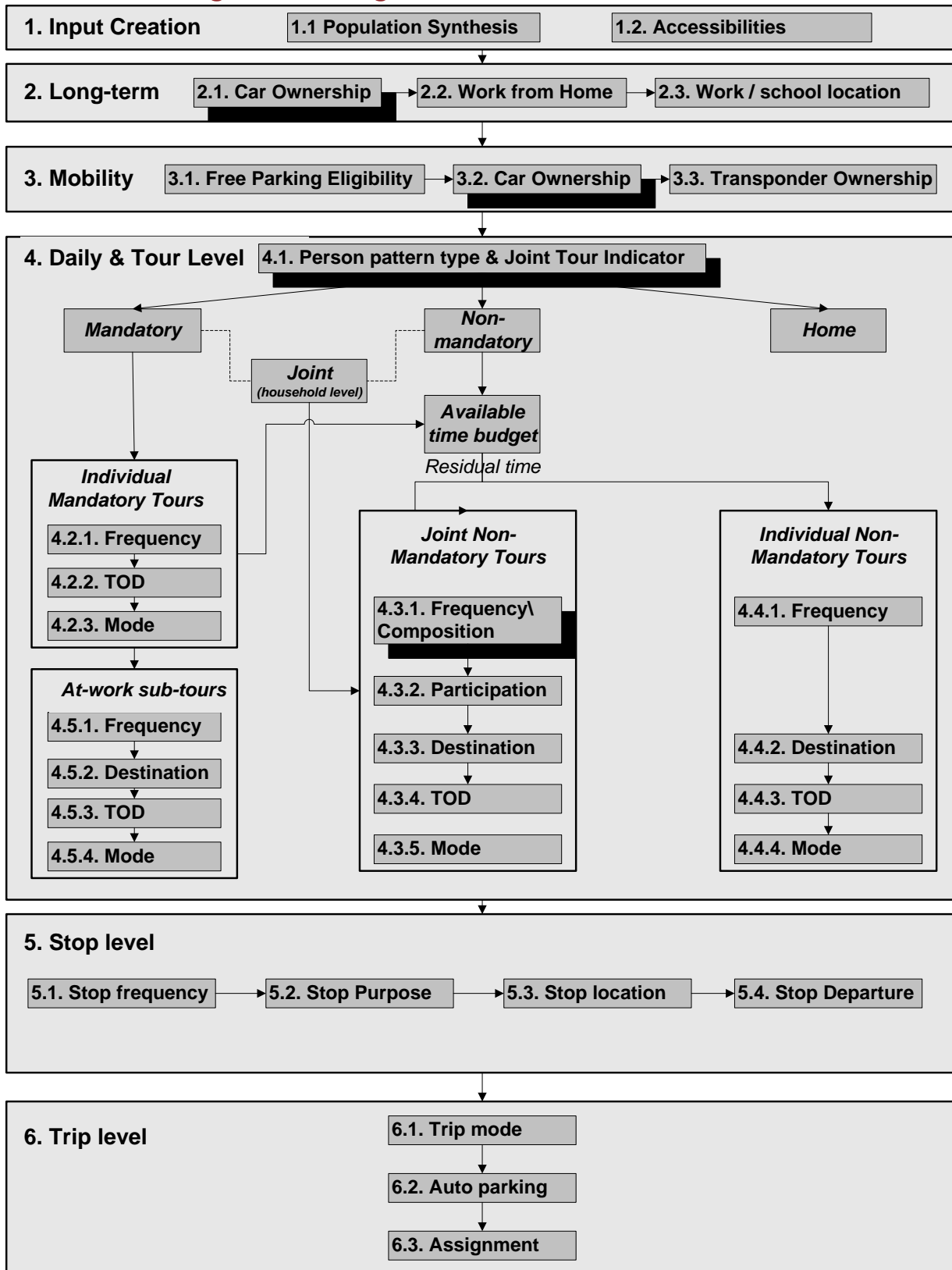
The general design of the SANDAG CT-RAMP model is presented in Figure T.3. The following outline describes the basic sequence of sub-models and associated travel choices:

1. Input Creation:
 - 1.1. Synthetic population creation
 - 1.2. Calculation of destination-choice accessibilities for use in mobility models and tour generation
2. Long term level:
 - 2.1. Household car ownership (based on household/person attributes and household accessibilities)
 - 2.2. Work from home model that indicates whether a worker's regular workplace is their home
 - 2.3. The location for each mandatory activity for each relevant household member (workplace/university/school)
3. Mobility Level:
 - 3.1. Free Parking Eligibility (determines whether workers pay to park if workplace is an MGRA with parking cost)
 - 3.2. Household car ownership (based on household/person attributes, household, and mandatory accessibilities)
 - 3.3. Transponder ownership for use of toll lanes
4. Daily pattern/schedule level:
 - 4.1. Daily pattern type for each household member (main activity combination, at home versus on tour) with a linkage of choices across various person categories, and generation of a joint tour indicator at the household level
 - 4.2. Individual mandatory activities/tours for each household member (note that locations of mandatory tours have already been determined in long-term choice model)
 - 4.2.1 Frequency of mandatory tours
 - 4.2.2 Mandatory tour time of day (departure/arrival time combination)
 - 4.2.3 Mandatory tour mode choice
 - 4.3. Joint travel tours (conditional upon the available time window left for each person after the scheduling of mandatory activities, and the presence of a joint tour indicated from Model 4.1)
 - 4.3.1 Joint tour frequency/composition, which predicts the exact number of joint tours (1 or 2), the purpose of each tour, and the composition of each tour (adults, children, or mixed)
 - 4.3.2 Person participation in each joint tour
 - 4.3.3 Primary destination for each joint tour
 - 4.3.4 Joint tour time of day (departure/arrival time combination)
 - 4.3.5 Joint tour mode choice

- 4.4. Individual non-mandatory tours (conditional upon the available time window left for each person after the scheduling of mandatory and joint non-mandatory activities)
 - 4.4.1 Individual non-mandatory tour frequency, applied for each person
 - 4.4.2 Individual non-mandatory tour primary destination
 - 4.4.3 Individual non-mandatory tour departure/arrival time
 - 4.4.4 Individual non-mandatory tour mode choice
- 4.5. At-work sub-tours (conditional upon the available time window within the work tour duration)
 - 4.5.1 At-work sub-tour frequency, applied for each work tour
 - 4.5.2 At-work sub-tour primary destination
 - 4.5.3 At-work sub-tour departure/arrival time
 - 4.5.4 At-work sub-tour mode choice
- 5. Stop level:
 - 5.1. Frequency of secondary stops
 - 5.2. Intermediate stop purpose
 - 5.3. Intermediate stop location choice
 - 5.4. Intermediate stop departure time choice
- 6. Trip level:
 - 6.1. Trip mode choice conditional upon the tour mode
 - 6.2. Auto trip parking location choice for parking constrained areas
 - 6.3. Trip assignment

Figure T.3

Basic Model Design and Linkage Between Sub-Models



Shaded boxes in Figure T.3 indicate choices that relate to the entire household or a group of household members and assume explicit modeling of intra-household interactions (sub-models 2.1, 3.2, 4.1, and 4.3.1). The other models are applied to individuals, though they may consider household-level influences on choices.

The model system uses synthetic household population as a base input (sub-model 1.1). Certain models also utilize destination-choice logsums, which are represented as MGRA variables (sub-model 1.2). Once these inputs are created, the travel model simulation begins.

An auto ownership model is run before workplace/university/school location choice in order to select a preliminary auto ownership level for calculation of accessibilities for location choice. The model uses the same variables as the full auto ownership model, with the exception of the work/university/school-specific accessibilities that are used in the full model. It is followed by long-term choices that relate to the workplace/university/school for each worker and student (sub-models 2.2 and 2.3). Medium-term mobility choices relate to free parking eligibility for workers in the CBD (sub-model 3.1), household car ownership (sub-model 3.2), and transponder ownership (sub-model 3.3).

The daily activity pattern type of each household member (model 4.1) is the first travel-related sub-model in the modeling hierarchy. This model classifies daily patterns by three types: (1) mandatory (that includes at least one out-of-home mandatory activity); (2) non-mandatory (that includes at least one out-of-home non-mandatory activity but does not include out-of-home mandatory activities); and (3) home (that does not include any out-of-home activity and travel). The pattern type model also predicts whether any joint tours will be undertaken by two or more household members on the simulated day. However, the exact number of tours, their composition, and other details are left to subsequent models. The pattern choice set contains a non-travel option in which the person can be engaged in in-home activity only (purposely or because of being sick) or can be out of town. In the model system application, a person who chooses a non-travel pattern is not considered further in the modeling stream, except that they can make an internal-external trip. Daily pattern-type choices of the household members are linked in such a way that decisions made by some members are reflected in the decisions made by the other members.

The next set of sub-models (4.2.1 - 4.2.3) defines the frequency, time-of-day, and mode for each mandatory tour. The scheduling of mandatory activities is generally considered a higher priority decision than any decision regarding non-mandatory activities for either the same person or for the other household members. "Residual time windows," or periods of time with no person-level activity, are calculated as the time remaining after tours have been scheduled. The temporal overlap of residual time windows among household members are estimated after mandatory tours have been generated and scheduled. Time window overlaps, which are left in the daily schedule after the mandatory commitment of the household members has been made, affect the frequency of joint and individual non-mandatory tours, and the probability of participation in joint tours. At-work sub-tours are modeled next, taking into account the time-window constraints imposed by their parent work tours (sub-models 4.5.1 - 4.5.4).

The next major model component relates to joint household travel. Joint tours are tours taken together by two or more members of the same household. This component predicts the exact number of joint tours by travel purpose and party composition (adults only, children only, or mixed) for the entire household (4.3.1), and then defines the participation of each household member in each joint household tour (4.3.2). It is followed by choice of destination (4.3.3) time-of-day (4.3.4), and mode (4.3.5).

The next stage relates to individual maintenance (escort, shopping, and other household-related errands) and discretionary (eating out, social/recreation, and other discretionary) tours. All of these tours are generated by person in model 4.4.1. Their destination, time of day, and mode are chosen next (4.4.2, 4.4.3, and 4.4.4).

The next set of sub-models relate to the stop-level details for each tour. They include the frequency of stops in each direction (5.2), the purpose of each stop (5.2), the location of each stop (5.3) and the stop departure time (5.4). It is followed by the last set of sub-models that add details for each trip including trip mode (6.1) and parking location for auto trips (6.2). The trips are then assigned to highway and transit networks depending on trip mode and time period (6.3).

Main sub-models and procedures of the core demand model

This section describes each model component in greater detail, including the general algorithm for each model, the decision-making unit, the choices considered, the market segmentation utilized (if any), and the explanatory variables used.

1.1 Population Synthesizer

The population synthesis procedure takes into account zonal and regional controls and includes a procedure to allocate households to MGRAs. A synthetic population is created using a modified open source PopSyn software originally designed for Atlanta Regional Commission (ARC). The ARC population synthesizer was developed by Parsons Brinckerhoff to be a flexible tool for creating synthetic populations for AB modeling. The population synthesizer inputs are U.S. Census data at the zonal- and regional-levels describing the distribution of households by various characteristics. The synthetic population is forced to match the zonal and regional characteristics. The ARC population synthesizer is being enhanced to consider person-level attributes in the population controls in order to match workers by occupation provided by PECAS.

The population synthesis approach includes the following steps:

- Create a sample of households in each TAZ (all households from the correspondent PUMA can be used in a simplified case).
- Balance the individual household weights to ensure the controlled totals across all person and household dimensions.
- Create a list of households by discretizing the individual weights.

The advantage of working with the list of households (compared to a multi-way distribution) is that both person and household variables can be incorporated. If only household or person attributes are controlled, the proposed procedure yields exactly the same multidimensional distribution as conventional matrix balancing. Also, the elimination of the drawing procedure allows for a theoretically closed formulation with no unnecessary empirical components.

General formulation

Since the procedure is applied for each TAZ separately, we formulate the model for a single TAZ. Introduce the following notation:

- $i = 1, 2, \dots, I$ = household and person controls,
- $n \in N$ = seed set of households in the PUMA (or any other sample),
- w_n = a priori weights assigned in the PUMA (or any other sample),
- A^i = zonal controls,
- $a_n^i \geq 0$ = coefficients of contribution of household to each control.

The principal flexibility of the procedure is that the contribution coefficients can take any non-negative value. In the conventional procedure, the contribution coefficients are implied to be Boolean incidence indicators (belong or not belong). An example is shown in Table T.6 for controls specified by household size and person age brackets.

Table T.6
Controls and Contribution Coefficients

HH ID	HH size				Person age				HH initial weight w_n
	1 $i = 1$	2 $i = 2$	3 $i = 3$	4+ $i = 4$	0-15 $i = 5$	16-35 $i = 6$	36-64 $i = 7$	65+ $i = 8$	
$n = 1$	1							1	20
$n = 2$		1			1	1			20
$n = 3$			1			1	2		20
$n = 4$				1		2	2		20
$n = 5$				1	1	3	2		20
....									...
Control	100	200	250	300	400	400	650	250	

The first household has one person of age 65+. The second household has two persons: one age 0-15, and one age 16-35. The third household has three persons: one age 16-35, and another two aged 36-64. The fourth household has four persons: two aged 16-35, and two aged 36-64. The fifth household has size persons: one person age 0-15, three persons aged 16-35, and two persons aged 36-64.

The balancing problem can be written as a convex mathematical program of the entropy-maximization type in the following way:

$$\min_{\{x_n\}} \sum_n x_n \ln \frac{x_n}{w_n}, \quad \text{Equation 1}$$

Subject to constraints:

$$\sum_n a_n^i x_n = A^i, (\alpha^i), \quad \text{Equation 2}$$

$$x_n \geq 0, \quad \text{Equation 3}$$

where α^i represents dual variables that give rise to balancing factors.

The objective function expresses the principle of using all households uniformly (proportionally to the assigned a priori weight). The constraints ensure matching the controls.

By forming the Lagrangian and equating the derivatives to zero we obtain the following solution:

$$x_n = k \times w_n \times \exp(\sum_i a_n^i \alpha^i) = w_n \times \prod_i [\exp(\alpha^i)]^{a_n^i} = w_n \times \prod_i (\hat{\alpha}^i)^{a_n^i}, \quad \text{Equation 4}$$

where $\hat{\alpha}^i$ represents balancing factors that have to be calculated. Note that the balancing factors correspond to the controls, not to households. For each household, the weight is calculated as a product of the initial weight by the relevant balancing factors exponentiated according to the participation coefficient. A zero participation coefficient automatically results in a balancing factor reset to 1 that does not affect the household weight.

Solution algorithm

The problem formulated in the previous section has a unique solution that can be achieved by the following iterative procedure:

Step 0: Set the iteration counter $k = 1$. Set zero-iteration weight $x_n(0,0) = w_n$.

For $k = 1$ to K (number of iterations):

For $i = 1$ to I (number of controls):

Step 1: Calculate balancing factor

$$\hat{\alpha}^i(k, i) = \frac{A^i}{\sum_n a_n^i x_n(k-1, i-1)}. \quad \text{Equation 5}$$

Step 2: Apply balancing factor (note exponentiation!)

$$x_n(k-1, i) = x_n(k-1, i-1) \times [\hat{\alpha}^i(k, i)]^{a_n^i}. \quad \text{Equation 6}$$

Step 3: Set starting weights for the next iteration

$$x_n(k, 0) = x_n(k-1, I). \quad \text{Equation 7}$$

Step 4: Calculate convergence criterion:

$$C(k) = \max_i \{ \text{abs}[\hat{\alpha}^i(k, i) - 1] \}. \quad \text{Equation 8}$$

If $C(k) \leq \varepsilon$ (degree of accuracy) or $k = K$ **Stop**.

Note that the solution is unique and independent of the order of controls. Normally, 100 iterations guarantee very good degree of convergence.

Base year controls

The population synthesizer first develops a “base year” population distribution using year 2000 Census or 2005-2009 ACS data, and a set of control attributes are defined. Census Summary File 1, Summary File 3, and the Census Transportation Planning Package information are used to develop single and multi-dimensional distributions of these attributes. These attributes, which are specified at the TAZ level in the base-year, include:

Household Controls:

- Housing Unit Type
- Household Size
- Household Income
- Number of Workers in Household
- Number of Units in Structure and Quality

Person Controls:

- Age
- Occupation

Once this distribution is established, the population synthesis tool samples PUMS records to create a fully enumerated representation of the population.

Household Controls

Each household is defined by eight dimensions. The dimensions are:

Household unit type (3)

Household
Non-Institutional Group Quarters
Institutional Group Quarters

Income in 2007 dollars (5)

<\$30k
\$30-60k
\$60-100k
\$100-150k
\$150k+

Household size (4)

1
2
3
4+

Number of Workers (4)

0
1
2
3+

Number of Units in Structure & Quality (8)

Single-Family Attached/Luxury
Single-Family Attached/Economy
Single-Family Detached/Luxury
Single-Family Detached/Economy
Multi-Family/Luxury
Multi-Family/Economy
Mobile Home
Military

Person Controls

Age (9)

0-17
18-24
25-34
35-49
50-64
65-69
80+

Occupation (7)

White collar labor
Work at home labor
Service labor
Health labor
Retail and food labor
Blue collar labor
Military labor

Group quarters residents are treated as a separate category of households. In the PUMS data, each group quarters resident has a record in the person format as well as a record in the household format representing a one-person pseudo-household containing only that individual. These fields are distinguished from the normal household records by the UNITTYPE field, which indicates if the record is a household record, a non-institutional group quarters record, or an institutional group quarters record. The UNITTYPE field is used to distinguish the type of household, and group quarters residents are otherwise treated just like any other household record. Institutional group quarters residents are generated so that the total population matches control totals. However, because institutional residents are not expected to travel, these records are not printed to the population output file used by the model system.

Combinations of the dimensions that are excluded or merged include:

- Illogical combinations of workers and household size are excluded.
- For group quarters, no distinctions are made by household income.
- For group quarters, no distinctions are made by household size.
- For group quarters, no distinctions are made by person dimensions.
- For group quarters, no distinction is made by the number of units in the structure.

Base-Year Control Totals

For the base-year application, the control totals are derived entirely from 2000 Census data tabulated at the block-group level and converted to a TAZ-level. The controls include:

- Households by Household Size (4 controls);
- Households by Household Size x Number of Workers (4x4=16 controls);
- Households by Household Income x Household Size (4x4=16 controls);
- Households by Household Income x Number of Workers (4x4=16 controls);
- Households By Household Income x Household Size x Number of Workers (3x4x4=48 controls);
- Households By Household Size x Number of Units (4x2=8 controls);
- Households By Number of Units (2 controls);
- Households By Group Quarters Type x Number of Workers (2x2=4 controls);
- Persons by age (9 controls); and
- Workers by occupation (7 controls).

Future-Year Control Totals

For the forecast years, a more limited set of control totals is available from PECAS. The forecast-year control totals from PECAS include:

- Housing type and quality (available at a TAZ level)
- Group Quarters (held constant except where known changes occur)
- Household income (available at an MGRA level, summarized to a TAZ level)
- Household size (will be available at a TAZ level)
- Workers per household (will be available at a TAZ level)
- Workers by occupation (available at a PECAS-zone level)
- Persons by age (county-level control)

This second IPF process results in a floating point future-year seed distribution for the 608 categories. That distribution is then converted to an integer seed distribution using a randomized rounding method. The randomized rounding works such that if a cell contains the value 0.14, it has an 86 percent chance of being rounded to 0, and a 14 percent chance of being rounded to 1. This randomized rounding is preferred because it avoids bias, but it does not guarantee that the total number of households in a TAZ exactly matches the targets. Households are drawn from the PUMS sample to fill this integer distribution and create the synthetic population. Any income values less than zero are set to zero prior to writing the population.

The forecast-year control totals are based on PECAS land-use model projections and other supplemental data (such as distributions of persons by age). PECAS operates at a 350 zone system, but also tracks certain data at the TAZ and parcel level. Housing type and quality, for example, are tracked at the TAZ level, while workers by occupation and place of residence are tracked at the PECAS-zone level. The distribution of persons by age is specified as a county-wide control.

The population synthesizer currently operates at the TAZ level. Every household is automatically assigned to a TAZ based on the marginal distributions generated for each TAZ. This model assigns an MGRA to each household as follows:

- The quantity of housing by type (single-family attached, single-family detached, multi-family, mobile-home, non-institutional group quarters, and military) will be summarized by MGRA (Q_h). This data is available at the parcel level.
- A probability for each housing type will be computed for each MGRA as the quantity of housing by type for the MGRA divided by the sum of housing by type across all MGRAs in the TAZ ($P_{i,h} = Q_{i,h} / \sum Q_h$).

A Monte-Carlo random number draw will be made for each synthetic household, and that household will select a residential MGRA based on its housing type and the probability distribution for that housing type across all MGRAs in the TAZ.

1.2 Accessibilities

All accessibility measures for the SANDAG ABM are calculated at the MGRA level. The auto travel times and cost are TAZ-based and the size variables are MGRA-based. This necessitates that auto accessibilities be calculated at the MGRA level. The SANDAG ABM requires accessibility indices only for non-mandatory travel purposes since the usual

location of work/school activity for each worker/student is modeled prior to the DAP, tour frequency, and tour destination choice for non-mandatory tours. In addition, school proximity to the residential MGRA, and travel time by transit for each student, can be used as an explanatory variable for escorting frequency.

The set of accessibility measures for the SANDAG ABM model is summarized in Table T.7.

Table T.7
Accessibility Measures for the SANDAG ABM

No.	Description	Model utilization	Attraction size variable S_j	Travel cost c_{ij}	Dispersion coefficient $-\gamma$
1	Access to non-mandatory attractions by SOV in off-peak	Car ownership	Total weighted employment for all purposes	Generalized SOV time including tolls	-0.05
2	Access to non-mandatory attractions by transit in off peak	Car ownership	Total weighted employment for all purposes	Generalized best path walk-to-transit time including fares	-0.05
3	Access to non-mandatory attractions by walk	Car ownership	Total weighted employment for all purposes	SOV off-peak distance (set to 999 if >3)	-1.00
4-6	Access to non-mandatory attractions by all modes except HOV	CDAP	Total weighted employment for all purposes	Off-peak mode choice logsums (SOV skims for ipersons) segmented by 3 car-availability groups	+1.00
7-9	Access to non-mandatory attractions by all modes except SOV	CDAP	Total weighted employment for all purposes	Off-peak mode choice logsums (HOV skims for interaction) segmented by 3 car-availability groups	+1.00
10-12	Access to shopping attractions by all modes except SOV	Joint tour frequency	Weighted employment for shopping	Off-peak mode choice logsum (HOV skims) segmented by 3 HH adult car-availability groups	+1.00
13-15	Access to maintenance attractions by all modes except SOV	Joint tour frequency	Weighted employment for maintenance	Off-peak mode choice logsum (HOV skims) segmented by 3 adult car-availability groups	+1.00
16-18	Access to eating-out attractions by all modes except SOV	Joint tour frequency	Weighted employment for eating out	Off-peak mode choice logsum (HOV skims) segmented by 3 adult HH car-availability groups	+1.00
19-21	Access to visiting attractions by all modes except SOV	Joint tour frequency	Total households	Off-peak mode choice logsum (HOV skims) segmented by 3 adult car-availability groups	+1.00
22-24	Access to discretionary attractions by all modes except SOV	Joint tour frequency	Weighted employment for discretionary	Off-peak mode choice logsum (HOV skims) segmented by 3 adult car-availability groups	+1.00
25-27	Access to escorting attractions by all modes except SOV	Allocated tour frequency	Total households	AM mode choice logsum (HOV skims) segmented by 3 adult car-availability groups	+1.00

Table T.7 Continued
Accessibility Measures for the SANDAG ABM

No.	Description	Model utilization	Attraction size variable S_j	Travel cost c_{ij}	Dispersion coefficient $-\gamma$
28-30	Access to shopping attractions by all modes except HOV	Allocated tour frequency	Weighted employment for shopping	Off-peak mode choice logsum (SOV skims) segmented by 3 adult car-availability groups	+1.00
31-33	Access to maintenance attractions by all modes except HOV	Allocated tour frequency	Weighted employment for maintenance	Off-peak mode choice logsum (SOV skims) segmented by 3 adult car-availability groups	+1.00
34-36	Access to eating-out attractions by all modes except HOV	Individual tour frequency	Weighted employment for eating out	Off-peak mode choice logsum (SOV skims) segmented by 3 car-availability groups	+1.00
36-39	Access to visiting attractions by all modes except HOV	Individual tour frequency	Total households	Off-peak mode choice logsum (SOV skims) segmented by 3 car-availability groups	+1.00
40-41	Access to discretionary attractions by all modes except HOV	Individual tour frequency	Weighted employment for discretionary	Off-peak mode choice logsum (SOV skims) segmented by 3 car-availability groups	+1.00
43-44	Access to at-work attractions by all modes except HOV	Individual sub-tour frequency	Weighted employment for at work	Off-peak mode choice logsum (SOV skims) segmented by adult 2 car-availability groups (0 cars and cars equal or greater than workers)	+1.00
45	Access to all attractions by all modes of transport in the peak	Work location, CDAP	Total weighted employment for all purposes	Peak mode choice logsums	+1.00
46	Access to at-work attractions by walk	Individual sub-tour frequency	Weighted employment for at work	SOV off-peak distance (set to 999 if >3)	+1.00
47	Access to all households by all modes of transport in the peak?		Total weighted households for all purposes	Generalized best path walk-to-transit time including fares	+1.00

Size Variables by Travel Purpose

Table T.8

Correspondence of LU Variables to Travel Purposes and Relative Attraction Rate

Employment by PECAS Model categories of Industry and other variables		Non-mandatory travel purpose in the ABM							
		4=escort	5=shop	6=main	7=eat	8=visit	9=disc	At-work	All
12	Retail Activity		3.194	0.776	0.325		0.098	0.154	3.970
13	Professional and Business Services			0.243				0.029	0.087
19	Amusement Services				0.089		0.364		0.407
20	Hotels Activity (479, 480)						0.318		
21	Restaurants and Bars		3.081		2.103	0.253	0.769	0.367	8.123
22	Personal Services Retail Based			0.500				0.054	0.999
23	Religious Activity						5.154		7.786
25	State and Local Government Enterprises Activity								
27	Federal Non-Military Activity			1.025					1.313
29/30	State and Local Non-Education Activity								0.214
	Total number of households	1.0				0.105	0.156		0.489

The size variable is calculated as a linear combination of the MGRA LU variables with the specified coefficients. The values of coefficients in the table have been estimated by means of an auxiliary regression model that used the LU variables as independent variables and expanded trip ends by travel purpose as dependent variables. The intercept was set to zero. The regressions were applied at the MGRA level (approximately 15,000 out of 33,334 MGRAs have non-zero values at least for some LU activity and/or observed trip ends).

The following travel cost functions are used in the accessibility calculations: generalized single-occupancy vehicle (SOV) time; generalized best path walk-to-transit time; SOV off-peak distance; off-peak mode choice logsum. These travel cost functions are explained.

- Generalized SOV time, including tolls and parking cost; time equivalent of tolls and operation cost should be included (approximately \$1 per 6 minutes, that is Value of Time (VOT)=\$10/h).
- Generalized best path walk-to-transit time including fares; this includes total in-vehicle time (reset to 10,000 if no transit path), walk, weight, transfer penalty, and time equivalent of fare according to the average VOT. It is suggested to use the relative in-vehicle and out-vehicle coefficients in the current mode choice model. First wait = 1.5, transfer wait = 3.0, Short walk (less than 1/4 mile) = 1.5, long walk (1/4 + miles) = 2.5, and there are additional transfer penalties equal to 2 minutes for the first transfer for LRT or Commuter rail only, and 15 minutes for all ride modes for the second transfer. The current cost coefficient is \$5.41/hour which is for the

middle income category; but I think we ought to use 1/2 of the average annual salary in San Diego in 2005 (which was \$43,824 according to BLS) divided by 2080 = \$10.53.

- SOV off-peak distance (set to 999 if distance>3) for non-motorized travel.
- Off-peak mode choice logsum calculated over 3 modes in trinary multinomial logit (auto/SOV skims, walk to transit, and non-motorized) segmented by 4 individual car-availability groups; the utility specifications are found in Table T.9.
- Off-peak mode choice logsum calculated over 3 modes in trinary multinomial logit (auto/HOV skims, walk to transit, and non-motorized) segmented by 3 household car-availability groups; the specifications are found in Table T.9. It should be noted that despite a large number of measures to be calculated (42), this set is not computationally intensive since the most detailed model portion (mode choice logsum) is calculated for only 7 different types (4 for individual activities and 3 for household joint activities). These 7 logsums are then combined with different size variables.

**Table T.9
Mode Utility Components for Accessibility Calculations**

Segment	Mode	Constant	Travel time		Cost, \$	
			Variable	Coefficient	Variable	Coefficient
U16 Adult, 0 cars	SOV*	-999	SOV time / off-peak	From the 4-step model	SOV toll plus operating cost plus parking	From the 4-step model
Adult, cars fewer than adults		2.0				
Adult, cars equal or greater than adults		3.5				
0 cars	HOV*	0.5	HOV time / off-peak	From the 4-step model	HOV toll plus operating cost plus parking	From the 4-step model
Cars fewer than adults		1.5				divided by 2
Cars equal or greater than adults		1.0				(if not scaled in the model)
U16 Adult, 0 cars	Transit (best path)	-0.5	Total in-vehicle time (10,000 if no transit path) plus weighted walk plus weighted wait plus transfer penalty as defined in the 4-step model	From the 4-step model	Fare	From the 4-step model
Adult, cars fewer than adults						
Adult, cars equal or greater than adults						
U16 Adult, 0 cars	Non-motorized		SOV off-peak distance (set to 999 if distance>3)	-1.00		
Adult, cars fewer than adults						
Adult, cars equal or greater than adults						

*Only one utility (SOV or HOV) is used at a time depending on the accessibility type as specified in Table T.7.

2.1 Pre-Mandatory Car Ownership Model

Number of Models:	1
Decision-Making Unit:	Household
Model Form:	Nested Logit
Alternatives:	Five (0, 1, 2, 3, 4++ autos)

The car ownership models predict the number of vehicles owned by each household. It is formulated as a nested logit choice model with five alternatives, including “no car”, “one car”, “two cars”, “three cars”, and “four or more cars.” The nesting structure is shown in Figure T.4.

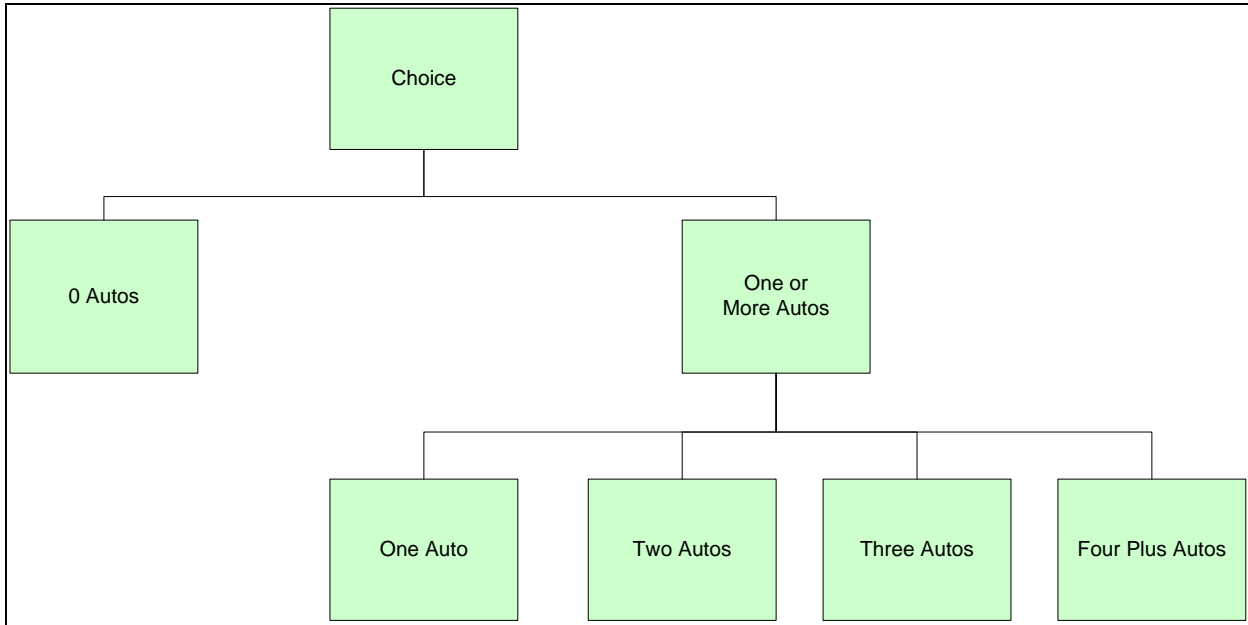
There are two instances of the auto ownership model. The first instance, model 2.1, is used to select a preliminary auto ownership level for the household, based upon household demographic variables, household ‘4D’ variables, and destination-choice accessibility terms created in sub-model 1.2 (see above). This auto ownership level is used to create mode choice logsums for workers and students in the household, which are then used to select work and school locations in model 2.2. The auto ownership model is re-run (sub-model 3.2) in order to select the actual auto ownership for the household, but this subsequent version is informed by the work and school locations chosen by model 2.2. All other variables and coefficients are held constant between the two models, except for alternative-specific constants.

The model includes the following explanatory variables:

- Number of driving-age adults in household
- Number of persons in household by age range
- Number of workers in household
- Number of high-school graduates in household
- Dwelling type of household
- Household income
- Intersection density (per acre) within one-half mile radius of household MGRA
- Population density (per acre) within one-half mile radius of household MGRA
- Retail employment density (per acre) within one-half mile radius of household MGRA
- Non-motorized accessibility from household MGRA to non-mandatory attractions (accessibility term #3)
- Off-peak auto accessibility from household MGRA to non-mandatory attractions (accessibility term #1)
- Off-peak transit accessibility from household MGRA to non-mandatory attractions (accessibility term #2)

Note that the model includes both household and person-level characteristics, ‘4D’ density measures, and accessibilities. The accessibility terms are destination choice (DC) logsums, which represent the accessibility of non-mandatory activities from the home location by various modes (auto, non-motorized, and transit). They are fully described under 1.2, above.

Figure T.4
Auto Ownership Nesting Structure



2.2 Work from Home Choice

Number of Models: 1
 Decision-Making Unit: Workers
 Model Form: Binary Logit
 Alternatives: Two (regular workplace is home, regular workplace is not home)

The work from home choice model determines whether each worker works from home. It is a binary logit model, which takes into account the following explanatory variables:

- Household income
- Person age
- Gender
- Worker education level
- Whether the worker is full-time or part-time
- Whether there are non-working adults in the household
- Peak accessibility across all modes of transport from household MGRA to employment (accessibility term #45, see section 1.2)

2.3 Mandatory (workplace/university/school) Activity Location Choice

Number of Models: 5 (Work, Preschool, K-8, High School, University)
 Decision-Making Unit: Workers for Work Location Choice; Persons Age 0-5 for Preschool, 6-13 for K-8; Persons Age 14-17 for High School; University Students for University Model
 Model Form: Multinomial Logit
 Alternatives: MGRAs

A *workplace location choice* model assigns a workplace MGRA for every employed person in the synthetic population who does not choose ‘works at home’ from Model 2.2. Every worker is assigned a regular work location zone (TAZ) and MGRA according to a multinomial logit destination choice model. Size terms in the model vary according to

worker occupation, to reflect the different types of jobs that are likely to attract different (white collar versus blue-collar) workers. There are six occupation categories used in the segmentation of size terms, as shown in Table T.2. Each occupation category utilizes different coefficients for categories of employment by industry, to reflect the different likelihood of workers by occupation to work in each industry. Accessibility from the workers home to the alternative workplace is measured by a mode choice logsum taken directly from the tour mode choice model, based on peak period travel (A.M. departure and P.M. return). Various distance terms are also used.

The explanatory variables in work location choice include:

- Household income
- Work status (full versus part-time)
- Worker occupation
- Gender
- Distance
- The tour mode choice logsum for the worker from the residence MGRA to each sampled workplace MGRA using peak level-of-service
- The size of each sampled MGRA

Since mode choice logsums are required for each destination, a two-stage procedure is used for all destination choice models in the CT-RAMP system for SANDAG in order to reduce computational time (it would be computationally prohibitive to compute a mode choice logsum for over 20,000 MGRAs and every tour). In the first stage, a simplified destination choice model is applied in which all TAZs are alternatives. The only variables in this model are the size term (accumulated from all MGRAs in the TAZ) and distance. This model creates a probability distribution for all possible alternative TAZs (TAZs with no employment are not sampled). A set of alternatives are sampled from the probability distribution and, for each TAZ, an MGRA is chosen according to its size relative to the sum of all MGRAs within the TAZ. These sampled alternatives constitute the choice set in the full destination choice model. Mode choice logsums are computed for these alternatives and the destination choice model is applied. A discrete choice of MGRA is made for each worker from this more limited set of alternatives. In the case of the work location choice model, a set of 40 alternatives is sampled.

The application procedure utilizes an iterative shadow pricing mechanism in order to match workers to input employment totals. The shadow pricing process compares the share of workers who choose each MGRA by occupation to the relative size of the MGRA compared to all MGRAs. A shadow price is computed which scales the size of the MGRA based on the ratio of the observed share to the estimated share. The model is re-run until the estimated and observed shares are within a reasonable tolerance. The shadow prices are written to a file and can be used in subsequent model runs to cut down computational time.

There are four school location choice models: (1) a pre-school model, (2) a grade school model, (3) a high school model, and (4) a university model.

The pre-school mandatory location choice model assigns a school location for pre-school children (person type 8) who are enrolled in pre-school and daycare. The size term for this model includes a number of employment types and population, since daycare and pre-school enrollment and employment are not explicitly tracked in the input land-use data. Explanatory variables include:

- Income
- Age
- Distance
- The tour mode choice logsum for the student from the residential MGRA to each sampled pre-school MGRA using peak levels-of-service
- Size of each sampled pre-school MGRA

The grade school location choice model assigns a school location for every K-8 student in the synthetic population; the size term for this model is K-8 enrollment. School district boundaries are used to restrict the choice set of potential school location zones based on residential location. The explanatory variables used in the grade school model include:

- School district boundaries
- Distance
- The tour mode choice logsum for the student from the residence MGRA to the sampled school MGRA using peak levels-of-service
- The size of the school MGRA

The high school location choice model assigns a school location for every high-school student in the synthetic population; the size term for this model is high school enrollment. District boundaries are also used in the high school model to restrict the choice set. The explanatory variables in the high school model include:

- School district boundaries
- Distance
- The tour mode choice logsum for the student from the residence MGRA to the sampled school MGRA using peak levels-of-service
- The size of the school MGRA

A university location choice model assigns a university location for every university student in the synthetic population. There are three types of college/university enrollment in the input land-use data file: (1) College enrollment, which measures enrollment at major colleges and universities; (2) other college enrollment, which measures enrollment at community colleges, and (3) adult education enrollment, which includes trade schools and other vocational training. The size terms for this model are segmented by student age, where students aged less than 30 use a 'typical' university size term, which gives a lower weight to adult education enrollment, while students age 30 or greater have a higher weight for adult education.

Explanatory variables in the university location choice model include:

- Student worker status
- Student age
- Distance
- Tour mode choice logsum for student from residence MGRA to sampled school MGRA using peak levels-of-service

3.1 Employer Parking Provision Model

Number of Models: 1
Decision-Making Unit: Workers whose workplace is in CBD or other priced-parking area (park area 1)
Model Form: Multinomial Logit
Alternatives: Three (Free on-site parking, parking reimbursement, and no parking provision)

The *Employer Parking Provision* Model predicts which persons have on-site parking provided to them at their workplaces and which persons receive reimbursement for off-site parking costs. The provision model takes the form of a multinomial logit discrete choice between free on-site parking, parking reimbursement (including partial or full reimbursement of off-site parking and partial reimbursement of on-site parking) and no parking provision.

It should be noted that free-onsite parking is not the same as full reimbursement. Many of those with full reimbursement in the survey data could have chosen to park closer to their destinations and accepted partial reimbursement. Whether parking is fully reimbursed will be determined both by the reimbursement model and the location choice model.

Persons with workplaces outside of *park area1* are assumed to receive free parking at their workplaces.

Explanatory variables in the provision model include:

- Household income
- Occupation
- Average daily equivalent of monthly parking costs in nearby MGRAs

3.2 Car Ownership Model

Number of Models: 1
Decision-Making Unit: Households
Model Form: Nested Logit
Alternatives: Five (0, 1, 2, 3, 4+ autos)

The car ownership model is described under 2.1, above. The model is re-run after work/school location choice, so that auto ownership can be influenced by the actual work and school locations predicted by model 3.1.

The explanatory variables in model 3.2 include the ones listed under 2.1 above, with the addition of the following:

- A variable measuring auto dependency for workers in the household based upon their home to work tour mode choice logsum
- A variable measuring auto dependency for students in the household based upon their home to school tour mode choice logsum
- A variable measuring the time on rail transit (light-rail or commuter rail) as a proportion of total transit time to work for workers in the household

- A variable measuring the time on rail transit (light-rail or commuter rail) as a proportion of total transit time to school for students in the household

The household mandatory activity auto dependency variable is calculated using the difference between the single-occupant vehicle (SOV) and the walk to transit mode choice logsum, stratified by person type (worker versus student). The logsums are computed based on the household MGRA and the work MGRA (for workers) or school MGRA (for students). The household auto dependency is obtained by aggregating individual auto dependencies of each person type (worker versus student) in the household. The auto dependency variable is calculated according to the following formula:

$$\text{Dependency}_{\text{auto}} = \min((\text{Logsum}_{\text{auto}} - \text{Logsum}_{\text{transit}}) / 3, 1.0) * \text{Factor}_{\text{non-motorized}}$$

Where:

$$\text{Factor}_{\text{non-motorized}} = 0.5 * \min((\max(\text{Distance}_{\text{home,work/school}}, 1.0), 3.0)) - 0.5$$

The non-motorized factor takes a value of 0 if the distance between home and work or school is less than one mile. If the distance between home and work/school is between one and three miles, the factor takes a value between 1.0 and 3.0. If the distance between home and work/school is greater than 3 miles (which serves as an upper cap on walkability), the non-motorized factor takes the maximum value of 1.0. The effect of this factor is to reduce the auto dependency variable if the work or school location is within walking distance of the residential MGRA.

The difference between auto and transit utility is divided by 3.0 to represent the resulting utility difference in units of hours (assuming an 'average' time coefficient of -0.05 multiplied by 60 minutes per hour). The difference is capped at 1.0, in effect representing the difference in scaled utility as a fraction between zero and one.

The household mandatory activity rail mode index is calculated using the ratio of the rail mode in-vehicle time over the total transit in-vehicle time for trips that used rail as part of their transit path, stratified by person type (worker versus student). The household rail mode index is obtained by aggregating individual rail indices of worker/student members in the household. All mandatory mode choice logsums and accessibilities are calculated using AM peak skims.

3.3 Toll Transponder Ownership Model

Number of Models:	1
Decision-Making Unit:	Households
Model Form:	Binomial Logit
Alternatives:	Two (Yes or No)

This model predicts whether a household owns a toll transponder unit. It was estimated based on aggregate transponder ownership data using a quasi-binomial logit model to account for over-dispersion. It predicts the probability of owning a transponder unit for each household based on aggregate characteristics of the zone.

The explanatory variables in the model include:

- Percent of households in the zone with more than one auto
- The number of autos owned by the household
- The straight-line distance from the MGRA to the nearest toll facility, in miles
- The average transit accessibility to non-mandatory attractions using off-peak levels-of-service (accessibility measure #2)

- The average expected travel time savings provided by toll facilities to work
- The percent increase in time to downtown San Diego incurred if toll facilities were avoided entirely

The accessibility terms are destination choice (DC) logsums, which represent the accessibility of non-mandatory activities from the home location by various modes (auto, non-motorized, and transit). They are fully described under 1.2, above.

The average expected travel time savings provided by toll facilities to work is calculated using a simplified destination choice logsum. The expected travel time savings of households in a zone z is:

$$\frac{\sum_d (AutoTime_{zd} - TollTime_{zd}) \cdot Employment_d \cdot \exp(-0.01 AutoTime_{zd})}{\sum_d Employment_d \cdot \exp(-0.01 AutoTime_{zd})}$$

The times are calculated in minutes and include both the AM peak travel time to the destination and the PM peak time returning from the destination. The percent difference between the AM non-toll travel time to downtown zone 3781 and the AM non-toll travel time to downtown when the general purpose lanes parallel to all toll lanes requiring transponders were made unavailable in the path-finder. This variable is calculated as:

$$\frac{Non\ Toll\ Time\ Avoiding\ Facility - Non\ Toll\ Time}{Non\ Toll\ Time}$$

4.1 Coordinated Daily Activity Pattern (DAP) Model

Number of Models:	1
Decision-Making Unit:	Households
Model Form:	Multinomial Logit
Alternatives:	691 total alternatives, but depends on household size (see Table T.10)

This model predicts the main daily activity pattern (DAP) type for each household member. The activity types that the model considers are:

- **Mandatory pattern (M)** that includes at least one of the three mandatory activities – work, university or school. This constitutes either a workday or a university/school day, and may include additional non-mandatory activities such as separate home-based tours or intermediate stops on the mandatory tours.
- **Non-mandatory pattern (N)** that includes only maintenance and discretionary tours. Note that the way in which tours are defined, maintenance and discretionary tours cannot include travel for mandatory activities.
- **At-home pattern (H)** that includes only in-home activities. At-home patterns are not distinguished by any specific activity (e.g., work at home, take care of child, being sick, etc.). Cases where someone is not in town (e.g., business travel) are also combined with this category.

Statistical analysis performed in a number of different regions has shown that there is an extremely strong correlation between DAP types of different household members, especially for joint N and H types. For this reason, the DAP for different household members should not be modeled independently, as doing so would introduce significant error in the types of activity patterns generated at the household level. This error has implications for a number of policy sensitivities, including greenhouse gas policies. Therefore, the model is applied across all household members simultaneously; the interactions or influences of different types of household members (e.g., the effect of a child who stays at home on the simulation day on the probability of a part-time worker also staying at home) is taken into account through a specific set of interaction variables.

The model also simultaneously predicts the presence of fully-joint tours for the household. Fully-joint tours are tours in which two or more household members travel together for all stops on the tour. Joint tours are only a possible

alternative at the household level when two or more household members have an active (M or N) travel day. The joint tour indicator predicted by this model is then considered when generating and scheduling mandatory tours, in order to reflect the likelihood of returning home from work earlier in order to participate in a joint tour with other household members.

The choice structure includes 363 alternatives with no joint travel and 328 alternatives with joint travel, totaling to 691 alternatives as shown in Table T.10. Note that the choices are available based on household size. There are also two facets of the model that reduce the complexity. First of all, mandatory DAP types are only available for appropriate person types (workers and students). Even more importantly, intra-household coordination of DAP types is relevant only for the N and H patterns. Thus, simultaneous modeling of DAP types for all household members is essential only for the trinary choice (M, N, H), while the sub-choice of the mandatory pattern can be modeled for each person separately.

Table T.10
Number of Choices in CDAP Model

Household Size	Alternatives – no Joint Travel	Alternatives with Joint Travel	All Alternatives
1	3	0	3
2	$3 \times 3 = 9$	$3 \times 3 - (3 \times 2 - 1) = 4$	13
3	$3 \times 3 \times 3 = 27$	$3 \times 3 \times 3 - (3 \times 3 - 2) = 20$	47
4	$3 \times 3 \times 3 \times 3 = 81$	$3 \times 3 \times 3 \times 3 - (3 \times 4 - 3) = 72$	153
5 or more	$3 \times 3 \times 3 \times 3 \times 3 = 243$	$3 \times 3 \times 3 \times 3 \times 3 - (3 \times 5 - 4) = 232$	475
Total	363	328	691

The structure is shown graphically in Figure T.5 for a three-person household. Each of the 27 daily activity pattern choices is made at the household level and describes an explicit pattern-type for each household member. For example, the fourth choice from the left is person 1 mandatory (M), person 2 non-mandatory (N), and person 3 mandatory (M). The exact tour frequency choice is a separate choice model conditional upon the choice of alternatives in the trinary choice. This structure is much more powerful for capturing intra-household interactions than sequential processing. The choice of 0 or 1+ joint tours is shown below the DAP choice for each household member. The choice of 0 or 1+ joint tours is active for this DAP choice because at least two members of the household would be assigned active travel patterns in this alternative.

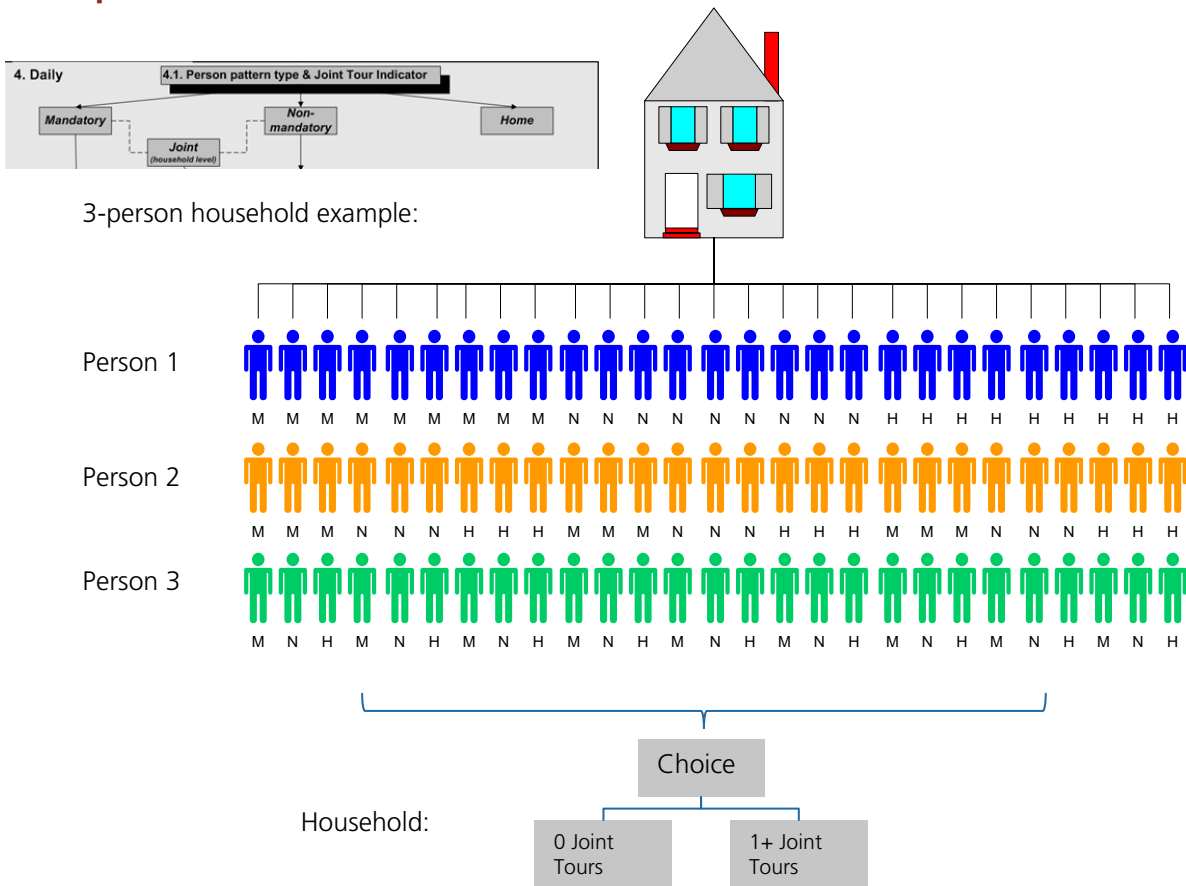
For a limited number of households of size greater than five, the model is applied for the first five household members by priority while the rest of the household members are processed sequentially, conditional upon the choices made by the first five members. The rules by which members are selected for inclusion in the main model are that first priority is given to any full-time workers (up to two), then to any part-time workers (up to two), then to children, youngest to oldest (up to three).

The CDAP model explanatory variables include:

- Household Size
- Number of Adults in household
- Number of children in household

- Auto Sufficiency
- Household Income
- Dwelling Type
- Person type
- Age
- Gender
- Usual Work location
- The tour mode choice logsum for the worker from the residential MGRA to each sampled workplace MGRA using peak levels-of-service
- The tour mode choice logsum for the student from the residential MGRA to each sampled school MGRA using peak levels-of-service
- Accessibility across all modes of transport from household MGRA to retail employment or non-mandatory locations (accessibility term #45, see section 1.2)

Figure T.5
Example of DAP Model Alternatives for a 3-Person Household



4.2.1 Individual Mandatory Tour Frequency

Number of Models: 1
Decision-Making Unit: Persons
Model Form: Multinomial Logit
Alternatives: 5 (1 Work Tour, 2 Work Tours, 1 School Tour, 2 School Tours, 1 Work/1 School Tour)

Based on the DAP chosen for each person, individual mandatory tours, such as work, school and university tours are generated at person level. The model is designed to predict the exact number and purpose of mandatory tours (e.g., work and school/university) for each person who chose the mandatory DAP type at the previous decision-making stage. Since the DAP type model at the household level determines which household members engage in mandatory tours, all persons subjected to the individual mandatory tour model implement at least one mandatory tour. The model has the following five alternatives: (1) 1 Work Tour, (2) 2 or more Work Tours, (3) 1 School Tour, (4) 2 or more School Tours, (5) 1 Work/1 School Tour.

DAPs and subsequent behavioral models of travel generation include these explanatory variables:

- Auto sufficiency
- Household income
- Non-family household indicator
- Number of preschool children in household
- Number of school aged children 6-18 years old in household NOT going to school
- Person type
- Gender
- Age
- Distance to work location
- Distance to school location
- Best travel time to work location
- HOV accessibility from household MGRA to employment (accessibility terms #25, 26, 27 (by auto sufficiency) , see section 1.2)

4.2.2 Individual Mandatory Tour Time of Day Choice

Number of Models: 3 (Work, University, and School)
Decision-Making Unit: Persons
Model Form: Multinomial Logit
Alternatives: 820 (combinations of tour departure half-hour and arrival half-hour back at home, with aggregation between 1 AM and 5 AM)

After individual mandatory tours have been generated, the tour departure time from home and arrival time back at home is chosen simultaneously. Note that it is not necessary to select the destination of the tour, as this has already been determined in Model 2.1. The model is a discrete choice construct that operates with tour departure from home and arrival back home time combinations as alternatives. The proposed utility structure is based on “continuous shift” variables, and represents an analytical hybrid that combines the advantages of a discrete choice structure (flexible in specification and easy to estimate and apply) with the advantages of a duration model (a simple structure with few parameters, and which supports continuous time). The model has a temporal resolution of one-half hour that is

expressed in 820 half-hour departure/arrival time alternatives. The model utilizes direct availability rules for each subsequently scheduled tour, to be placed in the residual time window left after scheduling tours of higher priority. This conditionality ensures a full consistency for the individual entire-day activity and travel schedule as an outcome of the model.

In the CT-RAMP model structure, the tour-scheduling model is placed after destination choice and before mode choice. Thus, the destination of the tour and all related destination and origin-destination attributes are known and can be used as variables in the model estimation.

For model estimation, the following practical rules can be used to set the alternative departure/arrival time combinations:

- Each reported/modeled departure/arrival time is rounded to the nearest half-hour. For example, the half-hour "17" includes all times from 10:45 A.M. to 11:14 A.M.
- Any times before 5 A.M. are shifted to 5 A.M., and any times after 1 A.M. are shifted to 1 A.M. This typically results in a shift for relatively few cases, and limits the number of half-hours in the model to 41.
- Every possible combination of the 41 departure half-hours with the 41 arrival half-hours (where the arrival half-hour is the same or later than the departure hour) is an alternative. This gives $41 \times 42/2 = 861$ choice alternatives.

The network simulations to obtain travel time and cost skims are implemented for five broad periods: (1) early A.M., (2) A.M. peak, (3) midday, (4) P.M. peak, and (5) night (evening, and late night) for the three mandatory tour purposes (work, university, and school).

The model includes the following explanatory variables:

- Household income
- Person type
- Gender
- Age
- Mandatory tour frequency
- Auto travel distance
- Destination employment density
- Tour departure time
- Tour arrival time
- Tour duration
- The tour mode choice logsum by tour purpose from the residence MGRA to each sampled MGRA location

4.2.3 Individual Mandatory Tour Mode Choice Model

Number of Models:	3 (Work, University, K-12)
Decision-Making Unit:	Person
Model Form:	Nested Logit
Alternatives:	26 (See Figure T.6)

This model determines how the “main tour mode” (used to get from the origin to the primary destination and back) is determined. The tour-based modeling approach requires a certain reconsideration of the conventional mode choice structure. Instead of a single mode choice model pertinent to a four-step structure, there are two different levels where the mode choice decision is modeled:

- The tour mode level (upper-level choice)
- The trip mode level (lower-level choice conditional upon the upper-level choice)

The tour mode choice model considers the following alternatives:

- Drive-alone
- Shared-Ride 2
- Shared-Ride 3+
- Walk
- Bike
- Walk-Transit
- Park-and-Ride Transit (drive to transit station and ride transit)
- Kiss-and-Ride Transit (drop-off at transit station and ride transit)
- School Bus (only available for grade school and high school tour purposes)

The mode of each tour is identified based on the combination of modes used for all trips on the tour, according to the following rules:

- If any trip on the tour is Park-and-Ride Transit, then the tour mode is Park-and-Ride Transit.
- If any trip on the tour is Kiss-and-Ride Transit, then the tour mode is Kiss-and-Ride Transit.
- If any trip on the tour is School Bus, then the tour mode is School Bus.
- If any trip on the tour is Walk-Transit, then the tour mode is Walk-Transit.
- If any trip on the tour is Bike, then the tour mode is Bike.
- If any trip on the tour is Shared-Ride 3+, then the tour mode is Shared-Ride 3+
- If any trip on the tour is Shared-Ride 2, then the tour mode is Shared-Ride 2.
- If any trip on the tour is Drive-Along, then the tour mode is Drive-Along.
- All remaining tours are Walk.

These tour modes create a hierarchy of importance that ensures that transit is available for trips on tours with transit as the preferred mode, and that high-occupancy vehicle lanes are available for trips on tours where shared-ride is the preferred mode. It also ensures that if drive-transit is utilized for the outbound trip on the tour, that mode is also available for the return journey (such that the traveler can pick up their car at the parking lot on the way home).

Modes for the tour mode choice model are shown in Figure T.6. The model is distinguished by the following characteristics:

- Segmentation of the HOV mode by occupancy categories, which is essential for modeling specific HOV/HOT lanes and policies.
- An explicit modeling of toll vs. non-toll choices as highway sub-modes, which is essential for modeling highway pricing projects and policies.
- Distinguishing between certain transit sub-modes that are characterized by their attractiveness, reliability, comfort, convenience, and other characteristics beyond travel time and cost (such as Express Bus, Bus-Rapid Transit, Light-Rail Transit, and Commuter Rail).
- Distinguishing between walk and bike modes if the share of bike trips is significant.

Note that free and pay alternatives for each auto mode provide an opportunity for toll choice as a path choice within the nesting structure. This requires separate free and pay skims to be provided as inputs to the model (where free paths basically “turn off” all toll and HOT lanes). Transit skims are segmented by local versus premium (express bus, BRT, LRT, and commuter rail) modes, but as described, the mode used for the longest segment of in-vehicle time is used to define the actual premium ride mode in the creation of transit level-of-service. Transit ride modes are based on a modal hierarchy in which modes that are ranked lower in the hierarchy are used as feeder modes to modes ranked higher. Table T.11 describes skims used in tour mode choice. A number of mode codes have been reserved for future use.

The tour mode choice model is based on the round-trip (outbound and return) level-of-service (LOS) between the tour anchor location (home for home-based tours and work for at-work sub-tours) and the tour primary destination. The tour mode choice model assumes that the mode of the outbound journey is the same as the mode for the return journey in the consideration of level-of-service information. This is a simplification that results in a model with a relatively modest number of alternatives, and also allows the estimation process to utilize data from an on-board survey in which the mode for only one direction is known. Only these aggregate tour modes are used in lower level model components such as stop frequency, stop location, and as constraints in trip mode choice.

However, the estimation and application process calculates utilities for a more disaggregate set of modes in lower level alternatives that are consistent with the more detailed modes in trip mode choice. This allows the tour mode choice model to consider the availability of multiple transit line-haul modes and/or managed lane route choices in the choice of tour mode, with their specific levels-of-service and modal constants. The more aggregate tour modes act as constraints in trip mode choice; for example, if walk-transit is chosen in tour mode choice, only shared-ride, walk, and walk-transit modes are available in trip mode choice. Ultimately, trips are assigned to networks using the more disaggregate trip modes.

The lower level nest mode choices (which are the same as the trip mode choice model alternatives) are:

- Drive-alone Free
- Drive-Alone Pay
- Shared-Ride 2 Free (General Purpose Lane)
- Shared-Ride 2 Free (HOV Lane)
- Shared-Ride 2 Pay
- Shared-Ride 3+ Free (General Purpose Lane)
- Shared-Ride 3+ Free (HOV Lane)
- Shared-Ride 3+ Pay
- Walk
- Bike
- Walk-Local Bus
- Walk-Express Bus
- Walk-Bus Rapid Transit
- Walk-Light Rail Transit
- Walk-Commuter Rail
- Park and Ride (PNR)-Local Bus
- PNR-Express Bus
- PNR-Bus Rapid Transit
- PNR-Light Rail Transit
- PNR-Commuter Rail
- Kiss and Ride (KNR)-Local Bus
- KNR-Express Bus
- KNR-Bus Rapid Transit
- KNR-Light Rail Transit
- KNR-Commuter Rail
- School Bus

The appropriate skim values for the tour mode choice are a function of the MGRA of the tour origin and MGRA of the tour primary destination. As described in the section on Treatment of Space, all transit level-of-service and certain non-motorized level of service (for MGRAs within 1.5 miles of each other) are computed “on-the-fly” in mode choice. Transit access and egress times are specifically determined via detailed MGRA-to-TAP distances computed within

Geographic Information System (GIS) software. Actual TAP-TAP pairs used for the MGRA-pair, and therefore actual transit levels-of-service, are based on a selection of the path with the best overall utility for each of five transit ride modes (local bus, express bus, bus rapid-transit, light-rail, and heavy rail).

