

NCTCOG Mode Choice Model Estimation

This report describes the estimation of mode choice models by Cambridge Systematics for the North Central Texas Council of Governments (NCTCOG). The logit mode choice models are applied to the trip table outputs of the NCTCOG trip distribution modeling process. The inputs to the mode choice modeling process are daily person trip tables by purpose for trips that are internal to the NCTCOG region. The outputs of the mode choice process are person trip tables by purpose and mode.

1.0 Technical Approach

1.1 Logit Mode Choice Model

The most common mode choice modeling procedure is the logit model. Logit models are discrete choice models, which attempt to explain the behavior of individuals making a choice between a finite number of separate alternatives, in this case travel modes. In the logit model, the probability of choosing a particular alternative i is given by the following formula:

$$P(i) = \exp(U_i) / \sum_j \exp(U_j)$$

where:

$P(i)$ = probability of choosing alternative i

U_i = utility of alternative i

\exp = exponential function

The utility function U_i represents the worth of alternative i compared to other alternatives and is expressed as a linear function:

$$U_i = B_{0i} + B_{1i}X_{1i} + B_{2i}X_{2i} + \dots + B_{ni}X_{ni}$$

where the X_{ki} variables represent attributes of alternative i , the decision maker, or the environment in which the choice is made and B_{ki} represents the coefficient reflecting the effect of variable X_{ki} on the utility of alternative i . The coefficients are estimated using statistical “maximum likelihood” methods using logit model estimation software such as ALOGIT. In the case of logit mode choice models, the alternatives are the modes while

the attributes may include attributes of the modes (e.g., travel time), the decision maker (e.g., auto ownership), and the environment (e.g., production zone population density).

1.2 Use of Nested Model Structures

It was desirable to attempt to develop a nested logit mode choice model for the NCTCOG travel model system. Over the past decade, advanced statistical estimation procedures have become available which allow nested models to be estimated efficiently using individual observations of tripmaking by mode. Such data are available from NCTCOG's household and transit on-board surveys.

A major advantage of nested structures is that similar modes, such as transit with auto access and transit with walk access, can be grouped as a subset, all branching from a common "composite mode." A nesting parameter, which represents the degree to which the elemental modes are more strongly related to each other than to any modes in other parts of the model, is estimated for this composite mode.

In the nested logit model, the probability of choosing an alternative i in nest n is given by:

$$P(i) = P(i | \text{nest } n) P(\text{nest } n)$$

where the probability of nest n is given by the logit formula presented in Section 1.1 and the alternatives are all other nests of alternatives at the same nesting level. The conditional probability $P(i | \text{nest } n)$ is given by the logit formula where the alternatives over which the exponentiated utilities are summed include only the other alternatives within nest n .

The primary advantage of nested logit models over non-nested multinomial logit models is that the nested logit models reduce the intensity of the "independence of irrelevant attributes" (IIA) property. The IIA property, which is characteristic of all multinomial logit models as well as the lowest level nests in nested logit models, assumes that the relative shares of any two modes are independent of the availability of other modes. For example, assume there are three modes: auto, bus, and rail. It might be reasonable to assume that the ratio of the choice probabilities of bus and rail is independent of whether auto is available. However, it would likely be unreasonable to assume that the ratio of the bus and auto shares is independent of the availability of rail; adding the rail mode to a market would likely draw more bus users than auto travelers.

1.3 Model Estimation and Testing Procedure

The model estimation approach included the development of the nested model structures using, to the greatest extent possible within the limitations of the available data, a maximum likelihood logit estimation process, which provided "full information" estimates of all model parameters simultaneously. This strategy maximized the travel behavior information represented in the expanded models and provided a flexible framework for model development.

In summary, the mode choice model estimation and application process included the following steps:

1. Assembly of model estimation data sets from the household survey and transit on-board survey trip records as previously assembled by NCTCOG, and highway and transit network skims and zonal data provided by NCTCOG;
2. Multinomial (non-nested) model estimation using all modal alternatives for each trip purpose using the ALOGIT choice behavior analysis package; and
3. Nested model estimation for each trip purpose.

2.0 Data for Model Estimation and Application

2.1 Survey Data Set

The best data source for estimation of a mode choice model is data from a local household travel survey which contains information about all trips made by respondent households. Besides the chosen mode, the following information is required for each trip:

- Trip purpose;
- Location of origin and destination;
- Information about the trip, such as the time of day it took place; and
- Characteristics of the household and the individual trip-maker, such as auto availability, income, number of persons and workers in the household, etc.

Although there were a large number of households represented in the 1996 NCTCOG household survey, a relatively small number of reported trips were made by transit. It was therefore necessary to enrich the estimation data set by including records from the 1998 DART and 1996 FWTA transit on-board surveys.

The survey data for each trip are supplemented by information about the travel time and cost between the origin and destination and about the origin and destination areas. The travel time and cost data, referred to as *level of service data*, are obtained by skimming the model system's highway and transit networks for the given origin and destination of each trip. The level of service data for transit modes included wait, transfer, walk access, auto access, and egress times; number of transfers; in-vehicle times by transit mode; and transit fares. The level of service data for highway modes included in-vehicle times, out-of-vehicle times, and distances.

The information about the origin and destination areas, known as *zonal data*, is usually developed on a zone-specific basis. Available variables included area type, levels of

population and employment by type, and parking costs. Other variables, such as population and employment density, were easily calculated using the land area for each zone. The complete set of variables tested in the mode choice models is described in Section 3.5.

