Appendix A Implementation of Model for Simulations

A.1 Calculation of Aggregate Totals

The model presented in chapter 8 operates at the level of individual households. For the California Energy Commission, demand simulations were required for the state as a whole rather than for individual households. The method for obtaining aggregate, statewide simulations from the household level model is described in this section.

Estimates of statewide totals are obtained by (1) simulating the demands of each household within a sample from the state, and (2) taking the weighted sum of the simulated household demands over all sampled households, with the weights reflecting the sampling proportion of each household. The details of this procedure are the following. Consider a sample of N households, with each household labeled $n = 1, \ldots, N$. Each sampled household has some weight associated with it, representing the number of households similar to it in the population. (This weight, for samples based on exogenous factors, is the inverse of the probability that the household was chosen for the sample.) Label the weight for household n as w_n. Note that if the sample is purely random, then w_n is the same for all n; if the sample is stratified random, then w_n is the same for all n within a stratum.

For each sampled household, the following probabilities are calculated with the model.

Vehicle Quantity Submodel:

P_0^n = probability that household n chooses to own no vehicles;

P_1^n = probability that household n chooses to own one vehicle;

P_2^n = probability that household n chooses to own one or two vehicles.

Class/Vintage Submodel for One Vehicle:

$P_{i|1}^n$ = probability that household n chooses to own a vehicle in class/vintage i, given that it chooses to own one vehicle (for each class/vintage).

Class/Vintage Submodel for Two Vehicles:

$P_{ij|2}^n$ = probability that household n chooses to own a vehicle in class/vintage i and a vehicle in class/vintage j, given that it chooses to own two vehicles (for each pair of class/vintages).

The VMT submodels calculate the vehicle miles traveled for each class and vintage of vehicle:

VMT^n_{i1} = vehicle miles traveled by household n, given that it chooses to own one vehicle of class/vintage i;

VMT^n_{ij2} = vehicle miles traveled by household n in a vehicle of class/vintage i, given that the household chooses to own two vehicles, with one being a vehicle in class/vintage j in addition to the vehicle in class/vintage i.

With these calculated numbers, the **expected** number of vehicles that household n will own of class/vintage i can be determined as follows:

N^n_i = expected number of vehicles that household n will own of class/vintage i

= (probability of owning one vehicle) × (probability of choosing class/vintage i, given one vehicle)

+ (probability of owning two vehicles) × (probability of choosing a pair of vehicles that includes one in class/vintage i)

$$= [P^n_1 \times P^n_{i|1}] + \left[P^n_2 \times \sum_j P^n_{ij|2}\right] + [P^n_2 \times P^n_{ii|2}],$$

where the last term allows for the possibility that the household can own two vehicles of the same class/vintage. Similarly, the expected VMT that household n will drive on vehicles of class/vintage i is

$$\mathrm{VMT}^n_i = [P^n_1 \times P_{i|1} \times \mathrm{VMT}^n_{i1}] + \left[P^n_2 \times \sum_j P^n_{ij|2} \times \mathrm{VMT}^n_{ij2}\right] + P^n_2 \times P^n_{ii|2} \times \mathrm{VMT}^n_{ii2}.$$

Estimates of statewide totals are calculated once the expected number of vehicles and VMT in each class/vintage have been calculated for each sampled household. The total number of vehicles of class/vintage i in the state is calculated as the weighted sum of the expected number for each sampled household:

$$\bar{N}_i = \sum_n w_n N^n_i.$$

$$\overline{\text{VMT}}_i = \sum_n w_n \text{VMT}_i^n.$$

Fuel consumption is calculated as the last step. The amount of fuel consumed by vehicles in class/vintage i is estimated as the number of miles traveled divided by average fuel efficiency (miles per gallon) for vehicles in that class/vintage:

F_i = statewide fuel consumption by personal use vehicles in class/vintage i

$$= \overline{\text{VMT}}_i / e_i,$$

where e_i is the fuel efficiency of vehicles in class/vintage i.

A.2 Inputs

Vehicles

The California Energy Commission, with assistance from Energy and Environmental Analysis, Inc., calculated the average characteristics of each class of vehicle built from 1971 to 1980 and projected the characteristics of vehicles that would be built in the simulation years 1980–2000. These data are given in appendix B. For the 1971–1980 vehicles, the price given in appendix B is the cost of buying that vehicle in the first simulation year, 1980 (e.g., the cost of buying a 1975 compact in 1980, given that it is five years old). For vehicles projected to be built after 1980, the price represents the cost of buying that vehicle new when it is produced (e.g., the cost of buying a 1990 compact in 1990). To represent the fact that vehicles decrease in price as they become older, depreciation rates were applied to the price of each vehicle in moving from one simulation year to another.

