Journal of Transportation Technologies, 2023, 13, 243-261
https://www.scirp.org/journal/jtts
ISSN Online: 2160-0481
ISSN Print: 2160-0473

Accounting for Heterogeneity in Stop Frequency Models of Work Tours Using Latent Class Poisson Models

Babak Mirzazadeh

KLD Engineering, P.C. Islandia, NY, USA
Email: bmirzazadeh@kldcompanies.com

How to cite this paper: Mirzazadeh, B. (2023) Accounting for Heterogeneity in Stop Frequency Models of Work Tours Using Latent Class Poisson Models. *Journal of Transportation Technologies*, **13**, 243-261.
https://doi.org/10.4236/jtts.2023.132012

Received: January 30, 2023
Accepted: April 8, 2023
Published: April 11, 2023

Abstract

Stop frequency models, as one of the elements of activity based models, represent an important part of travel behavior. Unobserved heterogeneity across the travelers should be taken into consideration to prevent biasedness and inconsistency in the estimated parameters in the stop frequency models. Additionally, previous studies on the stop frequency have mostly been done in larger metropolitan areas and less attention has been paid to the areas with less population. This study addresses these gaps by using 2012 travel data from a medium sized U.S. urban area using the work tour for the case study. Stop in the work tour were classified into three groups of outbound leg, work based subtour, and inbound leg of the commutes. Latent Class Poisson Regression Models were used to analyze the data. The results indicate the presence of heterogeneity across the commuters. Using latent class models significantly improves the predictive power of the models compared to regular one class Poisson regression models. In contrast to one class Poisson models, gender becomes insignificant in predicting the number of tours when unobserved heterogeneity is accounted for. The commuters are associated with increased stops on their work based subtour when the employment density of service-related occupations increases in their work zone, but employment density of retail employment does not significantly contribute to the stop making likelihood of the commuters. Additionally, an increase in the number of work tours was associated with fewer stops on the inbound leg of the commute. The results of this study suggest the consideration of unobserved heterogeneity in the stop frequency models and help transportation agencies and policy makers make better inferences from such models.

Keywords

Activity Based Model, Work Tour, Stop Frequency, Latent Class, Poisson

Regression Model

1. Introduction

Stop frequency models, as one of the segments of an activity-based travel model, predict the frequency or occurrence of stops that take place within a tour [1]. Sequencing the trips and adding stops to a tour can improve the efficiency of a journey for travelers [2]. Previous studies on the stop models have mainly assumed homogenous characteristics across the observations, subsequently estimated fixed parameters associated with each variable, and have not accounted for possible unobserved heterogeneity across the travelers. Unobserved heterogeneity could arise from the omission of certain variables in the model that may not have been captured through the data collection of the study. As an example of unobserved heterogeneity, households with similar socioeconomic characteristics may have significantly different stop making behavior within their daily tours. For instance, the travelers with similar socioeconomic characteristics may decide to add a stop for coffee within their work tours, however, another group of travelers with the same socioeconomic characteristics may not be interested in making such a stop while driving to or from workplace. This unobserved heterogeneity which may exist across many datasets [3] is difficult to measure or is unknown. Generally, ignoring the impact of unobserved heterogeneity can lead to biased and inefficient estimates of parameters and subsequently incorrect inferences from the model parameters [4] [5].

The main objective of this study is to model stop frequency on different legs of a work tour in a medium-sized U.S. metropolitan area while accounting for the potential unobserved heterogeneity. In this regard, Latent Class Poisson Regression Model (LCPRM) that accounts for unobserved heterogeneity is used. The predictive performance of the LCPRM is compared with that of a regular Poisson model that is used in the stop frequency literature such as in [6]. The goal is to evaluate whether there is a significant reduction in the residuals of a LCPRM in comparison to a regular Poisson model. This reduction would indicate a significant improvement in the predictive power of the latent class approach compared to the regular Poisson model. Only vehicle-based trips are considered for this study since the data contained over 97% of vehicle trips.

The major contribution of this research is to demonstrate how accounting for unobserved heterogeneity improves the predictive power of stop frequency models. Additionally, most stop frequency models have been developed for large metropolitan areas which might have different stop making behaviors within tours in comparison to small/medium-sized urban areas. By using data from a small/medium-sized urban area, this study provides parameters that could be used for similarly sized areas. For the remainder of this paper, the next section provides a brief review of related literature. Then, the methodology, data, and results are discussed. Finally, conclusions are made and directions for future research are discussed.

2. Stop Frequency Literature

The literature on stop models can be divided into those that evaluate the factors contributing to stop making on the entire tour and the studies that investigate specific legs of a tour. For example, stop models were developed for the entire tour as tour complexity in [7] [8], stop frequency, or stop making models in [9]-[14] and in some other studies they were developed for different legs of a tour such as outbound legs, stops at the primary destination *i.e.* middle of the day stops, and inbound legs [15]-[20]. Previous studies have discussed the variables contributing to the likelihood of stop making within a tour [11] [14] [15] [16] [18] and some of them such as [21] have discussed how these variables are associated with stops with different purposes across a tour.

Some of the variables contributing to stop making on tours include personal and household demographic variables [9] [15] [16] [17] [22] [23] [24]; environmental or land use related variables [9] [10] [15] [22] [25] [26]; activity related variables such as arrival/departure time; activity duration [11] [16] [22]; transportation related attributes such as mode choice [10] [11] [12] [16] [22], and travel time or distance [10] [11] [12] [25] [26].

In a study in New York city by Chu, ordered probit models were applied to predict how different explanatory variables contribute to the stop making of commuters on their way to work from home known as the morning commute, during subtours consisting of trips from work to work known as midday travel, and on their way to home from work known as the evening commute [16]. The stop categories considered in this study include no stop, one stop, and two or more than two stops [16]. The explanatory variables used in the models include individuals and household-related, land use measures, transportation-related attributes, and work schedule attributes [16]. The results indicated that variables such as high-income level and an increase in employment density at work were associated with an increased likelihood of making more stops on the midday and evening legs of commutes and people above 40 were associated with a higher likelihood of making more stops on all the three legs of a commute [16].

