Transit Modeling Update

Trip Distribution Review and Recommended Model Development Guidance

Contents

1 Introduction

The purpose of the Transit Modeling Update project is to specify, within FSUTMS and associated support systems, the changes necessary to improve the preparation of transit demand forecasts to a point consistent with federal expectations, and to incorporate state of the practice techniques and tools through a prototype model application. The Tallahassee Capital Region Transportation Planning Agency (CRTPA) model was chosen as the prototype FSUTMS model application.

For any individual origin-destination pair, the trip distribution model establishes the total number of person trips for the mode choice model. Since targets for transit demand in the mode choice model are usually established independently using surveys and counts of observed transit demand, it is vitally important that the distribution model provide a realistic number of total person trips from which the mode demand is calibrated. For example, an underestimation of person-trip demand from the distribution model will cause unrealistically high transit share targets. During calibration, the mode choice coefficients will be distorted as the model tries to match extreme mode share targets. In a broader sense, the trip distribution model's ability to estimate the right demand by volume and market and location is fundamental to a proper estimate of travel by mode, since the modal depends upon trip flows being placed properly with regard to the patterns of facilities and accessibilities within a region.

2 FSUTMS Trip Distribution Review

FSUTMS uses gravity models to distribute all trips. These trip distribution models are stratified by trip purpose, but not by household attributes (income or auto ownership, for example) or time period. Free-flow travel times are used to distribute all trips. As such, the FSUTMS trip distribution models rely on simplifying assumptions that most likely are no longer applicable to most urban areas in Florida. More importantly, the gravity model framework itself can respond inconsistently to changes in highway or transit levels of service.

The use of free-flow travel time impedances is an unnecessary, and in many instances, erroneous simplification. Free-flow travel times do not reflect the deterioration of accessibility caused by highway congestion. In application the model may over-estimate the attractiveness of distant destinations because it under-estimates the travel time to reach these destinations. Current best practice is to stratify the distribution models into two time periods, peak and off-peak, so that more appropriate travel times can be used in distribution.

The lack of trip market segmentation besides trip purpose can also lead to incorrect forecasts of trip distribution patterns. Persons of restricted mobility (low income, zero car households, for example) exhibit, on average, shorter trip lengths than other persons. Low income workers are more likely to be employed in retail and service occupations, while high income workers are more often professionals. Without appropriate trip market segmentation the distribution model is unable to link the right workers to the right jobs. Failing to reproduce person trip market patterns can be particularly severe for the mode choice model, often times making it simply impossible to calibrate logical and representative mode-specific constants.

A significant restriction of the FSUTMS trip distribution model is its reliance on the gravity model. The gravity analogy works for trip distribution in mono-centric urban regions where accessibility to transit plays little to no role in choice of destination. This is no longer the case in some Florida urban areas, where there may be more than one dominant attraction region, multiple and important suburban-to suburban trip flows, and where there is interest in understanding the contribution of transit towards achieving more sustainable urban development patterns. The gravity model often times exhibits incorrect demand elasticities; in particular, the model may respond illogically to changes in levels of service—improved accessibility to a given destination may cause a disproportionate increase in total trips, and/or an increase of trips using the mode(s) whose accessibility did not change. Both results are undesirable and may lead to erroneous assessments of the impact of transit (or highway) improvements.

This limitation of the gravity model is overcome by destination choice models. Like mode choice models, destination choice models are founded on the application of discrete choice theory to travel behavior. With the correct utility specification, consistency between changes in levels of service and changes in demand is guaranteed. Moreover, because the functional form of the destination choice utility is very flexible, accounting for singularities in the trip distribution pattern can be accomplished in intuitive ways. For example, rather than using a K-factor to inform the model of a natural barrier, such as a river, a term can be added to the utility equation and interpreted in terms of 'equivalent' minutes of travel time.

In fact, given appropriate assumptions, the gravity model can be shown to be a destination choice model. As such, the destination choice framework includes the gravity model itself plus a family of more flexible functional forms. The desired consistency with the mode choice model is obtained by the introduction of mode choice logsums in an appropriate way, as shown below.