The depreciation rates are given in table A.1. While the same depreciation rates were applied in all years to all classes of vehicles, the model is capable of handling different rates in different years and for different classes.

Socioeconomic Variables

The sample of households and projections of socioeconomic variables is described in section 9.2.

Fuel Prices

The projections of fuel price that the Commission specified are shown in table A.2. All prices are listed in 1978 cents/gallon except electricity, which

Table A.1
Depreciation rates used in simulations

Age of vehicle (years)	Price as proportion of price when new (both prices in 1978 dollars)
1	0.82
2	0.67
3	0.54
4	0.43
5	0.35
6	0.28
7	0.23
8	0.19
9+	0.12

Transit Trips per Capita

For each household, the model takes as input the number of transit trips per capita that are taken in the household's metropolitan area. As discussed in chapter 8, this variable is used as a proxy for the quality of transit in the household's area, which affects the household's ownership and usage of autos.

The Commission's projections were obtained from System Design Concepts, Inc. For each of the seven metropolitan areas with transit service the actual number in 1980 and the projected number for 2004 are given in table A.3; the figures for years 1981–2000 were calculated by linear interpolation. For households living in other areas of California, the number of transit trips per capita in the area was projected to be zero throughout the simulation period.

A.3 Sampling of Class/Vintage Pairs for Two-Vehicle Households

The class/vintage submodel for two-vehicle households calculates, for each household, the probability that the household will choose each pair of class/vintages. For a reasonable number of classes and vintages of vehicles, the number of pairs of class/vintages is very large. For example, with 25 classes of vehicles and 10 vintages (for a total of 250 class/vintages), there are over 30,000 pairs of class/vintages.

Table A.2
Projected fuel prices, as specified by the California Energy Commission (in 1978 cents per gallon or per KWH)

Year	Gas	Diesel	Electric	Methanol	LPG
1980	97.50	89.70	4.312		
1981	97.50	88.92	4.231		
1982	85.02	82.68	4.764	54.60	65.52
1983	89.70	87.36	4.816	57.72	68.64
1984	89.70	86.58	4.809	56.94	67.86
1985	89.70	87.36	4.985	57.72	67.86
1986	89.70	87.36	5.096	56.94	68.64
1987	89.70	87.36	5.174	57.72	68.64
1988	92.04	89.70	5.275	58.50	69.42
1989	93.60	91.26	5.378	60.06	71.76
1990	98.94	93.60	5.385	61.62	73.32
1991	97.50	95.16	5.478	62.40	74.10
1992	100.62	97.50	5.504	63.96	76.44
1993	102.96	99.84	5.541	65.52	78.00
1994	106.08	101.40	5.469	67.08	80.34
1995	109.20	105.30	5.698	70.20	82.68
1996	113.88	109.98	5.900	71.76	85.80
1997	118.56	113.88	6.006	74.88	89.70
1998	123.24	118.56	6.117	78.00	93.60
1999	128.70	124.02	6.201	81.12	96.72
2000	133.38	128.70	6.141	84.24	101.40
2001	140.40	134.16	6.234	87.36	105.30
2002	145.86	139.62	6.327	92.04	109.20
2003	152.10	145.08	6.423	95.94	114.66
2004	158.34	152.10	6.518	99.84	119.34

Table A.3
Number of transit trips per capita used in simulations

Metropolitan area	Actual for 1980	Projected by Commission for 2004
Anaheim/Santa Ana/Garden Grove	15.20	9.86
Fresno	16.49	14.50
Los Angeles/Long Beach	57.24	58.99
Sacramento	19.33	20.33
San Diego	20.96	18.76
San Francisco/Oakland	98.73	85.43
San Jose	19.96	19.41

While a large number of pairs of class/vintages poses no theoretical problem, it entails a practical problem due to the high cost of calculating probabilities for each pair of class/vintages. Early runs of the model, in which the probability of each class/vintage pair was calculated, cost much more money that was considered feasible to spend on a per run basis. Consequently, the code was rewritten to sample, for each household, every ninth class/vintage pair and calculate the probability of only the sampled pairs. This reduced the cost of running the entire model by a factor of nearly nine (since nearly all the costs were incurred in the class/vintage submodel for two-vehicle households).

To assure that all class/vintage pairs were represented, the sampling of pairs was cycled across households. That is, the first household was programmed to sample the first, tenth, nineteenth, and so on, class/vintage pairs. The second household sampled the second, eleventh, twentieth, and so on, class/vintage pairs. And so on, to the ninth household, with sampled the ninth, eighteenth, twenty-seventh, and so on pairs. The ninth household completed a cycle with each class/vintage pair selected once by one of the first nine households. The cycle was then repeated with the tenth household choosing the first, tenth, nineteenth, and so on class/vintage pairs.