Figure T.6
Tour Mode Choice Model Structure

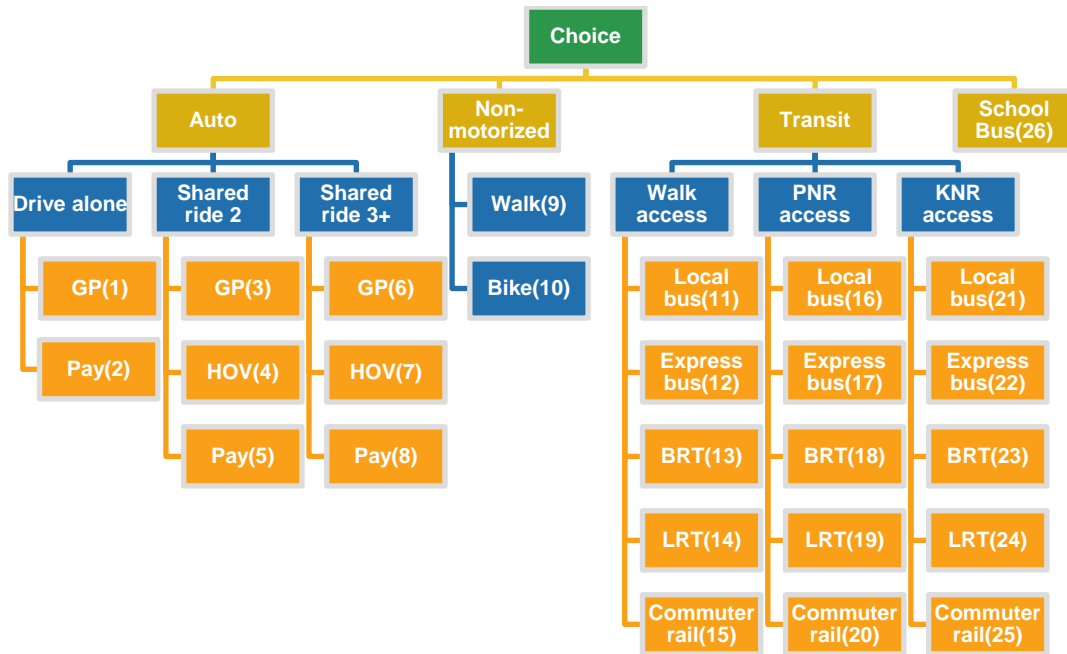


Table T.11

Skims Used in Tour Mode Choice

Mode	Skims
Drive-alone Non-Toll	All general purpose lanes available. HOV lanes, HOT lanes, and toll lanes unavailable. Toll bridges are available.
Drive-alone Toll	All general purpose lanes and toll lanes are available. HOV lanes are unavailable. HOT lanes are available for the SOV toll rate. Toll bridges are available.
Shared-2 Non-Toll, Non-HOV	All general purpose lanes available. HOV lanes, HOT lanes, and toll lanes unavailable. Toll bridges are available.
Shared-2 Non-Toll, HOV	All general purpose lanes available. 2+ occupancy HOV lanes available. Toll lanes unavailable. HOT lanes where 2+ occupant vehicles go free are available. Toll bridges are available.
Shared-2 Toll, HOV	All general purpose lanes available. 2+ occupancy HOV lanes and HOT lanes where 2+ occupant vehicles go free are available for free. Toll lanes and HOT lanes where 2-occupant vehicles are tolled at the 2-occupant toll rate. Toll bridges are available.
Shared-3+ Non-Toll, Non-HOV	All general purpose lanes available. HOV lanes, HOT lanes, and toll lanes unavailable. Toll bridges are available.
Shared-3+ Non-Toll, HOV	All general purpose lanes available. 2+ and 3+ occupancy HOV lanes available. Toll lanes unavailable. HOT lanes where 2+ or 3+ occupant vehicles go free are available. Toll bridges are available.
Shared-3+ Toll, HOV	All general purpose lanes available. 2+ and 3+ occupancy HOV lanes and HOT lanes where 2 or 3+ occupant vehicles go free are available for free. Toll lanes and HOT lanes where 3+ occupant vehicles are tolled at the 3+ occupant toll rate. Toll bridges are available.
Walk	Highway distance, excluding freeways, but allowing select bridges with sidewalks. This is used for any MGRA-pair whose distance is greater than 1.5 miles. The walk time for MGRA-pairs whose distance is less than 1.5 miles relies on the GIS-based walk distances.
Bike	Highway distance, excluding freeways, but allowing select bridges with bike lanes. This is used for any MGRA-pair whose distance is greater than 1.5 miles. The bike time for MGRA-pairs whose distance is less than 1.5 miles relies on the GIS-based bike distances.
Transit-Local	Local Bus TAP-to-TAP skims, including in-vehicle time, first wait time, transfer wait time, and fare.
Transit-Premium	Premium TAP-to-TAP skims, including in-vehicle time, first wait time, transfer wait time, and fare. These include local bus as a feeder mode, as well as express bus, bus rapid transit, light rail, and commuter rail. A premium mode designator is also included in the skim for each interchange, to identify which of the four premium ride-modes is used, based on the mode for which the greatest distance was travelled.

The individual mandatory tour mode choice model contains the following explanatory variables:

- Auto sufficiency
- Household size
- Age
- Gender
- In-vehicle time (auto and transit)
- Walk and bike time
- Auto operating cost
- Auto parking cost
- Auto terminal time
- Auto toll value
- Transit first wait time
- Transit transfer time
- Number of transit transfers
- Transit walk access time
- Transit walk egress time
- Transit walk auxiliary time
- Transit fare
- Transit drive access time
- Transit drive access cost
- Intersection density
- Employment density
- Dwelling unit density

4.3 Generation of Joint Household Tours

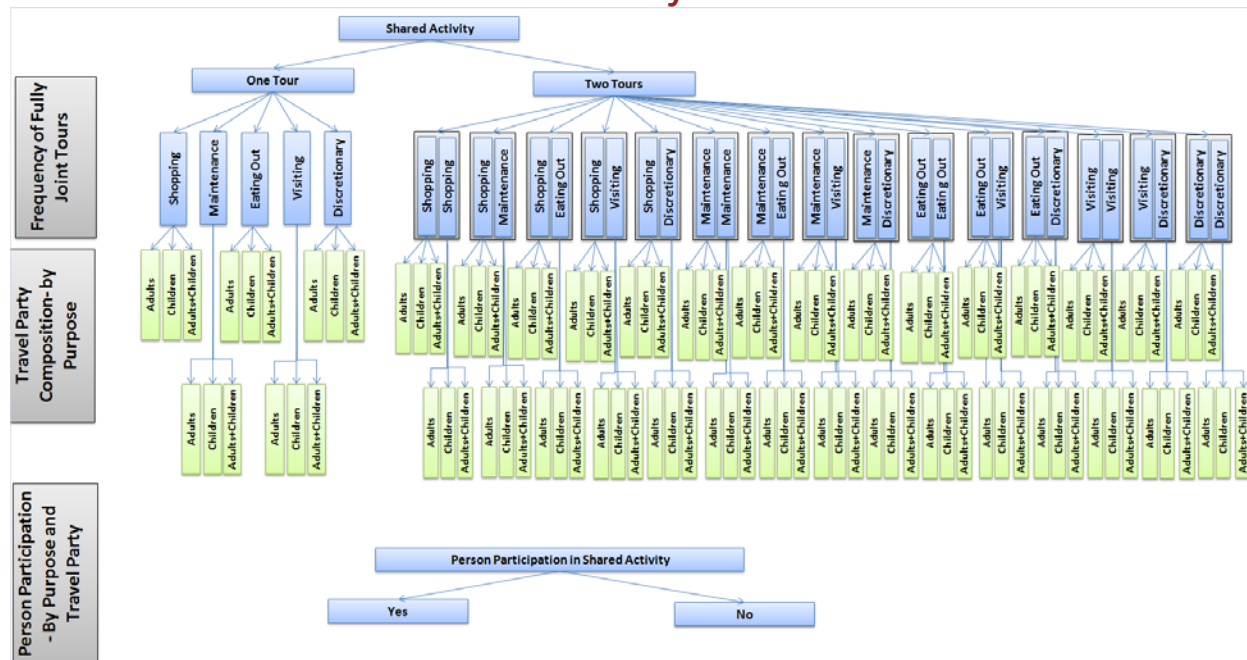
In the CT-RAMP structure, joint travel for non-mandatory activities is modeled explicitly in the form of fully joint tours (where all members of the travel party travel together from the beginning to the end and participate in the same activities). This accounts for more than 50 percent of joint travel.

Each fully joint tour is considered a modeling unit with a group-wise decision-making process for the primary destination, mode, frequency and location of stops. Modeling joint activities involves two linked stages – see Figure T.7.

- A tour generation and composition stage that generates the number of joint tours by purpose/activity type made by the entire household. This is the joint tour frequency model.

- A tour participation stage at which the decision whether to participate or not in each joint tour is made for each household member and tour.

Figure T.7
Model Structure for Joint Non-Mandatory Tours



Joint tour party composition is modeled for each tour. Travel party composition is defined in terms of person categories (e.g., adults and children) participating in each tour. Person participation choice is then modeled for each person sequentially. In this approach, a binary choice model is calibrated for each activity, party composition and person type. The model iterates through household members and applies a binary choice to each to determine if the member participates. The model is constrained to only consider members with available time-windows overlapping with the generated joint tour. The approach offers simplicity but at the cost of overlooking potential non-independent participation probabilities across household members. The joint tour frequency, composition, and participation models are described below.

4.3.1 Joint Tour Frequency and Composition

Number of Models: 1
 Decision-Making Unit: Households with a Joint Tour Indicator predicted by the CDAP model
 Model Form: Multinomial Logit
 Alternatives: 105 (1 Tour segmented by 5 purposes and 3 composition classes, 2 tours segmented by 5 purposes and 3 composition classes)

Joint tour frequencies (1 or 2+) are generated by households, purpose, and tour composition (adults only, children only, adults and children). Later models determine who in the household participates in the joint tour. The model is only applied to households with a joint tour indicator at the household level, as predicted by the CDAP model.

The explanatory variables in the joint tour frequency model include:

- Auto sufficiency
- Household income
- Number of full time workers in household
- Number of part time workers in household

- Number of university students in household
- Number of non-workers in household
- Number of retirees in household
- Number of driving age school children in household
- Number pre-driving age school children in household
- Number of preschool children in household
- Number of adults in household not staying home
- Number of children in household not staying home
- Shopping HOV Accessibility from household MGRA to employment (accessibility terms #10, 11, 12 (by auto sufficiency), see section 1.2
- Maintenance HOV Accessibility from household MGRA to employment (accessibility terms #13, 14, 15 (by auto sufficiency), see section 1.2
- Discretionary HOV Accessibility from household MGRA to employment (accessibility terms #22, 23, 24 (by auto sufficiency), see section 1.2
- Presence and size of overlapping time-windows, which represent the availability of household members to travel together after mandatory tours have been generated and scheduled

4.3.2 Joint Tour Participation

Number of Models:	1
Decision-Making Unit:	Persons
Model Form:	Multinomial Logit
Alternatives:	2 (Yes or No)

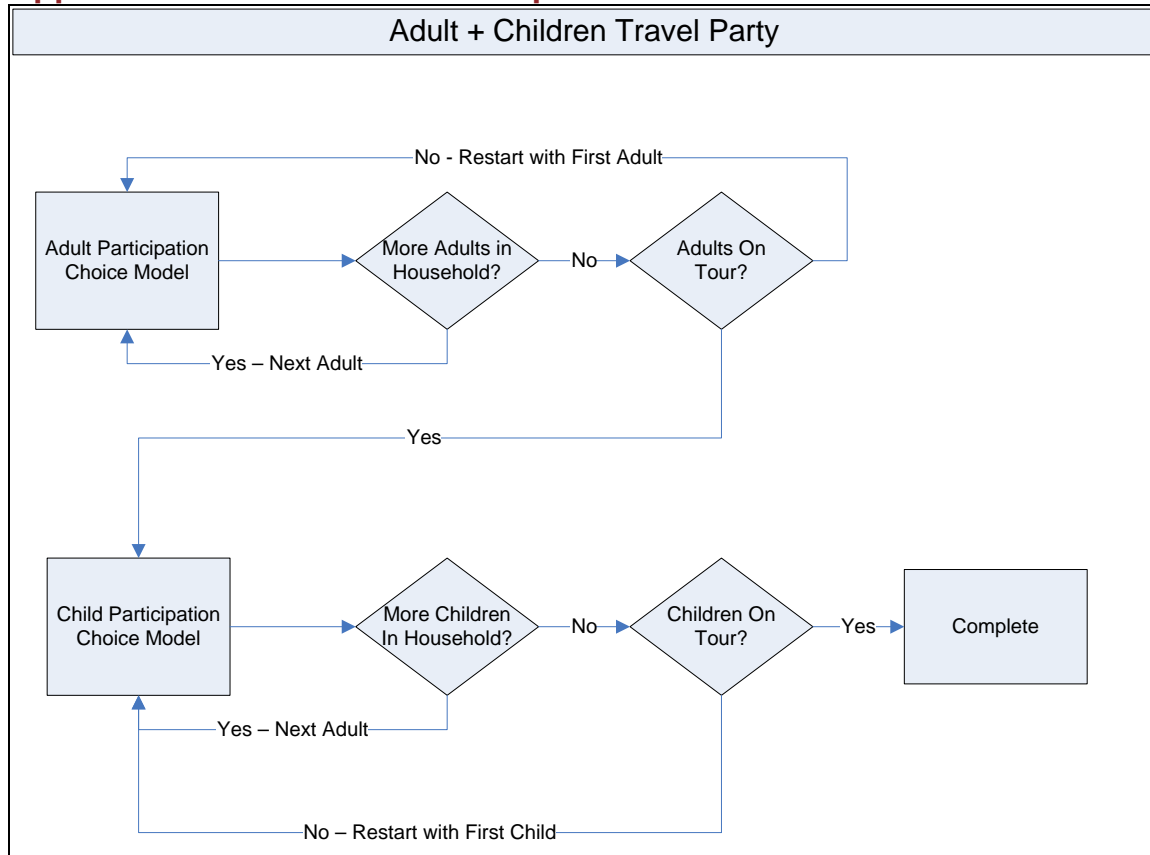
Joint tour participation is modeled for each person and each joint tour. If the person does not correspond to the composition of the tour determined in the joint tour composition model, they are ineligible to participate in the tour. Similarly, persons whose daily activity pattern type is home are excluded from participating. The model relies on heuristic process to assure that the appropriate persons participate in the tour as per the composition model. The model follows the logic depicted in Figure T.8.

The explanatory variables in the participation model include:

- Auto sufficiency
- Household income
- Frequency of joint tours in the household
- Number of adults (not including decision-maker) in household
- Number of children (not including decision-maker) in household
- Person type
- Maximum pair-wise overlaps between the decision-maker and other household members of the same person type (adults or children)

Figure T.8

Application of the Person Participation Model



4.3.3 Joint Tour Primary Destination Choice

Number of Models: 1
Decision-Making Unit: Tour
Model Form: Multinomial Logit
Alternatives: MGRAs

The joint tour primary destination choice model determines the location of the tour primary destination. The destination is chosen for the tour and assigned to all tour participants. The model works at an MGRA level, and sampling of destination alternatives is implemented in order to reduce computation time.

The explanatory variables for the joint tour primary destination choice model include:

- Household income
- Gender
- Age
- Maximum pair-wise overlaps between the decision-maker and other household members of the same person type (adults or children)
- Number of tours left over (including the current tour) to be scheduled
- Off-peak MGRA to MGRA distance
- The tour mode choice logsum for the person from the residence MGRA to each sampled MGRA location

- Non-mandatory HOV accessibility from household MGRA to employment (accessibility terms #7, 8, 9 (by auto sufficiency) (see section 1.2)
- The size of each sampled MGRA by tour purpose (see section 1.2)

4.3.4 Joint Tour Time of Day Choice

Number of Models: 1
 Decision-Making Unit: Persons
 Model Form: Multinomial Logit
 Alternatives: 861 (combinations of tour departure half-hour and arrival half-hour back at home)

After joint tours have been generated and assigned a primary location, the tour departure time from home and arrival time back at home is chosen simultaneously. The model is fully described under 4.2.2, above. However, a unique condition applies when applying the time-of-day choice model to joint tours. That is, the tour departure and arrival period combinations are restricted to only those available for each participant on the tour, after scheduling mandatory activities. Once the tour departure/arrival time combination is chosen, it is applied to all participants on the tour.

The model includes the following explanatory variables:

- Household income
- Person type
- Gender
- Age
- Mandatory tour frequency
- Auto travel distance
- Destination employment density
- Tour Departure time
- Tour Arrival time
- Tour duration
- The tour mode choice logsum by tour purpose from the residence MGRA to each sampled MGRA location

4.3.5 Joint Tour Mode Choice Model

Number of Models: 2 (Maintenance, Discretionary)
 Decision-Making Unit: Person
 Model Form: Nested Logit
 Alternatives: 23 (See Figure T.6 under the Individual Mandatory Tour Mode Choice Section)

Like the individual mandatory tour mode choice model, the joint tour model determines how the “main tour mode” (used to get from the origin to the primary destination and back) is determined.

The joint tour mode choices are (drive alone, and school bus is eliminated for this model):

- Shared-Ride 2 Free (General Purpose Lane)
- Shared-Ride 2 Free (HOV Lane)
- Shared-Ride 2 Pay

- Shared-Ride 3+ Free (General Purpose Lane)
- Shared-Ride 3+ Free (HOV Lane)
- Shared-Ride 3+ Pay
- Walk
- Bike
- Walk-Local Bus
- Walk-Express Bus
- Walk-Bus Rapid Transit
- Walk-Light Rail Transit
- Walk-Commuter Rail
- PNR-Local Bus
- PNR-Express Bus
- PNR-Bus Rapid Transit
- PNR-Light Rail Transit
- PNR-Commuter Rail
- KNR-Local Bus
- KNR-Express Bus
- KNR-Bus Rapid Transit
- KNR-Light Rail Transit
- KNR-Commuter Rail

The joint tour mode choice model contains the following explanatory variables:

- Auto sufficiency
- Household size
- Age
- Gender
- In-vehicle time (auto and transit)
- Walk and bike time
- Auto operating cost
- Auto parking cost
- Auto terminal time

- Auto toll value
- Transit first wait time
- Transit transfer time
- Number of transit transfers
- Transit walk access time
- Transit walk egress time
- Transit walk auxiliary time
- Transit fare
- Transit drive access time
- Transit drive access cost
- Intersection density
- Employment density
- Dwelling unit density

4.4.1 Individual Non-Mandatory Tour Frequency

Number of Models: 1

Decision-Making Unit: Households (at least one household member must have a DAP type of M or N)

Model Form: Multinomial Logit

Alternatives: Approximately 197 alternatives, composed of 0-1+ or 2+ tours of each type of maintenance activity (Escort, Shop, Other Maintenance, Eat Out, Visit, and Other Discretionary)

Allocated tours cover non-mandatory activities taken on by an individual on behalf of the household and include escort, shopping, other maintenance, eat out, visit, and other discretionary tours. They are generated by the household and later assigned to an individual in the household based on their residual time window. The choices include the number (0-2) and type of tours generated by each of the non-mandatory tour purposes. The explanatory variables include:

- Auto sufficiency
- Household income
- Dwelling type
- Number of full time workers in household
- Number of part time workers in household
- Number of university students in household
- Number of non-workers in household
- Number of retirees in household
- Number of driving age school children in household
- Number pre-driving age school children in household

- Number of preschool children in household
- Number of adults in household not staying home
- Number of children in household not staying home
- Gender
- Age
- Education level
- Indicator variable for whether person works at home regularly
- Number of individual/joint tours per person by tour purpose
- Population density at the origin
- Work Accessibility from household MGRA to employment (accessibility terms #45, see section 1.2)
- School Accessibility from household MGRA to employment (accessibility terms #45, see section 1.2)
- Escorting HOV Accessibility from household MGRA to employment (accessibility terms #25, 26, 27 (by auto sufficiency), see section 1.2)
- Shopping SOV/HOV Accessibility from household MGRA to employment (accessibility terms #10, 11, 12, 28, 29, 30 (by auto sufficiency), see section 1.2)
- Maintenance SOV/HOV Accessibility from household MGRA to employment (accessibility terms #13, 14, 15, 31, 32, 33 (by auto sufficiency), see section 1.2)
- Eating Out SOV/HOV Accessibility from household MGRA to employment (accessibility terms #16, 17, 18, 34, 35, 36 (by auto sufficiency), see section 1.2)
- Walk Accessibility from household MGRA to non-mandatory activities (accessibility terms #3 , see section 1.2)

4.4.2 Individual Non-Mandatory Tour Primary Destination Choice

Number of Models: 6 (Escort, Shop, Other Maintenance, Eat Out, Visit, and Other Discretionary)

Decision-Making Unit: Person

Model Form: Multinomial Logit

Alternatives: MGRAs

The six non-mandatory tour purposes are: (1) escorting, (2) shopping, (3) other maintenance, (4) eating out, (5) visiting, and (6) other discretionary. The non-mandatory tour primary destination choice model determines the location of the tour primary destination for each of the six non-mandatory tour purposes. The model works at an MGRA level, and sampling of destination alternatives is implemented in order to reduce computation time. Note that the mode choice logsum used is based on a ‘representative’ time period for individual non-mandatory tours, which is currently off-peak, since the actual time period is not chosen until model 4.4.3.

The explanatory variables in non-mandatory tour location choice models include:

- Household income
- Age of the traveler
- Gender

- Distance
- The tour mode choice logsum for the traveler from the residence MGRA to each sampled destination MGRA using off-peak level-of-service
- Time Pressure calculated as the log of the maximum time divided by number of tours left to be scheduled
- The size of each sampled MGRA

4.4.3 Individual Non-Mandatory Tour Time of Day Choice

Number of Models: 6 (Escort, Shop, Other Maintenance, Eat Out, Visit, and Other Discretionary)
 Decision-Making Unit: Person
 Model Form: Multinomial Logit
 Alternatives: 861 (combinations of tour departure half-hour and arrival half-hour back at home)

After individual non-mandatory tours have been generated, allocated, and assigned a primary location, the tour departure time from home and arrival time back at home is chosen simultaneously. The model is fully described under 4.1.2, above. The tour departure and arrival period combinations are restricted to only those available for each participant on the tour, after scheduling individual mandatory tours and joint tours.

The model includes the following explanatory variables:

- Household Income
- Person type
- Gender
- Age
- Mandatory tour frequency
- Joint tour indicator
- Auto travel distance
- Tour Departure time
- Tour Arrival time
- Tour duration
- Time Pressure calculated as the log of the maximum time divided by number of tours left to be scheduled
- The tour mode choice logsum by tour purpose from the residence MGRA to each sampled MGRA location

4.4.4 Individual Non-Mandatory Tour Mode Choice Model

Number of Models: 2 (Maintenance, Discretionary)
 Decision-Making Unit: Person
 Model Form: Nested Logit
 Alternatives: 25 (See Figure T.6 under the Individual Mandatory Tour Mode Choice Section)

Like the individual mandatory tour mode choice model, the individual non-mandatory tour model determines how the “main tour mode” (used to get from the origin to the primary destination and back) is determined.

The individual non-mandatory tour mode choices are (school bus is eliminated):

- Drive-alone Free
- Drive-Alone Pay
- Shared-Ride 2 Free (General Purpose Lane)
- Shared-Ride 2 Free (HOV Lane)
- Shared-Ride 2 Pay
- Shared-Ride 3+ Free (General Purpose Lane)
- Shared-Ride 3+ Free (HOV Lane)
- Shared-Ride 3+ Pay
- Walk
- Bike
- Walk-Local Bus
- Walk-Express Bus
- Walk-Bus Rapid Transit
- Walk-Light Rail Transit
- Walk-Commuter Rail
- PNR-Local Bus
- PNR-Express Bus
- PNR-Bus Rapid Transit
- PNR-Light Rail Transit
- PNR-Commuter Rail
- KNR-Local Bus
- KNR-Express Bus
- KNR-Bus Rapid Transit
- KNR-Light Rail Transit
- KNR-Commuter Rail

The individual non-mandatory tour mode choice model contains the following explanatory variables:

- Auto sufficiency
- Household size
- Age
- Gender
- In-vehicle time (auto and transit)
- Walk and bike time
- Auto operating cost
- Auto parking cost
- Auto terminal time
- Auto toll value
- Transit first wait time
- Transit transfer time
- Number of transit transfers
- Transit walk access time
- Transit walk egress time
- Transit walk auxiliary time
- Transit fare
- Transit drive access time
- Transit drive access cost
- Intersection density
- Employment density
- Dwelling unit density

4.5.1 At-Work Sub-Tour Frequency

Number of Models: 1

Decision-Making Unit: Persons

Model Form: Multinomial Logit

Alternatives: 6 (None, 1 eating out tour, 1 work tour, 1 other tour, 2 work tours, 2 other tours, and a combination of eating out, work, and other tours)

At-work based sub-tours are modeled last, and are relevant only for those persons who implement at least one work tour. These underlying activities are mostly individual (e.g., business-related and dining-out purposes), but may include some household maintenance functions as well as person and household maintenance tasks. There are seven alternatives in the model, corresponding to the most frequently observed patterns of at-work sub-tours. The alternatives define both the number of at-work sub-tours and their purpose.

The at-work sub tour frequency model includes the following explanatory variables:

- Household income
- Number of driving age adults
- Number of preschool children
- Person type
- Gender
- Number of individual and joint mandatory and non-mandatory tours generated in the day
- Employment density at the work place
- Mixed use category at the work place
- Non-motorized eating out accessibility from work MGRA to destination MGRA(accessibility terms #46, see section 1.2)

4.5.2 At-Work Sub-Tour Primary Destination Choice

Number of Models: 1
Decision-Making Unit: Person
Model Form: Multinomial Logit
Alternatives: MGRAs

The at-work sub-tour primary destination choice model determines the location of the tour primary destination. The model works at an MGRA level, and sampling of destination alternatives is implemented in order to reduce computation time. Note that the mode choice logsum used is based on a 'representative' time period for individual non-mandatory tours, which is currently off-peak, since the actual time period is not chosen until model 3.6.3. The model is constrained such that only destinations within a reasonable time horizon from the workplace are chosen, such that the tour can be completed within the total available time window for the sub-tour.

The explanatory variables in the at-work sub tour choice models include:

- Person type
- Distance
- The tour mode choice logsum for the traveler from the residence MGRA to each sampled destination MGRA using off-peak level-of-service
- The size of each sampled MGRA

4.5.3 At-Work Sub-Tour Time of Day Choice

Number of Models: 1
Decision-Making Unit: Person
Model Form: Multinomial Logit
Alternatives: 861 (combinations of tour departure half-hour and arrival half-hour back at home, with aggregation of time between 1 AM and 5 AM)

After at-work sub-tours have been generated and assigned a primary location, the tour departure time from workplace and arrival time back at the workplace is chosen simultaneously. The model is fully described under 3.1.2, above. The tour departure and arrival period combinations are restricted to only those available based on the time window of the parent work tour.

The model includes the following explanatory variables:

- Household Income
- Sub-tour purpose
- Auto travel distance
- Tour Departure time
- Tour Arrival time
- Tour duration
- Maximum Available Continuous Time Window (in hours) between 5 a.m. to 11 p.m. before this tour is scheduled
- The tour mode choice logsum from the work MGRA to each sampled MGRA location

4.5.4 At-Work Sub-Tour Mode Choice Model

Number of Models: 1

Decision-Making Unit: Person

Model Form: Nested Logit

Alternatives: 25 (See Figure T.6 under the Individual Mandatory Tour Mode Choice Section)

Like the individual mandatory tour mode choice model, the at-work sub-tour model determines the “main tour mode” used to get from the origin to the primary destination and back is determined.

The at-work sub-tour mode choices are (school bus is eliminated):

- Drive-alone Free
- Drive-Alone Pay
- Shared-Ride 2 Free (General Purpose Lane)
- Shared-Ride 2 Free (HOV Lane)
- Shared-Ride 2 Pay
- Shared-Ride 3+ Free (General Purpose Lane)
- Shared-Ride 3+ Free (HOV Lane)
- Shared-Ride 3+ Pay
- Walk
- Bike
- Walk-Local Bus
- Walk-Express Bus
- Walk-Bus Rapid Transit
- Walk-Light Rail Transit
- Walk-Commuter Rail

- PNR-Local Bus
- PNR-Express Bus
- PNR-Bus Rapid Transit
- PNR-Light Rail Transit
- PNR-Commuter Rail
- KNR-Local Bus
- KNR-Express Bus
- KNR-Bus Rapid Transit
- KNR-Light Rail Transit
- KNR-Commuter Rail

The at work sub-tour mode choice model contains the following explanatory variables

- Auto sufficiency
- Household size
- Age
- Gender
- In-vehicle time (auto and transit)
- Walk and bike time
- Auto operating cost
- Auto parking cost
- Auto terminal time
- Auto toll value
- Transit first wait time
- Transit transfer time
- Number of transit transfers
- Transit walk access time
- Transit walk egress time
- Transit walk auxiliary time
- Transit fare
- Transit drive access time
- Transit drive access cost

- Intersection density
- Employment density
- Dwelling unit density

5.1 Intermediate Stop Frequency Model

Number of Models: 9 (By purpose plus one model for at-work sub-tours)

Decision-Making Unit: Person

Model Form: Multinomial Logit

Alternatives: 16, with a maximum of 3 stops per tour direction, 6 total stops on tour

The stop frequency choice model determines the number of intermediate stops on the way to and from the primary destination. The SANDAG model allowed more than one stop in each direction (up to a maximum of three) for a total of eight trips per tour (four on each tour leg). An additional constraint placed on this model was that no stops were allowed on drive-transit tours. This was enforced to ensure that drivers who drive to transit picked up their cars at the end of the tour.

The stop frequency model was based on the following explanatory variables:

- Household income
- Number of full time workers in the household
- Number of part time workers in the household
- Number of non-workers in the household
- Number of children in the household
- Number of individual/joint mandatory and non-mandatory tours made by household
- Person type
- Age
- Tour mode
- Tour distance from anchor location (home) to primary destination
- Maintenance accessibility (#s31, 32, 33)
- Discretionary accessibility (#s40, 41,42)

The observed stop frequency distributions are given in Table T.12 and Table T.13.

Table T.12

Observed Stop Frequency Distribution by Tour Purpose

Inbound Stops	Outbound Stops	Primary Tour Purpose									Total
		Work	University	School	Escorting	Shopping	Maintenance	Eating Out	Visiting	Discretionary	
0	0	2,061	174	1,162	1,038	726	869	336	318	968	7,652
1	0	511	22	203	151	141	165	26	35	118	1,372
2	0	124	6	40	36	34	66	2	3	26	337
3	0	85	2	31	18	18	27	0	11	10	202
0	1	294	23	84	84	216	95	17	41	64	918
1	1	154	8	73	25	62	43	11	19	30	425
2	1	47	3	19	5	20	13	0	5	15	127
3	1	32	1	7	2	14	10	0	0	2	68
0	2	78	5	6	23	56	22	3	12	15	220
1	2	27	0	7	11	22	18	4	1	9	99
2	2	22	0	3	0	14	5	2	1	8	55
3	2	15	0	0	2	2	2	0	1	2	24
0	3	17	0	3	10	20	11	3	7	6	77
1	3	13	1	0	6	11	8	0	2	2	43
2	3	9	0	0	1	1	5	0	0	7	23
3	3	5	0	0	1	5	1	0	0	1	13
Total		3,494	245	1,638	1,413	1,362	1,360	404	456	1,283	11,655

**Table T.13
Observed Stop Frequency Distribution by
At-work Sub Tour Purpose**

Inbound Stops	Outbound Stops	At-Work Sub Tour Purpose
0	0	441
1	0	25
2	0	2
3	0	1
0	1	26
1	1	8
2	1	2
3	1	0
0	2	1
1	2	0
2	2	0
3	2	0
0	3	1
1	3	0
2	3	0
3	3	0
Total		507

5.2 Intermediate Stop Purpose Choice Model

Number of Models: 1
 Decision-Making Unit: Stop
 Model Form: Lookup Table
 Alternatives: 9 Stop Purposes (Work, University, School, Escort, Shop, Maintenance, Eating Out, Visiting, or Discretionary)

The stop purpose choice model will be a lookup table of probabilities based upon tour purpose, stop direction, departure time, and person type.

5.3 Intermediate Stop Location Choice Model

Number of Models: 1
 Decision-Making Unit: Person
 Model Form: Multinomial Logit
 Alternatives: MGRA

The stop location choice model predicts the location (the Master Geographic Reference Area, or MGRA) of each intermediate stop (each location other than the origin and primary destination) on the tour. In this model, a maximum of three stops in outbound and three stops in inbound direction are modeled for each tour. Since there are a large

number (over 33,000) of alternative destinations it is not possible to include all alternatives in the estimation dataset. A sampling-by-importance approach was used to choose a set of alternatives. Each record was duplicated 20 times, then different choice sets with 30 alternatives each were selected based on the size term and distance of the alternative destination. This approach is statistically equivalent to selecting 600 alternatives for the choice set. It is not straightforward to segment the model by purpose because size (or attraction) variables are related to purpose of the stop activity while impedance variables are strongly related to the tour characteristics – primary tour purpose, primary mode used for the tour, etc. Therefore, a single model is estimated with size variables based on stop purpose and utility variables based on both stop and tour characteristics.

The stop location choice model includes the following explanatory variables:

1. Household Income
2. Gender
3. Age
4. Mode choice logsum
5. Distance deviation or “out-of-the-way” distance for stop location when compared to the half-tour distance without detour for any stop
6. Distance of stop location from tour origin and destination is used to define closeness to tour origin or destination.
7. Stop purpose
8. Tour purpose
9. Tour mode
10. Stop Number
11. Direction of the half-tour
 - Size variables:
 - a. Employment by categories
 - b. Number of households
 - c. School enrollments – pre-school, K to 6th grade, and 7th to 12th grade, based on type of school child in the household
 - d. University and other college enrollments

5.4 Intermediate Stop Departure Model

Number of Models: 1
 Decision-Making Unit: Trips other than first trip and last trip on tour
 Model Form: Lookup Table
 Alternatives: 40 (stop departure half-hour time periods, with aggregation between 1 AM and 5 AM)

The stop departure model will be a lookup table of probabilities based upon tour purpose, stop direction, tour departure time, and stop number.

6.1 Trip Mode Choice Model

Number of Models: 6 (Work, University, K-12, Maintenance, Discretionary, and At-work subtours)
 Decision-Making Unit: Person
 Model Form: Multinomial Logit
 Alternatives: 26 (See Figure T.6)

The trip mode choice model determines the mode for each trip along the tour. Trip modes are constrained by the main tour mode. The linkage between tour and trip levels is implemented through correspondence rules (which trip

modes are allowed for which tour modes). The model can incorporate asymmetric mode combinations, but in reality, there is a great deal of symmetry between outbound and inbound modes used for the same tour. In particular, symmetry is enforced for drive-transit tours, by excluding intermediate stops from drive-transit tours.

The tour and trip mode correspondence rules are shown in Table T.14. Note that in the SANDAG trip mode choice model, the trip modes are exactly the same as the modes in the tour mode choice model. However, every trip mode is not necessarily available for every tour mode. The correspondence rules depend on a hierarchy with the following rules:

- The highest occupancy across all trips is used to code the occupancy of the tour.
- There are no mode switching on walk and bike tour modes.
- Shared-ride trips are allowed on walk-transit tours.
- Drive-alone is disallowed for walk-transit and KNR-transit tours, since driving on a trip leg in combination with walk-transit would imply PNR-transit as a tour mode.
- Walk trips are allowed on all tour modes with the exception of driving alone and biking, since these modes imply that the traveler is attached to the mode of transport (the auto or bike) for the entire tour.
- Note that cases in which a traveler parks at a lot and then walks to their destination are treated as a single trip in the context of trip mode choice. A subsequent parking location choice model will break out these trips into the auto leg and the walk leg, for trips to parking-constrained locations.
- An additional restriction on availability is imposed on work-based sub-tours, where drive-alone is disallowed if the mode to work is not one of the three auto modes (drive-alone, shared 2, or shared 3+).

The school bus tour mode, which is only available for the School tour purpose, implies symmetry – all trips on school bus tours must be made by school bus.

The trip mode choice model's explanatory variables include:

- Household Size
- Auto sufficiency
- Age
- Gender
- Tour mode
- Individual or joint tour indicator
- Number of outbound and return stops
- First and last stop indicators
- In-vehicle time (auto and transit)
- Walk and bike time
- Auto operating cost
- Auto parking cost

- Auto terminal time
- Auto toll value
- Transit first wait time
- Transit transfer time
- Number of transit transfers
- Transit walk access time
- Transit walk egress time
- Transit walk auxiliary time
- Transit fare
- Transit drive access time
- Transit drive access cost
- Intersection density
- Employment density
- Dwelling unit density

Table T.14
Tour and Trip Mode Correspondence Rules

Trip Mode	Tour Mode							
	Drive-Alone	Shared 2	Shared 3+	Walk	Bike	Walk-Transit	PNR-Transit	KNR-Transit
Drive-alone Free	A	A	A				A	
Drive-Alone Pay	A	A	A				A	
Shared-Ride 2 Free (GP Lane)		A	A			A	A	A
Shared-Ride 2 Free (HOV Lane)		A	A			A	A	A
Shared-Ride 2 Pay		A	A			A	A	A
Shared-Ride 3+ Free (GP Lane)			A			A	A	A
Shared-Ride 3+ Free (HOV Lane)			A			A	A	A
Shared-Ride 3+ Pay			A			A	A	A
Walk		A	A	A		A	A	A
Bike					A			
Walk-Local Bus						A	A	A
Walk-Express Bus						A	A	A
Walk-Bus Rapid Transit						A	A	A
Walk-Light Rail Transit						A	A	A
Walk-Commuter Rail						A	A	A
PNR-Local Bus							A	
PNR-Express Bus							A	
PNR-Bus Rapid Transit							A	
PNR-Light Rail Transit							A	
PNR-Commuter Rail							A	
KNR-Local Bus								A
KNR-Express Bus								A
KNR-Bus Rapid Transit								A
KNR-Light Rail Transit								A
KNR-Commuter Rail								A

School Bus Available for school bus tour mode only, on school tours.

A = Trip mode is available by that particular tour mode.

6.2 Parking Location Choice

Number of Models:	2 (Work, and other)
Decision-Making Unit:	Trips with non-home destinations in areas with paid parking
Model Form:	Multinomial Logit
Alternatives:	<i>In estimation</i> , lots sampled in the parking behavior survey <i>In application</i> , MGRAs within 3/4 mile of the destination MGRA

The parking location choice model determines where vehicles are parked at the terminal end of each trip with a destination in *park area 1* (downtown San Diego area). For work trips, the model will subtract the output from the employer parking reimbursement model from the daily price of parking at each alternative destination to determine the effective price borne by the individual. The output of the model will be used to obtain traffic assignments that are more accurate at small scales in the downtown area during the morning and afternoon peaks. The coefficients from the parking location choice model estimation are also used in defining the logsum-weighted average parking cost used in mode choice.

The parking location model explanatory variables include:

- Number of stalls available to the driver (size variable)
- Parking cost
- Walk distance to destination

Special Market Models

Cross border model

In 2010, SANDAG collected data on Mexican resident border crossings into the United States and their travel patterns within the United States. Data was collected at the three border crossing stations – San Ysidro, Otay Mesa, and Tecate. Based upon this data, PB is developing a travel demand model for Mexican residents. The purpose of this model is two-fold. The primary purpose of the model is to measure the impact of Mexican resident travel on the San Diego transport network. The model will account for Mexican resident demand (such as auto volume, transit boarding, and toll revenue) for transportation infrastructure in San Diego County. The other purpose of the model is to forecast border crossings at each current and potential future border crossing station.

The model flow and inputs are shown in Figure T.10, and described in detail in the following sections.

Cross Border Modeled Travel Dimensions

Cross Border Tour Purposes

There are six tour purposes for the Mexican resident model. They were coded based on the activity purposes engaged in by the traveler in the United States, according to a hierarchy of activity purposes as follows:

- **Work:** At least one trip on the tour is for working in the United States.
- **School:** At least one trip on the tour is made for attending school in the United States, and no work trips were made on the tour.
- **Cargo:** At least one trip on the tour was made for picking up or dropping off cargo in the United States, and no work or school trips were made on the tour.
- **Shop:** No trips on the tour were made for work, school, or cargo, and the activity with the longest duration on the tour was shopping in the United States.

- **Visit:** No trips on the tour were made for work, school, or cargo, and the activity with the longest duration on the tour was visiting friends/relatives in the United States.
- **Other:** No trips on the tour were made for work, school, or cargo, and the activity with the longest duration on the tour was other (collapsed escort, eat, personal, medical, recreation, sport, and other activity purposes).

Tour Mode

The tour mode is the mode used to cross the border, which conditions the mode used for all trips on the tour, including the trip from the border crossing to the first destination in the United States. The tour modes are defined by whether the border was crossed via auto or by foot, the occupancy if by auto, and also whether the SENTRI lane was used or not. SENTRI lanes offer expedited border crossings to pre-qualified citizens of the United States and Mexico. One must apply for a SENTRI pass, which requires extensive background checks. Mexican residents must have a valid United States Visa, Mexican passport, and contact number in the United States. This typically means that in order to obtain a pass, Mexican residents must be lawfully employed in the United States.

Trip Mode

The trip modes used in the Mexican resident travel model are the same modes available in the resident activity-based model. Note that toll and HOV usage was not asked as part of the survey. Usage of these facilities in the model will be based upon the characteristics of the trips/vehicle occupancies and income (value-of-time) of travelers, and validated along with resident demand models.

Treatment of Time

Every trip will be allocated to a half-hourly period consistent with the resident travel demand model treatment of time. Travel skims will also be consistent with resident travel demand models; currently there are three sets of skims, for the AM peak, PM peak, and off-peak periods. Eventually the off-peak period will be represented as four time periods (early AM, early midday, late midday, and evening), for a total of six sets of skims.

Treatment of Space

Every trip ending in San Diego County will be allocated to an MGRA. Within Tijuana, each border crossing origin will be assigned to a *colonia*, or neighborhood with which survey respondents identify. Population estimates are collected by the Instituto Nacional de Estadística y Geografía (INEGI) at the level of a basic geostatistical area (Area Geostadística Básica, or AGEB, roughly equivalent to U.S. Census Tracts). AGEBs and *colonia* largely overlap within Tijuana city boundaries (though there is no coherent spatial nesting scheme), and AGEB population estimates were redistributed to *colonia* based on a proportional area operation to operationalize *colonia* trip origins in the model. Outside of Tijuana, the origins will be distributed to a *localidad*, or locality. These units are similar to the Census Designated Place in the United States.

Figure T.9

Total Population of Áreas Geoestadísticas Básicas (AGEB) in Tijuana

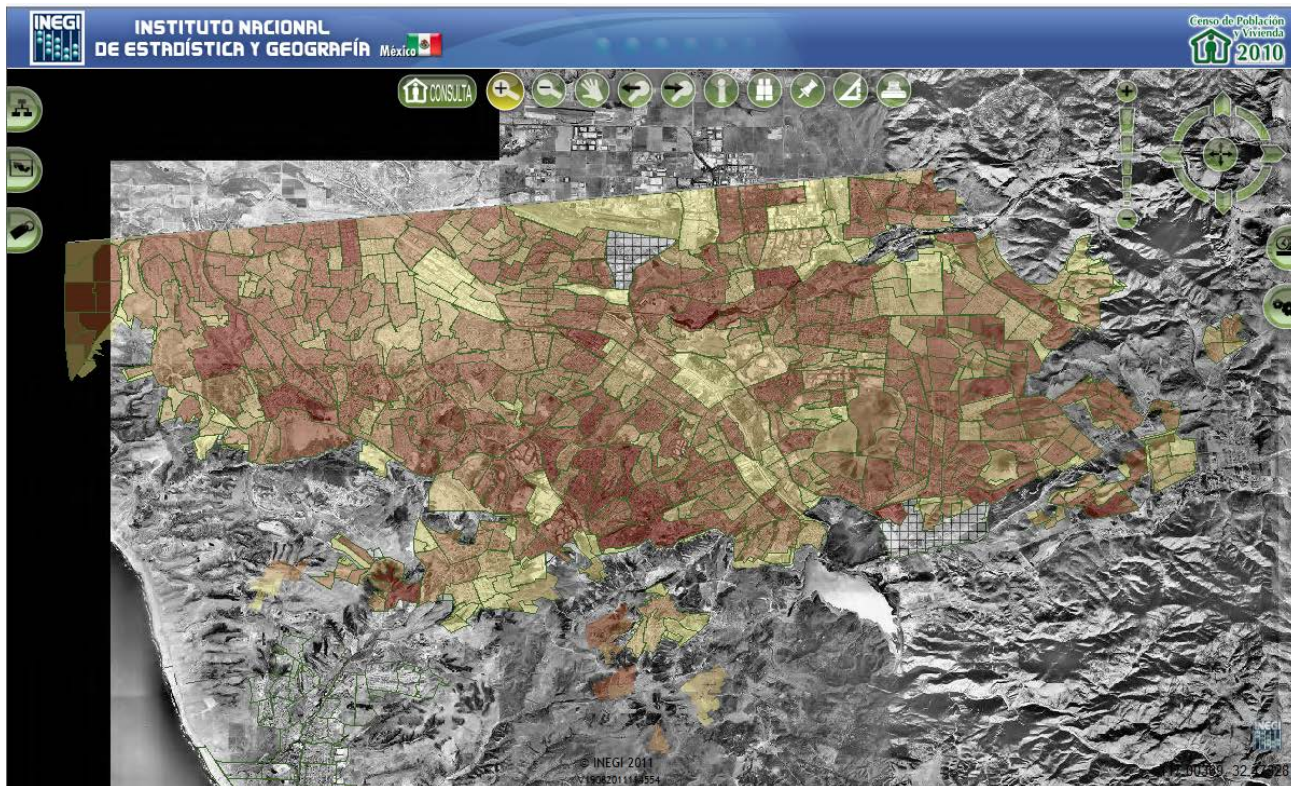
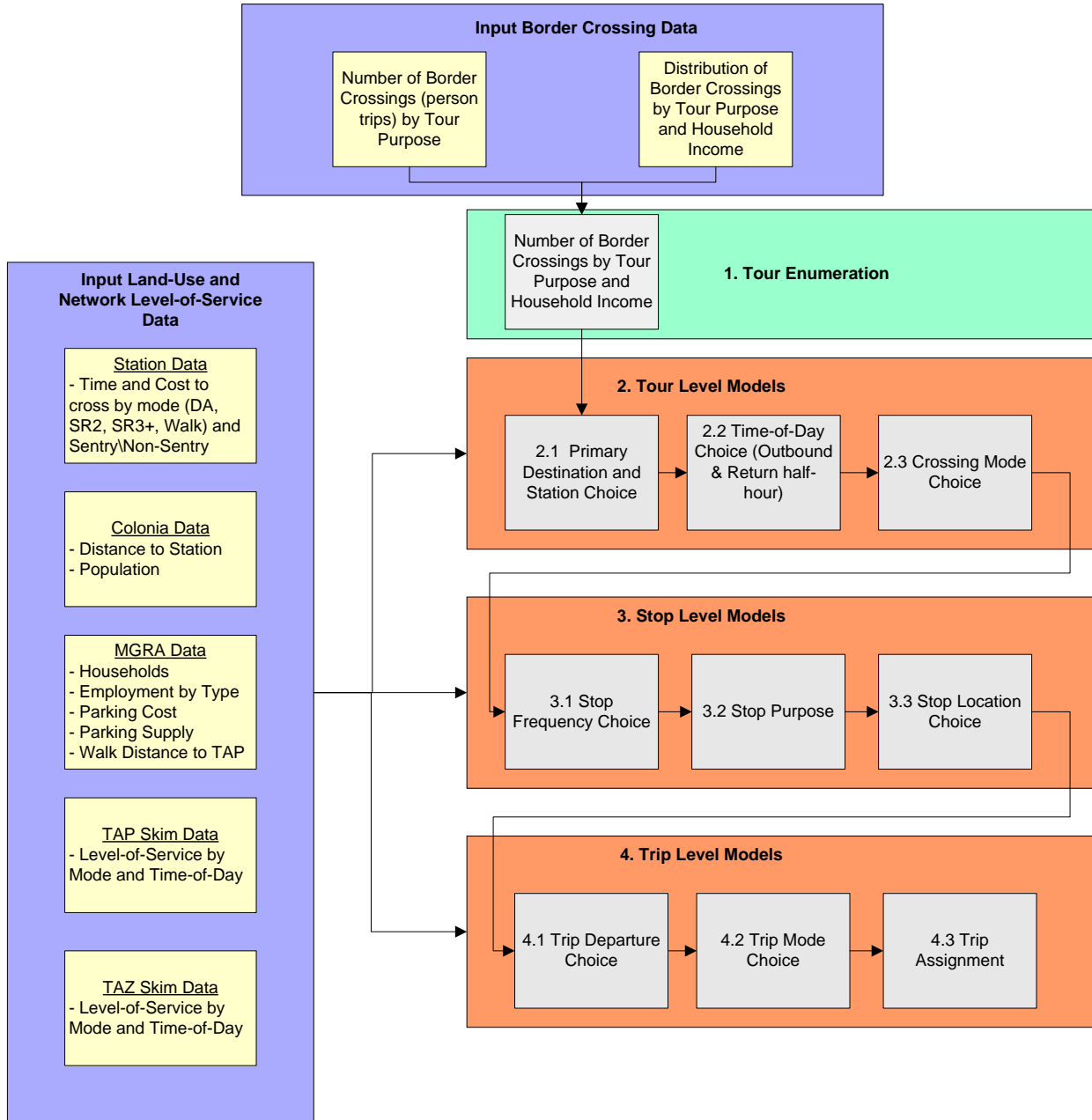


Figure T.10

Mexican Resident Travel Model



2.1 Border Crossing Primary Destination and Station Crossing Choice

Number of Models:	6 (Work, School, Cargo, Shop, Visit, Other)
Decision-Making Unit:	Person-tour
Model Form:	Multinomial Logit
Alternatives:	MGRAs and border crossing stations

The primary destination and border crossing choice model is a joint choice model of tour primary destination in the United States and border crossing station. Due to the number of alternatives in the model, a sampling will be used to select a sub-set (50) of primary destination MGRAs and border crossing pairs. The sampling procedure will be based upon a simplified destination choice model that takes into account:

- The weighted distance from the TAZ to all border crossing stations
- The time and cost that it takes to cross each station
- The accessibility of the border crossing station to persons in Mexico

The explanatory variables in the primary destination and station crossing choice model are:

- Distance
- Station index
- The tour mode choice logsum for the traveler from the residence MGRA to each sampled destination MGRA using a representative departure and arrival periods for each purpose
- Station accessibility to population
- The size of each sampled MGRA

2.2 Border Crossing Tour Time-of-Day Choice

Number of Models:	1-6 (Work, School, Cargo, Shop, Visit, Other – depending on significance of purpose-specific parameters)
Decision-Making Unit:	Person-tour
Model Form:	Lookup table
Alternatives:	820 (combinations of border entry half-hour and return half-hour, with aggregation between 1 AM and 5 AM)

The border crossing tour time of day choice model will be a lookup table of probabilities based upon tour purpose, tour entry and tour return.

2.3 Border Crossing Tour Mode Choice

Number of Models:	1-6 (Work, School, Cargo, Shop, Visit, Other – depending on significance of purpose-specific parameters)
Decision-Making Unit:	Person-tour
Model Form:	Multinomial or Nested Logit
Alternatives:	Tour Mode (4: Drive Alone, Shared 2 person, Shared 3+ person, Walk)

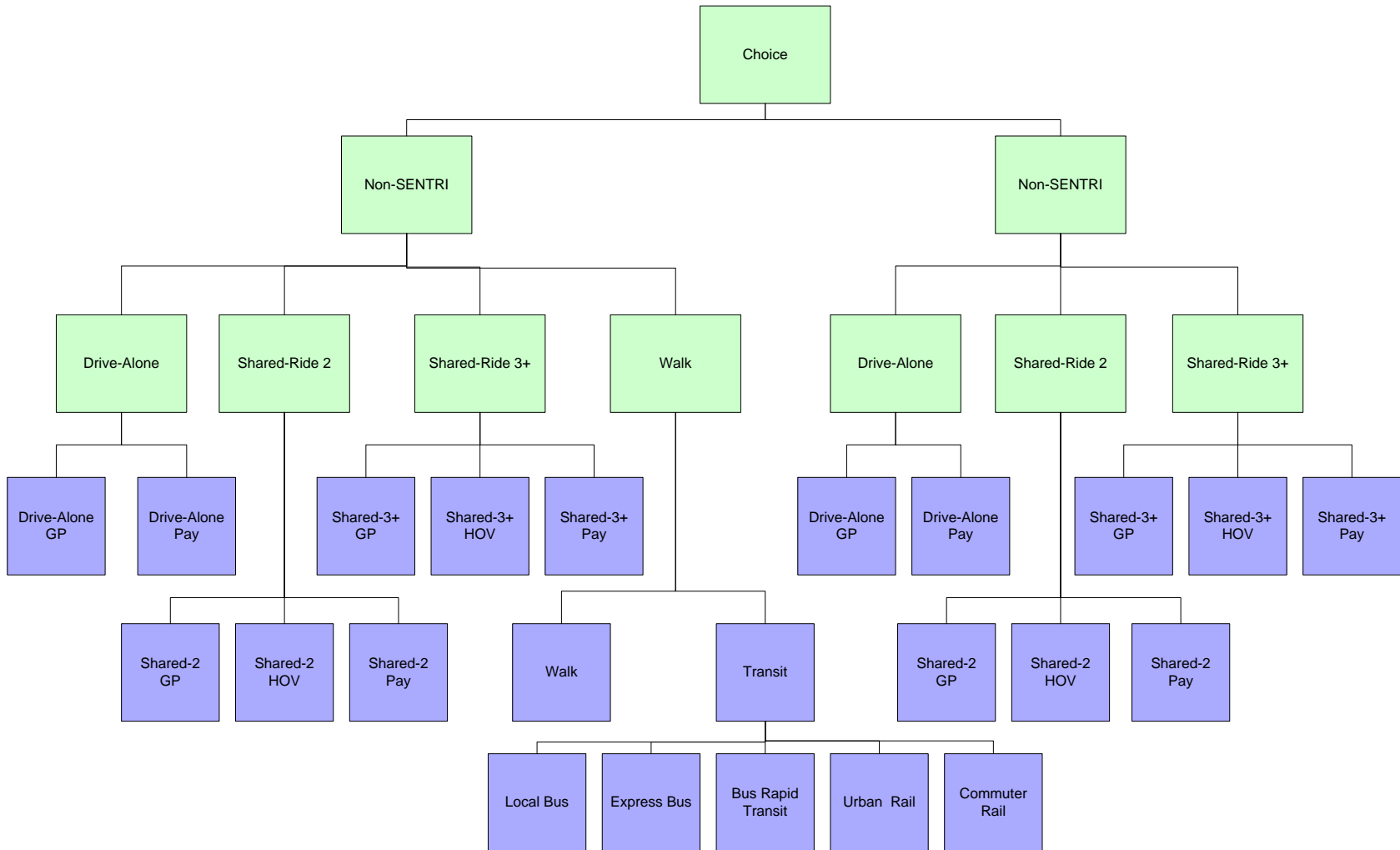
This model chooses tour mode based on a known tour destination, border crossing station, and entry/return time-of-day. The mode choice will be based upon a simplified trip mode choice logsum representing the accessibility of relevant modes for each border crossing mode as well as the time and cost of crossing the station by border crossing mode.

Figure T.11 shows a potential nesting structure of the border crossing mode choice model. In this figure, trip modes are shown in blue and border crossing modes are shown in green. The choice of mode will be made from one of the green (border crossing mode) alternatives, but the accessibility between the border crossing station and the tour primary destination by the trip modes shown in blue will influence the choice.

The border crossing tour mode choice model includes the following explanatory variables:

- Traveler has a SENTRI pass indicator
- Trip mode choice logsum by mode and stop direction
- Border wait time for automobiles
- Border wait time for travelers with a SENTRI pass
- Border wait time for pedestrians

Figure T.11
Potential Border Crossing Mode Choice Nesting Structure



3.1 Border Crossing Stop Frequency Choice

Number of Models:	1-6 (Work, School, Cargo, Shop, Visit, Other – depending on significance of purpose-specific parameters)
Decision-Making Unit:	Person-tour
Model Form:	Lookup table
Alternatives:	16 (0, 1, 2, 3 stops per half-tour)

The stop frequency choice model will be a lookup table of probabilities based upon tour purpose and duration.

3.2 Border Crossing Stop Purpose Choice

Number of Models:	1-6 (Work, School, Cargo, Shop, Visit, Other – depending on significance of purpose-specific parameters)
Decision-Making Unit:	Stop
Model Form:	Lookup table
Alternatives:	6 (Work, School, Cargo, Shop, Visit, Other)

The stop purpose choice model will be a lookup table of probabilities based upon tour purpose and number of stops on tour. The purpose-segmentation will be based on the tour purpose, if implemented.

3.3 Border Crossing Stop Location Choice

Number of Models:	1 (with stop-purpose-specific size terms)
Decision-Making Unit:	Stop
Model Form:	Multinomial Logit
Alternatives:	MGRAs

The stop location choice model predicts the location of stops along the tour other than the primary destination. The stop-location model is structured as a multinomial logit model using MGRA attraction size variable and route deviation measure as impedance. The alternatives are sampled from the full set of MGRAs, based upon the out-of-direction distance to the stop and the size of the MGRA. The sampling mechanism is also subject to certain rules based on tour mode. All destinations are available for auto tour modes, so long as there is a positive size term for the MGRA. Intermediate stops on walk tours must be within 3 miles of both the tour origin and primary destination MGRAs. The sampling for intermediate stops on walk-transit tours is based upon the MGRAs that are within walking distance of the boarding or alighting stops at the tour origin and primary destination.

The intermediate stop location choice model works by cycling through stops on tours. The level-of-service variables (including mode choice logsums) are calculated as the additional utility between the last location and the next known location on the tour. For example, the LOS variable for the first stop on the outbound direction of the tour is based on additional impedance between the tour origin and the tour primary destination. The LOS variable for the next outbound stop is based on the additional impedance between the previous stop and the tour primary destination. Stops on return tour legs work similarly, except that the location of the first stop is a function of the additional impedance between the tour primary destination and the tour origin. The next stop location is based on the additional impedance between the first stop on the return leg and the tour origin, and so on.

The cross border stop location choice model explanatory variables included were:

- Mode choice logsum
- Tour duration
- Distance of stop location from tour origin and destination is used to define closeness to tour origin or destination
- Stop Number
- Direction of the half-tour

- Stop frequency
- The size of each sampled stop location MGRA

4.1 Border Crossing Trip Departure Choice

Number of Models: 1
 Decision-Making Unit: Trips other than first trip and last trip on tour
 Model Form: Lookup from Probabilities
 Alternatives: Number of half-hour time after outbound period for outbound trips/number of half-hour time periods before return period for return trips

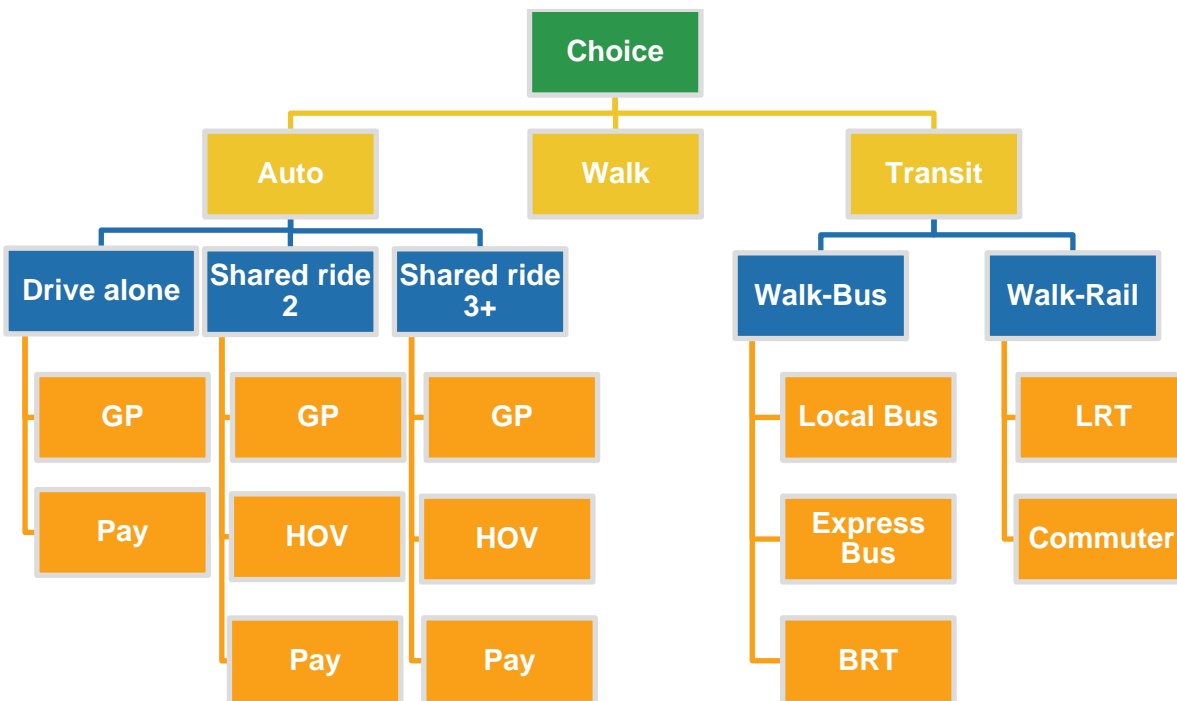
Each trip will be assigned to a trip departure time period. The first and last trips of the tour are set to the entry/return time periods from Model 2.2, respectively. Each intermediate trip departure time is calculated from a lookup table of probabilities that consider the number of remaining half-hour periods in the tour from the last scheduled trip and whether the stop is made on the outbound or return direction.

4.2 Border Crossing Trip Mode Choice

Number of Models: 1
 Decision-Making Unit: Trip
 Model Form: Nested Logit
 Alternatives: 13 Trip Modes

A trip mode is chosen for each trip on the tour. Trip modes are consistent with the resident travel model, as shown in Figure T.12, though certain modes (bike, drive-transit, and school bus) are unavailable for Mexican residents. The utility of each mode is a function of the time and cost of the mode for the period that the trip occurs in, and is influenced by the mode used to cross the border.

Figure T.12
Mexican Resident Trip Mode Choice Model for Travel in US



The border crossing trip mode choice model includes the following explanatory variables:

- Tour mode
- Tour purpose
- Traveler has a SENTRI pass indicator
- In-vehicle time (auto and transit)
- Walk and bike time
- Auto operating cost
- Auto parking cost
- Auto terminal time
- Auto toll value
- Transit first wait time
- Transit transfer time
- Number of transit transfers
- Transit walk access time
- Transit walk egress time
- Transit walk auxiliary time
- Transit fare

Airport ground access model

In 2008, San Diego International Airport (SDIA) conducted a survey of airport passengers in which data was collected on their travel to the airport prior to their departure. Based upon this data, PB is developing a model of travel to and from the airport for arriving and departing passengers. The purpose of this model is to capture the demand of airport travel on transport facilities in San Diego County. Additionally, the model will allow SANDAG to test the impacts of various parking price and supply scenarios at the airport.

The airport ground access model has the following features:

- A disaggregate micro-simulation treatment of air passengers, with explicit representation of duration of stay or trip in order to accurately represent costs associated with various parking and modal options.
- The full set of modes within San Diego County, including auto trips by occupancy, transit trips by line-haul mode (bus versus trolley), and toll/HOT/HOV lanes modes.
- Forecasts of airport ground access travel based upon the official SDIA enplanement projections.

The model flow and inputs are shown in Figure T.13, and described in detail in the following sections.

Airport Model Travel Dimensions

Airport Model Trip Purposes

There are four trip purposes and these were coded based on the resident status of air passengers and the purpose of air travel, as follows:

- **Resident Business:** Business travel made by San Diego County residents (or residents of neighboring counties who depart from SDIA).
- **Resident Personal:** Personal travel made by San Diego County residents (or residents of neighboring counties who depart from SDIA).
- **Visitor Business:** Business travel made by visitors to San Diego County (or a neighboring county).
- **Visitor Personal:** Personal travel made by visitors to San Diego County (or a neighboring county).

Airport Model Trip Mode

The model of airport ground access will be trip-based, since the survey did not collect the full tour from origin to airport. In addition, the survey only collected information on the trip to the airport before the passenger boarded their plane; information was not collected on the trip in which passengers arrived at the airport and traveled to a destination in San Diego County. Therefore, symmetry will be assumed for the non-reported trip. Finally, the survey did not collect data on whether an HOV lane or toll lane was used for the trip, so path-level mode cannot be determined. The number of survey respondents by trip mode is listed in Table T.15, and shares of respondents by trip mode are shown in Table T.16. Note that if private auto is used to access the airport, the choice of parking versus curbside pickup/dropoff is explicitly represented. For travelers that park, the chosen lot (terminal, airport remote lot, private remote lot) is explicit as well. Also note that auto occupancy is not a choice for airport ground access trips. Auto occupancy will be based upon travel party size, which will be simulated as part of the attribution of ground access trips.

Table T.15
Survey Respondents by Purpose and Trip Mode

Mode	Purpose*					Total
	Res-Bus	Res-Pers	Vis-Bus	Vis-Pers	External	
Unknown/Walk	1	7	7	5	2	22
Park-terminal	163	133	14	66	39	415
Park-SAN Offsite	61	74	3	4	12	154
Park-PVT Offsite	174	253	6	7	33	473
PUDO-Escort	39	100	18	53	27	237
PUDO-Curbside	627	1,538	310	1,096	205	3,776
Rental Car	9	9	695	761	123	1,597
Taxi/Limo/Towncar	155	227	520	288	15	1,205
Shuttle	66	143	301	259	21	790
Transit	13	32	10	38	8	101
Total	1,308	2,516	1,884	2,577	485	8,770

Table T.16

Frequency of Survey Respondents by Purpose and Trip Mode

Mode	Purpose					Total
	Res-Bus	Res-Pers	Vis-Bus	Vis-Pers	External	
Park-terminal	12%	5%	1%	3%	8%	5%
Park-SAN Offsite	5%	3%	0%	0%	2%	2%
Park-PVT Offsite	13%	10%	0%	0%	7%	5%
PUDO-Escort	3%	4%	1%	2%	6%	3%
PUDO-Curbside	48%	61%	16%	43%	42%	43%
Rental Car	1%	0%	37%	30%	25%	18%
Taxi/Limo/Towncar	12%	9%	28%	11%	3%	14%
Shuttle	5%	6%	16%	10%	4%	9%
Transit	1%	1%	1%	1%	2%	1%
Total	100%	100%	100%	100%	100%	100%

Airport Model Treatment of Time

Every trip will be allocated to a half-hour period consistent with the resident travel demand model treatment of time. Travel skims will also be consistent with resident travel demand models; currently there are three sets of skims, one each for the AM peak, PM peak, and off-peak periods. Eventually the off-peak period will be represented as four time periods (early AM, early midday, late midday, and evening), for a total of six sets of skims.

