Latent class analysis of residential and work location choices

Rajesh Paleti [a], Sabyasachee Mishra [b*], Khademul Haque [b], Mihalis M. Golias [b]

[a] Department of Civil and Environmental Engineering, Old Dominion University, Norfolk, VA 23529, United States.

[b] Department of Civil Engineering, University of Memphis, 3815 Central Avenue, Memphis, TN 38152, United States.

ABSTRACT

This paper developed a two-stage modeling framework for analyzing residential and work location choices with probabilistic choice sets. In the first stage, a household (or a worker) was assumed to select a neighborhood (such as central business district, urban, suburban *etc.*) to live (or work). In the second stage, the household (or worker) was assumed to choose a specific zone conditional on the selected neighborhood. The neighborhood choice model component takes the form of Manski model with latent choice sets. The model was used to analyze residential and work location decisions in Nashville, Tennessee. The model results indicate significant heterogeneity in the consideration probability of different neighborhood alternatives both in the residential and work location choices.

Keywords: residential location, work location, latent class models, Manski model

[*] Corresponding author.
Email-addresses: rpaleti@odu.edu (R. Paleti), smishra3@memphis.edu (S. Mishra), khaque@memphis.edu (K. Haque), mgkolias@memphis.edu (M. Golias)

1. BACKGROUND AND MOTIVATION

Residential and work location choices are medium-to-long term decisions that have a significant impact on day-to-day activity-travel decisions of people. These choices are typically modeled using discrete choice models that assume certain decision making mechanism. For instance, the Random Utility Maximization (RUM) rule is one such mechanism in which the decision maker is assumed to choose the alternative that provides the highest utility. Within the class of discrete choice models, the multinomial logit (MNL) and its generalizations (e.g., nested logit, cross nested logit *etc.*) are commonly used to analyze travel-related choices. In these models, the utilities of different alternatives are specified as a function of different observed variables collected from household survey data that can affect the choice being modeled. However, several important aspects including the attitudes and preferences, the consideration choice set, and the decision making mechanism are typically not observed in the survey data (Walker and Li, 2006). For instance, it is reasonable to assume that there are certain households/people who have greener life styles or tech-savvy attitudes from the rest of the population. People in these '*neo'* households are likely to have different residential and work location preferences compared to those in '*conventional'* households (Bhat and Guo, 2007). But, these attitudinal variables are not available in most revealed preference datasets. The effects of these unobserved factors can manifest in different ways. For instance, these factors can lead to heterogeneity in travel sensitivities across different population segments or lead to variation in the consideration choice set across decision makers. So, standard choice models such as the MNL model cannot control for these factors. In such scenarios, latent class models that can probabilistically classify households into latent classes (e.g., neo and conventional) are particularly useful. It is important to note that these groups or classes are not observed in the real world (and hence the name '*latent*').

Latent class choice models have been applied in various disciplines. Methodological development and model application is spread over multiple domains including marketing research (Dillon et al., 1994; Grover and Srinivasan, 1987; Russell and Kamakura, 1993; Swait, 1994; Swait and Sweeney, 2000), economics (Boxall et al., n.d.; Boxall and Adamowicz, 2002), transportation (Walker and Li, 2006), geography(Baerenklau, 2010; Hynes et al., 2008; Scarpa and Thiene, 2005), agriculture (Mitani et al., 2008) and health science (Bandeen-Roche et al., 2006; Bucholz et al., 1996; Jung and Wickrama, 2008; Lanza and Rhoades, 2011). Application of latent class models in transportation planning can be summarized into four categories. *First*, studies that focused on varying travel sensitivities and preferences where endogenous market segmentations are made based on intrinsic biases and responsiveness to level-of-service attributes (Bhat, 1997, 1998; Greene and Hensher, 2003). Recently, researchers also started to explore *attribute non-attendance* where some respondents only consider a subset of attributes during decision making (Martínez et al., 2009). These studies can also be grouped under the category of those dealing with varying travel sensitivity. *Second*, studies that analyze the variation in consideration choice sets across decision makers (Manski, 1977; Martínez et al., 2009). *Third*, studies that recognize that people might use alternate decision making mechanisms or decision rules such as RUM or Random Regret Minimization (RRM) while evaluating choice alternatives (Hess and Stathopoulos, 2013). *Fourth*, studies that considered all possible dependency pathways while modeling multiple choices simultaneously. For instance, work location decisions can be made conditional on residential location or vice-versa leading to two different dependency pathways (Waddell et al., 2007).

The current research belongs to the second group of studies that aim to uncover population segments with varying choice sets in residential and work location choices. Typically, location choices are undertaken at the zonal level (*i.e.*, traffic analysis zone, block, or parcel). The size of choice set in location choice models is typically large extending into thousands of alternatives. Even with moderately sized choice sets, it is difficult to identify more than 2-3 latent classes in most empirical applications. So, it can be quite challenging to uncover latent classes with large choice sets. While researchers have used sampling techniques to resolve the computational problem associated with large choice sets, the sampling mechanism itself might introduce some bias and make it further difficult to identify latent segments. To address this problem, the current study adopted a two-stage decision framework for location choices. In the first stage, a household (or a worker) will select a neighborhood (such as central business district, urban, suburban *etc.*)

to live (or work). In the second stage, the household (or worker) will choose a specific zone conditional on the selected neighborhood in the first stage. The latent class analysis is undertaken at the first stage which has a much smaller choice set compared to the conditional zonal choice model in the second stage. For instance, certain households might only consider high density neighborhoods while deciding where to reside leading to varying consideration choice sets in the neighborhood choice model. The two-stage modeling framework is also reasonable from a behavioral standpoint because households are very unlikely to consider all zones within the study area while making decisions regarding where to live and work. They are more likely to choose a neighborhood and then explore residential choices within the neighborhood. However, these two components are not completely independent. Better zonal alternatives within a neighborhood should increase the likelihood of choosing that neighborhood over others. This dependence between the neighborhood and zonal choice components is captured by using log-sum from conditional zonal choice model as an explanatory variable in the utility of the neighborhood choice model component.

The remainder of the paper is organized as follows. Literature review is presented in the next section followed by methodology. The case study and data section describes the study area and data used for model development. The result and discussion provides insights on case study findings and possible application of model results in transportation planning and demand modeling. The final section concludes the paper and outlines the scope of future research.

2. LITERATURE REVIEW

As briefly discussed in the introduction section, the literature review is presented along four themes to draw insights from past transportation research: (1) Endogenous market segmentation; (2) choice set variation; (3) heterogeneous decision rules, and (4) alternative dependency pathways.