In another study by Bhat and Singh on the data of an activity survey in the Boston Metropolitan area, an ordered response model was developed to estimate the stop propensity of commuters on the evening leg of commutes [17]. Several variables including socioeconomic ones, work schedule-related ones, and level of service measures such as travel time were used in the model [17]. The results of the model showed that commuters from single parent households, female and married commuters, and commuters who leave work before 4 p.m. were associated with a higher propensity of making stops on their way to home from work [17].

In another study conducted in New York City, a multivariate probit model structure, consisting of several simultaneous stop models, was applied to the data of the study area to capture the possible interdependencies of the stop making decisions among different legs of a work tour [18]. The results of the study revealed the existence of interdependency among the stops made across the different legs of the commute [18]. Additionally, variables such as age, gender, high-income level, and several work schedule-related variables appeared to contribute to the stop making of commuters on different legs of a work tour [18].

3. Methodology

In a Poisson distribution model, the dependent variable is assumed to follow a Poisson distribution with a λ parameter with a log-linear function which is defined using the explanatory variables in the model. The equations of the model and its λ parameter are [27]:

$$Prob\left(Y_i = y_i \mid x_i\right) = \frac{e^{-\lambda_i} * \lambda_i^{y_i}}{y_i!}, \; y_i = 0,1,2,\cdots \tag{1}$$

$$\lambda_i = e^{\beta * x_i} \tag{2}$$

where y_i is the frequency of activity for observation i, x is the vector of explanatory variables, β is the vector of coefficients, and the λ parameter is the mean rate of frequency for observation i. The Poisson model has an equi-dispersion assumption stating that the variance of the dependent variable y_i is equal to its mean. This property is written as follows [27]:

$$E\left[y_i \mid x_i\right] = Var\left[y_i \mid x_i\right] = \lambda_i = e^{\beta * x_i} \tag{3}$$

Unobserved heterogeneity in the data could cause the variance of the dependent variable to exceed its mean [28]. This leads to the violation of the equi-dispersion assumption and subsequently results in an invalid statistical test of the parameters in the Poisson regression model [28]. Therefore, to account for the unobserved heterogeneity as well as the counting nature of the data, LCPRM is used in this study to model the stop frequency on the work tours on the outbound leg, on the work based subtour as the middle of the day leg, and on the inbound leg of the commutes. In this approach, several discrete segments known as classes are considered in the model structure and a distinct vector of parameters is estimated for each class. In fact, homogenous subgroups in the data are identified in a latent class model, but due to its restrictive homogeneity assumption, variations within each homogenous subgroup are ignored [29]. In a latent class model, an individual's behavior depends on the observable attributes and on the latent heterogeneity which itself is dependent on unobservable factors [30].

In a latent class model, the probability that individual i chooses y_i on condition that class j is chosen is:

$$P\left(i \mid j\right) = Prob\left[Y_i = y_i \mid x_i, \beta_j\right], \; y_i = 0,1,2,\cdots, \; j = 1,\cdots,J \tag{4}$$

The probability of being in class j which sums up to 1 for all the classes is:

$$p_{ij} = \frac{e^{\theta_{ij}}}{\sum_{j=1}^{J} e^{\theta_{ij}}}, \; j = 1,\cdots,J, \;\; \theta_{iJ} = 0, \;\; \sum_{j=1}^{J} p_j = 1 \tag{5}$$

The log-likelihood function, L, in a latent class model which is maximized to obtain the βs and θs is:

$$\ln L = \sum_{i=1}^{N} \ln \sum_{j=1}^{J} p_{ij} * P(i \mid j) \tag{6}$$

Using Bayes theorem, the posterior estimate of the probabilities of membership in a specific class is calculated as:

$$P(j \mid i) = \frac{P(i, j)}{P(i)} = \frac{P(i \mid j) * p_{ij}}{\sum_{j=1}^{J} P(i \mid j) * p_{ij}} \tag{7}$$

Observation i is classified into the class with the highest posterior probability and the vector of the parameters of that class is used to calculate the predicted values and predicted probabilities for each observation. The predicted values representing the predicted number of stops are yielded via:

$$\lambda_{i|j} = \mathrm{e}^{\beta_j * x_i} \tag{8}$$

For stop making on a commute tour, travelers with the same observed characteristics may generate a different number of stops on each leg of a commute due to a latent heterogeneity that originates from unobserved factors varying across the commuters. To accommodate this heterogeneity, a latent class model allows the commuters to belong to distinct classes such that the commuters located in the same class follow the same choice behavior, namely, the same vector of parameters.

4. Data

The data used in this study to develop the models of stop frequency for outbound leg (OL), work based subtour leg (WBST), and inbound leg (IL) of the commutes were from a household survey conducted in 2012 for the Fargo-Moorhead metropolitan area, a small/medium-sized urban area with a population of about 225,830 in 2015 [31].

Work tours, which are the unit of analysis, are defined as activities with the primary purpose of work starting from and ending at home. If a traveler has a work or work-related activity during a tour, that tour was considered a commute or work tour. The number of activities or stops between home location and work location, when the home is the start of the trip, is considered the number of stops on the outbound leg of the commute. For example, in **Figure 1**, the orange arrows from home to work arriving at 8:00 a.m. represent this OL with two stops to drop off kids at school and to get coffee from a coffee shop. If the traveler leaves the work activity and then returns to a work or work-related activity, this is considered a work based subtour (WBST) represented by the blue arrows in **Figure 1**. The stops or activities within a work based subtour begin at 12:00 p.m. and have two stops. First the commuter goes for lunch at a restaurant and then picks up medications at a pharmacy before going back to work at 1:00 p.m. In fact, WBST represents the number of stops of all purposes between the first and