The survey, zonal, and level of service data were merged to provide estimation data sets for each trip purpose. These data sets consisted of the survey trip records, extended to include household and person variables from the surveys; zonal data for the zones of trip origin, destination, production, and attraction (as appropriate); and level of service data for all modes available between the trip end zones.

The estimation of disaggregate discrete choice models such as logit models does not require that data be weighted to reflect either the actual choice probabilities of the alternatives in the general population or any attribute differences between the decision makers represented in the data set and those in the general population. Therefore, survey data weights were not applied in the model estimation process. These weights are necessary only for aggregate model estimation (for example, the computation of the time of day factors).

2.2 Observation Exclusions

There are a total of 56,095 trip observations in the original combined data set provided by NCTCOG. The numbers of observations by survey data source are shown in Table 1. It was necessary to remove from the data set observations for which required data was unavailable. The following criteria were used to exclude observations that could not be used for estimation:

- Intrazonal trips (since there is no available level of service information from the network skims);
- Origin or destination zone missing;
- Chosen mode not available; or
- Household auto ownership information missing (for home based work trips).

The numbers of observations excluded from each data source for missing household or trip information, or because they represented intrazonal trips, are also shown in Table 1.

It was also necessary to remove observations where there was no valid mode or trip purpose coded. There were a total of 3,588 trip records where this occurred, and these records were also dropped from the data set. This left a total of 35,377 observations in the data set.

Table 1. Survey Data Sources

Data Source	Original Number of Observations	Missing Household Data	Missing Origin/ Destination or Intrazonal	Remaining Number of Observations
Household survey – RDD	31,928	433	7,189	24,306
Household survey – transit intercept	2,502	24	715	1,763
DART – bus	16,259	1,953	5,231	9,075
DART – rail	1,331	123	392	816
FWTA	4,075	255	815	3,005
Total	56,095	2,788	14,342	38,965

3.0 Estimation of the Mode Choice Model

3.1 Modes and Trip Purposes

The number of trips in the household survey data set by trip purpose and detailed travel mode, as well as the characteristics of the transportation system in the NCTCOG region, determined the modes and trip purposes that could be considered in the mode choice model. The trip purposes chosen for the mode choice model by NCTCOG are:

- Home Based Work (HBW);
- Home Based Non-Work (HBNW); and
- Non-Home Based (NHB).

The modes considered for inclusion in the model are:

- Auto - drive alone;
- Auto - two occupants;
- Auto - three or more occupants;
- Transit – Auto access; and
- Transit – Walk access.

The number of observations by chosen mode and purpose in the original estimation data set is provided in Table 2.

Table 2. Distribution by Chosen Mode and Purpose

Mode Description	HBW	HBNW	NHB
Auto - drive alone	5,376	4,773	3,109
Auto - two occupants	508	3,583	1,314
Auto - three or more occupants	137	3,469	1,039
Transit – Walk access	4,480	3,311	1,688
Transit – Auto access	206	128	100

3.2 Path Building and Skim Table Strategies

TransCAD path building and skimming procedures were used by NCTCOG to create the level of service (LOS) data. These programs were run once for the auto modes and once for each of the two transit modes. The TransCAD pathfinder path building procedure was used.

The average auto operating cost is assumed to be 7.3 cents per mile (in 1999 dollars), which was determined during earlier work by NCTCOG. This figure is similar to those used in some models recently developed by Cambridge Systematics. The average weekday parking costs (in 1999 dollars) were provided by NCTCOG. Parking cost was divided by two so that the cost was allocated to each leg of a round trip. Auto operating and parking costs were divided by auto occupancy for each auto mode. The average auto occupancy per person trip for the auto - three or more occupants for each purpose was estimated directly from unweighted household survey data by trip purpose:

- Home Based Work (HBW): 4.73
- Home Based Non-Work (HBNW): 3.80
- Non-Home Based (NHB): 4.36

It was necessary to define maximum walk access, auto access, and transit transfer times for the transit skims. These limits are:

- 1.0 miles (20 minutes) for walk access and transit transfers;
- 45 minutes for auto access; and
- 3 transit transfers per trip.

Weights to time and cost variables

Initial weights had to be given to time and cost variables within the generalized cost used to determine the shortest transit paths. For model application, these weights, for variables such as cost (i.e. value of time), transfer time, and wait time, will eventually be determined based on the mode choice model parameter estimates but had to be specified initially so that transit paths could be built for the model estimation data set.

Based on the weights (ratio of wait time coefficient to in-vehicle time coefficient) by trip purpose in the previous NCTCOG mode choice model, **it was recommended that 2.0 be used as the weight for wait time and walk time of HBW trips, and 2.5 for HNW and NHB trips.**

The weights for other time variables (in-vehicle time, transfer penalty, and dwell time) were set to 1.0. These variables either represent components of in-vehicle time or, in the case of the transfer penalty, do not represent actual time at all.