2.1. Endogenous Market Segmentation

Bhat (1997) recognized the need to accommodate differences in intrinsic mode biases (preference heterogeneity) and differences in responsiveness to level-of-service attributes (response heterogeneity) across individuals. He incorporated preference and response heterogeneity into the MNL when studying mode choice behavior from cross-sectional data in an intercity travel. The study found that endogenous segmentation model described causal relationship best and provided intuitively more reasonable results compared to traditional approaches (Bhat, 1997). Walker and Li (2006) conducted an empirical study of residential location choices and uncovered three lifestyle segments – suburban dwellers, urban dwellers, and transit riders with varying location preferences (Walker and Li, 2006). Similarly, Wen and Lai (2010) demonstrated using air carrier choice data that the latent class model outperforms the standard MNL model considerably (Wen and Lai, 2010). Arunotayanun and Polak (2011) identified three latent segments in the context of freight mode choice of shippers with three alternatives – small truck, large truck, and rail as a function of several attributes including transport time, cost, service quality and service flexibility (Arunotayanun and Polak, 2011). While the first segment was found to be highly sensitive to all attributes considered, the second segment preferred better service quality and the third segment preferred better service flexibility.

Wen *et al.* (2012) used a nested logit latent class model for high speed rail access in Taiwan and showed that flexible substitution patterns among alternatives and preference heterogeneity in the latent class nested logit model outperformed traditional models (Wen et al., 2012). More recently, several studies analyzed attribute non-attendance which may be considered as a variant of taste heterogeneity in which some respondents make their choices based on only a subset of attributes that described the alternatives at hand (Hensher, 2010). For example, it is possible that a portion of respondents do not care about time savings while making travel decisions. Scarpa (2009) showed that 90% of the respondents do not consider cost while choosing rock-climbing destination spots. Similarly, Campbell *et al.* (2011) revealed that 61% of respondents are not attending to cost while making environmental choices (Campbell et al., 2010). Hess and Rose (2007) proposed a latent class approach to accommodate attribute non-attendance (Hess and Rose, 2007), and a number of studies adopted similar approach thereafter (Hensher et al., 2011; Hensher and Greene, 2009; Hess and Rose, 2007). Hess *et al.* (2012) suggest that with this approach, different latent classes relate to different combinations of attendance and non-attendance across attributes (Hess et al., 2012). Model estimation is conducted to compute a non-zero coefficient, which is used in the attendance classes, while the attribute is not employed in the non-attendance classes, *i.e.* the coefficient is set to zero. In a complete specification, covering all possible combinations, this would thus lead to 2^K classes, with *K* being the number of attributes (Hess et al., 2012).

2.2. Choice Set Variation

Manski (1977) developed the theoretical framework for the two stage decision process that accounts for choice set heterogeneity (Manski, 1977). Decision makers were assumed to first construct their choice set in a non-compensatory manner and then make choice conditional on the generated choice set using a compensatory mechanism (e.g., RUM). The choice probability of an alternative is obtained as a weighted probability of choosing that alternative over all possible choice sets. Swait and Ben-Akiva (1987) and Ben-Akiva and Boccara (1995) build on Manski's framework and used explicit random constraints to determine the choice set generation probability (Ben-Akiva and Boccara, 1995; Swait and Ben-Akiva, 1987). Bierlaire *et al.* (2009) stated that earlier latent class choice set generation methods are hardly applicable to medium and large scale choice problems because of the computational complexity that arises from the combinatorial number of possible choice sets. If the number of alternatives in the universal choice set is *C*, the number of possible choice sets is ($2^C - 1$) (Bierlaire et al., 2009). Several heuristics that derive tractable models by approximating the choice set generation process were developed (Calastri et al., 2017; Chiang et al., 1998; Gilbride and Allenby, 2004; Hauser, 2014; Swait, 2001a, 2001b). The most promising heuristics are based on the use of penalties of the utility functions, and were proposed by Cascetta and Papola (2001) and further expanded by Martinez *et al.* (2009) (Cascetta and Papola, 2001; Martínez et al., 2009). These heuristics were recently further modified to closely replicate the Manski's original formulation (Paleti, 2015).

2.3. Heterogeneous Decision Rules

An increasing number of studies investigated the use of alternatives to random utility maximization (RUM) rule to explore which paradigm of decision rules best fits a given dataset as well as the variation in decision rules across respondents. Srinivasan *et al.* (2009) developed a latent class model that assigns respondents to either the utility maximizing or disutility minimizing segments probabilistically for analyzing mode choice decisions. This study found that only 32.5% respondents belong to the utility maximizing segment whereas a majority (67.5%) belonged to the disutility minimizing segment (Srinivasan et al., 2009). Along similar lines, Hess *et al*. (2013) developed latent class models that linked latent character traits to choice of decision rule between RUM and RRM. This found an almost even split between the shares of respondents that adopted the two decision rules in the context of commute mode choice (Hess and Stathopoulos, 2013). Zhang *et al.* (2009) examined different types of group decision-making mechanisms in household auto ownership choices using latent classes models (Zhang et al., 2009).

2.4. Alternate Dependency Pathways

Joint choice modeling can result in several pathways of dependency among the choice dimensions considered. However, one of the challenges is that as the number of choice dimensions in the integrated modeling framework increases, the number of possible dependency pathways among choice dimensions can explode very quickly. Specifically, there are $K!$ possible dependency structures in an integrated model with K choice dimensions. So, it is not always possible to estimate latent models with all possible dependency pathways. However, latent class models can be useful in empirical contexts where there are very limited dependency pathways. For instance, Waddel *et al.* (2007) used latent class models to estimate the proportion of households in which residential location choice is made conditional on workplace location choice and *vice versa* (Waddell et al., 2007). However, this study only considered single-worker households because of several possible permutations of work and home location choices in multi-worker households. Additionally, the authors also mention the complexity involved in modeling the interdependencies when dynamics among choice dimensions can change over time.

In summary, latent class models have proven useful with better policy insights and improved statistical fit in a wide array of empirical contexts within transportation. Moreover, these models have the same data requirements as standard un-segmented models. However, it might not be analytically tractable to estimate latent class models in certain choice contexts without making some simplifying assumptions.

3. METHODOLOGICAL FRAMEWORK

Let q be the index for the decision maker (household for residential location choice and individual employee for work location choice). As discussed earlier, in this study, a two-stage decision making process in which the decision maker first chooses the neighborhood in the first stage and then looks for a specific zone within the chosen neighborhood in the second stage was assumed. Both these two model components were estimated sequentially but the expected utility or logsum from the zonal destination choice was used as an explanatory variable in the neighborhood choice alternatives to link the two model components. Moreover, it is unlikely that all decision makers consider the full set of neighborhoods while making the first stage neighborhood choice. This variation in the consideration choice set of neighborhood choice is accounted using the latent choice set Manski model. Lastly, the universal choice set of zonal choice conditional on the neighborhood in the second stage comprises of all zones within the chosen neighborhood of the decision maker. The size of this choice set can still be quite large. So, importance sampling methods were used to construct the sampled choice set for zonal choice in the second stage. A brief overview of different modeling components is presented below.