In summary, the recommended trip distribution approach for FSUTMS is the implementation of destination choice models, stratified by two time periods, trip purpose, and household attributes. As discussed in the trip generation guidance, candidate household attribute stratification for FSUTMS is a combination of household income and car sufficiency:

i. Zero Car Households – All Income
ii. Cars less than Workers – Low Income
iii. Cars less than Workers - Medium Income
iv. Cars less than Workers – High Income
v. Cars equal to or greater than Workers - Low Income
vi. Cars equal to or greater than Workers - Medium Income
vii. Cars equal to or greater than Workers – High Income

The remainder of this memorandum discusses the structure of destination choice models and provides guidance for model estimation and calibration.

3 Proposed Trip Distribution Approach

3.1 Model Form and Utility Structure

The most common implementation of destination choice is the multinomial logit form. Travelers are hypothesized to choose the destination that maximizes their utility. The utility of a destination is a function of multi-modal accessibilities and preferences, the attractiveness of the destination zone, person and household attributes, and other unknown, un-included attributes of the trip maker or the destination zone. The probability that trip *m* produced in zone *i* chooses destination zone *j* is given by the utility of zone *j* and the utility of all other possible destinations.

$$P_{ijm} = \frac{e^{U_{ijm}}}{\sum_j e^{U_{ijm}}}$$

$$U_{ijm} = \theta \times L_{ijm} + \sum_k \beta^k D_{ij}^k + \sum_k \delta_m^k N_m^k D_{ij}^k + \sum_k \gamma_m^k M_i^k IZ_j + Ln\left(A_{jm}\right)$$

In this utility function L_{ijm} is the mode choice logsum of trip market *m*, $\beta^k D_{ij}^k$ are the terms of a distance polynomial; N_m^k represent attributes of the trip market such as income or auto availability, usually in the form of indicator variables; M_i^k represent attributes of the trip production zone, such as residential density that may be included to represent the utility of the intra-zonal destination; A_{jm} is the size variable; and θ, β^k, δ_m^k and γ_m^k are parameters to be estimated. This is a generic utility function; some of these terms may not apply to a particular region while others may be included. For example, in regions where transit service is quite limited, a common, acceptable simplification is to use highway time or generalized cost instead of the mode choice logsum. The specific composition of the distance polynomial, the distance interaction terms, and the size variable results from the model estimation and calibration processes. At a minimum the utility function is expected to have an impedance term and a size variable.

In the formulation above the mode choice logsum is the travel impedance; the model is informed of changes in the cost or travel time of any mode via the logsum. Because the mode choice model includes modal preferences, in the form of alternative-specific constants and income-stratified cost coefficients, these preferences affect how the distribution model responds. For example, a change in local bus accessibility will influence the distribution pattern of low income, zero car households more so than higher income households because persons from these households are more likely to use local bus than all other persons. In effect, by virtue of the inclusion of the mode choice logsum in the destination choice utility, the two models can be interpreted as a nested logit model, where the choice of mode is conditional on the choice of destination.

The size variable is equivalent to the attractions in a gravity model. In fact, it is often calculated in the same way, as a linear regression of trip attractions on various employment categories:

$$A_{jm} = \alpha_m^1 \times E_j^1 + \alpha_m^2 \times E_j^2 + \alpha_m^3 \times E_j^3 + \cdots + \alpha_m^n \times E_j^n$$

The size variable may include terms other than employment, such as population, households, school enrollment, or any other indicator of the attractiveness of a zone. The coefficients on these variables can be pre-calculated, and/or they can be estimated simultaneously with the other coefficients in the utility function. An equivalent formulation of the size term is obtained by scaling all coefficients. The scale factor can be conveniently chosen to be the coefficient on one of the employment terms, such that one of the employment categories has a coefficient of 1, and all other coefficients are scaled accordingly:

$$\varphi_m^k = \frac{\alpha_m^k}{\alpha_m^1}$$

$$A_{jm} = \alpha_m^1 \times \left(E_j^1 + \varphi_m^2 \times E_j^2 + \varphi_m^3 \times E_j^3 + \cdots + \varphi_m^n \times E_j^n\right)$$

The size term enters the utility function in logarithmic form, so the log of the scale coefficient α_m^1 becomes a constant added to all destinations. Since the logit probabilities do not change when the same constant is added to all the choices, the scale of the size term is, in effect, arbitrary.

The role of the distance polynomial in the utility function is to assist in reproducing the observed trip length frequencies. It is often difficult to reproduce the observed trip length frequencies using mode choice logsums and size terms alone in the utility function. The specific composition of the distance polynomial is not important, but the entire distance decay function has to be well-behaved, which typically means monotonically decreasing. It is preferable to use distance instead of time to fit the trip length frequencies, because the distance between destination choices does not change across transportation alternatives, while travel times do change.

