

文章编号:1007-6735(2012)05-0419-05

考虑重复观测之间关联性的信息响应建模

干宏程¹, 张 宁²

(1. 上海理工大学 超网络研究中心, 上海 200093; 2. 上海理工大学 管理学院, 上海 200093)

摘要: 以高架快速路上可变情报板发布高架和地面道路行程时间条件下改道决策的意向数据为建模对象, 进行信息响应的离散选择建模, 建立二元 Probit 横截面模型(CM)和考虑重复观测之间关联性的面板模型(PM), 分析两种模型的性能. 研究表明: 驾龄越长, 高架使用频率越高; 中等年龄段的驾驶员越愿意改道; 地面交叉口越少, 节省时间越多, 越愿意改道; 单位用车驾驶员越不愿意改道. PM 的拟合度明显高于 CM; PM 在模型系数上与 CM 相差不大, 但是其统计推断鲁棒性更好; CM 会明显高估驾驶员相关的解释变量系数的 t 检验值, 但是会略微低估出行情境相关的解释变量系数的 t 检验值.

关键词: 信息响应; PM; CM; 二元 Probit; 可变情报板; 行程时间; 改道意向

中图分类号: U 491 **文献标志码:** A

Modeling of Driver Response to Travel Information by Taking Account of Correlations Among Repeated Observations

GAN Hong-Cheng¹, ZHANG Ning²

(1. Center for Supernetworks Research, University of Shanghai for Science and Technology, Shanghai 200093, China;
2. Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: Urban freeway users' diversion response to VMS that display travel time on both urban freeway and local streets was explored. Stated preference data of diversion behavior was used to develop a cross-sectional binary probit model and a panel binary probit model for identifying factors that influence diversion response. Main findings regarding VMS impacts are travel time saving and drivers' years of driving experience serve as positive factors in diverting; number of traffic lights on the local street, using frequency of urban freeway; being a mid-age driver, and being an employer-provided car driver serve as negative factors in diverting. On the modeling aspect, it was shown that the panel model does not provide substantially different model coefficients but more robust statistical inferences for model coefficients as compared to the cross-sectional model, and the cross-sectional model tends to seriously overestimate t -test values for explanatory variables correlated with drivers (e. g. demographic characteristics) but slightly underestimate t -test values for explanatory variables correlated with scenarios (e. g. travel time savings). The findings have implications for better design and operation of advanced traveler information systems and for future

收稿日期: 2012-07-06

基金项目: 国家自然科学基金资助项目(51008195); 上海市重点学科建设资助项目(S30504); 上海市教委科研创新资助项目(09YZ205)

作者简介: 干宏程(1978-), 男, 副教授. 研究方向: 交通运输规划与管理、智能交通管理等. E-mail: hongchenggan@126.com

effort on survey design, data collection and model estimation.

Key words: *response to information; panel model; cross-sectional model; binary probit; variable message signs (VMS); travel time; diversion; stated preference data*

正确理解实时交通信息影响下的驾驶员响应(例如路径选择、改道)是对先进旅行者信息系统(ATIS)进行优化与评价的必然要求,因为 ATIS 的效果主要取决于驾驶员对信息的响应.国际上,驾驶员信息响应(driver response to information)是交通行为研究的一个热点领域^[1-10],属于管理科学、交通科学、系统科学、行为科学、经济学等学科的交叉研究.运用微观计量经济学中的离散选择理论进行信息响应的建模是该领域的一个前沿^[1-11].用离散选择理论建立的择路概率模型能够揭示择路概率与信息、出行者社会经济属性、出行特征、交通状况等因素之间的映射关系,为动态交通分配提供理论基础,以便评价路网层面的 ATIS 效益,为制定与优化 ATIS 策略提供依据和启示^[6-8,10].例如,交通管理者可以借助择路概率模型预估城市快速路发生异常情况下交通信息引起的分流比例,以便调整替换道路上信号灯控制方案,确保路网整体效率更高.国内,以往信息响应离散选择建模的文献报道不多,近几年才日益增多^[12-14].目前,国际上驾驶员信息响应建模的前沿研究新进展的一个方面是,考虑“同一被试重复观测的关联(correlations among repeated observations of the same individual)”问题,以使模型更精确和符合实际^[2,9].在 ATIS 信息响应的意向 SP 调查中,对同一被试做多次观测相当普遍,而信息响应模型是利用 SP 数据建立的,因此考虑重复观测关联尤为重要.国内,尚未见到考虑了重复观测之间关联性的信息响应研究报道.值得一提的是,一些学者用网络均衡分析(如交通分配)等方法考察交通信息下交通网络运行效率,得到了诸多有益成果^[15].这些工作(因侧重点不在驾驶员个体行为建模上)本质上是假设出行者遵守信息响应规则的前提下进行的,这从另一方面也显示了驾驶员信息响应建模的重要性,信息响应建模研究成果与以往工作有很好的互补性.