3.1. Neighborhood Choice Component

Let i be the index for neighborhood alternative and C denote the universal choice set of the neighborhood choice $C = \{1 = \text{CBD}, 2 = \text{URBAN}, 3 = \text{SUBURBAN}, 4 = \text{RURAL}\}$. It is very likely that decision maker q only considers a subset C_q (of C), known as the consideration choice set, while making the actual choice. In the multinomial logit (MNL) framework, the utility associated with alternative i can be written as:

$$U_q^i = V_q^i + \varepsilon_q^i = \boldsymbol{\beta}_i' \boldsymbol{X}_q^i + \varepsilon_q^i \quad (1)$$

Where $V_q^i = \boldsymbol{\beta}_i' \boldsymbol{X}_q^i$ is the observed part of the utility, $\boldsymbol{X}_q^i$ is the vector of explanatory variables, and $\boldsymbol{\beta}_i$ is the corresponding column vector of coefficients, and ε_q^i is standard gumbel random variable that captures all unobserved factors that is independent and identically distributed across alternatives and decision makers. The vector $\boldsymbol{X}_q^i$ also includes logsum from the conditional zonal destination choice model.

So, the probability of a decision maker q choosing an alternative 'i' from a set of mutually exclusive and exhaustive alternatives C_q is given by:

$$P_q(i|C_q) = \frac{e^{V_q^i}}{\sum_{j \in C_q} e^{V_q^j}} \tag{2}$$

However, the consideration set C_q is not observed by the analyst. To resolve this problem, past researchers have assumed that observed choice is an outcome of two latent (unobserved) steps – (1) formation of consideration set C_q from the universal choice set and (2) choice conditional on the consideration set C_q. So, the unconditional probability that decision maker q chooses neighborhood i is obtained as a weighted average across all possible consideration sets using Bayes' theorem as follows (Manski, 1977):

$$P_q(i) = \sum_{C_q \in C} P_q(i|C_q) \times P_q(C_q) \tag{3}$$

The consideration set formation step of the Manski model is viewed as a non-compensatory process whereas the second step is viewed as an outcome of compensatory mechanism (in our case, this is the Random Utility Maximization (RUM) principle in the MNL model). Consistent with this notion, the probability ϕ_q^i that decision maker q considers alternative i is specified as a binary logit model as follows:

$$\phi_q^i = \frac{e^{\gamma_i' z_q^i}}{1+e^{\gamma_i' z_q^i}} \tag{4}$$

where $\boldsymbol{Z}_q^i$ is the vector of variables that impact whether alternative i is considered by decision maker q or not and $\boldsymbol{\gamma}_i$ is the corresponding column vector of coefficients. The probability of different consideration sets can be computed using these individual consideration probabilities. For instance, the probability of decision maker q considers the choice set {CBD, URBAN} is given as follows:

$$P_q[(CBD, URBAN)_q] = \phi_q^{CBD} \times \phi_q^{URBAN} \times (1 - \phi_q^{SUBURBAN}) \times (1 - \phi_q^{RURAL}) \tag{5}$$

There are 15 possible consideration sets in the universal choice set comprising of four alternatives – CBD, URBAN, SUBURBAN, and RURAL, excluding the null choice set without any alternatives. However, all these subsets of alternatives are not intuitive from a behavioral standpoint. For instance, {CBD, RURAL} is one such possible subset of alternatives.It is difficult to justify why someone might consider both the extreme neighborhoods (CBD and RURAL) that have very different residential and employment composition but not intermediate options (URBAN and SUBURBAN). To avoid such instances of behavioral inconsistency, we only considered the following 10 feasible consideration choice sets that avoid discontinuity: {CBD, URBAN, SUBURBAN, RURAL}, {URBAN, SUBURBAN, RURAL}, {CBD, URBAN, SUBURBAN}, {CBD, URBAN}, {URBAN, SUBURBAN}, {SUBURBAN, RURAL}, {CBD}, {URBAN}, {SUBURBAN}, {RURAL}.

Lastly, to ensure that the sum of probabilities across all alternatives in the universal choice set add up to one, all the choice probabilities are re-scaled by the factor (1-probability of all infeasible choice sets).

3.2. Conditional Zonal Destination Choice Component

Let *s* denote the index for location *i.e.* Traffic Analysis Zone (TAZ). The observed part of the utility function for each alternative in the zonal choice set V_q^s can be written as follows:

$$V_q^s = LN\left(Size_q^s\right) + \boldsymbol{\pi}'\boldsymbol{W}_s + \boldsymbol{\alpha}' \times \boldsymbol{LOS}_{q,h,s} + \boldsymbol{\delta}' f\left(D_{h,s}\right) \tag{6}$$

where is $Size_q^s$ the size variable for destination zone *s* for decision maker *q* (zonal household population for residential location and industry-specific zonal employment for work location), $\boldsymbol{W}_s$ is vector of zonal variables describing zonal alternative *s* and $\boldsymbol{\pi}$ is the corresponding vector of coefficients, $\boldsymbol{LOS}_{q,ij}$ is the set of level-of-service variables between zone pair (*h,s*) where *h* is the home zone and their interaction with decision maker characteristics and $\boldsymbol{\alpha}$ is the corresponding vector of coefficients, $D_{i,s}$ is the network distance between home zone *h* and work zone alternative *s*, and $\boldsymbol{f}\left(D_{h,s}\right)$ is a vector of non-linear functions of $D_{i,s}$ (for example, linear, squared, cubic, and logarithmic) and $\boldsymbol{\delta}$ is the corresponding vector of coefficients. Please note that the last two components (LOS and distance-based impedance measures) are relevant only to the work location component where we assume that the home location is already known. Assuming *i.i.d.* standard Gumbel term assumption for the unobserved part of utilities will lead to the MNL model.

3.3 Sampling Destination Zones

As mentioned earlier, it is computationally difficult to consider all location alternatives within a neighborhood during model estimation. While a completely random sampling approach will produce consistent parameter estimates, it is not an efficient option. So, a sampling-by-importance model with TAZ activity-specific size terms (for both residential and work location models) and a coefficient of -0.1 for "Distance between home and work TAZ" variable (only for work location choice) was applied. During model estimation, a correction term equal to $\ln\left(\frac{n_i}{N \times q(i)}\right)$ was added to the utility function of the sampled alternative to account for the difference in the sampling probability and the frequency of the alternative in the sample. The sampling correction term represents natural logarithm of the ratio of the sampling frequency to selection probability for each alternative as was substantiated in the theory (Ben-Akiva and Lerman, 1985; McFadden, 1978). In this correction term, $q(i)$ is the selection probability (probability to be drawn) which is a function of size variable and simplified distance-based impedances, n_i is the selection frequency in the sample or the number of times an alternative is chosen, and *N* is the sample size (= 50 because we sample fifty TAZs).