Airport Model Treatment of Space

Every trip end in San Diego County will be allocated to an MGRA. The trips that are external to the region will be modeled by the internal-external travel model.

Airport Model Inputs

The model system requires the following exogenously-specified inputs (note that three additional data sets are required in addition to the data currently input to the resident activity-based models):

- **SDIA Enplanement Forecast:** The total number of yearly enplanements, without counting transferring passengers, at SDIA, and an annualization factor to convert the yearly enplanements to a daily estimate. This would be given for each simulation year. The data is available in the Aviation Activity Forecast Report³.
- **Traveler characteristics distributions:** There are a number of distributions of traveler characteristics that are assumed to be fixed but can be changed by the analyst to determine their effect on the results. These include the following:

The distribution of travelers by purpose (as per Table T.15)

The distribution of travelers by purpose and household income.

The distribution of travelers by purpose and travel party size.

The distribution of travelers by purpose and trip duration (number of nights).

The distribution of travelers by purpose, direction (arriving versus departing), and time period departing for airport.

- **MGRA data.** The population and employment (by type) in each MGRA, parking cost and supply, etc. This data provides sensitivity to land-use forecasts in San Diego County. These are the same data sets as are used in the resident activity-based model.
- **TAP skim data.** Transit network level-of-service between each transit access point (transit stop). This provides sensitivity to transit network supply and cost. These are the same data sets as are used in the resident activity-based model.
- **TAZ skim data.** Auto network level-of-services between each transportation analysis zone. This provides sensitivity to auto network supply and cost. These are the same data sets as are used in the resident activity-based model.

Airport Model Description

This section describes the model system briefly, followed by a more in-depth discussion of each model component.

1. Trip Enumeration and attribution: A total number of airport trips is created by dividing the input total enplanements (minus transferring passengers) by an annualization factor. The result will be divided by an average travel party size to convert passengers to travel parties. This will be converted into a list format that will then be exposed to the set of traveler characteristic distributions, as identified above, to attribute each travel party with the following characteristics:

- Travel purpose
- Party size
- Duration of trip
- Household income
- Trip direction (it will be assumed that 50 percent of the daily enplanements are arriving passengers and 50 percent are departing passengers)
- Departure time for airport

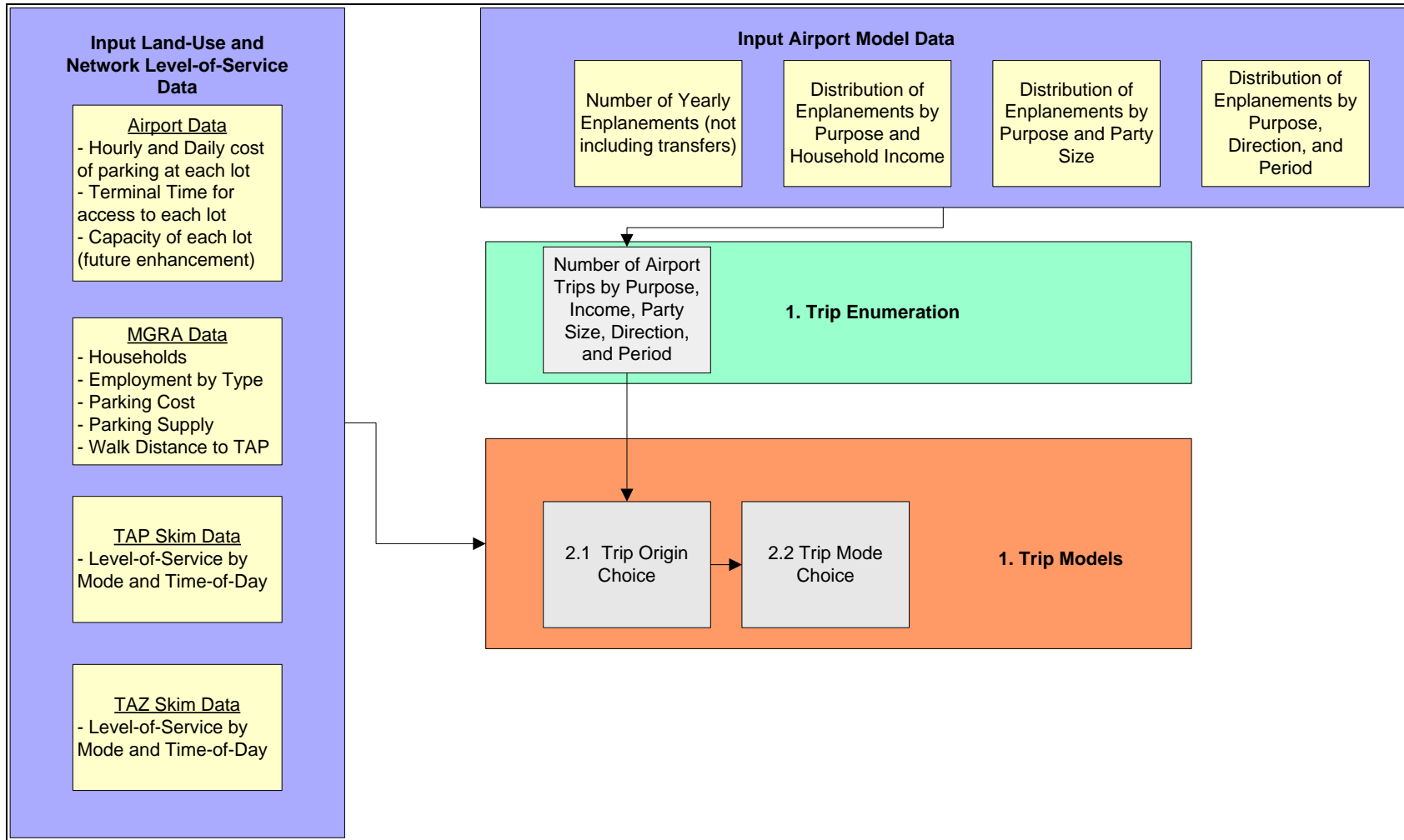
2. Trip Models

2.1 Trip origin: Each travel party will be assigned an origin MGRA.

2.2 Trip mode: Each travel party will be assigned a trip mode.

Figure T.13

SAN Airport Ground Access Travel Mode



2.1 Airport Destination Choice Model

Number of Models:	4 (Resident-Business, Resident- Personal, Visitor-Business, Visitor-Personal)
Decision-Making Unit:	Travel party
Model Form:	Multinomial Logit
Alternatives:	MGRAs

The airport destination choice model chooses the origin or destination MGRA, depending on whether the travel party is arriving or departing. The model is based upon the airport survey data, which collected the zip code of the origin location for trips made to the airport by departing passengers.

The airport destination choice model explanatory variables include:

- Trip Distance
- La Jolla destination indicator
- Size of sampled MGRA

To speed calculations, the TAZ will be sampled first since distance is the only measure of impedance used to represent accessibility of primary destination to station, and distance is represented at the TAZ level. Zone size in this case will be equal to the sum of the sizes of the MGRAs within the TAZ. Once the TAZ is sampled, an MGRA within the TAZ can be chosen based on the pre-calculated probability of the MGRA within the TAZ, which is based on the MGRA proportion of the TAZ size.

2.2 Airport Trip Mode Choice

Number of Models:	4 (Resident-Business, Resident-Personal, Visitor-Business, Visitor-Personal)
Decision-Making Unit:	Travel party
Model Form:	Nested Logit
Alternatives:	Trip Mode (see Table T.16)

Since the data in the ground access survey is too aggregate to estimate a mode choice model, trip mode was asserted based upon estimation work from applications of the similar models in other regions (for example, for Port of Portland).

Figure T.14 shows the nesting structure of the mode choice model.

The model explicitly represents the options of parking versus pick-up/drop-off for private vehicle trips. All trips are assigned either curbside (for pick-up/drop-off, taxi, shuttle/van/courtesy vehicle, and transit) or parking lot (terminal, off-site SAN lot, or off-site private lot). The choice of transit access and line-haul mode is not shown but will be modeled explicitly, as will the path choice for HOV or pay options for auto trips.

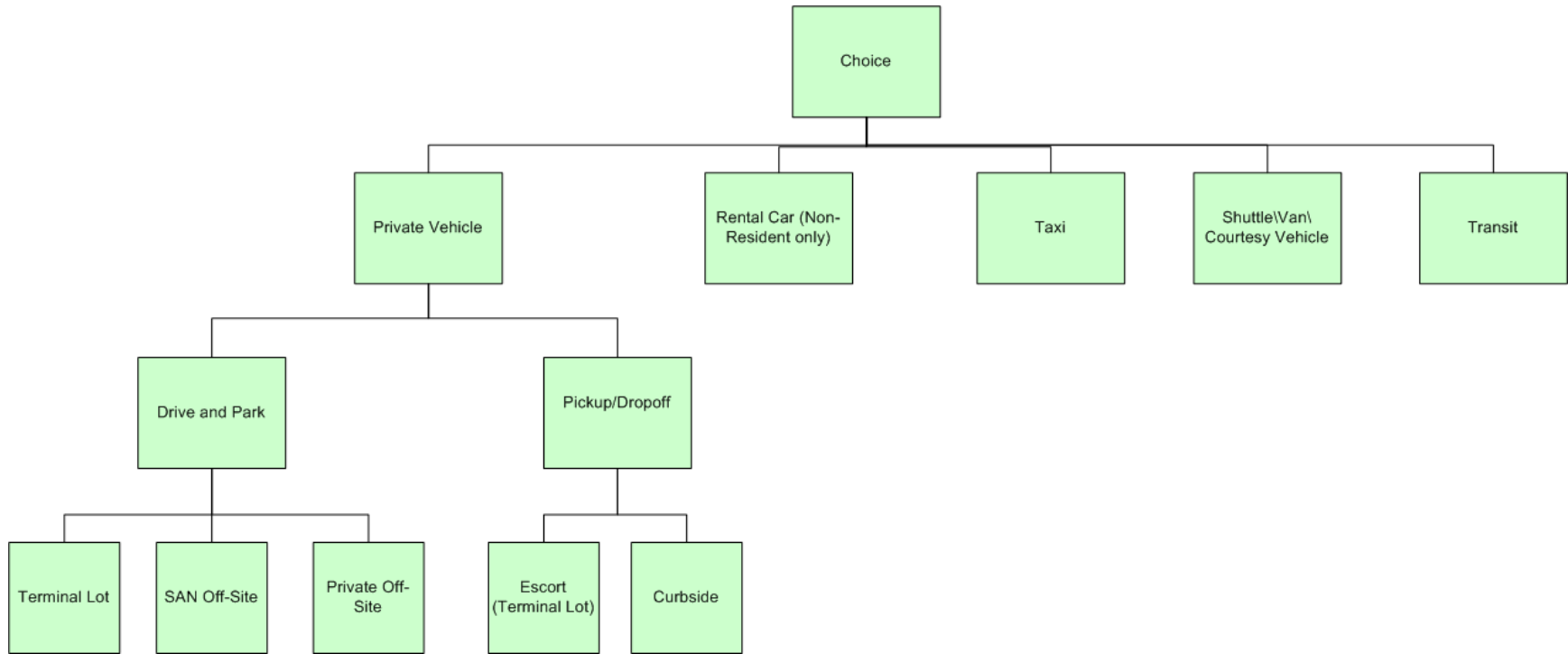
The explanatory variables included in the airport trip mode choice model are:

- Household income
- Resident status
- Party size
- Number of nights stayed in San Diego
- Trip mode
- Trip purpose
- Trip departure time

- Mode choice logsums
- In-vehicle time (auto and transit)
- Walk time
- Auto operating cost
- Auto toll value
- Transit first wait time
- Transit transfer time
- Number of transit transfers
- Transit walk access time
- Transit walk egress time
- Transit walk auxiliary time
- Transit fare
- Terminal parking walk time
- Terminal parking cost per day
- Terminal escorting cost per hour
- Off-site in-vehicle time
- Off-site walk time
- Off-site wait time
- Off-site parking cost per day
- Time to rental car center
- Walk time to rental car shuttle
- Wait time to rental car shuttle
- Rental car cost per day
- Taxi cost per trip

Figure T.14

SANDAG Airport Ground Access Mode Choice Nesting Structure



Visitor model

In 2011, the San Diego Association of Governments (SANDAG) conducted a survey of airport passengers and hotel guests in which data was collected on their travel while visiting San Diego. Based upon this data, a model of visitor travel was developed. The purpose of this model is to capture the demand of visitor travel on transport facilities in San Diego County.

The visitor model has the following features:

- A disaggregate micro-simulation treatment of visitors by person type, with explicit representation of party attributes
- Special consideration of unique visitor travel patterns, including rental car usage and visits to San Diego attractions like Sea World
- The full set of modes within San Diego County, including auto trips by occupancy, transit trips, non-motorized trips, and toll/HOT/HOV lanes modes

The model flow and inputs are shown in Figure T.15, and described in detail in the following sections.

1. Model Inputs

The model system requires the following exogenously-specified inputs (note that three additional data sets are required in addition to the data currently input to the resident activity-based models):

- **Traveler characteristics distributions.** There are a number of distributions of traveler characteristics that are assumed to be fixed but can be changed by the analyst to determine their effect on the results. These include the following:
 - Rates of visitor occupancy for hotels and separately for households
 - Shares of visitor parties by visitor segment for hotels and separately for households
 - The distribution of visitor parties by household income
 - The distribution of business segment travel parties by number of tours by purpose
 - The distribution of personal segment travel parties by number of tours by purpose
 - The distribution of visitor tours by tour purpose and party size
 - The distribution of visitor tours by tour purpose and auto availability
 - The distribution of visitor tours by outbound and return time-of-day and tour purpose
 - The distribution of visitor tours by frequency of stops per tour by tour purpose, duration, and direction
 - The distribution of stops by stop purpose and tour purpose
 - The distribution of stops on outbound tour legs by half-hour offset period from tour departure period and time remaining on tour
 - The distribution of stops on inbound tour legs by half-hour offset period from tour arrival period and time remaining on tour
- **MGRA data.** The population, employment (by type), and number of hotel rooms in each MGRA, parking cost and supply, etc. This data provides sensitivity to land-use forecasts in San Diego County. These are the same data sets as are used in the resident activity-based model.
- **TAP skim data.** Transit network level-of-service between each transit access point (transit stop). This provides sensitivity to transit network supply and cost. These are the same data sets as are used in the resident activity-based model.

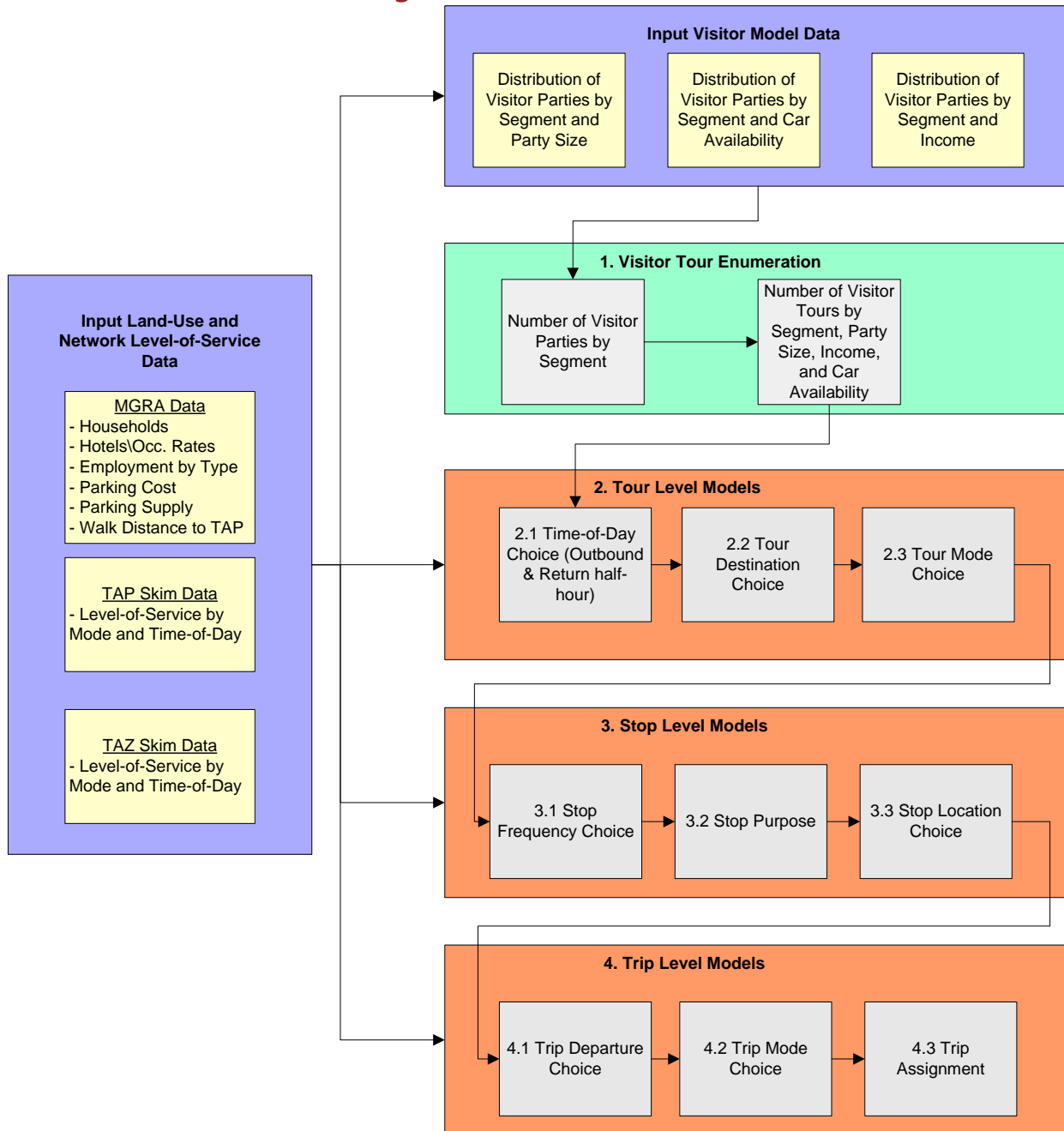
- **TAZ skim data.** Auto network level-of-services between each transportation analysis zone. This provides sensitivity to auto network supply and cost. These are the same data sets as are used in the resident activity-based model.

2. Model Description

This section describes the model system briefly, followed by a more in-depth discussion of each model component.

Figure T.15

SANDAG Visitor Model Design



1. **Visitor Tour Enumeration:** Visitor travel parties are created by visitor segment based upon input hotels and households. Travel parties are attributed with household income. Tours by purpose are generated for each party. Each tour is attributed with auto availability and party size. The tour origin MGRA is set to the MGRA where the tour was generated.
2. **Tour Level Models**
 - 2.1. *Tour Time of Day:* Each tour is assigned a time of day, based on probability distribution.
 - 2.2. *Tour Destination choice:* Each tour is assigned a primary destination, based on the coefficients estimated through a multinomial logit model.
 - 2.3. *Tour Mode Choice:* Each tour selects a preferred primary tour mode, based on an asserted nested logit model (the resident tour mode choice model).
3. **Stop Models**
 - 3.1. *Stop Frequency Choice:* Each tour is attributed with a number of stops in the outbound direction and in the inbound direction, based upon sampling from a distribution.
 - 3.2. *Stop Purpose:* Each stop is attributed with a purpose, based upon sampling from a distribution.
 - 3.3. *Stop Location Choice:* Each stop is assigned a location based upon a multinomial logit model (asserted based upon resident stop location choice models)
4. **Trip Level Models**
 - 4.1. *Trip Departure Choice:* Each trip is assigned a departure time period based upon sampling from distributions.
 - 4.2. *Trip Mode Choice:* Each trip within the tours selects a preferred trip mode, based on an asserted nested logit model.
 - 4.3. *Trip Assignment:* Each trip is assigned to the network.

Visitor Travel Parties and Tour Generation

This section describes the generation of visitor travel parties, the generation of tours, and the attribution of each.

Visitor Travel Party Generation

The number of visitors to San Diego in 2009, according to the San Diego Convention and Visitor Bureau, and are summarized by visitor segment in Table T.17.

Table T.17

Number of Visitors

	Count	Frequency
Business	2.5M	17%
Personal	11.8M	83%
Total	14.3M	100%

Visitors are generated for two visitor segment types:

- **Business:** Self-identified as business traveler, or self-identified as 'Both Business and Personal' but took at least one 'business' purpose trip on travel day.
- **Personal:** Self-identified as personal traveler, or self-identified as 'Both Business and Personal' but took no business purpose trips on travel day. A few self-identified Personal travelers have reported Work tours.

The distributions of visitors by segment in the visitor survey are shown in Table T.18. We assume that the share of business travelers in the San Diego Convention and Visitor Bureau data (17%) is more accurate than the visitor survey (25%) since the visitor survey was a place-based survey and likely did not capture a proportional share of visitors staying in households. Therefore the visitor data was re-weighted to the split of business versus personal travelers from the Convention and Visitor Bureau data.

Table T.18
Survey Respondents by Visitor Segment

Segment	Visitor Respondent Count	% of Total
Business	259	25%
Personal	769	75%
Total	1,028	100%

The model generates visitor parties by segment by applying separate occupancy rates to hotels and households, which were obtained from the San Diego Convention and Visitor Bureau. The occupancy rate for hotels is 70 percent, while the occupancy rate for households is 1.8 percent (a bit less than 2 out of every 100 households in San Diego County have visitors, on average). The model then applies separate distributions of visitor parties by segment to hotel visitor parties and household visitor parties separately. The frequencies used are shown in Table T.19.

According to the visitor survey, only 2 percent of overnight visitors stayed in a location that was not identified as a hotel or private residence. Of those, 54 percent stayed at a military base and 38 percent stayed at a vacation rental. For the purposes of this model, vacation rentals are included in the estimate of households. A small number of visitors could be allocated to the military base in the future, but this is not done currently in the model.

Table T.19
Share of Visitor Parties by Segment and Overnight Accommodation

Segment	Hotel	Household
Business	30%	4%
Personal	70%	96%
Total	100%	100%

Visitor parties are attributed with household income based upon the distribution of parties by visitor segment and income, as shown in Table T.20. Note that party size and auto availability are attributed on a tour-by-tour basis, since these attributes can change depending on which tour is undertaken and which day it is taken on.

Table T.20

Visitor Parties by Visitor Segment and Household Income

Income	Business	Personal
< \$30k	7%	34%
\$30-\$60k	29%	34%
\$60-\$100k	34%	20%
\$100-\$150k	16%	7%
\$150k+	14%	5%
Total	100%	100%

Tour Generation

Next, tours are generated by visitor parties and attributed with party size, auto availability, and income attributes. There are three tour purposes, which were coded based on the reported trip purpose in the survey, as follows:

- **Work:** Business travel made by Business travelers.
- **Recreational:** All other recreational purposes besides dining.
- **Dining:** Travel to eating establishments.

Tour purpose was coded according to a hierarchy of trip purposes, with work at the top and dining last. Tours by visitor segment are shown in Table T.21.

Table T.21

Tour Purpose by Visitor Segment

Person Type	Business	% of Total	Personal	% of Total
Work	154	59%	27	4%
Recreational	78	30%	691	90%
Dining	27	10%	51	7%
Total	259	100%	769	100%

Each travel party can generate one or more tours of each purpose on any given day. The tour generation rates are shown in Table T.22 (for the business segment) and Table T.23 (for the personal segment).

Table T.22
Tour Distribution, Business Parties

Work	Recreational	Dining	Total Tours	Frequency
1	0	0	1	40%
2	0	0	2	1%
0	1	0	1	29%
1	1	0	2	5%
0	2	0	2	1%
0	0	1	1	11%
1	0	1	2	10%
0	1	1	2	3%

Table T.23
Tour Distribution, Personal Parties

Work	Recreational	Dining	Total Tours	Frequency
1	0	0	1	3%
0	1	0	1	82%
1	1	0	2	0%
0	2	0	2	5%
0	0	1	1	6%
0	1	1	2	4%
0	2	1	3	0%
0	1	2	3	0%

The average size of the travel parties was obtained from both the San Diego Convention and Visitor Bureau numbers and the visitor survey. The visitor survey averages are slightly smaller.

Table T.24
Average Party Size

	Business	Personal
Convention and Visitor Bureau	1.3	2.4
Visitor Survey	1.04	2.2

Ultimately, the average party size observed in the survey is used in the model. The distribution of visitor tours by party size and tour purpose is shown in Table T.25.

Table T.25
Tours by Party Size and Tour Purpose

Party Size	Work	Recreate	Dining
1	31%	4%	10%
2	12%	12%	13%
3	18%	29%	33%
4	15%	23%	17%
5	3%	17%	6%
6	0%	5%	6%
7	20%	9%	14%
8	0%	0%	0%
9	0%	0%	0%
10	1%	0%	0%
Total	100%	100%	100%

Most visitors in the visitor survey did not have access to an automobile. Table T.26 shows the number and percentage of visitors who made complete tours by auto accessibility. If a person drove into San Diego, either in a personal or rental vehicle, they were assumed to have access to a car during their stay. If a person flew into San Diego and rented a car, they were also assumed to have access to a car. Persons who do not fit into either of those categories were assumed to have no vehicle. The model uses the distribution of tours by auto availability to attribute each tour with whether an auto is available, as shown in Table T.27.

Table T.26
Auto Availability by Segment

	Business	% of Total	Personal	% of Total
Drove into San Diego	18	8%	82	12%
Flew into SD, rented car	67	31%	173	25%
No Vehicle	197	60%	617	64%
Total	215	100%	699	100%

Table T.27
Auto Availability by Tour Purpose

Auto Available	Work	Recreate	Dining
Yes	38%	58%	53%
No	62%	42%	47%
Total	100%	100%	100%

2.1 Visitor Time of Day

Number of Models: 3 (Work, Recreational, Dining)
Decision-Making Unit: Tour
Model Form: Look up table
Alternatives: 40 half-hour time-of-day periods by purpose

This model selects an outbound and return half-hour period, based on a probability distribution created using the visitor survey observed tour arrival and departure data, by tour purpose. Model input is the observed percent of tours by purpose with each combination of departure and arrival time period.

2.2 Visitor Destination Choice

Number of Models: 3 (Work, Recreational, Dining)
Decision-Making Unit: Tour
Model Form: Multinomial logit
Alternatives: MGRAs

The tour destination choice model predicts the 'preferred' destination for the tour at the level of the Master Geographic Reference Area (MGRA). There are two stages involved in the estimation and application of the model. In the first phase, a list of sampled MGRAs is created. In the second phase, a multinomial logit form is applied to each sampled alternative and a destination MGRA is selected. The two-stage procedure is necessary in order to minimize the computational burden associated with computing mode choice logsums for each tour to 21633 MGRAs. In both estimation and application, 30 destination MGRAs were sampled.

Mode choice logsums used in this model were based upon an asserted mode choice model that was derived from the resident discretionary tour mode choice model.

The visitor survey contains information about the trips for each visitor during one day in San Diego. A handful of MGRAs, containing popular attractions like the Convention Center, Sea World, and the San Diego Zoo attracted a large number of the visitor trips, although destinations did occur throughout the model area.

Table T.28
Special Attractors

Location	Trips	Percent of Total
Convention Center	120	4%
Sea World	120	4%
San Diego Zoo	92	3%
Balboa Park	75	3%
Coronado Beach	62	2%
Legoland	32	1%
San Diego Zoo Safari Park	16	1%
Total trips in data frame	2,914	100%

The visitor destination choice model explanatory variables are:

- Mode choice logsum
- Distance
- Size of the sampled MGRA

2.3 Visitor Tour Mode Choice

Number of Models:	5 (Business-Work, Business-Recreational, Business-Dining, Personal-Recreational, Personal-Dining)
Decision-Making Unit:	Tour
Model Form:	Nested Logit
Alternatives:	15 Tour Modes (all the residential tour modes plus taxi except bike, park and ride, kiss-and-ride, and school bus)

This model chooses a tour mode based on a known trip origin and destination MGRAs, and travel party characteristics including purpose, party size, and income. Since the data in the visitor survey is too aggregate to estimate a mode choice model, the model was asserted based on the resident tour mode choice model. Figure T.16 shows a nesting structure of the mode choice model. Note that the modes are the same as are used in the resident model, with the addition of the taxi mode, which utilizes the same coefficients as auto modes (in-vehicle time and cost), though cost is based on an initial fare (meter-drop) and a cost-per-mile.

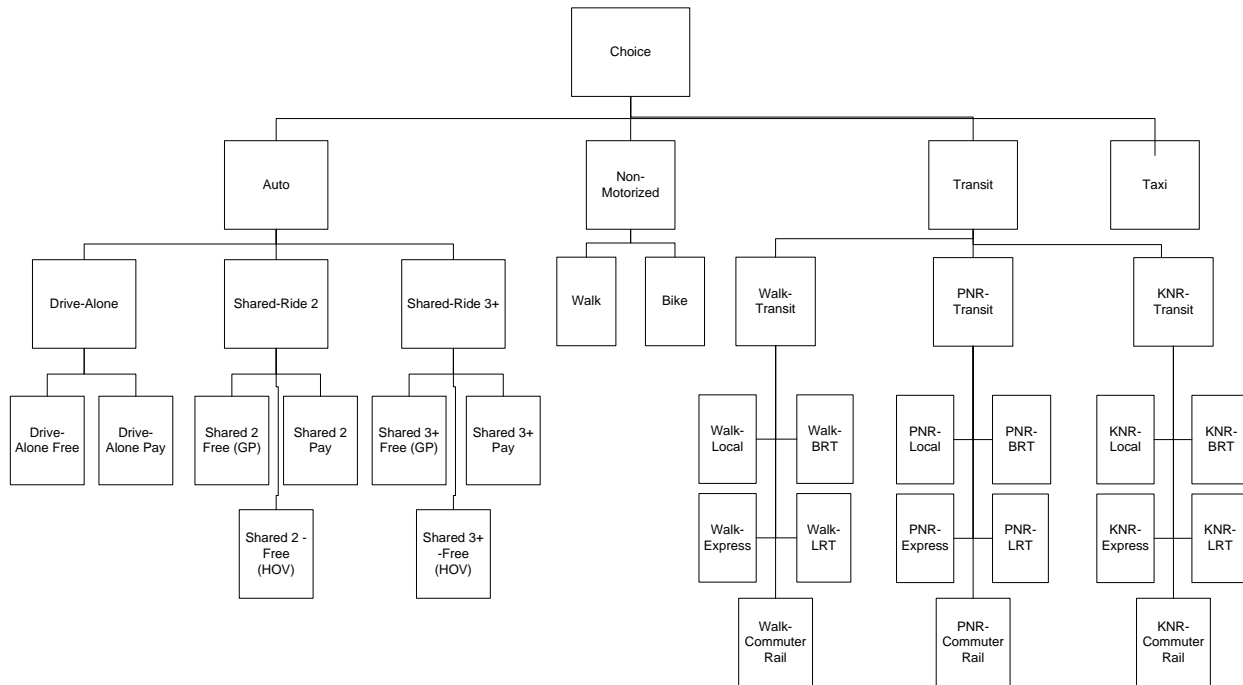
The visitor tour model explanatory variables include:

- Auto availability indicator
- In-vehicle time (auto and transit)
- Walk time
- Auto operating cost
- Auto parking cost
- Auto terminal time
- Auto toll value
- Transit first wait time
- Transit Transfer time
- Number of transit transfers
- Transit walk access time
- Transit walk egress time
- Transit walk auxiliary time
- Transit Fare

Table T.29
Survey Respondents by Purpose and Trip Mode

Mode	Tours				Trips on Tours			
	Business	% of Total	Personal	% of Total	Business	% of Total	Personal	% of Total
Personal Vehicle	49	19%	279	36%	75	11%	633	34%
Rental Car	75	29%	233	30%	160	24%	528	29%
Taxi	29	11%	184	24%	93	14%	458	25%
Walk/Bike	104	40%	68	9%	330	49%	201	11%
Bus	1	0%	4	1%	5	1%	18	1%
Trolley	1	0%	0	0%	5	1%	2	0%
Coaster	0	0%	0	0%	1	0%	4	0%
Other Mode	0	0%	1	0%	1	0%	5	0%
Total	259	100%	769	100%	670	100%	1849	100%

Figure T.16
SANDAG Nested Tour Mode Choice model structure



3.1 Visitor Stop Frequency

Number of Models: 1
 Decision-Making Unit: Tour
 Model Form: Probability Distribution Lookup table
 Alternatives: 0-1 outbound stops by 0-3 inbound stops

Number of stops per tour is determined by sampling from the observed distribution of number of stops per tour. The model input is the percentage of observed number of stops (both inbound and outbound) by purpose and tour duration. The input frequency table is too large for documentation, but the frequency of stops by tour purpose and segment is shown in Table T.30. For simplicity’s sake, intermediate stops are not allowed on non-motorized or transit tours (which speeds up and simplifies the intermediate stop sampling procedure).

Table T.30
Frequency of Stops on Tour

Number of Stops	Business Respondents			Personal Respondents	
	Work	Shop\Recreate	Eat Out	Shop\Recreate	Eat Out
0	31%	47%	100%	52%	94%
1	41%	44%	0%	42%	6%
2	28%	9%	0%	6%	0%
Total	100%	100%	100%	100%	100%

3.2 Visitor Stop Purpose

Number of Models: 1
 Decision-Making Unit: Stop
 Model Form: Probability Distribution Lookup Table
 Alternatives: Work, Other Recreational, Dining

Purpose of stops is determined by the observed purpose of stops in the visitor survey. The model input is the percentage of observed stops by purpose, stop number, number of stops on tour, and stop direction (inbound or outbound). The actual table used is too large to include in documentation, but the frequency of stops by tour purpose and segment is shown in Table T.31.

Table T.31
Stops by Purpose

Stop Purpose	Business Respondent			Personal Respondent	
	Work	Shop\Recreate	Eat Out	Shop\Recreate	Eat Out
Work	100%	63%	76%	5%	5%
Shop\Recreate	0%	37%	24%	95%	93%
Eat Out	0%	0%	0%	0%	2%

3.3 Visitor Stop Location

Number of Models:	1
Decision-Making Unit:	Stop
Model Form:	Multinomial Logit
Alternatives:	MGRAs

The stop location model was asserted, based on the discretionary purpose of the resident stop location choice model. See Section 5.3.

4.1 Visitor Trip Time of Day

Number of Models:	1
Decision-Making Unit:	Travel party
Model Form:	Lookup table
Alternatives:	Half-hour offsets

The stop time of day is chosen based on a distribution of observed stop durations from the survey. Distributions were prepared for stop duration for outbound and inbound stops, by purpose, and overall tour duration.

4.2 Visitor Trip Mode Choice

Number of Models:	5 (Business-Work, Business-Recreational, Business-Dining, Personal-Recreational, Personal-Dining)
Model Form:	Nested Logit
Alternatives:	15 Tour Modes (all the residential tour modes plus taxi except bike, park and ride, kiss-and-ride, and school bus)

As with tour mode choice, this model is asserted. It is based on the resident trip mode choice model, with the addition of a taxi mode.

The visitor trip mode selection is based on the following explanatory variables:

- Auto availability indicator
- In-vehicle time (auto and transit)
- Walk time
- Auto operating cost
- Auto parking cost
- Auto terminal time
- Auto toll value
- Transit first wait time
- Transit transfer time
- Number of transit transfers
- Transit walk access time
- Transit walk egress time
- Transit walk auxiliary time
- Transit fare

Special event model

A special event survey was conducted in San Diego in 2011, collecting information on trips to and from a number of sporting and cultural events throughout San Diego County. This data has been used to implement a special event model component for the San Diego region. The model shares similar features with the event model originally developed for Phoenix, including an origin choice and mode choice component. However, in keeping with the other model components developed as part of the activity-based model development project, the SANDAG special event model is a simulation-based framework. Travel parties to and from the special event are generated from input event attendance, and key characteristics (party size and income) are selected from observed distributions. Then each travel party chooses an explicit origin micro-zone (Master Geographic Reference Area) and an explicit mode. The model produces a list of trips to and from the event, with a number of records equal to twice the number of event attendees (one trip to the event and one trip from the event).

Survey Data

Table T.32 shows the special event data used to develop the model. There were six different events surveyed, including concerts, sporting events, the Gaslamp Mardi Gras festival, and the San Diego County Fair.

Table T.32
Special Events Surveyed, Location, Samples and Annual Attendance

Event Name	Type	MGRA	Number of Samples	Average Attendance Per Event-Day
Gaslamp Mardi Gras	Other	2351	219	40,000
Viejas Arena (SDSU)	Concert\Sporting	955	266 (186\80)	12,000
Petco Park (Baseball)	Sporting	2436	586	23,000
San Diego County Fair (Fairgrounds)	Other	12395	369	50,000
Cricket Wireless Amphitheater (Chula Vista)	Concert	6468	245	11,000
Qualcomm Stadium (Football)	Sporting	5855	183	65,000

Characteristics of the respondents at each event were analyzed to determine appropriate groupings of events. Figure T.17 shows survey respondents by age range. The chart shows that Gaslamp Mardi Gras attendees and attendees at Viejas Arena tend to be younger, while attendees at the County Fair tend to be older.

Figure T.17
Survey Respondents by Age Range and Event

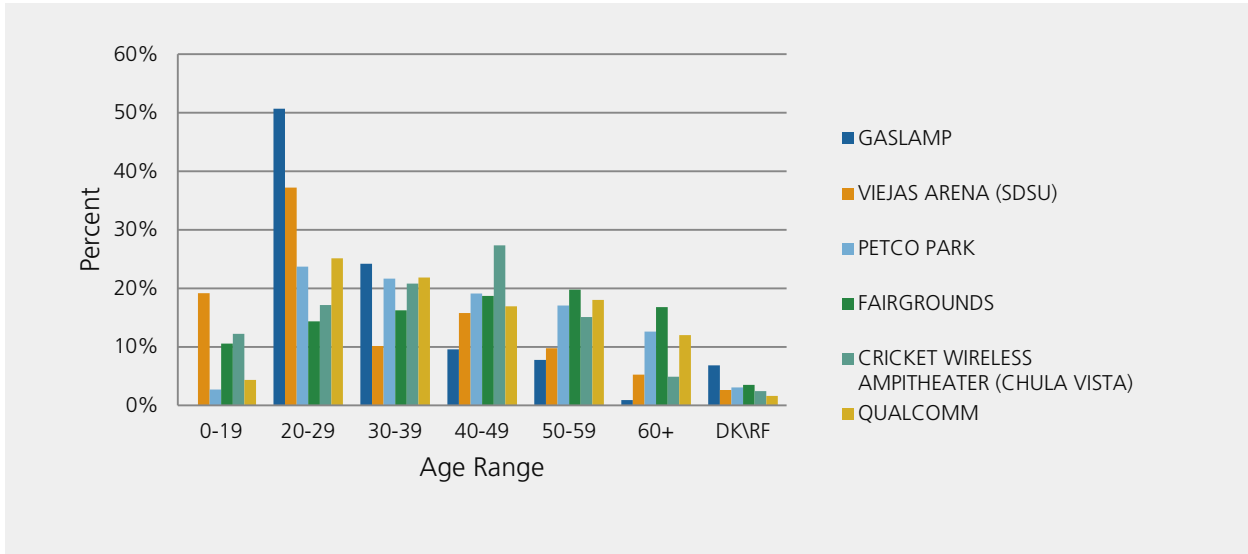


Figure T.18
Respondents by Party Size and Event

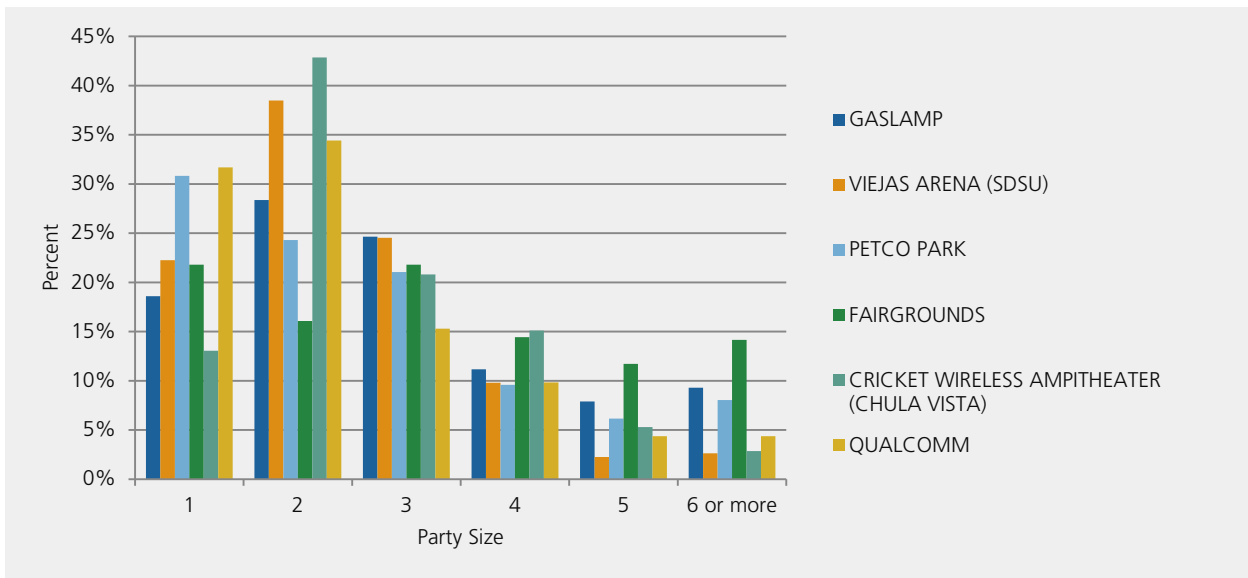


Figure T.18 shows survey respondents by party size and event. Party sizes at Petco Park tend to be smaller than other events, while party sizes at the Fairgrounds tend to be larger. The fairgrounds may attract school classes, resulting in a larger average party size than other events. Events at the Cricket Wireless Amphitheater and at Viejas Arena tend to attract a greater share of two-person parties.

Figure T.19
Survey Respondents by Household Income Range and Event

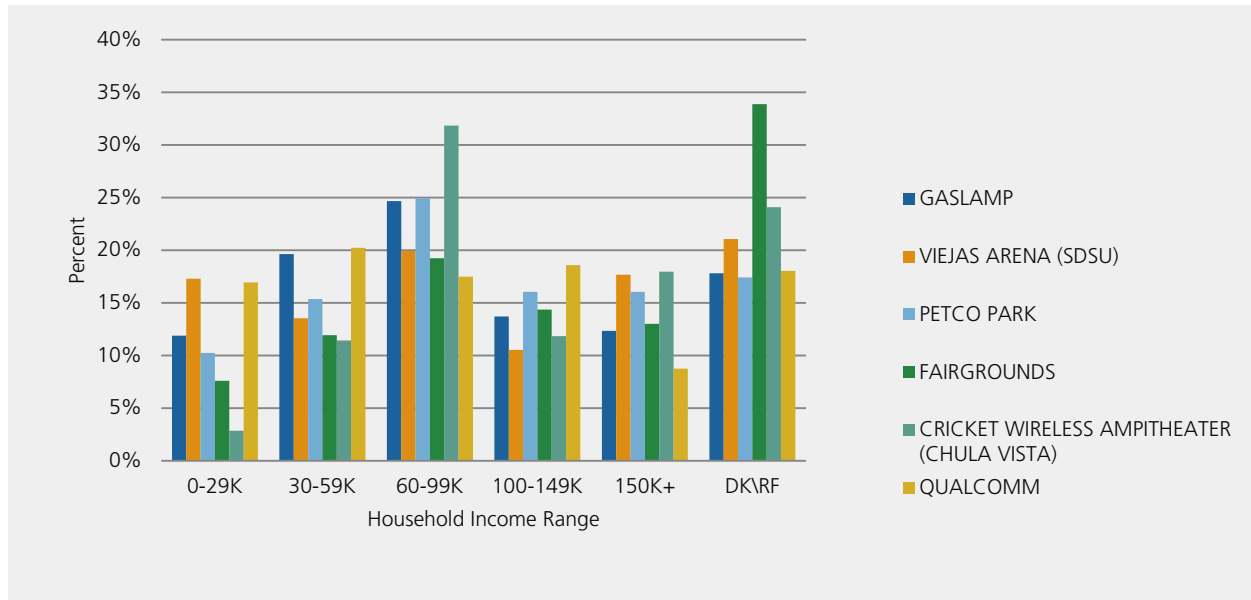


Figure T.19 shows survey respondents by household income and event. The most noticeable difference among events is that Chula Vista appears to attract a larger proportion of attendees in the medium (\$60k-99k) income range than other events, while events at Viejas Arena tend to attract relatively lower income attendees (as might be expected given that the venue is at a major university and is likely attended by a greater proportion of college students with limited income).

Figure T.20
Survey Respondents by Origin Place Type and Event

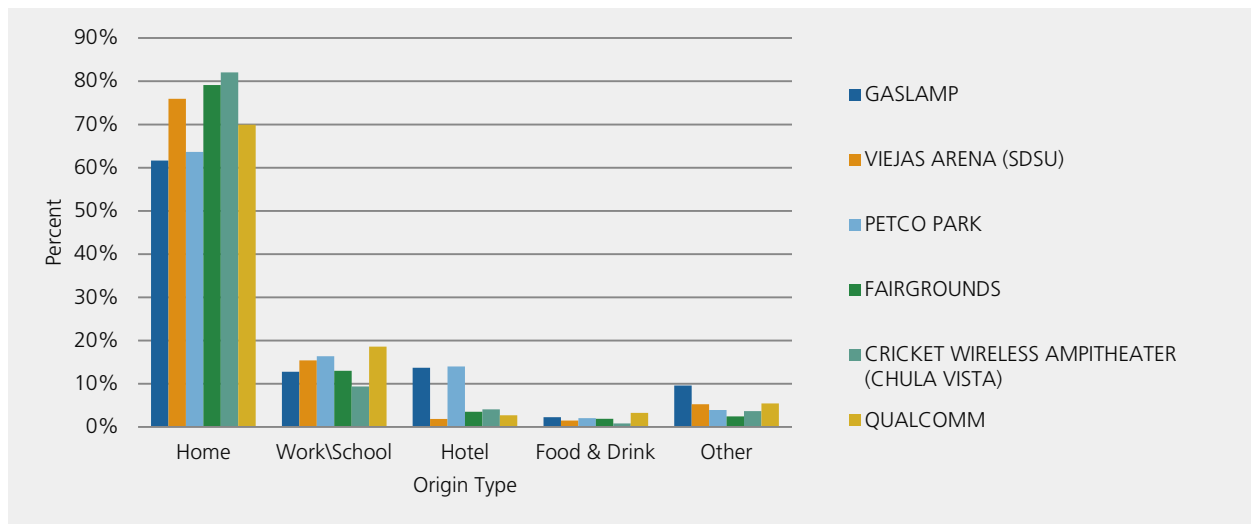


Figure T.20 shows survey respondents by origin place type and event. The chart shows that between 60 percent and 80 percent of attendees travel to the special event directly from home. Special events with a relatively lower share of attendees from home tend to have a higher share of attendees from hotel – including Gaslamp Mardi Gras and Petco Park. This is reasonable given that both of these venues are in the Gaslamp area, which has a high concentration of hotels. The special event survey also collected information on destination location after the event. As can be seen in Figure T.21, between 80 percent and 90 percent of respondents return to the same location after the event. The special events with lower shares of respondents who return to the same location after the event correspond to events

with lower shares of respondents whose origin is home, such as Gaslamp Mardi Gras and Petco Park. Again, given the high concentration of restaurants and bars around these venues, it is not surprising that a greater portion of attendees are not going directly home after these events.

Figure T.21
Respondents by Destination Location

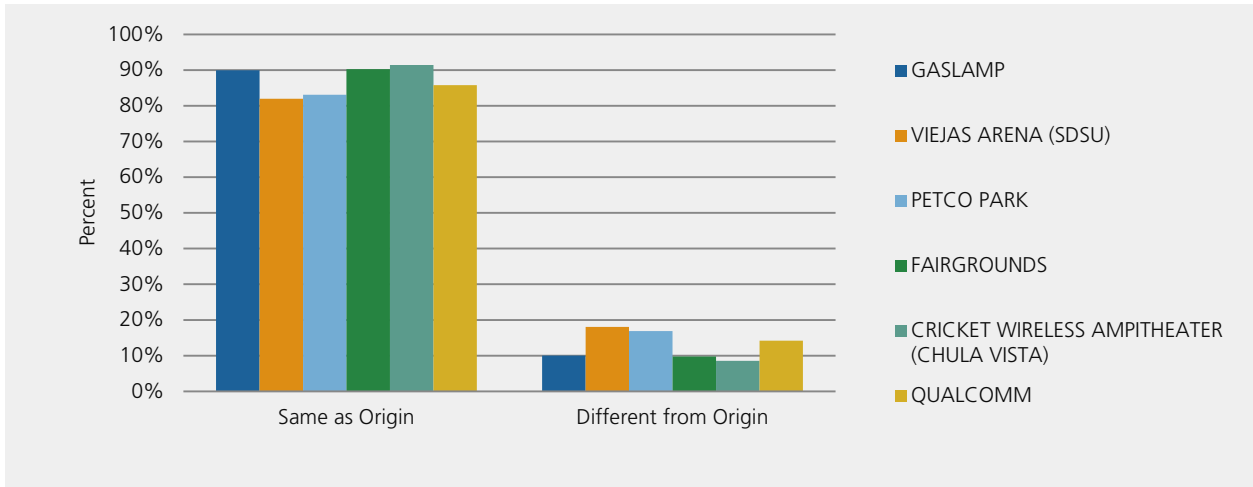


Figure T.22
Respondents by Distance to Special Event

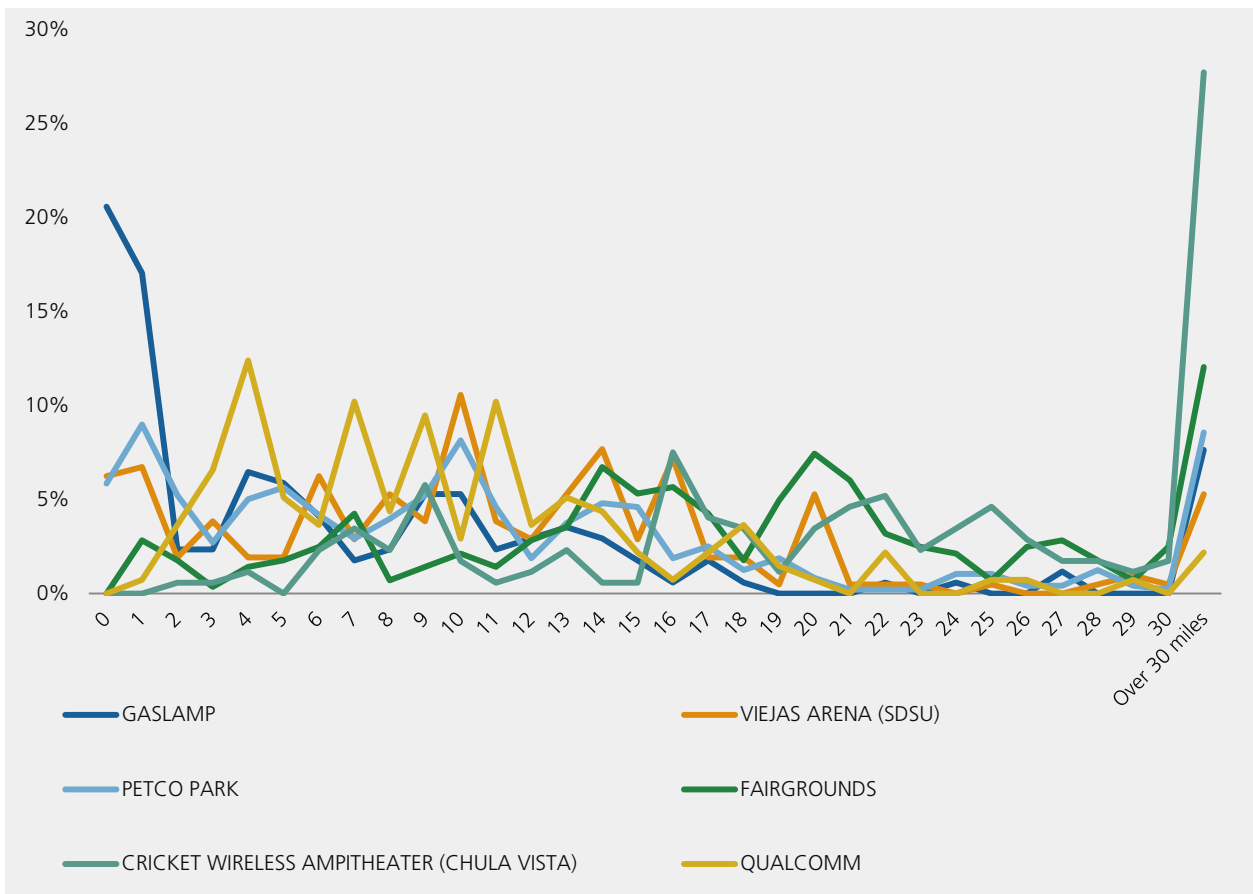


Table T.33
Average Distance by Special Event

Special Event	Average Distance
Gaslamp Mardi Gras	8.8
Viejas Arena (SDSU)	12.0
Petco Park (Baseball)	12.1
San Diego County Fair (Fairgrounds)	19.1
Cricket Wireless Amphitheater (Chula Vista)	24.6
Qualcomm Stadium (Football)	10.7
Total	14.4

Trip length frequency distributions to each event are shown in Figure T.22, and average distance to event is shown in Table T.33.

Trip length distributions are lumpy due to the relatively low number of observations at each event and the relatively high rate of non-geocodable observations (between 20 percent and 30 percent at each event). However, it is clear that certain events, such as the San Diego County Fair and the Cricket Wireless Amphitheater in Chula Vista, tend to draw from a much larger area than other, more centrally located events such as Gaslamp Mardi Gras and football games held at Qualcomm Stadium.

Figure T.23 shows survey respondents by mode to special events. Although carpool tends to be the dominant travel mode, there is considerable variation in mode usage across events. Qualcomm has the highest share of trips by transit, with the vast majority of transit trips on the trolley mode. Note that Qualcomm Stadium is well-served by a stop on the trolley Green Line, and MTS provides extra service on football game-days. Events at Petco Park, SDSU, and Gaslamp also have relatively higher shares of transit patronage than other events. In all cases in which transit shares are significant, trolley carries a much higher share of attendees than local bus.

Figure T.24 shows transit trips by mode of access and event. Transit trips to Gaslamp Mardi Gras, SDSU, and Petco Park tend to be more walk-access (45% to 55%), while transit trips to Qualcomm and the Fairgrounds tend to be more drive-access (50% to 55%). It appears that the transfer rate is low (only 5% to 10% of respondents report bus as their mode of access) as is kiss-and-ride (1% to 5% of transit trips). Note that there was only one observed transit trip to Cricket Wireless Amphitheater, which reported walk as the access mode, resulting in a 100% walk-transit share for that venue.

Table T.34 shows parking costs by venue, including the highest parking cost paid across all respondents, the average parking cost paid across all respondents and the average parking costs paid among respondents who paid for parking (not including respondents who walked, took transit, or were dropped off at the event). Not surprisingly, the events with higher parking costs also have relatively higher shares of transit use.

Figure T.23
Respondents by Mode to Event

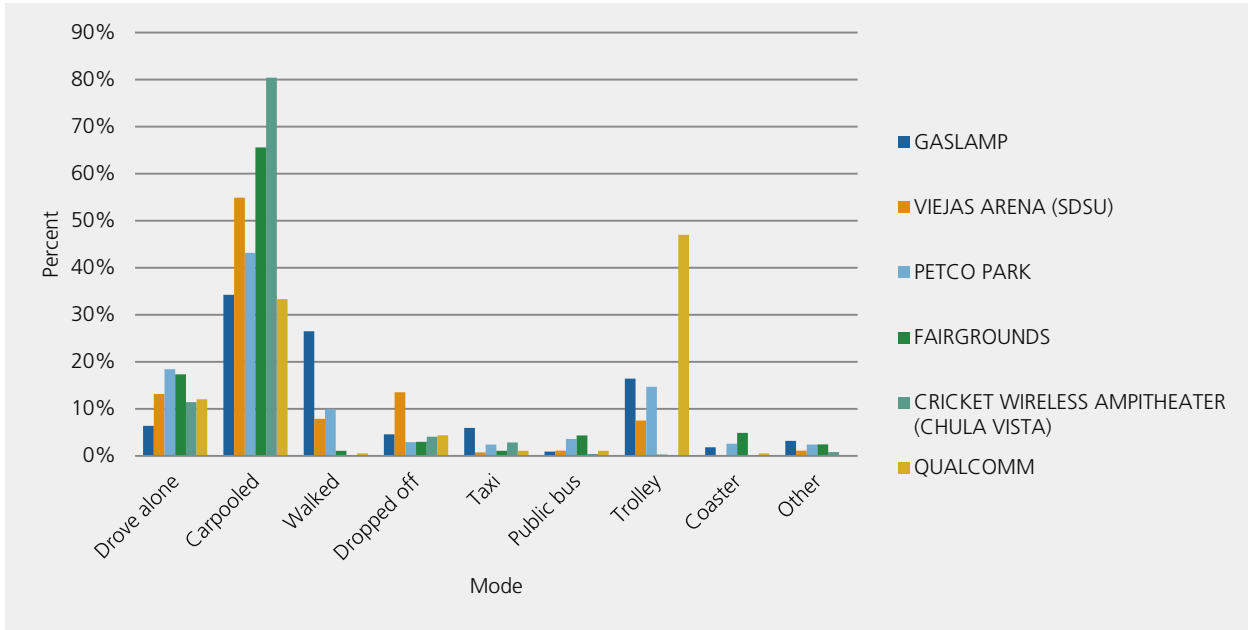


Figure T.24
Transit Trips by Mode of Access to Transit

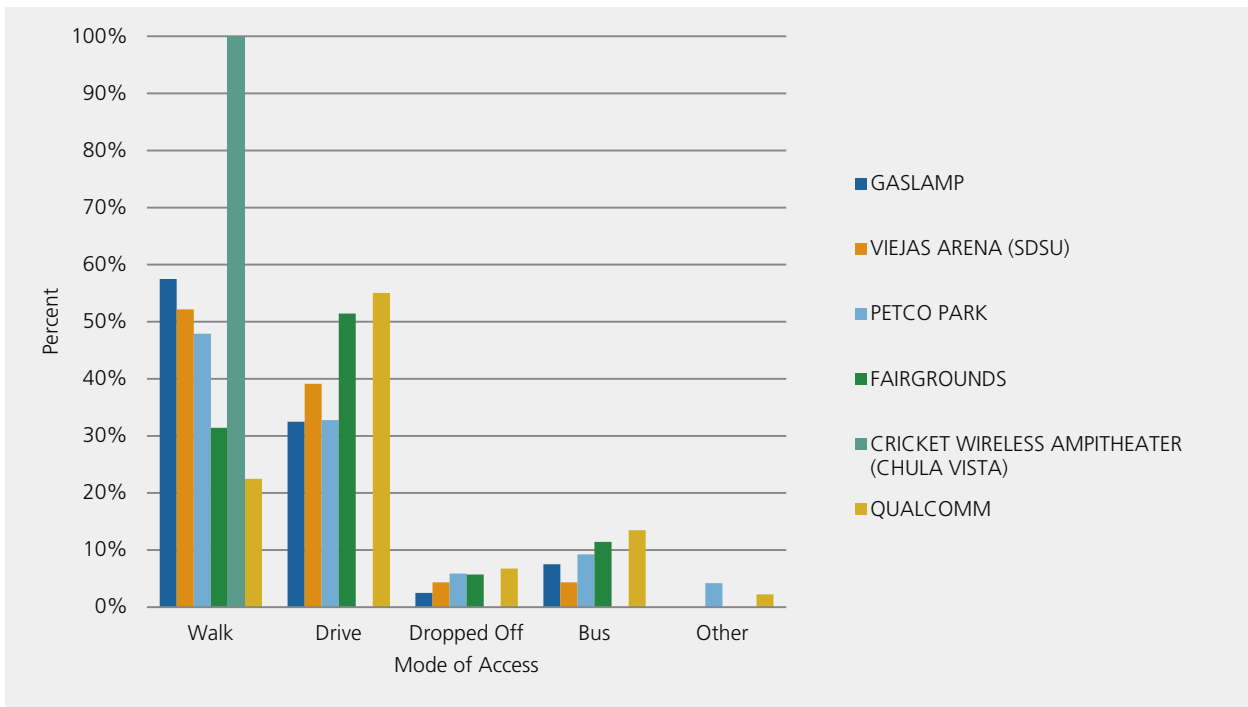


Table T.34
Parking Costs by Venue

Special Event	Average Parking Paid	Maximum Parking Cost	Average Parking Paid, of people who paid
Gaslamp Mardi Gras	\$4.07	\$36.00	\$15.29
Viejas Arena (SDSU)	\$3.94	\$16.00	\$9.04
Petco Park (Baseball)	\$6.26	\$32.00	\$12.62
San Diego County Fair (Fairgrounds)	\$7.13	\$20.00	\$10.73
Cricket Wireless Amphitheater (Chula Vista)	\$1.76	\$35.00	\$21.50
Qualcomm Stadium (Football)	\$7.23	\$40.00	\$24.35
Average	\$5.75	\$40.00	\$12.73

Model Form

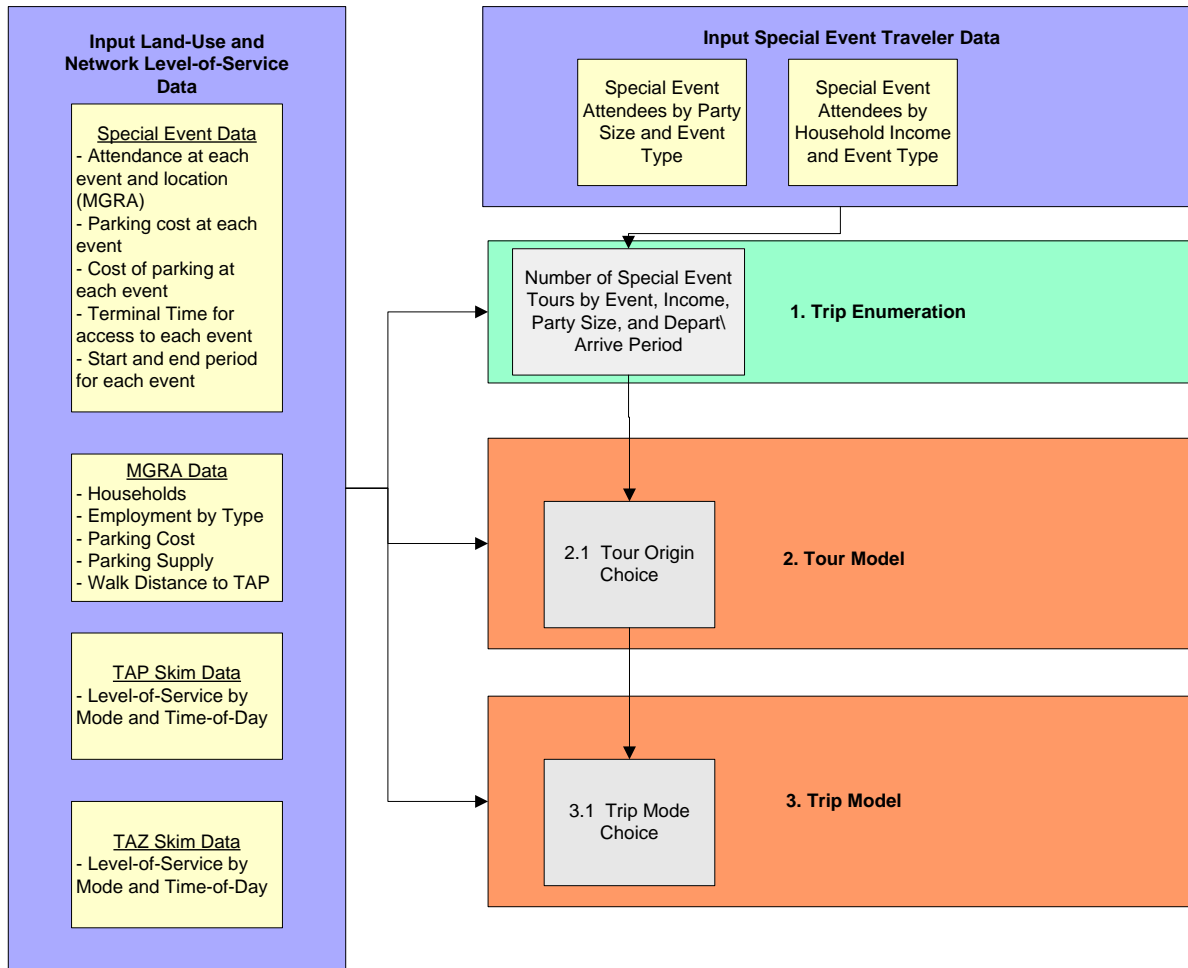
The special event model is formulated as an origin and mode choice model, as shown in Figure T.25. The model reads a special event input file (see Table T.35) and generates special event travel parties based upon the input attendance at each special event. Given that there is a fairly low share of trips to special events that do not return to the same location, an origin is chosen for every Special Event person tour. Trips are generated to and from the event for each tour, and the mode of each trip is modeled independently.