本文紧跟国际前沿,利用上海快速路驾驶员在 ATIS 信息影响下改道行为的 SP 数据,开展信息响应建模研究,建立考虑重复观测关联性的面板模型(panel model)和常规(不考虑关联性)的横截面模型(cross-sectional model),不仅揭示 ATIS 信息与改

道概率的关系,而且通过比较两种模型的性能得出可为进一步开展信息响应建模研究提供参考与启示的有益结论.

1 信息响应建模

1.1 用于建模的行为数据

信息响应建模的对象为驾驶员在快速路可变情报板(VMS)的行程时间信息影响下的改道行为.行为数据为2007年4月笔者研究组对上海驾驶员的问卷调查得到的 SP 数据.SP 调查构造的出行情境(见图1)为:行驶于上海高架快速路上的驾驶员在接收到分流点(出口匝道)上游一块 VMS 发布的行程时间信息后,做出“改道至地面道路”或“继续走高架”的决策.VMS 提供高架和地面道路的行程时间.

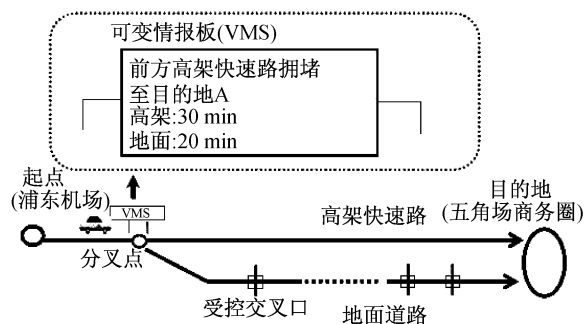


图1 SP 问卷调查中构造的交通出行情景

Fig.1 SP journey of the study

采用正交设计方法,SP 问卷中共构造 8 个改道决策的情境.建模用的样本量为 140 份问卷,一共有 1 120 个(140×8)样本用于数据建模.驾驶员个体属性统计情况:性别方面,男性占 74.3%;年龄方面,30 岁以下占 24.3%,30~40 岁之间占 41.4%,40~50 岁占 25.0%,50 岁以上占 9.3%;驾龄方面,2 a 以下占 6.4%,2~5 a 占 41.4%,6 a 以上占 52.1%;驾车类型方面,私家车占 42.1%,单位公车占 35%,出租车占 22.9%;高架使用频率方面,几乎每天占 49.3%,每周 2~3 d 占 23.6%,很少以及从来不使用占 27.2%.

1.2 建模

研究中 SP 数据为面板数据(每个被试回答 8 个

情境的问题),因此以下建立横截面模型与面板模型,分析和比较两种模型描述驾驶员信息响应行为的性能.

1.2.1 横截面模型

SP 调查中,驾驶员的决策为二元选择,故采用二元 probit 模型^[11]建立择路概率模型.决策者 i 选择路径时所基于的随机效用函数可表达为

$$U_i^* = \beta_0 + \mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i \quad (1)$$

