Advances in Applied Mathematics 应用数学进展, 2025, 14(8), 213-226
Published Online August 2025 in Hans. https://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2025.148384

Hans汉斯

基于嵌套Logit模型的居民出行行为研究

李朋昌，李　娜*，杨宇宁

内蒙古工业大学理学院，内蒙古　呼和浩特

收稿日期：2025年7月19日；录用日期：2025年8月12日；发布日期：2025年8月21日

摘　要

随着突发公共卫生事件进入常态化阶段，居民的出行行为模式发生了显著变化。为了深入探究个体属性、出行特征及家庭因素对居民出行选择行为的影响，本文基于效用最大化理论构建双层嵌套Logit模型(Nested Logit Model)，将出行方式作为上层决策变量，出行目的作为下层决策变量，对两者之间的联合选择行为进行建模。通过问卷调查获取全国范围内513份有效样本数据，并利用R软件开展相关性分析、主成分分析、多重共线性检验及模型参数标定。结果表明：性别和职业对出行方式具有显著影响，家庭中6岁以下儿童和60岁以上老年人口对出行目的具有显著影响，而是否拥有驾照及小汽车则同时对出行方式与目的产生显著影响。鉴于样本以城市青年为主，研究成果主要适用于与样本特征相似的群体，可为城市交通政策制定和差异化出行管理提供数据支持与理论依据。

关键词

巢式Logit模型，双层模型结构，居民出行选择行为，SP问卷调查

Research on Residents' Travel Behavior Based on Nested Logit Model

Pengchang Li, Na Li*, Yuning Yang

College of Science, Inner Mongolia University of Technology, Hohhot Inner Mongolia

Received: Jul. 19th, 2025; accepted: Aug. 12th, 2025; published: Aug. 21st, 2025

Abstract

Under the background of the normalization of the epidemic, the travel methods and purposes of residents have undergone significant changes. In order to deeply explore the influence of individual attributes, travel characteristics and family factors on residents' travel choice behavior, this paper constructs a two-layer nested Logit model based on the utility maximization theory, taking travel

*通讯作者。

文章引用: 李朋昌, 李娜, 杨宇宁. 基于嵌套 Logit 模型的居民出行行为研究[J]. 应用数学进展, 2025, 14(8): 213-226.
DOI: 10.12677/aam.2025.148384

mode as the upper-level decision variable and travel purpose as the lower-level decision variable. Model the joint choice behavior between the two. Through questionnaire surveys, 513 valid sample data from across the country were obtained, and correlation analysis, principal component analysis, multicollinearity test and model parameter calibration were carried out using R software. The results show that gender and occupation have a significant impact on the mode of travel. Children under 6 years old and the elderly population over 60 years old in the family have a significant impact on the purpose of travel. Meanwhile, whether one has a driver's license and a car has a significant impact on both the mode of travel and the purpose. Given that the sample is mainly composed of urban youth, the research results are mainly applicable to groups with similar characteristics to the sample, and can provide data support and theoretical basis for the formulation of urban traffic policies and differentiated travel management.

Keywords

Nested Logit Model, Double-Layer Model Structure, Residents' Travel Choice Behavior, SP Questionnaire Survey

Copyright © 2025 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

随着城市人口规模的增大，城市交通拥堵的问题也日益凸显，出行问题已经成为世界各国政府和普通群众最关注的焦点之一。研究居民的出行行为、挖掘居民出行时的行为特征成为了热门的研究话题。自 2020 年以来，人类社会经历了一场严重的全球性突发公共卫生事件——新型冠状病毒肺炎(COVID-19)，对中国乃至世界的经济与社会运行产生了深远影响。尽管相关防控政策已逐步优化调整，但该事件所带来的影响在短期内仍然存在，尤其在城市居民的日常生活中表现尤为明显。具体而言，居民在日常出行的方式、频率与目的等方面发生了显著变化。在突发公共卫生事件影响逐步常态化的背景下，系统探讨居民出行行为的变化特征，并识别其主要影响因素，对于理解城市交通行为机制、优化交通资源配置以及完善城市应急管理具有重要意义。

国内外学者基于该背景下的居民出行行为做出了一定的研究。国外学者 Koppelman [1]在 Logit 模型的基础上进行了改进，从而形成了 NL 模型，该模型可以将选择方案按照树状结构分层表示。由此 NL 模型被逐渐应用于交通领域的问题研究中来。Palma [2]运用嵌套 Logit 模型(NL 模型)对瑞士日内瓦市上下班出行方式的选择进行了实证分析，研究结果强调了除了模式特征(出行时间、成本和舒适度)之外，环境因素和个人因素的重要性。Bharat [3]等利用多项逻辑回归(Multinomial Logit, MNL)模型研究挪威市居民出行方式的选择行为，同时指出 MNL 模型在进行出行方式选择时可能存在的局限。Limtanakool [4]等对荷兰土地利用和运输系统的空间配置中长途旅行模式选择的影响进行了调查，发现土地利用和运输系统的空间配置在解释中长途出行方式的变化方面具有重要意义。Saleh [5]等对爱丁堡市一项关于拥堵收费对出行选择影响的程度进行调查，研究发现非工作活动以及工作时间表的灵活性都会影响上班旅程的出发时间选择。Ozbay [6]等建立出发/到达时间的经典时间分配模型对新泽西州收费公路分时定价方案的效率进行评价。国内学者同样将 NL 模型应用在交通领域中。叶倩文[7]从出行费用角度出发研究公共交通票价对出行方式选择的影响，应用 NL 模型从出行者个人属性和交通属性两方面挖掘了多个影响乘客出行选择的因素。杨励雅[8]基于随机效用最大化理论，分析了北京市居民出行样本数据，构建了一个双层

NL 模型，研究结果表明 NL 模型较传统的 MNL 模型具有更好的统计学特征。杨昌涛[9]将 NL 模型应用在交通方式选择行为研究中，并选择实例得到出行者对交通方式选择的概率。李佳玮[10]基于 SP 调查问卷获得的数据并运用 NL 模型从公共交通票价的角度出发，分析常规公共交通票价对出行者出行选择行为的影响。柳文燕[11]在已有研究的基础上，从收入水平角度出发，采用巢式 Logit 模型分析城市低收入人群的出行选择行为的影响因素，从而为特定人群出行方式的引导和管制提供参考依据。欧舟[12]构建 NL 模型研究居民出行方式选择行为和居民的日常出行活动类型，并发掘居民日常出行的内在规律。从上述文献可以看出，目前大多数的文献都将出行方式与出行目的分离开来考虑，忽视了两个因素之间存在的相互作用。肖光年等[13]基于贝叶斯网络构建出行方式识别模型，利用智能手机采集的 GPS 轨迹数据实现对五种方式的自动分类。通过引入低速点比例与平均方向改变等特征，有效提升了模型的识别准确率。本文将出行方式与出行目的利用 NL 模型联系起来，研究对居民日常出行产生影响的因素，从而为居民出行方式管制和引导提供参考依据。