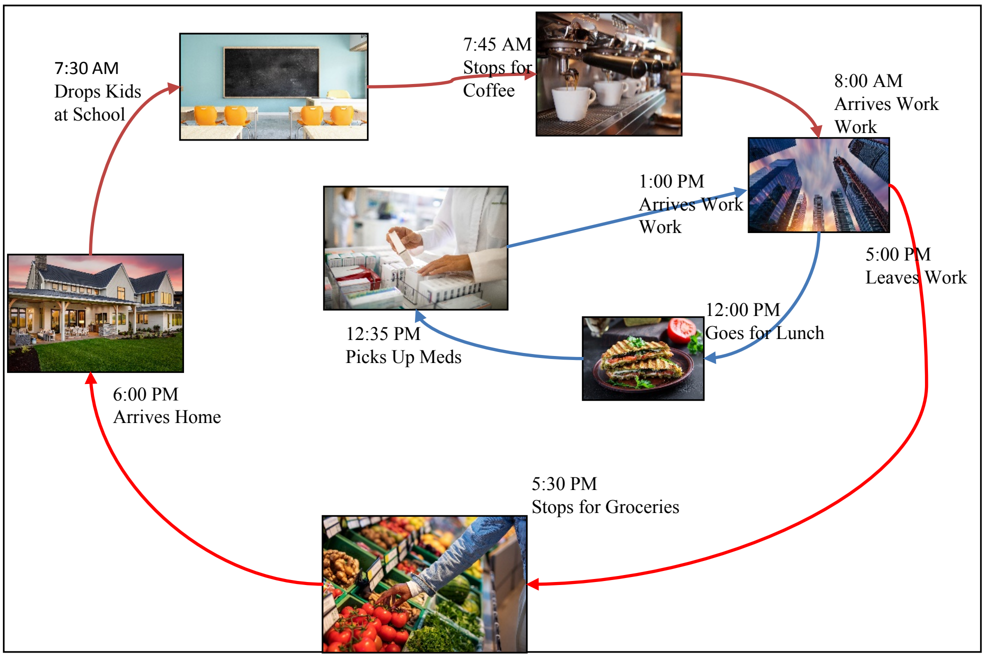

Figure 1. Example of a daily commute tour with OL, WBST, and IL.

last work activity. Once the commuter leaves work to home, all the activities or stops between work and home are considered as the number of stops on the inbound leg (IL) of the commute. As shown in **Figure 1**, this leg starts at 5:00 p.m., it has one stop at the grocery store, and ends with the commuter arriving back home at 6:00 p.m.

Although stops across a tour may have different purposes such as pick-up/drop-off, shopping, and recreation with various percentages as addressed in some of the previous studies [15] [16] [18] [32] [33] [34], in this study only the number of stops regardless of their purpose is considered.

The geographic unit of analysis is the Traffic Analysis Zones (TAZs) for the metropolitan area in this study. The study area in the Fargo-Moorhead area consists of 590 TAZs in which the travel survey was done in 2012. As the coordinates of the household locations and trip's origins and destinations were reported within a 500-feet tolerance in the survey, some of these data points were located in other surrounding TAZs instead of their actual TAZ if they were close to the boundary of a TAZ. Hence, to account for this inaccuracy, an optional 1-mile buffer was developed in ArcGIS around the centroid of each TAZ. Then, the attributes of all the TAZs that fully or partially fall into this buffer were summed up. These new attributes are assigned to the TAZ whose centroid was used to develop the buffer. Thus, the TAZs that are discussed and used in this study are modified TAZs whose new attributes were assigned to the home locations and to the trip's origins and destinations. The variables used in this study to predict the frequency of stops on each leg of the commutes are classified into four groups including personal, household, land use, and tour-related ones. **Table 1** displays the explanatory variables used in the modeling process.

Table 1. Model variables and their descriptions.

Variable Categories	Variable Description	Mean	Standard Deviation
Personal			
Male	If the traveler is male it takes value 1, otherwise 0, a binary variable	0.50	0.50
Age	Age of the traveler, a continuous variable	50.76	11.85
Household			
Income level between \$35K - \$100K, (MiddleInc.)	If the traveler's annual household income is more than or equal to \$35K and less than \$100K, a binary variable	0.60	0.48
Income level greater than \$100K, (HighInc.)	If the traveler's annual household income is more than or equal to \$100K, a binary variable	0.31	0.46
Number of cars in the household = 2, (Car2)	Two cars in the household, a binary variable	0.48	0.50
Number of cars in the household > 3, (Car ≥ 3)	More than two cars in the household, a binary variable	0.39	0.48
Number of driver's license holders in the household, (DLHH)	A continuous variable	2.11	0.69
Number of full-time workers in the household, (FULLHH)	A continuous variable	1.52	0.69
Number of part-time workers in the household, (PARTHH)	A continuous variable	0.33	0.56
Number of children under 18 in the household, (UND18HH)	A continuous variable	0.72	1.08
Land Use			
Retail jobs density in the home TAZ, (RJDH)	A continuous variable (employee per square mile)	292.40	297.11
Service jobs density in the home TAZ, (SJDH)	A continuous variable (employee per square mile)	750.67	786.38
Retail jobs density in the work TAZ, (RJDW)	A continuous variable (employee per square mile)	724.91	662.89
Service jobs density in the work TAZ, (SJDW)	A continuous variable (employee per square mile)	1438.45	1011.20
Tour			
Number of work tours, (NWT)	Number of work tours that a traveler makes during a day minus one, a binary variable indicating 1 or more than 1 work tours	0.27	0.44
Number of school tours, (NST)	Number of school tours that a traveler makes during a day, a continuous variable indicating 0, 1 or more than 1 school tours	0.02	0.17
Number of other tours, (NOT)	Number of other tours that are not work or school made by a traveler during a day, a continuous variable a continuous variable indicating 0, 1 or more than 1 other tours	0.36	0.57
Tour start time from home, (TS79AM)	A two-hour period during which the work tour starts. TS79AM means the tour starts between 7 and 9 a.m., a binary variable	0.54	0.49

Continued

Tour ending time, (TE57PM)	A two-hour period during which the work tour ends. TE57PM means the traveler arrives back at home between 5 and 7 p.m., a binary variable	0.36	0.48
Daily Activity Pattern, (DAP)	If travelers make at least one non-joint mandatory tour including, work, school and religious activities during a day and they do not participate in a joint tour or trip, a binary variable	0.80	0.39

5. Results

The results of the analyses are presented within the following three sections:

5.1) Number of Latent Classes

5.2) Comparison of the Residuals of Regular Poisson Model and LCPRM

5.3) LCPRM parameters and Marginal Effects

5.1. Number of Latent Classes

Information criteria such as Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) can be used to decide on the appropriate number of classes in a latent class model [35]. Usually, lower values of AIC and BIC of a model indicate that model is a better fit for the chosen number of latent classes [36]. **Table 2** shows the AIC, BIC, and the log-likelihood of the models with a different number of classes including the model without latent class application which is the model with one class.