式中,向量 $\mathbf{x}_i$ 由各种路径选择影响变量(如行程时间节省量、信控交叉口数量、驾驶员各项个体属性)组成; β_0 为待估计的常量; $\boldsymbol{\beta}$ 为与影响变量相对应的系数组成的向量; ε_i 是随机项,服从独立同分布的正态分布.只要 U_i^* 大于 0,则决策者 i 将选择改道,即

$$P(U_i^* > 0) = P(y_i = 1) \quad (2)$$

$$P(U_i^* < 0) = P(y_i = 0) \quad (3)$$

y_i 表示决策者 i 是否改道的哑变量($y_i = 1$,改道至地面; $y_i = 0$,继续走高架).将式(1)代入式(2),则有

$$P(y_i = 1) = \Phi(\beta_0 + \mathbf{x}_i\boldsymbol{\beta}) \quad (4)$$

$$P(y_i = 0) = 1 - \Phi(\beta_0 + \mathbf{x}_i\boldsymbol{\beta}) \quad (5)$$

进一步地,决策者 i 选择路径的概率函数可以公式化为

$$P(y_i) = [\Phi(\beta_0 + \mathbf{x}_i\boldsymbol{\beta})]^{y_i} [1 - \Phi(\beta_0 + \mathbf{x}_i\boldsymbol{\beta})]^{1-y_i} \quad (6)$$

1.2.2 面板模型

对于情境 t ,决策者 i 选择路径时所基于的随机效用函数可表达为

$$U_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \rho v_i + \varepsilon_{it} \quad (7)$$

式中,向量 $\mathbf{x}_{it}$ 由不随情境变化的影响变量(如性别、年龄)与不随驾驶员变化的影响变量(如行程时间节省量、信控交叉口数)组成; ρ 为是对应于 v_i 的系数,表征了不可观测随机变量的方差; v_i 为用于考虑那些影响同一被试选择行为的共通的不可观测的随机变量(如驾驶员习惯、偏好), ε_{it} 为随个体和情境变化而变化的随机变量,为便于比较面板模型和横截面模型,对式(7)作重新构造,得到

$$U'_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \rho v_i + \sqrt{1+\rho^2} \varepsilon_{it} \quad (8)$$

式中, v_i 和 ε_{it} 都为标准正态分布.于是效用函数的随机部分的方差为 1,与横截面模型中的方差相同.这样面板模型和横截面模型就能相互比较.不失一般性的,此时效用函数可以重构为

$$U'_{it} = \frac{\mathbf{x}_{it}\boldsymbol{\beta}}{\sqrt{1-\rho^2}} + \frac{\rho}{\sqrt{1-\rho^2}} v_i + \varepsilon_{it} \quad (9)$$

令 y_{it} 表示决策者 i 在情境 t 中是否改道的哑变量($y_{it} = 1$,改道; $y_{it} = 0$,继续走高架). $U_{it} > 0$ 时, $y_{it} = 1$; $U_{it} < 0$ 时, $y_{it} = 0$.于是得到关于 v_i 的条件概率函数

$$P(y_{it} | v_i) = [P(y_{it} = 1 | v_i)]^{y_{it}} \cdot [P(y_{it} = 0 | v_i)]^{(1-y_{it})} \quad (10)$$

全概率函数 P_i 则通过在 v_i 的概率密度函数的分布域上对 $P(y_{it} | v_i)$ 进行积分,得到

$$P_i = \int_{-\infty}^{+\infty} \phi(v_i) \prod_{t=1}^T P(y_{it} | v_i) dv_i \quad (11)$$

式中, $T = 8$,代表被试面临 8 个 SP 情境,利用 Gaussian-Hermite 积分法估计式(11),得到

$$P_i \approx \sum_{k=1}^K [w_k \phi(z_k) \prod_{t=1}^T P(y_{it} | z_k)] \quad (12)$$

式中,常数 K 取为 10; w_k 和 z_k 为权重和横坐标.最后,整个样本的对数似然函数构造为

$$LL = \sum_{i=1}^N \ln(P_i) \quad (13)$$

式中, N 为样本大小.

采用 Gauss 软件,模拟似然函数及其一阶导数的程序代码编写和模型参数估计.

2 模型估计结果分析

2.1 模型估计结果

共有 1 120(140×8)个样本用于模型估计.估计模型时所选的解释变量包括所有可能影响改道行为的因素:年龄、性别、驾龄、快速路使用频率、用车类型、行程时间节省量、信控交叉口数量、快速路延误原因.表 1(见下页)给出了模型估计结果.由表 1 可见,具有显著影响的变量有:常数 C 、年龄 A_1 、年龄平方 A_2 、驾龄 D 、很少使用快速路的哑变量 S 、几乎每天使用快速路的哑变量 E 、单位公车驾驶员的哑变量 p 、地面道路信控交叉口数量 n 、出行时间节省量 t_0 .高架延误原因和性别两个变量因所获系数不显著,没进入模型.