2. 问卷调查与数据筛选

2.1. 居民出行影响因素的选取

“理性行为理论”是由美国学者阿耶兹(Ajzen)和菲什拜因(Fishbein)于 1975 年提出的，该理论认为个体的行为在某种程度上可以由行为意向合理地推断，而个体的行为意向又是由对行为的态度和主观准则决定的。人的行为意志是人们打算从事某一特定行为的量度，而态度是人们对从事某一目标行为所持有的正面或负面的情感，它是由对行为结果的主要信念以及对这种结果重要程度的估计所决定的。主观规范(主观准则)指的是人们认为对其有重要影响的人希望自己使用新系统的感知程度，是由个体对他人认为应该如何做的信任程度以及自己对与他人意见保持一致的动机水平所决定的。这些因素结合起来，便产生了行为意向(倾向)，最终导致了行为改变。综上所述，在病毒常态化背景下研究影响居民出行的因素，可以将所有因素分为两类，如图 1 所示。

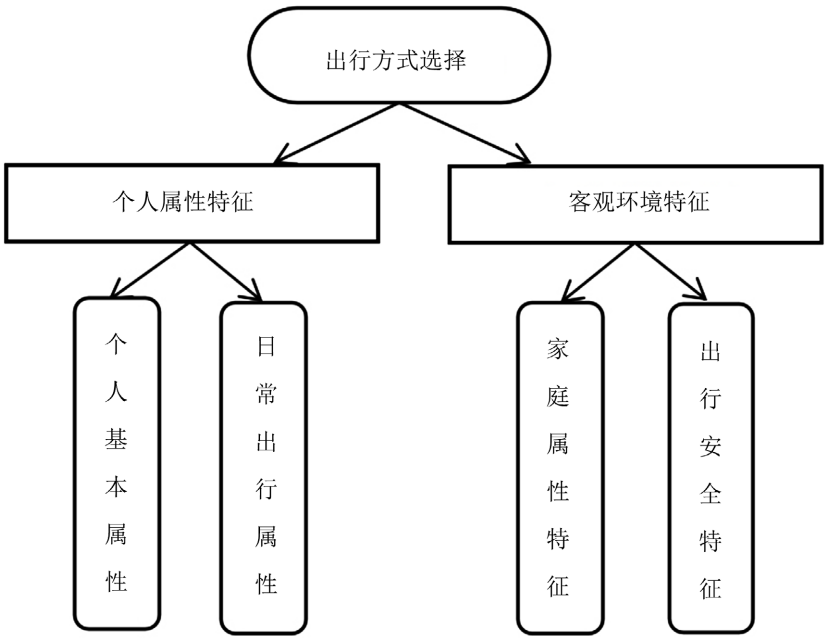

Figure 1. Analysis of residents’ travel choices under normal conditions
图 1. 常态化下居民出行选择分析

2.2. 居民出行数据收集与统计

本文通过问卷星平台发放病毒常态化下居民出行选择行为调查问卷，调查问卷的内容共涵盖个人基本属性、日常出行属性、家庭属性三个方面。为了得到更加普适性的意见和答案来契合本次的研究内容，本次问卷调查范围覆盖了全国 23 个省、市，获得的有效调查样本个数为 513 份。受访人群中，男性研究对象共有 264 人(所占比例为 51.46%)，女性研究对象一共有 249 人(所占比例为 48.54%)，研究对象的主要年龄阶段为 21~30 岁(所占的比例为 53.22%)。调查问卷中研究对象的受教育程度占比分别为专科(所占的比例为 39.38%)，本科(所占的比例为 31.97%)，高中及以下和研究生及以上占比较少(所占的比例为 28.65%)。在研究对象的职业中，企业职工的占比最高(所占的比例为 58.67%)，其次是学生(所占的比例为 16.18%)和个体经营户或私营企业主(所占的比例为 10.92%)。在月均收入方面，4001~6000 元占比最高(所占比例为 37.43%)，月均收入 8000 元以上的人数占比最少(所占比例为 2.14%)。原始样本特征可根据变量类型分为数值型变量和类别型变量两类，指标分类见下表 1。

Table 1. Classification of characteristic variable indicators
表 1. 特征变量指标分类

一级指标	二级指标	特征变量	变量类型
个人属性特征	个人信息	性别	类别变量
		年龄	有序类别变量
		受教育程度	有序类别变量
		职业	类别变量
		月收入水平	有序类别变量
	交通条件	是否有小汽车	类别变量
		是否有驾照	类别变量
		是否有公交卡	类别变量
客观环境特征	家庭情况	家庭共同居住人口数	数值变量
		家庭 60 岁以上人口数	数值变量
		家庭 6 岁以下孩子数	数值变量
	出行情况	出行时间	数值变量
		出行距离	数值变量
		出行目的	类别变量
		出行方式	类别变量
		出行费用	数值变量
		出行频率	数值变量
		弹性出行次数	数值变量
		非弹性出行次数	数值变量
	不同距离下的出行倾向	出行距离为 5 公里时采用的出行方式	类别变量
		出行距离为 10 公里时采用的出行方式	类别变量
		出行距离为 20 公里时采用的出行方式	类别变量

2.2.1. 居民出行行为描述性统计分析

1) 出行目的和出行频率分析

通过对病毒常态化下的居民出行选择行为数据中的出行目的和出行频率进行研究分析，可以获得该两个影响因素之间的堆叠柱状图，如图 2 所示。从图 2 中可以看出，当居民的出行目的是工作时，出行频率大部分以每周 4 天或每周 3 天为主。只有少部分居民的出行频率是在每周 2 天以下或 5 天以上，这表明大多数的工作人群每周平均需要工作 3 天或 4 天。当居民的出行目的是出门购物时，大部分居民会选择每周出行 3 天或 4 天，且出行 3 天或者 4 天的比例基本一致。当居民的出行目的是探亲访友或旅游时，大部分居民仅选择每周出行 2 天或者更少，这表明在病毒常态化阶段大家都对非必要出行持审慎的态度。