Household or production zone attributes that identify the trip market segment are entered in the utility function interacted with any of the distance terms. These attributes help to identify the propensity to make shorter or longer trips, all else equal. It is often found that zero car households exhibit shorter trip lengths, on average, than all other households; that trip length tends to increase with household income, and that trip length decreases with residential density.

Another variable that is often required is an intrazonal indicator variable, to help estimate the correct proportion of intrazonal trips. The intrazonal indicator can be interacted with density and/or stratified by zone size. In this way the prediction of intrazonal trips does not rely entirely on the intra-zonal time or distance.

Yet another class of variables that can be included in the utility equation is analogous to k-factors. These variables include bridge penalties, intra-region indicators, CBD indicators, or indicators for other special attractions. Like the trip market variables, these terms help to reproduce the base year observed trip patterns. Best practice is to use them parsimoniously, to avoid overwhelming the level of service variables. A good check on the reasonableness of these variables is to express them in terms of equivalent minutes of travel time.

3.2 Trip Market Stratification

As indicated above, the destination choice models must be able to support the proposed mode choice stratification, which is comprised of household income and car sufficiency segments. Full model segmentation, however, is not required in the destination choice model; some terms in the utility function, such as the mode choice logsum, are market segment specific, while others need not be. The models are usually estimated by pooling together observations from different market segments, and terms are added to identify market-specific effects, usually interacted with a distance term. For any such effects desired in the final model, a separate vector of trip productions must be generated. The application program can be designed such that the common terms in the utility function are calculated only once, to reduce model run times, and to write out only the required trip tables (rather than intermediate steps) to economize on model output storage.

A key market stratification issue is whether to stratify the size term. Where possible, model estimation should be used to inform this question. Good, but not indispensible, practice is to stratify the HBW size term by income group, particularly if the trip generation model is stratified by household income (as has been proposed for FSUTMS). Income stratification helps to link low income workers with low income jobs, etc. The linkage is somewhat confounded by the use of household income, instead of worker earnings, as the segmentation variable, but reasonable relationships often emerge—the coefficient on retail employment is proportionally more important for low income than high income workers, for example.

3.3 Time Period Stratification and Best Path Skims

The FSUTMS time period stratification proposed in Technical Memorandum #2 is tightly related to the need to produce representative travel time and cost skims for the trip distribution and mode choice models. In brief, it is proposed that the trip distribution and mode choice models be fully stratified into peak and off-peak periods, as shown to the right, and that travel time feedback be used for both AM and MD skims to ensure consistency between the assignment results and the demand model.

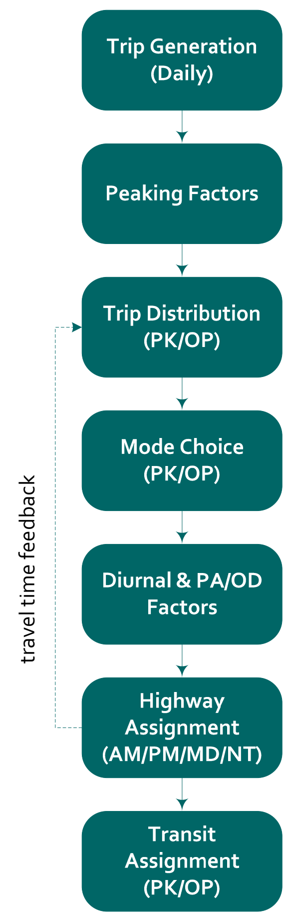

Figure 3-1: Proposed Time of Day Stratification

Consistency extends to the specification of the generalized cost function used to find best paths for creating travel time and cost skims. Currently FSUTMS builds best highway travel time paths. This practice ignores the cost of tolls—in effect it assumes that trips can take advantage of the increased speeds afforded by a toll road at no cost, and in doing so makes the toll road paths more attractive than they are. The generalized cost function often includes a distance term, which may be interpreted as related to the cost of gas and vehicle maintenance. This distance term makes longer paths less likely, even if they result in shorter travel times. The cost terms in the generalized cost function are expressed in time units using a regional value of time.

$$GC_{ij} = Travel\ Time_{ij} + \left(AOC \times Distance_{ij} + Toll_{ij}\right)/VOT$$

3.4 Attraction Constraining

A common practice when using gravity models is to doubly-constrain the person trip table: row sums and column sums are made to match productions and attractions, respectively. This is achieved by iteratively fitting the person trip table to productions and attractions. The equivalent procedure in destination choice models is to apply a shadow price:

$$U'_{ijm} = U_{ijm} + sp_j$$

The shadow price is calculated iteratively by comparing the person trip table attractions to pre-defined target attractions.