The following section describes each model.

Table T.35
Example Special Event Input File

Event Number	MGRA	Attendance	Event Type	Start Period	End Period	Parking Cost	Parking Time
1	2436	23,000	Sporting	28	37	1262	5
2	955	12,000	Sporting	28	37	904	10
3	18721	12,000	Cultural	26	35	904	10

Figure T.25
Special Event Model Form



Trip Enumeration

Special event tours are generated based on total attendance. Tours are attributed with party size (number of participants) and household income based upon input distributions which vary by event type (see Table T.35). The event types are flexibly defined in the input event table, but at a minimum, based upon observed data, we recommend Sporting and Cultural events as two distinct categories of events. Alternatively, one could input specific party size and income distributions for each event based upon the data shown in the above tables and charts. For purposes of calibration, we used data for Petco Park.

Tour Model

There is only one tour model, in which the origin is chosen.

Tour Origin Choice

Number of Models: 1 per event type (size terms estimated for each venue)
 Decision-Making Unit: Person-tour
 Model Form: Multinomial logit
 Alternatives: 21,633 MGRAs

The tour origin choice model predicts the origin for the Special Event tour at the level of the Master Geographic Reference Area (MGRA). There are two stages used for destination choice. In the first stage, a TAZ is chosen based upon distance to the venue and size term. In the second stage, an MGRA within the chosen TAZ is selected based

upon the relative size of the MGRA compared to the size of the TAZ. Note that mode choice logsums are not used in the origin choice model.

Origin choice model size terms were estimated for each venue. They are shown in Table T.36.

The size term variables are defined as follows:

- Low income households: \$0-\$30k
- Medium income households: \$30k- \$150k
- High income households: \$150k+
- Retail employment: Retail activity
- Service employment: Professional and Business Services, Personal Services Office-Based, and Personal Services Retail-Based
- Hotel Employment: Hotels and Models
- Restaurant Employment: Restaurants and Bars

A distance term was not estimated. However, an asserted term (-0.05) was compared to observed data and found to match trip length very well for Petco Park. The term may need to be adjusted somewhat for venues with longer trip lengths.

Table T.36
Estimated Size Terms (Exponentiated Coefficients)

Variable	All Venues	Gaslamp Mardi Gras	Viejas Arena	Petco Park	Fair-grounds	Cricket Amphitheater	Qualcomm
Low Income HH	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Med Income HH	1.60	2.17	0.77	0.50	3.46	5.05	43.82
High Income HH	2.80	0.00	2.01	1.68	1.49	4.22	0.28
Retail Employment	0.27	0.05	0.01	0.08	0.13	0.63	0.00
Service Employment	0.72	0.67	0.89	3.42	0.90	4.57	10.70
Hotel Employment	0.52	0.70	1.22	0.11	0.98	0.88	1.22
Restaurant Employment	0.86	0.34	0.13	0.34	0.44	1.16	4.81

Trip Mode Choice

Number of Models: 1
 Decision-Making Unit: Person-trip
 Model Form: Nested logit

An asserted nested logit mode choice model is used to model trip mode choice for trips to and from special events. The model utilizes the same functional form as the resident mode choice models, though many of the person and household specific variables are not available in the special event model and are therefore not used. The asserted mode choice parameters are shown in Table T.37.

Table T.37**Special Event Trip Mode Choice Model Asserted Parameters**

Coefficient	Value	Ratio to IVT (\$/hr)
In-vehicle time coefficient	-0.0300	1.0
First wait time coefficient	-0.0457	1.5
Transfer wait time coefficient	-0.0457	1.5
Walk access time coefficient	-0.0457	1.5
Walk egress time coefficient	-0.0457	1.5
Walk auxiliary time coefficient	-0.0457	1.5
Transfer penalty (above one)	-0.7200	24
Walk mode time coefficient	-0.0536	1.8
Transfer penalty - PNR	$c_{ivt} * 15$	24
Cost coefficient for income < \$30k	-0.0075	(\$2.40)
Cost coefficient for income \$30k-\$60k	-0.0033	(\$5.45)
Cost coefficient for income \$60k - \$100k	-0.0021	(\$8.57)
Cost coefficient for income > \$100k	-0.0009	(\$20.00)
Express bus IVT factor	0.9000	0.9000
BRT IVT factor	0.9000	0.9000
LRT IVT factor	0.8500	0.8500
Commuter rail IVT factor	0.7500	0.7500

External model

The external travel models predict characteristics of all vehicle trips and selected transit trips crossing the San Diego County border.

External Model Definition of Trip Type

The external-external, external-internal, and internal-external trips in San Diego County were segmented into the following trip types:

- **US-US:** External-external trips whose production and attraction are both in the United States, but not in San Diego County.
- **US-MX:** External-external trips with one trip end in the United States and the other in Mexico.
- **US-SD:** External-internal trips with a production elsewhere in the United States and an attraction in San Diego County.
- **MX-SD:** External-internal trips with a production in Mexico and an attraction in San Diego County (covered by the Mexican resident microsimulation model).
- **SD-US:** Internal-external trips with a production in San Diego and an attraction elsewhere in the United States.

- **SD-MX:** Internal-external trips with a production in San Diego County and an attraction in Mexico.

External Model Estimation of Trip Counts by Type

The total count of trips by production and attraction location was estimated in a series of steps:

1. The number of trips made by Mexican residents to attractions in San Diego was previously determined during development of the Mexican resident travel microsimulation model.
2. The trips in the resident travel survey were expanded to estimate the total number of trips made by San Diego residents to attractions in Mexico.
3. The number of MX-SD (1) and SD-MX (2) trips was subtracted from the total number of border-crossings to derive an estimate of the number of US-MX trips. The distribution of US-MX trips among external stations on the US-side of San Diego County will be assumed to be proportional to the total volume at each external station, regardless of the point of entry at the Mexican border.
4. The number of US-MX trips was then subtracted from the total number of trips in the SCAG cordon survey to arrive at an estimate of the combined total of US-US, US-SD, and SD-US trips with routes through San Diego County.
5. Finally, the actual amounts of US-US, US-SD, and SD-US trips at each external station were estimated from the remaining trips (4) according to their proportions in the successfully geocoded responses in the SCAG cordon survey.

External Model Design Overview

The behavioral characteristics of the different types of external trip were derived from the various data sources available as follows:

- **US-US trips:** A fixed external station OD trip matrix was estimated from the SCAG cordon survey.
- **US-MX trips:** A fixed external station OD trip matrix was estimated from the SCAG cordon survey, Customs and Border Protection vehicle counts, and Mexican resident border-crossing survey as described in the previous section.
- **US-SD trips:** Rates of vehicle trips per household for each external county were developed from the SCAG cordon survey, and the trips were distributed to locations in San Diego County according to a destination choice model estimated from the interregional survey.
- **MX-SD trips:** A microsimulation model of Mexican resident travel has been developed in a separate task.
- **SD-US trips:** A binary logit model for a person's making a trip as a function of accessibility to external stations and demographic characteristics was developed from the San Diego County resident survey, and the trips were distributed to external stations according to their market shares in the base year, which were estimated as described in the previous section.
- **SD-MX trips:** A binary logit model simulating an individual's decision to make a trip as a function of accessibility to external stations and demographic characteristics was developed from the San Diego County resident survey, and the trips were distributed to external stations according to their market shares in the base year, which were estimated as described in the previous section.

External-External (EE) Trips

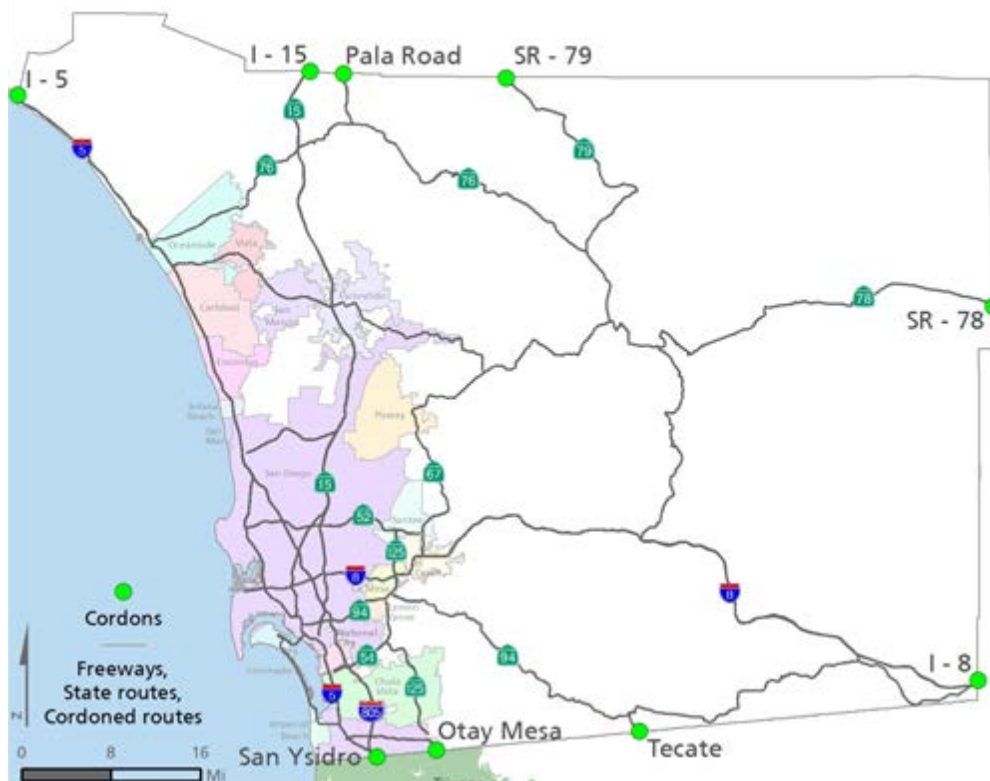
Number of Models: 1
 Model Form: Fixed Trip table

The EE trip matrix (covering US-US and US-MX trips) is estimated as described in the previous sections appears in Table T.38.

Table T.38
External-External Trip Matrix

Origin	Destination									
	San Ysidro	Otay Mesa	Tecate	I-8	CA-78	CA-79	Pala Road	I-15	I-5	Total
San Ysidro	-	-	-	167	17	25	37	1,563	1,527	3,336
Otay Mesa	-	-	-	42	4	6	9	396	387	844
Tecate	-	-	-	13	1	2	3	124	121	264
I-8	167	42	13	-	-	-	-	-	22	244
CA-78	17	4	1	-	-	-	-	-	-	22
CA-79	25	6	2	-	-	-	-	-	-	33
Pala Road	37	9	3	-	-	-	-	-	-	49
I-15	1,563	396	124	-	-	-	-	-	1,086	3,169
I-5	1,527	387	121	22	-	-	-	1,086	-	3,143
Total	3,336	844	264	244	22	33	49	3,169	3,143	11,104

Figure T.26
San Diego County Cordons



US-SD External-Internal (EI) Trips

The US-SD External-Internal trip model covers vehicle trips with destinations in San Diego made by persons residing in other areas of the United States. Intermediate stops and transit trips are not modeled in this segment due to the small contribution of these events to the total demand in the segment.

The US-SD model accepts as an input the total number of work and non-work vehicle trips from the SCAG cordon survey at each external station (Table T.39).

Table T.39
US-SD Trips by Production Location

Production	Trips
I-5	28,820
I-15	33,661
Pala Rd	813
CA-79	543
I-8	3,413
CA-78	344
Total	67,593

External-Internal Destination Choice Model

Number of Models: 2 (Work, Non-work)
Decision-Making Unit: Tour
Model Form: Multinomial logit
Alternatives: MGRAs

The external-internal destination choice model distributes the EI trips to destinations within San Diego County.

The EI destination choice model explanatory variables are:

- Distance
- The size of each sampled MGRA

Diurnal and vehicle occupancy factors (Table T.40 and Table T.41) are then applied to the total daily trip tables to distribute the trips among shared ride modes and different times of day.

Table T.40
US-SD Vehicle Occupancy Factors

Vehicle Occupancy	Percent
One	58%
Two	31%
Three or more	11%
Total	100%

Table T.41
US-SD Diurnal Factors

Time Period	Work Percent		Non-Work Percent	
	P to A	A to P	P to A	A to P
Early AM	26%	8%	25%	12%
AM Peak	26%	7%	39%	11%
Midday	41%	41%	30%	37%
PM Peak	6%	42%	4%	38%
Evening	2%	2%	2%	2%
Total	100%	100%	100%	100%

External-Internal Toll Choice Model

Number of Models: 2 (Work, Non-work)
 Decision-Making Unit: Tour
 Model Form: Multinomial logit
 Alternatives: MGRAs

The trips are then split among toll and non-toll paths according to a simplified toll choice model. The toll choice model included the following explanatory variables:

- In-vehicle-time
- Toll cost

Internal-External (IE) Trips

IE Trip Generation Model

Number of Models: 2 (Work, Non-work)
 Decision-Making Unit: Person
 Model Form: Binary logit
 Alternatives: 2 (Made a IE trip or not)

The internal-external trip generation model covers the SD-US and SD-MX trips.

The IE trip generation model explanatory variables are:

- Household income
- Vehicle ownership
- Age
- Accessibility to external stations which is defined by:

$$Access_h = \sum_z IePct_z \times \exp(-0.047 \times Dist_{h,z})$$

where h is the home zone, z ranges over external zones, $IePct_z$ is the percent of base-year IE trips that used the external station at zone z , and $Dist_{h,z}$ is the distance between the home and the external zone in miles.

IE Destination Choice Model

Number of Models: 1
Decision-Making Unit: Trip
Model Form: Multinomial logit
Alternatives: MGRAs

The IE trips are distributed to external stations with a destination choice model. The explanatory variables of the IE destination choice model are:

- Distance
- Size variable equal to the percent of IE trips using the external zone in the base year

IE Mode Choice Model

Number of Models: 1
Decision-Making Unit: Trip
Model Form: Multinomial logit
Alternatives: Trip Modes

After choosing an external station, the IE trip-maker chooses a mode according to an asserted nested logit mode choice model. The explanatory variables in the trip mode choice model are:

- Household income
- Gender
- In-vehicle time (auto and transit)
- Walk time
- Bike time
- Auto operating cost
- Auto Parking cost
- Auto toll value
- Transit first wait time
- Transit transfer time
- Number of transit transfers
- Transit walk access time
- Transit walk egress time
- Transit walk auxiliary time
- Transit fare
- Drive access to transit in-vehicle time
- Drive access to transit cost

Endnotes

- ¹ Note that although it is possible to include race/ethnicity as an explanatory variable in the model system, it has been decided to not include such variables in the San Diego implementation.
- ² Full-time employment is defined in the SANDAG 2006 household survey as at least 30 hours/week. Part-time is less than 30 hours/week but works on a regular basis.
- ³ Destination Lindbergh: The Ultimate Build-Out Aviation Activity Forecast, Jacobs Consultancy Team, February 2009, page 5-39 (Table 5.4-1).

SERIES 13 DEMOGRAPHIC AND ECONOMIC FORECASTING MODEL (DEFM)

**(DEFM Version 2013)
TECHNICAL DOCUMENTATION**



401 B Street, Suite 800
San Diego, CA 92101
(619) 699-1900 www.sandag.org

February 2013

MEMBER AGENCIES:

Cities of Carlsbad, Chula Vista, Coronado, Del Mar, El Cajon, Encinitas, Escondido, Imperial Beach, La Mesa, Lemon Grove, National City, Oceanside, Poway, San Diego, San Marcos, Santee, Solana Beach, Vista, and County of San Diego.

ADVISORY/LIAISON MEMBERS:

California Department of Transportation, U.S. Department of Defense, San Diego Unified Port District, San Diego County Water Authority, and Tijuana/Baja California/Mexico

ABSTRACT

TITLE: Demographic and Economic Forecasting Model (DEFM Version 2013) - Volume 2 - Technical Description

AUTHOR: San Diego Association of Governments

DATE: February 2013

SOURCE OF COPIES: San Diego Association of Governments
401 B Street, Suite 800
San Diego, CA 92101
(619) 699-1900

NUMBER OF PAGES: 62

ABSTRACT: This document discusses DEFM Version 2013.

TABLE OF CONTENTS

1	GENERAL DESCRIPTION.....	1
1.1	Introduction.....	1
1.2	Model Structure	1
1.3	Exogenous Variables	3
1.4	Units of Measurement.....	4
1.5	Binary Variables	4
1.6	Shift Variables	5
2	DEMOGRAPHICS.....	7
2.1	Introduction.....	7
2.2	Population Components.....	7
2.3	Components of Change.....	8
2.4	Labor Force.....	11
2.5	School Enrollment	11
2.6	Group Quarters Population	11
2.7	Household Population and Total Households.....	11
2.8	Control Values	12
2.9	Equations.....	12
3	INCOME.....	21
3.1	Income Equations.....	21
3.2	Income Equations.....	22
4	ECONOMIC ACTIVITY.....	27
4.1	Introduction.....	27
4.2	Industry Definitions	27
4.3	General Approach.....	30
4.4	Market Indexes.....	30
4.5	Employment Equations.....	34
4.6	Employment Aggregates.....	35
4.7	Output Equations	36
4.8	Output Aggregates	37
4.9	Unemployment Rate	38
4.10	Gross Regional Product	39
4.11	Orange County Gross Regional Product.....	39

5	CONSTRUCTION.....	41
5.1	Model Logic.....	41
5.2	Construction Equations.....	43
6	PUBLIC FINANCE	47
6.1	Revenue.....	47
6.2	Expenditures	48
6.3	Revenue Model Equations	49
7	PRICES.....	51
7.1	Discussion of Equations.....	51
7.2	Price Equations	52

LIST OF FIGURES AND TABLES

Figure 1-1	DEFM Sector Interactions	2
Figure 2-1	Population Components	8
Figure 2-2	Military Population	9
Figure 2-3	Change in Adjusted Civilian Population.....	10
Figure 5-1	Construction Sector Logic	42
Table 4-1	Industry Definitions	27

1 General Description

1.1 Introduction

This document describes the Demographic and Economic Forecasting Model (DEFM), in specific terms including the flows of information, equations, and theoretical assumptions. DEFM is a simultaneous, nonlinear econometric model designed to forecast population and economic variables for the San Diego region from a set of basic assumptions. The model produces medium to long-term forecasts.

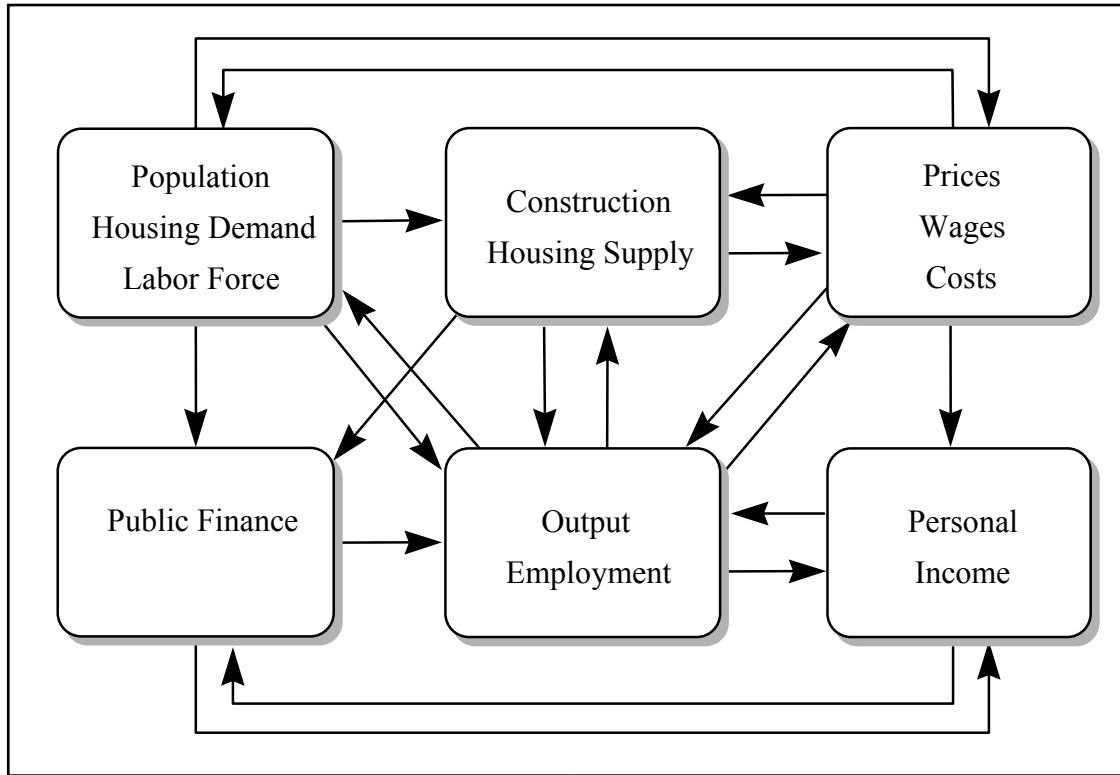
Historical data for the period 1970-2012 form the basis for DEFM equations and parameters. Regression analysis of these time-series data is used to determine the relationship between the exogenous variables that drive the forecast and the dependent variables that are determined within the model. Additionally, DEFM makes extensive use of cross-sectional data, including the U.S. decennial census of population and local industry input/output relationships. Current forecasts include data for a wide range of economic and demographic variables for the period 2013 - 2050.

In general, the model is a synthesis of two widely used techniques. On the demographic side, the model uses a cohort component method. This method considers such factors as birth rates, death rates, and the age, sex and ethnic distribution of the resident population to arrive at forecasts of demographic variables. On the economic side, time-series/regression methods are used to estimate economic relationships. The resulting econometric equations provide forecasts of employment, income and other economic variables based on assumptions about national, state and local growth patterns and local inter-industry relationships.

1.2 Model Structure

DEFM has several major sub models, and the general flow of information is as illustrated in Figure 1-1. As is evident from the figure, the model is simultaneous in nature, and most major variables are interrelated directly or indirectly. For example, the level of economic activity, as measured by employment or output, depends in part on the size of the local population and income levels. Income, in turn depends in part on employment and labor market conditions. Over time, the population responds to economic conditions through impacts on net migration levels. Thus, local activity depends on local population, but local population also depends on activity. This type of circularity is a central feature of simultaneous models, and the DEFM structure is designed to capture the main interdependencies and interactions that exist in a dynamic economy.

Figure 1-1: DEFM Sector Interactions



Some of the key interactions are as follows.

- The survived resident population by age, gender and ethnicity provides the basis for computing housing demand and the associated public finance. It also contributes to demand for goods and services and provides the basis for computing the size of the labor force.
- Local employment is determined by local, state, and national economic conditions. For each industry, demands include intermediate demands by other local industries, final demands of households and government agencies, and export demands of the national and international economies. Output is a function of employment and labor productivity.
- Construction levels are determined by the demand for housing relative to the current stock of housing and national construction industry conditions.
- Local housing prices and wage levels are determined by conditions in the local housing and labor markets relative to comparable national markets. The local consumer price index follows national trends with modification for local housing and energy prices.

- Personal income depends on employment, local wage rates, and estimates of income from other sources, including dividends, interest, and rental income as well as a variety of government transfer payments.
- Public finance covers government revenues from sales taxes, property taxes, state grants, and other sources. On the expenditure side, outlays are forecasted for payroll costs, educational expenses, and other spending programs.

As shown in Figure 1, the linkages between the demographic and economic sectors are numerous, and a full understanding of these interactions requires close study of the explicit model linkages. A simultaneous solution algorithm reconciles the relationships among population, income, and economic activity.

Output, employment, and labor productivity levels are forecast for 50 selected private-sector industries, identified by specific three-digit and four-digit North American Industrial Classification System (NAICS) codes. The industries are selected to represent the most important activities in San Diego and to allow aggregation of these detailed activities into a set of export-based economic clusters that have been defined for the county.

Most of the employment equations have a common form that specifies employment as a linear function of a composite market index that depends on a comprehensive set of explanatory variables. The demand component of the market index depends on employment levels in other local industries, indicators of household and government final demand, and indicators of national and international export demand for local goods. The supply component of the market index depends on local costs and productivity levels, providing a measure of the relative competitiveness of San Diego businesses. The market indexes are constructed using transaction weights and regional purchase coefficients from a San Diego input/output table constructed specifically for the 50 selected industries. The final importance of the each market index is determined by econometric estimation for each industry.

1.3 Exogenous Variables

The exogenous assumptions in DEFM influence the forecasts and can be controlled by users. Users may alter input data, demographic parameters and economic coefficients, or may introduce other changes (shifts) to alter the model outcome. A variety of shifts have been included in the model, and additional shifts can be added as needed by editing the text of the model equations. Since these equations are input to the model, these changes to the model can be made without any need to recompile the simulation code itself.

The presentation of each sector begins with a brief discussion of the main variables determined in that sector, followed by the model functional form. With each set of equations, the model mnemonics are presented and units of measurement are provided to facilitate coefficient interpretation.

1.4 Units of Measurement

The following abbreviations are used for units of measurement:

B	billions
F	fraction or ratio
M	millions
T	thousands
YRS	years of age
\$	dollars
\$2010	real dollars with a 2010 base
%	percent
A	actual count

In addition to definitions of measurement units, each variable is further identified, by code, as to type. The types are defined as follows:

Code	Variable Type
I	Endogenous variable determined within the sector being discussed (Inside Sector)
O	Endogenous variable determined outside the current sector (Outside Sector)
X	Exogenous forecast variable determined outside of the model
A_	Additive user-defined shift variable
M_	Multiplicative user-defined shift variable
P	User-defined demographic parameter

The model sectors and forecast equations are discussed in detail in the following sections. In general, each of the equations relates a dependent variable (Y) to a linear or nonlinear combination of independent variables (X's).

1.5 Binary Variables

Often in estimated equations it is desirable to include shift factors during the historical estimation period. This is accomplished by inclusion of binary variables as explanatory factors. Typically, these variables are zero up to some point and are one thereafter. This type of variable can be used to account for changes in definitions, to account for events that occur only in specific years, or to account for apparent structural shifts in economic relationships.

In DEFM, this type of variable is named to suggest the timing of the variable. For example:

B00 indicates a variable that is 1.0 in 2000 and zero otherwise.

B00Plus indicates a variable that is zero through 1999 and 1.0 in 2000 and later years.

B00to02 indicates a variable that is 1.0 between 2000 and 2002 and zero otherwise.

1.6 Shift Variables

As indicated in the table above, user-controlled shift variables are indicated by an initial prefix of A_ or M_. These variables are treated specially in DEFM. Normally, if a variable appears in an equation and that variable is not defined in a Data Table of Transformation Table, then the equation will evaluate to “missing” and this will propagate through the model to any other variables that are subsequently dependent. This is not the case with shift variables.

For additive shift variables, beginning with A_, if the variable is not defined the value will default to zero in all evaluations. These shifts should be included in expressions in an additive format. Then, leaving the shifts undefined or including a link to a string of zero values will have the same numerical consequences in the model.

For multiplicative shift variables, beginning with M_, if the variable is not defined the value will default to 1.0 in all evaluations. These shifts should be included in expressions in a multiplicative format.

In this way, it is only necessary to link to actual data values for the specific shifts that are active in a run. All shifts that are not active can be left unspecified, implying the default results discussed above. This allows inclusion of hundreds of shift variables in the model that are ready for the user to invoke without carrying a large database that would contain strings of zeros and ones for the majority of the shift items.

Note: Shift variables are not included in the general equation specifications listed below. Consult the actual model code for the exact equation specification used in the forecast.

2 Demographics

2.1 Introduction

Population is forecast for single-year categories by gender and ethnicity for civilian population and total population. (Uniformed military population and military dependents and the associated age, gender, and ethnic group distributions are exogenous data.) The eight ethnic groups are defined as: (1) Hispanic, (2) Non-Hispanic White, (3) Non-Hispanic Black, (4) Non-Hispanic American Indian, (5) Non-Hispanic Asian, (6) Non-Hispanic Hawaiian and Pacific Islander, (7) Non-Hispanic Other and (8) Non-Hispanic Two or More races.

Summary data for each ethnic group include total population, civilian population, births, deaths, total fertility rate, life expectancy, net migration, and population for each of several age group aggregates (0-4, 5-9, ... 85+). Age, gender, and ethnic details are also forecast for labor force and for school enrollment by five levels (nursery school, kindergarten, elementary, high school, and college). Detailed forecasts are also developed for group quarters population, household population, and heads of household.

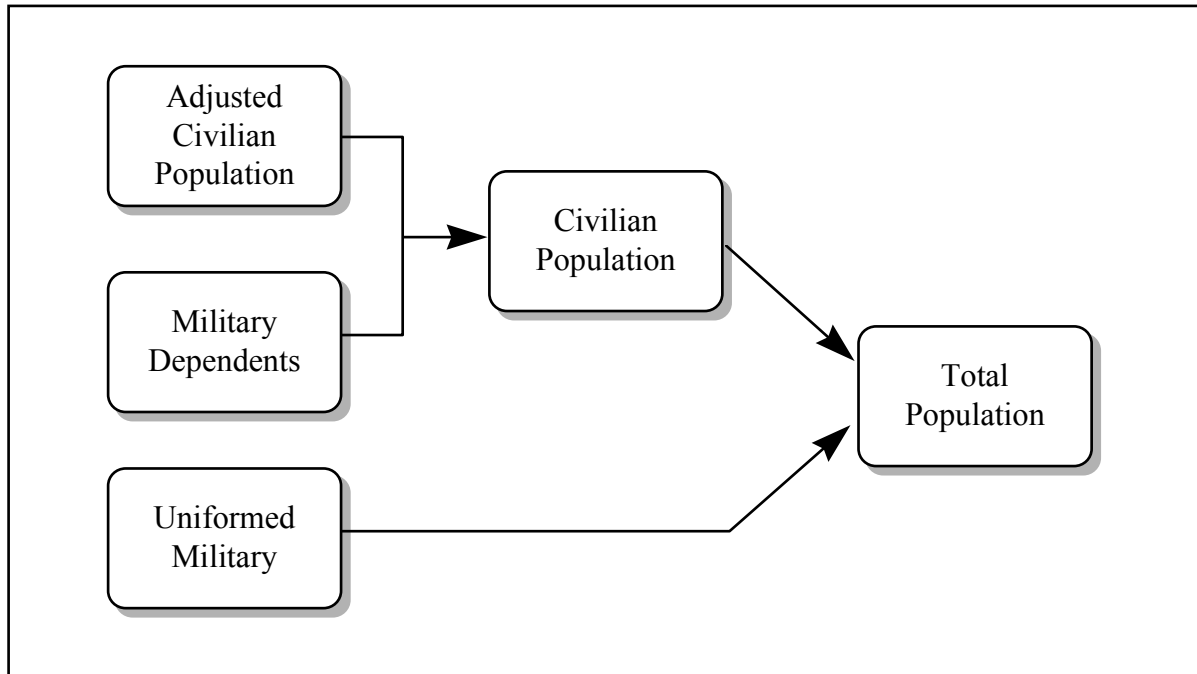
2.2 Population Components

The demographic forecasts are based on age, gender and ethnic detail of the population in the most recent decennial census year (2010). Initially, the census distribution is applied to the April 2010 population to develop starting values for each cohort. The civilian population is determined by subtracting number of military in uniform from the total population. The adjusted civilian population is determined by subtracting the military dependent population from the civilian population. This process results in the following definitions of population components.

- **Total Population.** The total population includes all residents including military in uniform and military dependents.
- **Military Population.** The military population is the sum of military in uniform and military dependents.
- **Civilian Population.** The civilian population is equal to the total population less military in uniform. As a result, the civilian population includes military dependents.
- **Adjusted Civilian Population.** The adjusted civilian population is equal to the total population less military in uniform and their dependents. Thus, adjusted civilian population plus military dependents equals the civilian population.

These relationships are depicted in Figure 2-1.

Figure 2-1: Population Components

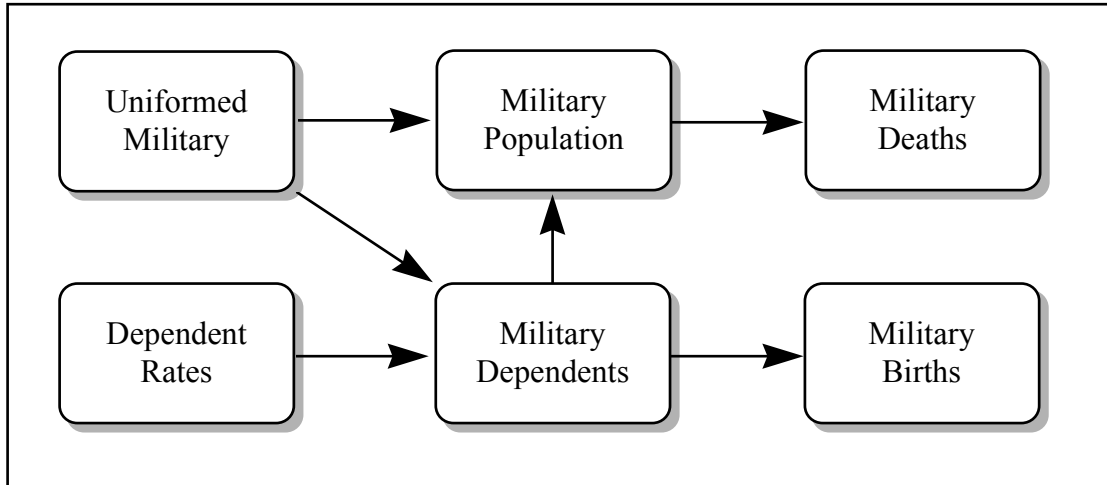


2.3 Components of Change

Military Population. As depicted in Figure 2-2, a top-down approach is used to compute the military population each year. In the upper-left of the diagram, the total number of uniformed military is an exogenous input. This total is allocated to age, gender, and ethnic groups based on a set of allocation factors. The military dependent population for each cohort is estimated as a proportion of the uniformed military population. Uniformed military and military dependents are summed to obtain total military population by age, gender, and ethnic group.

In each year, births to military dependents and deaths for the military population are computed and these values contribute to the county totals for births and deaths. However, reflecting the top-down approach, the military population is not “survived” from one year to the next. This is appropriate given the high levels of geographic mobility among uniformed personnel and their dependents, and mobility between military and civilian status. Finally, during the historical period, births to military families in each ethnic group are compared to control values based on vital statistics for the county. Calibration factors are computed as the ratio of the control value to the computed value. In the forecast period, these calibration factors are extended based on the value in the final calibration period.

Figure 2-2: Military Population



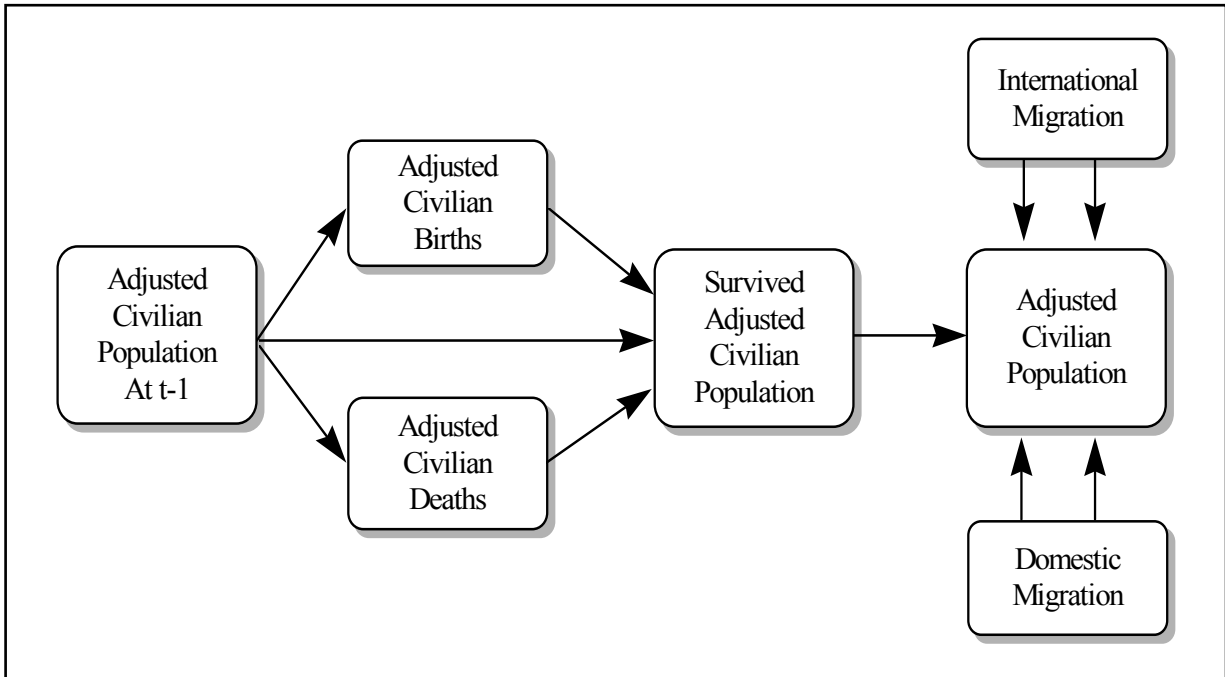
Adjusted Civilian Population. The adjusted civilian population is forecast using a modified cohort-survival method as illustrated in Figure 2-3. In each year, the civilian population is the civilian population in the previous year less civilian deaths, plus births, plus net civilian migration. The detailed survival and fertility computations are performed for population subgroups by age, gender and ethnic group. Summing across ethnic groups yields total population by age and gender.

Births. In each forecast year, natural increase is equal to births minus deaths. Adjusted civilian births (excluding births to military families) are forecast from last year's female population, using exogenous age-specific birth rates for women in each ethnic group. The births are split into male and female components using historical ratios estimated from vital statistics data for San Diego County.

During the historical period, births for each ethnic group are summed and compared with control total values for adjusted civilian births in the county. The ratio of the control value to the computed value is the calibration factor for that year. In the forecast period, these calibration factors are extended based on the value in the final calibration period.

Deaths. Deaths in a year are forecast from the civilian and military population in the preceding year, using exogenous survival rates by age, gender, and ethnicity. During the historical period, deaths for each ethnic group are summed and compared with control total values for civilian births in the county. The ratio of the control value to the computed value is the calibration factor for that year. In the forecast period, these calibration factors are extended based on the value in the final calibration period.

Figure 2-3: Change in Adjusted Civilian Population



Survived Population. Deaths are subtracted from the previous year's adjusted civilian population. Survived adjusted civilian population is computed by moving the survived population in each age group into the next age group. Births to the adjusted civilian population are adjusted for infant survival rates and are placed in the first age group (under 1) to complete the forecast of survived adjusted civilian population.

Migration. Net migration is modeled in two pieces. International net migration is modeled as a share of total US immigration. Domestic net migration is related to economic conditions in San Diego relative to the US and to recent employment gains. Economic conditions include the relative employment rate and the relative wage rate. The equation includes a first-order autoregressive term. The small estimated parameter on this term implies that the regression residual progresses to zero quickly in the forecast period.

Migration components are further divided by age, gender and ethnicity. First, the estimates of net migration are distributed to the eight ethnic/gender groups. Then age-specific migration rates for each ethnic group and gender are used to estimate the detailed age distribution within each group. The distribution factors and age-specific rates are exogenous to the model.

2.4 Labor Force

The natural civilian labor force is calculated from the civilian population by multiplying civilian population in each age, gender, and ethnic group by a natural labor force participation rate. The natural labor force participation rates are defined to represent the level of participation that occurs at full employment (defined by an unemployment rate of about 5%).

Actual labor force participation rates will be lower during periods of high unemployment (reflecting the fact that some potential workers are discouraged by the lack of job opportunities and stop actively looking for work). Similarly, under periods of low unemployment, labor force participation rates can rise above the natural rates, as additional people are drawn into the workforce. These variations are modeled based on a set of elasticity values for labor force participation with respect to employment rates, where employment rates reflect the level of employment relative to the natural labor force.

2.5 School Enrollment

Forecasts of school enrollment levels are based on the civilian population by age, gender, and ethnic group. Five school levels are covered: nursery school, elementary school, high school, and college. School participation rates are exogenous to the model. During the historical period, computed totals for each category of school enrollments are compared with control values for these totals. The ratio of the control value to the computed value is the calibration factor for that year. In the forecast period, these calibration factors should be extended based on the value in the final calibration period.

2.6 Group Quarters Population

The population living in group quarters is forecast by age, gender, and ethnic group for uniformed military living in barracks or onboard ship, college students living in dormitories, and for other persons in group quarters accommodations. Other group quarters population includes persons living in boarding houses, homes for the disabled, rest homes, jails and other group living situations. The age and gender specific rates for group quarters population are held constant throughout the forecast period. During the historical period, computed totals for each category of group quarters are compared with control values for these totals. The ratio of the control value to the computed value is the calibration factor for that year. In the forecast period, these calibration factors should be extended based on the value in the final calibration period.

2.7 Household Population and Total Households

Household population is calculated by subtracting the total group quarters population by age, gender, and ethnicity from the total population. Household headship rates, which give the fraction of population that is a household head, are applied to the household population by

age, gender and ethnicity to determine the number of households. Headship-rate parameters are exogenous to the model. During the forecast, these parameters are modified through a set of elasticity values that are applied to changes in the price of housing relative to income per household. These elasticity values are set to relatively small values, based on an analysis of the historical relationship between household size and housing prices.

During the historical period, computed totals for the number of households are compared with control values for these totals. The ratio of the control value to the computed value is the calibration factor for that year. In the forecast period, these calibration factors should be extended based on the value in the final calibration period.

2.8 Control Values

For each year between the census (2010) and the first forecast year (2013), detailed components of the population are calculated and controlled to totals contained in the input data base. The control values are used to compute calibration factors during the historical period, and forecasted values for these calibration factors are used beyond the historical period. Calibration factors are computed for the following variables.

- Civilian births by ethnicity
- Military births by ethnicity
- Deaths by ethnicity
- Total school enrollments by education level
- Total group quarters population by type of group quarters
- Total labor force
- Total households.

2.9 Equations

The equations below illustrate the computation of demographic components and characteristics. In some cases, the actual mathematical representation of the equation is listed. This is true for identities and for algebraic expressions for which the coefficients do not change. For other equations, the expression $f(,)$ is used to represent the general functional form. The following additional notation is used.

Subscript	Description	Range	Explanation
a	Age	0 to 98, 99+	Annual age groups to 99+
g	Gender	0 or 1	Male, Female
e	Ethnicity	1 to 8	Hispanic, et al
t	Time	NA	t = current year; t-1 = past year t-2 = two year lag, etc.

Military Population

$$MUNIF^t = \text{Exogenous}$$

$$MUNIF_{a,g,e}^t = MUNIF^t \times PMUNIF_{a,g,e}^t$$

$$MILDEP_{a,g,e}^t = MUNIF^t \times PMDEP_{a,g,e}^t$$

$$MILPOP_{a,g,e}^t = MUNIF_{a,g,e}^t + MILDEP_{a,g,e}^t$$

$$MILDEP^t = \sum_a \sum_g \sum_e MILDEP_{a,g,e}^t$$

$$MILPOP^t = \sum_a \sum_g \sum_e MILPOP_{a,g,e}^t$$

Mnemonic	Description	Units	Type
MILDEP	Military Dependents	A	I
MILPOP	Military Population	A	I
MUNIF	Military in Uniform	A	I
PMDEP	Military Dependent Distribution Factor	F	P
PMUNIF	Uniformed Military Distribution Factor	F	P

Natural Increase

$$\text{MILB}_{a,e}^t = k \text{MILB}_c^t \times \text{PFR}_{a,e}^t \times \text{MILDEP}_{a,1,e}^t$$

$$\text{MILB}_{a,0,e}^t = .512 \times \text{MILB}_{a,e}^t$$

$$\text{MILB}_{a,1,e}^t = .488 \times \text{MILB}_{a,e}^t$$

$$\text{MILD}_{a,g,e}^t = kD_e^t \times \text{MILPOP}_{a,g,e}^t \times (1 - \text{SURV}_{a,g,e}^t)$$

$$\text{ADJCIVB}_{a,e}^t = k \text{CIVB}_c^t \times \text{PFR}_{a,e}^t \times \text{ADJCIVPOP}_{a,1,e}^{t-1}$$

$$\text{ADJCIVB}_{a,0,e}^t = .512 \times \text{ADJCIVB}_{a,e}^t$$

$$\text{ADJCIVB}_{a,1,e}^t = .488 \times \text{ADJCIVB}_{a,e}^t$$

$$\text{ADJCIVD}_{a,g,e}^t = kD_e^t \times \text{ADJCIVPOP}_{a,g,e}^{t-1} \times (1 - \text{SURV}_{a,g,e}^t)$$

$$\text{TOTB}^t = \sum_a \sum_g \sum_e (\text{ADJCIVB}_{a,g,e}^t + \text{MILB}_{a,g,e}^t)$$

$$\begin{aligned} \text{TOTD}^t &= \sum_a \sum_g \sum_e (\text{ADJCIVD}_{a,g,e}^t + \text{MILD}_{a,g,e}^t) \\ &+ \sum_a \sum_g \sum_e (\text{ADJCIVB}_{a,g,e}^t + \text{MILB}_{g,e}^t) \times (1 - \text{INFSURV}_{g,e}^t) \end{aligned}$$

$$\text{NI}^t = \text{TOTB}^t - \text{TOTD}^t$$

$$\text{ADJCIVNI}^t = \text{ADJCIVB}^t - \text{ADJCIVD}^t$$

Mnemonic	Description	Units	Type
ADJCIVB	Births to Adjusted Civilian Population	A	I
ADJCIVD	Deaths of Adjusted Civilian Population	A	I
ADJCIVPOP	Adjusted Civilian Population	A	I
ADJCIVNI	Natural Increase of Adjusted Civilian Population	A	I
kCIVB	Calibration Constant for Civilian Births	F	P
kMILB	Calibration Constant for Military Births	F	P
kD	Calibration Constant for Deaths	F	P
MILB	Military Births	A	I
MILD	Military Deaths	A	I
MILPOP	Military Population	A	I
INFSURV	Infant survival rate	F	P
NI	Natural increase (Resident Births - Deaths)	A	I
PFR	Period Fertility Rates	F	P
SURV	Survival Rates	F	P
TOTB	Total Births	A	I
TOTD	Total Deaths	A	I

Migration

$$DMIG^t = f\left(\left(ENAG^{t-1} - ENAG^{t-3}\right), \frac{U}{UUS}, \frac{PH}{PHUS}\right)$$

$$DMIG_{g,e}^t = DMIG^t \times DMIGFRAC_{g,e}^t$$

$$DMIG_{a,g,e}^t = DMIG_{g,e}^t \times \frac{DMIGRATE_{a,g,e}^t \times ADJCIVPOP_{a,g,e}^{t-1}}{\sum_a DMIGRATE_{a',g,e}^t \times ADJCIVPOP_{a',g,e}^{t-1}}$$

$$INTMIG^t = f(INTMIGUS^t, SDPOP^t, USPOP^t)$$

$$INTMIG_{a,g,e}^t = INTMIG^t \times INTMIGFRAC_{a,g,e}^t$$

Mnemonic	Description	Units	Type
ENAG	Nonagricultural Employment	A	O
ENAGUS	US Nonagricultural Employment	M	X
U	Unemployment Rate	%	O
UUS	US Unemployment Rat	%	X
PH	Price of Housing	T	O
PHUS	US Price of Housing	T	X
DMIG	Domestic Migration	A	I

DMIGFRAC	Domestic Migration Distribution Factors by g, e	F	P
DMIGRATE	Domestic Migration Rate by a, g, e	F	P
INTMIG	International Migration	A	I
INTMIGUS	U.S. International Net Migration	M	X
MA2	Two year moving average		
INTMIGFRAC	International Migration Distrib Factor by a, g, e	F	X
P	Population	A	O
PUS	US Population	M	X
R_W	Real Wage	\$2010	O
R_WUS	US Real Wage	\$2010	X
SDPOP	San Diego Population	A	I
USPOP	US Population	M	X

Civilian and Total Population

$$ADJCIVSURV_{a,g,e}^t = ADJCIVPOP_{a-1,g,e}^{t-1} - ADJCIVD_{a-1,g,e}^t \quad a = 1-98$$

$$ADJCIVSURV_{99,g,e}^t = ADJCIVPOP_{98,g,e}^{t-1} - ADJCIVD_{98,g,e}^t + ADJCIVPOP_{99,g,e}^{t-1} - ADJCIVD_{99,g,e}^t \quad a = 99$$

$$ADJCIVSURV_{0,g,e}^t = ADJCIVB_{g,e}^t \times INFSURV_{g,e}^t \quad a = 0$$

$$ADJCIVPOP_{a,g,e}^t = ADJCIVSURV_{a,g,e}^t + DMIG_{a,g,e}^t + INTMIG_{a,g,e}^t$$

$$ADJCIVPOP^t = \sum_a \sum_g \sum_c ADJCIVPOP_{a,g,e}^t$$

$$CIVPOP_{a,g,e}^t = ADJCIVPOP_{a,g,e}^t + MILDEP_{a,g,e}^t$$

$$CIVPOP^t = \sum_a \sum_g \sum_e CIVPOP_{a,g,e}^t$$

$$TOTPOP_{a,g,e}^t = CIVPOP_{a,g,e}^t + MUNIF_{a,g,e}^t$$

$$TOTPOP^t = \sum_a \sum_g \sum_e TOTPOP_{a,g,e}^t$$

Mnemonic	Description	Units	Type
ADJCIVB	Births to Adjusted Civilian Population	A	I
ADJCIVD	Deaths of Adjusted Civilian Population	A	I
ADJCIVPOP	Adjusted Civilian Population	A	I
ADJCIVSURV	Adjusted Civilian Survived Population	A	I
CIVPOP	Civilian Population	A	I
DMIG	Domestic Migration	A	I
INTMIG	International Migration	A	I
MILDEP	Military Dependents	A	I
MUNIF	Military in Uniform	A	I
INFSURV	Infant survival rate	F	P
RMIG	Net Retirement Migration	A	I
TOTPOP	Total Population	A	I

Labor Force

$$NLF_{a,g,e}^t = LFPR_{a,g,e}^t \times CIVPOP_{a,g,e}^t$$

$$L_{a,g,e}^t = kLAB^t \times NLF_{a,g,e}^t \times \left(\frac{MA2(E^t/NLF^t)}{(E^{90}/NLF^{90})} \right)^{\alpha_{a,g,e}}$$

$$L^t = \sum \sum \sum L_{a,g,e}^t$$

Mnemonic	Description	Units	Type
CIVPOP	Civilian Population	A	I
kLAB	Calibration factor for Labor Force	F	X
L	Civilian Labor Force	A	I
LFPR	Labor Force Participation Rate	F	P
NLF	Natural (full employment) Labor Force	A	I
E	Resident Civilian Employment	A	X
MA2	Two Year Moving Average	na	Function
α	Labor Force Participation Elasticity	F	P

School Enrollment

$$ENRLNURS_{a,g,e}^t = kSCHNURS^t \times SPRN_{a,g,e}^t \times CIVPOP_{a,g,e}^t$$

School Enrollment (cont'd)

$$ENRLKIND_{a,g,e}^t = kSCHKIND^t \times SPRK_{a,g,e}^t \times CIVPOP_{a,g,e}^t$$

$$ENRLELEM_{a,g,e}^t = kSCH ELEM^t \times SPRE_{a,g,e}^t \times CIVPOP_{a,g,e}^t$$

$$ENRLHIGH_{a,g,e}^t = kSCHHIGH^t \times SPRH_{a,g,e}^t \times CIVPOP_{a,g,e}^t$$

$$ENRLCOL_{a,g,e}^t = kSCHCOL^t \times SPRC_{a,g,e}^t \times CIVPOP_{a,g,e}^t$$

$$ENRLNURS^t = \sum_a \sum_g \sum_e ENRLNURS_{a,g,e}^t$$

$$ENRLKIND^t = \sum_a \sum_g \sum_e ENRLKIND_{a,g,e}^t$$

$$ENRLELEM^t = \sum_a \sum_g \sum_e ENRLELEM_{a,g,e}^t$$

$$ENRLHIGH^t = \sum_a \sum_g \sum_e ENRLHIGH_{a,g,e}^t$$

$$ENRLCOL^t = \sum_a \sum_g \sum_e ENRLCOL_{a,g,e}^t$$

$$ENRLTOT^t = ENRLKIND^t + ENRLELEM^t + ENRLHIGH^t$$

$$ADA^t = f(ENRLTOT^t)$$

Mnemonic	Description	Units	Type
ADA	Average Daily Attendance – K to Junior College	A	I
CIVPOP	Civilian Population	A	I
ENRL	Enrollments (NURS, KIND, ELEM, HIGH, COL)	A	I
ENRLTOT	Total School Enrollments	A	I
kSCH	Adjustment Factors for School Enrollment (NURS, KIND, ELEM, HIGH, COL)	F	I
SPR	School Participation Rates (N, K, E, H, C)	F	P

Group-Quarters

$$GQM_{a,g,e}^t = PGQM_{a,g,e}^t \times MUNIF_{a,g,e}^t$$

Group-Quarters (cont'd)

$$GQS_{a,g,e}^t = PGQS_{a,g,e}^t \times ENRLCOL_{a,g,e}^t$$

$$GQO_{a,g,e}^t = PGQO_{a,g,e}^t \times CIVPOP_{a,g,e}^t$$

$$GQT_{a,g,e}^t = GQM_{a,g,e}^t + GQS_{a,g,e}^t + GQO_{a,g,e}^t$$

$$GQM^t = kGQMIL \times \sum_a \sum_g \sum_e GQM_{a,g,e}^t$$

$$GQS^t = kGQCOL^t \times \sum_a \sum_g \sum_e GQS_{a,g,e}^t$$

Mnemonic	Description	Units	Type
CIVPOP	Civilian Population	A	I
ENRLCOL	College Enrollment	A	X
GQM	Military Group Quarters Population	A	I
GQO	Other Group Quarters Population	A	I
GQS	Student (College) Group Quarters Population	A	I
GQT	Total Group Quarters Population	A	I
kGQMIL	Calibration Factor for Military Group Quarters	F	I
kGQCOL	Calibration Factor for Student Group Quarters	F	I
MUNIF	Military In-Uniform Population	A	X
PGQM	Military Group-Quarters Fractions	F	P
PGQO	Other Group Quarters Fractions	F	P
PGQS	College Dorm Fractions	F	P

Household Population and Households

$$HHP_{a,g,e}^t = POP_{a,g,e}^t - GQM_{a,g,e}^t - GQS_{a,g,e}^t - GQO_{a,g,e}^t$$

$$HHP^t = \sum_a \sum_g \sum_e HHP_{a,g,e}^t$$

$$HHCALC_{a,g,e}^t = HHRATE_{a,g,e}^t \times HHP_{a,g,e}^t$$

$$HH_{a,g,e}^t = kHH^t \times HHCALC_{a,g,e}^t \times \left(MA2 \left(\frac{PH^t}{Y^t / HH^t} \right) / \left(\frac{PH^{2000}}{Y^{2000} / HH^{2000}} \right) \right)^{\alpha_{a,g,e}}$$

$$HH^t = \sum_a \sum_g \sum_e HH_{a,g,e}^t$$

Mnemonic	Description	Units	Type
POP	Total Population	A	I
GQM	Military Group Quarters Population	A	I
GQS	Student (College) Group Quarters Population	A	I
GQO	Total Group Quarters Population	A	I
HH	Households	A	I
HHP	Household Population	A	I
kHH	Calibration Factor for Households	A	I
HHCALC	Households with Exogenous Headship Rates	A	I
HHRATE	Household Headship Rates	F	P
MA2	Two-Year Moving Average	F	P
PH	Price of Housing	T	O
Y	Personal Income	M	O
α	Household Headship Elasticity values	F	P

3 Income

Personal income forecasts for San Diego County residents are developed for seven separate income sources. These sources are as follows:

- payroll (wage and salary income),
- other labor income,
- proprietors' income,
- dividends, interest and rent,
- transfer payment sources,
- net of social security payments
- adjusted for place of residence.

3.1 Income Equations

The income equations provide estimated relationships between components of income and a variety of explanatory variables. Civilian payroll depends on forecasts of employment, wages, and the industry mix of employment. The industry mix variable is calculated using 2002 payroll per employee values for the 50 private-sector employment categories. In each year, these values are weighted by the fraction of employment in each industry, and the resulting weighted average is divided by the 2002 weighted average.

Forecasts of military payroll depend on the number of uniformed military and general wage trends in the private sector. Other labor income is assumed to depend on civilian payroll levels and US ratios of other labor income to payroll. Similarly, proprietor's income is assumed to depend on civilian payroll and US ratios of proprietor's income to payroll. Dividends, interest and rental income depend upon regional population and national income from the same sources.

Transfer income is forecast for several components. Real retirement income is estimated to depend on the population over 65 years of age. Real unemployment transfer income depends on the number of unemployed persons in the economy. Real public assistance transfer income and other transfer income are modeled to depend on total population.

The income components are combined to obtain regional personal income, after subtracting Social Security contributions. Social Security contributions are assumed to depend on the level of civilian payroll. A residence adjustment is made to account for the income of local

residents who work outside the area and for the income of people who work locally but live outside the region.

Disposable personal income is calculated by subtracting forecasts of personal tax payments from total personal income. Taxes paid out of personal income include federal income taxes, state income taxes and personal property taxes. Because the property tax variable includes both payments by households and businesses, only two thirds of the total is subtracted in computing disposable income. This fraction represents the part of total property tax payments paid by households against residential properties.

The income equations are estimated in either real (constant dollars) or nominal (current dollars) form. When an equation is estimated in nominal terms, a second equation is added to compute the real value, using the CPI based in 2010 as a deflator. When an equation is estimated in real terms, an additional equation is used to compute the nominal value, using the CPI based in 2010 as an inflator. The real variables are denoted by a prefix of R_.

3.2 Income Equations

Payroll

$$CPAY^t = f(PUS^t, L^t, QWandS, EMPMIX^t)$$

$$MPAY^t = f(W^t, MUNIF^t)^t$$

$$R_CPAY^t = CPAY^t / CPI2000^t$$

$$R_MPAY^t = MPAY^t / CPI2000^t$$

Mnemonic	Description	Units	Type
CPAY	Civilian Payroll	\$M	I
CPI2010	Consumer Price Index (2010 = 1.00)	F	O
PUS	US Payroll	\$M	X
L	Labor Force	A	O
QWandS	Output for Wage and Salary Jobs	\$M	O
EMPMIX	Employment Mix Index (2010 = 1.00)	F	X
MPAY	Military Payroll	\$M	I
MUNIF	Military In-Uniform Population	A	X
R_CPAY	Real Civilian Payroll	\$M2010	I
R_MPAY	Real Military Payroll	\$M2010	I

Other Earned Income

$$OLY^t = f(CPAY^t, OLYUS^t/PAYUS^t, SEDW^t)$$

$$PROPY^t = f(CPAY^t, PROPYUS^t/PAYUS^t)$$

$$DIR^t = f(DIRUS^t, P^t/PUS^t)$$

$$R_OLY^t = OLY^t / CPI2000^t$$

$$R_PROPY^t = PROPY^t / CPI2000^t$$

$$R_DIR^t = DIR^t / CPI2000^t$$

Mnemonic	Description	Units	Type
CPAY	Civilian Payroll	\$M	I
CPI2010	Price Index in 2010\$	F	O
DIR	Dividend, interest, and rent income	\$M	I
DIRUS	US Dividend, interest, and rent income	\$M	I
OLY	Other Labor Income	\$M	I
OLYUS	US Other Labor Income	\$M	I
PAYUS	US Private and Government Payroll	B\$	X
PROPY	Proprietors' Income	\$M	I
PROPYUS	US Proprietors' Income	\$M	I
P	Population	A	O
PUS	US Population	M	X
R_DIR	Real dividend, interest, and rent income	\$M2010	X
R_PROPY	Real proprietors income	\$M2010	X
R_OLY	Real other labor income	\$M2010	X
SEDW	Self Employed and Domestic	A	I

Transfer Income

$$R_RTRAN^t = f(P65^t)$$

$$R_UNTRAN^t = f(L^t, E^t)$$

$$R_PATRAN^t = f(P^t)$$

$$R_OTRAN^t = f(P^t)$$

$$RTRAN^t = R_RTRAN^t \times CPI2000^t$$

$$UNTRAN^t = R_UNTRAN^t \times CPI2000^t$$

$$PATRAN^t = R_PATRAN^t \times CPI2000^t$$

$$OTRAN^t = R_OTRAN^t \times CPI2000^t$$

$$TRAN^t = RTRAN^t + UNTRAN^t + PATRAN^t + OTRAN^t$$

Mnemonic	Description	Units	Type
E	Total Resident Employment	A	O
CPI2010	Price Index in 2010\$	F	O
L	Labor Force	A	O
OTRAN	Other Transfer Income	\$M	I
P	San Diego Population	A	O
P65	San Diego Population aged 65 and over	A	O
PATRAN	Public Assistance Transfers	\$M	I
RTRAN	Retirement and Disability Transfers	\$M	I
TRAN	Total Transfer Income	\$M	I
UNTRAN	Unemployment Transfers	\$M	I
R_OTRAN	Real Other Transfer Income	\$M2010	I
R_PATRAN	Real Public Assistance Transfers	\$M2010	I
R_RTRAN	Real Retirement and Disability Transfers	\$M2010	I
R_UNTRAN	Real Unemployment Transfers	\$M2010	I

Adjustments to Income

$$SSC^t = f\left(\left(\left(CPAY^t \times \frac{SSCUS^t}{PAYUS^t}\right)\right)^t\right)$$

$$RESADJ^t = f(CPAY^t)$$

Mnemonic	Description	Units	Type
CPAY	Civilian Payroll	\$M	I
PAYUS	US Civilian Payroll	B\$	X
RESADJ	Residence Adjustment	\$M	I
SSC	Social Security Contributions	\$M	I
SSCUS	US Social Security Contributions	B\$	X

Personal Income

$$Y^t = CPAY^t + MPAY^t + OLY^t + PROPY^t + DIR^t + TRAN^t - SSC^t + RADJ^t$$

Mnemonic	Description	Units	Type
CPAY	Civilian Payroll	\$M	I
DIR	Dividends, Interest and Rent Income	\$M	I
MPAY	Military Payroll	\$M	I
OLY	Other Labor Income	\$M	I
PROPY	Proprietors' Income	\$M	I
RADJ	Residence Adjustment	\$M	I
SSC	Social Security Contributions	\$M	I
TRAN	Total Transfer Income	\$M	I
Y	Personal Income	\$M	I

Personal Taxes and Disposable Personal Income

$$SYTAX^t = SYTAXR^t \times [Y^t - 0.67 \times RPTAX^t]$$

$$FYTAX^t = FYTAXR^t \times [Y^t - SYTAX^t - 0.67 \times RPTAX^t]$$

$$PERTAX^t = SYTAX^t + FYTAX^t + 0.67 \times RPTAX^t$$

$$YD^t = Y^t - PERTAX^t$$

Mnemonic	Description	Units	Type
FYTAX	Federal Income Tax Payments	\$M	I
FYTAXR	Federal Average Income Tax Rate	F	X
PERTAX	Personal Taxes (Income and Property)	\$M	I
RPTAX	Local Property Taxes	\$M	O
SYTAX	State Income Tax Payments	\$M	I
SYTAXR	State Average Income Tax Rate	F	X
Y	Personal Income	\$M	I
YD	Disposal Income		

Taxable Retail Sales

$$TRETS^t = f(YD^t)$$

Mnemonic	Description	Units	Type
TRETS	Taxable Retail Sales	\$M	I
YD	Disposable Income	\$M	I

4 Economic Activity

4.1 Introduction

The level of economic activity is modeled in terms of employment and output for each of 50 industries in San Diego County. For each industry, output provides a direct measure of the quantity of goods and services produced in that industry. Employment represents the main input into the production process. Other inputs, which are not represented here, include capital equipment, and raw materials, such as energy. For most industries, the approach taken in DEFM models employment directly based on indexes of demand and competitiveness. Output is derived based on employment levels and labor productivity forecasts.