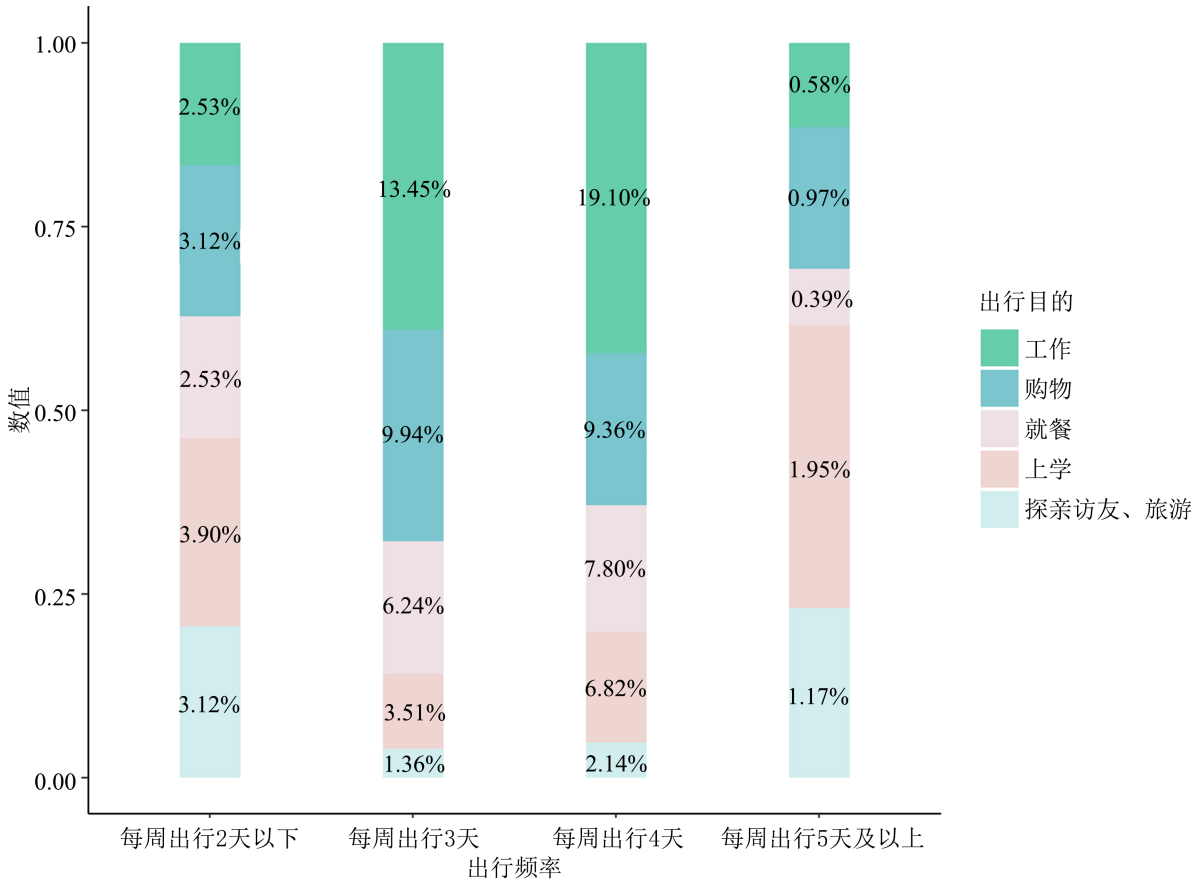

Figure 2. The relationship between travel purposes and travel frequencies
图 2. 出行目的与出行频率关系图

2) 出行时间和出行费用分析

利用 R 分析得出的出行时间和出行费用的堆叠柱状图如下图 3 所示。从图 3 中可以看出当出行时间位于 10~30 分钟区间且出行费用位于 6~10 元区间的占比为 27.88%，说明大多数居民更倾向于在时间在 10~30 分钟及费用在 6~10 元这个区间出行。当出行时间为 0~10 分钟时，居民的出行费用在 6~10 元或者 11~20 元居多，表明在出行距离较短时，居民可能更倾向于私人出行这种费用较高且更具安全性的出行方式。当居民需要出行的时间在一个小时或者更多的时候，大多数居民的出行费用还是以 6~10 元和 10~20 元为主，表明居民更愿意在出行时间较长时选择更加经济型的出行方式。

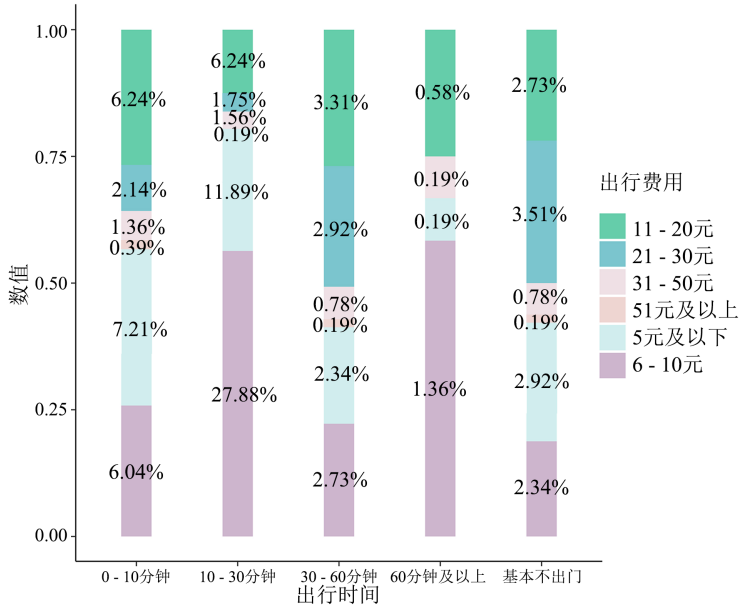

Figure 3. The relationship between travel time and travel cost
图 3. 出行时间和出行费用关系图

3) 出行方式和出行目的分析

出行方式共分为五类，出租车或网约车、地铁、自行车或步行、公共交通以及私家车。出行目的也共分为五类，分别是工作、购物、就餐、上学以及探亲访友、旅游。利用 R 语言得到的出行方式和出行目的的堆叠柱状图如下图 4 所示。从图 4 中可以看出，当居民选择以工作作为出行目的时，乘坐地铁或者公共交通的比例都较高，分别为 10.14%和 18.91%，表明大部分居民工作时更倾向于选择公共类的交通方式。而以探亲访友为出行目的的居民倾向于选择出租车或者网约车这种快速便捷的方式出行。