The key question to address is whether this level of constraining is truly required or even desirable. Doubly-constraining the person trip table can seriously distort the travel patterns, because it assumes that attraction rates are independent of accessibility. In reality, the number of trips attracted per job can vary widely depending on business location, type of business, and other factors that may not be accounted for in the size term. Accepted practice is to doubly-constrain the HBW models, so that attractions are proportional to zonal employment. Because less is known about the number of jobs in a zone that are filled by workers of a certain socio-economic background, than about the total number of jobs in the zone, the HBW shadow price is computed by pooling all trip market segments together. Shadow pricing is not recommended for purposes other than HBW.

In summary, where the level of confidence associated with the pre-calculated target attractions is relatively high, as is for total employment, shadow pricing is recommended. Otherwise it is often best not to doubly-constrain.

4 Model Estimation

To estimate the parameters of the destination choice model, an estimation file combining household survey data, skim data, and employment data must be constructed. The destination choice estimation file has a structure similar to that of a mode choice estimation file. The file is constructed by selecting the household survey trip records and keeping relevant information, such as the production zone, the chosen attraction zone, and household income and other trip market data items. Then, information about the choice set — which may consist of all or a subset of the TAZs in the region—is appended to the household survey trip record. This information includes highway travel times and distances, mode choice logsums corresponding to the trip market segment, and the employment estimates for each possible destination TAZ.

Trip record data	Trip Id Production TAZ (PTAZ) Attraction TAZ (choice) Trip purpose Household income
Destination choice set data	TAZ 1 Travel time PTAZ to TAZ 1 Travel distance PTAZ to TAZ 1 Mode choice logsum PTAZ to TAZ 1 Employment TAZ 1 … TAZ N Travel time PTAZ to TAZ N Travel distance PTAZ to TAZ N Mode choice logsum PTAZ to TAZ N Employment TAZ 3

It is impractical to include all possible TAZs in the region in the estimation file choice set. Instead, a sample of destinations is chosen for each TAZ. Several destination sampling schemes have been proposed. For a discussion of alternative sampling schemes see Ben-Akiva and Lerman (1985)[1]. The recommended approach for the development of Florida destination choice models is importance sampling.

4.1 Importance Sampling

The concept behind importance sampling is that the likelihood of choosing a zone for the destination set should be related to the likelihood that the zone is chosen as the trip destination. Operationally, an 'importance' function is defined and used to compute the selection probabilities. Using Monte Carlo simulation, a number of destinations are sampled with replacement from the importance probability distribution. Finally, appropriate sampling correction factors are used in estimation to maintain the statistical properties of the maximum likelihood estimator.

The importance function gives the probability of selecting a zone *j* for the choice set, and can be defined as:

$$W_j = A_j \times exp\left(-2D_{ij}/D\right)$$

$$P_j = \frac{W_j}{\sum_j W_j}$$

Where *D* is the regional observed average distance, A_j is a size variable, and D_{ij} is the distance to each zone. In essence, the importance function is itself a simple destination choice model based on pre-defined distance and size term parameters.

Given this importance function, 20 (or 30 or 40) zones are sampled with replacement based on the calculated selection probabilities. Sampling with replacement means that the same zone can be chosen more than once, so the number of unique zones in the destination set may be less than 20. The chosen zone is added to the destination set too.

In estimation, a term must be added to the utility function to correct for the sampling error and ensure the consistency of the estimator. This correction term represents the difference between the sampling probability and final estimated probability for each alternative.

$$U'_{ijm} = U_{ijm} + C_j$$

$$C_j = -Ln\left(P_j/F_j\right)$$

where F_j is the frequency of zone j in the destination set. When the selection frequencies F_j are equal to the probability of selection P_j the correction is zero. If a zone is selected with higher frequency than its probability suggests, the correction is negative, i.e., it is given a lower base utility. This correction factor is not part of the 'true' model, that is, it is not part of the utility function in application.

[1] Ben-Akiva, M. and S. Lerman (1985). Discrete Choice Analysis. Theory and Application to Travel Demand. MIT Press: Cambridge, MA.