A.4 Recalibration of Alternative-Specific Constants

Each of the submodels that are logit (i.e., all except the VMT submodels) take the form

$$P_i = \frac{e^{bz_i + a_i}}{\sum_j e^{bz_j + a_j}},$$

where z_i is a vector of variables relating to alternative i, b is a vector of parameters, and a_i is a constant term. The constant a_i represents the average impact of all variables that are not included in the model.

In estimation, the value of each a_i is determined along with the other model parameters, b. The estimated value of a_i is the average in the estimation sample of the unincluded terms. It is chosen by the estimation routine so that the number of households in the sample predicted to choose each alternative is exactly equal to the number in the sample that actually chose it.

In simulating demands for the California Energy Commission, a sample is used that is different from the estimation sample. (The simulation sample is from California, while the estimation sample is nationwide.) Since the average of unincluded variables is necessarily different for the simulation sample, new values of a_i needed to be calculated, in the manner described in section 6.3. For each submodel, new values of a_i were chosen so that the simulated number of households choosing alternative i in the first year of simulation, 1980, exactly equaled the number of households that actually chose that alternative in 1980. The procedure for calibrating each a_i is described in the following for each submodel.

Vehicle Quantity Submodel

Two alternative specific constants were estimated in the vehicle quantity submodel: a_1 for the alternative of owning one vehicle and a_2 for the alternative of owning two vehicles. As shown in table 8.1, the estimated values of a_1 and a_2 are -1.79 and -4.95, respectively.

Let A_1 denote the proportion of households in the simulation sample that chose to own one vehicle, and A_2 denote the proportion that chose two vehicles. The model was run with the original values of a_1 and a_2, labeled a_1^0 and a_2^0, and the number of households that would choose to own one and two vehicles was simulated. Let the simulated proportion of households choosing one and two vehicles be denoted S_1^0 and S_2^0, respectively, where the superscripts refer to the fact that the simulation is based on a_1^0 and a_2^0.

The model with its original values of a_1 and a_2 underpredicts the share of households choosing to own one vehicle if A_1 is greater than S_1^0 and overpredicts if S_1^0 is greater than A_1 (similarly for the share choosing two vehicles). To correct the misprediction (or, more precisely, to adjust the constants so that they represent the average of unincluded variables in the simulation sample), a_1 and a_2 are adjusted to new values using the formula

$$a_i^1 = a_i^0 + \ln(A_i/S_i^0), \qquad i = 1, 2.$$

Note that if A_i is greater than S_i^0 so that the model is underpredicting the share of households choosing i vehicles, then the adjustment increases a_i. Conversely, if S_i^0 is greater than A_i and the model is overpredicting, a_i is adjusted downward.

The adjustment just described completes the first iteration of the recalibration procedure. For the second iteration, the model is run with the new values of a_i (that is, with a_1^1 and a_2^1) and new simulation shares are obtained, labeled S_1^1 and S_2^1. If S_1^1 and S_2^1 are not equal to A_1 and A_2, respectively, then the constants are adjusted again, using the formula

$$a_i^2 = a_i^1 + \ln(A_i/S_i^1), \qquad i = 1, 2,$$

where a_i^2 is the twice-adjusted value of a_i. This process is continued, obtaining new values of a_1 and a_2 with each iteration, until the simulated proportion of households choosing each number of vehicles equals (or is very close to) the actual proportion in the sample.

In the sample of households used for the base case simulations, 34.3% owned one vehicle and 55.2% owned two vehicles. Using the procedure described, the constants were recalibrated to values of −9.795 for the alternative of owning one vehicle and −13.087 for the alternative of owning two vehicles. These values replace, in the simulation code, the values of table 8.1.

Class/Vintage Submodels

A separate constant for each class and vintage of vehicle was not calibrated, due to the unwieldy number of class/vintages that were considered available for the simulations. Rather, one constant was estimated for each class of vehicles and applied to all vintages within that class. The procedure for estimating these constants is the following. First, the model was run to simulate the number of vehicles in each class/vintage. Second, the simulated number of vehicles in each class was calculated by summing over vintages the number simulated in each class/vintage. Third, the class specific constants in the model were adjusted using the formula last displayed for the vehicle quantity submodel.