The values of time from the previous NCTCOG mode choice model (in 1984 \$/hour) were:

- Home based work: \$3.83
- Home based non-work: \$0.97
- Non-home based: \$1.66

To combine these different values, we used a weight average of the three trip purposes, using weights based on the trip purposes from the DART and FWTA on-board surveys. These weights were:

- Home based work: 0.354
- Home based non-work: 0.353
- Non-home based: 0.293

Using these weights resulted in a value of time of \$2.20/hour. This translates to \$3.33 in 1999 dollars, using Dallas area-specific inflation rates. Value of time is represented in TransCAD in \$/minute, and so **the recommended value of time was \$0.0555/minute.**

Maximums and minimums for level of service variables

The following values were used:

- Maximum number of transfers – Commonly, any number from 2 to 5 is used. Cambridge Systematics has successfully used a value of 3 in other applications, and so **a value of 3 was used.**

- Maximum wait time – The value of time is generally calculated as half of the headway, which assumes random arrivals at the transit stop. However, when headways are very long, some riders tend to know the schedule and do not start arriving until a shorter time before the arrival of the transit vehicle. In other areas, maximum wait time values of 15 to 20 minutes are typical. **A value of 15 minutes for was used for NCTCOG.**
- Minimum wait time – For a variety of reasons (that transit vehicles do not always stay exactly on schedule, that persons do not want to risk missing their vehicle, etc.), TransCAD allows the input of a minimum wait time. Cambridge Systematics has successfully used a value of 2 minutes in other applications, and so **a value of 2 minutes was used.**
- Layover time – Layover time does not appear to affect the skims from a demand standpoint, and so using a common sense value is adequate. **A value of 5 minutes was used.**
- Maximum walk access (and egress and transfer) time - Documentation of several travel models throughout the U.S. was reviewed. Most of the documentation reports available do not provide information on the maximum distances used for transit walk access. The following table shows the maximum walk distances to transit in several metropolitan areas for which information was available:

Urban Area	Maximum Walk Distance (miles)
Atlanta	0.50
Boston	2.00
Detroit	0.75
Miami/Tampa	0.33 - 1.00
New York (preliminary)	0.75
Norfolk	0.75
Salt Lake	0.40

Note that the value of 0.75 listed for Detroit was recently chosen by Cambridge Systematics based on the other urban areas in the list. While there is considerable variation in the distances used in other urban areas, the value of 0.75 seems to be an acceptable average, and did work well in Detroit. Through consultation with NCTCOG, a higher value of 1.00 miles was used so that, at an assumed 3.0 mph walking speed, the maximum walk time would be calculated as 20 minutes.

Since TransCAD uses maximum walk time, rather than distance, as the variable, the walk distance must be converted to minutes. At 3 mph, 1.00 miles works out to 20 minutes. **The recommended maximum walk access, egress, and transfer times were 20 minutes.**

- Maximum total trip time – This is another “common sense” figure that is usually set to be greater than the longest travel time by transit that would be expected in the region. **A value of 240 minutes was used.**

- Maximum total trip cost – This is another “common sense” figure that is usually set to be greater than the highest cost for a transit trip that would be expected in the region. **A value of \$99 was used.**

Other parameters

- Dwell time – Based on discussion with NCTCOG, values for buses varied by mode and range from 25 to 29 seconds (0.42 to 0.48 minutes) per coded bus stop. Dwell times were not used for train travel, since the scheduled stop-to-stop rail times that were input directly by NCTCOG included the dwell times.
- Layover time – Layover time does not appear to affect the skims from a demand standpoint, and so using a common sense value is adequate. A value of 5 minutes was used.

3.3 Unavailability of Modes

It was assumed that the auto modes were available to all travelers. Households without autos are assumed to be able to be auto passengers; this assumption was borne out by the fact that households without cars made auto trips in the household survey data set. It was also assumed that each transit submode (walk or auto access) was available to any trip where a transit skim was built in accordance with the limits described in Section 3.2. In the data set, transit in-vehicle and out-of-vehicle times and fares were set to zero or -99 for any trip for which a path could not be built within these limits. Therefore, a transit mode was deemed unavailable in any record where these times were zero or negative.

A total of 2,166 of the 35,377 observations were excluded automatically by ALOGIT because of the unavailability of the chosen mode, in accordance with the criteria described above. The number of observations available, after exclusions, for each trip purpose was as follows:

- Home based work – 10,707
- Home based non-work – 15,264
- Non-home based – 7,250

3.4 Model Estimation Process

The estimation of nested logit mode choice models for NCTCOG was a two-step process. Before estimating the more complicated nested models, multinomial logit models were estimated to explore initial structures and data problems. This information was then used to develop nested logit models.

This strategy was followed for each trip purpose. Since statistical model estimation is a trial-and-error process, a large number of model specifications were tested. The nesting

structures for the new mode choice models were prespecified with the two transit submodes included in a transit nest. Because there were only three modes in the model, there were no other logical nesting structures to test.

Before beginning model estimation, it was necessary to define the variables that would be tested as part of the utility functions. There were three main criteria for a variable to be usable in the utility functions:

1. Data for the variable must be available in the model estimation data set, with sufficient variability to provide statistically significant estimates of its coefficient.
2. The variable must be available for use in the model application, from skimming the highway or transit network, from the zonal data set, or from another source.
3. The variable must be able to be forecast so that it can be used in model application.

Variables based on household data, such as auto availability, cannot be used in non-home based models since the location of the home of the tripmaker is not known during model application.

