

doi:10.16055/j.issn.1672-058X.2015.0002.002

个人特质与大学生消费水平关系的研究*

——基于 Probit 模型和 ANOVA 模型

项舒畅¹, 王雅静², 祖璇¹, 毛军军^{1,3**}

(1. 安徽大学 数学科学学院, 合肥 230601; 2. 安徽大学 商学院, 合肥 230601;
3. 安徽大学 计算智能与信号处理教育部重点实验室, 合肥 230039)

摘要:在 512 份安徽大学大学生消费的调查问卷的基础上,先用有序 Probit 回归模型分析个人特质对大学生消费水平的影响;再用方差分析方法分析个人特质的不同水平对于消费支出的影响;结果显示,大学生所在的年级、生源地、家庭收入状况以及大学生的理财观念对大学生消费水平影响显著,大学生的性别、专业类型和兄弟姐妹个数影响不显著,而家庭收入对于大学生的消费起到决定性影响。

关键词:大学生;个人特质;消费水平;Probit 模型;ANOVA

中图分类号: O176 **文献标识码:** A **文章编号:** 1672-058X(2015)02-0005-06

郭奉元(1998)指出在众多大学生消费的影响因素中,大学生个人特质对大学生消费水平的影响是非常深刻的^[1]。曾本君(2008)指出,大学生所在的年级对大学生的月总消费、上网消费、人际消费等方面均有显著影响,并指出大学生总消费基本上随年纪呈现上升趋势^[2]。Li Sanshan 和 Li Yongjian(2010)^[3]、Zhao Yingwu(2004)^[4]分别举例分析了心理因素对于消费的影响,梁前德(2010)研究了大学生家庭特征对大学生消费水平的影响,指出大学生家庭收入对大学生消费水平和消费结构的影响是决定性的^[5]。张志辉(2002)研究了大学生生源地对其消费水平的影响,指出来自城市的大学生消费水平一般要高于农村大学生^[7]。

现有研究成果为研究提供了坚实的理论基础,但也存在明显的不足。前人普遍采用统计描述的方法对大学生消费水平的影响因素进行归纳,即使采用计量分析的方法,多数研究将因变量简单地判断为“是”“否”,导致样本信息不同程度的丢失,进而影响了结论的科学性。鉴于已有研究成果的不足,在 512 份安徽大学学生消费调查数据的基础上,参考王玉枝^[8]、胡雪梅^[9]等的研究方法,建立有序 Probit 模型,探究大学生个人特质对其消费水平的影响。另外,将消费的影响因素自然地划分为不同水平,对每个因素进行单因素方差分析(ANOVA),对有序 Probit 模型的结果进行佐证。

1 调研设计

调查以安徽大学在校的本科生为总体,采用多阶段分层抽样选取样本。方法是:先按照随机抽样在高

收稿日期:2014-05-10;修回日期:2014-06-25.

* 基金项目:国家自然科学基金(71371011);安徽省高等学校省级自然科学研究重点项目(KJ2013A033);大学生创新创业训练计划项目(201310357118).

作者简介:项舒畅(1992-),男,安徽舒城人,从事数理统计和计量经济研究.

** 通讯作者:毛军军(1973-),女,浙江杭州人,副教授,博士,从事多属性决策理论及方法研究.E-mail:maojunjun@ahu.edu.cn.

校中抽取 18 个院或独立系,然后以年级进行分层,从所抽到的院系中每个年级各抽 2 个班,每班学号前 4 名的学生纳入本次调查的样本。调查共发放问卷 600 份,其中有效问卷 512 份,有效回收率为 85.3%。问卷调查的主要内容包括大学生的个人及家庭特征、大学生消费水平及消费结构状况、大学生的消费理念等 3 个方面内容。调查样本的基本信息及各变量不同水平的占比见表 1。

表 1 大学生基本信息

项 目	类 别	样本数	比例/%	项 目	类 别	样本数	比例/%
性 别	男	283	55.3	生源地	城镇	200	39.1
	女	229	44.7		农村	312	60.9
年 级	一年级	115	22.5	专业类型	理工科	280	54.7
	二年级	152	29.7		文科	232	45.3
	三年级	133	26.0	独生子女	是	136	26.6
	四年级	112	21.9		否	376	73.4

