

文章编号:1007-6735(2014)01-0097-06

基于上海市消费者的汽车共享选择分析

周 彪¹, 周溪召¹, 李 彬²

(1. 上海海事大学 经济管理学院, 上海 201306; 2. 上海市交通港航发展研究中心, 上海 200025)

摘要: 利用 Mcfadden 提出的离散选择模型研究了消费者选择汽车拥有和汽车共享的主要影响因素. 基于对上海市消费者的问卷调查, 构建单次出行选择汽车共享的多项 Logit 模型和因汽车共享而放弃购车的多项 Logit 模型, 探讨影响消费者选择汽车共享因素的重要性, 分析消费者个人长期的汽车拥有行为.

关键词: 汽车共享; 影响因素; Logit 模型; SPSS 软件

中图分类号: U 49 **文献标志码:** A

Choice Analysis of Car Sharing in Consideration of the Consumers in Shanghai

ZHOU Biao¹, ZHOU Xi-zhao¹, LI Bin²

(1. School of Economic and Mangement, Shanghai Maritime University, Shanghai 201306, China;
2. Shanghai Transport & Port Research Center, Shanghai 200025, China)

Abstract: With the aid of Mcfadden's discrete choice models, the main factors influencing the consumer choice between car ownership and car sharing were studied. The multinomial Logit model to select car sharing and the multinomial Logit model to give up purchasing a car duo to car sharing were constructed based on the survey of Shanghai market. The importance of factors influencing the choice between car sharing and car ownership of long term was discussed.

Key words: car sharing; influencing factor; Logit model; SPSS software

20 世纪 80 年代末期, 汽车共享在欧洲出现了加速发展的趋势. 到目前为止, 欧洲有近 200 个汽车共享服务组织, 遍布瑞士、德国、英国、奥地利、荷兰、丹麦、瑞典及挪威等国家, 会员总数达到 12.5 万人. 汽车共享研究和实践在国外取得了一系列成果^[1], 但由于我国的交通基础设施条件、居民出行方式及城市基本结构等与欧美国家存在差异, 使得它们的研究成果并不能直接应用到我国的汽车共享之中.

在我国, 张小明^[2]研究了汽车共享出现的动因, 主要提出了私家车发展带来的 3 个问题: 环境问题、资源问题和堵车问题. 薛跃等^[3]在研究中阐明, 基于私人拥有为基础的汽车普及消费既是一种大幅度提高社会福利水平的生活方式, 又是一种非可持续性消费模式, 它不仅消耗大量自然资源, 且造成极大的环境危害. 程伟力^[4]在关于汽车共享服务的对象研究中指出, 汽车共享服务组织的服务对象是那些富有驾

收稿日期: 2013-04-05
基金项目: 教育部人文社科规划资助项目(11YJA790227); 上海市研究生教育创新计划资助项目(wk2013002)
第一作者: 周彪(1990-), 男, 硕士研究生. 研究方向: 运输管理. E-mail: ziyue0574@163.com
通讯作者: 周溪召(1963-), 男, 教授. 研究方向: 交通运输规划与管理. E-mail: xzzhou@shmtu.edu.cn

驶经验且遵守交通规则的人。

本文将基于消费者选择行为理论,通过问卷调查,明确汽车共享的潜在市场.通过建立模型来分析购车与汽车共享之间的效用差异,明确选择购车或汽车共享影响因素的权重,分析消费者个人长期的汽车拥有行为。

1 模型构建

1.1 问卷目的

通过有效的问卷调查描述各个因素对个人选择行为的影响.影响汽车拥有和汽车共享的主要因素包括:a.家庭经济条件,如家庭收入、家庭车辆拥有数、家庭房屋拥有情况、家庭成员及个人驾驶经验;b.相关费用,如现有汽车的使用费用、计划购车的费用、汽车共享的使用费用;c.用车需求,如汽车使用频率、家庭未成年人数、居住地公共交通条件;d.个人特征,如个人通勤及其它出行特征、个人环保意识、个人年龄、个人职业及受教育程度。