The following variables were available for testing in the model estimation process:

Mode-specific constants:

AUTO2	Auto-two occupants constant
AUTO3	Auto-three or more occupants constant
TRANAUTO	Transit-auto access constant
TRANWALK	Transit-walk access constant

Level of service variables (mode dependent—can be mode specific or generic):

AUTOIVTT	Auto travel time in minutes (auto modes only)
TRANIVTT	Transit travel time in minutes (includes time in auto or transit vehicles, including auto access time for transit trips with auto access) (transit modes only)
AUTOOVTT	Auto out-of-vehicle time (i.e. terminal time) in minutes (auto modes only)
WAACTIME	Walk access time at production end in minutes (transit walk access only)
WAITTIME	Initial wait time in minutes (transit modes only)
XFERTIME	Transit transfer time in minutes (transit modes only)

AUACTIME	Auto access time at production end in minutes (transit auto access only)
EGRTIME	Walk egress time at attraction end in minutes (transit modes only)
TRANOVTT	Transit out-of-vehicle time (sum of walk access, wait, transfer, and walk egress time) in minutes (transit modes only)
FARE	Fare in dollars (transit modes only)
AUOPCOST	Auto operating cost in dollars (auto modes only)
PARKCOST	Parking cost in dollars (auto modes only)
AUTOCOST	Total auto cost (sum of auto operating and parking costs) in dollars (auto modes only)
XFERS	Number of transit transfers (transit modes only)

Zonal variables (must be mode-specific):

POPDENP	Population density at the production zone in persons per acre
EMPDENA	Employment density at the attraction zone in employees per acre
CBD	Attraction zone is CBD area type
OBD	Attraction zone is OBD area type
SUBURB	Attraction zone is suburb area type
URBAN	Attraction zone is urban area type
RURAL	Attraction zone is rural area type

Household variables (must be mode-specific):

LOWINC	1 if household income is less than \$30,000, 0 otherwise
MEDINC	1 if household income is between \$30,000 and \$75,000, 0 otherwise
HIGHINC	1 if household income is greater than \$75,000, 0 otherwise
AUTOS	number of autos in household
HHSIZE	number of persons in household
AULTPER	1 if fewer autos in household than persons, 0 otherwise

3.5. Model Estimation Results

3.5.1 Home Based Work Trips

A number of model specifications were run on the estimation data set. The variables and their forms finally retained in the model were based on how well the estimated coefficients conformed to the expectations on their signs and relative magnitudes and on their statistical significance. The final multinomial model specification was used in testing the applicability of various nested structures. The estimation results for both the multinomial and nested model structures are described below.

Table 3 presents the coefficient estimates for the final multinomial model specification for home based work trips, with t-statistics in parentheses and other statistics shown at the bottom. It was necessary to constrain the model by setting the transit fare coefficient to -0.550 and the auto cost coefficient to -0.770. These values were chosen so as to obtain similar values, close to the national averages, for the auto and transit cost coefficients for the nested model. As Table 3 shows, all coefficients are statistically significant, and the rho-squared values indicate a good statistical fit. The implied values of in-vehicle time from the multinomial model are \$4.02 per hour for auto and \$3.01 per hour for transit. These are computed using the following formula:

$$\text{Value of time} = 60 (\text{In-vehicle time coefficient} / \text{Cost coefficient})$$

where the scaling factor of 60 converts the value of time from dollars per minute to dollars per hour (since the units of the time variables are in minutes).

Other models that were tested include the following:

- Models that did not constrain the cost coefficients. The estimated coefficients for cost were very low, resulting in extremely high values of time.
- Models with other variables, including AUTOOVTT, XFERS, HHSIZE, POPDENP, EMPDENA, and the other area type variables. These variables had coefficients that either were insignificant or had incorrect signs.
- Separating transit out-of-vehicle time into its components. This resulted in high ratios of out-of-vehicle time to in-vehicle time coefficients and low values of time.
- Separating the cost variable into its components. This resulted in insignificant or positive coefficients for some of the cost variables.

Table 3. Home Based Work Multinomial Logit Mode Choice Model Coefficients and Statistics

	Coefficient	T-Statistic
AUTO2	-3.457	-38.3
AUTO3	-5.116	-44.0
TRANWALK	0.410	2.8
TRANAUTO	-1.939	-12.3
AUTOIVTT (auto only)	-0.052	-6.3
TRANIVTT (transit only)	-0.028	-4.9
TRANOVT (transit only)	-0.065	-14.3
AUTOCOST (auto only)	-0.770	(constrained)
FARE (transit only)	-0.550	(constrained)
AULTPER (shared ride only)	1.197	12.6
AULTPER (transit only)	0.852	11.0
LOWINC (shared ride only)	0.439	3.3
LOWINC (transit-walk only)	3.076	34.6
LOWINC (transit-auto only)	2.207	13.3
HIGHINC (shared ride only)	-0.372	-3.8
HIGHINC (transit-walk only)	-5.743	-9.8
HIGHINC (transit-auto only)	-3.599	-4.9
CBD (transit only)	-0.267	-2.0
OBD (transit only)	0.282	2.9
Number of observations	10,707	
Rho-squared w.r.t. zero	0.6341	
Rho-squared w.r.t. constants	0.3765	

(See Section 3.5 for variable definitions.)

Several nesting structures were tested for the home based work model. The structure with the best fit included an auto nest consisting of the three auto modes. The two transit modes were not included in a nest (in effect, each was in its own nest). The final nesting structure is shown in Figure 1.