2 变量设定与模型选择

2.1 变量设定

在参考前人研究成果和调查问卷数据可得性的基础上,因变量选取大学生总体的消费水平。为了后文的有序 Probit 回归和方差分析的变量设置,将大学生消费水平共分为 5 个档次,分别为 600 元以下,600~900,900~1 200,1 200~1 500 和 1 500 元以上。自变量从大学生的个人特质进行构建,具体包括大学生性别(X_1)、年级(X_2)、生源地(X_3)、专业类型(X_4)、家庭兄弟姐妹个数(X_5)、家庭月收入状况(X_6)和大学生理财观念(X_7)7 个指标。而大学生理财观念用大学生是否有记账习惯进行替代。具体的变量名称、相关说明及变量的统计描述见表 2。

表 2 变量设置

变 量	变量含义及赋值	均 值	方 差
大学生消费水平(Y) [*]	1-2-3-4-5 五个档次	927.46	119 833.00
性别(X_1)	1=男,0=女	0.55	0.247
年级(X_2)	1=大一,2=大二,3=大三,4=大四	2.47	1.136
生源地(X_3)	1=城镇,0=乡村	0.39	0.238
专业类型(X_4)	1=理工科,0=文科	0.55	0.248
家庭兄弟姐妹个数(X_5)	最小值 0—3 最大值	0.38	0.235
家庭月收入状况(X_6)	1-2-3-4-5 五个档次	5 239.26	13 696 212.73
是否有记账习惯(X_7)	1=有,0=没有	0.386 1	0.237 0

2.2 模型选择

2.2.1 有序 Probit 模型

根据调查问卷的设计,确定大学生消费水平为顺序关系, Y_i 为有序变量,适合运用有序 Probit 模型进行分析。其结构线性形式如下:

$$Y^* = X'_i \beta + \varepsilon_i \quad i = 1, 2, \dots, N \quad (1)$$

式(1)中 i 为样本序号, β 为参数向量,是待求的一组参数, X_i 是自变量矢量,表示可能影响大学生消费

水平的一组解释变量的观测值, ε_i 为随机误差项。 Y^* 表示大学生消费水平的潜变量, 是一个不可观测的变量, 通过可观测的 Y_i 与 Y^* 之间的关系来代表。可观测变量 Y_i 与被解释变量 Y^* 存在的关系如下:

$$Y_i = - \begin{cases} 1(600 \text{ 以下}), & Y^* < \alpha_1 \\ 2(600 \sim 900), & \alpha_1 \leq Y^* < \alpha_2 \\ 3(900 \sim 1\,200), & \alpha_2 \leq Y^* < \alpha_3 \\ 4(1\,200 \sim 1\,500), & \alpha_3 \leq Y^* < \alpha_4 \\ 5(1\,500 \text{ 以上}), & Y^* \geq \alpha_4 \end{cases} \quad (2)$$

式(2)中 Y_i 是离散变量, 且取值为(1, 2, 3, 4, 5), 表示第 i 个样本消费水平。 α_i 是一组新参数, 是决定样本组别的分界线, 有 $\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4$ 。 Y^* 就被划分为 5 个互不重叠的区间, Y_i 表示某个具体的观察值落到了哪个区间, Y_i 取到一特定值 j 的概率为

$$P(Y_i = j) = P(\alpha_{j-1} \leq Y^* < \alpha_j) = P(\alpha_{j-1} - X'_i \beta \leq \varepsilon_i < \alpha_j - X'_i \beta) = F(\alpha_j - X'_i \beta) - F(\alpha_{j-1} - X'_i \beta) \quad (3)$$

式(3)中 F 为 ε_i 的累积分布函数, 其中 $1 \leq j \leq 5$ 。现假设误差项 ε_i 服从标准正态分布, 则 F 满足标准正态分布累积函数的条件, 有:

$$\varepsilon_i / X_i \sim (0, \delta^2) \quad (4)$$

如此把大学生消费水平观察值 Y_i 作为被解释变量, 建立标准有序 Probit 模型, 其对数似然函数为

$$\ln L = \sum_{i=1}^n \sum_{j=1}^5 Y_{ij} \ln [\varphi(\alpha_j - X'_i \beta) - \varphi(\alpha_{j-1} - X'_i \beta)] \quad (5)$$

式(5)中, φ 为标准正态分布的累积函数。通过最大化对数似然函数式, 即可估计出有序 Probit 模型中的系数 β 和参 α_i 数。估计所得的 β 值就是“大学生消费水平的”影响因素系数。