4.2 Industry Definitions

Activity is modeled for a list of industries defined specifically for the San Diego region. Each industry is a collection of 3-digit or 4-digit NAICS codes. An effort was made to provide as much detail as possible relevant to the definition of economic clusters that had been defined as part of Economic Opportunities project. However, to be included in the list, an industry had to have historical data and a significant activity level. The list of categories is presented in Table 4-1.

Table 4-1: Industry Definitions

Sector	Description
E	Total Civilian Employment
EAGMIN	Employment: Natural Resources & Mining
ECONS236	Employment: Construction of Buildings
ECONS237	Employment: Heavy and Civil Engineering Construction
ECONS238	Employment: Specialty Trade Contractors
EMFG323	Employment: Printing and Related Support Activities
EMFG3254	Employment: Pharmaceutical and Medicine Manufacturing
EMFGONDUR	Employment: OTHER NONDURABLES = 311+ 312+313+314+315+316+321 +322+324+(325-3254)+326+327
EMFG3336	Employment: Engine, Turbine, and Power Transmission Equipment Manufacturing

DEFM Technical Documentation

Sector	Description
EMFG3341	Employment: Computer and Peripheral Equipment Manufacturing
EMFG3342	Employment: Communications Equipment Manufacturing
EMFG3343	Employment: Audio and Video Equipment Manufacturing
EMFG3344	Employment: Semiconductor and Other Electronic Component Manufacturing
EMFG3345_3346	Employment: Navigation, Measuring and Magnetic & Optical = 3345+3346
EMFG3366	Employment: Ship and Boat Building
EMFG3391	Employment: Medical Equipment and Supplies Manufacturing
EMFG3399	Employment: Other Miscellaneous Manufacturing
EMFGODUR	Employment: OTHER DURABLES = 331 + 332 + 333 + 335 + (336-3366) + 337
EWTRD	Employment: Wholesale Trade
ERTRD	Employment: Retail Trade
ETWU481	Employment: Air Transportation
ETWU485	Employment: Transit and Ground Passenger Transportation
ETWUOTH	Employment: OTHER TRANS & WAREHOUSING AND UTILITIES = 482+483+484+486+487+488+492+493+22
EINFO511	Employment: Publishing Industries (except Internet)
EINFO512_515	Employment: Motion Picture and Broadcasting = 512 + 515
EINFO517	Employment: Telecommunications
EINFOOTH	Employment: INTERNET AND OTHER = 518 + 519
EFRE522	Employment: Credit Intermediation and Related Activities
EFRE523	Employment: Securities, Commodity Contracts, and Other Financial Investments and Related Activities
EFRE531	Employment: Real Estate
EFREOTH	Employment: OTHER FINANCE AND REAL ESTATE = 521 + 524 + 525 + 532 + 533
EPBS5411	Employment: Legal Services
EPBS5412	Employment: Accounting, Tax Preparation, Bookkeeping, and Payroll Services
EPBS5413	Employment: Architectural, Engineering, and Related Services
EPBS5415	Employment: Computer Systems Design and Related Services
EPBS5416	Employment: Management, Scientific, and Technical Consulting Services
EPBS5417	Employment: Scientific Research and Development Services
EPBS55	Employment: Management of Companies and Enterprises
EPBS5613	Employment: Employment Services

Sector	Description
EPBS5615	Employment: Travel Arrangement and Reservation Services
EPBSOTH	Employment: OTHER PROFESSIONAL & BUSINESS SERVICES = 5414 + 5418 + 5419 + (56 - 5613 - 5615)
EEDHS621_622	Employment: Ambulatory Health & Hospitals = 621 + 622
EEDHS623	Employment: Nursing and Residential Care Facilities
EEDHSOTH	Employment: OTHER EDUCATION AND HEALTH SERVICES = 61 +624
ELH711_712	Employment: Performing Arts & Museums = 711 + 712
ELH713	Employment: Amusement, Gambling, and Recreation Industries
ELH721	Employment: Accommodation
ELH722	Employment: Food Services and Drinking Places
EOSV81	Employment: Other Services (except Public Administration)
EGOVFED	Employment: Federal Government
EGOVSTLCL	Employment: State & Local Government = GVS+GVL
ECONS	Employment: CONSTRUCTION (Super Sector (SS))
EMFG	Employment: MANUFACTURING (SS)
ETWU	Employment: TRANSPORTATION, WAREHOUSING, UTILITIES (SS)
EINFO	Employment: INFORMATION (SS)
EFRE	Employment: FINANCIAL & REAL ESTATE (SS)
EPBS	Employment: PROFESSIONAL & BUSINESS SERVICES (SS)
EEDHS	Employment: EDUCATION & HEALTH (SS)
ELH	Employment: LEISURE AND HOSPITALITY (SS)
EOSV	Employment: OTHER SERVICES (SS)
EGOV	Employment: GOVERNMENT (SS)
EOTH	Employment: Other
ESEDW	Employment: Self Employed and Domestic
JOBS	Employment: Civilian W & S + SEDW

4.3 General Approach

With a few exceptions, specifically agriculture, construction and government categories, the activity forecasting equations have a common form. This form specifies activity in each industry (i) as a linear function of a market index (MIND) which reflects demand conditions and regional competitiveness factors.

Each index provides a measure of the market for goods and services produced in San Diego by a specific industry, and the construction of these indexes is discussed below. For most sectors, the econometric equations take the following general form.

$$E_i^t = \beta_{i,0} + \beta_{i,1} \times \text{MIND}_i^t + \nu_i^t$$

where E_i is an estimate of San Diego employment for industry i ,
 MIND_i is the market index for goods and services produced by industry i ,
 ν_i is the residual for industry i .

4.4 Market Indexes

For most local industries, the level of economic activity is determined by a combination of factors reflecting local, state and national economic conditions. These factors may include local population and income, output and employment levels in closely related industries, levels of government spending, cost conditions, and so on. Since the data for a single area will not support estimation of full set interrelationships, prior structuring is used to impose the relative importance of the multiple explanatory factors. Although structure is imposed, the ultimate importance of the full set of explanatory factors is left to an econometric step.

The market indexes have two parts. The first part captures trends in demand for output produced in San Diego. The second part captures the relative competitiveness of goods and services produced in San Diego. Together, the combined index reflects a combination of demand and supply factors that are expected to determine the size of the market for San Diego output, and the local share of that market.

The demand portion of the index is constructed using weights from an estimated Input/Output table for 2001 for San Diego County. The specific table used to compute the index weights for this purpose was purchased from IMPLAN, which provides regional tables based on the most recent national tables and a variety of relevant regional information. Estimated I/O tables for 2001 were purchased for San Diego, California and the US.

The I/O transactions table gives the estimated gross dollar flows for each industry in the region. The IMPLAN tables provide data for a total of 528 industries, which have been aggregated to give data for the 50 industry groups represented in DEFM. For a given industry row, the dollar flows represent the amount of local output estimated to be purchased by each

local industry and by each of several export and final demand categories, including local residents, state and local governments, the Federal government, the rest of the US, and the international economy. For each industry row, these gross dollar flows are converted to fractions that sum to one, giving an indication of the relative importance of each demand category.

These weights are used to compute a composite demand index for each industry. The demand index is a fixed-weight index that combines growth in inter-industry industry demands and final demands into a single measure for the strength of demand for the output of an industry in a particular year. In constructing the indexes, a weighted average of current and lagged values for each demand factor is used. This implies that half of the direct impact from a change in a demand factor will occur in the year of the change and the remaining half will occur in the following year. For purposes of simplification, the index is presented below with only the current values for each variable. With this simplification, the index for industry i is indicated by $DIND_i$, and is defined generally as follows.

$$DIND_i^t = \sum_{j=1}^{50} \frac{E_j^t}{E_j^{2001}} F_{ij} + \sum_{k=1}^6 \frac{LFD_k^t}{LFD_k^{2001}} \times F_{ik} + \sum_{n=1}^5 \frac{X_n^t}{X_n^{2001}} \times F_{in}$$

where $DIND_i$ is the demand index for industry i

E_j is the estimated employment of industry j in San Diego and

LFD_k is an indicator of activity in local final demand sector k , and

X_n is an indicator of activity in export sector n , and

F_{ij} are the transactions weights for inter-industry demands,

F_{ik} are the transactions weights for local final demands, and

F_{in} are the transactions weights for export final demands.

The ratios (E/E^{2001}) provide a measure of the employment level in each industry relative to its 2001 level. For example, if employment for an industry is 20% above its 2001 level, this ratio will be 1.2. The local final demand (LFD) terms capture changes in local consumption, investment, and government demand. A set of six drivers are assigned to serve as proxies for demand in these categories, as summarized in the table below. For example, for consumption demand, the driver variable is real disposable income. The export demand (X) terms capture changes in California, US and international demand. A set of five drivers are assigned to serve a proxies for demand in these categories. For example, real gross state product is assigned as the driver for California final demand.

As indicated above, the demand index is a fixed-weight index. The weights for a specific source of demand represent the fraction of estimated transactions with that source in the 2001 I/O table. For example, if a final demand category represents 5% of the transactions in the 2001 table, then changes in the driver for this final demand category will get a 5% weight in construction of the industry demand index.

The variables used to indicate the level of activity in each final demand sector depend on the sector involved. For some final demand sectors, different variables are used for different types of local industries. For example, government demand for output from the service industries is most closely related to the level of government activity in the local economy. However, government purchases of manufactured goods may be more strongly related to national spending levels. Similarly, San Diego exports for manufacturing industries are most directly related to the level of activity nationally in each specific industry. This implies the use of industry-specific activity measures for the export sector for these categories.

The variables used in DEFM for constructing demand indexes are summarized below.

Final Demand Sector	Demand Index Variable (X)
SD Households	Real Disposable Income (R_YD)
S&L Govt Education	San Diego Government Spending on Education (R_GEd)
S&L Govt NonEd	San Diego Government Spending Noneducation (R_GLoc)
Fed Govt Military	Military in Uniform (MUNIF)
Fed Govt NonMilitary	San Diego Output for Federal Sector other than military (QFEDNMIL)
Investment	Real Gross Regional Product (R_GRP) × US Investment Per Dollar of Output (IUS/GDPUS)
Exports – Orange Cnty	Real Gross Regional Product for Orange County (R_GRP_OC)
Exports – California	Real Gross State Product for California (R_GSP_CA)
Exports – Rest of U.S.	Real Gross Domestic Product (R_GDPUS)
Exports – International	Real US Exports (R_EX_US)
Exports – Mexico	Real US Exports (R_EX_US)

The regional competitiveness index is computed to represent the relative cost of business in San Diego. This index is computed using relative wage rates, relative productivity levels, relative capital costs, and relative fuel costs.

$$CIND_i^t = \left(\frac{R_W^t / LABPROD_i^t}{R_WUS^t / LABPRODUS_i^t} \right)^{a_i} \times \left(\frac{CAPCOST^t}{CAPCOSTUS^t} \right)^{b_i} \times \left(\frac{FUELCOST^t}{FUELCOSTUS^t} \right)^{c_i}$$

where R_W is the real wage rate in San Diego
 R_WUS is the real US wage rate
 CAPCOST is the San Diego cost of capital,
 CAPCOSTUS is the US cost of capital,
 FUELCOST is the San Diego cost of fuel (gas and electricity) and
 FUELCOSTUS is the US cost of fuel (gas and electricity)
 LABPROD is Labor Productivity
 LABPRODUS is Labor Productivity, US

The parameters in the competitiveness index (a_i , b_i , and c_i) give the elasticity of overall relative cost with respect to each cost component. These elasticity values are computed to represent the relative importance of each cost category. Estimates for these coefficients are based on data for San Diego from the IMPLAN system. The coefficient for labor costs represents employee compensation plus self employment income as a share of value added. The coefficient for fuel cost represents gross purchases from the utility sector as a fraction of total value added. The coefficient on capital costs represents the cost share of all remaining items other than labor and fuel.

Increases in the San Diego competitiveness index will occur when any of the local cost elements decrease relative to US costs, or when local labor productivity increases relative to national productivity. Conversely, if local costs outstrip national costs or if local productivity gains lag behind national gains, local competitiveness will decline. When these types of changes occur, demand will be focused toward or away from the local economy, depending on the degree to which the industry involved is competitive.

In what follows, the demand index and competitiveness indexes are combined into a single market index. In forming this combination, the importance of the competitiveness index is assumed to depend on the degree to which there is competition with outside providers of the goods and services produced by an industry. For industries that are largely locally based, such as legal services, the competitiveness index will be a relatively unimportant factor, and the elasticity of demand with respect to local costs and productivity will be small. For industries that are export based, the reverse is true, and the elasticity of demand for San Diego output with respect to local costs and productivity will be large.

The impacts of local competitiveness are most obvious for export-based manufacturing industries. In this case, local companies are competing directly in export markets and are also competing with outside suppliers for local business. It is also true, however, for service based industries, to the extent that local purchasers of local services have alternatives outside the local region. Rather than pick an arbitrary cutoff point that specifies some industries as being competitive and some as being noncompetitive, a sliding scale was developed.

For export demands in all industries, it is assumed that the competitiveness index has full impact. For inter-industry demands and local final demands, a sliding scale is used to reflect the extent to which each industry is locally based. This is determined based on the Regional Purchase Coefficient estimates from the IMPLAN database. These coefficients give the fraction of regional consumption that is produced locally. For each industry, it is assumed that twice the non-local fraction is competitive, and the remainder is noncompetitive.

The elasticity coefficients for regional competitiveness index are based on an overall relative profitability elasticity published in Treyz, Rickman, and Shao (The REMI Economic-Demographic Forecasting and Simulation Model, *International Regional Science Review*, Vol 14, No. 3, pp 221-253, 1992). This paper provides an estimated overall elasticity of 1.83, which is estimated on a geometric distributed lag of current and past relative profitability. With this overall elasticity, the response to a specific component of cost depends on the share of total cost represented by the component and on the fraction of local purchases that are

competitive. For example, if labor costs are 67% of total costs, then the elasticity of the regional output share respect to labor costs is $.67 \times 1.83$ in the competitive portion of the market.

Combining the components of the demand index with the regional competitiveness index gives the following expression for the overall market index:

$$\begin{aligned} \text{MIND}_i^t = & \left(\sum_{j=1}^{52} \frac{Q_j^t}{Q_j^{2002}} \times F_{ij} + \sum_{k=1}^6 \frac{\text{LFD}_k^t}{\text{LFD}_k^{2002}} \times F_{ik} \right) \times \left(\text{CIND}^t \right)^{-1.83 \times \text{CFRAC}} \\ & + \left(\sum_{n=1}^4 \frac{X_n^t}{X_n^{2002}} \times F_{in} \right) \times \left(\text{CIND}^t \right)^{-1.83} \end{aligned}$$

where MIND_i is the overall market index, and

CFRAC is the competitive fraction for local purchases based on the estimated RPC.

4.5 Employment Equations

For most sectors, the equations for economic activity are estimated using econometric equations that contain the market index variables as the main explanatory factor in the form:

$$E_i^t = \beta_{i,0} + \beta_{i,1} \times \text{MIND}_i^t + \upsilon_i^t$$

where E_i is an estimate of San Diego employment for industry i ,

MIND_i is the market index for goods and services produced by industry i ,

υ_i is the residual for industry i .

Exceptions to this general specification are as follows:

- For mining, employment is forecasted based on an estimated time trend term. Fluctuations in mining activity were unrelated to fluctuations in the demand index over the historical period.
- Construction activity and employment is assumed to depend on the size of the economy and the level of construction activity (authorizations). The size of the economy drives the demand for maintenance and remodeling services, and authorizations indicate the level of new construction.

- Federal government employment and the related contribution to regional product are assumed to depend on national federal government employment (for federal non-defense), and uniformed military (for federal defense).
- State and local government employment and the related contribution to regional product are assumed to depend on state university enrollments (for state education) and population (for other state and local employment).

4.6 Employment Aggregates

Output and employment are forecasted separately for each of 50 industries, listed in Table 4-1, including the government sectors. The following variables represent aggregate employment:

$$\begin{aligned}
 E &= EAG + ENAG + EOTHER \\
 ECON &= ECONS236 + ECONS237 + ECONS238 \\
 EMFG &= EMFG323 + EMFG3254 + EMFGONDUR + EMFG3336 + \\
 &EMFG3341 + EMFG3342 + EMFG3343 + EMFG3344 + \\
 &EMFG3345_3346 + EMFG3366 + EMFG3391 + EMFG3399 + \\
 &EMFGODUR \\
 ETRD &= EWTRD + ERTRD \\
 ETWU &= ETWU481 + ETWU485 + ETWUOTH \\
 EINFO &= EINFO511 + EINFO512_515 + EINFO517 + EINFOOTH \\
 EFRE &= EFRE522 + EFRE523 + EFRE531 + EFREOTH \\
 EPBS &= EPBS5411 + EPBS5412 + EPBS5413 + EPBS5415 + EPBS5416 + \\
 &EPBS5417 + EPBS55 + EPBS5613 + EPBS5615 + EPBSOTH \\
 EEDHS &= EEDHS621_622 + EEDHS623 + EEDHS_OTH \\
 ELH &= ELH711_712 + ELH713 + ELH721 + ELH722 \\
 EOSV &= EOSV81 \\
 EGOV &= EGOVFED + EGOVSTLCL \\
 ENAG &= ECON + EMFG + ETRD + ETWU + EINFO + EFRE + EPBS + \\
 &EEDHS + ELH + EOSV + EGOV
 \end{aligned}$$

$$EWandS = ENAG + EAGMIN$$

$$JOBS = EWandS + SEDW$$

Mnemonic	Description	Units	Type
E	Total Resident Employment	A	I
EAGMIN	Employment – Agriculture and Mining	A	I
ECON	Employment - Construction	A	I
EMFG	Employment - Manufacturing	A	I
ETRD	Employment - Trade	A	I
ETWU	Employment – Transportation & warehousing	A	I
EINFO	Employment - Information	A	I
EFRE	Employment – Finance and Real Estate	A	I
EPBS	Employment – Professional & Business Services	A	I
EEDHS	Employment – Education & Health Services	A	I
ELH	Employment – Leisure & Hospitality	A	I
EOSV	Employment – Other Services	A	I
EGOV	Employment – Government	A	I
ENAG	Employment – Non Ag	A	I
EWANDS	Employment – Wage and Salary	A	I
EOTH	Employment – Other	A	I
JOBS	Employment – Jobs (Place of Work)	A	I
SEDW	Employment – Self employed and domestics	A	I

4.7 Output Equations

The equations above give estimates of employment by industry. The employment forecasts for each industry are translated into output forecasts based on the projected values for labor productivity. That is, for each industry:

$$Q_i^t = EMP_i^t * LABPROD_i^t$$

where EMP_i is an San Diego employment in industry i ,

Q_i is San Diego output for industry i , and

$LABPROD_i$ is the San Diego labor productivity for industry i .

A full model of labor productivity would account for changes in technology, investments in physical capital, human capital characteristics of the San Diego population, and investments in local infrastructure. Because there are no reliable data for any of these variables, a simple exogenous approach is taken in DEFM. Specifically, labor productivity is modeled as an exogenous factor that depends on national estimates of labor productivity and User-supplied shift variables. Specifically:

$$\text{LABPROD}_i^t = f(\text{LABPRODUS}_i^t)$$

where LABPRODUS_i is US labor productivity for industry i , in thousand \$2010 per employee.

4.8 Output Aggregates

Output (Q) is forecast separately for each of 50 private sector industries, listed in Table 4-1, and 7 government categories. The following variables represent aggregate output:

$$\begin{aligned} Q &= QAGMIN + QNAG + QOTH \\ QCON &= QCONS236 + QCONS237 + QCONS238 \\ QMFG &= QMFG323 + QMFG3254 + QMFGONDUR + QMFG3336 + \\ &QMFG3341 + QMFG3342 + QMFG3343 + QMFG3344 + \\ &QMFG3345_3346 + QMFG3366 + QMFG3391 + QMFG3399 + \\ &QMFGODUR \\ QTRD &= QWTRD + QRTRD \\ QTWU &= QTWU481 + QTWU485 + QTWUOTH \\ QINFO &= QINFO511 + QINFO512_515 + QINFO517 + QINFOOTH \\ QFRE &= QFRE522 + QFRE523 + QFRE531 + QFREOTH \\ QPBS &= QPBS5411 + QPBS5412 + QPBS5413 + QPBS5415 + QPBS5416 + \\ &QPBS5417 + QPBS55 + QPBS5613 + QPBS5615 + QPBSOTH \\ QEDHS &= QEDHS621_622 + QEDHS623 + QEDHS_OTH \\ QLH &= QLH711_712 + QLH713 + QLH721 + QLH722 \\ QOSV &= QOSV81 \\ QGOV &= QGOVFED + QGOVSTLCL \\ QNAG &= QCON + QMFG + QTRD + QTWU + QINFO + QFRE + QPBS + \\ &QEDHS + QLH + QOSV + QGOV \\ QWANDS &= QNAG + QAGMIN \end{aligned}$$

Mnemonic	Description	Units	Type
Q	Total Resident Employment	A	I
QAGMIN	Output – Agriculture and Mining	A	I
QCON	Output - Construction	A	I
QMFG	Output - Manufacturing	A	I
QTRD	Output - Trade	A	I
QTWU	Output – Transportation & warehousing	A	I
QINFO	Output - Information	A	I
QFRE	Output – Finance and Real Estate	A	I
QPBS	Output – Professional & Business Services	A	I
QEDHS	Output – Education & Health Services	A	I
QLH	Output – Leisure & Hospitality	A	I
QOSV	Output – Other Services	A	I
QGOV	Output – Government	A	I
QNAG	Output – Non Ag	A	I
QWANDS	Output – Wage and Salary	A	I
QOTH	Output – Other	A	I

4.9 Unemployment Rate

The unemployment rate is computed from data on the civilian labor force and total resident employment. The civilian labor force is computed in the demographic section of the model, as discussed in Section 3. Total resident employment is computed above as the sum of wage and salary employment levels in San Diego and a residual term, EOther. This residual term captures employment other than wage and salary employment, such as self employed and domestic workers. It also captures the adjustment between employment by place of work and employment by place of residence. That is, the employment total (E) is the total number of San Diego residents who are employed, regardless of their place of work. In contrast, all of the wage and salary data are for employment by place of work. In the construction of the data, EOther is computed as a residual, and therefore implicitly contains the net adjustment for the difference between San Diego residents who are employed elsewhere and the number of individuals who are employed in San Diego, but who reside elsewhere. As a result, the unemployment rate provides an estimate of the fraction of San Diego residents who are in the labor force but who are not currently employed.

$$UNEMP = L - E$$

$$U = \frac{L - E}{L}$$

Mnemonic	Description	Units	Type
E	Total Resident Employment	A	I
L	Civilian Labor Force	A	X
UNEMP	Unemployed Residents	A	I
U	Unemployment Rate	%	I

4.10 Gross Regional Product

The estimates of output by industry can be summed up to provide an estimate of Gross Regional Product. This bottom-up estimate will not necessarily agree with the data series for GRP developed by the San Diego Chamber of Commerce. These data are developed based on estimates of Gross State Product for California and personal income data for San Diego and California. The following bridging equation is used to calibrate the bottom-up estimates from the DEFM industries with the aggregate value from the Chamber.

$$R_GRP_t = f(QW\&S_t)$$

$$GRP_t = R_GRP_t \times PGDPUS_t$$

Mnemonic	Description	Units	Type
GRP	San Diego Gross Regional Product	\$M	I
PGDPUS	US Deflator for Gross Domestic Product	F	X
QW&S	Value Added by Wage and Salary	\$M2010	I
R_GRP	Real San Diego Gross Regional Product	\$M2010	I

4.11 Orange County Gross Regional Product

One of the final demand sectors represents exports from San Diego to neighboring Orange County. Although a detailed model of the Orange County economy is not included here, interactions between the two regions are modeled. This is accomplished by including a simple model of gross regional product in Orange County, modified by changes in the San Diego economy. This is accomplished in three steps:

- Orange County Base Forecast. The base model relates real output in Orange County to the California economy.
- San Diego County Base Forecast. The base model for San Diego County relates real output in San Diego to the California Economy.
- Modified Orange County Forecast. The modified Orange County forecast for real output is defined by the base forecast modified through an elasticity value for the deviation between the San Diego base forecast and the DEFM forecast.

Because Orange County GRP feeds back into the activity equations as a source of final demand, this provides an interaction between the two county economies. That is, any increase in the San Diego forecast will feed into an increased forecast for Orange County, which will, in turn, feedback to further increase the San Diego forecast.

$$R_GRPBASESD_t = f(R_GSPCA_t)$$

$$R_GRPBASEOC_t = f(R_GSPCA_t)$$

$$R_GRP_OC^t = R_GRPBASEOC^t \times \left(\frac{R_GRP^t}{R_GRPBASESD^t} \right)^{.05}$$

$$GRPOC_t = R_GRPBASEOC_t \times PGDPUS_t$$

Mnemonic	Description	Units	Type
GRP_OC	Orange County Gross Regional Product	\$M	I
PGDPUS	US Deflator for Gross Domestic Product	F	X
R_GRP	Real San Diego Gross Regional Product	\$M2010	I
R_GRPBASEOC	Time Series Forecast for R_GRPOC	\$M2010	I
R_GRPBASESD	Time Series Forecast for R_GRPSD	\$M2010	I
R_GRP_OC	Orange County Real Gross Regional Product	\$M2010	I
R_GSPCA	California Real Gross State Product	\$B2010	X

5 Construction

5.1 Model Logic

The construction sector covers the level of residential and nonresidential building activity. Activity levels are based on local demand for homes and nonresidential buildings and national economic conditions, captured by national construction levels. The overall flow of logic in this sector is depicted in Figure 5-1.

Housing Supply

Housing units are forecast by structure type as a function of the previous year's stock plus the number of last year's residential permit authorizations that were completed. For single and multifamily housing stocks, an historically estimated relationship translates new authorizations into stock changes. This relationship implicitly adjusts for permit realization rates that are less than 100% as well as for replacement of demolitions. The stock of manufactured housing units (mobile homes) is forecast to increase at the average historic rate in recent years, based on an estimated time trend. This estimated trend indicates that the addition of about 5,000 units per decade.

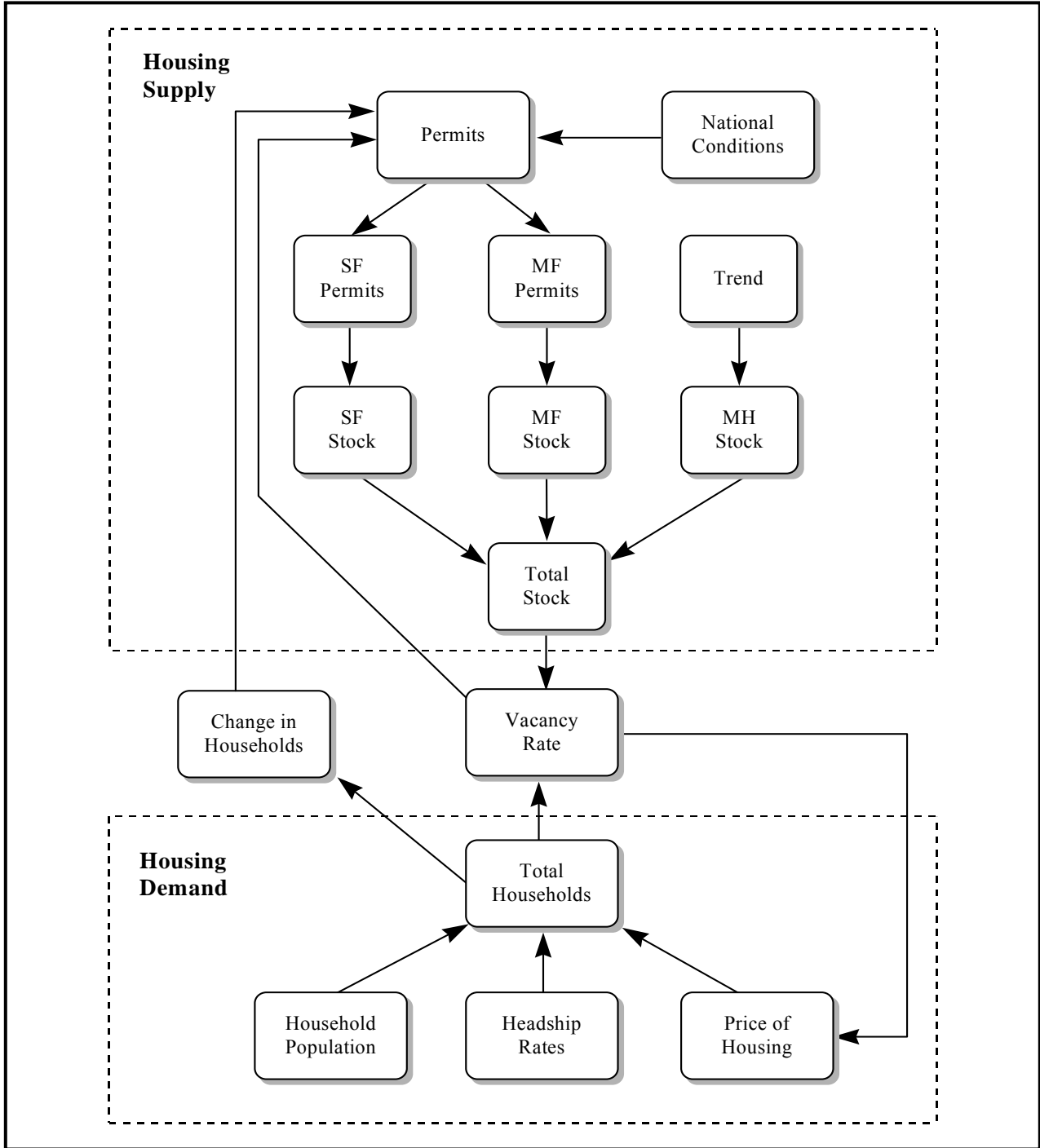
Vacancy Rates

Housing demand is determined by the number of households, which is computed in the Demographic sub-model based on exogenous housing preference factors, household headship rates, and elasticity values with respect to housing prices. Vacancy rates are computed from the supply and demand for housing units.

Housing Unit Authorizations

The number of new housing units authorized by building permits in a forecast year is positively related to local demand for new homes and to national housing market conditions, and is inversely related to the housing stock vacancy rate. Local demand depends on the increase in the number of households and the replacement of old housing stock. No consideration is given to local mortgage market conditions since secondary mortgage markets rather than local deposit inflows serve as the main source of mortgage funds. Given these secondary markets, the national housing market variables are assumed to account for national and therefore local mortgage market conditions.

Figure 5-1: Construction Sector Logic



Permit activity is allocated between single and multifamily authorizations based on an estimated equation that depends on the single family share for national construction.

Construction Value

Residential. The real value of authorizations for new residential units is influenced by the level of single family and multifamily permits. The nominal value of authorizations depends on the real value and the level of construction costs.

Nonresidential. Valuation of new nonresidential construction is forecast by a regression equation relating activity to the gain in local employment and to national construction market conditions. National conditions are captured by the value of nonresidential construction investment in the US.

Construction Sector Shifts

A variety of additive and multiplicative shift factors are included. For residential construction, the main shifts appear in the construction authorization equations, which flow through to the housing stock and construction valuation equations, since unit authorizations appear in these equations. For nonresidential construction, shifts appear directly in the valuation equations.

5.2 Construction Equations

Housing Unit Authorizations

$$NAFDU^t = f(HH^t, HH^{t-1}, HS^{t-1}, HSTARTSUS^T, HS^{t-1}, HSUS^{t-1}, (.05/VRATE^t))$$

$$SFSHARE^t = f\left(\frac{HSTARTSFUS^t}{HSTARTSUS^t}\right)$$

Mnemonic	Description	Units	Type
HH	San Diego Households	A	O
HS	San Diego Housing Stock	A	I
HSTARTSUS	US Housing Starts	M	X
HSTARTSSFUS	US Single Family Housing Starts	M	X
HSUS	U.S. Housing Stock	M	X
NAFDU	Number of Authorized Family Dwelling Units	A	I
NAFDUSF	Number of Authorized Single Family Units	A	I
NAFDUMF	Number of Authorized Multi Family Units	A	I
SFSHARE	San Diego Single Family Share of Authorizations	F	I
VRATE	Residential Vacancy Rate	%	I
α	Vacancy Rate Elasticity	F	I

Housing Supply

$$HSSF^t = HSSF^{t-1} + 0.95 \times NAFDUSF^t$$

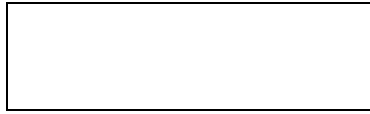
$$HSMF^t = HSMF^{t-1} + 0.90 \times NAFDUMF^t$$

$$HSMOB^t = f\left(\frac{(YEAR - 1990)}{10}\right)$$

$$HS^t = HSSF^t + HSMF^t + HSMOB^t$$

Mnemonic	Description	Units	Type
HS	Total Housing Stock	A	I
HSMF	Multi-Family Housing Stock	A	I
HSMOB	Stock of Mobile Homes	A	I
HSSF	Single-Family Housing Stock	A	I
NAFDUMF	Multi-Family Authorizations	A	I
NAFDUSF	Single Family Authorizations	A	I
YEAR	Value of t, for example, 2012	-	X

Vacancy Rates



Mnemonic	Description	Units	Type
HH	Total Housing Demand (Households)	A	O
HS	Total Housing Stock	A	I
VRATE	Residential Vacancy Rate	F	I

Value of Construction Authorizations

$$R_VRCA^t = f(NAFDU^t, HS^t)$$

$$R_VNRCA = f\left(\left(ENAG^t - ENAG^{t-2}\right), \left(\frac{VNRCIUS^t}{CCUS^t} \times \frac{ENAG^t}{ENAGUS^t}\right)\right)$$

$$VRCA^t = CC2000^t \times R_VRCA^t$$

$$VNRCA^t = CC2000^t \times R_VNRCA^t$$

$$VACA^t = VRCA^t + VNRCA^t$$

Mnemonic	Description	Units	Type
CC2010	Construction Cost Index (1.00 in 2010)	F	O
CCUS	US Construction Cost Index (1.00 in 2010)	F	O
ENAG	Nonagricultural Wage and Salary Employment	A	O
EUSNAG	US Nonag Wage and Salary Employment	M	X
HS	Total Housing Stock	A	O
NAFDU	Number of Authorized Family Dwelling Units	A	I
VNRCIUS	Value of US Non-Residential Construction	B\$2010	X
R_VRCA	Real Value of Res Construction Authorizations	\$M2010	I
R_VNRCA	Real Value of NonRes Const Authorizations	\$M2010	I
VACA	Value of All Construction Authorizations	\$M	I
VNRCA	Value of Non-Residential Const Authorizations	\$M	I
VRCA	Value of Residential Construction Authorizations	\$M	I

6 Public Finance

DEFM contains data on revenues and expenditures of the County, the Cities, and School Districts. Revenues include property taxes, retail sales tax, federal and state grants, and other revenues (fines, service fees, assessments, etc.). Expenditures include educational expenditures and other local government expenditures. Revenues and expenditure forecasts are based on regression models that relate the level of dollar flows to measures of the scale of the region and the level of economic activity.

6.1 Revenue

Local government revenues are modeled in five parts. First, property tax revenues are modeled as a function of assessor's market value. The remaining components are state and Federal grants, sales tax revenues, and other revenues.

Assessor's Market Value

The tax base for property taxes is termed "assessor's market value" (AMV). Because of limitations imposed by the Gann Appropriation limits, the response of AMV to market conditions occurs with a lag. In times of price increase, the associated increase in assessed market values is limited unless there is a change in ownership. However, over time, as property changes hands, the AMV will eventually reflect the market value as captured by the transaction price. Also, in times of declining value, AMV will follow with some lag as properties turn over or as owners request assessment updates.

A variety of lag structures were examined to model the expected lag between market values and AMV. The best fit was obtained with a moving average of 5 years for residential property and 3 years for nonresidential property. The resulting equation specifies total AMV as a function of the moving average of housing prices interacted with the housing stock and a moving average of construction costs interacted with nonagricultural employment, which serves as a proxy for the stock of nonresidential structures.

Property Tax Revenues

Property tax revenues of local government agencies are forecast from the assessor's market value tax base. Historically, the overall average rate for total property taxes collected (since Prop. 13) reflects the one percent full value tax and additional debt service in voter-approved tax rate areas. The regression equation indicates a slope of about 1.16 percent as the average realized tax rate.

Grants

Federal revenue sharing, which was a source of local government revenues through the mid-1980s has been discontinued. Other Federal grants and State grants are a major source of funding for education and health and welfare programs. Total revenues from these sources are modeled as a function of San Diego population and average daily attendance.

Total Retail Sales Tax

Local revenues from retail sales taxes are set to 2.25% of Taxable Retail Sales.

Other Revenues

Other revenues include fines, fees, service assessments, etc. These revenues represent most of the unconstrained local revenues not affected by spending limitations. Since 1979, these revenues grew proportionately faster than the other sources of revenues and now make up the majority of non-education related local government revenue. Since these revenues are closely linked to local economic and demographic conditions, an historically estimated equation is used for forecasting purposes. This equation relates the level of revenues to population and economic activity.

6.2 Expenditures

Local government expenditures are modeled in two pieces. The first covers educational expenditures and the second covers other local government expenditures. Educational expenditures are forecasted based on an estimated relationship between real expenditures and average daily attendance (ADA). Local government expenditures are modeled based on the historical relationship between these real expenditures and real revenues less education expenditures.

6.3 Revenue Model Equations

$$AMV^t = f(MA5(PH), HS^t, MA3(CC), ENAG^t)$$

$$RPTAX^t = f(AMV^t)$$

$$R_RGRANT^t = f((P^t)^4, (ADA^t)^6)$$

$$RSALES^t = f(.0225) \times TRETS^t$$

$$R_ROTH^t = f(CP^t)$$

$$ROTH^t = R_ROTH^t \times CPI2000^t$$

$$RGRANT^t = R_GRANT^t \times CPI2000^t$$

$$R^t = RPTAX^t + RSALES^t + RGRANT^t + ROTH^t$$

Mnemonic	Description	Units	Type
AMV	Assessor's Market Value	\$M	I
ADA	Average Daily Attendance	A	I
CP	Civilian Population	A	O
ENAG	Non Agricultural Employment	A	O
HS	Total Housing Stock	A	O
MA5(PH)	Five year moving average of housing price	M	I
MA3(CC)	Three year moving average of construction cost	T	I
CPI2010	Consumer Price Index (2010 = 1.00)	F	O
P	San Diego Population	A	O
R	Total Government Revenues	\$M	I
RGRANT	Total Subvention and Grant Revenues	\$M	I
ROTH	Other Revenues	\$M	I
RPTAX	Property Tax Revenues	\$M	I
RSALES	Sales Tax Revenues	\$M	I
R_GRANT	Total Subvention and Grant Revenues in 2010\$	\$M2010	I
R_ROTH	Other Revenues in 2010\$	\$M2010	I
TRETS	Taxable Retail Sales	\$M	O

Expenditure Equations

$$ADA^t = f(ENRLTOT^t)$$

$$R_GED^t = f(ADA^t)$$

$$R_REVNONED^t = R_R^t_RGED^t$$

$$R_GLOC^t = f(R_REVNONED^t)$$

$$GED^t = R_GED^t \times CPI2000^t$$

$$GLOC^t = R_GLOC^t \times CPI2000^t$$

$$G = GED + GLOC$$

Mnemonic	Description	Units	Type
CPI2010	Price Index in 2010\$	F	O
ENRLTOT	Total Enrollment	A	I
GED	School Districts Expenditures	\$M	I
GLOC	Local Government Non Education Expenditures	\$M	I
G	Total Local Government Expenditures	\$M	I
R_GED	School Districts Expenditures in 2010\$	\$M2010	I
R_GLOC	Local Government Non Education Expenditures	\$M2010	I
R_REVNONED	Non Education Revenue	\$M	I
ADA	Average Daily Attendance	A	O
R_R	Total Local Government Revenues	\$M2010	I

7 Prices

7.1 Discussion of Equations

Four price variables appear in the DEFM model, including the consumer price index, the average price of a single family home, a construction cost index, and a wage rate index. All real income and price data in DEFM are based on 2010 constant dollars. All price deflators are computed from the U.S. Department of Commerce Consumer Price Index, with 1982-1984 average = 100.00. For some series, the US Gross Domestic Product deflator, which is an exogenous variable in the model, is used.

Construction Cost Index

Construction costs are modeled as a function of a national cost index modified for relative wage levels in San Diego. A second factor measures the interaction of national costs with a measure of the level of construction activity in San Diego. This measure is computed based on dwelling unit authorizations per household in the county. When this ratio is high, construction costs will be slightly higher, reflecting high levels of demand. When this ratio is low, slack conditions will result in reduced construction costs.

Housing Prices

Housing price level forecasts are based on an historically estimated relationship with an adjusted income variable. Adjusted income is defined to include average household income plus a dynamic wealth term reflecting past capital gains on homes. This dynamic term is based on the change in housing prices over the last 6 years.

The adjusted income term is multiplied by a vacancy rate adjustment, centered around 5%. When vacancy rates are above 5%, this will result in reduced housing price levels. When vacancy rates are below 5%, this will place additional upward pressure on housing prices. The vacancy rate effect is modeled using elasticity on the inverse vacancy rate. The elasticity is asymmetric and has the following effect. At 5%, the vacancy rate multiplier is 1.0. At lower vacancy rates, the elasticity is .14, implying an overall multiplier of about 1.10 at a 3% vacancy rate. At higher vacancy rates, the elasticity is .07, implying an overall multiplier of about .95 at a 12% vacancy rate. The coefficient on the combined term, which includes adjusted income and the vacancy rate multiplier, is estimated econometrically.

Consumer Price Index

The local CPI is estimated to follow its national counterpart up to the influence of differential changes in local and national housing prices. The elasticity on the relative price of housing is set to 12.5%, which is the weight for housing costs in the consumer price index. By using this elasticity, a 10% increase in the relative cost of housing in San Diego will cause a 1.25% increase in the San Diego CPI.

Wage Rate

The wage rate index for San Diego is modeled as a function of the US earnings index and the difference between unemployment rates in San Diego and the US. The model is structured so that each percentage point difference between local and national unemployment rates takes 1.5% off of the explanatory variable index. For example, if the local unemployment rate is 5% and the national rate is also 5%, the modifier for unemployment rates is a multiplier of 1.0 and has no effect. If the local unemployment rate is 6% and the national rate is 4%, this 2 percentage point difference reduces the modifier to .97, indicating less pressure on San Diego wages relative to the national wage rate. Conversely, if the local unemployment rate is 4% and the national rate is 6%, this 2 percentage point difference increases the modifier to 1.03, indicating upward pressure on San Diego wages relative to national wages. The coefficient on the product of the national wage rate and the unemployment modifier is estimated based on historical relationships in the data.

7.2 Price Equations

Construction Cost

$$CC2000_t = f\left(CCUS \times \frac{W_t}{WUS_t}, CCUS \times \frac{NAFDU_t}{HH_t} \right)$$

Mnemonic	Description	Units	Type
CC2010	Construction Cost Index (2010 = 1.0)	F	I
CCUS	U.S. Construction Cost Index (2010 = 1.0)	F	X
HH	Households	A	O
NAFDU	Number of Authorized Dwelling Units	A	O
W	San Diego Earnings Index	F	I
WUS	US Earnings Index	F	I

Housing Price

$$PH_t = f \left(\left(ADJY_t \times \left(\frac{.05}{VRATE_t} \right)^{PHELAS} \right), PHUS \right)$$

$$ADJY_t = \frac{Y_t}{HH_t} + .75 \times \left(\frac{PH_{t-1} - PH_{t-6}}{5} \right)$$

$$PHELAS = (VRATE_t > .05) \times .07 + (VRATE_t \leq .05) \times .14$$

Mnemonic	Description	Units	Type
ADJY	Income per HH adjusted for capital gains	T\$	I
HH	Households	A	O
PH	Average Price of Housing	A	I
PHUS	US Price of Housing	T	X
PHELAS	Elasticity with respect to the vacancy rate	F	P
VRATE	Housing Vacancy Rate	F	O
Y	Personal Income	\$M	I

Consumer Price Index

$$CPI2000_t = f \left(CPIUS2000_t \times \left(\frac{PH_t}{PHUS_t} \right)^{.125} \right)$$

Mnemonic	Description	Units	Type
CPI2010	Consumer Price Index in 2010\$	F	I
CPIUS2010	U.S. Consumer Price Index 2010\$	F	I
PH	Average Price of Single Family Home	A\$	I
PHUS	U.S. Average Price of Housing	T\$	X

Wage Rate

$$W = f(WUS_t, (1 - (MA3(U_t) - MA3(UUS_t))))$$

$$R_W_t = \frac{W_t}{CPI2000_t}$$

Mnemonic	Description	Units	Type
CPI2010	Consumer Price Index in 2010\$	F	I
W	San Diego Earnings Index	F	I
WUS	US Earnings Index	F	I
U	San Diego Unemployment Rate	%	O
UUS	U.S. Unemployment Rate	%	X
MA3	Three year moving average function		
R_W	Real Wages in 2010\$	F	I

SERIES 13 SUBREGIONAL ALLOCATION

TECHNICAL DOCUMENTATION

July 2015



401 B Street, Suite 800
San Diego, CA 92101
(619) 699-1900
www.sandag.org

MEMBER AGENCIES:

Cities of Carlsbad, Chula Vista, Coronado, Del Mar, El Cajon, Encinitas, Escondido, Imperial Beach, La Mesa, Lemon Grove, National City, Oceanside, Poway, San Diego, San Marcos, Santee, Solana Beach, Vista, and County of San Diego.

ADVISORY/LIAISON MEMBERS:

California Department of Transportation, U.S. Department of Defense, San Diego Unified Port District,

ABSTRACT

TITLE: Series 13 Subregional Allocation
Technical Description

AUTHOR: San Diego Association of Governments

DATE: July 2015

SOURCE OF COPIES: San Diego Association of Governments
401 B Street, Suite 800
San Diego, CA 92101
(619) 699-1900

ABSTRACT: This document provides a detailed technical description of the programs used for the Series 13 subregional allocation

TABLE OF CONTENTS

SECTION 1 – OVERVIEW.....	1
1.1 Geographic Specification.....	1
1.2 Nomenclature.....	1
1.3 Generalized Model Structure.....	3
SECTION 2 – 11 MODEL COMPONENTS.....	5
2.1 GPALL.....	5
2.2 Capacity.....	11
2.3 UDM Module 1 – Preprocess Capacity Data.....	14
2.4 UDM Module 2 – Civilian Employment Allocation.....	17
2.5 Module 3 – Housing Stock Allocation.....	30
2.6 Module 4 – Detailed Characteristics.....	46
SECTION 3 – PROGRAMMING.....	69
3.1 Introduction.....	69
3.2 AccessWeights.....	70
3.3 Capacity.....	72
3.4 GPALL.....	74
3.5 UDM.....	76
Appendix A Household Income Distribution.....	79
Appendix B Generalized Development Type Codes.....	81

LIST OF TABLES

Table 1	Development Code Assignments	5
Table 2	Household Income Allocation – Initial Matrix	62
Table 3	Household Income Allocation – LUZ Decrement Accounting	63
Table 4	Household Income Allocation – MGRA Decrement Accounting.....	64
Table 5	Household Income Allocation – MGRA 1022 Increment Accounting.....	65
Table 6	Household Income Allocation – MGRA 1024 Increment Accounting.....	65
Table 7	Series 11 Forecast Software and Data Organization	70
Table 8	AccessWeights Input and Output	70
Table 9	CAPACITY Input and Output.....	72
Table 10	GPALL Input and Output.....	74
Table 11	UDM Input and Output	76

LIST OF FIGURES

Figure 1	Subregional Modeling Flow of Information	4
Figure 2	GPALL Flow of Information.....	6
Figure 3	GPALL Detail.....	7
Figure 4	AccessWeights Flow of Information	71
Figure 5	CAPACITY Flow of Information.....	73
Figure 6	GPALL Flow of Information.....	75
Figure 7	UDM Flow of Information	77

SECTION 1 – OVERVIEW

This document provides a detailed technical description of the models used in the subregional allocation of population, housing, income, and employment for the SANDAG Series Forecasts. It discusses the four modules that constitute the Urban Development Model (UDM), as well as the ancillary programs CAPACITY and GPALL.

1.1 GEOGRAPHIC SPECIFICATION

Beginning with Series 13 and the 2005 Current Estimates, SANDAG redefined its atomic level of geography for land use, housing and employment to be the land-use polygon as implemented in the agency's geographic information system (GIS). The Landcore Polygon – LANDCORE is SANDAG's primary GIS land-use database table - may represent an individual parcel or in many cases, a split parcel. While the complete parcel file is always changing, a set of static Landcore polygons, which reflect data at one point in time, is used to complete a forecast and current estimates series. Each Landcore polygon is identified by a unique identifier called an **LCKEY**.

SANDAG uses a multilevel, nested geographic reference system. The foundation of the reference system is the Master Geographic Reference Area (MGRA). The 23,002 MGRAs are the combinations of census blocks, census tracts, community planning areas, city boundaries, spheres of influence, and zip codes. Blocks also are split using other criteria (e.g. ridgelines) to develop traffic analysis zones. Accordingly, MGRAs can aggregate to these standard geographic units, including subregional areas and major statistical areas, which are aggregations of census tracts. Because MGRAs are small, their aggregation can closely approximate any nonstandard or user-specified areas, such as school districts or fire districts. Along with the MGRAs, another geographic boundary system used in the subregional allocation process is the 229 Land Use Zones (LUZ). In general, LUZs represent groups of census tracts. Within the City and County of San Diego, they adhere to community plan area boundaries. For the other jurisdictions, they adhere to current (2012) city boundaries. Figure 1-2 shows the LUZs used in the 2050 Regional Growth Forecast.

The MGRA replaces the MGRA that was used for previous SANDAG forecasts and estimates. The main advantage offered by these two new data definitions is that MGRAs remain static through time. Every LCKEY is assigned to a MGRA. If the political boundary that includes an LCKEY changes, it simply gets a new MGRA identifier. The result is more consistent comparisons of data between forecast periods and more accurate data within a forecast.

UDM uses land-use records input, but most of the modeling is accomplished at an aggregate geographic unit called the Land Use Zone (LUZ). LUZs are aggregations of MGRAs, that were built around practical and functional planning boundaries.

1.2 NOMENCLATURE

The Series 13 modeling system consists of hundreds of data elements, many program procedures and source code written in the Microsoft C#.net programming language. To make the technical

description of the model more readable, this discussion uses a shorthand notation for variables that is parallel to the actual model source code.

C#.net is an object-oriented programming language. In general, variables are stored in collections of data types called system classes. These are similar to structures in the C language which was first used to develop UDM and the other programs. Each class is defined according to the data that it represents and may contain a variety of different data types (integers, character strings, flags and so on). Classes may also be nested. That is, a class may contain other classes that comprise other classes. For example, the class "z." refers to all of the data for a Land Use Zone (LUZ) – one of the basic geographic units used in the modeling process. In practice, a general class like "z" also has a dimension representing the number of elements, in this case 229 LUZs. Class and element nomenclature is of the form class.element with the "." character being a separator. In the discussion that follows class dimensions are ignored for expositional simplicity; the elements of z[N] are represented by z.

For example, the class "site." identifies a collection of site-specific data. Within the "site" class there is another class "ac." that contains details about acreage. The "ac." class contains an array of 6 types of employment land uses represented by the symbol "ae[7]." This particular array is dimensioned to 7 because the C#.net arrays start at 0. In some cases, this apparent mismatch is used to retain the numbering sequence 1-n, rather than refer to the cases as 0 – n-1. Hence, the constructed variable **z.site.ac.ae₁** refers to the LUZ (z) site-specific (site) acreage (ac) for employment land (ae) type 1. Similarly, **z.site.ac.ae_{lute}** represents LUZ site-specific employment land uses for all types.

The main classes used in this technical description are shown below.

- z = LUZ data
- reg = regional totals
- rc = regional controls
- base = base year
- fcst = forecast year
- trans = transactions class
- site = site-specific data
- hs = housing stock
- hh = occupied units (households)
- capacity = housing and employment capacity
- ac = acres
- e = employment
- m = MGRA data

The algorithms discussed below also use C#.net programming notation for arithmetic operations with replacement. Specifically:

- + = : addition with replacement (a += b is the same as a = a + b);
- = : subtraction with replacement (a -= b is the same as a = a – b);
- * = : multiplication with replacement (a *= b is the same as a = a * b); and
- / = : division with replacement (a /= b is the same as a = a / b).

1.3 GENERALIZED MODEL STRUCTURE

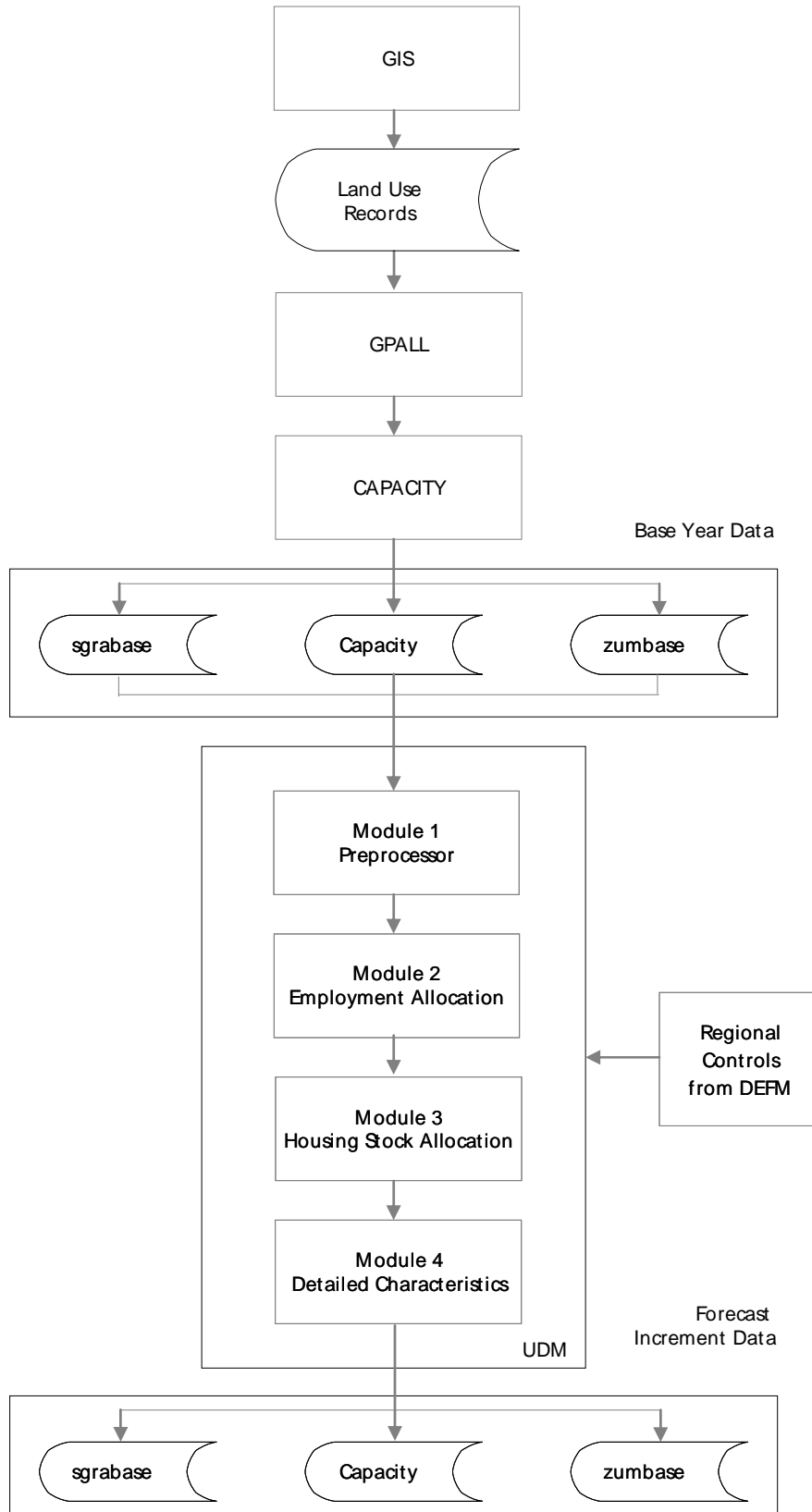
Figure 1 illustrates the generalized structure and sequence of the subregional forecast for the economic and demographic modeling components. The modeling steps accomplish one principal effect; they decide where and when housing and employment changes will occur on parcels, according to various land use criteria and planning assumptions.

The basic data structures of Series 13 are contained in several tables that are linked through a unique LCKEY. For simplification purposes, the entire process can be thought of as modifying a table of land use records. This table is initially derived from the GIS process described in a separate document. GPALL comprises a series of logic tests to assign a development code based on various existing and planned land use and planning criteria. The program CAPACITY computes holding capacity for employment and housing stock for each land use record based on the development code, the land use, an assigned density, and a forecast phase variable. At this point the table of land use records is called the **capacity** table. CAPACITY also builds summary tables that aggregate population, housing, employment and land use by acreage to MGRAs and LUZs.

UDM is the last of three programs that provide the subregional allocation. There are four main modules. Within UDM, the individual land use records of the capacity table are referred to as **transactions**. Module 1 is a preprocessor for the capacity table, building temporary data structures and tables that are used by the other modules. These include site-specific activity, capacity, and acreage information. Transaction or parcel-level site-specific accounting is done while accumulating LUZ information. Module 2 computes the civilian employment allocation for LUZs, which are then distributed to transactions. Within the housing stock module 3, UDM allocates multiple family units, followed by single family units, and then mobile homes. Within each housing type, UDM does a forecast for LUZs, whose forecasts are then allocated to transactions. Finally, module 4 computes detailed household characteristics, completes the allocation to MGRAs, and updates all output tables. All regionwise population, housing, and employment controls are derived from DEFM output.

This modular execution is possible because the principal output table, the capacity table, is updated as the model progresses. After Module 1, the capacity table reflects updated site-specific information. After Module 2, the capacity table also contains the updated employment information needed for the residential allocation. Similarly, after Module 3, the capacity table contains the updated employment and housing stock information needed for the other characteristics allocation in Module 4.

**Figure 1.
Subregional Modeling Flow of Information**



SECTION 2 – 11 MODEL COMPONENTS

2.1 GPALL

The GPALL program is a series of logic tests to assign a development type code (devCode) to each land use record according to the specific existing and planned land uses, the location and the planned phase for development. The devCode is used in the CAPACITY program and in UDM to determine where activity can occur during the forecast period. At this writing, there are 16 devCodes as listed in Table 1. The output of GPALL is a capacity table template that is the primary input to CAPACITY.

Input

1. Land-use input table
2. Default density parameters by sphere

Output

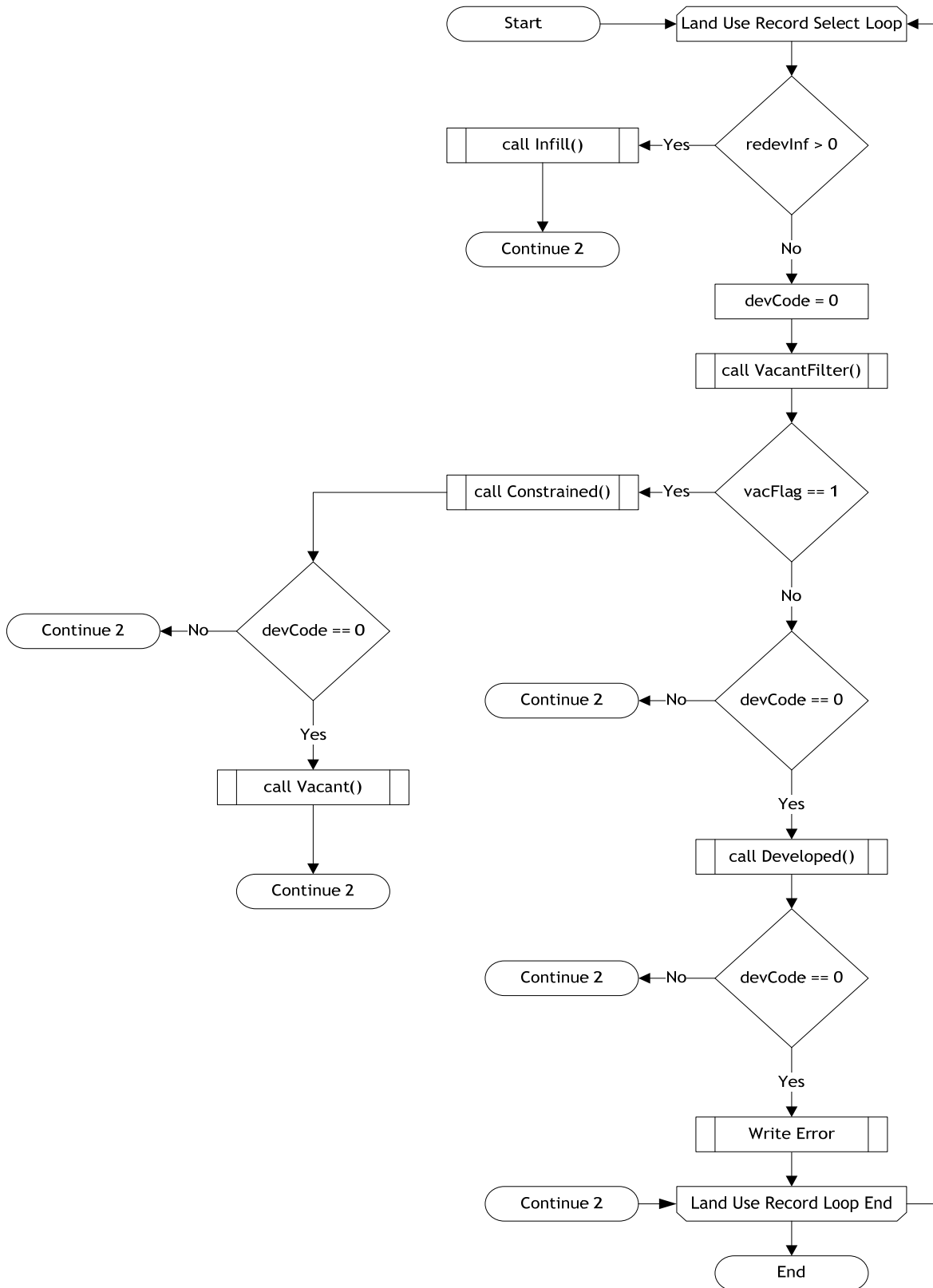
1. Initial base year capacity table with development code assignments

The basic GPALL logic is presented in Figure 2. The detail indicated in Figure 2 is presented in the multi-page diagram in Figure 3.