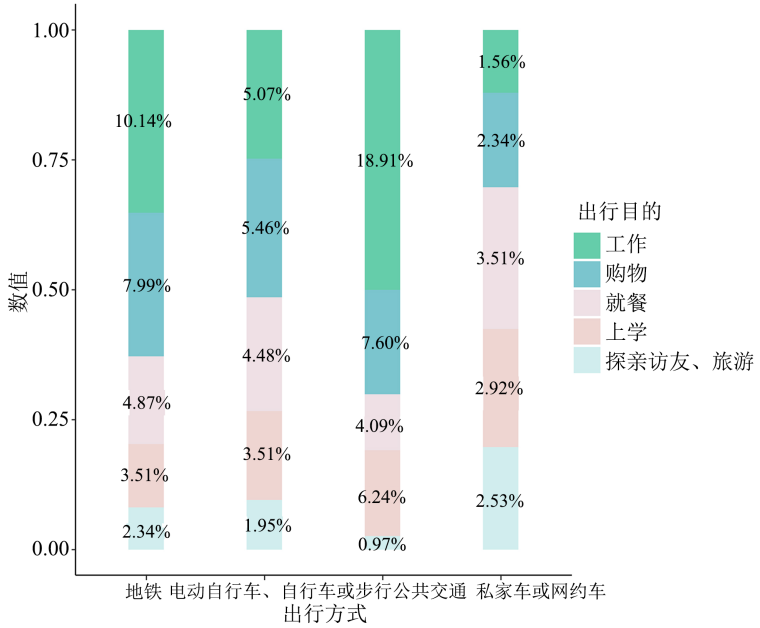

Figure 4. Diagram of the relationship between travel purposes and travel modes
图 4. 出行目的和出行方式关系图

2.2.2. 相关性分析

相关性分析指对两个或多个具备相关性的元素进行分析，从而得到各个元素之间的相关性密切程度。在居民出行行为选择影响因素的相关性分析中，采用皮尔森相关系数计算影响因素之间的相关性程度，见图 5。该图展现的是影响因素之间的皮尔森相关系数的热力图，皮尔森相关系数(也即 R)，根据 R 的大小，可以判断测试两个变量之间的相关关系。其中 0.8~1.0 极强相关；0.6~0.8 强相关；0.4~0.6 中等程度相关；0.2~0.4 弱相关；0.0~0.2 极弱相关或不相关。由图 5 可以看出，家中人口数量与家中 60 岁以上老人的数量和家中包含 6 岁以下孩子的数量分别具有 76%和 68%的相似度，这意味着总体来说，家里人口数量的增多代表着家里有更多的 60 岁以上老人或者 6 岁以下孩子。其次，出行的目的地和年龄之间也有着高于 40%的相似度，这意味着不同年龄段在病毒常态化下出行的目的地也有显著的区别。年龄段在 31~50 之间的调查样本出行目的更多集中在购物、探亲和旅游，而年龄段在 21~30 之间的调查样本其主要出行目的为工作。20 岁及以下的调查样本出行目的通常为上学，相比之下 51 岁以上的调查样本出行目的多为就餐。影响因素之间也有呈现负相关趋势的，调查样本显示年龄和工作职位以及受教育程度呈相反趋势，这代表在所调查的样本中，很多年龄段较高的人依然有着较低的职位和较低的受教育程度。除此之外，出行目的地与工作职位呈负相关趋势，但与收入成正相关，这意味着工作的职位和收入会显著影响出行的目的地，例如职位和收入的高低也会侧面影响调查样本出行的目的地。

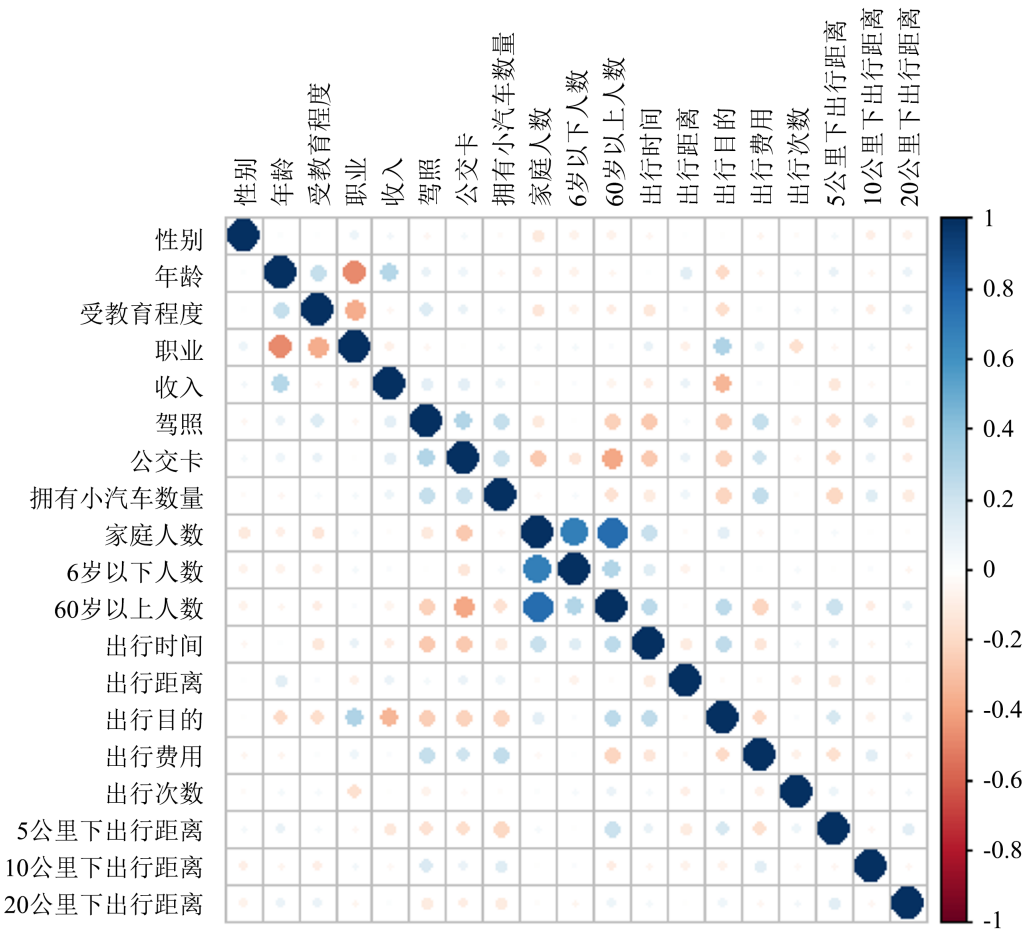

Figure 5. The heat map of the correlations among various influencing factors
图 5. 各影响因素之间的相关性热图