2.2.2 方差分析模型

为了研究 7 个解释变量各自在不同水平下对大学生消费的不同影响, 采用了单因素方差分析模型。课题将各变量按照表 2 进行划分, 而各变量的不同水平占比按表 1 进行划分, 单因素方差分析模型如下:

$$\begin{cases} x_{ij} = \mu + \alpha_i + \varepsilon_{ij}, & i = 1, 2, \dots, r, \quad j = 1, 2, \dots, n_i \\ \varepsilon_{ij} \sim N(0, \sigma^2), & \text{相互独立} \\ \sum_{i=1}^r n_i \alpha_i = 0 \end{cases} \quad (6)$$

将水平 $A_i (i=1, 2, \dots, r)$ 下的实验结果 $x_{i1}, x_{i2}, \dots, x_{in}$ 看做来自第 i 个正态总体 $X_i \sim N(\mu_i, \sigma^2)$ 的样本观测值, 每个总体 X_i 相互独立, ε_{ij} 是相应的实验误差。比较因素 A 的 r 个总体的均值, 即检验:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_r, \quad H_0: \mu_1, \mu_2, \dots, \mu_r \text{ 不全相等}$$

这里 $\mu = \frac{1}{n} \sum_{i=1}^r n_i \mu_i$, $n = \sum_{i=1}^r n_i$, $\alpha_i = \mu_i - \mu$, μ 表示总和的均值, α_i 表示水平 A_i 对指标的效应。

家庭收入 (X_6) 的划分标准如下:

$$X_{6i} = - \begin{cases} 1(\text{很低}) & X_6^* \leq 1\,000 \\ 2(\text{较低}) & 1\,000 < X_6^* \leq 3\,000 \\ 3(\text{正常}) & 3\,000 < X_6^* \leq 5\,000 \\ 4(\text{较高}) & 5\,000 < X_6^* \leq 8\,000 \\ 5(\text{很高}) & X_6^* > 8\,000 \end{cases} \quad (7)$$

式(7)中 X_6^* 表示家庭收入的潜变量, X_{6i} 是家庭收入划分等级。

3 实证分析

对于有序 Probit 模型,研究运用 R 软件对模型进行了估计,模型 1 分析了大学生个人特征变量对大学生消费水平的影响。结果表明,在模型 1 中,部分解释变量不显著。因此,在模型 1 回归结果的基础上,剔除不显著的变量,进行二次回归估计,得到优化模型 2。具体回归结果见表 3。对于方差分析模型,研究运用 R 软件,先对每个变量进行单因素方差分析,得到模型一;在模型一的基础上,再进行 Kruskal-Wallis 秩和检验,得到模型二,结果见表 4。

表 3 Probit 回归结果

变 量	模型 1		模型 2	
	回归系数	P-value	回归系数	P-value
性别(X_1)	-0.402 398	0.118 1		
年级(X_2)	0.144 667**	0.049 8	0.123 452*	0.053 1
生源地(X_3)	0.063 414*	0.051 2	0.084 472**	0.039 7
专业类型(X_4)	0.055 843	0.741 1		
家庭兄弟姐妹个数(X_5)	-0.020 875*	0.099 0	-0.010 779	0.102 1
家庭月收入状况(X_6)	0.507 186***	0.000 0	0.623 127***	0.000 29
是否有记账习惯(X_7)	0.126 305 2***	0.001 6	0.139 844 1***	0.002 7

注:***、**、* 分别表示在 1%、5%和 10%的显著性水平下显著

表 4 方差分析结果

变 量	模型 1		模型 2	
	F 值	P-value	K-WF 卡方值	P-value
性别(X_1)	0.279 0	0.597 0	1.813 2	0.178 1
年级(X_2)	3.222**	0.022 4	10.661 2**	0.013 7
生源地(X_3)	29.610***	0.000 0	25.601 5***	0.000 0
专业类型(X_4)	3.304 0*	0.069 7	3.150 4*	0.075 9
家庭兄弟姐妹个数(X_5)	0.494 0	0.483 0	0.055 1	0.814 4
家庭月收入状况(X_6)	39.680***	0.000 0	112.797 8***	0.000 0
是否有记账习惯(X_7)	15.75***	0.000 14	11.222***	0.000 8