Table 1
Development Code Assignments

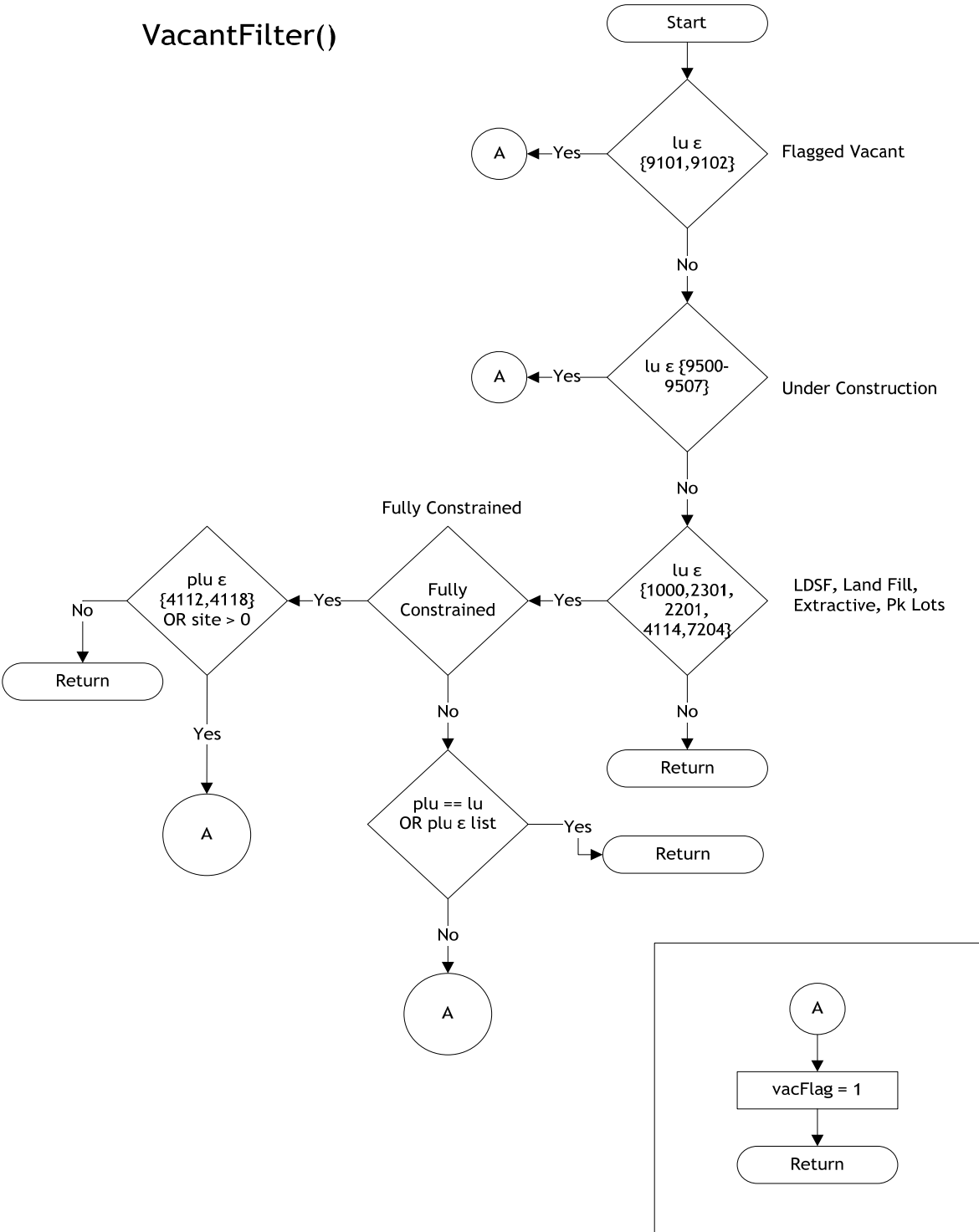
Code	Description
1	Developed
2	Constrained or Unusable
3	Vacant Developable
4	Employment Infill
5	Single Family Infill
6	Multi-Family Infill
7	Residential to Employment Redevelopment
8	Single Family to Multi-Family Redevelopment
9	Mobile Home to Other Residential Redevelopment
10	Agricultural Redevelopment
11	Employment to Residential Redevelopment
12	Employment to Employment Redevelopment
13	Residential to Road or Freeway
14	Employment to Road or Freeway
15	Employment or Residential to Mixed Use Redevelopment
16	Vacant to Mixed Use Development

**Figure 2
GPALL Flow of Information**

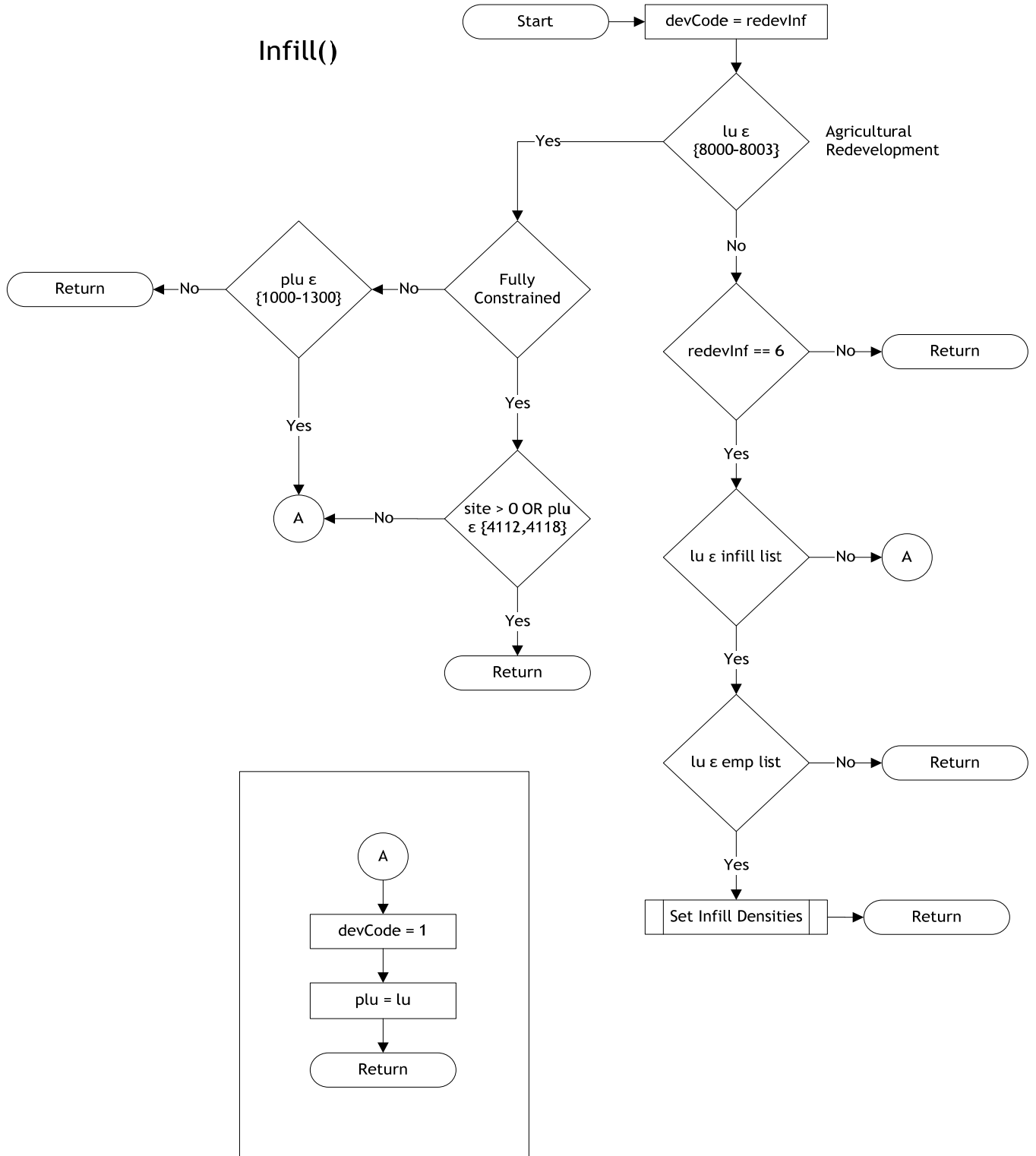


**Figure 3
GPALL Detail**

VacantFilter()

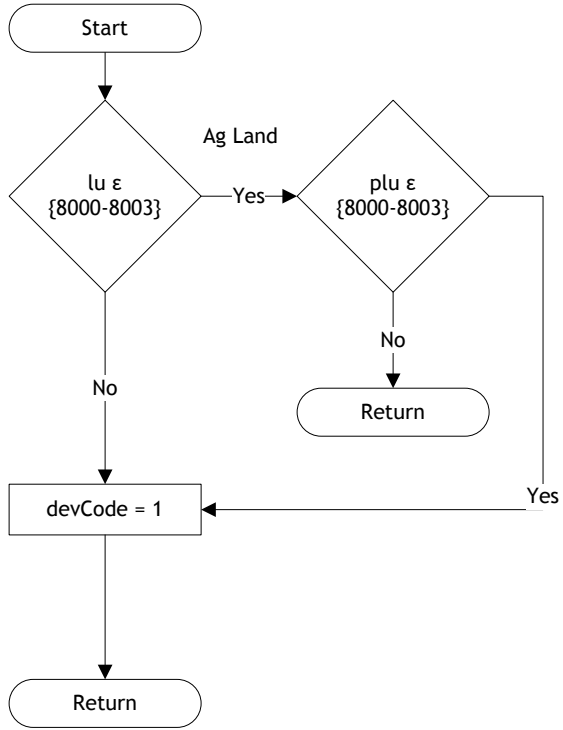


**Figure 3
GPALL Detail (cont'd)**

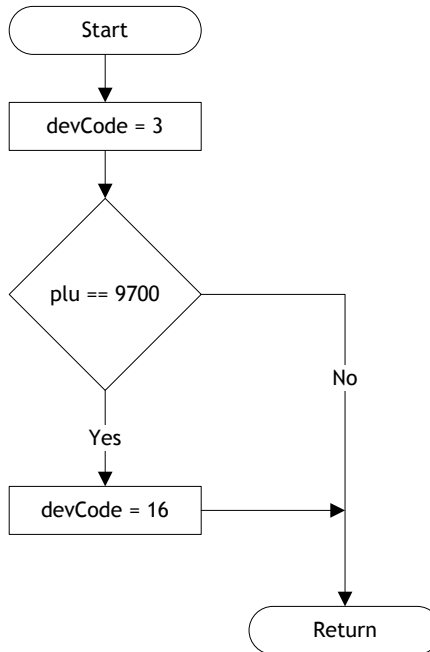


**Figure 3
GPALL Detail (cont'd)**

Developed()

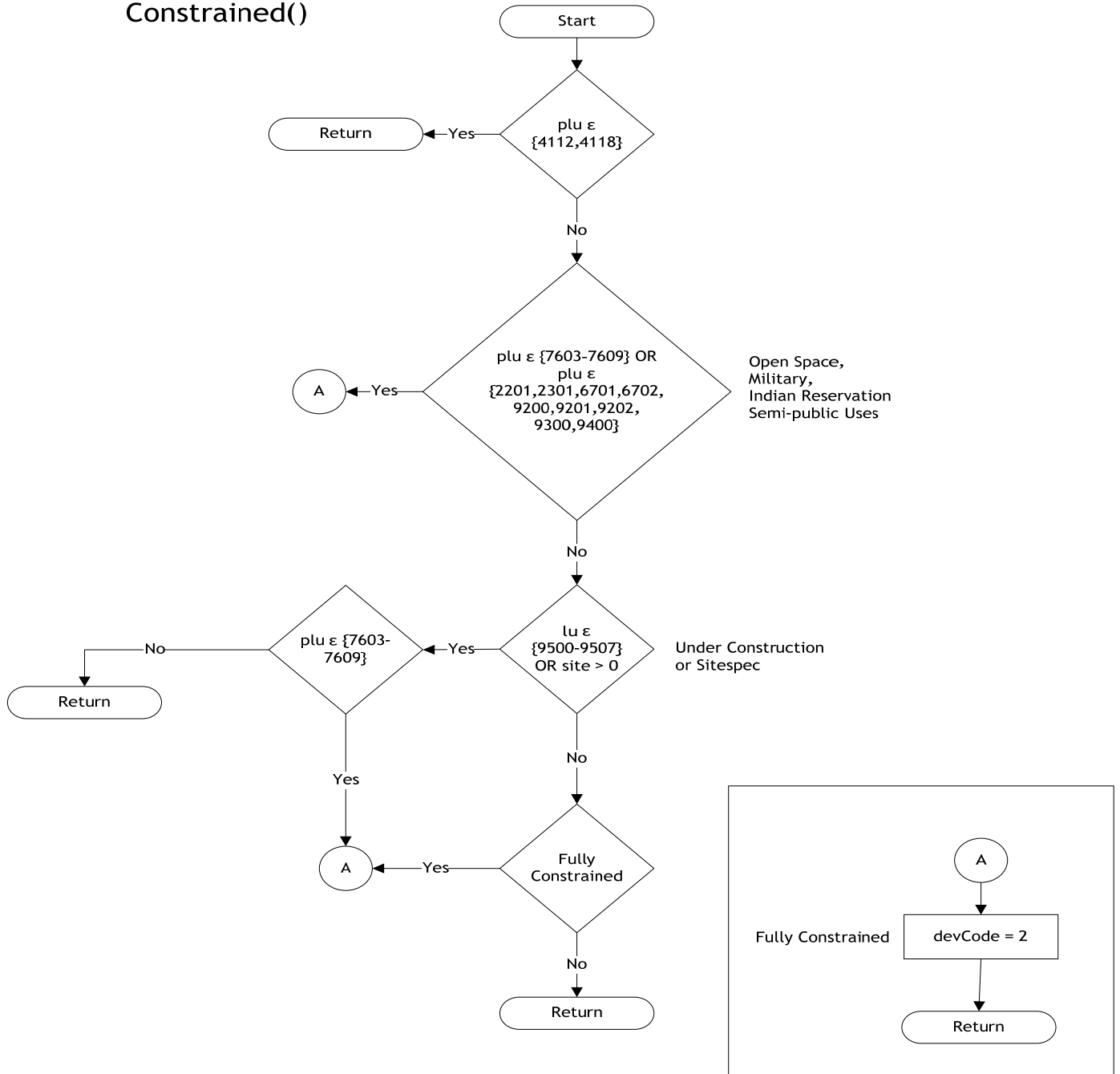


Vacant()



**Figure 3
GPALL Detail (cont'd)**

Constrained()



2.2 CAPACITY

The capacity table contains the land use input used in the subregional allocation. Each record of the capacity table, called a transaction, represents a unique land-use polygon (parcel or split parcel). Capacity is defined as the incremental (or decremental) number of jobs or housing units that a transaction can accommodate and is a function of the specific land use type, the development type and the associated employment or housing density. Each transaction also includes the base year employment and number of housing units by structure type. These are used to control the number of jobs or housing units that can be eliminated during employment or residential redevelopment activities.

Input

1. Initial capacity table
2. Default density parameters by sphere
3. Overrides for specific density or capacity changes
4. Employment densities
5. Base year population, housing and income distribution data
6. Modified site-specific table

Output

1. Base year capacity
2. Base year mgrabase
3. Base year LUZbase

PCAP

Every transaction also contains measures of the housing and employment capacity that has been developed. These values, PCAP_hs and PCAP_emp, permit proper accounting of acreage for partially developed transactions. At the beginning of a forecast increment, a transaction's PCAP values may range from 0.00 (total capacity available) to 1.00 (no capacity remaining). PCAP values between 0 and 1 indicate partially developed transactions (some capacity remaining).

Employment Density

Regional parameters are used to calculate employment capacities. Base year employment is sited to individual parcels and the employment densities are computed from regional averages derived by aggregating employment by land use across all parcels. Since employment densities are usually higher in the urban core, records within downtown San Diego are separated and computations are made for both datasets. There are other employment density overrides provided for the unincorporated areas of San Diego County. Employment and land use acres are aggregated to 14 employment categories and then summed for each land use type.

Residential Density

Default density assignments are used to calculate residential development for vacant and redevelopment/infill land. These values represent the point within the density range where development would occur. For instance, if the planned minimum and maximum densities are

3 units/acre and 6 units/acre, respectively, a .40 density range fraction means that development occurs at 40 percent of the density range or 4.2 units/acre. This fraction varies by community but the defaults are established at .50 for vacant and .75 for redevelopment/infill land. The CAPACITY program accommodates user-supplied overrides for residential densities

Site Specific Activity

The Sitespec table contains information for specific development sites, identified by a unique identifier called siteid. These sites logically contain many parcels. The total projected activity, employment and housing units, must therefore be distributed among the many land-use polygons that each site includes. CAPACITY does this distribution based the distribution of acreage among the composite parcels.

Capacity Calculations

In general, transactions with development codes (devCode) 3 through 16 are eligible for employment and/or housing unit capacity. Developed land and constrained land, devCodes 1 and 2, have employment and/or housing unit capacity of zero. There are exceptions as noted below. The capacity calculations are described according to development type. These are computations for records that do not involve site-specific development, which are described elsewhere.

- Vacant Employment Land (devCode = 3 and a planned employment land use). Vacant employment density capacity is calculated as the product of acres, employment density and the default vacant employment multiplier minus existing civilian employment. If the computed capacity is less than a minimum value of 10% of the existing civilian employment, capacity is reset to that minimum value.
- Vacant Residential Land (devCode = 3 and a planned residential land use). Potential vacant residential capacity is calculated as the product of acres and the appropriate residential density.
- Employment Infill (devCode = 4 or devCode = 1 with employment land uses and appropriate densities). Employment infill capacity is computed for a subset of the potential infill candidates. Furthermore, only a selected number of employment land uses are considered appropriate for infill activities. Also, there are some additional exclusions for infill land uses in downtown San Diego. Once the employment land use exclusions are eliminated, the base year employment density is calculated as the employment divided by acreage. This value is compared to the regional employment density. For developed land (devCode = 1 and employment land use), if the existing density is between 10% and 95% of the regional number, the transaction is allowed to infill up to the regional density. For records identified as employment infill (DevCode = 4), infill capacity is constrained to 50% of the base year employment and at least 10% of the base year employment. Any infill records with a computed capacity of 0 are reset to developed land status (devCode = 1). Any potential infill records that originate as developed land and have positive infill capacities are reset as infill records (devCode = 4).
- Single-family Infill (devCode = 5). Single-family infill capacity is computed as the product of (infill_density minus base year density) and acreage. For records with computed capacity ≤ 0 , the capacity is set to zero and the record is marked as developed land (devCode = 1).
- Multi-family Infill (devCode = 6). Multi-family infill capacity is computed as the product of (infill_density minus base year density) and acreage. For records with computed capacity ≤ 0 ,

the capacity is set to zero and the record is marked as developed land (devCode = 1). Furthermore, in this model, single-family units and mobile homes are lost during multi-family infill. The single-family and mobile homes capacities for multi-family infill records are set equal to the negative of the base year values for those structure types.

- Residential to Employment Redevelopment (dev-code = 7). Residential to employment redevelopment capacity is calculated as the product of acres and employment density minus existing civilian employment. Computed capacity is constrained to $\geq 10\%$ of the base year employment. For records with computed capacity > 0 , all residential units are removed; the structure capacities are set to the negative of the base year values.
- Single-family to Multi-Family Redevelopment (devCode = 8). Redevelopment density is computed as the low density range plus the product of the redevelopment multiplier and the density range (high density – low density). The multi-family capacity is computed as the product of the redevelopment density and acreage less the existing multi-family stock. Capacity was constrained $\geq 10\%$ of the base year stock. Any single-family units and mobile homes are removed; the associated capacities are set to the negative of the base year stock.
- Mobile Home to Residential Redevelopment (devCode = 9). Redevelopment density is computed as the low density range plus the product of the density range multiplier and the density range (high density – low density). The single-family (planned land use = 1100) or multi-family (planned land use = 1200) redevelopment capacity is computed as the product of the redevelopment density and acreage less the existing stock. Capacity was constrained $\geq 10\%$ of the base year stock. Mobile homes are removed; the capacity is set to the negative of the base year stock.
- Agriculture Redevelopment (devCode = 10). Agricultural redevelopment is handled exactly as vacant residential land or vacant employment land, depending on the planned land use.
- Employment to Residential Redevelopment (devCode = 11). Redevelopment density is computed as the low density range plus the product of the density range multiplier and the density range (high density – low density). The single-family (planned land use = 1100), multi-family (planned land use = 1200), or mobile home (planned land use = 1300) redevelopment capacity is computed as the product of the redevelopment density and acreage. For records with a computed residential capacity = 0, the record is marked as developed land.
- Employment to Employment Redevelopment (devCode = 12). The employment redevelopment potential is computed as the product of the employment density, acreage and the vacant employment land multiplier. Redevelopment capacity is computed as potential less existing employment. Capacity is constrained to $\geq 10\%$ of base year employment. For records with computed capacity < 0 , the capacity is set to zero and the record is marked as developed land.
- Residential to Road (devCode = 13). All existing housing stock is removed from this land use record (residential capacity is negative).
- Employment to Road (devCode = 14). All employment is removed from this land use record (employment capacity is negative).
- Employment or Residential to Mixed Use (devCode = 15). This is a mixed use record, (both employment and multifamily housing units). Capacity is computed as density * acres for both categories. Any existing single family units on mixed use land are removed.

- Vacant to Mixed Use (devCode = 16). This is a mixed use record, (both employment and multifamily housing units). Capacity is computed as density * acres for both categories.
- Site-Specific Records (Any devCode and SITE 10 > 0). Site-specific records, regardless of development code, are treated as a separate development class. First, any specified employment change, either positive or negative, is defined as the capacity for that record. Military employment change can only be entered as a site-specific change. Similarly, any projected changes in housing stock are defined as stock capacity. As site-specific record is processed, the planned land use (PLU) is set equal to the site land use (siteLU) and all residential densities are set to 0. Furthermore, for site-specific records defined as future roads or freeways, any existing employment is removed (the capacity is set equal to the negative of the existing employment). All existing housing stock is also removed for any site-specific records with vacant land (devCode = 3) or agricultural land (devCode = 10) converting to roads. All residential units are also removed for any residential to employment redevelopment (devCode = 7). Single-family units are removed for single-family to multi-family redevelopment (devCode = 8). Existing mobile homes are removed for mobile homes to residential redevelopment records (devCode = 9).

2.3 UDM MODULE 1 – PREPROCESS CAPACITY DATA

Introduction

The preprocessing of the capacity file accumulates information needed for the LUZ allocations and reports; that is, (1) site-specific activity, (2) employment capacities, (3) residential capacities; and 4) developable acres. It creates an output file used in the remaining UDM modules, which minimizes the number of times the capacity file must be accessed. In addition, module 1 also handles the transaction accounting of most the site-specific activities, except civilian and military group quarters population. Some equations show subscripts *lute*, *lutsf*, and *lutmf*, which denote the generalized code specifying the type of developable land for employment, single family, and multiple family activities.

Input

1. Base year capacity table

Output

1. LUZ site-specific
2. LUZ employment capacities
3. LUZ residential capacities
4. LUZ developable acreage
5. Intermediate capacity file, updated for site-specific activities

LUZ Site-Specific Activities, Capacities, and Developable Acres

In the capacity file, site-specific employment and housing unit change and capacities are stored in the same variables and a SITEID (trans.site) distinguishes between them. Some site-specific activity and capacity variables may contain negative values representing unit or employment losses due to redevelopment. Aggregations of site-specific activities and capacities involve transactions where

trans.phase is less than the forecast year (fyear) - they exclude areas designated for development in future forecast intervals. Developable acreage aggregations include all developable transactions regardless of the phase year; although, site-specific transactions dealing with future roads are excluded from its acreage aggregations. These general variables are: 1) site-specific activities; 2) employment and residential capacities; 3) developable employment acres (total and site-specific); and 4) developable residential acres (total and site-specific).

Aggregate to LUZs

Initialize all accumulation variables to zero

IF (trans.site \neq 0 (site-specific transactions))

IF (trans.site \neq 99) (accumulate site-specific developable acres, except for future roads)

z.site.ac.ae_{lute} += trans.acres_{lute} (civilian employment or CIVEMP)

z.site.ac.asf_{lutsf} += trans.acres_{lutsf} (single family or SF)

z.site.ac.amf_{lutmf} += trans.acres_{lutmf} (multi-family or MF)

IF (trans.phase < fyear) (accumulate site-specific activities)

z.site.civ += trans.cap_emp_civ (civilian employment)

z.site.mil += trans.cap_emp_mil (uniformed military employment)

z.site.sf += trans.cap_hs_sf (single family units)

z.site.mf += trans.cap_hs_mf (multi-family units)

z.site.mh += trans.cap_hs_mh (mobile homes)

z.site.gq_civ += trans.cap_gq_civ (civilian group quarters)

z.site.gq_mil += trans.cap_gq_mil (military group quarters)

IF (trans.site = 0 AND base.trans.devcode > 2) (accumulate non site-specific developable acres)

z.capacity.ac.ae_{lute} += trans.acres_{lute} (employment acres)

z.capacity.ac.asf_{lutsf} += trans.acres_{lutsf} (single family acres)

z.capacity.ac.amf_{lutmf} += trans.acres_{lutmf} (multi-family acres)

IF (trans.phase < fyear) (accumulate capacities)

z.capacity.e_{lute} += trans.cap_emp_civ_{lute} (civilian employment)

z.capacity.sf_{lutsf} += trans.cap_hs_sf_{lutsf} (single family units)

z.capacity.mf_{lutmf} += trans.cap_hs_mf_{lutmf} (multi-family units)

Create Additional LUZ Variables

z.capacity.ac.ae_{lute} += z.site.ac.ae_{lute} (developable acres, CIVEMP)

z.capacity.ac.asf_{lutsf} += z.site.ac.asf_{lutsf} (developable acres, SF)

z.capacity.ac.amf_{lutmf} += z.site.ac.amf_{lutmf} (developable acres, MF)

z.site.ac.tae = Σ z.site.ac.ae_{lute} (total site-specific developable acres, CIVEMP)

z.capacity.ac.tae = Σ z.capacity.ac.ae_{lute} (total developable acres, CIVEMP)

z.site.ac.tasf = Σ z.site.ac.asf_{lutsf} (total site-specific developable acres, SF)

z.capacity.ac.tasf = Σ z.capacity.ac.asf_{lutsf} (total developable acres, SF)

z.site.ac.tamf = Σ z.site.ac.amf_{lutmf} (total site-specific developable acres, MF)

z.capacity.ac.tamf = Σ z.capacity.ac.amf_{lutmf} (total developable acres, MF)

z.capacity.tote = Σ z.capacity.e_{lute} (total employment capacity)

z.capacity.totsf = Σ z.capacity.sf_{lutsf} (total single family capacity)

z.capacity.totmf = Σ z.capacity.mf_{lutmf} (total multi-family capacity),

where Σ is the sum over the land use types.

Create Regional Variables

reg.site.civ = $\Sigma z.site.civ$ (site-specific CIVEMP)
reg.site.mil = $\Sigma z.site.mil$ (site-specific uniformed military)
reg.site.sf = $\Sigma z.site.sf$ (site-specific SF)
reg.site.mf = $\Sigma z.site.mf$ (site-specific MF)
reg.site.mh = $\Sigma z.site.mh$ (site-specific mobile homes)
reg.site.gq_civ = $\Sigma z.site.gq_civ$ (site-specific civilian group quarters)
reg.site.gq_mil = $\Sigma z.site.gq_mil$ (site-specific military group quarters)
reg.capacity.e_{lute} = $\Sigma z.capacity.e_{lute}$ (civilian employment capacity)
reg.capacity.sf_{lutsf} = $\Sigma z.capacity.sf_{lutsf}$ (SF capacity)
reg.capacity.mf_{lutmf} = $\Sigma z.capacity.mf_{lutmf}$ (MF capacity)
reg.capacity.ac.ae_{lute} = $\Sigma z.capacity.ac.ae_{lute}$ (developable acres, CIVEMP)
reg.capacity.ac.asf_{lutsf} = $\Sigma z.capacity.ac.asf_{lutsf}$ (developable acres, SF)
reg.capacity.ac.amf_{lutmf} = $\Sigma z.capacity.ac.amf_{lutmf}$ (developable acres, MF)
reg.site.ac.ae_{lute} = $\Sigma z.site.ac.ae_{lute}$ (site-specific developable acres, CIVEMP)
reg.site.ac.asf_{lutsf} = $\Sigma z.site.ac.asf_{lutsf}$ (site-specific developable acres, SF)
reg.site.ac.amf_{lutmf} = $\Sigma z.site.ac.amf_{lutmf}$ (site-specific developable acres, MF),
where Σ is the sum over the LUZs

reg.capacity.ac.tae = $\Sigma reg.capacity.ac.ae_{lute}$ (total developable acres, CIVEMP)
reg.capacity.ac.tasf = $\Sigma reg.capacity.ac.asf_{lutsf}$ (total developable acres, SF)
reg.capacity.ac.tamf = $\Sigma reg.capacity.ac.amf_{lutmf}$ (total developable acres, MF)
reg.site.ac.tae = $\Sigma reg.site.ac.ae_{lute}$ (total site-specific developable acres, CIVEMP)
reg.site.ac.tasf = $\Sigma reg.site.ac.asf_{lutsf}$ (total site-specific developable acres, SF)
reg.site.ac.tamf = $\Sigma reg.site.ac.amf_{lutmf}$ (total site-specific developable acres, MF)
reg.capacity.tote = $\Sigma reg.capacity.e_{lute}$ (total employment capacity)
reg.capacity.totsf = $\Sigma reg.capacity.sf_{lutsf}$ (total single family capacity)
reg.capacity.totmf = $\Sigma reg.capacity.mf_{lutmf}$ (total multi-family capacity),
where Σ is the sum over the land use types

Write Site-Specific, Capacity, and Acres Information

1. LUZ, z.site.sf, z.site.mf, z.site.mh, z.site.civ, z.site.mil, z.site.gq_civ, z.site.gq_mil (site-specific)
2. LUZ, z.capacity.tote, z.capacity.e_{lute} (employment capacities)
3. LUZ, z.capacity.totsf, z.capacity.sf_{lutsf}, z.capacity.totmf, z.capacity.mf_{lutmf} (residential capacities)
4. LUZ, z.capacity.ac.tae, z.capacity.ac.ae_{lute}, z.site.ac.tae, z.site.ac.ae_{lute} (total and site-specific developable acres for employment)
5. LUZ, z.capacity.ac.tasf, z.capacity.ac.asf_{lutsf}, z.site.ac.tasf, z.site.ac.asf_{lutsf} (total and site-specific developable acres, SF)
6. LUZ, z.capacity.ac.tamf, z.capacity.ac.amf_{lutmf}, z.site.ac.tasf, z.site.ac.amf_{lutmf} (total and site-specific developable acres, SF)

Site-Specific Accounting in the Transactions

As noted earlier, the capacity is updated as the model progresses. The site-specific accounting is the first of five updates. The other four are: (1) civilian employment, (2) multiple family housing stock, (3) single family housing stock, and (4) mobile homes. These updates involve changes to several variables on the transaction record; for site-specific activities, there are no partially developed

transactions. Changes to site-specific transactions occur during the accumulation of the LUZ-level data. During the transaction processing, the base year employment and housing variables are updated and reflect the resulting forecast. In addition, the pcap variables, LU, development code, and capacity variables change. The labels **base.trans** (input file) and **fcst.trans** (updated file) clarify the capacity file updating process. In addition, the label **fcst.trans.chg** indicates the change variables contained in the capacity file, rather than having to define new variables. The updating process for the site-specific transactions ($\text{trans.site} \neq 0$ and $\text{trans.phase} < \text{fyear}$) is shown below.

Adjust Codes and Variables

$\text{fcst.trans.pcap_hs} = 1$ (percent housing acres developed)
 $\text{fcst.trans.pcap_emp} = 1$ (percent employment acres developed)
 $\text{fcst.trans.devcode} = 1$ (development code = developed)
 $\text{fcst.trans.lu} = \text{base.trans.plu}$ (base year land use code)
 $\text{fcst.trans.emp_civ} = \text{base.trans.emp_civ} + \text{base.trans.cap_emp_civ}$; (CIVEMP)
 $\text{fcst.trans.emp_mil} = \text{base.trans.emp_mil} + \text{base.trans.cap_emp_mil}$ (military employment)
 $\text{fcst.trans.hs_sf} = \text{base.trans.hs_sf} + \text{base.trans.cap_hs_sf}$ (single family)
 $\text{fcst.trans.hs_mf} = \text{base.trans.hs_mf} + \text{base.trans.cap_hs_mf}$ (multi-family)
 $\text{fcst.trans.hs_mh} = \text{base.trans.hs_mh} + \text{base.trans.cap_hs_mh}$ (mobile homes)
 $\text{fcst.trans.cap_emp_civ} = 0$ (employment capacity)
 $\text{fcst.trans.cap_emp_mil} = 0$ (uniformed military capacity)
 $\text{fcst.trans.cap_hs_sf} = 0$ (single family capacity)
 $\text{fcst.trans.cap_hs_mf} = 0$ (multi-family capacity)
 $\text{fcst.trans.cap_hs_mh} = 0$ (mobile home capacity)
Note: the site-specific military and civilian group quarters variables are not set to zero, because they are needed in module 4 to allocate the group quarters population to MGRAs.

Check For Negative Forecast

IF ($\text{fcst.trans.hs_xx} < 0$) $\text{fcst.trans.hs_xx} = 0$, for $\text{xx} = \text{sf, mf, mh}$ (Write warning message)
IF ($\text{fcst.trans.emp_civ} < 0$) $\text{fcst.trans.emp_civ} = 0$ (Write warning message)
IF ($\text{fcst.trans.emp_mil} < 0$) $\text{fcst.trans.emp_mil} = 0$ (Write warning message)

Create Change Variables

$\text{fcst.trans.chg_emp_civ} = \text{fcst.trans.emp_civ} - \text{base.trans.emp_civ}$ (CIVEMP)
 $\text{fcst.trans.chg_emp_mil} = \text{fcst.trans.emp_mil} - \text{base.trans.emp_mil}$ (uniformed military)
 $\text{fcst.trans.chg_hs_xx} = \text{fcst.trans.hs_xx} - \text{base.trans.hs_xx}$, for $\text{xx} = \text{sf, mf, mh}$ (housing stock)

2.4 UDM MODULE 2 – CIVILIAN EMPLOYMENT ALLOCATION

Introduction

Module 2 deals with the allocation of total civilian employment, first to LUZs and then to transactions in the capacity file. The LUZ allocation uses a non-linear spatial interaction equation while transactions are developed in order of their accessibility to activities. Land consumption and the loss of residential units to employment redevelopment is determined during the transaction processing. Some equations show subscript *lute* which denotes the generalized code specifying the type of developable land for employment. This code appears as a variable on the capacity file and Appendix B describes how they are determined.

Input

1. Base year LUZ- Population, Housing, Employment, Income
2. LUZ History- Civilian Employment, SF and MF Housing Stock
3. Updated Capacity File from module 1
4. Travel Times (Home to Work Peak A.M. Highway Trips)
5. Regionwide Employment Forecast
6. Kfactors
7. User Parameters and Overrides
8. LUZ Site-Specific, Employment Capacities, and Developable Acres
9. MGRA Accessibility Weights

Output

1. Intermediate capacity file, updated for site-specific activities and the civilian employment allocation
2. Kfactors that match the adjusted employment change

User Parameters

1. Minimum constraints switch (minsw) (0- no minimum constraint, 1- impose minimum constraint)
2. Minimum constraint parameter (minparm) specifies the percent below the base year value. The same minimum constraint is used for all LUZs. For example, 0 will not allow an employment loss and -5.0 indicates up to a five percent permissible loss of employment. Minparm cannot exceed zero.
3. Controlling switch (ctlparme) (0- exempt civilian employment overrides from controlling, 1- control civilian employment overrides)
4. Kfactor attenuation parameter (kfacparm) specifies the proportion of the Kfactor used (e.g., 0 means no Kfactor and 1 means the entire Kfactor)
5. Five parameters for the total civilian employment equation (ALPHA, BETA, employment attractor, employment capacity, lagged employment term)

Overrides

1. Civilian employment change (z.eo.adj)
2. Civilian employment change by land use type (z.eo.elulute)

Kfactors

Kfactors represent an additive adjustment to the total civilian employment equation. Initially derived during calibration, kfactors are updated during for each forecast period. The civilian employment overrides are defined as employment change excluding site-specific activity. There are two overrides. One controls the civilian employment change in a LUZ (z.eo.adj) and the other controls the allocation of employment to the six land use types (z.eo.elu_{lute}). The second override is used to alter land consumption patterns.

Base Year Activities

Variables: Civilian Employment (z.base.e.sectors_{sec}), Uniformed Military (z.base.e.mil) and Travel Times

$z.base.e.civ = \Sigma z.base.e.sectors_{secr}$, where Σ is the sum over all sectors (total civilian employment)

$z.base.e.tot = z.base.e.civ + z.base.e.mil$ (total employment)

Regional totals are created.

$reg.base.e.mil = \Sigma z.base.e.mil$ (uniformed military)

$reg.base.e.sectors_{sec} = \Sigma z.base.e.sectors_{secr}$ (civilian employment by sector)

where Σ is the sum over the LUZs

$reg.base.e.civ = \Sigma reg.base.e.sectors_{secr}$, where Σ is the sum over the sectors (total civilian employment)

$reg.base.e.tot = reg.base.e.civ + reg.base.e.mil$ (total employment)

Historical Employment

Variable: (z.hist_e.I5)

Historical trends in total civilian employment help: (1) judge the reasonableness of the forecast; (2) identify areas whose change is radically different from the past; and (3) insure a smooth trend to the forecast. The time points of the historical data are defined in reference to the forecast base year and, thus, change with each forecast period

LUZ-level changes are computed.

$z.hist_e.c5 = z.base.e.civ - z.hist_e.I5$ (prior 5-year numeric change)

$z.hist_e.pct5 = z.hist_e.c5 / z.hist_e.I5 * 100$ (prior 5-year percent change)

Regional totals are then created.

$reg.hist_e.c5 = \Sigma z.hist_e.c5$, where Σ is the sum across LUZs (prior 5-year numeric change)

Site-Specific Activities, Capacities, Developable Acres

Variables: Site-specific (z.site.civ, z.site.mil), Capacities (z.capacity.tote, z.capacity.e_{lute}), Developable Acres (z.capacity.ac.tae, z.capacity.ac.ae_{lute}, z.site.ac.tae, z.site.ac.ae_{lute})

Create regional totals for variables needed in module 2.

$reg.site.civ = \Sigma z.site.civ$ (civilian site-specific)

$reg.site.mil = \Sigma z.site.mil$ (military site-specific)

$reg.capacity.e_{lute} = \Sigma z.capacity.e_{lute}$ (civilian employment capacity, land use type)

where Σ is the sum over the LUZs

$reg.capacity.tote = \Sigma \Sigma z.capacity.e_{lute}$, where $\Sigma \Sigma$ is the double summation over LUZs and land use types (civilian employment capacity)

Regionwide Employment Forecast

$rc.fcst.e.civ = \sum rc.fcst.e.sectors_{sec}$, where Σ is the sum over all sectors (total civilian employment)

$rc.fcst.e.tot = rc.fcst.e.civ + rc.fcst.e.mil$ (total employment forecast)

$rc.fcst.e.adj = rc.fcst.e.civ - reg.site.civ$ (civilian employment adjusted for site-specific)

Control total changes are created.

$rc.fcst.ei.tot = rc.fcst.e.tot - rc.base.e.tot$ (total employment)

$rc.fcst.ei.civ = rc.fcst.e.civ - rc.base.e.civ$ (civilian employment)

$rc.fcst.ei.adj = rc.fcst.ei.civ - reg.site.civ$ (civilian employment adjusted for site-specific activities)

$rc.fcst.pct.civ = rc.fcst.ei.civ / reg.base.e.civ * 100$ (% change in civilian employment)

$rc.fcst.pct.tot = rc.fcst.ei.tot / reg.base.e.tot * 100$ ((% change in total employment)

Employment Share Forecast

A single equation develops the civilian employment forecast for a LUZ, excluding site-specific activity. This equation deals with the forecast endpoint and not the change in employment. UDM solves and processes a single equation. This equation produces the LUZ's share of the region's employment forecast, rather than the number of jobs. Its formulation is based on the EMPAL model developed by S.H. Putman and Associates, which is a modified version of a standard singly-constrained spatial interaction model. The modifications are: (1) a multivariate, multiparametric attractiveness function; (2) a lagged variable included outside the spatial interaction formulation; and (3) a procedure that allows for LUZ specific constraints. The non-linear equation class used in UDM is given below.

$z.fcst.share_j = \lambda * \sum_i z.base.hh_i * A_i * W_j * c_{i,j}^\alpha * \exp(\beta c_{i,j}) + (1 - \lambda) * z.base.e.civ$, where

$W_j = z.base.e.civ^a * z.capacity.e.etot^b$

$A_i = [\sum_i z.base.e.civ^a * z.capacity.e.etot^b * c_{i,j}^\alpha * \exp(\beta c_{i,j})]^{-1}$

i = origin LUZ

j = destination or workplace LUZ

l = notation for LUZ, regardless of origin or destination

c = travel time between origin and destination LUZs

$\lambda, \alpha, \beta, a,$ and k are empirically derived parameters

The civilian employment share forecast is a function of four factors: (1) base year households in the origin LUZ; (2) base year employment in the workplace LUZ; (3) capacity for additional civilian employment in the workplace LUZ; and (4) the travel time between the origin and workplace LUZs.

Five parameters are estimated for the equation. Two, a and b , are in the W or attractiveness term; b is associated with the employment capacity variable and a is related to the employment potential term or base year employment. λ indicates the importance of the base year employment location versus the spatial allocation term of the model. EMPAL uses a two parameter travel time function and the calibration parameters α and β define the shape of that curve. The A inverse term balances the equation. The balancing term insures a consistent allocation both regionally and for each LUZ. A simple example of a residential allocation illustrates the function of the A term. Assume that a given workplace LUZ has 100 employees and those employees are allocated to ten residential

destination LUZs. The balancing term guarantees that the employed residents generated from that workplace LUZ summed over the 10 destination LUZ equals 100.

Calibration

The CALIB program, also developed by S.H. Putman and Associates, estimates the five parameter values in the equation based on the 2000 to 2004 period. Because of the linear equation class in multiple regression, it is possible to solve directly for the best or optimal fit parameters (i.e., regression coefficients) using Ordinary Least Squares techniques. In nonlinear parameter estimation, however, a best fit for a set of parameters only exists if there is one optimum on the mathematical surface defined by the equations and data set. CALIB uses a gradient search method based on the partial derivative for each parameter to locate the optimum. The partial derivatives of the likelihood criterion function (L) are calculated with respect to each parameter in the equation. This criterion function defines a perfect fit between the observed and estimated employment that yields a value of $L = O$, which is the maxima being searched for on the gradient surface. The gradient of this likelihood function, when evaluated at a specific point on its surface, indicates the direction of the steepest slope on the surface from that point and the steepness of that slope. The maxima or optimum point is detected during CALIB's search of the surface when the partial derivative changes sign or reaches a relatively small value, say under 100. The best fit parameters occur when the optimum is found for each variable in the equation.

After CALIB performs the gradient search, it computes statistics that measure the equation's goodness-of-fit. In linear regression, best fit is measured by the R2 criterion and CALIB shows this measure for comparison purposes. The most appropriate measure of fit for UDM's equation class is the Best/Worst Likelihood Ratio (ϕ) derived from the maximum likelihood criterion. This measure relates three different likelihood quantities: (1) L is probability of observing the estimated civilian employment in a LUZ derived from the calibrated equation; (2) LW is the worst estimate and represents the probability of observing the mean of the civilian employment in every LUZ; and (3) LB is the best estimate in that it represents the observed employment in each LUZ. The Best/Worst Ratio is shown below

$$\phi = 1.0 - ((L - LB) / LW)$$

This ratio has a value of 1.0 for a model that perfectly estimates the observed value of employment (i.e., $L = LB$). For a model with less than a perfect fit, L is greater than LB, but less than LW. In this case, the fractional term value is less than 1.0 and falls between zero and one. CALIB also produces significance tests for the equation's parameters. T-tests are used in regression to determine statistical significance, but they are only valid if the assumptions underlying the equation class are met. These assumptions are not valid for UDM's nonlinear class, so the usual t-tests are not appropriate. Instead, CALIB calculates asymptotic t-tests that are interpreted in a manner similar to the more traditional t-tests. Although, the decision rules for asymptotic t-tests are not as precise and the associated probabilities are not exact.

Finally, k factors are computed that represent the absolute difference between the observed and estimated civilian employment shares for each LUZ.

$$z.kfac = (z.base.e.civ / rc.base.e.civ) - (z.base.e.civ(est.) / rc.base.e.civ)$$

Kfactors and Civilian Employment Change

Kfactors represent an additive adjustment to the employment share that corrects for errors in calibration. They can be altered by the user to reflect LUZs with unique locational properties. Kfactors are incorporated into the forecast.

$$z.fcst.ashare = z.fcst.share + (z.kfac * kfacparm)$$

Z.fcst.ashare is normalized to 1.0, replacing any negative shares with the base year share.

$$\text{IF } (z.fcst.ashare) < 0.0) \ z.fcst.ashare = z.base.e.civ / rc.base.e.civ$$

$$\text{TEMP} = \sum z.fcst.ashare, \text{ where } \sum \text{ is the sum over LUZs}$$

$$z.fcst.ashare = z.fcst.ashare * (1 / \text{TEMP})$$

The employment share is converted to an employment count by multiplying the share times the regional control total adjusted for site-specific activity.

$$z.fcst.e.adj = z.fcst.ashare * rc.fcst.e.adj$$

The civilian employment forecast is expressed as a change (increment or decrement) for the remainder of the processing.

$$z.fcst.ei.adj = z.fcst.e.adj - z.base.e.civ$$

Civilian Employment Overrides and Normalization

If there are no civilian employment overrides, the LUZ-level changes (z.fcst.ei.adj) are consistent with the regional site-specific adjusted control total, within rounding error. If overrides are used, the program checks their consistency (i.e., override cannot exceed capacity) and then normalizes the LUZ changes to the regional control.

$$\text{IF } (z.eo.adj > z.capacity.tote) \ z.eo.adj = z.capacity.tote \text{ (Write warning message when this occurs)}$$

$$z.fcst.ei.adj = z.eo.adj$$

A Plus-Minus controlling procedure adjusts the z.fcst.ei.adj to match the regional control total (rc.fcst.ei.adj). Rounding error after the application of the Plus-Minus routine is corrected using a random adjustment to the employment vector, adding or subtracting one until the control total is satisfied. If overrides to civilian employment are allowed to change during the controlling procedure (ctlparme = 1), the Plus-Minus routine is shown below.

$$\text{ABSUM} = \sum(\text{abs}(z.fcst.ei.adj)), \text{ where } \sum \text{ is the sum across LUZs and abs is the absolute value}$$

$$\text{SUM} = \sum z.fcst.ei.adj$$

$$\text{POSADJ} = (\text{ABSUM} + (\text{rc.fcst.ei.adj} - \text{SUM})) / \text{ABSUM}, \text{ where POSADJ is the adjustment for LUZs with employment growth}$$

$$\text{NEGADJ} = (\text{ABSUM} - (\text{rc.fcst.ei.adj} - \text{SUM})) / \text{ABSUM}, \text{ where NEGADJ is the adjustment for LUZs with employment loss}$$

IF (z.fcst.ei.adj \geq 0.0) z.fcst.ei.adj *= POSADJ
IF (z.fcst.ei.adj < 0.0) z.fcst.ei.adj *= NEGADJ

If the total overrides are exempt from the controlling procedure (ctlparme = 0), z.fcst.ei.adj is normalized using three modifications to Plus-Minus approach: 1) create ABSUM and SUM only from employment changes for the non-overridden LUZs; 2) create a regional sum of the employment changes from the overridden LUZs; and 3) subtract this regional sum from the regional control (rc.fcst.ei.adj) during the calculation of POSADJ and NEGADJ. The controlling factors are written.

POSADJ and NEGADJ sum to 2.0 and both must be positive. If (rc.fcst.ei.adj - SUM) is larger than ABSUM, one of the adjustment factors will be negative. If this condition occurs, the signs on z.fcst.ei.adj reverse when multiplied by the negative factor. Translation solves this problem by adding or subtracting a constant to the employment changes until both factors are positive. The program calculates a translation value to increase the negative adjustment factor to at least .10. Once the normalization is complete, the constant is subtracted from z.fcst.ei.adj to return to the scale of the original values.

Minimum and Maximum Constraints

UDM allows for two kinds of constraints: 1) increasing the employment change in LUZs that fall below a minimum threshold (minimum constraint); and 2) decreasing the employment change in LUZs that exceed their capacity (maximum constraint). While the imposition of a minimum constraint is optional, LUZs are always checked for their consistency with their maximum capacity.

Minimum Constraint Adjustment

The minimum constraint adjustment is used to lessen the impact of excessive employment loss. Its threshold value is defined with a user-defined regional parameter.

$$z.min_emp = minparm / 100 * z.base.e.civ, \text{ where } minparm \leq 0$$

The adjustment for the minimum constraint is invoked under the following conditions.

$$minsw = 1 \text{ AND } (z.fcst.ei.adj < z.min_emp)$$

The employment deficit is made up by lowering the employment change in the closest or target LUZ and increasing the change in the source LUZ. To avoid an excessive reduction in the employment change for any one target LUZ, the employment change in the target LUZ is lowered by 10% until the employment deficit in the source LUZ is made up or the target LUZ reaches its minimum constraint. If there is still an employment deficit, the program finds the next closest LUZ above its minimum constraint and repeats the adjustments. If all target LUZs are adjusted and the deficit still remains, the adjustment process is repeated, again starting with the closest target LUZ. These iterations continue until the deficit is eliminated. The closest target LUZ is defined as the shortest highway travel time from the source LUZ. This reallocation is restricted to LUZs in the urbanized area of the region. Other LUZs with a civilian employment override can receive reallocations if ctlparme = 1. For reporting purposes, the program stores the cumulative amount of employment added to the source LUZ and the amount of employment removed from the target LUZ.

Maximum Constraint Adjustment

The maximum constraint adjustment insures that a LUZ's employment increment does not exceed its holding capacity. This adjustment occurs when

$$z.fcst.ei.adj > z.capacity.tote.$$

The employment excess is removed by raising the employment change in the closest or target LUZ and lowering the change in the source LUZ. Employment change in the target LUZ is increased until all excess employment is reallocated or the target LUZ reaches capacity. If there are still excess jobs in the source LUZ, the program finds the next closest LUZ with capacity and repeats the process. This reallocation is restricted to LUZs in the urbanized area of the region. Other LUZs with a civilian employment override can receive reallocations if `ctlparme = 1`. For reporting purposes, the program stores the amount of employment removed from the source LUZ and the amount of employment added to the target LUZ.

Additional Forecast Variables

The true civilian and total employment change and forecast are computed.

$$\begin{aligned} z.fcst.ei.civ &= z.fcst.ei.adj + z.site.civ \text{ (civilian employment change)} \\ z.fcst.e.civ &= z.base.e.civ + z.fcst.ei.civ \text{ (civilian employment forecast)} \\ z.fcst.pct.civ &= z.fcst.ei.civ / z.base.e.civ * 100 \text{ (\% change in civilian employment)} \\ z.fcst.ei.tot &= z.fcst.ei.civ + z.site.mil \text{ (total employment change)} \\ z.fcst.e.tot &= z.base.e.tot + z.fcst.ei.tot \text{ (total employment forecast)} \\ z.fcst.pct.tot &= z.fcst.ei.tot / z.base.e.tot * 100 \text{ (\% change in total employment)} \end{aligned}$$

For reporting purposes, the program calculates variables representing positive and negative employment changes, excluding site-specific activity.

$$\begin{aligned} \text{IF } (z.fcst.ei.adj \geq 0) \text{ } z.gain &= z.fcst.ei.adj \\ \text{IF } (z.fcst.ei.adj < 0) \text{ } z.loss &= z.fcst.ei.adj \end{aligned}$$

Regional totals are created.

$$\begin{aligned} \text{reg.gain} &= \Sigma z.gain \\ \text{reg.loss} &= \Sigma z.loss, \\ \text{where } \Sigma &\text{ is the sum across LUZs} \end{aligned}$$

Kfactor Output

Output kfactors will likely differ from the input kfactors, because output kfactors reflect any user-defined overrides to civilian employment. They serve as inputs for the subsequent forecast period, which helps produce a more consistent forecast from increment to increment. Output kfactors also make it easier and quicker to do scenario evaluation. Output kfactors are produced in three steps: 1) create an adjusted (less site-specific activities) civilian employment forecast; 2) create the LUZ's share of the regional adjusted employment forecast; and 3) subtract the model share from the share determined in step 2. Finally, the output kfactors are written to an external data base.

$$z.temp.fcst = z.fcst.e.civ - z.site.civ \text{ (adjusted civilian employment forecast)}$$

$z.temp.share = z.temp.fcst / \Sigma z.temp.fcst$, where Σ is the sum over all LUZs (adjusted civilian employment forecast share)
 $z.fcst.kfac = z.temp.share - z.fcst.share$, where $z.fcst.share$ is the share from civilian employment equation

Distribution to Land Use Types

If $z.fcst.ei.adj > 0$, the employment increment is distributed to the six land use types ($z.fcst.ei.eland_{lute}$) according to their relative capacities, unless land use type overrides ($z.eo.elu_{lute}$) are specified.

$z.fcst.ei.eland_{lute} = z.capacity.e_{lute} * (z.fcst.ei.adj / z.capacity.tote)$, subject to the following constraints, which are satisfied using a ± 1 vector rounding and redistribution routine

$z.fcst.ei.eland_{lute} \leq z.capacity.e_{lute}$ AND
 $\Sigma z.fcst.ei.eland_{lute} = z.fcst.ei.adj$, where Σ is the sum over the land use types

If specified, the program implements land use type overrides.

IF ($z.eo.elu_{lute} > 0$) THEN
 $z.fcst.ei.eland_{lute} = z.eo.elu_{lute}$

Land use overrides also are subject to the above two constraints; that is, they cannot exceed the maximum constraint by land use type and should sum to the employment increment. The program checks the land use overrides. If either constraint is violated, the default capacities by land use type are not overridden. The program warns the user when the override was not implemented and which constraint was violated.

The remaining capacity and percent of capacity used is then computed.

$z.remcap.e_{lute} = z.capacity.e_{lute} - z.fcst.ei.eland_{lute}$ (remaining capacity, land use type)
 $z.remcap.tote = \Sigma z.remcap.e_{lute}$, where Σ is the sum over the land use types (total remaining capacity)
 $z.usecap.pe_{lute} = z.fcst.ei.eland_{lute} / z.capacity.e_{lute} * 100$ (% of capacity used, land use type)
 $z.usecap.ptote = z.fcst.ei.adj / z.capacity.tote * 100$ (% of total capacity used)

Regional totals are created.

$reg.remcap.e_{lute} = \Sigma z.remcap.e_{lute}$ (remaining capacity, land use type)
 $reg.fcst.ei.eland_{lute} = \Sigma z.fcst.ei.eland_{lute}$, where Σ is the sum across the LUZs (capacity used, land use type)
 $reg.usecap.pe_{lute} = reg.fcst.ei.eland_{lute} / reg.capacity.e_{lute} * 100$ (% of capacity used, land use type)
 $reg.usecap.ptote = rc.fcst.ei.adj / reg.capacity.tote * 100$ (% of total capacity used)

Outliers

An outlier or potentially unreasonable employment forecast is defined in reference to historical trends and to the relationship between change in a LUZ and the region. $Z.eout.r5$ captures these relationships by comparing two ratios: 1) the ratio of the forecast change to the 5-yr. historical

change in the LUZ; and 2) the same ratio for the region. The computation of the outlier ratio for LUZs is shown below.

$$\begin{aligned} \text{reg.hist_e.r5} &= \text{rc.fcst.ei.civ} / \text{reg.hist_e.c5} \text{ (region, forecast change to history change)} \\ \text{z.hist_e.r5} &= \text{z.fcst.ei.civ} / \text{z.hist_e.c5} \text{ (LUZ, forecast change to history change)} \\ \text{z.eout.r5} &= \text{z.hist_e.r5} / \text{reg.hist_e.r5} \text{ (LUZ to region ratio of changes)} \end{aligned}$$

A z.eout.r5 of 1 indicates that the forecast is perfectly consistent with its historical trend, while a z.eout.r5 less than 1 indicates that the LUZ forecast is lower than its historical trend. Conversely, a z.eout.r5 above 1 indicates that the LUZ forecast is higher than its historical trend.

LUZs with small employment changes in both the historical and forecast periods are not considered as potential outliers. Small employment changes are defined as ± 200 jobs. This criteria is illustrated using LUZ variables follows.

$$-200 \geq \text{z.fcst.ei.civ} \leq 200 \text{ AND } -200 \geq \text{z.hist_e.c5} \leq 200$$

Several rules determine outliers in LUZs (or aggregations areas) with larger employment changes. The first test examines whether the forecast change has a different sign than the 5-year historical change.

$$\begin{aligned} &\text{z.fcst.ei.civ} < 0 \text{ AND } \text{z.hist_e.c5} > 0 \text{) OR} \\ &\text{z.fcst.ei.civ} > 0 \text{ AND } \text{z.hist_e.c5} < 0 \end{aligned}$$

Areas defined by this criteria are given an outlier reason code (z.eout.outcode) of 1. z.eout.r5 is used to determine the status of the remaining LUZs. If the z.eout.r5 meets the criteria defined below, z.eout.outcode = 2.

$$\text{z.eout.r5} \leq .70 \text{ OR } \text{z.eout.r5} \geq 1.20$$

For analysis, an additional civilian employment change variable is computed for areas with an z.eout.outcode > 1. This estimate is based on the upper and lower threshold values, and the difference is found between it and the civilian employment change, including site-specific employment, produced by UDM.

IF (z.eout.outcode > 1)

$$\begin{aligned} \text{IF (z.eout.r5} < .70) \text{ z.eout.inc5} &= (\text{reg.hist_e.r5} * .70) * \text{z.hist_e.c5} \\ \text{IF (z.eout.r5} > 1.20) \text{ z.eout.inc5} &= (\text{reg.hist_e.r5} * 1.20) * \text{z.hist_e.c5} \\ \text{z.eout.diff5} &= \text{z.fcst.ei.civ} - \text{z.eout.inc5} \end{aligned}$$

Civilian Employment to Transactions

This part of module 2 allocates the LUZ-level civilian employment change, not including site-specific activities (z.fcst.ei.adj), to transactions. This allocation is done in sequence based on the accessibility of the transaction within a LUZ; that is, the transaction with the greatest accessibility is examined first and, if necessary, the transaction with the least accessibility is examined last. Transactions are not developed if the phase year (trans.phase) on the capacity file is greater than or equal to the forecast year input parameter (fyear). There are two main branches in the allocation of non-site-specific civilian employment to transactions - LUZ decrement and LUZ increment.

Employment Decrement

Employment losses generated from the model occur on the following developed land uses identified with the 4-digit base year land use code (LU).

- Industrial
- Commercial/Services
- Office
- Agricultural

Transactions with a base.trans.devcode of 1 (developed) or 4 (infill) are potential candidates for employment losses, however, site specific transactions cannot lose employment. Once a candidate transaction is identified, two checks are made: (1) LUZ control (LUZ_control) is > 0; and 2) the transaction's base year employment is > 0. If a transaction passes these checks, its base year civilian employment is reduced; otherwise go to the next transaction in sorted order. The reduction in base year employment is the minimum of the transaction's base year employment or remaining LUZ_control. When LUZ_control is the minimum, the employment decrement is completely allocated. The LUZ_control is adjusted and the employment capacity is increased by the amount of the employment loss. If all candidate transactions are processed and the LUZ_control is not completely allocated, the program writes a warning message showing the LUZ and unallocated decrement. Transactions that lose employment have their capacity increased so they can recoup the loss in subsequent forecast years. The base.trans.devcode is changed from 1 to a fcst.trans.devcode of 4 and the fcst.trans.pcap is set to zero or the fcst.trans.devcode is kept at 4 if the transaction was already classified as employment infill. The employment decrement algorithm is represented below.

```
DO IF (z.fcst.ei.adj < 0)
  LUZ_control = z.fcst.ei.adj * -1

IF (LUZ_control > 0 and base.trans.emp_civ > 0 and trans.site = 0) AND LAND USE TYPE OK AND
(base.trans.devcode = 1 or 4)) THEN

  TEMP = minimum(base.trans.emp_civ, LUZ_control)
  fcst.trans.emp_civ = base.trans.emp_civ - TEMP (civilian employment forecast)
  fcst.trans.chg_emp_civ = TEMP (civilian employment change)
  LUZ_control = LUZ_control - TEMP (adjust LUZ control)
  fcst.trans.cap_emp_civ = base.trans.cap_emp_civ + TEMP (add capacity)
  IF (base.trans.devcode = 1) (adjust codes when converting developed land)
    fcst.trans.devcode = 4 (development code)
    fcst.trans.pcap = 0.0 (% of acres developed)
  ELSE GO TO NEXT CANDIDATE TRANSACTION
  IF AFTER LAST TRANSACTION (LUZ_control > 0) WRITE warning message
```

Employment Increment

For LUZs with a civilian employment increment, transactions are developed in their accessibility order within the six land use type categories. The accessibility order determines the type of residential to employment redevelopment land consumed and whether development occurs on vacant and/or agricultural land. The accounting within each land use type is done in the same way. All of the transaction's capacity is developed, until the remaining LUZ increment (control) is less

than the capacity of a transaction. That transaction is partially developed and the `fcst.trans.pcap` (percentage of capacity used) is computed. If a partial record is created in a vacant school site (`lute = 6`), a warning message is written indicating that a partial record was created and the amount of the excess capacity. The `fcst.trans.pcap` determines acreage computations for the MGRABASE file and the activity lost in redevelopment areas. A running total is kept of units lost due to redevelopment by LUZ, and after the processing of all transactions it is written to a file for use in module 3.

If the employment capacity is less than or equal to the remaining LUZ increment (completely developed transaction),

1. Set the employment capacity to zero;
2. Set the `fcst.trans.pcap` and `fcst.trans.devcode` to 1; and
3. Update the base year employment, employment change, land use code, and developable acres used.

For partially developed transactions (employment capacity greater than the remaining LUZ increment);

1. Update the base year employment, employment change, land use code, employment capacity, and developable acres used;
2. Do not change the `trans.devcode`; and
3. Compute the `fcst.trans.pcap`.

In addition, for residential to employment redevelopment areas,

1. Update the base year housing units and housing change variables based on the demolished units; and
2. Update the maximum number units that can be demolished.