通过问卷调查,分析各个影响因素对个人消费者选择汽车共享的权重,在给定消费者情况时,能够迅速发现其是否为汽车共享的潜在顾客群。

1.2 问卷设计

问卷主要从上述主要因素的具体表现形式来设计问题,框架如下:第一部分为个人特征部分,主要包括性别、年龄及受教育程度等情况;第二部分是家庭经济情况,主要是收入情况、家庭车辆拥有情况、家庭成员及个人驾车经验情况;第三部分主要为用车需求调查,主要搜集的是日常出行方式、现有交通出行情况、出行(用车)频率等方面信息;第四部分为相关费用部分,主要是了解被访者车辆的价格、每月开销等情况;第五部分是收集被访者对汽车共享的了解情况。

1.3 样本描述

问卷采取的是网络调查与纸质问卷调查相结合的方式.网络调查部分:将设计好的问卷放在问卷星网站上,通过邮件、QQ、MSN、微博及电话等方式邀请同学、同事、朋友、网友以及通过他们转发来完成.纸质问卷调查部分:在上海市一些住宅小区内发放问卷完成.本次问卷调查总共有250位被调查者,有效问卷213份,被调查对象的个人特征分布如表1所示。

1.4 Logit 模型

通过对汽车共享服务和拥有私家车两者的要素分析,希望能分析出在何种情况下,消费者会选择汽车共享,在何种情况下消费者会选择拥有私家车.由于McFadden提出的离散选择模型应用广泛^[5-6],本文采用该模型。

表1 个人特征部分调查结果

Tab.1 Personal characteristics of the survey results

分类	选项	人数	比例/%
性别	男	118	55.4
	女	95	44.6
年龄	20~30岁	114	53.5
	30~40岁	84	39.4
	40~50岁	13	6.1
	50~60岁	2	1.0
	高中以下(含高中)	5	2.3
	大专	10	4.7
受教育程度	本科	124	58.2
	硕士	69	32.4
	博士	5	2.4

假设 a. Logit 模型的因变量 y_i 是二分变量,它只能取值 0,1;

b. Logit 模型中的因变量和各自变量之间的关系是非线性的。

通常将事情分为可控部分和不可控部分,因此,二分变量的回归方程为

$$y_i = \alpha + \sum_{i=1}^k \beta_i x_i + \epsilon_i$$

式中, α, β_i 为参数; x_i 为自变量; y_i 为二分变量; ϵ_i 为误差项,假设 ϵ_i 服从 Logistic 分布,分布函数

$$F(\epsilon) = \frac{1}{1 + e^{-\frac{\epsilon - \mu}{\sigma}}}, \text{ 其中, } \mu, \sigma \text{ 分别为 Logistic 分布}$$

的位置参数和尺度参数^[7]。

将事件发生的概率标注为 $p(y_i = 1 | x_i) = p_i$,可以计算出事件发生的概率为

$$p_i = e^{\alpha + \sum_{i=1}^k \beta_i x_i} / (1 + e^{\alpha + \sum_{i=1}^k \beta_i x_i})$$

这就是著名的 Logit 模型。

2 SPSS 软件运算分析

问卷调查的目的是得到2个Logit模型:a.放弃私家车而选择汽车共享出行模式的Logit模型;b.无车一族选择汽车共享出行模式的Logit模型.因此,分割被调查对象的一个最重要因素就是:有车与否。

2.1 无车一族选择汽车共享出行模式的 Logit 模型

在本次问卷调查中,无车一族有125人,占58.69%。将已获得数据输入SPSS18.0,选择二元Logistic回归分析,输出结果如下:

本次SPSS操作共处理了125个案例,在参加问卷调查后愿意尝试汽车共享这种出行方式的共有77人,占62.4%。在SPSS中,默认将二分类因变量中出现次数较多的赋值为1,本文将“之后不愿意尝试汽车共享”赋为0,“之后愿意尝试汽车共享”赋为1。

从表2中的显著性指标取值可以知道,除了单独

纳入 Teenager(家庭中未成年人的个数)的模型没有统计意义之外,其余模型都有显著的统计学意义.