For all the models, the BICs of the two-class models have a lower value compared to those of the models with the classes of other numbers. Notably, AIC used in this study is the corrected value of AIC which increases the penalty for the number of extra parameters in the model [36]. For the OL and WBST models, although AICs of the three-class models are lower than those of the models with the classes of other numbers, the entire estimated parameters in some of the classes are insignificant for three-class models. Hence, the number of 2 latent classes is chosen for all the models of this study.

5.2. Comparison of the Residuals of Regular Poisson Model and LCPRM

To evaluate the predictive performance of the LCPRM model in comparison to a regular Poisson model, the residuals for the model with 1 class (Poisson model) and those of the model with two classes (the LCPRM model) are compared together. The residuals are calculated as the absolute value of the difference between the predicted values of the number of stops using Equation (8) and the observed number of stops in the data for each data point. The goal is to demonstrate if there is a significant reduction in the residuals for the two-class models in comparison to the one-class models. This reduction would indicate a significant improvement in the predictive power of the latent class approach compared to the regular Poisson model. A one-sided statistical test is conducted to ascertain

Table 2. Comparison of results between different latent classes.

Models	OL			WBST			IL		
No. of classes \ Criteria	AIC	BIC	Log-likelihood	AIC	BIC	Log-likelihood	AIC	BIC	Log-likelihood
1	894.3	970.2	−430.16	1534.7	1574.9	−758.3	1497.4	1586.7	−728.6
2	789.3	**945.6**	−359.66	1016.2	**1101.1**	−489.1	**1255.20**	**1438.30**	−586.60
3	**767.2**	1003.9	−330.57	**1003.1**	1132.6	−472.5	1272.3	1549.20	−574.12
4	804.8	1121.9	−331.41	1016.4	1190.6	−469.22	1301.1	1671.7	−567.52

whether a significant reduction in the sample mean has occurred [37]. The equations from [37] that are applied for the test in this study are shown in Equations (9) and (10):

$$Z_d = \frac{\bar{x}_1 - \bar{x}_2}{S_y} \tag{9}$$

$$S_y = \sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}} \tag{10}$$

where, $\bar{x}_1$ is the average of residuals for the one-class model, $\bar{x}_2$ is the average of residuals for the two-class model, S_y is the pooled standard deviation of the residuals differences, Z_d is the standard normal distribution equivalent to the observed difference in the residuals, S_1 and S_2 are respectively the standard deviations for the residuals of the one-class and two-class models, and N_1 and N_2 are respectively sample sizes of the one-class and two-class models that are equal in this study. **Table 3** shows the results of the statistical test on the residuals of the models. From the table of standard normal distribution, the Z value that is equal to 95% confidence interval is 1.645. If the $Z_d > Z$ is held, it shows that the reduction in the average of residuals after using the two-class model is significant. The results in **Table 3** show that $Z_d > Z$ for all three models in this study indicating that LCPRMs significantly improves the predictive power of the stop frequency models.

5.3. LCPRM Parameters and Marginal Effects

The results of the models are shown in **Table 4**. There are two sets of coefficients for each variable for each model. Some variables have opposite signs such as the number of part-time workers in the IL model. Some variables have a different significance level across the classes such as SJDW variable in the WBST model. These examples are the indications of heterogeneity. Due to this heterogeneity, marginal effects are used to interpret how the variables contribute to making stops on each leg of the commute.

Marginal effects, ME, show how much a unit of increase in a variable increases or decreases the average number of stops, and it is calculated at the means of the Xs. The marginal effects are estimated using Equation (11) shown below:

Table 3. Comparison of the Poisson model with the LCPRM.

Models	OL		WBST		IL	
No. of Class / Statistics	One-class model	Two-class model	One-class model	Two-class model	One-class model	Two-class model
$\bar{x}$	0.357	0.267	0.635	0.277	0.634	0.296
S	0.510	0.397	0.825	0.477	0.622	0.531
N	643	643	643	643	643	643
S_y	0.025		0.037		0.032	
Z_d	3.60		9.67		10.56	

Table 4. Coefficients of the latent class Poisson regression models.

Variables	**OL**		**WBST**		**IL**	
Class	**1**	**2**	**1**	**2**	**1**	**2**
Constant term	−2.2530	0.4738	**−4.1900*****	−0.5710	−0.6270	−0.0120
Personal						
Male	−0.1990	−0.5199	−0.3074	**0.4964****	−0.3770	**−0.5808*****
Age	0.0138	−0.0052	**0.0344****	0.0159*	0.0136	0.0033
Household						
MiddleInc.	1.1180	−2.4918			−1.4268	0.4389
HighInc.	1.4590*	−2.7258	0.1843	−0.0546	1.4355	−0.1320
Car2			−0.7466	0.3440	−2.3989*	0.6186*
Car ≥ 3	−0.5550	0.1332	−0.0524	0.4258	**−4.1948****	0.5806
DLHH	**−1.0410****	0.6666			**−1.8109****	0.0548
FullHH	0.4500	−0.1797			2.3414*	−0.2007
PartHH	**1.1510*****	−0.2212			**3.4045****	**−0.4736****
Und18HH	**0.4430*****	0.6610*			0.5424	−0.0691
Land Use						
RJDH	0.0004	0.0027*			0.0020	−0.0009*
SJDH	−0.0002	−0.0009*	−0.0006*	**−0.0004****	−0.0013	0.0002
SJDW	0.0002	0.0009*	**0.0004****	0.0002	0.0002	0.3628×10^{-4}
RJDW	−0.0004	0.0002			0.0004	0.2482×10^{-4}
Tour						
TS79AM	0.4010	−1.3661*	0.5302	0.0351		
TE57PM					−0.4866	−0.2592
NWT	0.1455	−0.9547*			−1.4380*	**−0.6776*****
NST					−0.8007	0.3485
NOT					**−0.7984****	−0.1015
DAP	**−0.9587*****	−1.2758*			**−1.0234****	−0.3995*