4. CASE STUDY DATA

The data for this study is derived from the 2012 household travel survey data conducted in Nashville metropolitan area. In addition to geo-coded location information, the data include detailed socio-economic and demographic data and activity travel diary data of all respondents. The travel skims and network related variables were gathered form the Nashville Travel Demand Model (TDM). Instead of using the standard definition of spatial unit of location choices (census tract or TAZ), this paper employs neighborhood categories based on household and employment density to characterize location choices. This helps make the definition of choice alternatives clear and manageable and more effectively captures the notion that people are looking for a built environment (land use density) that suits their mobility and lifestyle preferences. In other words, people are not choosing between TAZ A or B directly, but rather between a unit that offers a built environment of certain attributes versus another unit that offers a different built environment. Residence and workplace locations are categorized into four possible alternatives or neighborhoods based on a combination of population and employment density (population and employment in the half mile radius).

Only workers with work location outside home were considered in our analysis. One of the key variables in the work location choice model is industry in which the worker is employed. The Nashville travel demand model (TDM) uses work industry definition with five categories – agriculture and mining, retail, manufacturing, transportation, and office. The disaggregate work industry variable in the survey data was grouped together into these five categories to be consistent with the regional TDM. Several explanatory variables were considered in this study including age and gender composition, worker characteristics, household income, educational attainment, housing type, housing tenure, auto and bike ownership, typical commute mode choice and average daily trip frequency. In addition, distance, auto travel time, transit availability, and transit generalized cost were obtained from the network skims. Also, Hansen-type accessibility measures that indicate a zone's accessibility to different types of activity opportunities and mode choice log-sums were calculated using zonal data and network skim files.

After extensive data cleaning, the final estimation sample includes 4,344 households and 3,992 employed individuals without any missing information on all explanatory variables used in this study. The distribution of individuals in the four residential neighborhood alternatives was - 8.90% rural, 29.74% suburban, 60.36% urban, and 1.00% CBD. The distribution of individuals with respect to work neighborhood was 2.96% rural, 17.41% suburban, 65.88% urban, and 13.75% CBD. In the final sample, the share of respondents who live in CBD was quite low. So, the estimation of latent choice set model where people considered CBD alternative probabilistically is difficult with such small sample size. So, in the choice set formation stage of decision-making, the respondents are assumed to consider the CBD and URBAN alternatives as a bundle but not separately. So, the set of feasible choice sets is reduced to the six possibilities: {CBD, URBAN, SUBURBAN, RURAL}, {CBD, URBAN, SUBURBAN}, {CBD, URBAN}, {SUBURBAN, RURAL}, {SUBURBAN}, {RURAL}. For the same reasons, the SUBURBAN and RURAL alternatives are assumed to be considered as bundle in the latent choice set formation component of the work neighborhood choice model.

<<Figure 1 Here>>

5. RESULTS AND DISCUSSION

The location choice models comprise of two components – neighborhood choice and zonal choice conditional on neighborhood. For brevity, only results of the final models are presented in this study. For the Manski models with probabilistic choice sets, each explanatory variable was tested both in the utility specification as well as the alternative consideration probability specification for each alternative and the specification that provided better data fit as chosen.

5.1 Neighborhood Choice Component

Tables 1 and 2 present the results of the neighborhood choice components of residential location and work location models respectively.

5.1.1 Residential Neighborhood Choice

The CBD alternative was chosen as the reference alternative. Given that there are several other variables in the model, the constants in the model do not have substantive behavioral interpretation. Notwithstanding this, the relative magnitude of constants suggests that people, on average, prefer URBAN, SUBURBAN, and RURAL neighborhoods (and in that order) compared to CBD areas. Households with higher trip frequency are more likely to reside in the URBAN and SUBURBAN regions of the study area. As expected, households with children are more likely to reside in SUBURBAN and RURAL areas. Interestingly, households with more jobs (*i.e.,* workers) are less likely to live in URBAN areas. Households with more female members are more inclined to reside in less dense neighborhoods. Households with higher number of licensed drivers tend to live in the suburban and rural neighborhoods. The high positive parameter estimates on single family detached households show that these households almost certainly do not live in CBD neighborhood. Households with zero vehicles are most likely group to live in the CBD whereas households with more cars than driving age adults are more inclined to live in less dense neighborhoods. Households with more than $75K income and higher educational attainment (bachelor degree and higher) are less likely to reside in low density neighborhoods.

Among the four alternatives, the two low density options – SUBURBAN and RURAL were found to be considered probabilistically. Specifically, owner-occupied households are more likely to consider SUBURBAN and RURAL households. Also, while higher bicycle ownership levels are associated with higher chances of considering SUBURBAN neighborhood, it reduces the chances of considering RURAL neighborhood. This result is probably indicative of inadequate bicycle and pedestrian infrastructure in RURAL areas. As expected, households with higher average age are more likely to consider RURAL neighborhood compared to relatively younger households.

For comparison purposes, a MNL model that assumes that all households consider all the four neighborhood options was also estimated. The log-likelihood of the MNL and Manski models are -3,509.7 and -3,502.1, respectively. The Manski model has 7 additional parameters in the latent choice set component compared to the MNL model. The Likelihood Ratio (LR) test statistic of comparison between the two models is 19.30 that is significantly greater than 14.07 which is the critical chi-squared value corresponding to 7 degrees of freedom at 95 percent confidence level. This underscores the importance of accounting for latent choice sets in residential neighborhood choices.

<<Table 1 Here>>

5.1.2 Work Neighborhood Choice

The RURAL alternative was chosen as the reference alternative. Workers with disability are more likely to be employed in the URBAN and SUBURBAN neighborhoods compared to CBD and RURAL areas. This is intuitive because disabled workers do not prefer longer trips typically associated with RURAL neighborhood as well as crowded environment of CBD neighborhood. Workers who use auto mode for commute are less likely to be employed in CBD and URBAN areas. On the other hand, workers who have flexible work schedule and work five days a week are more likely to be employed in CBD and URBAN neighborhoods. Industry type was found to have a strong impact on work neighborhood choice. For instance, workers in agriculture, manufacturing, and transportation industries are more likely to be working in RURAL neighborhood which is consistent with the land use in these areas (e.g., agricultural land, factories, construction sites *etc*). On the other hand, workers employed in the office sector with desk jobs tend to work in CBD and URBAN neighborhoods. There was strong dependence between residential and work neighborhood choices with workers who reside in denser neighborhoods being more inclined to work in denser neighborhoods. Workers with lower education levels are more likely to work in low density neighborhood and less inclined to work high density CBD neighborhood.