表 2 不在方程中的变量 1
Tab.2 Variables not in the equation 1

步骤	变量	得分	自由度	显著性指标
步骤 0	性别	4.188	1	0.041
	年龄	7.804	1	0.005
	受教育程度	10.704	1	0.001
	有驾照人数	7.868	1	0.005
	无驾照人数	4.828	1	0.028
	未成年人数	0.533	1	0.465
	年收入	7.077	1	0.008
	从家到地铁站的时间	9.930	1	0.002
	每周工作天数	6.335	1	0.012
	费用低	9.647	1	0.002
总统计量		18.198	10	0.052

表 3 表示的是将上述变量纳入模型后模型的全局检验结果,共采用了 3 种检验方法,分别是步与步间的相对似然比检验(步骤),块之间的相对似然比检验(块)和模型间的相对似然比检验(模型).本例采取强行一次进入,一次将所有变量纳入模型,所以,3 种检验方法的结果是一致的,模型具有显著的统计学意义.

表 3 模型系数的综合检验 1
Tab.3 Comprehensive test of model coefficients 1

步骤	检验方法	卡方	df	显著性指标
步骤 1	步骤	20.053	10	0.029
	块	20.053	10	0.029
	模型	20.053	10	0.029

表 4 主要给出 - 2 对数似然值和两类决定系数, - 2 对数似然值越小,两类决定系数越接近 1,说明模型的拟合程度越好.

表 4 模型汇总
Tab.4 Model summary

步骤	- 2 对数似然值	Cox & Snell R方	Nagelkerke R方
步骤 1	85.147	0.382	0.563

当自变量数量增加时,尤其是连续自变量纳入模型后,皮尔逊卡方不再适用于估计拟合优度,此时引进 Hosmer 和 Lemeshow (HL)研究的对 Logistic 回归模型拟合优度的检验方法,如表 5 所示.

HL 检验根据预测概率值将数据分为大致相同规模的 10 个组,随机进行分析,预测值越接近观测值,说明模型拟合得越好.通过数据可以再次验证,此次模型拟合得不算好,但是,也具有一定的代表性.

表 6 是将变量全部纳入模型后模型的分

类预测值,此时模型预测的准确度只有 79.2%.从案例残差分析来看,只有 2 个案例(54 号和 58 号)存在异常,剩余的 123 个案例均在正态分布之内,无异常.

表 7(见下页)是 Logistic 模型的拟合结果.

表 5 随机性表
Tab.5 Randomness table

步骤 1	无意愿		有意愿		总计
	已观测	期望值	已观测	期望值	
1	10	7.775	3	5.225	13
2	7	6.914	6	6.086	13
3	5	6.258	8	6.742	13
4	6	5.412	7	7.588	13
5	5	4.882	8	8.118	13
6	2	4.369	11	8.631	13
7	5	3.953	8	9.047	13
8	3	3.370	10	9.630	13
9	3	2.621	10	10.379	13
10	1	0.688	7	7.312	8

表 6 分类表 1
Tab.6 Classification table 1

	已观测	已预测		百分率校正
		无意愿	有意愿	
步骤 1	无意愿	19	28	68.1
	有意愿	11	67	85.9
总计百分率				79.2

2.2 放弃私家车而选择汽车共享出行模式的 Logit 模型

对于已有私家车的被调查者的信息,用同样的方式带入 SPSS 运算,输出结果如下:

对于已有私家车的被调查者数据共有 88 个,占 41.3%.在 Logistic 模型运算中,系统自动赋值 1 给“之后愿意尝试汽车共享”的人群,赋值 0 给“之后不愿意尝试汽车共享”的人群,在本次调查中,之后愿意尝试汽车共享出行模式的人占 55.7%.

在进行无车一族 Logistic 回归分析时,由于各个因素对于最后是否愿意尝试汽车共享的影响不是很大,如果初始假设参加汽车共享的概率为 0.5,通过本次调查所得到的数据来检验,预测的准确度为 62.4%,因此,Logistic 回归将会失去意义,所以,在无车一族的 Logistic 回归分析时,剔除了常数的影响,直接在各个自变量中寻找某种联系.

在分析已有私家车这类人群,选择汽车共享这种出行方式时,常数本身对模型的影响不大.如表 8 所示(见下页),系统自动赋予常量 0.228,显著性指标的取值为 0.287,模型不够显著.如表 9 所示(见下页).从不在模型中变量的显著性指标取值可以知道,有些变量对于模型是没有统计意义的,在此选用向后剔除法,即开始将所有变量纳入模型中,每次将

最不显著的一个变量剔除,直到模型中剩余变量都 足够显著.