Continued

Class Probability	**0.843*****	**0.157*****	**0.866*****	**0.134*****	**0.460*****	**0.540*****
Log likelihood function	−359.66		−489.11		−586.60	
McFadden Pseudo R-squared	0.163		0.355		0.194	
Observation	643		643		643	

***significance level p < 0.01; **significance level p < 0.05; *significance level p < 0.10.

$$\frac{\partial E\left[y_i \mid x_i\right]}{\partial x_i} = \lambda_i \beta \tag{11}$$

Table 5 shows the ME of the one-class Poisson regression model for each of the three models. These ME are shown to compare them with the ME of the LCPRMs to yield a better understanding of the differences between the two methods. The numbers in bold in the following tables represent the significance level of 1% and 5%. In the interpretation of the results, 1% and 5% levels are considered significant.

In latent class models, the marginal effect of a variable is the summation of the calculated marginal effect of each class weighted by the posterior latent class probabilities [38]. Table 6 shows the marginal effects for each of the three models:

Although the parameters of the second class of the OL model were not significant at the 5% level, the two-class model was kept as its results showed an improvement in the predictive power of the model in comparison to a regular or one-class Poisson model. In the next part, the results based on the two-class models, in Table 6, are discussed and compared with those of one class Poisson model, in Table 5.

6. Discussion

6.1. Personal Characteristics Variables

According to Table 6, being male was associated with fewer stops compared to being female, but the ME are insignificant. However, for the one class Poisson model, being male was significantly related to fewer stops for the IL model. Thus, applying the latent class approach reveals that gender does not significantly contribute to the stop making activity. Previous studies show that females were found to be associated with a higher likelihood of making complex commute tours than males in [11] [12]. They are expected to undertake the main maintenance responsibilities of a household [15] [16] [17] [18].

An increase in the age of commuters only contributes to a significant increase in the number of stops made on the work based subtours according to Table 5 and Table 6. Likewise, previous studies have shown that commuters above 40 were found to be significantly associated with an increase in the likelihood of taking part in the activities of the middle of the day leg of commutes [16] [18]. However, age is an insignificant variable in the OL and IL models of this study.

Table 5. Estimated marginal effects of the one-class Poisson regression model.

Variables	OL	WBST	IL
Personal			
Male	−0.0770*	0.0935*	**−0.2789*** **
Age	0.0018	**0.0073*** **	0.0016
Household			
MiddleInc.	0.0051		0.0129
HighInc.	0.1077	−0.0389	0.0243
Car2		0.0536	0.0461
Car ≥ 3	−0.0773*	0.1234	−0.0802
DLHH	**−0.1534*** **		−0.0482
FullHH	0.0656		0.0253
PartHH	**0.2010*** **		−0.0382
Und18HH	**0.1102*** **		−0.0118
Land Use			
RJDH	**0.0003*** **		−0.0002
SJDH	**−0.0001*** **	**−0.0001*** **	0.9791×10^{-5}
SJDW	**0.8484×10^{-4}*** **	**0.5984×10^{-4}** **	0.5832×10^{-6}
RJDW	-0.4830×10^{-4}		0.3936×10^{-4}
Tour			
TS79AM	−0.0256	**0.1661*** **	
TE57PM			**−0.1212** **
NWT	−0.0313		**−0.3316*** **
NST			−0.0330
NOT			**−0.1432** **
DAP	**−0.2755*** **		**−0.2779*** **

***significance level $p < 0.01$; **significance level $p < 0.05$; *significance level $p < 0.10$.

Table 6. Estimated marginal effects of the LCPRMs.

Variables	OL	WBST	IL
Personal			
Male	−0.0328	−0.0250	−0.0745*
Age	0.0014	**0.0040** **	0.0012
Household			
MiddleInc.	0.0727		−0.0641
HighInc.	0.1058	0.0191	0.0900
Car2		−0.0754	−0.1176
Car≥3	−0.0589	0.0014	−0.2471*
DLHH	**−0.1018** **		−0.1228*

Continued

FullHH	0.0463		0.1480
PartHH	**0.1231****		0.2004*
Und18HH	**0.0628*****		0.0324
Land Use			
RJDH	0.9692×10^{-4}		0.6230×10^{-4}
SJDH	-0.4013×10^{-4}	$\mathbf{-0.7100} \times 10^{-4}$**	-0.6921×10^{-4}
SJDW	$\mathbf{0.3613} \times 10^{-4}$**	$\mathbf{0.4500} \times 10^{-4}$**	0.1938×10^{-4}
RJDW	-0.4031×10^{-4}		0.2677×10^{-4}
Tour			
TS79AM	0.0163	0.0580	
TE57PM			−0.0557
NWT	−0.0035		**−0.1572*****
NST			−0.0275
NOT			**−0.0645****
DAP	**−0.1326*****		**−0.1050*****

***significance level $p < 0.01$; **significance level $p < 0.05$; *significance level $p < 0.10$.

6.2. Household Variables

Income variables were insignificant in stop making behavior in contrast to previous studies in large metropolitan areas [15] [16] [17] [18]. It is suggested that the size of the metro area and comparatively higher accessibilities due to a lower average travel time could explain this.

