1. Model Framework

In developing a framework that captures the influence of latent individual modal preferences, or modality styles, on travel mode choice, we make use of a latent class model framework, shown in Figure 1, where individual modality styles are represented as latent classes. Class membership is hypothesized to be a function of observable household and individual characteristics, medium and long-term travel and activity decisions, and the consumer surplus offered by different modality styles. The disturbances denote unobserved factors that influence class membership. Travel mode choices are conditioned on individual modality styles and on observable attributes of the different modal travel modes. Consistent with travel demand models employed by the SFMTC, six modal travel modes are defined: drive alone, shared ride, walk, bike, walk to transit and drive to transit. Heterogeneity across modality styles includes both the travel modes considered and the sensitivity to different travel mode attributes. The disturbances reflect unobserved factors that influence mode choice. Separate class-specific models are estimated for mandatory and non-mandatory tours.

Latent class choice models comprise two components: a class membership model and a class-specific choice model. We begin with the class-specific choice model, which predicts the probability that individual n over tour k chooses travel mode j, conditional on the individual belonging to latent class s, and is written as:

$$P\left(y_{nkj} = 1 | q_{ns} = 1\right) \tag{1}$$

, where y_{nkj} equals one if individual n over tour k chose travel mode j, and zero otherwise, and q_{ns} equals one if individual n belongs to latent class s, and zero otherwise.

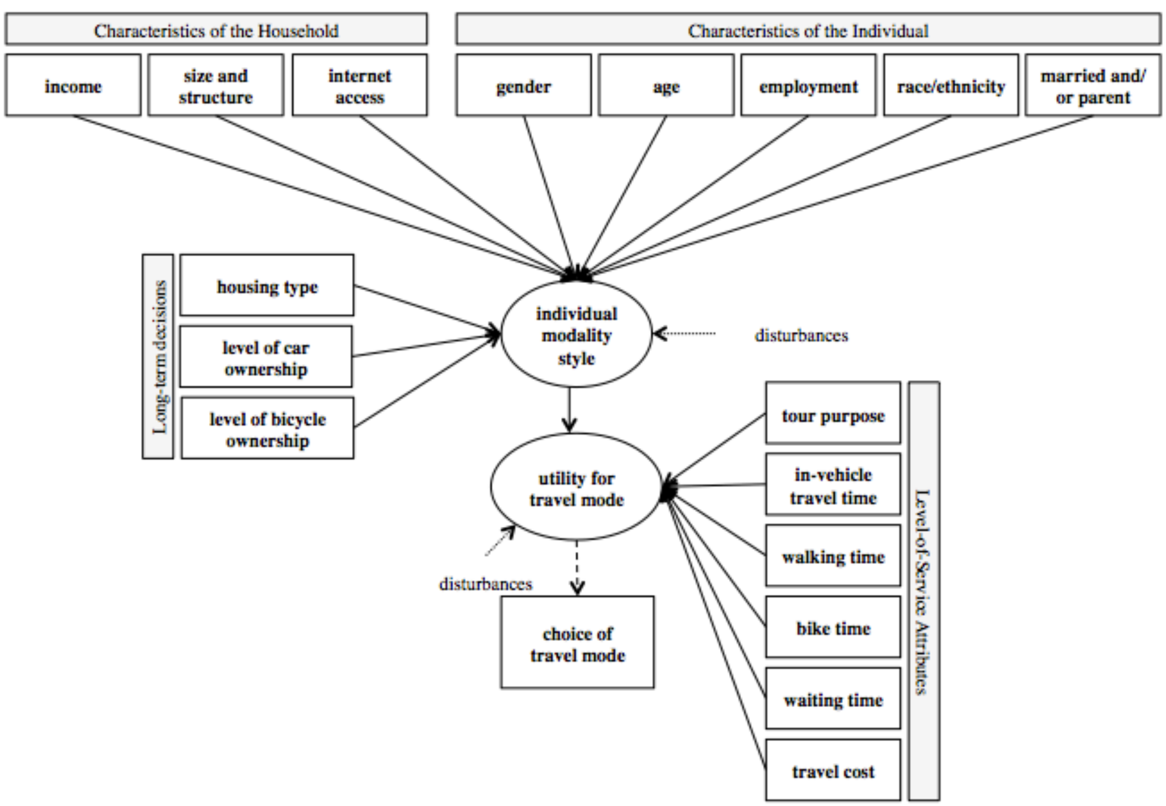

Figure 1: The influence of individual modality styles on travel mode choice

Let $u_{nkj|s}$ be the utility of travel mode j over tour k for individual n given that the individual belong to latent class s, which may be expressed as follows:

$$u_{nkj|s} = \mathbf{x}'_{\mathbf{nkj}}\boldsymbol{\beta}_{\mathbf{s}} + \varepsilon_{nkj|s} \tag{2}$$