表 7 方程中的变量 1

Tab.7 Variables in the equation 1

步骤	变量	变量的系数值	标准误差	卡方	自由度	显著性指标	期望值
步骤 1	性别	0.032	0.392	0.007	1	0.934	1.033
	年龄	0.005	0.032	0.026	1	0.871	1.005
	受教育程度	0.512	0.265	3.713	1	0.054	1.668
	有驾照人数	0.268	0.239	1.255	1	0.263	1.308
	无驾照人数	0.039	0.187	0.043	1	0.836	1.039
	未成年人数	-0.414	0.489	0.717	1	0.397	0.661
	年收入	0.001	0.061	0	1	0.991	1.001
	从家到地铁站的时间	-0.436	0.211	4.252	1	0.039	0.647
	每周工作天数	0.039	0.028	1.933	1	0.164	1.040
	费用低	0.198	0.200	0.987	1	0.321	1.220

在表 10 中,由于在这里采用的是向后剔除法,先一次性将所有变量纳入模型,再逐次剔除最不显著的变量,所以,3种检验方法的结果不大一致,步

表 8 方程中的变量 2

Tab.8 Variables in the equation 2

步骤 0	变量的系数值	标准误差	卡方	自由度	显著性指标	期望值
常量	0.228	0.215	1.131	1	0.287	1.256

表 9 不在方程中的变量 2

Tab.9 Variables not in the equation 2

步骤	变量	得分	自由度	显著性指标
步骤 0	性别	0.483	1	0.487
	年龄	0.031	1	0.861
	受教育程度	0.189	1	0.664
	是否有驾照	1.629	1	0.202
	有驾照人数	0.129	1	0.719
	无驾照人数	0.123	1	0.726
	未成年人数	2.74	1	0.098
	年收入	0.017	1	0.897
	从家到地铁站的时间	0.063	1	0.802
	每月行驶里程数	0.565	1	0.452
	费用	2.435	1	0.119
	是否听过汽车共享	1.106	1	0.293
	费用低	4.289	1	0.038
总统计量		17.256	13	0.188

表 10 模型系数的综合检验 2

Tab.10 Comprehensive test of model coefficients 2

步骤	检验方法	卡方	自由度	显著性指标
步骤 1	步骤	15.415	12	0.220
	块	15.415	12	0.220
	模型	15.415	12	0.220
步骤 4	步骤	-0.261	1	0.609
	块	14.796	9	0.097
	模型	14.796	9	0.097

与步间的相对似然比检验显示模型的统计学意义不够显著,但是,块间的相对似然比检验和模型间的相对似然比检验结果显示模型是越来越显著的.在表 11 和表 12 中,HL 检验显示,随着变量的减少,模型的拟合度下降比较小,在步骤 4 结束时,模型的拟合度还是不错的.

表 11 Hosmer 和 Lemeshow 检验

Tab.11 Hosmer and Lemeshow test

步骤	卡方	自由度	显著性指标
步骤 1	5.379	8	0.826
步骤 4	4.703	8	0.789

表 12 Hosmer 和 Lemeshow 检验的随机性表

Tab.12 Randomness table of the Hosmer and Lemeshow tests

步骤	无意愿		有意愿		总计	
	已观测	期望值	已观测	期望值		
步骤 1	1	7	7.237	2	1.763	9
	2	7	5.922	2	3.078	9
	3	5	5.039	4	3.961	9
	4	3	4.640	6	4.36	9
	5	4	4.248	5	4.752	9
	6	3	3.652	6	5.348	9
	7	5	3.111	4	5.889	9
	8	3	2.405	6	6.595	9
	9	1	1.987	8	7.013	9
	10	1	0.759	8	6.241	7
步骤 4	1	7	7.208	2	1.792	9
	2	6	5.913	3	3.087	9
	3	5	5.074	4	3.926	9
	4	4	4.537	5	4.463	9
	5	4	4.197	5	4.803	9
	6	4	3.629	5	5.371	9
	7	3	3.137	6	5.863	9
	8	3	2.521	6	6.479	9
	9	2	1.955	7	7.045	9
	10	1	0.830	6	6.170	7