The number of cars in a household is not significantly associated with stop frequencies. The number of driver's license holders in a household is associated with a lower number of stops made by travelers. This result can be an indication of more independent trips and mode choices for household members on the outbound leg of work tours. For smaller metropolitan areas that are very car-centric, most households own a car, and most work trips are made using cars. It has been shown that compared to other regions in the U.S., the Midwest has the highest number of vehicles per household [39].

The number of full-time workers in a household is not significantly associated with stop frequencies. However, an increase in the number of part-time workers in a household is associated with a significant increase in the stops made by the commuters from that household on the outbound leg of commutes according to Table 6. This result is contrary to previous studies in which more number of employed adults [15] [16] or more adults in a household [12] [14] were associated with a lowered likelihood of stop making during a commute due to sharing the household responsibilities with other household members [12]. As an explanation for this difference, part-time workers in the Fargo-Moorhead area might perhaps be younger adults or teens who may not drive themselves and have to be dropped off at or picked up from their workplace.

An increase in the number of children under 18 years old in a household results in a significant increase in the number of stops on the outbound leg of the commutes according to **Table 6**. Parents of these children may be dropping them off at school on their outbound commute leg. This result is in line with that of previous studies like in [14] [16] [18] [40]. However, on the inbound leg of the commutes, this variable is insignificant suggesting that parents in families with young children want to go home from work to take care of them as shown in [16] [18].

6.3. Land Use Variables

On the outbound leg of the commute, an increase in the service job density around the work zone of the commuter is associated with an increased stop number. The service job density around the residence of the commuters is negatively associated with stop making on the outbound leg of the commutes, though insignificant. These results suggest that commuters are more prone to making stops related to service jobs around their workplace rather than their residence on the outbound leg of their commute. However, job density variables were insignificant for the IL model of this study. This shows that land use which is represented by retail job density and service job density does not significantly contribute to the stop making behavior of the commuters on their inbound leg. This perhaps indicates that individuals in small/mid-sized urban areas may be more likely to go home and then go shopping instead of combining the two trips on their way home. Generally, higher employment density whether at residence [16] [18] [41] or at work location [16] [18] was shown to be a significant factor in the stop making propensity of commuters in larger metropolitan areas. A comparison between the marginal effects of service job density at the work TAZ and that in the residence TAZ shows that commuters in the study area are more likely to make stops related to service jobs on their work based subtours around their workplace rather than around their residence. This makes sense as the destinations of commuters for the middle of the day tours are mainly closer to their workplace than to their residences are. In the one class Poisson model, the RJDH and SJDH variables were significant on the outbound leg of the commutes, however, they are no longer significant once the latent class approach is applied.

6.4. Tour Variables

For this study, tour starting time between 7 a.m. to 9 a.m. for OL and WBST models and tour ending time between 5 p.m. to 7 p.m. for the IL model appeared to be insignificant according to **Table 6**. This is an indication that the stops of different purposes do not necessarily occur or increase during these time intervals. This might stem from the nature of the occupation and smaller size of the area of study where travel time is not as important as that in larger metro areas. However, these variables are significant for the one-class Poisson model. Thus, TS79AM and TE57PM variables do not significantly contribute to stop making when the latent class approach is applied.

An increase in the number of work tours and tours of purposes other than work or school during a day is associated with fewer stops made by commuters on their inbound commute leg. This shows that people with multiple jobs, commuters who commute to their work several times a day, or people who make tours with different purposes other than work and school have a lower likelihood of adding stops to their commutes. Commuters may consolidate their stops to one of the work tours. Similar to some of the previous variables, the ME for these variables has a larger magnitude when the one-class approach is applied.

DAP variable shows if the commuters participate in at least one non-joint mandatory activity during a day and do not participate in joint activities such as carpooling, they are associated with fewer stops on the outbound and inbound legs of their commutes. Therefore, independent mandatory tours such as driving oneself to work during a day and not making joint tours or trips provide the ground for fewer stops on the outbound and the inbound legs of the commutes.

7. Conclusions

Past studies about stop frequency of commuters within work tours have mostly been conducted in larger metropolitan areas that have a different travel pattern compared to smaller metropolitan areas. Besides, accommodation of unobserved heterogeneity across the tours during which the stops take place has not gained enough attention in the literature. Not accounting for unobserved heterogeneity results in estimated biased and inefficient model parameters. This could lead to incorrect results and subsequently incorrect inferences from the model parameters. To account for the counting nature of the data and unobserved heterogeneity, Latent Class Poisson Regression Models were applied to the data of this study. Stop models were developed for different legs of commute tours including OL, WBST, and IL. Using AIC, BIC in addition to the significance level of the coefficients of the developed model parameters, two classes were chosen as the optimal number of latent classes for this study.

Each model had two vectors of parameters and the difference in the sign or the significance of the parameters of the same variable showed the existence of unobserved heterogeneity across the observations. Additionally, the statistical analysis indicated that using two-class models significantly resulted in a better prediction compared to the one-class models. This was an important result implying that using latent class models improved the accuracy of stop frequency models.

The results of the marginal effects revealed several outcomes that were different from previous studies in predicting tour frequencies. Gender and household incomes were insignificant in predicting stop frequencies and age was only significant in the work based subtour model. It is notable that contrary to the previous studies, income did not emerge to be a significant variable influencing the stop making decision of the commuters.

In this study, the employment density variables were separated based upon the two general occupations of retail and service. The results showed that employment density of service jobs at the workplace was positive and significant in predicting stops of OL and WBST models, whereas it was negative and significant in the WBST model at the home location. The results of the land use variables suggested that more employment density of service-related jobs at a zone is associated with more stops on the outbound leg of the commutes and associated with more stops on the work based subtours of the commuters in that zone. In this study, land use variables along with age appeared to be the significant contributors to stop making on the work based subtour of commutes. Contrary to this study, high income level was one of the contributors to stop making on work based subtour of commutes in other studies conducted in larger metropolitan areas. Among the tour related variables, an increase in the number of work tours, number of tours with purposes other than work or school, and being a commuter who participates in non-joint mandatory activities during a day with no joint activities were found to be associated with decreased stops on the inbound commute leg.