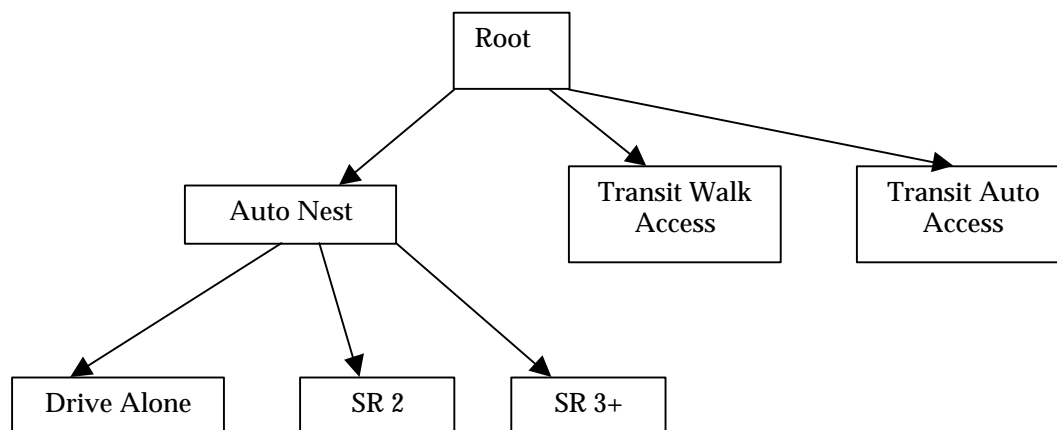
Figure 1. Home Based Work Nesting Structure

Table 4 presents the coefficient estimates for the final nested model specification for home based work trips, with t-statistics in parentheses and other statistics shown at the bottom. This table shows the coefficient values (for the auto modes) with the nesting coefficient for the auto modes already factored in. The variables are the same as in the multinomial model, and the coefficient values are similar to those in the multinomial model, which is not surprising given the relatively high nest coefficient. The implied value of auto in-vehicle time from the nested model is \$5.91, and the implied value of transit in-vehicle time is \$2.73.

The nest coefficient is 0.725 with a t-statistic of 6.6. While this indicates a significant variable (coefficient significantly different from zero), the associated standard error is 0.109. To test whether the nest coefficient is significantly different from one (the value at which a multinomial structure would be indicated), the t-statistic must be computed as:

$$\text{t-statistic} = (1 - \text{coefficient}) / \text{standard error}$$

This computation yields a t-statistic of 2.5, which is significant at the 95% confidence level.

Table 4. Home Based Work Nested Logit Mode Choice Model Coefficients and Statistics

	Coefficient	T-Statistic
AUTO2	-2.518	-38.3
AUTO3	-3.724	-44.0
TRANWALK	0.414	2.8
TRANAUTO	-1.926	-12.3
AUTOIVTT (auto only)	-0.055	-4.5
TRANIVTT (transit only)	-0.025	-4.4
TRANOVTT (transit only)	-0.064	-14.2
AUTOCOST (auto only)	-0.558	(constrained)
FARE (transit only)	-0.550	(constrained)
AULTPER (shared ride only)	0.873	12.6
AULTPER (transit only)	0.813	10.4
LOWINC (shared ride only)	0.329	3.4
LOWINC (transit-walk only)	3.056	34.6
LOWINC (transit-auto only)	2.183	13.1
HIGHINC (shared ride only)	-0.265	-3.7
HIGHINC (transit-walk only)	-5.665	-9.7
HIGHINC (transit-auto only)	-3.527	-4.8
CBD (transit only)	0.257	1.0
OBD (transit only)	0.276	2.8
Number of observations	10,707	
Rho-squared w.r.t. zero	0.6343	
Rho-squared w.r.t. constants	0.3769	

(See Section 3.5 for variable definitions.)

The final utility functions are therefore given by:

$$\text{Drive alone utility} = -0.055*(\text{AUTOIVTT}) - 0.558*(\text{AUTOCOST})$$

$$\text{Auto 2 person utility} = -0.055*(\text{AUTOIVTT}) - 0.558*(\text{AUTOCOST}) + 0.873*(\text{AULTPER}) + 0.329*(\text{LOWINC}) - 0.265*(\text{HIGHINC}) - 2.518$$

$$\text{Auto 3+ person utility} = -0.055*(\text{AUTOIVTT}) - 0.558*(\text{AUTOCOST}) + 0.873*(\text{AULTPER}) + 0.329*(\text{LOWINC}) - 0.265*(\text{HIGHINC}) - 3.724$$

$$\text{Transit-walk access utility} = -0.025*(\text{TRANIVTT}) - 0.064*(\text{TRANOVTT}) - 0.550*(\text{FARE}) + 0.813*(\text{AULTPER}) + 0.257*(\text{CBD}) + 0.276*(\text{OBD}) + 3.056*(\text{LOWINC}) - 5.665*(\text{HIGHINC}) + 0.414$$

$$\text{Transit-auto access utility} = -0.025*(\text{TRANIVTT}) - 0.064*(\text{TRANOVTT}) - 0.550*(\text{FARE}) + 0.813*(\text{AULTPER}) + 0.257*(\text{CBD}) + 0.276*(\text{OBD}) + 2.183*(\text{LOWINC}) - 3.527*(\text{HIGHINC}) - 1.926$$

3.5.2 Home Based Non-Work Trips

As with the home based work models, a number of model specifications were run on the estimation data set. The variables and their forms finally retained in the model were based on how well the estimated coefficients conformed to the expectations on their signs and relative magnitudes and on their statistical significance.

Table 5 presents the coefficient estimates for the final multinomial model specification for home based non-work trips, with t-statistics in parentheses and other statistics shown at the bottom. It was necessary to constrain the in-vehicle time coefficient to obtain reasonable results. The in-vehicle time coefficient for auto was constrained to -0.016, and the in-vehicle time coefficient for transit was constrained to -0.008. In addition, the auto and transit cost variables were constrained to have the same coefficient, as has been done in many other mode choice models. As Table 5 shows, all estimated coefficients are statistically significant, and the rho-squared values indicate a good statistical fit. The implied value of auto in-vehicle time from the multinomial model is \$4.94, and the implied value of transit in-vehicle time is \$2.47.