, where $\mathbf{x}_{\mathbf{nkj}}$ is a vector of attributes of travel mode j over tour k for individual n, $\boldsymbol{\beta}_{\mathbf{s}}$ is a vector of parameters specific to the class s and $\varepsilon_{nkj|s}$ is the stochastic component of the utility specification. Assuming that all individuals are utility maximizers, the class-specific choice model may then be formulated as:

$$P(y_{nkj} = 1|q_{ns} = 1) = P(u_{nkj|s} \geq u_{nkj'|s} \forall j' \in \mathbf{C}_{\mathbf{nk|s}}) \tag{3}$$

, where $\mathbf{C}_{\mathbf{nk|s}}$ is the choice set for tour k for individual n given that the individual belongs to latent class s. Depending upon the distributional assumptions regarding $\varepsilon_{nkj|s}$, equation (3) can be reduced to different functional forms. We assume $\varepsilon_{nkj|s}$ to be i.i.d. Extreme Value across individuals, tours, travel modes and latent classes with mean zero and variance $\pi^2/6$, in which case equation (3) yields the familiar multinomial logit model:

$$P(y_{nkj} = 1|q_{ns} = 1) = \frac{\exp(\mathbf{x}'_{\mathbf{nkj}}\boldsymbol{\beta}_{\mathbf{s}})}{\sum_{j' \in \mathbf{C}_{\mathbf{nsk}}} \exp\left(\mathbf{x}'_{\mathbf{nkj'}}\boldsymbol{\beta}_{\mathbf{s}}\right)} \tag{4}$$

Equation (4) may be combined iteratively over travel modes and tours to yield the following conditional probability of observing the vector of choices $\mathbf{y}_{\mathbf{n}}$ for individual n:

$$P(\mathbf{y}_{\mathbf{n}}|q_{ns} = 1) = \prod_{k=1}^{K_n} \prod_{j \in \mathbf{C}_{\mathbf{nsk}}} \left[\frac{\exp(\mathbf{x}'_{\mathbf{nkj}}\boldsymbol{\beta}_{\mathbf{s}})}{\sum_{j' \in \mathbf{C}_{\mathbf{nsk}}} \exp\left(\mathbf{x}'_{\mathbf{nkj'}}\boldsymbol{\beta}_{\mathbf{s}}\right)}\right]^{y_{nkj}} \tag{5}$$

, where K_n denotes the number of distinct tours observed for individual n. The second piece to the latent class choice model is the class membership model, which predicts the probability that individual n belongs to latent class s, and is written as:

$$P(q_{ns} = 1) \tag{6}$$

Let u_{ns} be the utility of latent class s for individual n, which may be expressed as follows:

$$u_{ns} = \mathbf{z}'_{\mathbf{n}}\boldsymbol{\gamma}_{\mathbf{s}} + \varepsilon_{ns} \tag{7}$$

, where $\mathbf{z}_{\mathbf{n}}$ is a vector of characteristics of individual n, $\boldsymbol{\gamma}_{\mathbf{s}}$ is a vector of parameters and ε_{ns} is the stochastic component of the utility specification. Since individuals are utility maximizing, the class membership model may be stated as:

$$P(q_{ns} = 1) = P(u_{ns} \geq u_{ns'} \forall s' = 1, \dots, S) \tag{8}$$

, where S is the number of latent classes in the sample population. If ε_{ns} is assumed to be i.i.d. Extreme Value across individuals and latent classes with mean zero and variance $\pi^2/6$, then equation (10) may be reduced to the following multinomial logit model:

$$P(q_{ns} = 1) = \frac{\exp(\mathbf{z}'_{\mathbf{n}}\boldsymbol{\gamma}_{\mathbf{s}})}{\sum_{\mathbf{s}'=\mathbf{1}}^{S} \exp(\mathbf{z}'_{\mathbf{n}}\boldsymbol{\gamma}_{\mathbf{s}'})} \tag{9}$$