Generally, the results of this study showed improved results by accounting for unobserved heterogeneity. The size of the study area as a small/medium-sized U.S. metropolitan area has potentially had its impact on the results of the study. The results of this study could be used by policymakers and planners, especially for the towns of similar sizes that are developing activity-based models or need to have an evaluation on the impact of different variables in the stop making behavior of commuters in their region. The results broaden the understanding of the variables contribution to the stop making behavior within a work tour by taking the unobserved aspect of the commuters into account. For example, the results reveal that an increase in the number of work tours is associated with fewer stops on the inbound leg of a commute, however, the magnitude of this impact declines noticeably when the unobserved heterogeneity is taken into consideration. This study has been done for a small/medium-sized metropolitan area and can be developed for larger areas. Besides, for future studies, the development of stop prediction models using a latent class model can be applied to the stops of different purposes to realize whether unobserved heterogeneity exists for the stops of a certain purpose.

Acknowledgements

The author expresses his gratitude for the support from KLD Engineering, P.C. and from the Advanced Traffic Analysis Center (ATAC), a program of the Upper Great Plains Transportation Institute (UGPTI) at North Dakota State University.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Castiglione, J., Bradley, M. and Gliebe, J. (2015) Activity-Based Travel Demand Models: A Primer. Transportation Research Board, Washington DC. https://doi.org/10.17226/22357

[2] Schmöcker, J.-D., Su, F. and Noland, R.B. (2010) An Analysis of Trip Chaining among Older London Residents. *Transportation*, **37**, 105-123. https://doi.org/10.1007/s11116-009-9222-z

[3] Han, X. and Shao, Y. (2018) Genome-Wide Likelihood Ratio Tests under Heterogeneity. *Open Journal of Statistics*, **8**, 468-478.

[4] Shaheed, M.S. and Gkritza, K. (2014) A Latent Class Analysis of Single-Vehicle Motorcycle Crash Severity Outcomes. *Analytic Methods in Accident Research*, **2**, 30-38. https://doi.org/10.1016/j.amar.2014.03.002

[5] Mannering, F.L., Shankar, V. and Bhat, C.R. (2016) Unobserved Heterogeneity and the Statistical Analysis of Highway Accident Data. *Analytic Methods in Accident Research*, **11**, 1-16. https://doi.org/10.1016/j.amar.2016.04.001

[6] Shi, X. (2017) Tour Complexity, Variability and Pattern Using Longitudinal GPS Data. Ph.D. Dissertation, University of California Irvine, Irvine. https://escholarship.org/uc/item/9cn6g9dx

[7] Paleti, R., Pendyala, R.M., Bhat, C.R. and Konduri, K.C. (2011) A Joint Tour-Based Model of Tour Complexity, Passenger Accompaniment, Vehicle Type Choice, and Tour Length. https://repositories.lib.utexas.edu/bitstream/handle/2152/21085/Bhat_JointTourAttributes.pdf?sequence=2&isAllowed=y

[8] Yang, M., Wang, W., Chen, X., Wan, T. and Xu, R. (2007) Empirical Analysis of Commute Trip Chaining: Case Study of Shangyu, China. *Transportation Research Record*, **2038**, 139-147. https://doi.org/10.3141/2038-18

[9] Daisy, N.S., Liu, L. and Millward, H. (2018) Trip Chaining Propensity and Tour Mode Choice of Out-of-Home Workers: Evidence from a Mid-Sized Canadian City. *Transportation*, **47**, 763-792. https://doi.org/10.1007/s11116-018-9915-2

[10] Daisy, N.S., Millward, H. and Liu, L. (2018) Trip Chaining and Tour Mode Choice of Non-Workers Grouped by Daily Activity Patterns. *Journal of Transport Geography*, **69**, 150-162. https://doi.org/10.1016/j.jtrangeo.2018.04.016

[11] Xianyu, J.C. (2013) An Exploration of the Interdependencies between Trip Chaining Behavior and Travel Mode Choice. *Procedia-Social and Behavioral Sciences*, **96**, 1967-1975. https://doi.org/10.1016/j.sbspro.2013.08.222

[12] Liu, H. (2013) What Affects the Number of Non-Work Stops Made during Commute Tours? A Study Based on the 2009 US National Household Travel Survey in Large Metropolitan Areas. Master's Thesis, University of California Los Angeles, Los Angeles.

[13] Xian-Yu, J.-C., Juan, Z.-C., Gao, L.-J., Ni, A.-N., Zhang, W. and Wu, B. (2011) Empirical Analysis of Commuters' Nonwork Stop-Making Behavior in Beijing, China. *Journal of Transportation Engineering*, **137**, 360-369. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000228

[14] Wang, R. (2015) The Stops Made by Commuters: Evidence from the 2009 US National Household Travel Survey. *Journal of Transport Geography*, **47**, 109-118. https://doi.org/10.1016/j.jtrangeo.2014.11.005

[15] Chu, Y.-L. (2005) Modeling Workers' Daily Nonwork Activity Participation and Duration. *Transportation Research Record*, **1926**, 10-18.

https://doi.org/10.1177/0361198105192600102

[16] Chu, Y.-L. (2003) Empirical Analysis of Commute Stop-Making Behavior. *Transportation Research Record*, **1831**, 106-113. https://doi.org/10.3141/1831-12

[17] Bhat, C.R. and Singh, S.K. (2000) A Comprehensive Daily Activity-Travel Generation Model System for Workers. *Transportation Research Part A*: *Policy and Practice*, **34**, 1-22. https://doi.org/10.1016/S0965-8564(98)00037-8

[18] Chu, Y.-L. (2004) Daily Stop-Making Model for Workers. *Transportation Research Record*, **1894**, 37-45. https://doi.org/10.3141/1894-05

[19] Wu, Z.Z. and Ye, X. (2008) Joint Modeling Analysis of Trip-Chaining Behavior on Round-Trip Commute in the Context of Xiamen, China. *Transportation Research Record*, **2076**, 62-69. https://doi.org/10.3141/2076-07