3. 巢式 Logit 模型

巢式 Logit 模型是由多项逻辑回归(MNL)模型演变得到的，是非集计模型中具有代表性的一种。巢式 Logit 模型克服了多项逻辑回归模型的不相关选择项独立性假设(IIA)，该假设可以提供一个合理的个体偏好决定或者个人选择行为，在 IIA 假设下，选择某个方案的概率不会因为选择个数的减少或者增加而受到影响，即任何两个选择概率的比值与其他选择无关。而巢式 Logit 模型存在分层次的巢式结构(如图 6 所示)，允许每个巢内的选择肢存在相关性，并考虑到了不同选择方案之间的相似性，这种特点使得巢式 Logit 模型可以有效克服 IIA 特性，因此巢式 Logit 模型也在出行领域中有着广泛的应用。

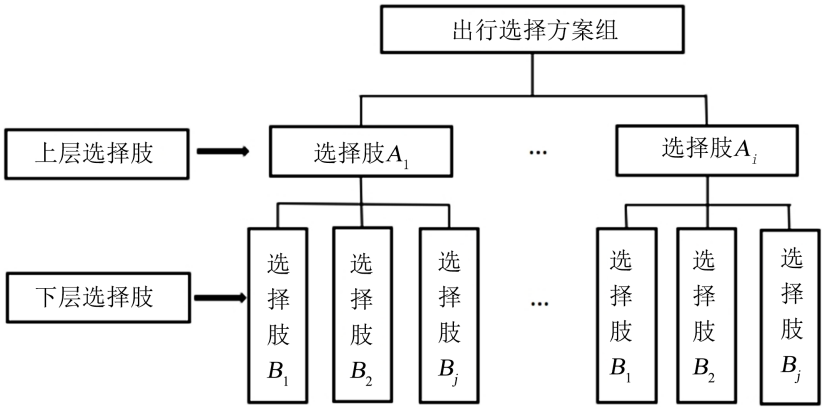

Figure 6. Conceptual diagram of the nested Logit model
图 6. 巢式 Logit 模型概念图

3.1. 巢式 Logit 模型原理

依照图 6 所示的巢式结构图形，上层选择肢共有 $A_I\left(A_1,A_2,\cdots,A_i\right)$ 个可选择类别，下层选择肢共有 $B_J\left(B_1,B_2,\cdots,B_j\right)$ 个可选择类别，将上下层选择肢可选择的类别组合在一起就一共有 $C_N=A_I\times B_J$ 种可选择的组合方案，分别表示为 $\left\{\left(A_1,B_1\right),\left(A_1,B_2\right),\cdots,\left(A_1,B_j\right),\cdots,\left(A_i,B_1\right),\left(A_i,B_2\right),\cdots,\left(A_i,B_j\right)\right\}$。记 U_{AB} 是出行者对上下层选择肢联合选择时的效用函数，令

$$U_{AB}=V_A+V_B+V_{AB}+\varepsilon_A+\varepsilon_B+\varepsilon_{AB} \tag{1}$$

其中 V_A 表示出行者选择上层方案 A 时可观测到的效用，该效用是仅随方案 A 变化的固定项。V_B 表示出行者选择上层方案 B 时可观测到的效用，该效用是仅随方案 B 变化的固定项。V_{AB} 表示出行者选择上层方案 A 和下层方案 B 时可观测到的效用，效用是随着 A 和 B 的组合而变化部分的固定项。ε_A，ε_B，ε_{AB} 分别表示出行者出行选择对应方案时的效用函数中的概率项。在一般的出行选择巢式 Logit 模型中 ε_B 的数量级都比较小可忽略不计，所以上式可简化为如下式子：

$$U_{AB}=V_A+V_B+V_{AB}+\varepsilon_A+\varepsilon_{AB} \tag{2}$$

巢式 Logit 模型是由一般 Logit 模型发展而来，所以具备一般 Logit 模型所遵循的理论原则，其中就包括效用最大化原则。效用最大化原则是出行者对若干出行方案的选择，在每一种出行方案边际效用相等时实现最大的总效用，即为效用最大化原则。则根据效用最大化理论可得到出行者选择上层选择肢 A 的概率为

$$\begin{aligned}P(A)&=P\left(\max U_{AB}\geq\max U_{A'B}\right)\\&=P\left(V_A+\varepsilon_A+\max\left(V_A+V_{AB}+\varepsilon_{AB}\right)\geq V_{A'}+\varepsilon_{A'}+\max\left(V_B+V_{A'B}+\varepsilon_{A'B}\right)\right),\forall A'\in A_I,A'\neq A\end{aligned} \tag{3}$$

假设出行者出行选择对应方案时的效用函数中的概率项$\varepsilon_A,\varepsilon_{AB}$相互独立，且$\varepsilon_A \sim G\left(0,u^A\right)$，$\varepsilon_{AB} \sim G\left(0,u^B\right)$，则$V_B+V_{AB}+\varepsilon_{AB} \sim G\left(V_B+V_{AB},u^B\right)$，$\max\left(V_B+V_{AB}+\varepsilon_{AB}\right) \sim G\left(\frac{1}{u^B}\ln\max \mathrm{e}^{(V_B+V_{AB})u^B},u^B\right)$，其中$\eta=\frac{1}{u^B}\ln\max \mathrm{e}^{(V_B+V_{AB})u^B}$为位置参数，则出行者选择上层选择肢 A 的概率可以表示为：

$$P(A)=\frac{\mathrm{e}^{U^A(V_A+\eta)}}{\sum_{A'=1}^{A_J}\mathrm{e}^{u^A(V_{A'}+\eta')}} \tag{4}$$

出行者的上层选择肢选择完成后，再选择下层选择肢 B 的概率为：

$$P(B\mid A)=\frac{\mathrm{e}^{U^B(V_B+V_{AB})}}{\sum_{B'=1}^{B_J}\mathrm{e}^{u^B(V_{B'}+V_{AB'})}} \tag{5}$$

因此，根据上述出行者选择上下层选择肢的概率可以得到出行者联合选择的概率可表示为：

$$\begin{aligned}P(AB)&=P(B\mid A)\cdot P(A)\\&=\frac{\mathrm{e}^{U^B(V_B+V_{AB})}}{\sum_{B'=1}^{B_J}\mathrm{e}^{u^B(V_{B'}+V_{AB'})}}\cdot\frac{U^A(V_A+\eta)}{\sum_{A'=1}^{A_J}\mathrm{e}^{u^A(V_{A'}+\eta')}}\end{aligned} \tag{6}$$