Other models that were tested include the following:

- Models including other variables, including AUTOOVTT, XFERS, POPDENP, EMPDENA, and other area type dummy variables. These variables had coefficients that either were insignificant or had incorrect signs.
- Separating transit out-of-vehicle time into its components. This resulted in either insignificant values for the IVTT coefficient or positive values for some level of service variable coefficients.
- Separating the auto cost variable into its components. This resulted in a positive coefficient for the fare variable.
- Models that had different or no constraints on level of service variables. These models invariably had at least one positive coefficient for a level of service variable.

Table 5. Home Based Non-Work Multinomial Logit Mode Choice Model Coefficients and Statistics

	Coefficient	T-Statistic
AUTO2	-0.992	-12.9
AUTO3	-4.464	-36.9
TRANWALK	1.836	13.8
TRANAUTO	-0.772	-4.7
AUTOIVTT (auto only)	-0.016	(constrained)
TRANIVTT (transit only)	-0.008	(constrained)
TRANOVTT (transit only)	-0.055	-13.6
AUTOCOST (auto only)	-0.194	-6.0
FARE (transit only)	-0.194	-6.0
AULTPER (shared ride only)	0.935	14.9
AULTPER (transit only)	0.202	1.8
MEDINC (shared ride only)	-0.099	-1.7
MEDINC (transit only)	-2.847	-34.3
HIGHINC (shared ride only)	-0.230	-3.7
HIGHINC (transit only)	-7.517	-12.9
CBD (shared ride only)	-0.447	-1.8
CBD (transit only)	1.345	4.5
HHSIZE (shared ride 2 only)	0.099	3.1
HHSIZE (shared ride 3+ only)	1.111	28.2
HHSIZE (transit only)	0.530	11.3
Number of observations	15,264	
Rho-squared w.r.t. zero	0.2725	
Rho-squared w.r.t. constants	0.2415	

(See Section 3.5 for variable definitions.)

Several nesting structures were also tested for the home based non-work model. The structure with the best fit had a nest including the two shared ride modes and another nest including the two transit modes. The drive alone mode was not included in either nest (in effect, it was in its own nest). The final nesting structure is shown in Figure 2.

Table 6 presents the coefficient estimates for the nested logit model specification for home based non-work trips, with t-statistics in parentheses and other statistics shown at the bottom. This table shows the coefficient values (for the auto modes) with the nesting coefficients already factored in. The variables are the same as in the multinomial model, and the coefficient values are similar to those in the multinomial model, which is not surprising given the relatively high nest coefficients. All coefficients are significant at the 90 percent level except the CBD coefficient for shared ride. The implied value of auto in-vehicle time from the nested model is \$4.07, and the implied value of transit in-vehicle time is \$1.94.

Figure 2. Home Based Non-Work Nesting Structure

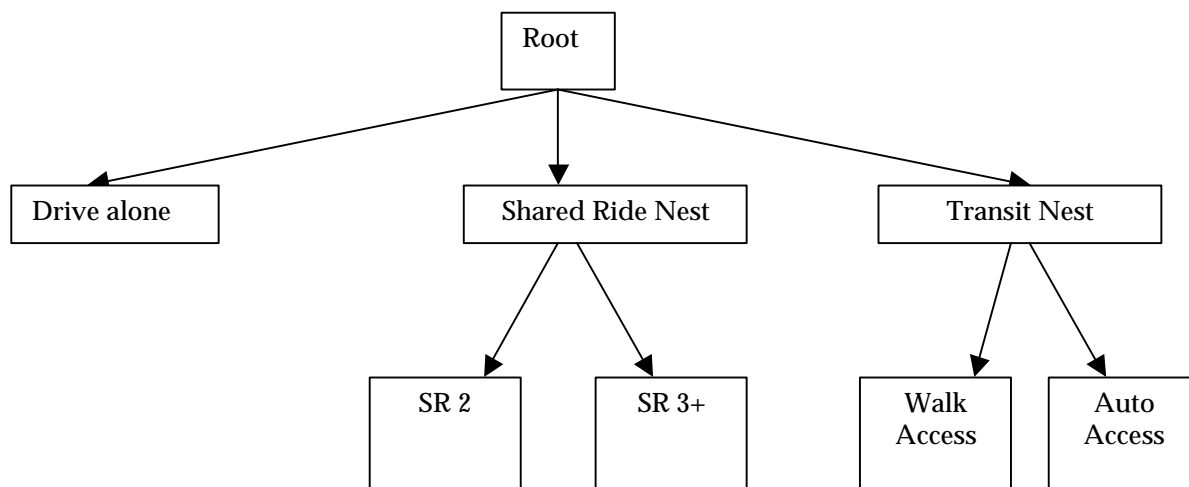


Table 6. Home Based Non-Work Nested Logit Mode Choice Model Coefficients and Statistics

	Coefficient	T-Statistic
AUTO2	-1.113	-3.8
AUTO3	-3.528	-10.4
TRANWALK	1.858	9.7
TRANAUTO	-0.469	-2.2
AUTOIVTT (auto only)	-0.011	(constrained)
TRANIVTT (transit only)	-0.007	(constrained)
TRANOVTT (transit only)	-0.053	-10.1
AUTOCOST (auto only)	-0.162	-5.0
FARE (transit only)	-0.217	-5.0
AULTPER (shared ride only)	0.935	4.6
AULTPER (transit only)	0.210	1.9
MEDINC (shared ride only)	-0.108	-1.7
MEDINC (transit only)	-2.852	-9.6
HIGHINC (shared ride only)	-0.241	-2.9
HIGHINC (transit only)	-7.533	-7.9
CBD (shared ride only)	-0.331	-1.4
CBD (transit only)	1.581	4.3
HHSIZE (shared ride 2 only)	0.212	2.2
HHSIZE (shared ride 3+ only)	0.916	8.3
HHSIZE (transit only)	0.527	7.7
Number of observations	15,264	
Rho-squared w.r.t. zero	0.2726	
Rho-squared w.r.t. constants	0.2416	

(See Section 3.5 for variable definitions.)