Among the four alternatives, the two low density options – SUBURBAN and RURAL were found to be considered probabilistically. But, as discussed earlier, these alternatives were assumed to be considered as a bundle in the latent choice component of the Manski model. Workers with flexible work schedule are more likely to consider these low density neighborhoods compared to workers with fixed work schedule. Also, workers employed in retail industrial sector are less likely to consider low density neighborhoods in their work neighborhood choices.

Again, the log-likelihood of the MNL and Manski work neighborhood models are -3,531.8 and -3,511.2, respectively. The LR test statistic of comparison between the two models is 41.09. This value is considerably larger than 7.82 which is the critical chi-squared value corresponding to 3 degrees of freedom at 95 percent confidence level. This indicates superior data fit in the Manski model.

5.2. Zonal Destination Choice Component

Table 3 presents the results of conditional zonal destination choice model components of residential and work location choice models.

5.2.1. Zonal Residential Location Choice

The coefficient on the size variable – natural logarithm of the "total number of households in the TAZ' is fixed to one to ensure that individual destination zone preferences sum up to zonal control totals.

Accessibility to different types of employment opportunities was found to a significant determinant of zonal residential location decisions. To be specific, households are less likely to reside in zones with high manufacturing accessibility which is expected given that these zones tend to have higher pollution levels and limited infrastructure for recreational activities. Interestingly, owner-occupied households tend to reside in areas with lower total employment accessibility compared to rental households. On the other hand, households with more children and senior adults prefer zones with better accessibility. Also, zones with higher total employment accessibility attract households with lower auto ownership levels, higher income, and higher educational attainment.

5.2.2. Zonal Work Location Choice

The coefficient on size variable – natural logarithm of "zonal employment in the industry of the worker' was fixed to one for the reasons alluded to above. Zones that are closer to home TAZ and with shorted auto travel times from home TAZ are more likely to be chosen compared to farther alternatives. Also, presence of transit service between home TAZ and destination TAZ was found to significantly enhance the likelihood of the person working in that zonal alternative. Women, workers in households with young children, and workers with higher educational attainment tend to prefer zonal alternatives that are in closer proximity to home TAZ. Also, workers with young children, varying work location, and those who work more than 40 hours per week are more sensitive to inter-zonal travel time between home and destination TAZ indicative of relatively higher time pressure on these individuals. Lastly, workers who reside in sub-urban neighborhood are more sensitive to travel time compared to those who reside in CBD, URBAN, and RURAL neighborhoods.

<<Table 3 Here>>

5.3 Elasticity Effects

The elasticity effects were computed as a percentage change in the aggregate shares of four different neighborhood alternatives due to a unit change in the explanatory variable. The unit change in the case of indicator variables is from 0 to 1 whereas in case of ordinal variables, the variable value was increased by one unit. The results of elasticity analysis for the residential and work neighborhood choice models are presented in Tables 4 and 5, respectively. For instance, in table 4 it can be noticed that households with more cars than adults at legally driving age (defined as “high auto-sufficiency”) are 142.3% more likely to live in RURAL neighborhood compared to households with fewer or same number of cars as driving age adults. Similarly, One additional child between 6 and 10 years makes a household 37% and 28% more likely to reside in RURAL and SUBURBAN neighborhoods and 16% and 9.4% less inclined to locate in CITY and URBAN neighborhoods, respectively. This result may be indicative of parents preferring greener neighborhoods with more open areas (recreational parks and play grounds) for supporting their children. Also, suburban areas tend to have more (and in some cases, better) schools that attract households with children. Other numbers in the table can be interpreted similarly. Based on the relative magnitude of elasticity effects in Table 4, the demographic groups most likely to reside in CBD, URBAN, SUBURBAN, and RURAL neighborhoods are households with high educational attainment (graduate degree), households with zero vehicles, single family detached households, and owner-occupied households, respectively. Similarly, from Table 5, the worker segments most likely to be employed in CBD, URBAN, SUBURBAN, and RURAL neighborhoods are workers who live in CBD neighborhood, workers who live in URBAN neighborhood, workers who use the auto mode for commute, and workers employed in the agriculture industry, respectively.

<<Table 4 Here>>
<<Table 5 Here>>

6. CONCLUSION

Latent choice modeling has served as a valuable modeling method for identifying population segments with significant behavioral heterogeneity, varying consideration choice sets, decision rule heterogeneity, and alternate dependency pathways among inter-dependent choices. However, studies that used latent choice methods in the context of location choices are relatively rare. This is primarily because of large choice sets in zonal-level destination choice models that make it unwieldy for estimating latent class models. This paper developed a latent class model that explicitly accounts for probabilistic nature of choice sets by using a two-stage modeling framework that assumes people first pick a neighborhood and then look for specific locations within the chosen neighborhood. The expected utility from the second stage zonal choice model component was used as an explanatory variable in the utility specification of neighborhood choice model to link the two models. The model was used to analyze residential and work location decisions in Nashville, Tennessee. The model results indicate significant heterogeneity in the consideration probability of different neighborhood alternatives both in the residential and work location choices. Also, the latent class neighborhood models were found to outperform standard MNL models that assume all decision makers consider the universal choice set in their decision making. The model applicability was demonstrated by calculating elasticity effects and identifying demographic groups with considerably different residential and work neighborhood preferences.

7. ACKNOWLEDGEMENT

This research was funded by Tennessee Department of Transportation (TDOT). The authors would like to thank TDOT for providing data and support during the research period. Also, computational facilities at University of Memphis are greatly acknowledged. Any opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the above-mentioned agencies.

8. REFERENCES

Arunotayanun, K., Polak, J.W., 2011. Taste heterogeneity and market segmentation in freight shippers' mode choice behaviour. Transportation Research Part E: Logistics and Transportation Review 47, 138–148. doi:10.1016/j.tre.2010.09.003

Baerenklau, K.A., 2010. A Latent Class Approach to Modeling Endogenous Spatial Sorting in Zonal Recreation Demand Models. Land Economics 86, 800–816. doi:10.3368/le.86.4.800

Bandeen-Roche, K., Xue, Q.-L., Ferrucci, L., Walston, J., Guralnik, J.M., Chaves, P., Zeger, S.L., Fried, L.P., 2006. Phenotype of Frailty: Characterization in the Women's Health and Aging Studies. J Gerontol A Biol Sci Med Sci 61, 262–266.

Ben-Akiva, M., Boccara, B., 1995. Consideration setsDiscrete choice models with latent choice sets. International Journal of Research in Marketing 12, 9–24. doi:10.1016/0167-8116(95)00002-J

Ben-Akiva, M.E., Lerman, S.R., 1985. Discrete choice analysis: theory and application to travel demand. MIT press.

Bhat, C.R., 1998. Accommodating variations in responsiveness to level-of-service measures in travel mode choice modeling. Transportation Research Part A: Policy and Practice 32, 495–507.