The number of classes S is determined exogenously by estimating models with different number of classes and using a combination of goodness-of-fit measures, such as the Bayesian Information Criterion and the Akaike Information

Criterion, and behavioral interpretation to select the most appropriate model. Equations (5) and (9) may now be combined to yield the marginal probability $\mathrm{P}(\mathbf{y_n})$ of observing the vector of choices $\mathbf{y_n}$ for individual n:

$$\mathrm{P}(\mathbf{y_n}) = \sum_{s=1}^{S} \frac{\exp(\mathbf{z}_n'\boldsymbol{\gamma}_s)}{\sum_{s'=1}^{S}\exp(\mathbf{z}_n'\boldsymbol{\gamma}_{s'})} \prod_{k=1}^{K_n} \prod_{j\in \mathbf{C}_{nsk}} \left[\frac{\exp(\mathbf{x}_{nkj}'\boldsymbol{\beta}_s)}{\sum_{j'\in \mathbf{C}_{nsk}} \exp\left(\mathbf{x}_{nkj'}'\boldsymbol{\beta}_s\right)}\right]^{y_{nkj}} \tag{10}$$

Equation (10) may be combined iteratively over all individuals to give the likelihood function for the sample population as follows:

$$\mathrm{L}(\boldsymbol{\beta},\boldsymbol{\gamma};\mathbf{y},\mathbf{X},\mathbf{Z}) = \prod_{n=1}^{N}\sum_{s=1}^{S} \frac{\exp(\mathbf{z}_n'\boldsymbol{\gamma}_s)}{\sum_{s'=1}^{S}\exp(\mathbf{z}_n'\boldsymbol{\gamma}_{s'})} \prod_{k=1}^{K_n} \prod_{j\in \mathbf{C}_{nsk}} \left[\frac{\exp(\mathbf{x}_{nkj}'\boldsymbol{\beta}_s)}{\sum_{j'\in \mathbf{C}_{nsk}} \exp\left(\mathbf{x}_{nkj'}'\boldsymbol{\beta}_s\right)}\right]^{y_{nkj}} \tag{11}$$

The unknown parameters $\{\boldsymbol{\beta},\boldsymbol{\gamma}\}$ may be estimated by maximizing equation (11) using the EM algorithm.

2. The EM Formulation

The first step to using the EM Algorithm is to write the complete log-likelihood to uncover the form of the M step estimates as well as the sufficient statistics required for the E step. Assuming that the individual modality styles are no longer latent but observable variables, the complete likelihood $\mathrm{L_c}$ can be written as:

$$\mathrm{L_c} = \left[\prod_{n=1}^{N}\prod_{s=1}^{S}\left[\frac{\exp(\mathbf{z}_n'\boldsymbol{\gamma}_s)}{\sum_{s'=1}^{S}\exp(\mathbf{z}_n'\boldsymbol{\gamma}_{s'})}\right]^{q_{ns}}\right]\left[\prod_{n=1}^{N}\prod_{s=1}^{S}\prod_{k=1}^{K_n}\prod_{j\in \mathbf{C}_{nsk}}\left[\frac{\exp(\mathbf{x}_{nkj}'\boldsymbol{\beta}_s)}{\sum_{j'\in \mathbf{C}_{nsk}} \exp\left(\mathbf{x}_{nkj'}'\boldsymbol{\beta}_s\right)}\right]^{y_{nkj}q_{ns}}\right]$$

Taking the logarithm, it can be seen that the log-likelihood function $\mathcal{L}_c$ breaks apart quite conveniently into two separate components, one each corresponding to the two endogenous variables:

$$\Rightarrow \mathcal{L}_c = \sum_{n=1}^{N}\sum_{s=1}^{S} q_{ns}\log\left[\frac{\exp(\mathbf{z}_n'\boldsymbol{\gamma}_s)}{\sum_{s'=1}^{S}\exp(\mathbf{z}_n'\boldsymbol{\gamma}_{s'})}\right] + \sum_{n=1}^{N}\sum_{s=1}^{S}\sum_{k=1}^{K_n}\sum_{j\in \mathbf{C}_{nsk}} y_{nkj}q_{ns}\log\left[\frac{\exp(\mathbf{x}_{nkj}'\boldsymbol{\beta}_s)}{\sum_{j'\in \mathbf{C}_{nsk}} \exp\left(\mathbf{x}_{nkj'}'\boldsymbol{\beta}_s\right)}\right] \tag{12}$$