Importance sampling can be combined with an exploded sample to augment the number of zones in the choice set without increasing the length of the estimation file record. Exploding the sample means that each trip record is repeated a number of times, and a different choice set is drawn for each exploded trip record. Experience with destination choice model estimation indicates that a sufficiently large destination choice set can be constructed by exploding the sample 10 times and choosing 20 to 40 zones per exploded record. Larger choice sets do not yield substantially better parameters, where better is defined as parameters exhibiting smaller standard errors. While the statistical properties of sample explosion have not been proven, tests with different choice set sizes indicate that sample explosion gives results that are nearly statistically equivalent to using a choice set size equal to the sum of destinations across the exploded samples. For example, if each trip record is exploded 10 times and a choice set of 30 zones is chosen for each exploded record, the estimated parameters and standard errors are nearly the same as if a destination choice set of 300 zones were used for each original trip record.

Once a process is in place to build the estimation file, it is a relatively simple task to rebuild the file using different choice set sizes and even different importance functions. It is recommended that the final model be estimated with different choice sets to verify its robustness.

4.2 Size Term Pre-Calculation

As indicated above, to build the estimation file an estimate of the size term is needed. One option is to form the size term using obvious and relevant components: total employment for HBW trips, retail employment for shop trips, student enrollment for school and university trips, and a combination of employment and households for the other purposes. This is usually sufficient for the importance function. A more detailed specification of the size term can be explored as part of the model estimation.

Alternatively the size terms can be pre-calculated on the basis of household survey data or Public Use Micro-Sample (PUMS) data. These data are helpful to specify income-stratified size terms. The household survey data can be used to develop linear regression models of trip attractions that depend on various employment categories. Separate models can be constructed for low, medium and high income attractions. As is the case with attraction models, these models must be specified as regressions through the origin. Even if no income-stratification is desired, these models help to understand correlation among various employment categories and to explore various ways of developing relevant employment classes.

The PUMS data can be used to develop HBW size terms. A simple tabulation of workers by industry and household income shows the relative importance of each industrial category in drawing workers from each household income group. Table 4-1shows this tabulation for all of Florida, while Table 4-2 shows the same tabulation for the three PUMAs that contain the CRTPA region (Gadsden, Jefferson, Leon and Wakulla counties). The row percents normalized to retail employment can be used as the size term coefficients.

These pre-calculated size term coefficients may be re-estimated along with the rest of the model, or may be kept constant while the other utility term parameters are estimated. Where possible it is recommended that re-estimation be attempted.

Table 4-1: Size Term Coefficients, Florida Workers

	Share of Workers					Size Coefficients			
	Household Income					Household Income			
Employment Industry	Low	Med	High	Very High	All	Low	Med	High	Very High
Retail Trade	13%	28%	24%	35%	100%	1.00	1.00	1.00	1.00
Professional	9%	23%	22%	46%	100%	0.72	0.81	0.91	1.32
Information	7%	20%	25%	47%	100%	0.54	0.73	1.03	1.37
Financial, Insurance, Real Estate	7%	21%	23%	50%	100%	0.51	0.75	0.93	1.44
Public Administration	4%	21%	25%	50%	100%	0.33	0.74	1.02	1.44
Education	10%	23%	23%	44%	100%	0.73	0.82	0.96	1.28
Wholesale Trade	7%	23%	24%	46%	100%	0.53	0.82	0.99	1.32
Manufacturing	8%	25%	25%	42%	100%	0.62	0.90	1.02	1.20
Agriculture	19%	30%	24%	27%	100%	1.46	1.05	1.00	0.79
Transportation	7%	24%	26%	43%	100%	0.56	0.84	1.09	1.24
Construction	10%	29%	25%	36%	100%	0.77	1.03	1.05	1.03
Recreation	18%	30%	24%	29%	100%	1.36	1.07	0.97	0.83
Other Services	14%	29%	25%	32%	100%	1.10	1.04	1.02	0.92

Source: American Community Survey, 2005-2009 Release. Sample size: 453,000 workers.