如表 13 和表 14 所示,在未剔除任何变量时,模型的准确度达到86.4%,但是,当模型进行到步骤 4,剔除了 3 个变量后,模型的准确度降低为 83%.虽然模型的准确度有些许降低,但是,其引起的误差还是在可接受范围内.进行到步骤 4 时,模型中被剔除的变量为:受教育程度、家庭中未持有驾照的成年

表 13 分类表 2
Tab.13 Classification table 2

已观测	已预测			
		无意愿	有意愿	百分率校正
步骤 1	无意愿	30	9	76.9
	有意愿	3	46	93.9
总计百分率				86.4
步骤 4	无意愿	29	10	74.7
	有意愿	5	44	89.5
总计百分率				83.0

人数和从家步行到最近地铁站的时间.

表 14 不在方程中的变量 3
Tab.14 Variables not in the equation 3

步骤	变量	得分	自由度	显著性指标
步骤 4	受教育程度	0.262	1	0.609
	无驾照人数	0.261	1	0.609
	从家到地铁站的时间	0.064	1	0.801
总统计量		0.618	3	0.892

在这组数据的运算中,不存在残差案例,所有的案例均落在正态分布内.从表 15 可以清楚地看出每个变量的系数以及其对是否选择汽车共享出行模式的影响程度.从表 15 中可以看到,已有私家车的潜在顾客群在是否选择汽车共享出行模式时,首先考虑的是成本问题,其次考虑的是家庭中未成年人数.

表 15 方程中的变量 3
Tab.15 Variables in the equation 3

步骤	变量	变量的系数值	标准误差	卡方	自由度	显著性指标	期望值
步骤 1	性别	- 0.678	0.597	1.290	1	0.256	0.507
	年龄	- 0.081	0.057	2.029	1	0.154	0.922
	受教育程度	- 0.240	0.452	0.282	1	0.596	0.787
	有驾照人数	0.459	0.439	1.094	1	0.296	1.583
	无驾照人数	0.182	0.340	0.286	1	0.593	1.199
	未成年人数	1.055	0.567	3.463	1	0.063	2.871
	年收入	- 0.026	0.037	0.507	1	0.476	0.974
	从家到地铁站的时间	- 0.010	0.034	0.079	1	0.779	0.991
	每月行驶里程数	0	0	0.733	1	0.392	1
	费用	0	0	0.880	1	0.348	1
	是否听过汽车共享	1.217	0.727	2.800	1	0.094	3.378
	费用低	0.694	0.316	4.819	1	0.028	2.001
	常量	- 1.129	2.354	0.230	1	0.632	0.323
步骤 4	性别	- 0.745	0.577	1.668	1	0.197	0.475
	年龄	- 0.081	0.056	2.128	1	0.145	0.922
	有驾照人数	0.389	0.042	0.939	1	0.333	1.476
	未成年人数	1.029	0.556	3.424	1	0.064	2.798
	年收入	- 0.040	0.030	1.780	1	0.182	0.961
	每月行驶里程数	0	0	0.562	1	0.454	1
	费用	0	0	1.455	1	0.228	1
	是否听过汽车共享	1.196	0.715	2.799	1	0.094	3.306
	费用低	0.657	0.308	4.550	1	0.033	1.930
	常量	- 1.268	1.852	0.469	1	0.494	0.281

3 结果分析

国外的诸多研究都表明,汽车共享的主要客户是 20~40 岁之间的青少年,这类人的通勤出行、购物出行需求较高,因此,问卷调查的主要对象集中在 20~40 岁之间的青少年群体.

通过本次问卷调查可以发现,在汽车共享不够发达的中国,对于出行方式的选择,大部分被调查者选择的是公共交通,占到 70.3%,而私家车出行只占到 10.9%.通过对私家车拥有的调查,很清楚可以看到,有 41.3% 的被调查者拥有私家车,由此可见私家车拥有者中有 73.6% 的人平时的出行仍然不选择私家车,而是选择更具经济性的公共交通.