The transit nest coefficient is 0.892 with a t-statistic of 10.0; the associated standard error is 0.089. The test of whether the nest coefficient is significantly different from one yields a t-statistic of 1.2, which is not significant at the 90% level. While this indicates that the transit nesting has relatively little effect on the model results, it is recommended that the transit nest be retained since the coefficient is valid and it makes sense to nest the two transit alternatives together.

The auto nest coefficient is 0.666 with a t-statistic of 4.8; the associated standard error is 0.138. The test of whether the nest coefficient is significantly different from one yields a t-statistic of 2.4, which is significant at the 95% level. The auto nest is therefore retained.

The final utility functions are therefore given by:

$$\text{Drive alone utility} = -0.011*(\text{AUTOIVTT}) - 0.162*(\text{AUTOCOST})$$

$$\begin{aligned} \text{Auto 2 person utility} = & -0.011*(\text{AUTOIVTT}) - 0.162*(\text{AUTOCOST}) + 0.935*(\text{AULTPER}) \\ & - 0.108*(\text{MEDINC}) - 0.241*(\text{HIGHINC}) + 0.212*(\text{HHSIZE}) - 0.331*(\text{CBD}) - 1.113 \end{aligned}$$

$$\begin{aligned} \text{Auto 3+ person utility} = & -0.011*(\text{AUTOIVTT}) - 0.162*(\text{AUTOCOST}) + 0.935*(\text{AULTPER}) \\ & - 0.108*(\text{MEDINC}) - 0.241*(\text{HIGHINC}) + 0.916*(\text{HHSIZE}) - 0.331*(\text{CBD}) - 3.528 \end{aligned}$$

$$\begin{aligned} \text{Transit-walk access utility} = & -0.007*(\text{TRANIVTT}) - 0.053*(\text{TRANOVTT}) - 0.217*(\text{FARE}) \\ & + 0.210*(\text{AULTPER}) + 1.581*(\text{CBD}) + 0.527*(\text{HHSIZE}) - 2.852*(\text{MEDINC}) \\ & - 7.533*(\text{HIGHINC}) + 1.858 \end{aligned}$$

$$\begin{aligned} \text{Transit-auto access utility} = & -0.007*(\text{TRANIVTT}) - 0.053*(\text{TRANOVTT}) - 0.217*(\text{FARE}) \\ & + 0.210*(\text{AULTPER}) + 1.581*(\text{CBD}) + 0.527*(\text{HHSIZE}) - 2.852*(\text{MEDINC}) \\ & - 7.533*(\text{HIGHINC}) - 0.469 \end{aligned}$$

3.5.3 Non-Home Based Trips

As with the home based models, a number of model specifications were run on the estimation data set. The variables and their forms finally retained in the model were based on how well the estimated coefficients conformed to the expectations on their signs and relative magnitudes and on their statistical significance.

Table 7 presents the coefficient estimates for the final multinomial model specification for non-home based trips, with t-statistics in parentheses and other statistics shown at the bottom. It was necessary to constrain both the in-vehicle time and cost coefficients to obtain reasonable results. The in-vehicle time coefficient for auto was constrained to -0.011, and the in-vehicle time coefficient for transit was constrained to -0.007. These are the same values as in the nested home based non-work model. The auto and transit cost variables were constrained to -0.200. As Table 7 shows, all estimated coefficients are statistically significant, and the rho-squared values indicate a good statistical fit. The implied value of auto in-vehicle time from this model is \$3.30, and the implied value of transit in-vehicle time is \$2.10.

Other models that were tested include the following:

- Models including other variables, including AUTOOVTT, XFERS, POPDENP, EMPDENA, and other area type dummy variables. These variables had coefficients that either were insignificant or had incorrect signs.
- Separating transit out-of-vehicle time into its components. This resulted in either insignificant values for the IVTT coefficient or positive values for some level of service variable coefficients.
- Separating the auto cost variable into its components. This resulted in a positive coefficient for the fare variable.
- Models that had different or no constraints on level of service variables. These models invariably had at least one positive coefficient for a level of service variable.

Table 7. Non-Home Based Multinomial Logit Mode Choice Model Coefficients and Statistics

	Coefficient	T-Statistic
AUTO2	-0.902	-27.1
AUTO3	-1.178	-32.5
TRANWALK	0.770	6.9
TRANAUTO	-1.119	-6.9
AUTOIVTT (auto only)	-0.011	(constrained)
TRANIVTT (transit only)	-0.007	(constrained)
TRANOVTT (transit only)	-0.036	-8.4
AUTOCOST (auto only)	-0.200	(constrained)
FARE (transit only)	-0.200	(constrained)
CBD (shared ride only)	-0.754	-4.6
CBD (transit only)	1.323	10.8
Number of observations	7,250	
Rho-squared w.r.t. zero	0.1598	
Rho-squared w.r.t. constants	0.0721	

(See Section 3.5 for variable definitions.)