Table A.4
Class distribution of vehicles owned in California in 1980

Class		Percent
2	Domestic gas subcompact car	8.7
3	Domestic gas compact car	12.7
4	Domestic gas large car	35.2
5	Foreign regular gas car	22.3
6	Foreign luxury gas car	3.2
10	Domestic diesel compact car	0.3
11	Foreign regular diesel car	0.1
12	Foreign luxury diesel car	0.3
18	Small gas pickups and utility vehicles	5.1
19	Large gas pickups and utility vehicles	8.0
21	Large diesel pickups and utility vehicles	0.05
23	Large gas vans and other vehicles	4.0

Rather than using, as the basis for recalibration, the share of vehicles in each class in the sample, the Commission provided information on the share of personal use vehicles in each class in California. By using the true statewide proportions for 1980 rather than the sample proportions, the effect of sampling errors are mitigated.

The proportion of personal use vehicles owned in each class that was available in 1980 is given in table A.4. The values of the constants that were obtained for each class of vehicle available in 1980 are given in table A.5. For classes of vehicles that were not available in 1980 (such as methanol vehicles), the procedure described previously cannot be used to obtain constants. Rather, constants for these vehicles must be assigned on the basis of reasonable notions concerning the similarity of different classes of vehicles.

Recall that the constant for each class of vehicles captures the average effect of all variables that are not not included in the model. The task, therefore, is to determine which class of vehicles available in 1980 most closely corresponds, with respect to factors not included in the model, to each new class of vehicle. The constant for the corresponding class of existing vehicles is assigned to the new class of vehicles.

For example, mini gas cars were not available in 1980, and so a constant for that class could not be estimated. It seems reasonable, however, to

Table A.5
Constants calibrated for each vehicle class

Classes of vehicles available in 1980		Constant
2	Domestic gas subcompact car	1.378
3	Domestic gas compact car	1.749
4	Domestic gas large car	2.298
5	Foreign regular gas car	2.276
6	Foreign luxury gas car	1.798
10	Domestic diesel compact car	−1.244
11	Foreign regular diesel car	−1.468
12	Foreign luxury diesel car	−0.169
18	Small gas pickups and utility vehicles	−0.038
19	Large gas pickups and utility vehicles	0.997
21	Large diesel pickups and utility vehicles	−2.347
23	Large gas vans and other vehicles	−4.955

assume that the unincluded factors affecting the desirability of subcompact gas cars, which were available in 1980 and for which a constant was calibrated, are similar to those of mini gas cars (or at least more similar than any other class of vehicle available in 1980). Consequently, mini gas cars were assigned a constant equal to the value calibrated for subcompact gas cars.

The assignment of constants for each new class of vehicle that was used in producing the base case simulations is given in table A.6. Note that each new, non-gas-powered vehicle is assigned the constant for an existing diesel class. This assignment reflects the fact that the unobserved factors that will probably be most important in determining the demand for these new vehicles—namely, uncertainty by the public about new types of vehicles, questions regarding the difficulty of customers' obtaining the alternative fuels, and so on—are currently being experienced for diesel vehicles.

Submodels for Proportion of VMT in Each Category

The Commission provided the following estimates of the proportion of VMT in each category:

1. short intracity work trips, 0.20;
2. short intracity nonwork trips, 0.54;

Table A.6
Constants assigned to vehicle classes not available in 1980

New class	Class available in 1980 that was judged most similar in unincluded factors	Value of constant
1 Gas mini cars	2 Gas subcompact cars	1.378
7 Diesel mini cars	11 Foreign regular diesel cars	−1.468
8 Domestic diesel subcompact cars	11 Foreign regular diesel cars	−1.468
9 Domestic diesel compact cars	11 Foreign regular diesel cars	−1.468
13 Electric cars	11 Foreign regular diesel cars	−1.468
14 Methanol compact cars	11 Foreign regular diesel cars	−1.468
15 Methanol large cars	10 Large diesel cars	−1.244
16 LPG compact cars	11 Foreign regular diesel cars	−1.468
17 LPG large cars	10 Large diesel cars	−1.244
20 Small diesel pickups and utility vehicles	21 Large diesel pickups and utility vehicles	−2.347
22 Small gas vans	23 Large gas vans and other vehicles	−4.955
24 Small diesel vans	23 Large gas vans and other vehicles	−4.955
25 Large diesel vans	23 Large gas vans and other vehicles	−4.955

3. other work trips, 0.20;
4. other nonwork trips, 0.06.

These proportions were taken as the actual proportions of VMT in each category, and the procedure described previously for the vehicle quantity submodel was used to calibrate an alternative-specific constant for each of these categories. The resulting constants are

1. short intracity work trips, 5.414;
2. short intracity nonwork trips, 6.806;
3. other work trips, 2.024;
4. other nonwork trips, 0.0.