2.2 关于驾驶员信息响应的发现及讨论

如表 1 所示,不论是面板模型还是横截面模型,各影响变量的系数都获得了符合逻辑的符号.模型揭示的规律有:

a. 年龄

年龄系数为负,而年龄平方系数为正,说明年轻的和年长的驾驶员更可能受 VMS 影响而改道.计算可知,36~37 岁的驾驶员改道的可能性最小.对这

一现象可能的解释是,年轻的驾驶员耐性较差,对延误更加敏感,而年长的驾驶员对延误引起的油耗增加更加不能容忍,因而,在 VMS 告知延误后更愿意改道.

b. 驾龄
驾龄系数为正,说明驾龄越长,改道可能性越大.这是因为驾车经验越丰富的驾驶员可能更加熟悉替换道路,在事故等异常情况发生时更能从容应对,因而在接收 VMS 信息后改道可能性更大.

c. 快速路使用频率
每天使用快速路的哑变量系数为负,很少使用快速路的哑变量系数为正,说明经常使用快速路者更不愿意改道,而很少使用快速路者更可能受 VMS 信息影响而改道.

d. 单位公车驾驶员
单位公车驾驶员的哑变量系数为负,说明单位公车驾驶员相比社会车辆和出租车司机更不愿意改道.可能的原因是:由于他们的车辆出行费用可以由单位报销,因而对延误引起的油耗增加更不敏感,从而在接收 VMS 信息后,改道可能性更小.

e. 地面道路信控交叉口数量
地面道路信控交叉口数量系数为负,说明地面道路信控交叉口越多,改道可能性越小.这一现象符合直觉判断.

f. 出行时间节省量
出行时间节省量系数为正,说明如果改道后能节省的行程时间越多,驾驶员改道的可能性越大.

表 1 面板模型和横截面模型的估计结果
Tab.1 Estimation results of cross-sectional and panel models

变量	横截面模型		面板模型	
	β	t 检验值	β	t 检验值
C	1.961 2	2.335	2.155 5	1.601
A_1	-0.129 7	-2.941	-0.140 1	-1.955
A_2	0.001 8	3.014	0.001 9	1.997
D	0.020 5	2.053	0.021 5	1.309
S	0.220 6	1.949	0.236 8	1.269
E	-0.495 8	-5.073	-0.512 1	-3.188
p	-0.358 4	-4.074	-0.335 0	-2.348
n	-0.056 0	-6.963	-0.054 6	-7.765
t_0	0.166 0	10.285	0.161 6	11.096
ρ	-	-	0.652 4	16.005
$L(\beta)$	-662.29		-602.74	
$L(c)$	-774.72		-774.72	
Adj. $\rho^2(c)$	0.134 8		0.210 4	
N	1 120		140 $\times$ 8	

2.3 关于两种建模方法性能的发现及讨论
面板模型中, ρ 值为 0.652 4,且统计意义上影响显著(t 检验值为 16.005),说明存在不可观测的随机因素对被试的改道决策产生影响.面板模型的似然率指标为 0.210 4,较横截面模型(0.134 8)有显著改善,面板模型的拟合度更好,模型性能更优.

由表 1 可知,面板模型与横截面模型的系数大小差别并不明显,这可能是因为两种模型在理论上都能给出统计意义上一致的估计结果.但是可以发现,面板模型的系数统计推断明显有别于横截面模型.那些随驾驶员而变化的解释变量(年龄、驾龄、高架使用频率、用车类型)的系数 t 检验值明显小于横截面模型中的 t 检验值.在横截面模型中影响显著的驾龄变量和很少使用快速路的哑变量在面板模型中似乎不显著.另一方面,面板模型中,那些随 SP 情境而变化的变量(信控交叉口数量、行程时间节省量)的系数 t 检验值较之横截面模型的略有增加.

面板模型相比横截面模型,模型系数的 t 检验值有显著不同,是因为考虑了随机效应的面板模型能够更精确地估计系数估计值的标准偏差,提供鲁棒性更好的关于系数估计值的统计推断.但是,两种模型的系数估计值的数量是十分接近的.