Bhat, C.R., 1997. An Endogenous Segmentation Mode Choice Model with an Application to Intercity Travel. Transportation Science 31, 34–48. doi:10.1287/trsc.31.1.34

Bhat, C.R., Guo, J.Y., 2007. A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. Transportation Research Part B: Methodological 41, 506–526. doi:10.1016/j.trb.2005.12.005

Bierlaire, M., Hurtubia, R., Flötteröd, G., 2009. A comparative analysis of implicit and explicit methods to model choice set generation. Presented at the 9th Swiss Transport Research Conference.

Boxall, P.C., Adamowicz, W.L., 2002. Understanding heterogeneous preferences in random utility models: a latent class approach. Environmental and resource economics 23, 421–446.

Boxall, P.C., Adamowicz, W.L., Boxall, P.C., n.d. Understanding Heterogeneous Preferences in Random Utility Models: A Latent Class Approach. Environmental and Resource Economics 23, 421–446. doi:10.1023/A:1021351721619

Bucholz, K.K., Heath, A.C., Reich, T., Hesselbrock, V.M., Krarner, J.R., Nurnberger, J.I., Schuckit, M.A., 1996. Can We Subtype Alcoholism? A Latent Class Analysis of Data from Relatives of Alcoholics in a Multicenter Family Study of Alcoholism. Alcoholism: Clinical and Experimental Research 20, 1462–1471. doi:10.1111/j.1530-0277.1996.tb01150.x

Calastri, C., Hess, S., Choudhury, C., Daly, A., Gabrielli, L., 2017. Mode choice with latent availability and consideration: Theory and a case study. Transportation Research Part B: Methodological. doi:10.1016/j.trb.2017.06.016

Campbell, D., Lorimer, V., Aravena, C., Hutchinson, G., 2010. Attribute processing in environmental choice analysis: implications for willingness to pay. Agricultural Economics 29, 31.

Cascetta, E., Papola, A., 2001. Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand. Transportation Research Part C: Emerging Technologies 9, 249–263. doi:10.1016/S0968-090X(00)00036-X

Chiang, J., Chib, S., Narasimhan, C., 1998. Markov chain Monte Carlo and models of consideration set and parameter heterogeneity. Journal of Econometrics 89, 223–248. doi:10.1016/S0304-4076(98)00062-1

Dillon, W.R., Böckenholt, U., De Borrero, M.S., Bozdogan, H., De Sarbo, W., Gupta, S., Kamakura, W., Kumar, A., Ramaswamy, B., Zenor, M., 1994. Issues in the Estimation and Application of Latent Structure Models of Choice. Marketing Letters 5, 323–334.

Gilbride, T.J., Allenby, G.M., 2004. A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules. Marketing Science 23, 391–406. doi:10.1287/mksc.1030.0032

Greene, W.H., Hensher, D.A., 2003. A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research Part B: Methodological 37, 681–698.

Grover, R., Srinivasan, V., 1987. A Simultaneous Approach to Market Segmentation and Market Structuring. Journal of Marketing Research 24, 139–153. doi:10.2307/3151504

Hauser, J.R., 2014. Consideration-set heuristics. Journal of Business Research 67, 1688–1699.

Hensher, D.A., 2010. Attribute processing, heuristics and preference construction in choice analysis. State-of Art and State-of Practice in Choice Modelling, Emerald Press, UK 35–70.

Hensher, D.A., Greene, W.H., 2009. Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification. Empir Econ 39, 413–426. doi:10.1007/s00181-009-0310-x

Hensher, D.A., Rose, J.M., Greene, W.H., 2011. Inferring attribute non-attendance from stated choice data: implications for willingness to pay estimates and a warning for stated choice experiment design. Transportation 39, 235–245. doi:10.1007/s11116-011-9347-8

Hess, S., Rose, J.M., 2007. A latent class approach to recognising respondents' information processing strategies in SP studies.

Hess, S., Stathopoulos, A., 2013. A mixed random utility — Random regret model linking the choice of decision rule to latent character traits. Journal of Choice Modelling, Issues in Choice Modelling: selected papers from the 13th International Conference on Travel Behaviour Research 9, 27–38. doi:10.1016/j.jocm.2013.12.005

Hess, S., Stathopoulos, A., Campbell, D., O'Neill, V., Caussade, S., O'Neill, V., Campbell, D., Hess, S., Stathopoulos, A., 2012. It's not that I don't care, I just don't care very much: confounding between attribute non-attendance and taste heterogeneity. Transportation 40, 583–607. doi:10.1007/s11116-012-9438-1

Hynes, S., Hanley, N., Scarpa, R., 2008. Effects on Welfare Measures of Alternative Means of Accounting for Preference Heterogeneity in Recreational Demand Models. Am. J. Agr. Econ. 90, 1011–1027. doi:10.1111/j.1467-8276.2008.01148.x

Jung, T., Wickrama, K. a. S., 2008. An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling. Social and Personality Psychology Compass 2, 302–317. doi:10.1111/j.1751-9004.2007.00054.x

Lanza, S.T., Rhoades, B.L., 2011. Latent Class Analysis: An Alternative Perspective on Subgroup Analysis in Prevention and Treatment. Prev Sci 14, 157–168. doi:10.1007/s11121-011-0201-1

Manski, C.F., 1977. The structure of random utility models. Theor Decis 8, 229–254. doi:10.1007/BF00133443

Martínez, F., Aguila, F., Hurtubia, R., 2009. The constrained multinomial logit: A semi-compensatory choice model. Transportation Research Part B: Methodological 43, 365–377. doi:10.1016/j.trb.2008.06.006

McFadden, D., 1978. Modeling The Choice of Residential Location. Transportation Research Record.

Mitani, Y., Shoji, Y., Kuriyama, K., 2008. Estimating economic values of vegetation restoration with choice experiments: a case study of an endangered species in Lake Kasumigaura, Japan. Landscape Ecol Eng 4, 103–113. doi:10.1007/s11355-008-0049-0

Paleti, R., 2015. Implicit choice set generation in discrete choice models: Application to household auto ownership decisions. Transportation Research Part B: Methodological 80, 132–149. doi:10.1016/j.trb.2015.06.015

Russell, G.J., Kamakura, W.A., 1993. Measuring brand value with scanner data. International journal of research in marketing 10, 9–22.