Table 4-2: Size Term Coefficient, CRTPA Workers

	Share of Workers					Size Coefficients			
	Household Income					Household Income			
Employment Industry	Low	Med	High	Very High	All	Low	Med	High	Very High
Retail Trade	25%	29%	20%	26%	100%	1.00	1.00	1.00	1.00
Professional	15%	23%	19%	43%	100%	0.59	0.80	0.92	1.68
Information	11%	21%	26%	42%	100%	0.45	0.73	1.27	1.63
Financial, Insurance, Real Estate	12%	22%	21%	45%	100%	0.49	0.77	1.02	1.75
Public Administration	8%	23%	25%	44%	100%	0.30	0.81	1.25	1.69
Education	19%	26%	21%	34%	100%	0.75	0.91	1.06	1.31
Wholesale Trade	8%	25%	29%	38%	100%	0.33	0.88	1.42	1.46
Manufacturing	12%	27%	20%	41%	100%	0.49	0.94	0.99	1.58
Agriculture	18%	33%	26%	23%	100%	0.70	1.16	1.30	0.88
Transportation	8%	26%	29%	37%	100%	0.31	0.92	1.43	1.43
Construction	13%	30%	25%	32%	100%	0.51	1.04	1.26	1.23
Recreation	39%	28%	15%	18%	100%	1.52	0.97	0.76	0.71
Other Services	17%	23%	23%	37%	100%	0.66	0.81	1.12	1.45

Source: American Community Survey, 2005-2009 Release. Sample size: 10,500 workers.

4.3 Agglomeration Effects

As indicated above, it is desirable to apply the HBW model in a constrained fashion, that is, shadow pricing so as to allocate attractions in proportion to total employment. In estimation, however, the model is unconstrained. Unconstrained estimation ignores agglomeration effects—locations in clusters of employment are more attractive partly because they offer better accessibility to other jobs. This effect can be introduced in estimation by using zonal accessibility indices. These indices take the form of destination choice logsums and represent the summation of attractions across all destinations, weighted by travel impedances.

$$Acc_j = \ln\left(\sum_k Emp_k \times e^{-\alpha \times Imp_{jk}}\right)$$

Note that in application these indices get absorbed by the shadow price. In estimation however they help the model to identify whether locations in clusters of certain types of employments are more or less attractive, and adjust the other model parameters accordingly. For example, in the estimation of a workplace location choice model for San Diego, two accessibility indices were found to explain workplace location: a total employment index, with positive coefficient, and a retail employment index, with negative coefficient. Without these indices the model would under-estimate the attractiveness of total employment clusters, which tend to be located far from residential areas, and would over-estimate the attractiveness of retail employment clusters, which tend to be located near residential areas[2].

4.4 Interpretation of Estimation Results

As is the case with any model estimation effort, sound a priori expectations about the sign and magnitude of the parameters to be estimated is critical for interpreting the estimation results. Two parameters in particular are discussed here, the mode choice logsum coefficient and the distance decay function.

Mode Choice Logsum Coefficient. A destination choice model that includes mode choice logsums can be thought of as a nested destination/mode choice model. As such, the mode choice logsum coefficient is in effect a nesting coefficient, and therefore should exhibit a value between 0 and 1. Often times estimated mode choice logsum parameters are greater than 1. This can be a sign that the nesting is incorrect, that is, that the choice of destination is conditional on the choice of model, and not vice-versa as is commonly assumed. One possible solution is to apply the two models simultaneously. In practice, however, the destination choice model is typically applied before the mode choice model. In such cases, mode choice logsum coefficients that assume a value greater than 1 are often constrained to a value between 0 and 1. Appropriate values for HBW are in the range of 0.5 to 0.8, and for other purposes in the range of 0.7 to 1.0. The lower value of the HBW mode choice logsum coefficient indicates that these trips are less elastic with respect to travel time and cost—there are fewer substitutes for the choice of

---

[2]Vovsha, P. et al. Workplace Choice Models: Insights into Spatial Patterns of Commuting in Three Metropolitan Regions. Presented at the 13th TRB National Transportation Planning Applications Conference, Reno, NV: May 8-12, 2011.

HBW location than HB Shop location for example, and less willingness to switch jobs due to changes in travel time or cost.

Distance Decay. The specific composition of the distance decay function is not important. The material point is that the function should be monotonically decreasing, indicating that all else equal, destinations that are far away are less likely than near destinations. A sample set of distance decay functions estimated for various worker segments (full and part time, gender and income segments) in San Diego is shown in Figure 4-1. The steeper functions indicate a propensity for shorter trip lengths, all else equal. In some cases the combined distance terms are monotonic only over a subset of the entire range of distances in a region. Possible solutions are to cap the distance decay term, or to estimate a step-wise linear term (assuming sufficient observations in the long distance range). The model must be applied consistently to how it was estimated, although some adjustment of the distance decay terms is often required to improve the fit to the observed trip length frequencies.