The computations for the civilian employment allocation follow. Note, in redevelopment areas (`lute = 1` and `base.trans.devcode = 7`) the maximum units that could be lost are stored as negative values in the residential capacity variables. For ease of presentation, the land use type subscript (`lute`) is not shown, but it is implied. `lute` is only used with acreage or capacity variables aggregated to LUZs that are stratified by land use categories.

```
DO IF (z.fcst.ei.eland > 0)
control = z.fcst.ei.eland (set LUZ control)
z.usecap.ac.aelute = 0 (developable acres used)
IF(base.trans.cap_emp_civ ≤ control AND base.trans.site = 0 AND base.trans.phase < fyear)
(completely developed transaction)
```

```
control = control - base.trans.cap_emp_civ (adjust LUZ control)
IF (control = 0) Allocation Complete
fcst.trans.pcap = 1 (% of capacity used)
fcst.trans.devcode = 1 (development code)
fcst.trans.LU = base.trans.plu (future land use code)
fcst.trans.emp_civ = base.trans.emp_civ + base.trans.cap_emp_civ (civilian employment
forecast)
```

fcst.trans.chg_emp_civ = base.trans.cap_emp_civ (civilian employment change)
z.usecap.ac.ae_{lute} += base.trans.acres (developable acres used)
fcst.trans.cap_emp_civ = 0 (civilian employment capacity)

IF (trans.lute = 1 and base.trans.devcode = 7) (residential to employment redevelopment)
IF (abs(base.trans.cap_hs_sf) > base.trans.hs_sf) base.trans.cap_hs_sf = base.trans.hs_sf * -1 (don't allow negative SF forecast)
IF (abs(base.trans.cap_hs_mf) > base.trans.hs_mf) base.trans.cap_hs_mf = base.trans.hs_mf * -1 (don't allow negative MF forecast)
IF (abs(base.trans.cap_hs_mh) > base.trans.hs_mh) base.trans.cap_hs_mh = base.trans.hs_mh * -1 (don't allow negative MH forecast)
fcst.trans.hs_sf = base.trans.hs_sf + base.trans.cap_hs_sf (SF forecast)
fcst.trans.hs_mf = base.trans.hs_mf + base.trans.cap_hs_mf (MF forecast)
fcst.trans.hs_mh = base.trans.hs_mh + base.trans.cap_hs_mh (MH forecast)
fcst.trans.chg_hs_sf = base.trans.cap_hs_sf (SF change)
fcst.trans.chg_hs_mf = base.trans.cap_hs_mf (MF change)
fcst.trans.chg_hs_mh = base.trans.cap_hs_mh (MH change)
fcst.trans.cap_hs_sf = 0 (all SF units removed)
fcst.trans.cap_hs_mf = 0 (all MF units removed)
fcst.trans.cap_hs_mh = 0 (all MH units removed)

IF (base.trans.cap_emp_civ > control AND base.trans.site = 0 AND base.trans.phase < fyear)
(Partially developed transaction, Allocation Complete)

fcst.trans.emp_civ = base.trans.emp_civ + control (civilian employment forecast)
fcst.trans.chg_emp_civ = control (civilian employment increment)
fcst.trans.cap_emp_civ = base.trans.cap_emp_civ – control (adjust capacity)
TEMP1 = base.trans.cap_emp_civ / (1 – base.trans.pcap) (employment capacity in the base year)
fcst.trans.pcap = 1 - (fcst.trans.cap_emp_civ / TEMP1) (% of developable acres used)
IF (base.trans.devcode = 4) fcst.trans.pcap = 0 (PCAP not computed for infill)
z.usecap.ac.ae_{lute} = z.usecap.ac.ae_{lute} + (base.trans.acres* (fcst.trans.pcap – base.trans.pcap)) (accumulate developable acres used)
IF (lute = 6) (vacant school site)
Write warning message showing LUZ and fcst.trans.cap_emp_civ

IF (lute = 1 and base.trans.devcode = 7) (residential to employment redevelopment)

TEMP2 = minimum of (((abs(base.trans.cap_hs_sf) / (1 – base.trans.pcap)) * (fcst.trans.pcap – base.trans.pcap)), base.trans.hs_sf) (SF units lost during increment, now a positive number)

TEMP3 = minimum of (((abs(base.trans.cap_hs_mf) / (1 – base.trans.pcap)) * (fcst.trans.pcap – base.trans.pcap)), base.trans.hs_mf) (MF units lost during increment, now a positive number)

TEMP4 = minimum of (((abs(base.trans.cap_hs_mh) / (1 – base.trans.pcap)) * (fcst.trans.pcap – base.trans.pcap)), base.trans.hs_mh) (MH units lost during increment, now a positive number)

$fcst.trans.hs_sf = base.trans.hs_sf - TEMP2$ (SF forecast)
 $fcst.trans.hs_mf = base.trans.hs_mf - TEMP3$ (MF forecast)
 $fcst.trans.hs_mh = base.trans.hs_mh - TEMP4$ (MH forecast)
 $fcst.trans.chg_hs_sf = TEMP2 * -1$ (SF change)
 $fcst.trans.chg_hs_mf = TEMP3 * -1$ (MF change)
 $fcst.trans.chg_hs_mh = TEMP4 * -1$ (MH change)
 $fcst.trans.cap_hs_sf = base.trans.cap_hs_sf + TEMP2$ (Adjust SF units lost)
 $fcst.trans.cap_hs_mf = base.trans.cap_hs_mf + TEMP3$ (Adjust MF units lost)
 $fcst.trans.cap_hs_mh = base.trans.cap_hs_mh + TEMP4$ (Adjust MH units lost)

Do When Last LUZ Complete

Additional LUZ variables are created.

$z.usecap.ac.ae_{lute} += z.site.ac.ae_{lute}$ (developable acres used, including site-specific, by land use type)
 $z.remcap.ac.ae_{lute} = z.capacity.ac.ae_{lute} - z.usecap.ac.ae_{lute}$ (developable acres remaining by land use type)
 $z.usecap.ac.tae = \sum z.usecap.ac.ae_{lute}$ (total developable acres used)
 $z.remcap.ac.tae = \sum z.remcap.ac.ae_{lute}$ (total developable acres remaining),
 where Σ is the sum across land use types

Regional totals are created.

$reg.capacity.ac.ae_{lute} = \sum z.capacity.ac.ae_{lute}$ (developable acres by land use type)
 $reg.usecap.ac.ae_{lute} = \sum z.usecap.ac.ae_{lute}$ (developable acres used by land use type)
 $reg.remcap.ac.ae_{lute} = \sum z.remcap.ac.ae_{lute}$ (developable acres remaining by land use type), where
 Σ is the sum across LUZs

 $reg.capacity.ac.tae = \sum z.capacity.ac.ae_{lute}$ (total developable acres)
 $reg.usecap.ac.tae = \sum reg.usecap.ac.ae_{lute}$ (total developable acres used)
 $reg.remcap.ac.tae = \sum reg.remcap.ac.ae_{lute}$ (total developable acres remaining)
 where Σ is the sum across land use types

2.5 MODULE 3 – HOUSING STOCK ALLOCATION

Introduction

Module 3 deals with the allocation of housing stock, first to LUZs and then to transactions in the capacity file. The LUZ allocation uses a non-linear spatial interaction equation known as PLUM, while transactions are developed in order of their accessibility to activities. Land consumption and the loss of single family and multiple family units to residential redevelopment is determined during the transaction processing. Module 3 first allocates multiple family units to determine any losses of single family and mobile home units. It next processes single family units to determine if any mobile home units are lost. Mobile homes are the last housing class type allocated. Module 3 equations use the terminology class.xx or class_xx, where xx = sf, mf, and mh for single-family, multi-family or mobile homes, respectively. Some variables do not include all class types. For these variables, the unit types are shown in parentheses following the definition. Some equations show subscripts lutsf

and lutmf, which denote the generalized code specifying the type of developable land for single family, and multiple family activities. These codes appear as a variable on the capacity file and Appendix B describes how they are determined.

Input

1. Base year LUZ
2. LUZ History
3. Updated Capacity File From module 2
4. Transportation
 - Work to Home p.m. Highway Trip Matrix
 - Work to Home p.m. Transit Trip Matrix
 - Proportion of Workers Using Transit
5. Regionwide Population and Housing Forecast
6. User Parameters and Overrides
7. LUZ site-specific and residential capacities

Output

1. Intermediate capacity file, updated for site-specific activities, civilian employment allocation, and housing stock allocation
2. LUZ rates: vacancy, persons per household, and employed residents per household

User Parameters

1. Controlling switch (ctlparmhs) (0- exempt housing stock overrides from controlling, 1- control housing stock overrides)
2. Travel time curve parameters for 3 trip types: (1) work to home, auto, (2) work to home, transit, and (3) work to home, transit (San Diego Central Business District (CBD)). Each trip type requires 3 parameters: (1) median travel time, (2) travel time standard deviation, and (3) non-linear adjustment for lognormal curve.

Overrides

1. Residential attractor for single and multiple family units (z.ho.attr_sf, z.ho.attr_mf)
2. LUZ housing stock overrides, single family, multiple family, mobile homes (z.ho.xx)
3. Single family housing units by land use type (z.ho.sflu_{lute})

Attractors and Housing Stock Overrides

UDM allocates single and multiple family housing stock according to a probability distribution based on travel time accessibility and the capacity for additional development, or attractors. Like former versions of the residential allocation model, these attractors can be overridden using z.ho.attr_sf or z.ho.attr_mf to alter the allocation of single and multiple family units, respectively. UDM also allows direct overrides to LUZ housing stock changes, including mobile homes (z.ho.xx). Housing stock overrides deal with problems resulting from excessive employment losses and reallocation of units from zones over capacity. The single-family land use override (z.ho.sflu_{lutsf}) is more specific and controls the single family land consumption patterns, especially the split between vacant urban and low density land.

Base Year Activities

Variables: Housing Stock (z.base.hs.xx), Employed Residents (z.base.p.er), Households (z.base.hh.xx), Household Population (z.base.p.hhp), and Transportation Factors

Housing stock and households by class type, employed residents, and household population are input. These variables create the vacancy rates and the employed residents per household rate used to convert employed residents to housing stock. Although not used in module 3, persons per household rates (base and forecast year) are developed for module 4.

$$\begin{aligned}z.base.hs.tot &= \sum z.base.hs.xx \text{ (total housing stock)} \\z.base.hh.tot &= \sum z.base.hh.xx, \text{ where } \Sigma \text{ is the sum over the class types (total occupied units)}\end{aligned}$$

Regional totals are created.

$$\begin{aligned}reg.base.p.hhp &= \sum z.base.p.hhp \text{ (household population)} \\reg.base.hs.xx &= \sum z.base.hs.xx \text{ (housing stock)} \\reg.base.hh.xx &= \sum z.base.hh.xx \text{ (households)} \\reg.base.p.er &= \sum z.base.p.er, \text{ where } \Sigma \text{ is the sum over the LUZs}\end{aligned}$$

$$\begin{aligned}reg.base.hs.tot &= \sum reg.base.hs.xx \text{ (total housing stock)} \\reg.base.hh.tot &= \sum reg.base.hh.xx, \text{ where } \Sigma \text{ is the sum over the class types (total occupied units)}\end{aligned}$$

Historical Housing Stock

Variable: z.hist_sf.I5, z.hist_mf.I5

Historical housing stock figures only pertain to single and multiple family class types. LUZ-level numeric and percent changes 5 years prior to the base year are computed. The five-year housing stock change (z.hist_xx.c5) is the difference between the base year and the beginning of the previous 5 year period.

$$\begin{aligned}z.hist_xx.c5 &= z.base.hs.xx - z.hist_xx.I5 \\z.hist_xx.pct5 &= z.hist_xx.c5 / z.base.hs.xx * 100 .\end{aligned}$$

Regional totals of the five year numeric changes are created.

$$reg.hist_xx.c5 = \sum z.hist_xx.c5, \text{ for } (xx = sf, mf), \text{ where } \Sigma \text{ is the sum across LUZs}$$

Site-Specific Activities, Capacities, and Developable Acres

Variables: Site-specific (z.site.xx), Capacities (z.capacity.totsf, z.capacity.sf_{Iutmf}, z.capacity.totmf, z.capacity.mf_{Iutmf}), and Developable Acres (z.capacity.ac.tasf, z.capacity.ac.asf_{Iutmf}, z.site.ac.tasf, z.site.ac.asf_{Iutmf}, z.capacity.ac.tamf, z.capacity.ac.amf_{Iutmf}, z.site.ac.tamf, z.site.ac.amf_{Iutmf}) (From module 1 Output Files)

For clarification, the attractor variables are defined as total capacity multiplied by a housing cost factor.

$$\begin{aligned}attrSF &= z.capacity.totalSF * z.homePriceIndex \\attrMF &= z.capacity.totalMF * z.homePriceIndex\end{aligned}$$

Finally, regional totals are developed.

$\text{reg.site.xx} = \Sigma \text{z.site.xx}$, where Σ is the sum over the LUZs (site-specific housing stock)
 $\text{reg.capacity.totsf} = \Sigma \Sigma \text{z.capacity.sf}_{\text{lutsf}}$ (single family capacity)
 $\text{reg.capacity.totmf} = \Sigma \Sigma \text{z.capacity.mf}_{\text{lutmfr}}$, where $\Sigma \Sigma$ is the double summation over LUZs and land use types (multiple family capacity)

LUZ Information from the Capacity File

Two pieces of LUZ information are accumulated from the capacity file's change variable. One is the civilian employment change and the other is the housing units lost to employment redevelopment activities. The employment change accumulates all transactions showing a change in employment, while units lost are only aggregated from non site-specific transactions. Any lost units due to site-specific activity are input to module 3. At this point in the model, housing stock change variables, in non site-specific transactions, are ≤ 0 , and unit losses only occur in transactions where the employment allocation required residential to employment redevelopment land.

$\text{z.fcst.ei.civ} = 0$ (civilian employment change)
 $\text{z.fcst.le.xx} = 0$ (units lost during residential to employment redevelopment)
 $\text{z.fcst.ei.civ} += \text{fcst.trans.chg_emp_civ}$
IF (trans.site = 0) $\text{z.fcst.le.xx} += \text{fcst.trans.chg_hs_xx}$

Regional totals for the lost units are created.

$\text{reg.fcst.le.xx} = \Sigma \text{z.fcst.le.xx}$, where Σ is the sum over the LUZs

Regionwide Forecast

Regional forecasts of housing stock and households by class type, employed residents, and household population, produced by the Demographic Economic Forecasting Model (DEFM), are input. Households and employed residents are used to compute vacancy rates and the employed residents per household rate, and the household population is used to compute persons per household. At this point, the correct multiple family change control total is established. The controls for single family units and mobile homes require additional adjustments for units lost due during the multiple family and single family allocations; they are adjusted later in the model.

$\text{rc.fcst.hsi.xx} = \text{rc.fcst.hs.xx} - \text{rc.base.hs.xx}$ (housing stock change)
 $\text{rc.fcst.hsi.tot} = \Sigma \text{rc.fcst.hsi.xx}$ (total housing stock change)
 $\text{rc.fcst.hs.tot} = \Sigma \text{rc.fcst.hs.xx}$, where Σ is the sum over the class types (total housing stock)
 $\text{rc.fcst.hsi.xx_adj} = \text{rc.fcst.hsi.xx} - \text{reg.site.xx} - \text{reg.fcst.le.xx}$ (housing stock change adjusted for site-specific and units lost to employment redevelopment)

Vacancy Rates, Employed Residents per Household, Persons per Household

Vacancy rates and employed residents per household rates are needed to convert the allocation of employed residents to housing stock. The forecasted rates in module 3 are interim and because of overrides might be different than the final rates used in module 4. While module 3 uses only vacancy rates for single and multiple family units and employed residents per household, mobile

home vacancy rates and persons per household are also developed in this routine. Base year and forecast year rates by LUZ and for the region are written for use in module 4.

The general forecasting procedure computes base year rates by LUZ and then adjusts them according to the regional change. Persons per household are forecast in a similar manner to the other rates, but also take into account that the rate of change in household size varies according to the LUZs household size relative to the region's household size. Rates for LUZs with little activity are generally unstable, while rates for LUZs with no activity cannot be computed. Some of these LUZs, however, may experience significant growth. So, regional rates are used for the forecast in LUZs with < 50 units in the base year.

Base Year Rates-LUZ

Initialize all base year rates to zero

IF (z.base.hs.xx \neq 0)

z.base.r.v_xx = (1 - (z.base.hh.xx / z.base.hs.xx)) * 100 (vacancy rates)

IF (z.base.hh.tot \neq 0)

z.base.r.er_hh = z.base.p.er / z.base.hh.tot (LUZ) (employed residents per household)

z.base.r.hhs = z.base.p.hhp / z.base.hh.tot (LUZ) (persons per household)

Base Year Rates-Region

IF (reg.base.hh.xx > 0)

reg.base.r.v_xx = (1 - (reg.base.hh.xx / reg.base.hs.xx)) * 100 (vacancy rates)

IF (reg.base.hh.tot > 0)

reg.base.r.er_hh = reg.base.p.er / reg.base.hh.tot (employed residents per household)

reg.base.r.hhs = reg.base.p.hhp / reg.base.hh.tot (persons per household)

Forecast Rates-Region

reg.fcst.r.v_xx = (1 - (reg.fcst.hh.xx / reg.fcst.hs.xx)) * 100 (vacancy rates)

reg.fcst.r.er_hh = reg.fcst.p.er / reg.fcst.hh.tot (employed residents per household)

reg.fcst.r.hhs = reg.fcst.p.hhp / reg.fcst.hh.tot (persons per household)

IF (reg.base.r.v_xx > 0) (adjustment factors)
 $v_factor_xx = reg.fcst.r.v_xx / reg.base.r.v_xx$ (vacancy rates)
 IF (reg.base.r.er_hh > 0)
 $er_factor = reg.fcst.r.er_hh / reg.base.r.er_hh$ (employed residents per household)
 IF (reg.base.r.hhs > 0)
 $hhs_factor = (reg.fcst.r.hhs / reg.base.r.hhs) - 1$ (persons per household)

Forecast Rates-LUZ

$z.fcst.r.v_xx = z.base.r.v_xx * v_factor_xx$ (vacancy rates)
 IF ($z.fcst.r.v_xx > 100$)
 $z.fcst.r.v_xx = 100$ (constrain the vacancy rate)
 $z.fcst.r.er_hh = z.base.r.er_hh * er_factor$ (employed residents per household)

IF ($hhs_factor \geq 0$) (persons per household)
 $z.fcst.r.hhs = z.base.r.hhs * (1 + (hhs_factor / (z.base.r.hhs / reg.base.r.hhs)))$
 IF ($hhs_factor < 0$)
 $z.fcst.r.hhs = z.base.r.hhs * (1 + (hhs_factor * (z.base.r.hhs / reg.base.r.hhs)))$
 The regional override is implemented for LUZs with less than 50 base year units.

IF ($z.base.hs.tot < 50$)
 $z.fcst.r.v_xx = reg.fcst.r.v_xx$
 $z.fcst.r.er_hh = reg.fcst.r.er_hh$
 $z.fcst.r.hhs = reg.fcst.r.hhs$
 Write message indicating LUZs where regional override was used.

Write output file (LUZs and region)

$z.base.r.v_xx$, $z.base.r.er_hh$, $z.base.r.hhs$,
 $z.fcst.r.v_xx$, $z.fcst.r.er_hh$, $z.fcst.r.hhs$,
 $reg.base.r.v_xx$, $reg.base.r.er_hh$, $reg.base.r.hhs$,
 $reg.fcst.r.v_xx$, $reg.fcst.r.er_hh$, $reg.fcst.r.hhs$

Residential Attractors

Residential attractors are the additional housing units that could be build in a LUZ. Separate attractors are calculated for single and multiple family housing units. Single family attractors (attr_sf) represent the sum of the capacities over the four land use types: (1) redevelopment, (2) infill, (3) vacant and agriculture-low density, and (4) vacant and agriculture-urban density. Multiple family attractors (attr_mf) represent the sum of the capacities over the three land use types: (1) redevelopment, (2) infill, and (3) vacant and agriculture. In both single family and multi-family, the residential attractors are adjusted by a residential price weighting, or attractiveness, factor that reflects the desirability of a neighborhood.

Probability Functions

The allocation of employees from their place-of-work to their place-of-residence is dependent upon the probability of working in one LUZ and living in that LUZ or some other LUZ. This probability function is a composite of accessibility, or interzonal and intrazonal travel times, and housing opportunities, or residential attractors. Accessibility is defined by LUZ-to-LUZ travel times grouped into 5-minute time bands according to a travel time probability distribution. A modified lognormal curve defines the travel time probability distribution as a function of three parameters: (1) median

travel time, (2) travel time standard deviation, and (3) a non-linear calibration factor. The calibration factor corrects for the fact that the standard lognormal curve underestimates short travel times and overestimates long travel times. Three probability functions representing travel modes are used in the allocation of residential activity: (1) work to home peak p.m. auto trips, (2) work to home peak p.m. transit trips from downtown San Diego, and (3) work to home peak p.m. transit trips not from downtown San Diego.

Two subroutines, PROB1 and PROB2, calculate the composite allocation probability matrix for the LUZs. PROB1 determines the probability of making the work to home trip in 5 minute time intervals, beginning with 0 to 5 minutes and ending with 110 to 115 minutes. For example, a 4 minute trip is assigned the probability of the 0 to 5 minute band based on the travel time probability distribution defined by the three parameters. A band's probability is defined by subtracting the cumulative probability of the lower time band and from the cumulative probability of the adjacent higher time band. For example, if the cumulative probabilities of making 30 and 35 minute work to home auto trips are .35 and .39, respectively, the probability of making a 30 to 35 minute home to work trip is .04 or .39 - .35. For the relatively few trips over 115 minutes, the very small probability is the subtraction of the cumulative probability of a 115 minute trip from 1.0.

PROB2 determines, from any work place or origin LUZ, the probability of a worker residing in any residential or destination LUZ. This probability is based on travel time accessibility weighted by the residential attractor, which can be overridden by z.ho.attr_sf or z.ho.attr_mf. PROB2 first groups the destination LUZs into five minute time bands based on the travel time probability function. This grouping uses a LUZ to LUZ travel time impedance matrix or skim tree. The initial allocation probability of each destination LUZ is computed by multiplying the residential attractor by the travel time probability of its band. This weighted holding capacity is accumulated as the multiplication is repeated for each LUZ. A LUZ's final allocation probability is normalized to 1.0 by dividing its weighted holding capacity by the regional accumulation of the weighted holding capacities.

These computations are performed from each workplace LUZ to that LUZ and all other LUZs, resulting in a LUZ-by-LUZ probability matrix that sums to one for each work place or origin LUZ. The allocation of employees to their places of residence uses probabilities for both auto (alloc_prob1) and transit (alloc_prob2) modes. Each set of probabilities is determined as previously described, except that transit riders use a impedance matrix derived from the transit network.

LUZ Multiple Family Housing Stock Forecast

Employed Residents

Separate allocations are done for employees living in single and multiple family houses. UDM computes the civilian employees living in multiple family units at the work place LUZ based on the multiple family share of the regionwide housing unit change. The change in employed residents at the workplace is assumed equal to the change in employment.

$$z.fcst.eri.tot = z.fcst.ei.civ$$

$$z.fcst.eri.mf_alloc = z.fcst.eri.tot * (rc.fcst.hsi.mf / rc.fcst.hsi.tot)$$

z.fcst.eri.mf_alloc is split into workers using transit or autos based on the LUZ-level proportion of workers using transit, or fractee.

z.fcst.eri.mf_transit = z.fcst.eri.mf_alloc * fractee (transit)
z.fcst.eri.mf_auto = z.fcst.eri.mf_alloc – z.fcst.eri.mf_transit (auto)

For each work place LUZ, the allocation of employees living in multiple family units to their residential LUZs is based on the weighted average of the workers by mode and their associated probabilities of living in the destination or residential LUZ.

DO FOR EACH WORKPLACE LUZ

dests_{r,w} = (z.fcst.eri.mf_auto_w * alloc_prob_{1,r}) + (z.fcst.eri.mf_transit_w * alloc_prob_{2,r}), where r is the destination or residential zone and w is the origin or workplace zone

These computations result in a LUZ by LUZ matrix containing the change of employed residents living in multiple family units in each residential (destination) LUZ based on workers at each employment (origin) LUZ. For example, element (150, 50) represents the change of employed residents living in multiple family units in LUZ 150 based on employees in LUZ 50. The final allocation of employed residents living in multiple family units for the destination LUZ accumulates the allocations from all work place LUZs.

er_inc_r = Σdests_{r,w}, where Σ is the sum across all origin LUZs denoted by the w subscript.

Housing Stock

An initial forecast is developed for multiple family housing stock change, not including site specific units and units lost due to employment redevelopment. This forecast is determined from the employed residents change, the multiple family vacancy rate forecast, and the employed residents per household rate.

z.fcst.hsi.mf_adj = er_inc / z.fcst.r.er_hh / (1 – z.fcst.r.v_mf)

If excessive employment losses occur, the employed residents allocation could be negative, which would imply a loss of housing units. In prior forecasts, unit losses were set to zero because the forecasting process only allows unit losses in areas explicitly designated for redevelopment. Consequently, these LUZs could never gain units during the forecast period without hard coding values into the program. Two procedures address this problem. First, negative unit changes are replaced by the prior 5 year change, adjusted for regional differences between the historical and forecasted periods.

IF (z.fcst.hsi.mf_adj < 0)
z.fcst.hsi.mf_adj = z.hist_mf.c5 * (rc.fcst.hsi.mf / reg.hist_mf.c5)
IF (z.fcst.hsi.mf_adj < 0)
z.fcst.hsi.mf_adj = 0 (prevent unit losses if 5 year history shows a decline)

Second, a direct override to the housing stock change can be invoked.

IF (z.ho.mf > z.capacity.totmf) (check for consistency with capacity)
z.ho.mf = z.capacity.totmf
Write warning to user
z.fcst.hsi.mf_adj = z.ho.mf

The plus-minus controlling procedure described in module 2 normalizes z.fcst.hsi.mf_adj to the regional control total (rc.fcst.hsi.mf_adj). If ctlparmhs = 0, overrides to the multiple family housing stock are not adjusted during the controlling procedure. The controlling factors are written out.

Maximum Constraints

After normalization, the multiple family unit change is checked for consistency with the maximum capacity. Reallocation is invoked if z.fcst.hsi.mf_adj > z.capacity.mf. The routine, described in module 2, reallocates multiple family units from LUZs that exceed capacity. Unlike the employment allocation routine, reallocations of housing units are not restricted to LUZs inside the urbanized area. LUZs with a multiple family housing stock override can only receive reallocations if the ctlparmhs = 1. For reporting purposes, the program stores the number of multiple family units removed from the source zone and the number added to the target zone.

Additional LUZ Variables

The true multiple family housing unit change and forecast are computed. The true change includes the UDM allocated change, site-specific activities, and units lost due to multiple family to employment redevelopment.

$$\begin{aligned} z.fcst.hsi.mf &= z.fcst.hsi.mf_adj + z.site.mf + z.fcst.le.mf \text{ (multiple family housing change)} \\ z.fcst.hs.mf &= z.base.hs.mf + z.fcst.hsi.mf \text{ (multiple family housing forecast)} \end{aligned}$$

Distribution to Land Use Types

Although the allocation of multiple family activity to land use types cannot be not overridden, like the civilian employment or single family allocations, it is more straight forward to handle the transaction allocations in the same general fashion; that is, specifying LUZ controls by land use type. Like for employment, this distribution of the multiple family increment is based on relative capacities.

$$z.fcst.hsi.mfland_{lutmf} = z.capacity.mf_{lutmf} * (z.fcst.hsi.mf_adj / z.capacity.totmf), \text{ subject to the following constraints, satisfied using a } \pm 1 \text{ vector rounding and redistribution routine}$$

$$\begin{aligned} z.fcst.hsi.mfland_{lutmf} &\leq z.capacity.mf_{lutmf} \text{ AND} \\ \sum z.fcst.hsi.mfland_{lutmf} &= z.fcst.hsi.mf_adj, \text{ where } \sum \text{ is the sum over the land use types} \end{aligned}$$

The remaining capacity by land use type is computed.

$$\begin{aligned} z.remcap.mf_{lutmf} &= z.capacity.mf_{lutmf} - z.fcst.hsi.mfland_{lutmf} \\ z.remcap.totmf &= \sum z.remcap.mf_{lutmf}, \text{ where } \sum \text{ is the sum over the land use types} \end{aligned}$$

Regional totals are then created.

$$\begin{aligned} reg.remcap.mf_{lutmf} &= \sum z.remcap.mf_{lutmf} \\ reg.fcst.hsi.mfland_{lutmf} &= \sum z.fcst.hsi.mfland_{lutmf}, \text{ where } \sum \text{ is the sum across the LUZs} \\ reg.remcap.totmf &= \sum reg.remcap.mf_{lutmf}, \text{ where } \sum \text{ is the sum over the land use types} \end{aligned}$$

Outliers

Using the same routine described in module 2, the outlier ratio, outlier criteria code, and historical trend forecast are developed for multiple family units. These variables are shown below.

z.mfout.r5 (multiple family outlier ratio)
z.mfout.outcode (multiple family outlier criteria code)
z.mfout.inc5 (multiple family housing stock change in line with 5 year history)
z.mfout.diff5 (difference between historical trend and model forecasts)

Allocate Multiple Family Housing Stock to Transactions

This part of module 3 allocates the LUZ-level multiple family housing stock change (z.fcst.hsi.mfland) to transactions, not including site-specific activities or units lost to employment development. The allocation uses essentially the same logic and updating procedures described in module 2, although it only involves positive multiple family unit changes. Transactions are developed in their accessibility order within the three land use type categories. The accessibility order determines the type of redevelopment land consumed (either single family or mobile homes) and whether development occurs on vacant and/or agricultural land. A running LUZ total is kept of single family or mobile home units lost due to redevelopment or infill. The computations for the multiple family allocation are shown below. Remember, in redevelopment areas and infill areas (lutmf = 1 and 2), the potential loss of single family and mobile home units are represented as negative values in their residential capacities.

Adjust Transaction Variables

control_{lutmf} = z.fcst.hsi.mfland_{lutmf}
z.fcst.lh.sf = 0 and z.fcst.lh.mh = 0 (accumulate SF and mobile home units lost by LUZ)
z.usecap.ac.amf_{lutmf} = 0 (accumulate developable acres used)

IF (z.capacity.mf_{lutmf} ≤ control_{lutmf} AND trans.site = 0 AND trans.phase < fyear) (completely developed transaction)

control_{lutmf} = control_{lutmf} - z.capacity.mf_{lutmf} (adjust LUZ control)

IF (control_{lutmf} = 0) Allocation Complete

fcst.trans.pcap = 1

fcst.trans.devcode = 1

IF (base.trans.lu ≠ base.trans.plu)

fcst.trans.lu = base.trans.plu

fcst.trans.hs_mf = base.trans.hs_mf + base.trans.cap_hs_mf

fcst.trans.chg_mf = base.trans.cap_hs_mf

z.usecap.ac.amf_{lutmf} += base.trans.acres

fcst.trans.cap_hs_mf = 0

IF (lutmf = 1 or 2) (redevelopment or infill)

IF (abs(base.trans.cap_hs_sf) > base.trans.hs_sf)

base.trans.cap_hs_sf = base.trans.hs_sf * -1

IF (abs(base.trans.cap_hs_mh) > base.trans.hs_mh)

base.trans.cap_hs_mh = base.trans.hs_mh * -1

fcst.trans.hs_sf = base.trans.hs_sf + base.trans.cap_hs_sf (single family)

fcst.trans.hs_mh = base.trans.hs_mh + base.trans.cap_hs_mh (mobile homes)

$fcst.trans.chg_sf = base.trans.cap_hs_sf$
 $fcst.trans.chg_mh = base.trans.cap_hs_mh$
 $z.fcst.lh.sf += base.trans.cap_hs_sf$
 $z.fcst.lh.mh += base.trans.cap_hs_mh$
 $fcst.trans.cap_hs_sf = 0$
 $fcst.trans.cap_hs_mh = 0$

IF (base.trans.cap_hs_mf > control_{lutmf} AND trans.site = 0 AND trans.phase < fyear) (partially developed transaction) (Allocation Complete)

$fcst.trans.hs_mf = base.trans.hs_mf + control_{lutmf}$
 $fcst.trans.cap_hs_mf = base.trans.cap_hs_mf - control_{lutmf}$
 $fcst.trans.chg_mf = control_{lutmf}$
 $temp1 = base.trans.cap_hs_mf / (1 - base.trans.pcap)$ (multiple family capacity in 2004)
 $fcst.trans.pcap = 1 - (fcst.trans.cap_hs_mf / temp1)$
 $z.usecap.ac.amf_{lutmf} += base.trans.acres * (fcst.trans.pcap - base.trans.pcap)$

IF (lutmf = 1 or 2) (redevelopment or infill)

$cap_diff = fcst.trans.pcap - base.trans.pcap$
 $cap_ratio = 1 - base.trans.pcap$
 $msf = abs(base.trans.cap_hs_sf) / cap_ratio * cap_diff$
 $mmh = abs(base.trans.cap_hs_mh) / cap_ratio * cap_diff$
 $temp_sf = \text{minimum of } (msf, base.trans.hs_sf)$ (SF units lost during increment, now a positive number)
 $temp_mh = \text{minimum of } (mmh, base.trans.hs_mh)$ (MH units lost during increment, now a positive number)

$fcst.trans.hs_sf = base.trans.hs_sf - temp_sf$
 $fcst.trans.hs_mh = base.trans.hs_mh - temp_mh$
 $fcst.trans.chg_sf = temp_sf * -1$
 $fcst.trans.chg_mh = temp_mh * -1$
 $z.fcst.lh.sf -= temp_sf$
 $z.fcst.lh.mh -= temp_mh$
 $fcst.trans.cap_hs_sf = base.trans.cap_hs_sf + temp_sf$
 $fcst.trans.cap_hs_mh = base.trans.cap_hs_mh + temp_mh$

DO WHEN LAST LUZ COMPLETE

Create Additional LUZ Variables

$z.usecap.ac.amf_{lutmf} += z.site.ac.amf_{lutmf}$ (developable acres used, including site-specific)
 $z.remcap.ac.amf_{lutmf} = z.capacity.ac.amf_{lutmf} - z.usecap.ac.amf_{lutmf}$ (developable acres remaining)
 $z.usecap.ac.tamf = \sum z.usecap.ac.amf_{lutmf}$ (total developable acres used)
 $z.remcap.ac.tamf = \sum z.remcap.ac.amf_{lutmf}$, where \sum is the sum across land use types (total developed acres remaining)

Create Regional Variables

$reg.fcst.lh.sf = \sum z.fcst.lh.sf$ (SF units lost to MF development)
 $reg.fcst.lh.mh = \sum z.fcst.lh.mh$ (MH units lost to MF development)
 $reg.capacity.ac.amf_{lutmf} = \sum z.capacity.ac.amf_{lutmf}$ (developable acres by land use type)
 $reg.usecap.ac.amf_{lutmf} = \sum z.usecap.ac.amf_{lutmf}$ (developable acres used by land use type)

$reg.remcap.ac.amf_{lutmf} = \Sigma z.remcap.ac.amf_{lutmf}$, where Σ is the sum across LUZs (developable acres remaining by land use type)

$reg.capacity.ac.tamf = \Sigma reg.capacity.ac.amf_{lutmf}$ (total developable acres)

$reg.usecap.ac.tamf = \Sigma reg.usecap.ac.amf_{lutmf}$ (total developable acres used)

$reg.remcap.ac.tamf = \Sigma reg.remcap.ac.amf_{lutmf}$, where Σ is the sum across land use types (total developable acres remaining)

LUZ Single Family Housing Stock Forecast

The material in this section essentially repeats the algorithms described for the multiple family allocation, although there are a few minor differences.

Employed Residents

The single family adjusted regional control total is adjusted a second time for any units lost during the multiple family allocation.

$$rc.fcst.hsi.sf_adj = rc.fcst.hsi.sf_adj - reg.fcst.lh.sf$$

Workers living in single family units are the difference between the employment change and workers living in multiple family units.

$$z.fcst.eri.sf_alloc = z.fcst.eri.tot - z.fcst.eri.mf_alloc$$

$z.fcst.eri.sf_alloc$ is further split into workers who use transit or autos.

$$z.fcst.eri.sf_transit = z.fcst.eri.sf_alloc * fractee (transit)$$

$$z.fcst.eri.sf_auto = z.fcst.eri.sf_alloc - z.fcst.eri.sf_transit (auto)$$

For each work place LUZ, the allocation of employees living in single family units to their residential LUZs is based on the weighted average of the workers by mode and their associated probabilities of living in the destination or residential LUZ.

DO FOR EACH WORKPLACE LUZ

$destr_{r,w} = (z.fcst.eri.sf_auto_w * alloc_prob_{1,r}) + (z.fcst.eri.sf_transit_w * alloc_prob_{2,r})$, where r is the destination or residential zone and w is the origin or workplace zone

These computations result in a LUZ-by-LUZ matrix containing the change of employed residents living in single family units in each residential (destination) LUZ based on workers at each employment (origin) LUZ. For example, element (150, 50) represents the change of employed residents living in single family units in LUZ 150 based on employees in LUZ 50. The final allocation of employed residents living in single family units for the destination LUZ accumulates the allocations from all work place LUZs.

$er_inc_r = \Sigma destr_{r,w}$ where Σ is the sum across all origin LUZs denoted by the w subscript.

Housing Stock

An initial forecast is developed for the single family housing stock change, not including site specific units or any unit losses due to the employment or multiple family allocations. This forecast is determined from the employed residents change, the single family vacancy rate forecast, and the employed residents per household ratio.

$$z.fcst.hsi.sf_adj = er_inc / z.fcst.r.er_hh / (1 - z.fcst.r.v_sf)$$

If the single family change is negative, it is replaced by the prior 5 year change adjusted for regional differences between the historical and forecasted periods.

```
IF (z.fcst.hsi.sf_adj < 0)
  z.fcst.hsi.sf_adj = z.hist_sf.c5 * (rc.fcst.hsi.sf / reg.hist_sf.c5)
IF( z.fcst.hsi.sf_adj < 0
  z.fcst.hsi.sf_adj = 0 (prevent unit losses if 5 year history shows a decline)
```

If specified, a direct override to the housing stock increment is implemented.

```
IF (z.ho.sf > z.capacity.totsf (check for consistency with capacity)
  z.ho.sf = z.capacity.totsf
  Write warning to user
  z.fcst.hsi.sf_adj = z.ho.sf
```

The plus-minus controlling procedure normalizes $z.fcst.hsi.sf_adj$ to the regional control total ($rc.fcst.hsi.sf_adj$). If $ctlparmhs = 0$, overrides to the single family housing stock are not adjusted during the controlling procedure. The controlling factors are written out.

Maximum Constraints

After normalization, the single family unit change is checked for consistency with the maximum capacity. This adjustment is invoked if $z.fcst.hsi.sf_adj > z.capacity.totsf$, using the same routine that reallocates multiple family units. LUZs with a single family housing stock override can only receive reallocations if the $ctlparmhs = 1$. For reporting purposes, the program stores the number of single family units removed from the source zone and the number added to the target zone.

Additional LUZ Variables

The true single family housing unit change and forecast are computed. The true change includes the UDM allocated change, site-specific activities, the units lost due to single family to employment redevelopment and during the multiple family housing stock allocation.

$$z.fcst.hsi.sf = z.fcst.hsi.sf_adj + z.site.sf + z.fcst.le.sf + z.fcst.lh.sf \text{ (single family housing change)}$$
$$z.fcst.hs.sf = z.base.hs.sf + z.fcst.hsi.sf \text{ (single family housing forecast)}$$

Distribution to Land Use Types

The distribution of the single family increment to land use types also is based on relative housing unit capacities. Unlike the multiple family distribution, the user can override the allocation of single family units to land use types.

$z.fcst.hsi.sfland_{lutsf} = z.capacity.sf_{lutsf} * (z.fcst.hsi.sf_adj / z.capacity.totsf)$, subject to the following constraints, satisfied using a ± 1 vector rounding and redistribution routine

$z.fcst.hsi.sfland_{lutsf} \leq z.capacity.sf_{lutsf}$ AND

$\sum z.fcst.hsi.sfland_{lutsf} = z.fcst.hsi.sf_adj$, where Σ is the sum over the land use types.

If specified, land use type single family unit increment overrides are implemented.

IF ($z.ho.sflu_{lutsf} > 0$)

$z.fcst.hsi.sfland_{lutsf} = z.ho.sflu_{lutsf}$

Land use overrides also are subject to the above two constraints; that is, they cannot exceed the maximum constraint by land use type and should sum to the adjusted single family housing unit increment. The program checks the land use overrides. If either constraint is violated, the default capacities by land use type are not overridden. The program warns the user when the override was not implemented and which constraint was violated.

The remaining capacity by land use type is then computed.

$z.remcap.sf_{lutsf} = z.capacity.sf_{lutsf} - z.fcst.hsi.sfland_{lutsf}$

$z.remcap.totsf = \sum z.remcap.sf_{lutsf}$, where Σ is the sum over the land use types

Regional totals are then created.

$reg.remcap.sf_{lutsf} = \sum z.remcap.sf_{lutsf}$

$reg.fcst.hsi.sfland_{lutsf} = \sum z.fcst.hsi.sfland_{lutsf}$, where Σ is the sum across the LUZs

$reg.remcap.totsf = \sum reg.remcap.sf_{lutsf}$, where Σ is the sum over the land use types

Outliers

Using the same routine described in module 2, the outlier ratio, outlier criteria code, and historical trend forecast are developed for single family units. These variables are shown below.

$z.sfout.r5$ (single family outlier ratio)

$z.sfout.outcode$ (single family outlier criteria code)

$z.sfout.inc5$ (single family housing stock change in line with 5 year history)

$z.sfout.diff5$ (difference between historical trends and model forecasts)

Allocate Single Family Housing Stock to Transactions

This part of module 3 allocates the LUZ-level single family housing stock change ($z.fcst.hsi.sfland$) to transactions, not including site-specific activities or units lost to employment development or during the multiple family allocation. The allocation uses essentially the same logic and updating procedures described for multiple family units. Transactions are developed in their accessibility order within the four land use type categories. The accessibility order determines whether development occurs on vacant and/or agricultural land. A running LUZ total is kept of mobile home units lost due to redevelopment. The computations for the single family allocation are shown below. Remember, in redevelopment areas ($lutsf = 1$), the potential loss of mobile home units is represented as a negative value in its residential capacity.

Adjust Transaction Variables

$control_{lutsf} = z.fcst.hsi.sfland_{lutsf}$

$z.fcst.lh.mh = 0$ (accumulate mobile home units lost by LUZ)

$z.usecap.ac.asf_{lutsf} = 0$ (accumulate developable acres used)

IF ($z.capacity.sf_{lutsf} \leq control_{lutsf}$ AND $trans.site = 0$ AND $trans.phase < fyear$) (completely developed transaction)

$control_{lutsf} = control_{lutsf} - base.trans.cap_hs_sf$ (adjust LUZ control)

IF ($control_{lutsf} = 0$) Allocation Complete

$fcst.trans.pcap = 1$

$fcst.trans.devcode = 1$

IF ($base.trans.LU \neq base.trans.plu$)

$fcst.trans.LU = base.trans.plu$

$fcst.trans.hs_sf = base.trans.hs_sf + base.trans.cap_hs_sf$

$fcst.trans.chg_sf = base.trans.cap_hs_sf$

$z.usecap.ac.asf_{lutsf} += base.trans.acres$

$fcst.trans.cap_hs_sf = 0$

IF ($lutsf = 1$) (mobile home redevelopment)

IF ($abs(base.trans.cap_hs_mh) > base.trans.hs_mh$)

$base.trans.cap_hs_mh = base.trans.hs_mh * -1$

$fcst.trans.hs_mh = base.trans.hs_mh + base.trans.cap_hs_mh$

$fcst.trans.chg_mh = base.trans.cap_hs_mh$

$z.fcst.lh.mh += base.trans.cap_hs_mh$

$fcst.trans.cap_hs_mh = 0$

IF ($base.trans.cap_hs_sf > control_{lutsf}$ AND $base.trans.site = 0$ AND $base.trans.phase < fyear$) (partially developed transaction) (Allocation Complete)

$fcst.trans.hs_sf = base.trans.hs_sf + control_{lutsf}$

$fcst.trans.cap_hs_sf = base.trans.cap_hs_sf - control_{lutsf}$

$fcst.trans.chg_sf = control_{lutsf}$

$temp1 = base.trans.cap_hs_sf / (1 - base.trans.pcap)$ (single family capacity in 2004)

$fcst.trans.pcap = 1 - (fcst.trans.cap_hs_sf / temp1)$

$z.usecap.ac.asf_{lutsf} += (base.trans.acres * (fcst.trans.pcap - base.trans.pcap))$

IF ($lutsf = 1$) (mobile home redevelopment)

$cap_diff = fcst.trans.pcap - base.trans.pcap$

$cap_ratio = 1 - base.trans.pcap$

$mmh = abs(base.trans.cap_hs_mh) / cap_ratio * cap_diff$

$temp_mh = \text{minimum of } (mmh, base.trans.hs_mh)$ (MH units lost during increment, now a positive number)

$fcst.trans.hs_mh = base.trans.hs_mh - temp_mh$

$fcst.trans.chg_mh = temp_mh * -1$

$z.fcst.lh.mh -= temp_mh$

$fcst.trans.cap_hs_mh = base.trans.cap_hs_mh + temp_mh$

DO WHEN LAST LUZ COMPLETE

Create Additional LUZ Variables

$z.usecap.ac.asf_{lutsf} += z.site.ac.asf_{lutsf}$ (developable acres used, including site-specific)
 $z.remcap.ac.asf_{lutsf} = z.capacity.ac.asf_{lutsf} - z.usecap.ac.asf_{lutsf}$ (developable acres remaining)
 $z.usecap.ac.tasf = \Sigma z.usecap.ac.asf_{lutsf}$ (total developed acres used)
 $z.remcap.ac.tasf = \Sigma z.remcap.ac.asf_{lutsf}$, where Σ is the sum across land use types (total developed acres remaining)

Create Regional Variables

$reg.fcst.lh.mh = \Sigma z.fcst.lh.mh$ (MH units lost to SF development)
 $reg.capacity.ac.asf_{lutsf} = \Sigma z.capacity.ac.asf_{lutsf}$ (developed acres by land use type)
 $reg.usecap.ac.asf_{lutsf} = \Sigma z.usecap.ac.asf_{lutsf}$ (developed acres used by land use type)
 $reg.remcap.ac.asf_{lutsf} = \Sigma z.remcap.ac.asf_{lutsf}$, where Σ is the sum across LUZs (developed acres remaining by land use type)

$reg.capacity.ac.tasf = \Sigma reg.capacity.ac.asf_{lutsf}$ (total developable acres)
 $reg.usecap.ac.tasf = \Sigma reg.usecap.ac.asf_{lutsf}$ (total developable acres used)
 $reg.remcap.ac.tasf = \Sigma reg.remcap.ac.asf_{lutsf}$, where Σ is the sum across land use types (total developable acres remaining)

Mobile Home Housing Stock Forecast

LUZ Forecast

The mobile regional control total is adjusted a second time for any units lost during the multiple family and single family allocations.

$$rc.fcst.hsi.mf_adj = rc.fcst.hsi.mf_adj - reg.fcst.lh.mh$$

The mobile home change for each LUZ is forecast by factoring adjusted base year mobile homes to the adjusted regional change. Base year mobile homes are adjusted for site-specific activity, unit losses in the current forecast period, and unit losses that could occur in subsequent forecast periods. This last adjustment ($z.fcst.pl.mh$) prevents changing mobile homes that may be demolished in later stages of the forecast. The transaction accounting process requires that mobile homes in these areas not be changed during the factoring process. $Z.fcst.pl.mh$ is created by aggregating the mobile home capacity variable ($fcst.trans.cap_hs_mh$) after processing the single family transactions.

$$\begin{aligned} temp10 &= z.base.hs.mh + z.site.mh + z.fcst.le.mh + z.fcst.lh.mh + z.fcst.pl.mh \text{ (adjusted base year)} \\ temp11 &= \Sigma temp10, \text{ where } \Sigma \text{ is the sum across LUZs (denominator of regional factor)} \\ z.fcst.hsi.mh_adj &= temp10 * rc.fcst.hsi.mf_adj / temp11 \end{aligned}$$

If there are no mobile home housing unit overrides, the allocation is complete; any rounding error is corrected by the ± 1 vector rounding routine. Otherwise implement the override; that is, ($z.fcst.hsi.mh_adj = z.ho.mh$) and re-control the change to ($rc.fcst.hsi.mf_adj$). Controlling is done with the same routine used in the single and multiple family housing stock allocations. If $ctlparmhs = 0$, overrides to the mobile home housing stock are exempted from adjustments during the controlling procedure. The controlling factors are written out.

Additional LUZ Variables

The true mobile housing unit change and forecast are computed. The true change includes the UDM allocated change, site-specific activities, the units lost due to mobile home to employment redevelopment and during the single and multiple family housing unit allocations.

$$\begin{aligned}z.fcst.hsi.mh &= z.fcst.hsi.mh_adj + z.site.mh + z.fcst.le.mh + z.fcst.lh.mh \\z.fcst.hs.mh &= z.base.hs.mh + z.fcst.hsi.mh\end{aligned}$$

Total housing stock variables are now created.

$$\begin{aligned}z.fcst.hsi.tot &= z.fcst.hsi.sf + z.fcst.hsi.mf + z.fcst.hsi.mh \\z.fcst.hs.tot &= z.base.hs.tot + z.fcst.hsi.tot\end{aligned}$$

Allocation of Mobile Homes to Transactions

If the adjusted mobile home change from each LUZ ($z.fcst.hsi.mh_adj$) is not zero, it is allocated to the transactions in proportion to their adjusted base year mobile homes, using the ± 1 Pachinko method. At this point, a transaction's mobile homes ($trans.fcst.hs_mh$) have been updated for site-specific activity and for units lost during the allocations of employment, single family units, and multiple family units. Some transactions with mobile homes are excluded from this allocation; that is, site-specific transactions ($trans.site > 0$), transactions in redevelopment areas ($trans.fcst.devcode \geq 6$ and ≤ 9), or transactions outside the phase (i.e., $phase \geq forecast\ year$). Therefore, within a LUZ, the sum of the eligible transactions' updated base year mobile homes will equal the LUZ's adjusted base year mobile homes ($temp10$). The Pachinko method uses a cumulative probability distribution (expressed as an integer from 1 to 100) computed from the transaction's share of the LUZ's adjusted base year mobile homes. The cumulative probability of the last eligible transaction in a LUZ should be 100. If not, a warning message is written.

In general, Pachinko distributes the mobile home change using a random number generated between 1 and 100. This random number in conjunction with the cumulative probability distribution determines the particular transaction to update. Depending on the direction of the mobile home change, the adjusted base year mobile homes are increased or decreased by 1, and the mobile home change variable is adjusted by the same amount. The procedure is repeated as many times as necessary to completely allocate the LUZ control. Pachinko gives essentially the same results as a simple raking procedure, but eliminates rounding error. Rounding error is usually a problem with small geographic areas containing relatively small activity counts.

2.6 MODULE 4 – DETAILED CHARACTERISTICS

Introduction

Module 4 uses the results of modules 2 and 3 as controls for the allocation of detailed characteristics. Again, it does the LUZs first and then allocates the LUZ forecast to the smaller level of geography, which is the MGRA. The detailed characteristics forecast in module 4 are: 1) occupied units by class type, household and group quarters population, household income distributions, and civilian employment for 14 NAICS groups. In addition, module 4 finished the update to the capacity file for civilian and military group quarters and also updates other key input files used in the next

forecast cycle (LUZBASE, MGRABASE, and LUZHIS). Module 4 equations use the terminology struct.xx or struct_xx, where xx = sf, mf, and mh for single-family, multi-family or mobile homes, respectively. Some variables do not include all class types.

Inputs

1. Base Year LUZ
2. Base Year MGRA
3. LUZ Site-Specific
4. LUZ History File
5. Intermediate Capacity File from module 3
6. Regionwide Population, Housing, and Employment Forecast
7. User Parameters and Overrides
8. Base Year and Forecast Rates
9. Employment Sector Distribution by Land Use Look-Up Table
10. Fixed Household Income Distribution Factors

Outputs

1. Final files: CAPACITY, MGRABASE, LUZBASE
2. Updated LUZ history file
3. Printed Reports

User Parameters

1. Rate controlling switch (ctlparmr) (0- exempt rate overrides from controlling, 1- control rate overrides)
2. Employment sector controlling switch (ctlparms) (0- exempt employment sector overrides from controlling, 1- control sector overrides)
3. Household income category controlling switch (ctlparmi) (0- exempt household income categories overrides from controlling, 1 – control category overrides)

Overrides

1. Vacancy rates by class type (z.ro.v_xx)
2. Employed resident and population per household rates (z.ro.er_hh, z.ro.hhs)
3. Household income distribution parameters (z.ro.income_median, z.ro.asd, z.ro.nla)
4. Employment sector increments (z.eso.sectors_adj_{sec})
5. Income distribution switch (z.ro.income_switch) determines the initial income distribution forecast, default is modified lognormal curve (1- use regional distribution, 2- use base year distribution, 3- use default)

LUZ Inputs

Overrides

Overrides influence the LUZ forecast of occupied units, household population, household income distribution, and civilian employment sectors. To simplify the programming, a complete set of overrides for each category of variables is always used; that is, (1) three vacancy rates, (2) employed

residents per household and household size, (3) three income parameters, and (4) all civilian employment sectors. Consistency checks are performed on the rate overrides here. Consistency checks for the employment overrides are done in the employment sector routine.

Vacancy rates

```
IF (z.ro.v_xx < 0)
  z.ro.v_xx = 0
ELSE IF (z.ro.v_xx > 50)
  z.ro.v_xx = 50
```

Write warning to user

Household size

```
IF (r.zo.hhs < 1.5)
  r.zo.hhs = 1.5
ELSE IF (r.zo.hhs > 4)
  r.zo.hhs = 4
```

Write warning to user

Employed residents per household

```
IF (z.ro.er_hh < 1)
  z.ro.er_hh = 1
ELSE IF (z.ro.er_hh > 2)
  z.ro.er_hh = 2
```

Write warning to user

Base Year Activities

Variables: Housing Stock (z.base.hs.xx), Households (z.base.hh.xx), Employed Residents (z.base.p.er), Household Population (z.base.p.hhp), Employment (z.base.e.sectors_{sec}, z.base.e.mil), Group Quarters (z.base.p.gq_civ, z.base.p.gq_mil), Household Income Distribution (z.base.i.hh_{inc}), Median Household Income (z.base.i.median)

Module 3 produced base and forecast year vacancy rates by class type, but some reports require a total vacancy rate. Base year total vacancy rates for LUZs and the region are computed here. The regional total vacancy rate forecast is computed from the regional inputs, and the LUZ forecast is done prior to checking the consistency of the forecast rates.

```
z.base.hs.tot = Σz.base.hs.xx (total housing stock)
z.base.hh.tot = Σz.base.hs.xx, where Σ is the sum over the class types (total occupied units)
z.base.p.gq = z.base.p.gq_civ + z.base.p.gq_mil (total group quarters (GQ))
z.base.e.civ = Σz.base.e.sectorssec, where Σ is the sum over the employment sectors (total civilian employment)
z.base.p.pop = z.base.p.hhp + z.base.p.gq (total population)
z.base.r.v = 0 (base year total vacancy rate)
IF (z.base.hs.tot > 0)
  z.base.r.v = (1 - (z.base.hh.tot / z.base.hs.tot)) * 100
```

Regional totals are created.

reg.base.p.pop = $\Sigma z.base.p.pop$ (total population)
reg.base.hs.tot = $\Sigma z.base.hs.tot$ (total housing stock)
reg.base.p.hhp = $\Sigma z.base.p.hhp$ (household population)
reg.base.hh.xx = $\Sigma z.base.hh.xx$ (occupied units by class type)
reg.base.p.gq_civ = $\Sigma z.base.p.gq_civ$ (civilian GQ)
reg.base.p.gq_mil = $\Sigma z.base.p.gq_mil$ (military GQ)
reg.base.i.hh_{inc} = $\Sigma z.base.i.hh_{inc}$ (household income distribution)
reg.base.e.sectors_{sec} = $\Sigma z.base.e.sectors_{sec}$ (civilian employment by sector)
reg.base.p.er = $\Sigma z.base.p.er$, where Σ is the sum over the LUZs (employed residents)
reg.base.hh.tot = $\Sigma reg.base.hh.xx$, where Σ is the sum over the class types (total occupied units)
reg.base.p.gq = reg.base.p.gq_civ + reg.base.p.gq_mil (total GQ)
reg.base.i.median = median_income(reg.base.i.hh_{inc}), where median_income is a function for computing median income (base year median income)
reg.base.r.v = $(1 - (reg.base.hh.tot / reg.base.hs.tot)) * 100$ (base year total vacancy rate)

Income Distribution Factors

Variables: Standard Deviation (z.asd), Exponent (z.nla), Adjustment to Base Year Distribution (z.fcst.income_adj)

Three parameters used to generate the LUZ income distribution do not change during the forecast; they are stored separately from the median income, which does change. Z.fcst.income_adj is the numeric difference between the observed base year proportionate income distribution and the estimated proportionate income distribution calibrated from the modified lognormal curve.

Site-Specific Group Quarters

Variables: (z.site.gq_civ, z.site.gq_mil)

The only site-specific variables not dealt with in the previous modules are civilian and military group quarters. Regional totals are created.

reg.site.gq_civ = $\Sigma z.site.gq_civ$
reg.site.gq_mil = $\Sigma z.site.gq_mil$, where Σ is the sum over the LUZs

Base and Forecast Year Rates, LUZ and Region

Base Year Variables: Vacancy Rates by Class Type (z.base.r.v_xx, reg.base.r.v_xx), Employed Residents per Household (z.base.r.er_hh, reg.base.r.er_hh), Persons Per Household (z.base.r.hhs, reg.base.r.hhs)

Forecast Year Variables: Vacancy Rates by Class Type (z.fcst.r.v_xx, reg.fcst.r.v_xx), Employed Residents per Household (z.fcst.r.er_hh, reg.fcst.r.er_hh), Persons Per Household (z.fcst.r.hhs, reg.fcst.r.hhs)

These forecast rates are interim and will change. The first change is the implementation of user overrides, shown below. In addition, the final rates reflect the normalized LUZ forecast of housing stock, occupied units, and household population.

z.fcst.r.v_xx = z.ro.v_xx (vacancy rate overrides)
z.fcst.r.er_hh = z.ro.er_hh (employed residents per household override)
z.fcst.r.hhs = z.ro.hhs (persons per household override)

BASE YEAR MGRA Inputs

Base year inputs for MGRAs are identical to LUZ inputs, except median income is not part of the MGRABASE. These inputs are read and stored for every MGRA. The variables are listed below.

housing stock, total and by class type (m.base.hs.tot, m.base.hs.xx)
occupied units, total and by class type (m.base.hh.tot, m.base.hh.xx)
employed residents (m.base.p.er)
household population (m.base.p.hhp)
civilian employment by sector (m.base.e.sectors)
uniformed military employment (m.base.e.mil)
civilian and military group quarters (m.base.p.gq_civ, m.base.p.gq_mil)
household income distribution (m.base.i.hh_{inc})

m.base.p.gq = m.base.p.gq_civ + m.base.p.gq_mil (total GQ)
m.base.e.civ = $\sum m.base.e.sectors_{sec}$ where Σ is the sum over the employment sectors (total civilian employment)

Housing Stock and Site-Specific Group Quarters

Housing stock changes are needed for both MGRAs and LUZs. In addition, MGRAs require military and civilian group quarters change from site-specific activities. For LUZs, these group quarters variables are inputs. In module 1, site-specific group quarters transactions were processed (i.e., fcst.trans.pcap, fcst.trans.devcode, and fcst.trans.LU were updated), but the group quarters information was not changed (i.e., set to zero). This routine sets these variables to zero, which completes the final update to the capacity file. Site-specific records are identified using the trans.site, trans.phase, and fyear codes.

DO FOR ALL TRANSACTIONS

z.fcst.hsi.xx = 0 (LUZ housing stock changes)
m.fcst.hsi.xx = 0 (MGRA housing stock changes)
m.site.gq_mil = 0 (MGRA site-specific military group quarters)
m.site.gq_civ = 0 (MGRA site-specific civilian group quarters)
m.fcst.hsi.xx += base.trans.hs_xx
z.fcst.hsi.xx += base.trans.hs_xx

IF (trans.site \neq 0 AND (trans.phase < fyear) AND (base.trans.gq_mil \neq 0 OR base.trans.gq_civ \neq 0))

m.site.gq_mil += base.trans.gq_mil
m.site.gq_civ += base.trans.gq_civ
fcst.trans.gq_civ = 0 (zero out civilian GQ variable)
fcst.trans.gq_mil = 0 (zero out military GQ variable)

Employment

Employment change is aggregated to MGRAs as three groups: (1) civilian site-specific; (2) civilian non-site-specific; and (3) uniformed military. (Note, regional changes in the uniformed military are handled entirely as site-specific activities.) This routine also distributes the site-specific employment change into the 14 civilian employment sectors. Site-specific civilian employment increments are distributed in each transaction and then the employment by sector is summed to MGRAs. The first step matches a transaction's proposed land use code to a look-up table containing a relative employment sector distribution for each 4-digit employment land use. The initial estimate of employment by sector is done by applying the relative distribution to the site-specific change. A ± 1 Pachinko is applied to correct rounding error. (The section documenting the allocation of household income to MGRAs describes the Pachinko method in detail.) Site-specific employment decrements are distributed to sectors after aggregation to MGRAs. Employment is removed according to the MGRA's base year relative employment sector distribution, using the same steps describe for site-specific employment increments. The MGRA employment aggregation routine is shown below.

DO FOR ALL TRANSACTIONS

```
m.fcst.ei.adj = 0 (MGRA civilian non site-specific employment change)
m.site.civ = 0 (MGRA site-specific total civilian employment change)
m.site.sectorssec = 0 (MGRA site-specific total civilian employment change by sector)
m.site.mil = 0 (site-specific uniformed military )
IF (trans.site = 0) (non site-specific transactions)
    m.fcst.ei.adj += fcst.trans.chg_emp_civ
IF (trans.site ≠ 0) (site-specific)
    m.site.mil += fcst.trans.chg_emp_mil
    m.site.civ += fcst.trans.chg_emp_civ
    IF (fcst.trans.chg_emp_civ > 0) (site-specific employment gain)
        Match fcst.trans.plu to look-up table and get distsec (share of employment by sector by 4
        digit land use code)
        emp_tempsec = distsec * fcst.trans.chg_emp_civ (initial distribution of site-specific
        employment gains into sectors)
        emp_tempsec = Pachinko(emp_tempsec) (correct rounding error)
        m.site.sectorssec =+emp_tempsec (final employment by sector)
        IF ( $\Sigma$ m.site.sectorssec ≠ fcst.trans.chg_emp_civ), where  $\Sigma$  is the sum across employment
        sectors
            Write error message
```

WHEN COMPLETE MGRA AGGREGATION

```
IF (m.site.civ < 0) (site-specific employment loss)
    m.base.edistsec = m.base.e.sectorssec / m.base.e.civ (base year relative sector distribution)
    emp_tempsec = m.base.edistsec * m.site.civ (initial distribution of site-specific employment
    losses into sectors)
    m.site.sectorssec = Pachinko(emp_tempsec)
    IF ( $\Sigma$ m.site.sectorssec ≠ m.site.civ) or (abs(m.site.sectorssec) > m.base.e.sectorssec), where  $\Sigma$  is
    the sum across employment sectors
        Write error message.
```

LUZ aggregations are created directly from the MGRAs.

$z.fcst.ei.adj = 0$ (total non site-specific civilian employment change)
 $z.site.sectors_{sec} = 0$ (site-specific civilian employment change by sector)
DO FOR ALL MGRAs
 $z.fcst.ei.adj += m.fcst.ei.adj$
 $z.site.sectors_{sec} += m.site.sectors_{sec}$

Finally, additional LUZ and regional totals are created.

$z.site.civ = \Sigma z.site.sectors_{sec}$, where Σ is the sum over the sectors (LUZ total site-specific change)
 $reg.site.civ = \Sigma z.site.civ$ (regional total site-specific civilian employment change)
 $reg.fcst.ei.adj = \Sigma z.fcst.ei.adj$ (regional total non site-specific civilian employment change)
 $reg.site.sectors_{sec} = \Sigma z.site.sectors_{sec}$, where Σ is the sum over the LUZs (regional site-specific civilian employment change by sector)

Regionwide Forecast

Variables: (rc.fcst.e.sectors_{sec}, rc.fcst.e.mil, rc.fcst.hh.xx, rc.fcst.p.er, rc.fcst.hs.tot, rc.fcst.p.hhp, rc.fcst.p.gq_civ, rc.fcst.p.gq_mil, rc.fcst.i.hh_{inc}, rc.fcst.i.median)

The form of the regional control total (i.e., totals or changes) varies in module 4. Households, household population, employed residents, and the income distribution are treated as totals. Civilian employment by sector and civilian and military group quarters are expressed as changes. This distinction is required for the allocation routines. Site-specific adjustments are only required for the employment by sector and group quarters controls because the LUZ-level housing stock forecasts, from which the household-related forecasts are derived, already include site-specific activities. The incremental controls are computed.

$rc.fcst.pi.gq_civ = rc.fcst.p.gq_civ - rc.base.gq_civ$ (civilian GQ change)
 $rc.fcst.pi.gq_civ_adj = rc.fcst.pi.gq_civ - reg.site.gq_civ$ (civilian GQ change adjusted for site-specific activity)
 $rc.fcst.pi.gq_mil = rc.fcst.p.gq_mil - rc.base.p.gq_mil$ (military GQ change)
 $rc.fcst.pi.gq_mil_adj = rc.fcst.p.gq_mil - reg.site.gq_mil$ (military GQ change adjusted for site-specific activity)
IF (rc.fcst.pi.gq_mil_adj \neq 0) Write Error Message

$rc.fcst.ei.sectors_{sec} = rc.fcst.e.sectors_{sec} - rc.base.e.sectors_{sec}$ (employment change by sector)
 $rc.fcst.ei.sectors_adj_{sec} = rc.fcst.ei.sectors_{sec} - reg.site.sectors_{sec}$ (employment change by sector adjusted for site-specific activity)
IF ($\Sigma rc.fcst.ei.sectors_adj_{sec} \neq reg.fcst.ei.adj$), where Σ is the sum across employment sectors
Write Error Message

Additional regional variables are calculated.