Scarpa, R., Thiene, M., 2005. Destination Choice Models for Rock Climbing in the Northeastern Alps: A Latent-Class Approach Based on Intensity of Preferences. Land Economics 81, 426–444. doi:10.3368/le.81.3.426

Srinivasan, K., Naidu, G., Sutrala, T., 2009. Heterogeneous Decision Rule Model of Mode Choice Incorporating Utility Maximization and Disutility Minimization. Transportation Research Record: Journal of the Transportation Research Board 2132, 59–68. doi:10.3141/2132-07

Swait, J., 2001a. A non-compensatory choice model incorporating attribute cutoffs. Transportation Research Part B: Methodological 35, 903–928. doi:10.1016/S0191-2615(00)00030-8

Swait, J., 2001b. Choice set generation within the generalized extreme value family of discrete choice models. Transportation Research Part B: Methodological 35, 643–666. doi:10.1016/S0191-2615(00)00029-1

Swait, J., 1994. A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data. Journal of Retailing and Consumer Services 1, 77–89. doi:10.1016/0969-6989(94)90002-7

Swait, J., Ben-Akiva, M., 1987. Incorporating random constraints in discrete models of choice set generation. Transportation Research Part B: Methodological 21, 91–102. doi:10.1016/0191-2615(87)90009-9

Swait, J., Sweeney, J.C., 2000. Perceived value and its impact on choice behavior in a retail setting. Journal of Retailing and Consumer Services 7, 77–88.

Waddell, P., Bhat, C., Eluru, N., Wang, L., Pendyala, R., 2007. Modeling interdependence in household residence and workplace choices. Transportation Research Record: Journal of the Transportation Research Board.

Walker, J.L., Li, J., 2006. Latent lifestyle preferences and household location decisions. J Geograph Syst 9, 77–101. doi:10.1007/s10109-006-0030-0

Wen, C.-H., Lai, S.-C., 2010. Latent class models of international air carrier choice. Transportation Research Part E: Logistics and Transportation Review, Selected Papers from the 12th ATRS Conference 46, 211–221. doi:10.1016/j.tre.2009.08.004

Wen, C.-H., Wang, W.-C., Fu, C., 2012. Latent class nested logit model for analyzing high-speed rail access mode choice. Transportation Research Part E: Logistics and Transportation Review 48, 545–554.

Zhang, J., Kuwano, M., Lee, B., Fujiwara, A., 2009. Modeling household discrete choice behavior incorporating heterogeneous group decision-making mechanisms. Transportation Research Part B:

Methodological, Modeling Household Activity Travel Behavior 43, 230–250. doi:10.1016/j.trb.2008.05.002

Table 1. Residential neighborhood choice model results.

Variables Description	**Urban**		**Sub-Urban**		**Rural**	
(Base Alternative: CBD)	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
Constant	3.914	7.275	2.375	4.146	1.191	1.843
Children per household aged 6 to 10 years			0.804	2.261	1.138	2.721
Children per household aged 11 to 15 years					0.472	1.785
Jobs per household			-0.167	-2.568		
Number of females per household	0.821	2.918	0.821	2.918	0.821	2.918
Number of licensed drivers per household			0.562	4.497	1.158	4.122
Residence type: single-family detached	6.791	29.560	7.888	27.126	7.888	27.126
Auto sufficiency: Zero vehicle	-1.444	-3.065	-2.334	-4.193	-1.883	-2.339
Auto sufficiency: High			0.707	3.707	2.206	4.502
Household income: More than 75k	-0.886	-2.226	-0.672	-1.624	-1.212	-2.554
Highest Education Attainment in household: Bachelor Degree	-1.182	-2.239	-1.644	-3.005	-1.990	-3.341
Highest Education Attainment in household: Graduate Degree	-1.743	-3.300	-2.352	-4.267	-3.348	-5.225
Latent Choice Set Component						
Constant			-0.615	-2.007	-1.6117	-3.757
Housing Tenure: Own			0.7991	3.654	0.5891	1.877
Number of bikes owned per household			0.311	1.819	-0.297	-1.589
Average age per household					0.018	3.164
Number of Observations	4344					
Number of Parameters Estimated	14					
Mean log-likelihood at convergence	-0.806					
Log-likelihood	-3502.072					
Chi-Square	19.296					
Critical Chi-Square $(df = 3, \alpha = 0.05)$	7.815					

Table 2. Work neighborhood choice model results.

Variables Description	**CBD**		**Urban**		**Sub-Urban**	
(Base Alternative: Rural)	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
Constant	-0.969	-2.155	0.141	0.349	0.668	3.471
Worker has disability (Yes = 1, No = 0)			0.815	2.616	0.868	2.241
Worker uses auto to travel to work (Yes = 1, No = 0)	-2.348	-8.960	-1.148	-4.653		
Worker has ability to change work schedule (Yes = 1, No = 0)	2.493	7.333	1.948	5.860		
Number of working days:5 days per week (Yes = 1, No = 0)	0.985	6.187	0.403	3.257		
Industry type of the worker: Agriculture	-2.708	-4.574	-1.758	-4.293	-1.225	-3.368
Industry type of the worker: Manufacturing	-2.639	-3.736	-0.599	-1.603	-0.599	-1.603
Industry type of the worker: Transportation	-1.915	-4.046	-0.658	-1.973		
Industry type of the worker: Retail	-0.966	-4.529				
Education Attainment of the Worker: Grade 12 or High school graduate					0.423	2.758
Education Attainment of the Worker: College credit or associate or technical school degree	-0.421	-3.433				
Residential neighborhood choice: CBD	1.447	3.720				
Residential neighborhood choice: Urban	2.903	8.995	2.781	10.332	1.504	5.918
Residential neighborhood choice: Sub-Urban	1.925	5.813	1.727	6.326	1.504	5.778
Latent Choice Set Component						
Constant			-0.320	-1.597	-0.320	-1.597
Worker has ability to change work schedule (Yes = 1, No = 0)			5.518	9.267	5.518	9.267
Industry type of the worker: Retail			-0.530	-2.905	-0.530	-2.905
Number of Observations						3992
Number of Parameters Estimated						16
Log-composite likelihood at convergence						-0.880
Log-likelihood						-3511.227
Chi-Square						41.094
Critical Chi-Square $(df = 3, \alpha = 0.05)$						7.815

Table 3. Zonal residential and work location choice components.