在调查人们选择出行方式考虑的主要因素时,有83.6%的被调查者选择了经济与快捷.在对私家车的调查中,被调查者普遍认为私家车能够提供出行更强的便捷性、舒适性、灵活性,但是,在私家车使用过程中,寻找停车位是一件很困难的事情.此外,私家车的使用成本高、带来严重的环境污染,而且,被调查者都不认为拥有私家车是一种社会地位的表现,因此,在消费者不认为私家车是社会地位的表现的前提下,汽车共享如果真正实行,将减少道路车辆,缓解道路拥堵,解决停车难的问题,使出行更加便捷;另一方面,汽车共享对比私家车,有使用费更低的优势,从这个角度出发,又能满足消费者对于经济性的要求.

对上海市场的个人消费者而言,汽车共享更能吸引那些目前还没有私家车,但有用车需求的人群,特别是已经计划购买私家车的人群,这一点和国外的研究结果非常接近,因此,这个阶段的营销重点应该在于那些有购车计划的人群.

经过调查,被调查者在考虑选择汽车共享时,汽车共享的价格是一个最主要的影响因素,取车、还车的手续简便也是一个重要因素,汽车的新旧程度的影响稍弱.亲友是否使用,对于消费者选择汽车共享的影响很小,基本可以忽略.由此可以看出,要成功推广汽车共享,价格的制定一定要合理,这一点至关重要.

在对有车和无车两组被调查对象的数据进行Logistic回归分析后,发现20~40岁的青少年的收入水平比较稳定,对于是否选择汽车共享的影响作用不是很大.从Logit函数可以看出,目前青少年的收入水平处于S形曲线的尾部,以其现有的水平可以负担起汽车共享的费用,所以,收入的变动对于消费者是否选择汽车共享的影响不是很大.

在无车一族中,通过Logistic回归,可以清楚地发现最重要的两个因素是:受教育程度和从家步行至最近地铁站的时间.受教育程度越高的消费者,更加倾向尝试汽车共享.这一调查结果与国外许多研究结果一致,汽车共享企业在开拓市场时,对这一部分人群制定专门的营销计划是必不可少的.

在有车一族中,最重要的两个影响因素是:汽车共享的价格和家庭中未成年人数.汽车共享的价格对是

否选择汽车共享,这在经济学中关于需求有解释.关于家庭中的未成年人数,在已有未成年人的家庭中,出行的很大一部分原因是送孩子上学,在这种情况下,同一小区各个家庭的孩子可以共车去上学,从而减少各个家庭使用私家车的费用,这样既便捷又经济.

此次问卷调查取得的数据基本无残差案例,都落在正态分布内,能说明一定情况.各个自变量对于是否选择汽车共享的影响是合理的,但是,由于被调查者数量的限制,各自变量对于消费者是否选择汽车共享的影响程度可能存在误差,需要进一步研究.

4 结束语

基于上海市场消费者作问卷调查,进行Logit回归分析,探讨影响消费者选择汽车共享因素的重要性,但是依然存在很多不足之处.首先是由于时间关系,上海市场消费者调研的样本数量还很小;其次在调查问卷的设计上也存在一些缺陷,尚有一些对消费者选择汽车共享有显著影响的因素未纳入模型;第三,对于有车一族与无车一族选择汽车共享的Logit模型的拟合度有待于提高.

参考文献:

- [1] 夏凯旋,何明升.国外汽车共享服务的理论与实践[J].城市问题,2006(4):87-92.
- [2] 张小明.共享汽车——一种全新的消费模式[J].世界汽车,2001(7):21-22.
- [3] 薛跃,杨同宇,温素彬.汽车共享消费的发展模式及社会经济特性分析[J].技术经济与管理研究,2005(1):54-58.
- [4] 程伟力.北美小汽车共用的发展状况[J].交通与运输,2007(增刊):34-37.
- [5] McFadden D. Conditional Logit analysis of qualitative choice behavior [M] // Zarembka P. Frontiers in Econometrics. New York: Academic Press, 1974.
- [6] 迈克尔·弗洛里安,吴稼豪,贺曙光.多层Logit结构多模式变量需求网络平衡模型[J].上海理工大学学报1999,21(3):191-202.
- [7] 王济川,郭志刚. Logistic 回归模型——方法与应用[M].北京:高等教育出版社,2001.

(编辑:石 瑛)