No nesting structures were found to give reasonable results for non-home based trips. The final utility functions are therefore given by:

$$\text{Drive alone utility} = -0.011*(\text{AUTOIVTT}) -0.200*(\text{AUTOCOST})$$

$$\text{Auto 2 person utility} = -0.011*(\text{AUTOIVTT}) -0.200*(\text{AUTOCOST}) -0.754*(\text{CBD}) -0.902$$

$$\text{Auto 3+ person utility} = -0.011*(\text{AUTOIVTT}) -0.200*(\text{AUTOCOST}) -0.754*(\text{CBD}) -1.178$$

$$\text{Transit-walk access utility} = -0.007*(\text{TRANIVTT}) -0.036*(\text{TRANOVTT}) -0.200*(\text{FARE}) \\ +1.323*(\text{CBD}) +0.770$$

$$\text{Transit-auto access utility} = -0.007*(\text{TRANIVTT}) -0.036*(\text{TRANOVTT}) -0.200*(\text{FARE}) \\ +1.323*(\text{CBD}) -1.119$$

4.0 Mode Choice Model Validation

The results of the NCTCOG mode choice model estimation were compared to models estimated in other urban areas in the United States, as compiled in a recent publication¹. Tables 8 and 9 present the results of this comparison. It should be noted that most of the models in other areas did not consider level of service variables with separate coefficients for auto and transit modes, and the auto and transit values tend to be closer to each other as a result. Nevertheless, the NCTCOG models compare favorably with the other models, as shown in the tables.

At the time of this writing, additional model calibration and validation is being performed by NCTCOG staff. The changes to Mode Choice include adjustments to ALOGIT's alternative-specific constants and introduction of rail and express bus constants. The changes to the transit skimming and assignment setups have included:

- Incorporation of bus-to-rail and rail-to-bus timed transfer points (pulse points).
- Modification of the maximum auto drive time to transit from 45 to 30 minutes.
- Modification of the maximum wait time for transit from 15 to 20 minutes.

¹ Rossi, Thomas F. and Maren L. Outwater. "How Valid Is It to Transfer Mode Choice Model Parameters?" Proceedings of the Seventh National Conference on Transportation Planning Applications, Transportation Research Board, Boston, Massachusetts, 1999.

Table 8. Comparison of NCTCOG Model Parameters to Other Urban Areas

Home Based Work	NCTCOG	Average	Range
Auto IVTT (min)	-0.055	-0.029	-0.01 to -0.05
Auto Cost (\$)	-0.558	-0.545	-0.3 to -1.3
Transit IVTT (min)	-0.025	-0.027	-0.01 to -0.05
Transit OVTT (min)	-0.064	-0.052	-0.03 to -0.07
Transit Cost (\$)	-0.550	-0.587	-0.3 to -1.3
Ratio: OVTT/IVTT	2.6	1.9	1.5 to 3
Value of Time (Auto)	\$5.91	\$3.24	\$2 to \$5
Value of Time (Transit)	\$2.73	\$2.72	\$2 to \$5

Home Based Non-Work	NCTCOG	Average	Range
Auto IVTT (min)	-0.011	-0.019	-0.01 to -0.03
Auto Cost (\$)	-0.162	-0.827	-0.2 to -1.3
Transit IVTT (min)	-0.007	-0.017	-0.004 to -0.05
Transit OVTT (min)	-0.053	-0.056	-0.007 to -0.07
Transit Cost (\$)	-0.217	-0.858	-0.2 to -1.3
Ratio: OVTT/IVTT	7.4	3.2	2 to 6
Value of Time (Auto)	\$4.07	\$1.38	\$0.5 to \$5
Value of Time (Transit)	\$1.94	\$1.20	\$0.5 to \$5

Non-Home Based	NCTCOG	Average	Range
Auto IVTT (min)	-0.011	-0.019	-0.01 to -0.03
Auto Cost (\$)	-0.200	-1.026	-0.2 to -1.3
Transit IVTT (min)	-0.007	-0.019	-0.01 to -0.05
Transit OVTT (min)	-0.036	-0.073	-0.02 to -0.20
Transit Cost (\$)	-0.200	-0.993	-0.2 to -1.3
Ratio: OVTT/IVTT	5.2	3.8	2 to 7
Value of Time (Auto)	\$3.30	\$1.11	\$0.2 to \$5
Value of Time (Transit)	\$2.10	\$1.16	\$0.2 to \$5

Table 9. Comparison of NCTCOG Trip Purpose Models to Other Urban Areas

Ratio Of HBNW To HBW	NCTCOG	Average	Range
Auto IVTT (min)	0.20	0.65	0.3 to 2
Auto Cost (cents)	0.29	1.52	0.3 to 2
Transit IVTT (min)	0.28	0.65	0.3 to 2
Transit OVTT (min)	0.83	1.08	0.1 to 1.4
Transit Cost (cents)	0.39	1.06	0.1 to 1.4
Ratio: OVTT/IVTT	2.8	1.7	1 to 3
Value of Time (Auto)	0.23	0.43	0.2 to 1
Value of Time (Transit)	0.45	0.44	0.2 to 1

Ratio Of NHB To HBW	NCTCOG	Average	Range
Auto IVTT (min)	0.20	0.64	0.3 to 2
Auto Cost (cents)	0.36	1.88	0.3 to 2
Transit IVTT (min)	0.28	0.72	0.3 to 2
Transit OVTT (min)	0.56	1.39	0.1 to 1.4
Transit Cost (cents)	0.36	1.69	0.1 to 1.4
Ratio: OVTT/IVTT	2.0	1.7	1 to 3
Value of Time (Auto)	0.81	0.43	0.2 to 1
Value of Time (Transit)	0.77	0.44	0.2 to 1