从上式出行者选择上下层选择肢和联合选择的概率可以看出，在进行出行选择时上下层选择肢之间会产生一定的联系，而位置参数η则可以连接上下层选择肢。将位置参数η定义为 logsum 变量，其系数λ可表示为：$\lambda=\frac{u^A}{u^B}$。一般来看如果巢式 Logit 结构模型搭建合理，则λ的取值范围为(0, 1)。

3.2. 巢式 Logit 模型的参数标定方法

巢式 Logit 模型的参数标定方法可分为极大似然估计法和分阶段估计法，本文采用极大似然估计法进行参数标定。在参数标定之前，首先确定各层选择肢的可观测效用和极大似然函数。上层选择肢的效用函数可表示为：

$$u_A=\sum_{i=1}^{n_2}\alpha_i\cdot X_i+\lambda\cdot\eta \tag{7}$$

则下层选择肢的效用函数可表示为：

$$u_B=\sum_{i=1}^{n_1}\beta_i\cdot X_i \tag{8}$$

其中，α_i，β_i分别是影响上下层选择肢选择的变量系数，λ为η变量的系数。则极大似然函数可表示为：

$$L=\prod_{n=1}^{N}\prod_{A_i=1}^{A_J}\prod_{B_j=1}^{B_J}\left(P(B\mid A)\cdot P_n(A)\right)^{y_{AB}} \tag{9}$$

对上式取对数可得：

$$L'=\ln L=\sum_{n=1}^{N}\sum_{B_j=1}^{B_J}y_{B_j}\cdot\ln P(B\mid A)+\sum_{n=1}^{N}\sum_{A_i=1}^{A_J}y_{A_i}\cdot\ln P(A)=L_1'+L_2' \tag{10}$$

其中，L_1',L_2'可看作为各层的对数极大似然函数，y_{AB},y_A,y_B均为(0, 1)变量，可表示为如下式子：

$$y_{AB}=\left\{\begin{array}{c}0\\1\ \text{出行者选择出行方案}\left(A_i,B_j\right)\end{array}\right\} \tag{11}$$

$$y_{A_n}=\left\{\begin{array}{c}0\\1\ \text{出行者选择出行方案}\ A_i\end{array}\right\} \tag{12}$$

$$y_{B_n} = \begin{Bmatrix} 0 \\ 1 \text{ 出行者选择出行方案 } B_j \end{Bmatrix} \tag{13}$$

本文采用 R 语言进行巢式 Logit 模型的编写过程以及进行参数标定，标定过程采用自模型下层到上层的标定顺序。在软件中对巢式 Logit 模型参数标定的步骤为：

1) 首先根据极大似然函数标定影响下层选择方案的变量系数 β_i；

2) 将已标定的估计值 β_i 带入 η 变量表达式中，得到 η 变量；

3) 将 η 变量作为上层选择方案的一个变量，再利用极大似然估计函数对影响上层选择方案的变量系数 α_i 和 η 变量系数 λ 进行标定；

4) 最后，运用显著性水平为 0.05 的 t 检验法对变量的显著性进行检验。

4. 数据分析

4.1. 模型建立及变量选择

经过前期的设计和统计分析，最终决定建立以出行方式为上层、出行目的为下层的巢式 Logit 模型(如图 7 所示)，其中出行方式包含私家车或网约车、公共交通、地铁、电动自行车、自行车或步行四种类型，出行目的包含上学、工作、就餐、购物、探亲访友、旅游五种类型。建立出行方式和出行目的互相影响的巢式 Logit 模型。根据对居民出行意愿的前期问卷调查所获得的相关数据，选择个人属性特征(性别、年龄、受教育程度、职业、月收入水平)、客观环境特征(小汽车、驾照、公交卡、家庭人口数、60 岁以上人口数、6 岁以下人口数)以及出行属性特征(出行时间、出行距离、出行目的、出行方式、出行费用、出行频率、弹性出行次数、非弹性出行次数和不同距离下出行时采用的出行方式)三个方面对居民日常出行选择行为进行分析，其中具体的变量说明如表 2 所示。

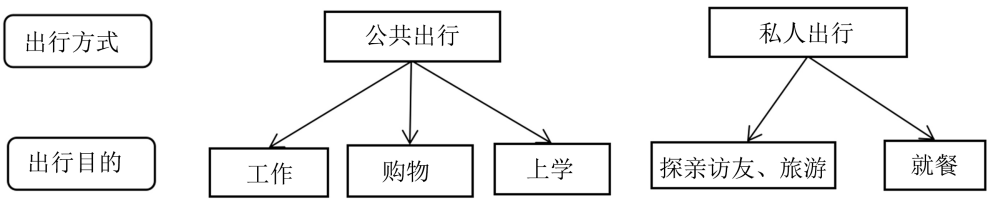

Figure 7. Structure of nested Logit model
图 7. 巢式 Logit 模型结构

Table 2. Detailed index classification of explanatory variables
表 2. 解释变量详细指标分类

类别	变量名	详细说明
因变量	出行方式	1. 公共出行(402) 2. 私人出行(111)
个人出行基本属性	性别	1. 男(264) 2. 女(249)
	年龄	1. 20 岁以下(46) 2. 21~30 岁(273) 3. 31~50 岁(171) 4. 50 岁及以上(23)
	受教育程度	1. 高中及以下(114) 2. 专科(202) 3. 本科(164) 4. 研究生及以上(33)
	职业	1. 个体经营户或私营企业主(56) 2. 公务员(24) 3. 企业职工(301) 4. 事业单位职员(49) 5. 学生(83)
	收入	1. 2000 元及以下(139) 2. 2001~4000 元(87) 3. 4001~6000 元(192) 4. 6001~8000 元(84) 5. 8000 元及以上(11)
	驾照	1. 无(127) 2. 有(386)
	公交卡	1. 无(119) 2. 有(394)

续表

出行属性	出行时间	1. 0~10 分钟(120) 2. 11~30 分钟(254) 3. 31~60 分钟(63) 4. 61 分钟及以上(12) 5. 基本不出门(64)
	出行距离	1. 5 公里及以下(365) 2. 6~10 公里(126) 3. 11~20 公里(18) 4. 21~30 公里(2) 5. 31 公里及以上(2)
	出行目的	1. 工作(183) 2. 购物(120) 3. 就餐(87) 4. 上学(83) 5. 探亲访友、旅游(40)
	出行费用	1. 5 元及以下(126) 2. 6~10 元(207) 3. 11~20 元(98) 4. 21~30 元(53) 5. 31~60 元(24) 6. 51 元及以上(5)
	出行频率	1. 每周出行 2 天及以下(78) 2. 每周出行 3 天(177) 3. 每周出行 4 天(232) 4. 每周出行 5 天及以上(26)
家庭属性	小汽车	1. 无(149) 2. 有(364)
	家庭 6 岁以下孩子数量	1. 0 人(233) 2. 1 人(179) 3. 2 人(87) 4. 3 人及以上(14)
	家庭 60 岁以上人口数量	1. 0 人(383) 2. 1 人(48) 3. 2 人(47) 4. 3 人及以上(35)