From above, it is clear that q_{ns} is the only sufficient statistic required for estimating all of the unknown parameters. Let $\boldsymbol{\theta} = (\boldsymbol{\beta},\boldsymbol{\gamma})$ denote the vector of unknown parameters to be estimated. Then the expectation of q_{rh} may be calculated as follows:

$$\begin{aligned} \mathrm{E}[q_{ns}|\mathbf{y};\boldsymbol{\theta}] &= \mathrm{P}(q_{ns} = 1|\mathbf{y};\boldsymbol{\theta}) = \mathrm{P}(q_{ns} = 1|\mathbf{y_n};\boldsymbol{\theta}) \\ &= \frac{\mathrm{P}(\mathbf{y_n}|q_{ns} = 1;\boldsymbol{\theta})\mathrm{P}(q_{ns} = 1|\boldsymbol{\theta})}{\mathrm{P}(\mathbf{y_n}|\boldsymbol{\theta})} \end{aligned}$$

$$\Rightarrow q_{ns}^{(t+1)} = \frac{\mathrm{P}\left(\mathbf{y_n}|q_{ns} = 1;\boldsymbol{\theta}^{(\mathbf{t})}\right)\mathrm{P}\left(q_{ns} = 1|\boldsymbol{\theta}^{(\mathbf{t})}\right)}{\mathrm{P}(\mathbf{y_n}|\boldsymbol{\theta}^{(\mathbf{t})})} \tag{13}$$

The function $\mathrm{P}\left(\mathbf{y_n}|q_{ns} = 1;\boldsymbol{\theta}^{(\mathbf{t})}\right)$ is given by equation (5), $\mathrm{P}\left(q_{ns} = 1|\boldsymbol{\theta}^{(\mathbf{t})}\right)$ is given by equation (9), and $\mathrm{P}\left(\mathbf{y_n}|\boldsymbol{\theta}^{(\mathbf{t})}\right)$ is given by equation (10). Having derived expressions for the updates in the E-Step for the sufficient statistic, we can proceed now to the M-Step. Taking the derivative of the complete log-likelihood with respect to the unknown parameters, we get the following updates for the M-Step:

$$\boldsymbol{\gamma}^{(t+1)} = \underset{\boldsymbol{\alpha}}{\operatorname{argmax}} \sum_{h=1}^{H} \sum_{r=1}^{R} q_{ns}^{(t+1)} \log\left[\frac{\exp(\mathbf{z}_{\mathbf{n}}' \boldsymbol{\gamma}_{\mathbf{s}})}{\sum_{\mathbf{s}'=\mathbf{1}}^{S} \exp(\mathbf{z}_{\mathbf{n}}' \boldsymbol{\gamma}_{\mathbf{s}'})}\right] \tag{14}$$

$$\boldsymbol{\beta}_{\mathbf{s}}^{(t+1)} = \underset{\boldsymbol{\beta}_{\mathbf{s}}}{\operatorname{argmax}} \sum_{\mathbf{n}=\mathbf{1}}^{\mathbf{N}} \sum_{k=1}^{K_n} \sum_{j \in \mathbf{C}_{\mathbf{nsk}}} y_{nkj} q_{ns}^{(t+1)} \log\left[\frac{\exp\left(\mathbf{x}_{\mathbf{nkj}}' \boldsymbol{\beta}_{\mathbf{s}}\right)}{\sum_{j' \in \mathbf{C}_{\mathbf{nsk}}} \exp\left(\mathbf{x}_{\mathbf{nkj}'}' \boldsymbol{\beta}_{\mathbf{s}}\right)}\right] \tag{15}$$

Equations (14) and (15) are both weighted multinomial logit models that can be solved fairly efficiently. The EM algorithm iterates between the E-step of equation (13) and the M-step of equations (14)-(15), until the convergence criterion is satisfied.