Variables Description	Coefficient	t-stat
Residential Location		
Size Variable: LN(Total number of households in TAZ)	1.0000	-
Correction Factor (Fixed)	1.0000	-
TAZ Attributes		
Accessibility in TAZ: Manufacturing sector	-0.4325	-4.25
Socio-economic Attributes (interacting with **Total Accessibility**)		
Household residence ownership: Owned (Yes=1 or No = 0)	-0.6498	-5.52
Presence of children in household (Yes=1 or No = 0)	-0.1869	-1.54
Presence of senior adults in household (Yes=1 or No = 0)	-0.4830	-4.21
Presence of disabled person in household (Yes=1 or No = 0)	0.3204	1.83
Household vehicle-ownership: Zero Vehicles (Yes=1 or No = 0)	0.8543	3.41
Household vehicle -ownership: One Vehicles (Yes=1 or No = 0)	0.5021	4.33
Household vehicle -ownership: Two Vehicles (Yes=1 or No = 0)	0.1488	1.32
Household income: $50K- $75K (Yes=1 or No = 0)	0.3508	2.63
Household income: More than $75K (Yes=1 or No = 0)	0.4644	3.35
Highest education attainment in household: Bachelor degree (Yes=1 or No = 0)	0.4642	4.02
Highest education attainment in household: Graduate degree (Yes=1 or No = 0)	0.7185	5.48
Mean Log-likelihood at convergence	-16,963.32	
Work Location	**Coefficient**	**t-stat**
Size Variable: Total number of employment in the industry of individual's employment	1.0000	-
Correction Factor	1.0000	-
Commuting Factors		
Commute distance	-3.4292	-3.21
Commute time by Auto during AM peak	-6.4880	-14.57
Presence of at least one type of transit (Yes=1 or No = 0)	8.2346	20.23
Socio-economic Attributes (interacting with **Distance**)		
Gender (Female=1, Male=0)	-4.6012	-4.50
Household vehicle -ownership: Three Vehicles (Yes=1 or No = 0)	2.6157	2.34
Presence of children in household (Yes=1 or No = 0)	-3.7878	-2.10
Education attainment of worker: Bachelor degree or higher (Yes=1 or No = 0)	-3.6382	-3.40
Socio-economic Attributes (interacting with **Travel Time by Auto**)		
Presence of children in household (Yes=1 or No = 0)	-1.0959	-2.47
Varying work location (Yes=1 or No = 0)	-1.3374	-3.94
Works more than 40 hours per week (Yes=1 or No = 0)	-1.2144	-3.87
Residential Location (interacting with **Travel Time by Auto**)		
Neighborhood type: Suburban (Yes=1 or No = 0)	-0.7219	-2.32
Log-likelihood at convergence	-7,110.875	

Table 4. Elasticity effects of residential neighborhood choice model.

Variables	**CBD**	**Urban**	**Sub-Urban**	**Rural**
Children per household aged 6 to 10 years	-9.386	-16.177	28.050	36.966
Children per household aged 11 to 15 years	-0.675	-1.578	-4.437	26.163
Jobs per household (Increased by 1)	1.287	2.814	-8.064	4.648
Number of females per household (Increased	-56.540	0.795	0.179	0.084
Number of licensed drivers per household (Increased by 1)	-6.941	-12.509	14.512	46.233
Residence type: single-family detached house	-99.902	-17.018	74.245	39.017
Auto sufficiency: Zero vehicle	275.740	13.128	-37.526	-1.705
Auto sufficiency: High	-10.659	-18.349	10.056	142.294
Household income: More than 75k	120.381	-4.104	13.820	-20.975
Highest Education Attainment in household: Bachelor Degree	192.059	8.578	-15.109	-25.359
Highest Education Attainment in household: Graduate Degree	395.559	12.147	-16.129	-51.502
Housing Tenure: Own	-6.063	-18.136	41.664	117.854
Number of bikes owned per household	-0.883	-2.565	9.662	-11.504
Average age per household (Increased by 1)	-0.049	-0.180	0.023	1.240

Table 5. Elasticity effects of work neighborhood choice model.

Variables	CBD	Urban	Sub-Urban	Rural
Worker has disability (Yes = 1, No = 0)	-48.804	9.355	12.719	-50.948
Worker uses auto to travel to work (Yes = 1, No = 0)	-62.647	13.062	129.132	110.544
Worker has ability to change work Schedule (Yes = 1, No = 0)	63.909	-3.432	-21.663	-9.171
Number of working days:5 days per week (Yes = 1, No = 0)	67.316	-1.891	-21.593	-19.604
Industry type of the worker: Agriculture	-59.743	-2.344	16.930	260.275
Industry type of the worker: Manufacturing	-83.631	13.855	2.962	78.995
Industry type of the worker: Transportation	-68.743	1.495	44.809	38.987
Industry type of the worker: Retail	-53.779	16.654	-14.135	-13.218
Education Attainment of the Worker: Grade 12 or High school graduate	-4.589	-5.002	26.778	-16.877
Education Attainment of the Worker: College credit or associate or technical school degree	-28.580	5.512	3.119	2.809
Residential neighborhood choice: CBD	165.156	-28.179	-17.894	-15.808
Residential neighborhood choice: Urban	33.856	22.064	-41.107	-85.956
Residential neighborhood choice: Sub-Urban	22.015	2.487	-1.465	-76.092

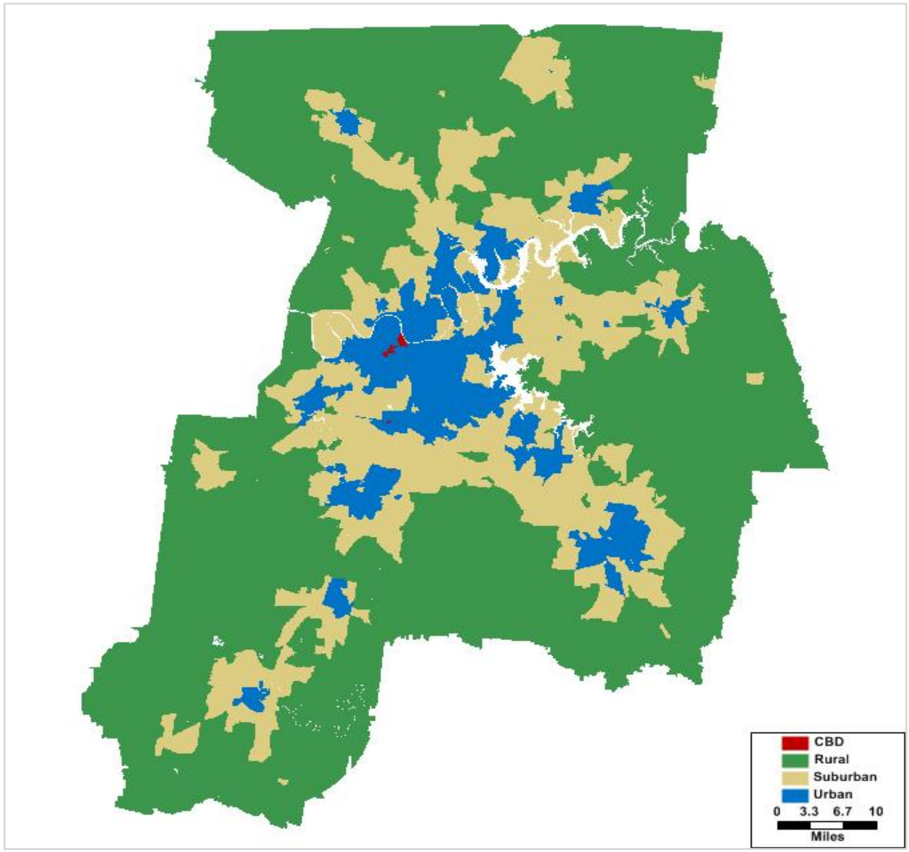

Figure 1. Neighborhood definition based on residential & employment density.