4.2. 主成分分析

主成分分析是通过原有变量的线性组合以及各个主成分的求解来实现变量降维，利用少量的数据来代替原始数据。尽管主成分分析法可能会降低原始数据的精度，但是它的简便计算的方法却常用于分析大量的数据分析中。鉴于调查问卷选取了较多类型的影响因素，从初步的分析结果来看，并不是所有的影响因素都与因变量出行方式存在显著性的相关关系，所以可以通过主成分分析法来对影响因素进行降维分析。经过初步分析变量后，选取性别、年龄、受教育程度、职业、月均收入、出行时间、出行距离、出行目的、出行费用、出行频率、家中 60 岁以上人口数 11 个影响因素进行主成分分析。验证构建的模型是否符合主成分分析法的原则是选取的影响因素要通过 KMO (Kaiser-Meyer-Olkin)检验和 Bartlett 球形检验，KMO 检验通过的标准是 KMO 检验值要大于 0.5，而 Bartlett 球形检验通过的标准是 P 值要小于 0.05。运用 R 软件做 KMO 检验和 Bartlett 球形检验，得到分析结果 KMO 检验值 KMO = 0.61，Bartlett 球形检验中 P = 0.000 < 0.005，说明该模型适合主成分分析方法。但是在研究选取的每个因素对原始变量的方差贡献率时(如图 8)可以看出，方差贡献率曲线后期并没有明显的平缓趋势，碎石图也呈现出同样的趋势。所以无法通过主成分分析方法来确定应该选取哪几个因子来替代原始变量。所以尽管 KMO 检验和 Bartlett 球形检验的两个结果都满足主成分分析的需求，但是仍无法采用主成分分析法来筛减变量。

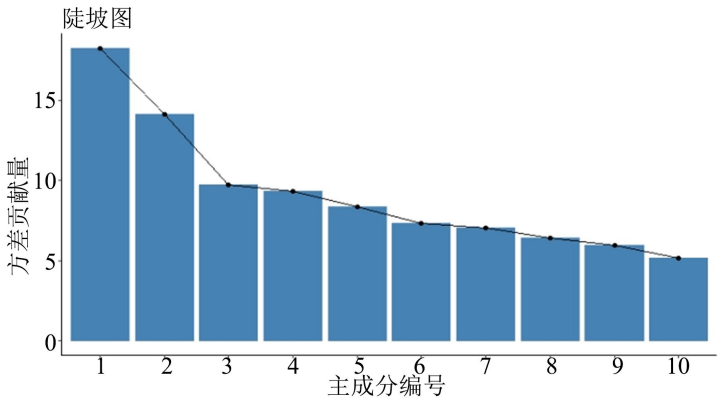

Figure 8. The variance contribution corresponding to each principal component
图 8. 每个主成分对应的方差贡献量

4.3. 模型参数标定及结果分析

运用问卷调查中居民出行选择行为的数据进行分析，并鉴于问卷调查数据量庞大，所以采用分阶段估计法进行分析。分阶段估计法是先将下层模型当做独立的多项式 Logit 模型进行参数估计，根据下层参数的估计结果计算包容值然后将其作为影响上层模型的一个因素，之后再对上层模型进行参数估计。本文采用 R 语言进行函数编写以及对前期建立的巢式 Logit 模型进行参数标定。因为变量数量过多，且采用主成分分析方法无法对数据进行降维处理，并考虑到居民出行选择行为的特点，对出行方式进行二分类处理。在进行模型分析的过程中，将地铁、公共交通和出租车网约车作为一类，重新定义为公共出行。将私家车、电动自行车、自行车和步行作为一类，重新定义为私人出行。如果将表 1 中含有的所有变量均作为居民出行方式和出行目的选择行为模型中的变量。得到的模型参数标定结果中的有一些变量的检验值会显著偏低，说明该变量对出行方式和出行目的选择行为的影响较小，且鉴于同一变量对不同选择肢的影响也不尽相同，所以应根据居民日常出行行为的选择特点，同时考虑到对出行方式和出行目的有影响的主要因素。在参数标定过程中，出行方式选择以私人出行作为参考类别。出行目的选择以上学作为参考类别。出行方式选择经过定性分析和试算，得到了模型参数标定的结果，如表 3 和表 4 所示。

Table 3. Calibration results of model parameters for travel mode selection
表 3. 出行方式选择模型参数标定结果

变量	公共出行
性别	
男性	−1.652 (0.037)
职业	
事业单位职员	1.043 (0.017)
企业职工	1.084 (0.000)
公务员	0.074 (0.085)
学生	16.643 (0.066)
是否有驾照	
有	−0.856 (0.001)
是否有小汽车	
有	−0.674 (0.005)
Log likelihood = −720.91958	
Prob > chi2 = 0.000	

Table 4. Calibration results of model parameters for travel purpose selection
表 4. 出行目的选择模型参数标定结果

变量	就餐	工作	探亲访友、旅游	购物
家庭 6 岁以下孩子数	−7.371 (0.081)	0.763 (0.061)	12.888 (0.432)	−2.922 (0.062)
家庭 60 岁以上人口数	−2.385 (0.049)	1.092 (0.016)	−4.627 (0.330)	−4.078 (0.076)
是否有小汽车	−1.470 (0.043)	5.626 (0.042)	3.714 (0.035)	−4.159 (0.073)
是否有驾照	6.756 (0.122)	12.502 (0.009)	4.035 (0.041)	4.963 (0.050)
Log likelihood = −719.10924				
Prob > chi2 = 0.1109				

从表 3 出行方式选择模型的参数标定结果可以看出：就个人属性来看，性别、职业、是否有驾照三个因素对居民出行选择行为有显著的影响。在性别方面，相对于女性来说，男性似乎对公共出行的方式并不感兴趣，他们更倾向于选择私人出行的出行方式。职业选择中，相较于私人出行的出行方式，事业单位职员、企业职员、公务员、学生都更加青睐于公共出行，且其中以学生最盛，系数达到了 16.643，显著高于其他三项职业人员，这也与学生的经济能力相关。在是否有驾照方面，有驾照的居民系数为负数，表明他们更倾向于私人出行的出行方式。就家庭属性来看，是否有小汽车这个因素对居民出行选择行为有显著的影响，有小汽车的居民参数标定结果为负数，表明有小汽车的居民较公共出行更偏向于私人出行。