从文献看,无论国内外,多数研究者与几乎所有实践人员处理 SP 数据时仅建立横截面模型,因此所建的信息响应模型可能有局限性.本文工作有利于启发研究人员更多地关注面板模型以便建立更有效的信息响应模型,为交通信息效益的分析与评价提供更可靠的理论依据.

3 结束语

对高架快速路 VMS 提供高架与地面道路行程时间条件下的驾驶员改道行为开展离散选择建模研究,建立了考虑重复观测之间关联性的面板模型和不考虑关联性的横截面模型,研究表明:

a. 驾驶员对VMS信息响应方面揭示的规律有中等年龄段,驾龄越大,使用高架快速路频率越高,改道可能性越大;单位公车驾驶员相比社会车辆和出租车司机,改道可能性更小;地面道路信控交叉口越少,改道能节省的行程时间越多,改道可能性越大.

b. 信息响应建模方法方面,面板模型性能更优,拟合度明显高于横截面模型;面板模型在模型系数上与横截面模型相差不大,但是得到的关于模型系数的统计推断鲁棒性更好.横截面模型会严重高

估驾驶员相关的解释变量(年龄、驾龄、高架使用频率、用车类型)系数的 t 检验值,但是会略微低估出行情境相关的解释变量(信控交叉口数量、行程时间节省量)系数的 t 检验值。

本文工作有别于以往研究,首次比较了横截面模型与面板模型用于刻画信息响应行为的性能与效果。下一步试图将信息响应模型与网络交通流仿真模型结合,考察VMS信息在路网层面的效益,为交通管理与控制措施与政策的设计提供更有有力支持。

参考文献:

- [1] Jou R C, Lam S H, Liu Y H, et al. Route switching behavior on freeways with the provision of different types of real-time traffic information[J]. Transport Res A, 2005, 39(5): 445 - 461.
- [2] Abdel-Aty M, Abdalla P. Modeling drivers diversion from normal routes under ATIS using generalized estimating equations and binomial probit link function[J]. Transportation, 2004, 31(3): 327 - 348.
- [3] Chatterjee K, Hounsell N B, Firmin P E, et al. Driver response to variable message sign information in London[J]. Transport Res C, 2002, 10(2): 149 - 169.
- [4] Chorus C G. Traveler response to information[D]. Delft: TU Delft, 2007.
- [5] Gan H C. Graphical route information panel for the urban freeway network in Shanghai, China[J]. IET Intel Transp Syst, 2010, 4(3): 212 - 220.
- [6] Abdel-Aty A, Kitamura M R, Jovanis P. Using stated preference data for studying the effect of advanced traffic information on drivers' route choice[J]. Transport Res C, 1997, 5(1): 39 - 40.
- [7] Mahmassani H S, Liu Y H. Dynamics of commuting decision behavior under advanced traveler information systems[J]. Transport Res C, 1999, 7(2): 97 - 107.
- [8] Peeta S, Ramos J L, Pasupathy R. Content of variable message signs and on-line driver behavior[J]. Transp Res Rec, 2000, 1725: 102 - 108.
- [9] Srinivasan K K, Mahmassani H S. Analyzing heterogeneity and unobserved structural effects in routeswitching behavior under ATIS: a dynamic kernel logit formulation[J]. Transport Res B, 2003, 37(9): 793 - 814.
- [10] Wardman M, Bonsall P W, Shires J D. Driver response to variable message signs: a stated preference investigation[J]. Transport Res C, 1997, 5(6): 389 - 405.
- [11] Ben-Akiva M, Lerman S R. Discrete choice analysis[M]. Cambridge: MIT Press, 1985.
- [12] 干宏程, 孙立军, 陈建阳. 提供交通信息条件下的途中改道行为研究[J]. 同济大学学报, 2006, 34(11): 1484 - 1488.
- [13] 干宏程. VMS诱导信息影响下的路径选择行为分析[J]. 系统工程, 2008, 26(3): 11 - 16.
- [14] 石京, 陶立. 实时交通信息提供对驾驶员路径选择行为影响量化分析[J]. 武汉理工大学学报(交通科学与工程版), 2010, 34(4): 639 - 643.
- [15] 吴文祥, 黄海军. 平行路径网络中信息对交通行为的影响研究[J]. 管理科学学报, 2003, 6(2): 12 - 16.