$rc.fcst.ii.median = rc.fcst.i.median - rc.base.i.median$ (median income change)
 $rc.fcst.ii.hhinc = rc.fcst.i.hhinc - rc.base.i.hhinc$ (household income distribution change)
 $rc.fcst.pi.hhp = rc.fcst.p.hhp - rc.base.p.hhp$ (household population change)
 $rc.fcst.pi.gq = rc.fcst.pi.gq_civ + rc.fcst.pi.gq_mil$ (total GQ change)
 $rc.fcst.p.gq = rc.fcst.p.gq_civ + rc.fcst.p.gq_mil$ (total GQ forecast)

$rc.fcst.p.pop = rc.fcst.p.hhp + rc.fcst.p.gq$ (total population forecast)
 $rc.fcst.pi.pop = rc.fcst.p.pop - rc.base.p.pop$ (total population change)
 $rc.fcst.hh.tot = \sum rc.fcst.hh.xx$, where Σ is the sum over the class types (occupied units forecast)
 $rc.fcst.hhi.tot = rc.fcst.hh.tot - rc.base.hh.tot$ (occupied units change)
 $rc.fcst.r.v = (1 - (rc.fcst.hh.tot / rc.fcst.hs.tot)) * 100$ (total vacancy rate forecast)
 $rc.fcst.pct.hhp = rc.fcst.pi.hhp / rc.base.p.hhp * 100$ (% change in household population)
 $rc.fcst.pct.hh = rc.fcst.hhi.tot / rc.base.hh.tot * 100$ (% change in occupied units)
 $rc.fcst.pct.gq = rc.fcst.pi.gq / rc.base.p.gq * 100$ (% change in total GQ)
 $rc.fcst.pct.pop = rc.fcst.pi.pop / rc.base.p.pop * 100$ (% change in total population)
 $rc.fcst.pct.sectors_{sec} = rc.fcst.ei.sectors_{sec} / rc.base.e.sectors_{sec} * 100$ (% change in civilian employment by sector)
 $rc.fcst.pct.income_median = rc.fcst.ii.median / rc.base.i.median * 100$ (% change in median income)
 $rc.fcst.pct.income_hh_{inc} = rc.fcst.ii.hh_{inc} / rc.base.i.hh_{inc} * 100$ (% change in the household income distribution)

LUZ Characteristics Forecast

Occupied Units by Class Type

Occupied units are done by class type and added together to get total occupied units, using the same procedure for each class type. First, the housing stock forecast is derived by adding changes from the capacity file to the base year value. Second, a class type specific vacancy rate forecast is applied to the housing stock forecast. Finally, the initial forecast is factored to the regional control using a raking procedure and the controlling factor is written. If the controlling switch (ctlparmr) is zero, LUZs with user overrides to the vacancy rates are not adjusted during the controlling process. A ± 1 vector rounding routine corrects rounding errors, insuring the occupied units do not exceed total units. The routine shown below assumes that all LUZs are controlled (ctlparmr = 1).

$z.fcst.hs.xx = z.base.hs.xx + z.fcst.hsi.xx$ (housing stock forecast)
 $z.fcst.hh.xx = z.fcst.hs.xx * (1 - (z.fcst.r.v_{xx} / 100))$ (occupied units)
 $lsum_xx = \sum z.fcst.hh.xx$, where Σ is the sum over the LUZs
 $hhfac_xx = rc.fcst.hh.xx / lsum_xx$ (controlling factor)
 $z.fcst.hh.xx = z.fcst.hs.xx * hhfac_xx$ (controlled occupied units)
 IF ($z.fcst.hh.xx > z.fcst.hs.xx$)
 $z.fcst.hh.xx = z.fcst.hs.xx$

APPLY VECTOR ROUNDING ROUTINE

If override zones are exempt from controlling, the adjustments to the routine are: (1) lsum_xx only includes LUZs without overrides; (2) a new regional variable accumulates the households from the overridden LUZs; and (3) this new variable is subtracted from the regional control (rc.fcst.hh.xx) prior to calculating the factor.

Additional LUZ variables related to households and housing stock are developed.

$z.fcst.hhi.xx = z.fcst.hh.xx - z.base.hh.xx$ (occupied units by class type change)
 $z.fcst.hs.tot = \sum z.fcst.hs.xx$ (total housing stock forecast)
 $z.fcst.hh.tot = \sum z.fcst.hh.xx$ (total occupied units forecast)

$z.fcst.hhi.tot = \sum z.fcst.hhi.xx$, where Σ is the sum over the class types (total occupied units change)
 $z.fcst.pct.hh = z.fcst.hhi.tot / z.base.hh.tot * 100$ (% change in total occupied units)

Household Population, Group Quarters, Employed Residents, and Total Population

Household population and employed residents are computed from the normalized occupied units forecast by multiplying it by the persons per household and employed residents per household forecasts, respectively. The controlling process for these variables is the same used for the occupied units and these controlling factors are written. UDM insures that at least one person lives in every household; that is, the household population forecast is greater than or equal the occupied units forecast.

UDM also forecasts both civilian and military group quarters. Changes to military group quarters are completely reflected in the site-specific data.

$$z.fcst.pi.gq_mil = z.site.gq_mil$$

The non site-specific civilian group quarters change is developed in the same way as the mobile home housing stock change, using a raking procedure in combination with the ± 1 vector rounding routine.

$z.base.p.gq_civ_extra = z.base.p.gq_civ + z.site.gq_civ$ (base year plus site-specific activity)
 $lsum_gq = \sum z.base.p.gq_civ_extra$, where Σ is the sum across LUZs (denominator of regional adjustment factor)
 $pt = z.base.p.gq_civ_extra * (rc.fcst.pi.gq_civ_adj / lsum_gq)$ (adjusted civilian group quarters change)

Additional LUZ level variables related to population are developed.

$z.fcst.pi.gq_civ = pt + z.site.gq_civ$ (civilian GQ change)
 $z.fcst.pi.gq = z.fcst.pi.gq_civ + z.fcst.pi.gq_mil$ (total GQ change)
 $z.fcst.p.gq_civ = z.base.p.gq_civ + z.fcst.pi.gq_civ$ (civilian GQ forecast)
 $z.fcst.p.gq_mil = z.base.p.gq_mil + z.fcst.pi.gq_mil$ (military GQ forecast)
 $z.fcst.p.gq = z.fcst.p.gq_civ + z.fcst.p.gq_mil$ (total GQ forecast)
 $z.fcst.pi.hhp = z.fcst.p.hhp - z.base.p.hhp$ (household population change)
 $z.fcst.p.pop = z.fcst.p.hhp + z.fcst.p.gq$ (total population forecast)
 $z.fcst.pi.pop = z.fcst.pi.hhp + z.fcst.pi.gq$ (total population change)
 $z.fcst.pct.hhp = z.fcst.pi.hhp / z.base.p.hhp * 100$ (% change in household population)
 $z.fcst.pct.gq = z.fcst.pi.gq / z.base.p.gq * 100$ (% change in total GQ)
 $z.fcst.pct.pop = z.fcst.pi.pop / z.base.p.pop * 100$ (% change in total population)

Compute Final Forecast Rates

Vacancy rate, persons per household, and employed residents per household forecasts are recomputed from the normalized data.

$z.fcst.r.v_xx = 0$ (vacancy rate by class type)
 $z.fcst.r.v = 0$ (total vacancy rate)

z.fcst.r.er_hh = 0 (employed residents per household)

z.fcst.r.hhs = 0 (persons per household)

IF (z.fcst.hs.xx > 0)

z.fcst.r.v_xx = (1 - (z.fcst.hh.xx / z.fcst.hs.xx)) * 100

IF (z.fcst.hs.tot > 0)

z.fcst.r.v = (1 - (z.fcst.hh.tot / z.fcst.hs.tot)) * 100

IF (z.fcst.hh.tot > 0)

z.fcst.r.er_hh = z.fcst.p.er / z.fcst.hh.tot

z.fcst.r.hhs = z.fcst.p.hhp / z.fcst.hh.tot

z.fcst.ri.v_xx = z.fcst.r.v_xx - z.base.r.v_xx (vacancy rate by class type change)

z.fcst.ri.v = z.fcst.r.v - z.base.r.v (overall vacancy rate change)

z.fcst.ri.hhs = z.fcst.r.hhs - z.base.r.hhs (persons per household change)

z.fcst.ri.er_hh = z.fcst.r.er_hh - z.base.r.er_hh (employed residents per household change)

z.fcst.pct.v_xx = z.fcst.ri.v_xx / z.base.r.v_xx * 100 (% change in vacancy rate by class type)

z.fcst.pct.v = z.fcst.ri.v / z.base.r.v * 100 (% change in total vacancy rate)

z.fcst.pct.hhs = z.fcst.ri.hhs / z.base.r.hhs * 100 (% change in persons per household)

z.fcst.pct.er_hh = z.fcst.ri.er_hh / z.base.r.er_hh * 100 (% change in employed residents per household)

Outlier Checks - Persons per Household and Employed Residents per Household

Outlier checks are based on the LUZ's change relative to the region's change in persons per household and in employed residents per household. If any of the following three conditions occur, a warning is written along with evaluation information: (1) a LUZ rate change in the opposite direction of the regional rate change; (2) a LUZ percent change that differs from the regional percent change by more than 5%; and (3) a LUZ with fewer than 50 household population in the base year and a household population forecast greater than 50. A LUZ with fewer than 50 household population in both the base and forecast years is not checked for outliers. The algorithm below illustrates the outlier checking mechanism for the persons per household.

IF (z.base.p.hhp < 50 and z.fcst.p.hhp < 50)SKIP CHECK

flag = 0

label = pph check

t9 = rc.fcst.r.hhs / rc.base.r.hhs (regional change in persons per household)

t10 = z.fcst.r.hhs / z.base.r.hhs (LUZ change in persons per household)

t11 = t10 / t9 (LUZ to regional change)

Check 1

IF (t10 > 1 AND t9 ≤ 1) OR (t10 < 1 AND t9 ≥ 1))

flag = 1

Check 2

IF (t11 < .95 OR t11 > 1.05)

flag = 2

Check 3

IF (z.base.p.hhp < 50 and z.fcst.p.hhp > 50)

flag = 3

IF (FLAG > 0)

Write warning: LUZ, flag, label, z.base.r.hhs, z.fcst.r.hhs, z.base.r.er_hh, z.fcst.r.er_hh, t9, t10, t11

The algorithm for the employed residents per household consistency check is essentially the same, with the temporary variables (t) representing this rate. The label variable changes to er_hh check. The same variables written for the household size check also are written for the employed residents per household check.

Outlier Checks - Vacancy Rate by Class Type

Four criteria are used to evaluate the vacancy rate for each class type. These criteria are: (1) a LUZ rate change in the opposite direction of the regional rate change; (2) a LUZ percentage change that differs from the regional percent change by more than 20%; (3) a LUZ with < 25 base year housing units and a housing unit forecast greater than 25; and (4) a vacancy rate that falls to zero or rises to 100 during the forecast interval. A LUZ with fewer than 25 units in both the base and forecast years is not checked for outliers. The algorithm below illustrates the outlier checking mechanism for the single family vacancy rate.

IF (z.base.hs.sf < 25 and z.fcst.hs.sf < 25) SKIP CHECK

flag1 = 0

label = sfvac check

t12 = rc.fcst.r.v_sf / rc.base.r.v_sf (regional change in vacancy rate)

t13 = z.fcst.r.v_sf / z.base.r.v_sf (LUZ change in vacancy rate)

t14 = t13 / t12

Check 1

IF (t13 > 1 and t12 ≤ 1) OR (t13 < 1 AND t12 ≥ 1))

flag1 = 1

Check 2

IF (t14 < .80 OR t14 > 1.20)

flag1 = 2

Check 3

IF (z.base.hs.sf < 25 and z.fcst.hs.sf > 25)

flag1 = 3

Check 4

IF (z.base.r.v_sf > 0 AND z.base.r.v_sf < 100) AND

(z.fcst.r.v_sf = 0 OR z.fcst.r.v_sf = 100)

flag1 = 4

IF (flag1 > 0)

Write warning: LUZ, flag1, label, z.base.r.v_sf, z.fcst.r.v_sf, z.base.r.v_mf, z.fcst.r.v_mf, z.base.r.v_mh, z.fcst.r.v_mh, t12, t13, t14

The algorithm for the other class types is essentially the same, with the temporary variables (t) representing their vacancy rates. The label variable changes to mfvac check and mhvac check. Vacancy rates for all class types are written regardless of the class type being checked.

Household Income Distribution

LUZ household income distributions are forecast as follows. First, the household forecast is split into an initial income distribution (z.fcst.i.distinc) in one of three ways, depending on z.ro.income_switch. If z.ro.income_switch = 1, the regional income relative distribution forecast is used. If z.ro.income_switch = 2, the LUZ's base year relative distribution is used. Otherwise, the modified lognormal curve generates the initial relative distribution.

Three LUZ-level parameters are needed to solve the modified lognormal curve: (1) median income, (2) standard deviation factor, and (3) calibration exponent. The latter two parameters are the inputs, z.asd and z.nla. The median income forecast is the base year median adjusted for the change in the regional median.

$$z.fcst.i.median = z.base.i.median * (rc.fcst.i.median / rc.base.i.median)$$

User overrides to these parameters are implemented.

$$\begin{aligned} z.fcst.i.median &= z.ro.income_median \text{ (median income forecast)} \\ z.fcst.asd &= z.ro.asd \text{ (lognormal curve standard deviation)} \\ z.fcst.nla &= z.ro.nla \text{ (lognormal curve non-linear adjustment parameter)} \end{aligned}$$

The modified lognormal-curve determines the proportionate distribution of households across income categories. These proportions are adjusted by a factor that keeps the forecast consistent with the base year distribution.

$$z.fcst.i.dist_{inc} = do_income_dist(z.fcst.i.median, z.fcst.asd, z.fcst.nla), \text{ where } do_income_dist \text{ is the modified lognormal curve function (unadjusted proportionate income distribution)}$$

$$\begin{aligned} z.fcst.i.dist_{inc} &+= z.fcst.income_adj_{inc} \text{ (adjusted proportionate income distribution)} \\ z.fcst.i.hh_{inc} &= z.fcst.i.dist_{inc} * z.fcst.hh.tot \text{ (income distribution forecast)} \end{aligned}$$

The plus-minus iterative proportionate adjustment method controls the initial LUZ household income distributions to both the LUZ occupied units forecast (z.fcst.hh.tot) and the regional income distribution forecast (rc.fcst.i.hh_{inc}). The controlling of household income involves only positive values, but this routine also can handle controlling of both negative and positive values. Controlling is an iterative sequential raking of the initial LUZ income distribution matrix, controlling alternatively to LUZ occupied units and to the regional income distribution. This adjustment sequence is continued until convergence to both controls and an unchanging matrix (z.fcst.i.hh_{inc}). Override LUZs are not changed if ctlparmi = 0. Rounding error is handled using a ± 1 matrix adjustment routine. Once the income distribution is normalized, the median income forecast is recomputed along with other variables.

$$\begin{aligned} z.fcst.ii.hh_{inc} &= z.fcst.i.hh_{inc} - z.base.i.hh_{inc} \text{ (household income distribution change)} \\ z.fcst.i.median &= median_income(z.fcst.i.hh_{inc}) \text{ (median income forecast)} \\ z.fcst.ii.median &= z.fcst.i.median - z.base.i.median \text{ (median income change)} \\ z.fcst.pct.income_hh_{inc} &= z.fcst.ii.hh_{inc} / z.base.i.hh_{inc} * 100 \text{ (% change in household income distribution)} \\ z.fcst.pct.income_median &= z.fcst.ii.median / z.base.i.median * 100 \text{ (% change in median income)} \end{aligned}$$

Civilian Employment Sectors

At this point, site-specific civilian employment is distributed into sectors for each LUZ and the regional control totals are adjusted. To complete the employment allocation, the non site-specific LUZ change is distributed into the 14 civilian employment sectors ($z.fcst.ei.sectors_adj_{sec}$). An initial distribution is computed and controlled, using the plus-minus iterative procedure, to the adjusted regional change ($rc.fcst.ei.sectors_adj_{sec}$) and the LUZ total non site-specific civilian employment change ($z.fcst.ei.adj$). After controlling, the employment increment for each sector is obtained by summing the site-specific and non-site-specific changes.

Several conditions determine the initial distribution of the LUZ non site-specific employment change into sectors. First, sectors with a regional decrement are distributed using the LUZ's base year employment in that sector relative to the region's base year employment in that sector. To maintain the proper LUZ control total, the LUZ change is increased by subtracting the allocated decrements from it. The distribution of the adjusted LUZ control into the remaining sectors depends on direction of the LUZ control ($z.fcst.ei.adj$) prior to the adjustment made for the allocation of regional decrements. If that control is negative, the LUZ's base year relative employment distribution is used. If it is positive, a weighted average of the LUZ's base year relative distribution and regional change relative distribution is used. The weights are based on the size of the LUZ increment to its base year civilian employment. For a LUZ with little or no base year employment or an employment increment equal to or greater than its base year employment, this method in effect defaults to the regional distribution. The allocation to civilian employment sectors, for a LUZ with a small increment relative to its base year employment, is influenced primary by the base year employment sector distribution.

```
t16 = z.fcst.ei.adj (temporary storage for LUZ control)
z.fcst.ei.sectors_adj_sec = 0 (non site-specific civilian employment by sector)
flag2 = 0
DO FOR ALL SECTORS
  IF (rc.fcst.ei.sectors_adj_sec < 0)
    flag2 =1 (check for regional decrements)
  END SECTOR
IF (flag2 = 0)
  GOTO ALLOCATE LUZ CONTROL
IF (flag2 = 1)
  ALLOCATE REGIONAL DECREMENTS (sec subscript on pertains only to employment sectors
  with a regional decrement)
  t17_sec = (z.base.e.sectors_sec / rc.base.e.sectors_sec) (distribution factor)
  z.fcst.ei.sectors_adj_sec = t17_sec * rc.fcst.ei.sectors_adj_sec (LUZ decrement)
  t18 =  $\Sigma z.fcst.ei.sectors\_adj_{sec}$ , where  $\Sigma$  is the sum over all sectors
  t16 -= t18 (LUZ control adjusted for employment loss due to regional decrement)
```

At this point, $z.fcst.ei.sectors_adj_{sec}$ contains negative values for all sectors with a regional decrement and zero for the other sectors.

```
ALLOCATE LUZ CONTROL (sec subscript on pertains only to employment sectors with a regional
increment)
```


$t19_{sec} = rc.fcst.ei.sectors_adj_{sec} / \sum rc.fcst.ei.sectors_adj_{sec}$, where Σ is the sum across all sectors with a positive regional change (proportionate distribution of the regional change)

$t20_{sec} = z.base.e.sectors_{sec} / \sum z.base.e.sectors_{sec}$, where Σ is the sum across all sectors with a positive regional change (LUZ base year sector proportionate distribution)

$wt = z.fcst.ei.adj / z.base.e.civ$ (weight for computing weighted average)

IF ($wt > 1$)

$wt = 1$

IF ($wt < 0$)

$wt = 0$

$t21_{sec} = (t19_{sec} * wt) + (t20_{sec} * (1 - wt))$ (weighted average distribution factors)

IF ($z.fcst.ei.adj < 0$) (LUZ change negative)

$z.fcst.ei.sectors_adj_{sec} = t16 * t20_{sec}$ (use base year distribution to allocate LUZ decrements)

IF ($z.fcst.ei.adj \geq 0$) (LUZ change positive)

$z.fcst.ei.sectors_adj_{sec} = t16 * t21_{sec}$ (use weighted average distribution to allocate LUZ increments)

At this point, $z.fcst.ei.sectors_adj_{sec}$ contains the initial distribution of all employment sectors for non-site-specific activity.

Users can override the initial allocation of the employment sectors ($z.eso.sectors_adj_{sec}$). Overrides are checked for consistency against the LUZ control and adjusted if necessary

$t22 = \sum z.eso.sectors_adj_{sec}$

IF ($t22_{LUZ} \neq z.fcst.ei.adj$)

Normalize overrides to control using ± 1 vector adjustment routine

Write a warning to the user

$z.fcst.ei.sectors_adj_{sec} = z.eso.sectors_adj_{sec}$

$z.fcst.ei.sectors_adj_{sec}$ is controlled to both the LUZ control ($z.fcst.ei.adj$) and regional total ($rc.fcst.ei.sectors_adj_{sec}$) using the plus-minus iterative proportionate adjustment procedure. The plus-minus method should preserve the direction of the change indicated in the initial matrix. Rounding error is handled using a ± 1 matrix adjustment routine. Override LUZs are not changed during the controlling if $ctlparms = 0$. After controlling, other variables are computed.

$z.fcst.ei.sectors_{sec} = z.site.sectors_{sec} + z.fcst.ei.sectors_adj_{sec}$ (civilian employment sector change including site-specific)

$z.fcst.e.sectors_{sec} = z.base.e.sectors_{sec} + z.fcst.ei.sectors_{sec}$ (civilian employment sector forecast)

$z.fcst.ei.civ = \sum z.fcst.ei.sectors_{sec}$ (total civilian employment change)

$z.fcst.e.civ = \sum z.fcst.e.sectors_{sec}$, where Σ is the sum over the sectors (total civilian employment forecast)

$z.fcst.pct.sectors_{sec} = z.fcst.ei.sectors_{sec} / z.base.e.sectors_{sec} * 100$ (% change in civilian employment sectors)

$z.fcst.pct.civ = z.fcst.ei.civ / z.base.e.civ * 100$ (% change in total civilian employment)

MGRA Characteristics Forecast

Occupied Units by Class Type

Occupied units are calculated by multiplying the housing stock forecast by class type-specific occupancy rates (1 – vacancy rate). The housing stock class type is forecast by adding the change aggregated from the transactions to the base year housing stock. The vacancy rate forecast is derived by adjusting the MGRA's base year vacancy rate by the change in the LUZ rate. If there are 20 or fewer base year units, the LUZ vacancy rate forecast is used. UDM insures that the occupied unit allocation does not exceed the housing stock and does not violate the LUZ control. The LUZ control total is reduced as the MGRA occupied units are developed. If the LUZ control is not zero after all MGRAs are processed, a ± 1 vector adjustment corrects the rounding error. The occupied units routine is shown below.

```
control_xx = z.fcst.hh.xx (LUZ occupied units control)
t23_xx = z.fcst.r.v_xx / z.base.r.v_xx (change in LUZ vacancy rate)
m.fcst.hs.xx = m.base.hs.xx + m.fcst.hsi.xx (housing stock forecast by class type)
IF (m.base.hs.xx > 0)
    m.base.r.v_xx = (1 - (m.base.hh.xx / m.base.hs.xx)) (base year vacancy rate)
IF (m.base.hs.xx > 20)
    m.fcst.r.v_xx = minimum(m.base.r.v_xx * t23_xx, 1) (vacancy rate forecast cannot exceed 1)
IF (m.base.hs.xx ≤ 20)
    m.fcst.r.v_xx = z.fcst.r.v_xx (use LUZ vacancy rate forecast)
m.fcst.hh.xx = m.fcst.hs.xx * (1 - m.fcst.r.v_xx) (occupied units forecast)
control_xx -= m.fcst.hh.xx (adjusted LUZ control)
IF (ALL MGRAs PROCESSED and control_xx ≠ 0)
    IMPLEMENT VECTOR ADJUSTMENT
```

```
m.fcst.hs.tot = Σm.fcst.hs.xx (total housing stock forecast)
m.fcst.hh.tot = Σm.fcst.hh.xx, where Σ is the sum across class types (total occupied units forecast)
m.fcst.hhi.xx = m.fcst.hh.xx - m.base.hh.xx (occupied units change by class type)
m.fcst.hhi.tot = m.fcst.hh.tot – m.base.hh.tot (total occupied units change)
```

Household Population and Employed Residents

The MGRA allocation of household population and employed residents per household use logic similar to the occupied units routine. UDM insures that every occupied unit has at least one resident (i.e., a persons per household of 1) and that the household population and employed resident allocations do not violate the LUZ control. To insure these consistencies, all MGRAs with a non-zero occupied units forecast are processed through this routine, even after the LUZ control reaches zero. If this happens, the adjusted LUZ control is negative and the vector adjustment removes the extra household population or employed residents. The persons per household and employed residents per household routines are shown below.

```
control = z.fcst.p.hhp (LUZ household population control)
t24 = z.fcst.r.hhs / z.base.r.hhs (change in LUZ persons per household)

IF (m.base.hh.tot > 0)
    m.base.r.hhs = m.base.p.hhp / m.base.hh.tot (base year persons per household)
```

IF (m.base.hh.tot > 20)
 m.fcst.r.hhs = maximum(m.base.r.hhs * t24, 1) (at least 1 person in each household)
 IF (m.base.hh.tot ≤ 20)
 m.fcst.r.hhs = z.fcst.r.hhs (use LUZ persons per household forecast)

m.fcst.p.hhp = m.fcst.hh.tot * m.fcst.r.hhs (household population forecast)
 control -= m.fcst.p.hhp (adjusted LUZ control)
 IF (ALL MGRAs PROCESSED and control ≠ 0)
 IMPLEMENT VECTOR ADJUSTMENT

control = z.fcst.p.er (LUZ employed residents control)
 t25 = z.fcst.r.er_hh / z.base.r.er_hh (change in LUZ employed residents per household)

IF (m.base.hh.tot > 0)
 m.base.r.er_hh = m.base.p.er / m.base.hh.tot (base year employed residents per household)
 IF (m.base.hh.tot > 20)
 m.fcst.r.er_hh = maximum(m.base.r.er_hh * t25, .5) (employed residents per households at least .5)
 IF (m.base.hh.tot ≤ 20)
 m.fcst.r.er_hh = z.fcst.r.er_hh (use LUZ employed residents per household forecast)

m.fcst.p.er = z.fcst.hh.tot * m.fcst.r.er_hh (employed residents forecast)
 control -= m.fcst.p.er (adjusted LUZ control)
 IF (ALL MGRAs PROCESSED and control ≠ 0)
 IMPLEMENT VECTOR ADJUSTMENT

Group Quarters Population

Civilian and military group quarters forecasts are also produced for the MGRAs. Military group quarters are forecast by adding site-specific group quarters to the base year military group quarters.

$$m.fcst.p.gq_mil = m.base.p.gq_mil + m.site.gq_mil$$

The civilian group quarters are done in a similar fashion to the LUZ forecast; that is, by factoring the MGRA's base year civilian group quarters adjusted for site-specific activity to the forecasted LUZ change. The adjustment factor is produced by dividing the LUZ forecasted change adjusted for site-specific activities (pt) by the LUZ adjusted base year civilian group quarters population (z.base.p.gq_civ_extra). These LUZ variables are the same ones computed during the allocation of civilian group quarters to LUZs. This factoring process yields an initial value for the change in the MGRA's civilian group quarters (m.fcst.pi.gq_civ_adj) and the ± 1 vector adjustment routine corrects rounding error.

Additional population variables are computed. The forecast of civilian group quarters is the sum of the base year group quarters, change not including site-specific activities, and change due to site-specific activities.

$$m.fcst.p.gq_civ = m.base.p.gq_civ + m.fcst.pi.gq_civ_adj + m.site.gq_civ \text{ (civilian group quarters forecast)}$$

$$m.fcst.p.gq = m.fcst.p.gq_civ + m.fcst.p.gq_mil \text{ (total group quarters forecast)}$$

$$m.fcst.p.pop = m.fcst.p.hhp + m.fcst.p.gq \text{ (total population forecast)}$$

Household Income Distribution

The MGRA household income distribution forecast requires several passes through the MGRAs, treats losses and gains separately, and employs the Pachinko method for determining allocation probabilities. A set of tables and equations accompany the narrative to illustrate these calculations. This example, for a single LUZ, involves three income classes rather than the ten classes modeled in UDM.

Table 2 shows the matrix prior to starting the income allocation. The marginals of the matrix represent change and the body of the table (in a smaller font) represents the MGRA base year income distribution ($m.base.i.hh_{inc}$). The row marginals (MGRA control total) are changes in occupied units ($m.fcst.hhi.tot$), while the column marginals (LUZ control total) represent changes in the income distribution ($z.fcst.ii.hh_{inc}$). In order for the income allocation to work, the sum over the row and column marginals must be equal.

$t26 = \sum m.fcst.hhi.tot$, where Σ is the sum over the MGRAs
 $t27 = \sum z.fcst.ii.hh_{inc}$, where Σ is the sum over the income categories
 IF ($t26 \neq t27$)
 Write Error Message

Table 2
Household Income Allocation – Initial Matrix

MGRA	Base Year Income Distribution			MGRA Change
	< \$30k	\$30-59k	\$60k+	
122	10	15	20	5
524	30	40	50	80
1022	60	70	80	-60
LUZ Change	10	-35	50	25

Once the matrix is assembled, the LUZ household income distribution change is allocated in three steps. First, LUZ income category decrements are processed. UDM then processes MGRAs with occupied unit decrements. Finally, positive LUZ changes and positive MGRA changes are allocated. UDM changes base year MGRA values during these allocations, so when UDM completes the allocation the base year distribution is transformed into the forecast distribution. Subsequent tables are labeled **Base to Forecast** to reflect this accounting process.

LUZ Income Category Decrements

UDM checks the LUZ change for negative values. For LUZ decrements, UDM adjusts the base year MGRA data until the LUZ decrement(s) reach zero. It does this with a -1 Pachinko method, which essentially removes units in proportion to the MGRA's share of LUZ occupied units in the particular income category. This routine subtracts one from the selected MGRA's base year units and makes as many passes through the MGRAs as are needed to satisfy the LUZ control. To insure proper accounting, the MGRA change in occupied units is increased by the number of units removed from

the MGRA income category. Table 3 illustrates this section of the routine; the bold numbers show the results after the distribution of the LUZ decrement. Income category 2 has a LUZ decrement of 35. In this example, base year MGRA households for income category 2 are reduced by 4, 11, and 20 respectively, reflecting the MGRAs' share of the LUZ occupied units.

Table 3
Household Income Allocation – LUZ Decrement Accounting

MGRA	Base to Forecast Income Distribution						MGRA Change	
	< \$30k		\$30-59k		\$60k+			
122	10	10	15	11	20	20	5	9
524	30	30	40	29	50	50	80	91
1022	60	60	70	50	80	80	-60	-40
LUZ Change	10	10	-35	0	50	50	25	60

MGRA Occupied Unit Decrements

For MGRA occupied unit decrements, UDM adjusts the base year MGRA data until the decrement(s) reach zero. This adjustment also is made using the –1 Pachinko method. To preserve the correct control totals, UDM increases the LUZ changes by the number of units removed from each income category. When UDM finishes accounting for MGRA decrements, it sums the adjusted MGRA controls and compares it to the sum of the adjusted LUZ controls. These two must match in order to finish the income allocation. UDM also checks the sum across MGRAs for each income category plus the LUZ increment against the LUZ forecast total. These also must match. The bold numbers in Table 4 illustrate the effect of these computations and the non bold numbers represent the results from Table 3. MGRA 1022 has a decrement of 40 occupied units. In this example, the income categories are reduced by 13, 10, and 17 respectively, reflecting the share of the MGRA's units in each income category.

$t_{28} = \sum m.fcst.hhi.tot$, where Σ is the sum over the MGRAs adjusted for LUZ income category and MGRA occupied unit decrements

$t_{29} = \sum z.fcst.ii.hh_{inc}$, where Σ is the sum over the income categories adjusted for LUZ income category and MGRA occupied unit decrements

IF ($t_{28} \neq t_{29}$)

Write Error Message

$t_{30_{inc}} = \sum z.base.i.hh_{inc}$, where Σ is the sum over the MGRAs

IF ($(t_{30_{inc}} + z.fcst.ii.hh_{inc}) \neq z.fcst.i.hh_{inc}$)

Write Error Message

Table 4
Household Income Allocation – MGRA Decrement Accounting

MGRA	Base to Forecast Income Distribution						MGRA Change	
	< \$30k		\$30-59k		\$60k+			
122	10	10	11	11	20	20	9	9
524	30	30	29	29	50	50	91	91
1022	60	47	50	40	80	63	-40	0
LUZ Change	10	23	0	10	50	67	60	100

The following consistency checks from Table 4 assume a LUZ forecast for the three income class of 110, 90, and 200 respectively.

$$(9 + 91 + 0 = 100) = (23 + 10 + 67 = 100) \text{ (marginals)}$$

$$(10 + 30 + 47) + 23 = 110 \text{ (income category 1)}$$

$$(11 + 29 + 40) + 10 = 90 \text{ (income category 2)}$$

$$(20 + 50 + 63) + 67 = 200 \text{ (income category 3)}$$

Positive Marginal Controls

Finally, positive increments are distributed during a final pass through the MGRAs using the +1 Pachinko method. Pachinko begins by calculating a cumulative probability distribution of the LUZ increments. This distribution is expressed as integers that range from 1 to 100. For the example shown in Table 4, the values for the three income categories are 23 (23 / 100 * 100), 33 ((23 + 10) / 100 * 100), and 100; the last income group always has a value of 100. This distribution determines the allocation of MGRA units into the income categories. A loop controls this allocation. The loop begins at 1 and ends with the MGRA units to be allocated. Each time through the loop a random number between 1 and 100 is selected and is compared to the values in the probability distribution. If the random number is less than or equal to the probability distribution value, a unit is allocated to that income category. For example, a random number of 60 points to the highest income category, a random number of 30 points to the middle income category, and a random number of 15 points to the lowest income category.

Once UDM determines the income category, it adds one to that category's units and subtracts one from both the MGRA control and the LUZ control. All MGRA units are allocated to their proper income categories when UDM finishes the loop. If any LUZ control reaches zero during the loop, the probability distribution is recalculated so that the remaining units are allocated to income categories with a non-zero LUZ control. UDM repeats this process for each MGRA with a positive control total. The bold numbers in Table 5 and Table 6 illustrate these computations for the two MGRAs with positive occupied unit increments. The non bold numbers in Table 5 represent the results from Table 4 and in Table 6 they represent the results from Table 5. For MGRA 122, the probability distribution applied to the 9 occupied units increment adds 2 units (.23 * 9), 1 unit ((.33 - .23) * 9), and 6 units (9 - (2 + 1)) to the income categories. For MGRA 524, the probability distribution applied to the 91 occupied units increment adds 21 units (.23 * 92), 9 units ((.33 - .23) * 92) and 61 units (91 - (21 + 9)) to the income categories.

**Table 5
Household Income Allocation – MGRA 1022 Increment Accounting**

MGRA	Base to Forecast Income Distribution						MGRA Change	
	< \$30k		\$30-59k		\$60k+			
122	10	12	11	12	20	26	9	0
524	30	30	29	29	50	50	91	91
1022	47	47	40	40	63	63	-0	0
LUZ Change	23	21	10	9	67	61	100	100

**Table 6
Household Income Allocation – MGRA 1024 Increment Accounting**

MGRA	Base to Forecast Income Distribution						MGRA Change	
	< \$30k		\$30-59k		\$60k+			
122	12	12	12	12	26	26	0	0
524	30	51	29	38	50	111	91	0
1022	47	47	40	40	63	63	-0	0
LUZ Change	21	0	9	0	61	0	91	0

The LUZ income and MGRA occupied units marginal control totals are now zero, which indicates that both have been completely allocated. The forecast in bold now matches the LUZ income distribution forecast and the MGRA occupied units forecast as shown below, assuming the occupied units forecast for the three MGRAs is 50, 200, and 150 respectively. The adjusted MGRA base year income distribution shown in table 5 is equivalent to the MGRA income distribution forecast (m.fcst.i.hh_{inc}).

- (12 + 51 + 47) = 110 (LUZ forecast income category 1)
- (12 + 38 + 40) = 90 (LUZ forecast income category 1)
- (26 + 111 + 63) = 200 (LUZ forecast income category 1)
- (12 + 12 + 26) = 50 (MGRA 1022 household forecast)
- (51 + 38 + 111) = 200 (MGRA 1024 household forecast)
- (47 + 40 + 63) = 150 MGRA 3768 household forecast)

Non Site-Specific Civilian Employment Sectors

The allocation of the LUZ non site-specific civilian employment sectors to the MGRAs also uses the income distribution algorithm. The row marginals of the initial matrix represent the non site-specific total civilian employment change for the MGRAs (m.fcst.ei.adj), while the column marginals represent the LUZ change for non site-specific civilian employment by sector (z.fcst.ei.sectors_adj_{sec}).

The base year civilian employment adjusted for site-specific changes is used in the body of the initial matrix.

$$\text{temp_MGRAs}_{\text{sec}} = \text{z.base.e.sectors}_{\text{sec}} + \text{z.site.sectors}_{\text{sec}}$$

When the allocation is completed, the adjusted base represents the correct forecast of civilian employment by sector ($\text{m.fcst.e.sectors}_{\text{sec}}$). The equations that follow represent the consistency checks done at the beginning of the allocation and after the distribution of the LUZ sector decrements and MGRA employment decrements.

Beginning of allocation

$t31 = \Sigma \text{m.fcst.ei.adj}$, where Σ is the sum over the MGRAs

$t32 = \Sigma \text{z.fcst.ei.sectors_adj}_{\text{sec}}$, where Σ is the sum over the employment sectors

IF ($t31 \neq t32$)

Write Error Message

After distribution of LUZ and MGRA Decrements

$t33 = \Sigma \text{temp_MGRAs}_{\text{sec}}$, where Σ is the sum over the MGRAs adjusted for LUZ employment sector and MGRA civilian employment decrements

$t34 = \Sigma \text{z.fcst.ei.sectors_adj}_{\text{sec}}$, where Σ is the sum over the employment sectors adjusted for LUZ employment sector and MGRA civilian employment decrements

IF ($t33 \neq t34$)

Write Error Message

$t35_{\text{sec}} = \Sigma \text{temp_MGRAs}_{\text{sec}}$, where Σ is the sum over the MGRAs

IF ($(t35_{\text{sec}} + \text{z.fcst.ei.sectors_adj}_{\text{sec}}) \neq \text{z.fcst.e.sectors}_{\text{sec}}$)

Write Error Message

To complete the MGRA employment forecast, the uniformed military is computed by adding the site-specific change to its base year value. Additional MGRABASE variables are computed.

$\text{m.fcst.e.mil} = \text{m.base.e.mil} + \text{m.site.mil}$ (uniformed military)

$\text{m.fcst.e.civ} = \Sigma \text{m.fcst.e.sectors}_{\text{sec}}$, where Σ is the sum across sectors (total civilian employment forecast)

$\text{m.fcst.e.tot} = \text{m.fcst.e.civ} + \text{m.fcst.e.mil}$ (total employment forecast)

Updating Data Bases

This part of module 4 creates the base year data for the next forecast period—MGRABASE, LUZBASE, LUZHIS.

MGRABASE

The procedures just described in module 4 creates the demographic and economic variables contained in MGRABASE. Its land use variables are aggregated from the capacity file, taking into account the trans.pcap . The creation of the MGRABASE land use information is a classification problem rather than a true aggregation, since it places land uses within each MGRA in their appropriate category. Complicating this process is the fcst.trans.pcap , which requires division of the

acres and the use of both fcst.trans.LU and fcst.trans.plu to determine their proper location. The code is too extensive to describe here.

LUZBASE and LUZHIS

Although the LUZ information for LUZBASE is directly available in module 4, to insure consistency the LUZBASE is created by aggregating the MGRA-level demographic and economic forecast.

LUZHIS (historical data) is updated by using the base year data as the new 5-year history. For example, for the 2004 to 2010 time period, the 2004 value represents the history .

z.hist_sf.I5 = z.base.hs.sf (single family units 5 years prior)

z.hist_mf.I5 = z.base.hs.mf (multiple family units 5 years prior)

z.hist_e.I5 = z.base.e.civ (civilian employment 5 years prior)

[THIS PAGE INTENTIONALLY LEFT BLANK]

SECTION 3 – PROGRAMMING

3.1 Introduction

This section provides general information for the Series 13 Forecast program suite, including source files, data and utilities. The discussion covers the current program organizations and operations, and provides general detail to assist in making changes to source code and to prepare executable programs. Additional technical information describing the various relationships and equations is available in the Section 2. Detailed program documentation is provided in each project and source file as described below. This discussion covers only those programs used to complete the sub-regional allocation and does not include programs used for the land-use input maintenance.

All of the Series 13 model components are maintained in the Microsoft .Net development environment and are coded in C#. While program names in this document may be listed in all capitals (CAPACITY, UDM), the actual program names may be all lower case or mixed case.

- <file_name>.csproj : C#.net project files containing source files, forms and other overhead files (See the .Net documentation for a complete description)
- <file_name>.cs : identifies C# source files

In the paragraphs that follow, the appropriate directory path is identified. Unless stated otherwise, files are assumed to reside in the listed directory. The symbol "~" signifies the parent directory path for the components of the modeling system. File and table names may include an identifier for forecast increment and scenario. Five-year forecast increments are identified as _YYYY where YYYY = 2004, 2010, 2020 and so on. Scenarios are labeled with two-character identifiers _AA such as "EP": (existing policies).

The Series 13 Forecast programs are maintained on SANDAG's shared network. All of the input data and most output are maintained in the SQL Server 2000 relational database management system. We use two different SQL Servers. **LIS** hosts the GIS components of the Series 13 Forecast. **Pila** hosts the sub-regional modeling components. When referring to databases and tables the convention used here is **server.database.tablename**. For example, **LIS.forecast.lcgpall** refers to the server **LIS**, database **forecast** and table **lcgpall**.

Table 7 summarizes the organization of the Series 13 Forecast development and operational software.

**Table 7
Series 13 Forecast Software and Data Organization**

Label	Description
M:\res\estimates & forecast\UDM (~)	Root directory path
~\AccessWeights	Access Weights program
~\CapacitySR13	Capacity program
~\docs	Formal documentation
~\GPALL	Gpall program
~\temp	Temporary output directory
~\UDM	UDM model components
Pila	SQL Server hosting sub-regional allocation components of the Series 13 Forecast
LIS	SQL Server hosting the GIS (land use records) components of the Series 13 forecast
pila.sr13	Sub-regional allocation components database for Series 13 forecast
LIS.forecast	GIS components data base for Series 13 forecast

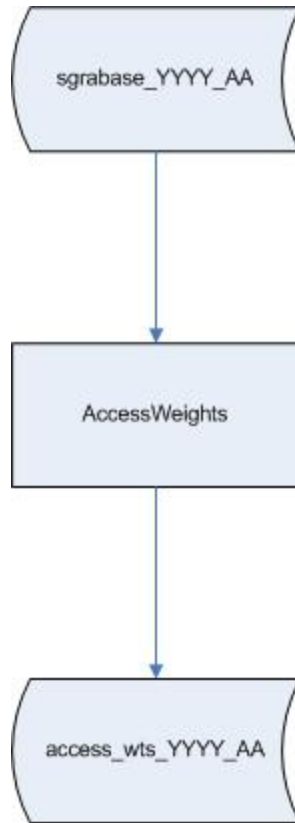
3.2 AccessWeights

AccessWeights creates the distribution weights used in UDM to distribute housing and employment forecasts to individual transaction in the capacity table. The program is run once before each increment and computes the distribution weights based on the sum of employment and housing units within a specified radius from a given MGRA. The program flow of information is illustrated in Figure 4. Input is from the MGRABase table for the specified increment and alternative. Computed access weights are output to a database table as listed in Table 8.

**Table 8
AccessWeights Input and Output**

Component	Description
MGRABase_YYYY_AA	Database input table - Civilian employment and housing stock input
access_wts_YYYY_AA	Database output table – derived allocation weights by MGRA

Figure 4
AccessWeights Flow of Information



3.3 CAPACITY

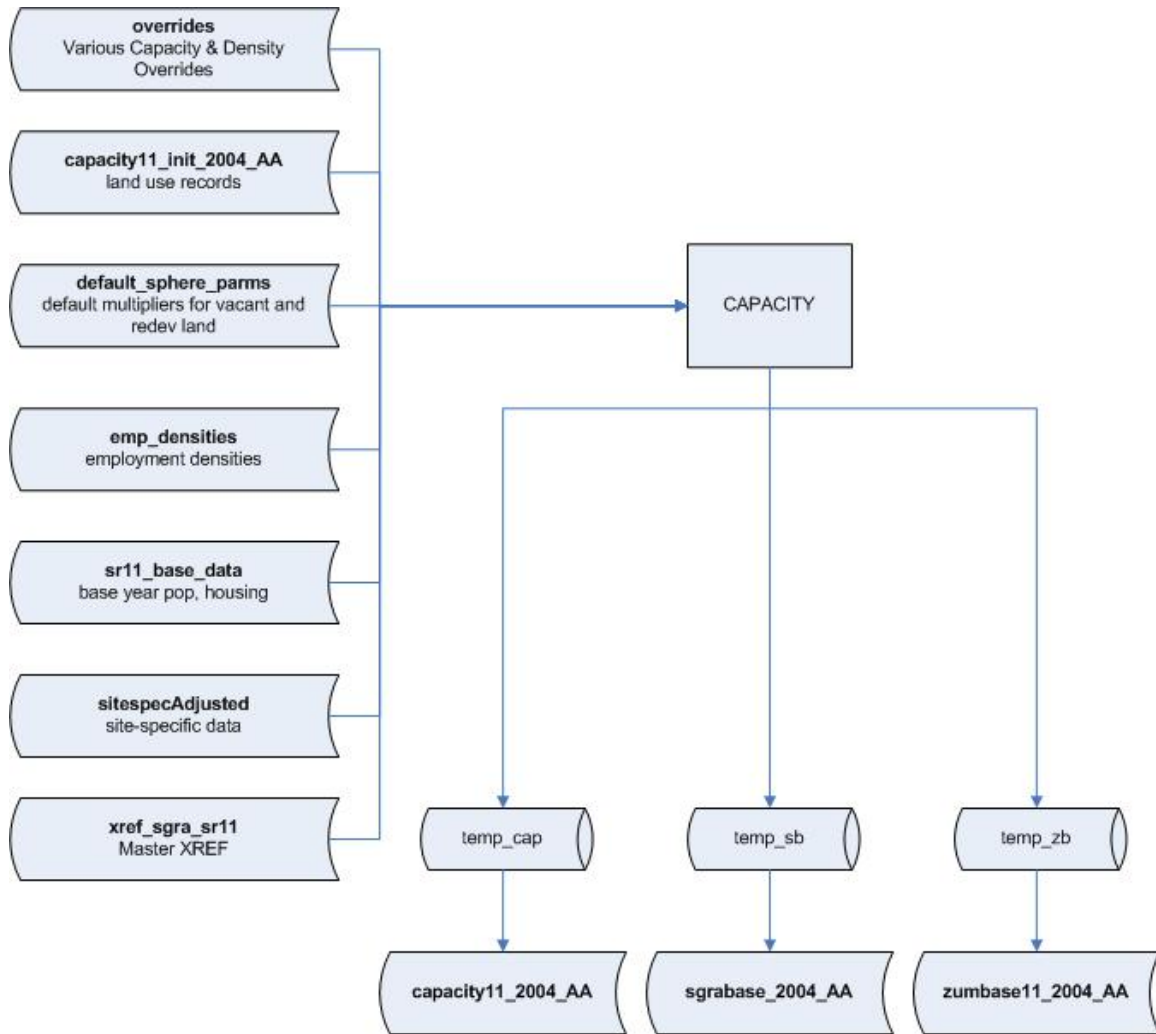
CAPACITY performs front-end processing for the main UDM sub-regional allocation model. The program accepts as input the regional land use file, computes employment and dwelling unit capacity for each MGRA and land use combination and builds the capacity, LUZbase and MGRABase tables. Its flow of information is illustrated in Figure 5.

Input and output for CAPACITY are database tables and ASCII files that are bulk-loaded to database tables as identified in Table 9.

**Table 9
CAPACITY Input and Output**

File/Table	Description
capacity11_2004_AA	Database output table - base year capacity, scenario AA
capacity11_init_2004_AA	Database input table – raw land use input data
default_sphere_parms	Database input table - default sphere multipliers for vacant land, redevelopment land and gross/net flags
emp_densities	Database input table- employment densities by land use category
overrides_carlsbad	Database input table - Carlsbad overrides for capacity calculations
overrides_ccdc	Database input table – San Diego Center City overrides for capacity calculations
overrides_county	Database input table – San Diego County overrides for capacity calculations
overrides_lkey	Database input table - lkey overrides for capacity calculations
overrides_nationalcity	Database input table – National City overrides for capacity calculations
overrides_sdcity	Database input table – San Diego city overrides for capacity calculations
overrides_spheres	Database input table - sphere overrides for capacity calculations
MGRABase_2004_AA	Database output table - base year MGRA data, scenario AA
sitespec_update	Database output table – used to adjust capacity table with site-specific data
sitespecAdjusted	Database input table – Adjusted site-specific input
sr13_base_data	Database input table – base year population, housing and income distribution
temp_cap	Temporary ASCII output file for capacity records
temp_sb	Temporary ASCII output file for MGRABase records
temp_zb	Temporary ASCII output file for LUZbase records
xref_MGRA_sr13	Database input table – SR13 cross-reference table
LUZ_inc_parms	Database input table - LUZ income calibration parameters
LUZbase11_2004_AA	Database output table - base year LUZ data, scenario AA

Figure 5
CAPACITY Flow of Information



3.4 GPALL

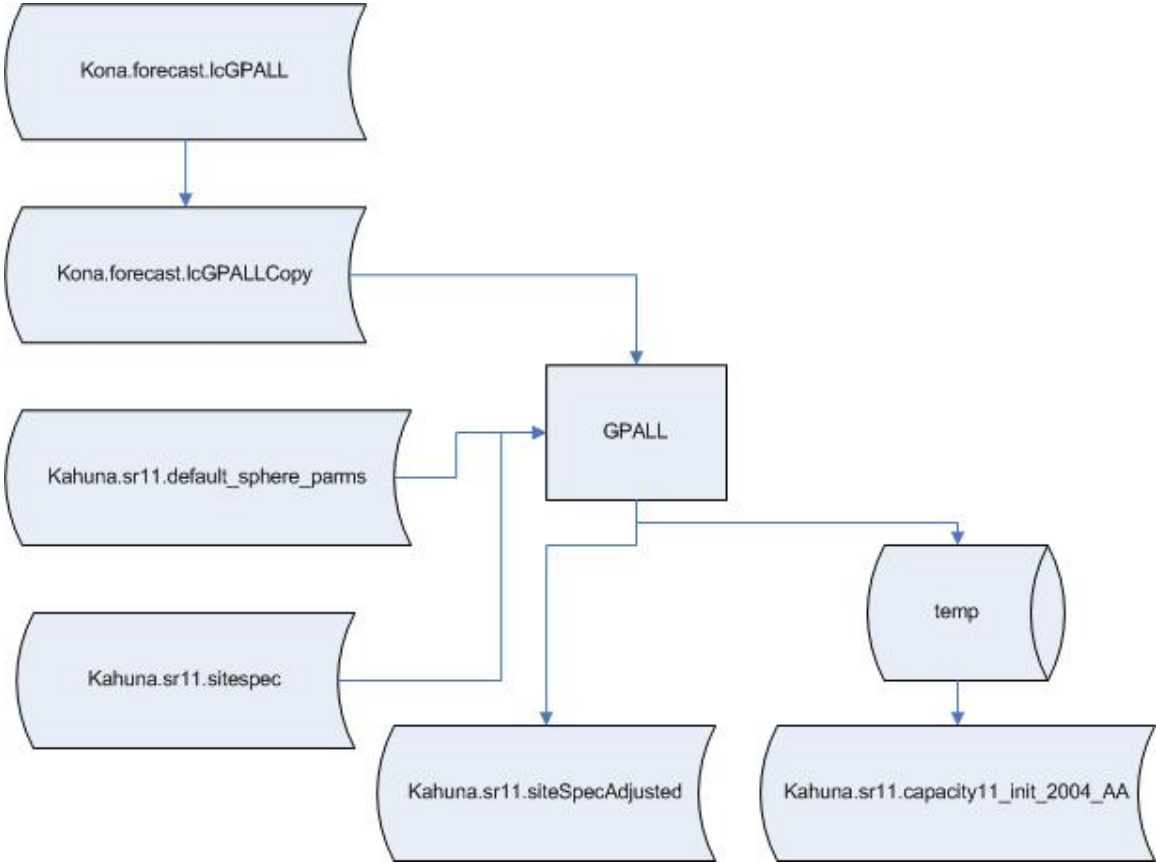
GPALL is the program that assigns the initial development codes to the land use records. Its flow of information is illustrated in Figure 6.

GPALL input come from database tables. The land-use records are resident on the GIS SQL Server LIS, in the **forecast** database. A copy of the table is used because some initial modifications are done to densities. The output is to the sub-regional allocation components data base Pila.SR13. GPALL writes an ASCII file and bulk loads the resulting initial capacity table. GPALL input and output are listed in Table 10.

**Table 10
GPALL Input and Output**

File/Table	Description
LIS.forecast.lcgpall	Land-use records
LIS.forecast.lcgpallcopy	Database input table – copy of lcgpall
Pila.sr13.default_sphere_parms	Database input table – default multipliers for density by land use and sphere
Pila.sr13.sitespec	Database input table – initial site-specific table
Pila.sr13.sitespecAdjusted	Database output table - adjusted
Pila.sr13.capacity11_init_2004_AA	Database output table – land use records with development code assignments

**Figure 6
GPALL Flow of Information**



3.5 UDM

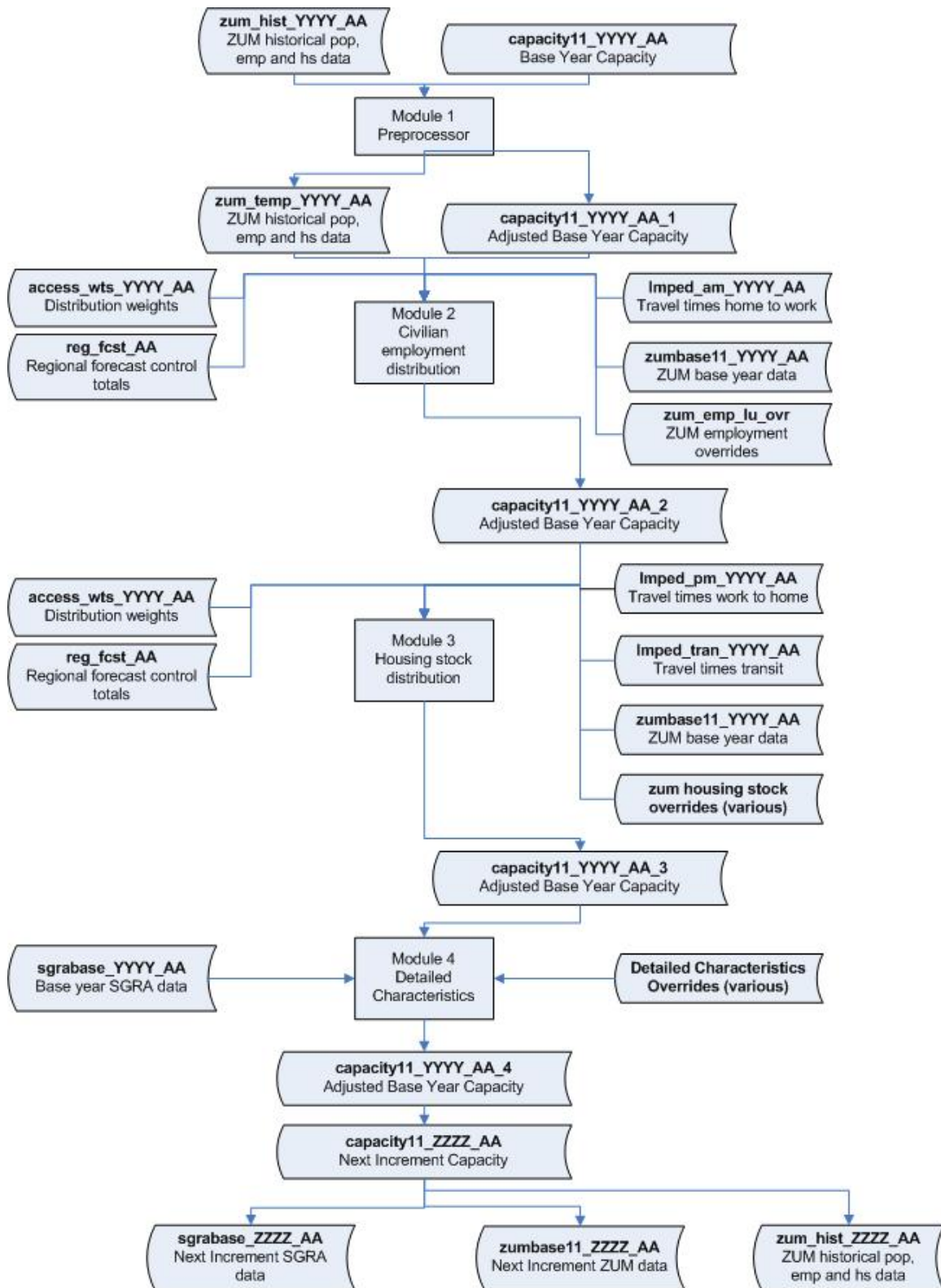
UDM is the principal sub-regional allocation model. It consists of four modules that may be run independently. The UDM flow of information is illustrated in Figure 7.

UDM input and output include database tables as identified in Table 11.

Table 11
UDM Input and Output

Table	Description
access_wts_YYYY_AA	Access weights used to distribute activity to transactions in UDM
capacity11_YYYY_AA	Capacity records by LCKEY
fractee_YYYY_AA	Percent of travel by transit by LUZ
imped_am_YYYY_AA	AM travel times
imped_pm_YYYY_AA	PM travel times
imped_tran_YYYY_AA	Transit travel times
reg_fcst_AA	DEFM regional controls
MGRAbase_YYYY_AA	Population, housing, employment, income and land use by MGRA
xref_MGRA_sr13	Master Series 13 geographic cross-reference table by MGRA
LUZ_attr_hs_ovr	LUZ housing stock attractors overrides
LUZ_dc_emp_sector_ovr	LUZ employment sector overrides
LUZ_dc_er_ovr	LUZ employed residents parms overrides
LUZ_dc_hhs_ovr	LUZ household size parms overrides
LUZ_dc_inc_ovr	LUZ income distribution parms overrides
LUZ_dc_vac_ovr	LUZ vacancy rates parms overrides
LUZ_emp_lu_ovr	LUZ employment and land use overrides
LUZ_hist_YYYY_AA	LUZ historical hs and employment data
LUZ_inc_parms	Income calibration parameters by LUZ
LUZ_mf_ovr	LUZ MF housing stock overrides
LUZ_mh_ovr	LUZ MH housing stock overrides
LUZ_sf_ovr	LUZ SF housing stock overrides
LUZ_temp_YYYY_AA	Intermediate (Module 1) data
LUZbase_YYYY_AA	Population, housing, income and employment by LUZ

**Figure 7
UDM Flow of Information**



[THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A HOUSEHOLD INCOME DISTRIBUTION

INTRODUCTION

Regional controls for critical economic and demographic data are an input to UDM. The Demographic and Economic Forecasting Model (DEFM) provides these controls. Household income distribution is computed in a post-processing module that was added to DEFM for the Series 13 forecast. This module produces income distribution forecasts for ten income groups based on assumptions about future changes in median income, the standard deviation of the regional income distribution, and a non-linear shape parameter.

MODIFIED LOGNORMAL CURVE

The income distribution module of DEFM employs a mathematical model, known as the modified lognormal curve, to estimate this distribution. The classical lognormal curve fits the middle of the income distribution well, but it underestimates the proportion of low-income units and overestimates the proportion of high-income units. A modification is made to the lognormal curve to remedy this problem. As a foundation of the development of the modified lognormal curve, the lognormal model is presented first.

A variate x , income, is lognormally distributed if the natural logarithm of x follows a normal curve. The lognormal distribution, y , is minimally specified by its mean and standard deviation:

$$\text{med}(y) = \ln[\text{med}(x)]; \text{ and}$$

$$\text{sd}(Y) = [2 * (\ln[\text{mean}(x)] - \ln[\text{med}(x)])]^{1/2}, \text{ where } \ln \text{ is the natural logarithm; med is the median; sd is the standard deviation; and mean is the average.}$$

Cumulative probabilities for units with incomes less than X are computed by: $y_i = \ln(x_i)$;

$$z_i = (1 / [\text{sd}(y) * \sqrt{2}]) * [y_i - \text{med}(y)]; \text{ and}$$

$$\text{cump}_i = (1 + \text{erf}[z_i / \sqrt{2}]) / 2, \text{ where } i \text{ is income class; } y \text{ is the upper limit of the income class; } z \text{ is the lognormal standard score; cump is the cumulative probability for income less than class } x_i; \text{ and erf is the error function of the normal curve.}$$

The modification of the lognormal curve is made to the z_i equation by:

$$\text{parm} = 1 / [\text{sd}(y) * \sqrt{2}] \text{ and;}$$

$$z_i = \text{parm} * y_i^{\text{exp}} * [y_i - \text{med}(y)], \text{ where } i \text{ is the income class; parm reflects the constant terms; and } y_i^{\text{exp}} \text{ is a non-linear adjustment parameter.}$$

The addition of y_i^{exp} to the lognormal curve helps corrects its estimation problems. Parm and exp are determined empirically by iteration to minimize the index of dissimilarity between the estimated curve and the observed income distribution (see the Calibration section below).

As the z_i equation shows, three parameters are needed to estimate the household income distribution—median income, parm, and exp. The cumulative probabilities (cump) are translated into a probability distribution, p , of a household being in an individual income class by:

$p_1 = cump_1$, where 1 is the lowest income category;

$p_i = cump_{j+1} - cump_j$, where i ranges from 2 to $n-1$; j ranges from 1 to n ; and n is the number of income categories;

$p_n = 1 - cump_{n-1}$; and

$\sum p_i = 1$, where \sum is the summation over the income groups.

Households in each income class are determined by multiplying the probability distribution by the total households. This method insures that the sum over the households in each income class will equal the total households for a geographic area. The calculation of the household income distribution is done by:

$estinc_i = hh * p_i$, where i is the income class; $estinc$ is the estimated household income distribution; and hh is the number of households.

CALIBRATION

The modified lognormal curve is calibrated to the 2000 regional income distribution. Using the median income and income distribution as inputs, the calibration routine estimates the other two parameters needed to fully specify the modified lognormal income distribution model (parm and exp). An iterative method is used to find the parameters that minimize the error between the observed and the estimated income distributions. This error is measured by the index of dissimilarity, which is computed by comparing two percentage distributions as follows:

$iod = \sum |act_i - est_i| * .5$, where \sum is the summation across income groups; $|$ is the absolute value; act is the observed percent of households in an income group; and est is the estimated percent based on the modified lognormal curve.

IOD is interpreted as the percentage that the estimated distribution would have to change in order to match the actual distribution.

APPENDIX B GENERALIZED DEVELOPMENT TYPE CODES

UDM uses the generalized development type categories to determine the kind of land development requires (i.e., vacant, redevelopment, or infill). It determines LUZ-level housing unit and employment capacities for these categories from a code on the capacity file assigned using the development type code (trans.devcode) and the 4-digit proposed land use code (trans.plu). The specific codes that create the generalized development type categories are shown below.

The codes for the six employment categories are:

Category 1 combines two kinds of redevelopment: 1) residential to employment redevelopment (trans.devcode = 7); and 2) employment to employment redevelopment (trans.devcode = 12);

Category 2 represents employment infill (trans.devcode = 4); and

Categories 3 to 6 partition vacant land and agricultural redevelopment land (trans.devcode 3 and trans.devcode 10, respectively) into four categories: industrial, commercial, office, and schools using the trans.plu to generalized land-use category cross-reference table (see Table 2.3-2).

The four generalized single family categories are:

Category 1 combines two kinds of redevelopment: 1) mobile home to single family redevelopment (trans.devcode = 9 and trans.plu = 1100); and 2) employment to single family redevelopment (trans.devcode = 11 and trans.plu = 1100);

Category 2 represents infill development (trans.devcode = 5); and

Categories 3 and 4 partition vacant land and agricultural redevelopment land (trans.devcode 3 and trans.devcode 10 respectively) using trans.plu = 1000 for low density and trans.plu = 1100 for urban density.

The three generalized multiple family categories are:

Category 1 combines three kinds of redevelopment: 1) mobile home to multiple family redevelopment (trans.devcode = 9 and trans.plu = 1200); 2) single family to multiple family (trans.devcode = 8); and 3) employment to multiple family redevelopment (trans.devcode = 11 and trans.plu = 1200);

Category 2 represents infill development (trans.devcode = 6); and

Category 3 partitions vacant land and agricultural redevelopment land (trans.devcode 3 and trans.devcode 10 respectively) using trans.plu = 1200.