从表 4 出行目的选择模型的参数标定结果可以看出：就个人属性来看，有驾照的居民相较于没有驾照的居民更倾向于各种目的的出行，且参数标定结果都较大，表明出行意愿更加强烈。出行就家庭属性来看，家庭 6 岁以下孩子数和家庭 60 岁以上人口数以及是否有小汽车三个因素对居民出行选择行为有显著的影响。在家庭中有 6 岁以下孩子的团体中，出行目的仍以工作和探亲访友、旅游居多，居民一般不倾向于就餐和购物。有 60 岁以上的老人的家庭的出行目的相较于上学依旧是多以工作为主，就餐、探亲访友、旅游、购物的参数都为负数，表明这些家庭不倾向于选择这几种因素作为出行目的。有小汽车的家庭的出行目的相较于上学多以工作和探亲访友、旅游为主。可以看出，在影响出行目的选择行为的模型中，有 6 岁以下孩子或者 60 岁以上老人或者有小汽车的家庭的就餐和购物的参数标定结果都为负数，表明当下居民不倾向于选择就餐和购物作为自己的出行目的，这与当下的整体大环境息息相关。

4.4. 研究数据的局限性

本研究采用在线 SP 调查问卷的方式获取数据，具有一定的样本选择性偏差。一方面，样本主要集中在具备网络访问能力和答题意愿的城市青年人群中，难以覆盖年龄偏大、网络使用能力较弱或居住在农村地区的群体；另一方面，SP 调查基于假设情境模拟，受访者的回答未必完全反映真实出行行为，可能存在假设性偏差。这些因素可能影响模型的外推能力，使得研究结论更适用于特定群体，需在后续研究中结合更大范围、更具代表性的样本数据进一步验证。

5. 结论

城市居民的出行选择是影响城市交通的重要因素。我们选取多种出行特征因素可以将出行机制清晰化，以便更准确地解释居民的实际出行行为。巢式 Logit 模型为我们研究出行方式和出行目的之间的联系提供了可靠的方法。根据模型分析结果可以看出，性别、职业、驾照和是否拥有小汽车都对出行方式有显著的影响。而家庭中 6 岁以下孩子数和 60 岁以上人口数以及驾照和是否拥有小汽车四个因素都对出行目的有一定的影响。因此，在实际交通管理过程中，结合本文的参数标定结果，可以相应地采取有针对性的倾向性管理措施和鼓励性引导政策：鼓励人群多乘坐公交车出行，可适当地增加公交班次，优化公交站点和公交线路。同时加强公交专用车道的治理，确保公交的正常有序运行。通过确定影响居民出行方式选择的因素，可以采取更具针对性的交通管理措施，科学引导城市青年群体的出行行为，满足其出行需求，进而提升整体城市的交通运行效率。需注意，本文基于网络问卷所得数据，研究结论主要适用于城市年轻群体，对其他人群的适用性仍需进一步验证。

基金项目

本研究获得以下项目资助：内蒙古自治区自然科学基金(项目编号：2024MS07002)，内蒙古工业大学“行人疏散建模与仿真”研究启动项目(项目编号：DC2200000934)，以及内蒙古自治区重点研发成果转化计划(2025)“低空经济高效视频传输系统”(项目编号：2025YFHH0215)。

参考文献

[1] Koppelman, F.S. (1981) Non-Linear Utility Functions in Models of Travel Choice Behavior. *Transportation*, **10**, 127-146. https://doi.org/10.1007/bf00165262

[2] Palma, A.D. and Rochat, D. (2000) Mode Choices for Trips to Work in Geneva: An Empirical Analysis. *Journal of Transport Geography*, **8**, 43-51. https://doi.org/10.1016/s0966-6923(99)00026-5

[3] Bhatta, B.P. and Larsen, O.I. (2011) Errors in Variables in Multinomial Choice Modeling: A Simulation Study Applied to a Multinomial Logit Model of Travel Mode Choice. *Transport Policy*, **18**, 326-335. https://doi.org/10.1016/j.tranpol.2010.10.002

[4] Limtanakool, N., Dijst, M. and Schwanen, T. (2006) The Influence of Socioeconomic Characteristics, Land Use and Travel Time Considerations on Mode Choice for Medium- and Longer-Distance Trips. *Journal of Transport Geography*, **14**, 327-341. https://doi.org/10.1016/j.jtrangeo.2005.06.004

[5] Saleh, W. and Farrell, S. (2005) Implications of Congestion Charging for Departure Time Choice: Work and Non-Work Schedule Flexibility. *Transportation Research Part A*: *Policy and Practice*, **39**, 773-791. https://doi.org/10.1016/j.tra.2005.02.016

[6] Ozbay, K. and Yanmaz-Tuzel, O. (2008) Valuation of Travel Time and Departure Time Choice in the Presence of Time-Of-Day Pricing. *Transportation Research Part A*: *Policy and Practice*, **42**, 577-590. https://doi.org/10.1016/j.tra.2007.12.002

[7] 叶倩文, 陈旭梅, 柏喜红. 基于 NL 模型的北京公共交通票价对出行方式影响研究[J]. 山东科学, 2019, 32(2): 79-88, 114.

[8] 杨励雅, 邵春福, HAGHANIA. 出行方式与出发时间联合选择的分层 Logit 模型[J]. 交通运输工程学报, 2012, 12(2): 76-83.

[9] 杨昌涛, 靳文舟, 范雪婷. 基于巢式 Logit 模型的交通方式选择行为研究[J]. 公路与汽运, 2011(4): 62-65.

[10] 李佳玮. 基于非集计模型的常规公交票价对居民出行行为影响研究[D]: [硕士学位论文]. 兰州: 兰州交通大学, 2021.

[11] 柳文燕, 傅忠宁, 李万慧, 翟涛涛. 基于巢式 Logit 模型的低收入人群出行行为分析——以上海市智能手机调查数据为例[J]. 交通运输工程与信息学报, 2020, 18(3): 40-49.

[12] 欧舟. 城市居民出行行为及特征研究[D]: [硕士学位论文]. 长沙: 长沙理工大学, 2012.

[13] 肖光年, 隽志才, 张春勤. 基于贝叶斯网络和 GPS 轨迹数据的出行方式识别[J]. 统计与决策, 2017(6): 75-79.