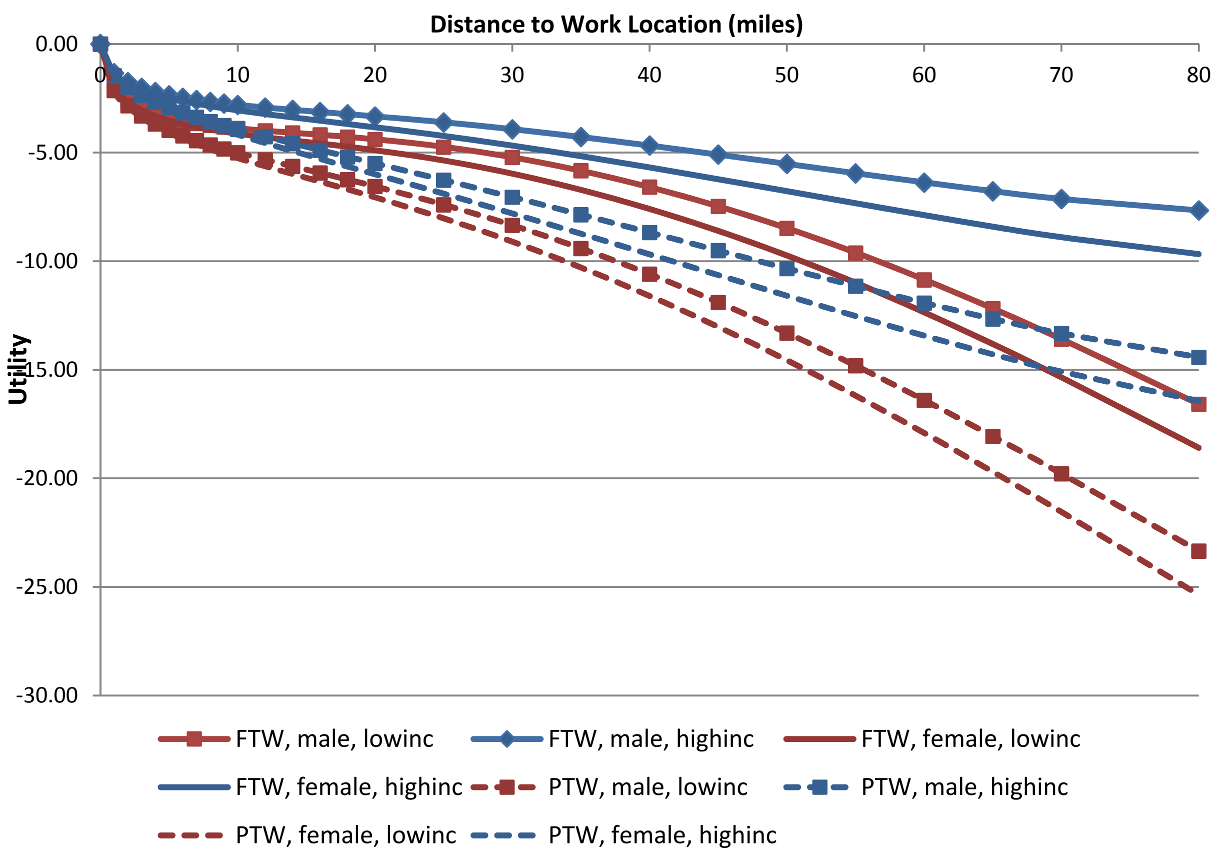

Figure 4-1: SANDAG Workplace Location Distance Decay

5 Model Calibration

The calibration of a destination choice model is no different than the calibration of a gravity model, in that the goal is to reproduce the observed trip length frequencies and district-level flow patterns. Possible calibration actions include adjusting the coefficients of the distance polynomial, the coefficient for intrazonal trips, and possibly the size term coefficients. In addition, the flexibility of the model structure allows accounting for systematic differences that may emerge when the model is applied, and that may lead to additional utility function terms.

Typical destination choice calibration targets include:

- Average trip lengths
- Trip length frequency distribution
- Percent of intra-zonal trips
- Coincidence ratio
- ACS worker flow pattern (county level or smaller geography)
- Trip flow pattern (20-district geography)
- Implied transit mode shares

Any of these targets may be stratified by trip purpose, time period and trip market segment, as allowed by the sample size.