[20] Luan, K., Juan, Z.C. and Tang, J. (2009) Empirical Analysis of Commuter Stop-Making Behavior Based on Ordered Porbit Model. *Proceedings of* 2009 *International Conference on Electronic Commerce and Business Intelligence*, Beijing, 6-7 June 2009, 423-426. https://doi.org/10.1109/ECBI.2009.69

[21] Garikapati, V.M. (2014) A Tour Level Stop Scheduling Framework and a Vehicle Type Choice Model System for Activity Based Travel Forecasting. https://keep.lib.asu.edu/_flysystem/fedora/c7/124407/Garikapati_asu_0010E_14444.pdf

[22] Chowdhury, T. and Scott, D.M. (2018) Role of the Built Environment on Trip Chaining Behavior: An Investigation of Workers and Non-Workers in Halifax, Nova Scotia. *Transportation*, **47**, 737-761. https://doi.org/10.1007/s11116-018-9914-3

[23] Schneider, F., Ton, D., Zomer, L., Daamen, W., Duives, D., Hoogendoorn-Lanser, S. and Hoogendoorn, S. (2021) Trip Chain Complexity: A Comparison among Latent Classes of Daily Mobility Patterns. *Transportation*, **48**, 953-975. https://doi.org/10.1007/s11116-018-9914-3

[24] Verma, A., Verma, M., Sarangi, P., Yadav, V. and Manoj, M. (2021) Activity Participation, Episode Duration and Stop-Making Behavior of Pilgrims in a Religious Event: An Exploratory Analysis. *Journal of Choice Modelling*, **38**, Article ID: 100267. https://doi.org/10.1016/j.jocm.2021.100267

[25] Zhu, P. and Guo, Y. (2022) Telecommuting and Trip Chaining: Pre-Pandemic Patterns and Implications for the Post-Pandemic World. *Transportation Research Part D*: *Transport and Environment*, **113**, Article ID: 103524. https://doi.org/10.1016/j.trd.2022.103524

[26] Chu, Y.-L. (2022) A Copula-Based Approach to Accommodate Intra-Household Interaction in Workers' Daily Maintenance Activity Stop Generation Modeling. *Journal of Traffic and Transportation Engineering* (*English Edition*), **9**, 59-68. https://doi.org/10.1016/j.jtte.2021.07.001

[27] Greene, W.H. (2002) Econometric Analysis. 5th Edition, Prentice Hall, Upper Saddle River, New Jersey.

[28] Jang, T.Y. (2005) Count Data Models for Trip Generation. *Journal of Transportation Engineering*, **131**, 444-450. https://doi.org/10.1061/(ASCE)0733-947X(2005)131:6(444)

[29] Mannering, F.L. and Bhat, C.R. (2014) Analytic Methods in Accident Research: Methodological Frontier and Future Directions. *Analytic Methods in Accident Research*, **1**, 1-22. https://doi.org/10.1016/j.amar.2013.09.001

[30] Greene, W.H. and Hensher, D.A. (2003) A Latent Class Model for Discrete Choice Analysis: Contrasts with Mixed Logit. *Transportation Research Part B*: *Methodo-*

logical, **37**, 681-698. https://doi.org/10.1016/S0191-2615(02)00046-2

[31] Ulteig and McKibben Demographics Research (2012) Demographic Forecast Study for the FM Metropolitan Area. http://www.fmmetrocog.org/application/files/5815/7591/1745/2012_Demographic_Forecast_Final_Report_Approved.pdf

[32] Currie, G. and Delbosc, A. (2011) Exploring the Trip Chaining Behaviour of Public Transport Users in Melbourne. *Transport Policy*, **18**, 204-210. https://doi.org/10.1016/j.tranpol.2010.08.003

[33] Hatcher, G.S. and Mahmassani, H.S. (1992) Daily Variability of Route and Trip Scheduling Decisions for the Evening Commute. *Transportation Research Record*, **1357**, 72-81.

[34] Kitamura, R. and Susilo, Y.O. (2006) Does a Grande Latte Really Stir Up Gridlock?: Stops in Commute Journeys and Incremental Travel. *Transportation Research Record*, **1985**, 198-206. https://doi.org/10.1177/0361198106198500122

[35] Porcu, M. and Giambona, F. (2017) Introduction to Latent Class Analysis with Applications. *The journal of Early Adolescence*, **37**, 129-158. https://doi.org/10.1177/0272431616648452

[36] Scarpa, R., Thiene, M. and Tempesta, T. (2007) Latent Class Count Models of Total Visitation Demand: Days Out Hiking in the Eastern Alps. *Environmental and Resource Economics*, **38**, 447-460. https://doi.org/10.1007/s10640-007-9087-6

[37] Roess, R.P., Prassas, E.S. and McShane, W.R. (2010) Traffic Engineering. 4th Edition, Pearson, New Jersey.

[38] Xie, Y., Zhao, K. and Huynh, N. (2012) Analysis of Driver Injury Severity in Rural Single-Vehicle Crashes. *Accident Analysis and Prevention*, **47**, 36-44. https://doi.org/10.1016/j.aap.2011.12.012

[39] Thakuriah, P. and Liao, Y. (2005) Analysis of Variations in Vehicle Ownership Expenditures. *Transportation Research Record*, **1926**, 1-9. https://doi.org/10.1177/0361198105192600101

[40] Cao, X., Mokhtarian, P.L. and Handy, S.L. (2008) Differentiating the Influence of Accessibility, Attitudes, and Demographics on Stop Participation and Frequency during the Evening Commute. *Environment and Planning B*: *Urban Analytics and City Science*, **35**, 431-442. https://doi.org/10.1068/b32056

[41] Cheng, L., Chen, X. and Yang, S. (2016) An Exploration of the Relationships between Socioeconomics, Land Use and Daily Trip Chain Pattern among Low-Income Residents. *Transportation Planning and Technology*, **39**, 358-369. https://doi.org/10.1080/03081060.2016.1160579