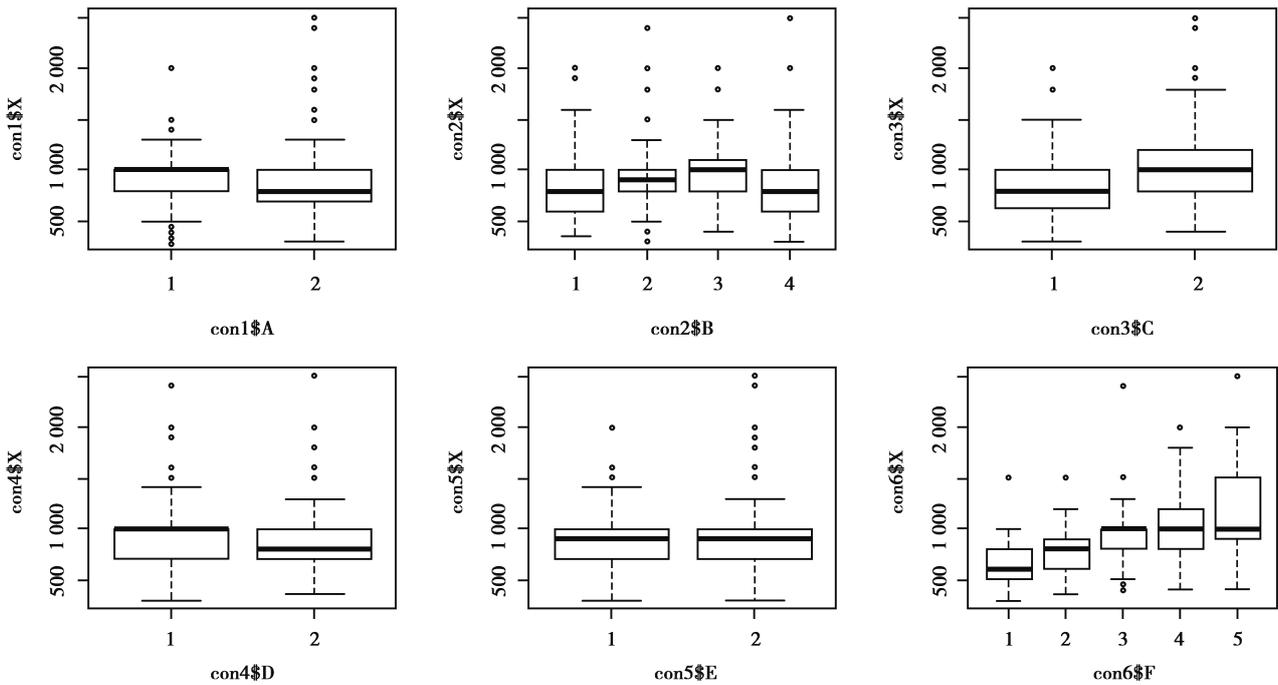
注:***、**、* 分别表示在 1%、5%和 10%的显著性水平下显著

对于前 6 个解释变量在如上不同水平下画出了箱线图,便于直接观察不同水平下消费均值的差异,见图 1。对于存在显著性差异且水平个数大于等于 3 的变量 X_2, X_6 进行均值多重比较,计算得到如下的均值多重 t 检验的 P 值表 5、6:

表 5 年级(X_2) 均值多重 t 检验结果

年级 X_2	消费均值/元	P_{ij}			
大一	865.304 3	1.000 0	0.026**	0.010***	0.648
大二	960.197 4	0.026**	1.000 0	0.653	0.085*
大三	978.571 4	0.010***	0.653	1.000 0	0.037**
大四	886.160 7	0.648	0.085*	0.037**	1.000 0

注:***、**、* 分别表示在 1%、5%和 10%的显著性水平下显著



注: 横坐标代表各变量 ($X_1 \sim X_6$) 的不同水平, 纵坐标代表消费支出金额

图 1 六变量不同水平下的箱线图

表 6 家庭收入 (X_6) 均值多重 t 检验结果

家庭收入 X_6	消费均值/元	P_{ij}				
1	657.500 0	1.000 0	0.177 52	0.000 35***	0.000 0***	0.000 0***
2	754.935 1	0.177 52	1.000 0	0.652 3	0.000 0***	0.000 0***
3	915.254 2	0.000 35***	0.652 3	1.000 0	0.000 25***	0.000 0***
4	1 059.890 1	0.000 0***	0.000 0***	0.000 25***	1.000 0	0.000 17***
5	1 242.857 1	0.000 0***	0.000 0***	0.000 0***	0.000 17***	1.000 0

注: ***、**、* 分别表示在 1%、5% 和 10% 的显著性水平下显著

表 3 的回归结果显示,两个模型都比较好的通过了异方差、方程显著性等基本的计量经济学检验。表 4 的方差分析结果显示,无论是模型一单因素方差分析还是模型二 Kruskal-Wallis 秩和检验,结论是一致的,与方差分析结论也大致相同。进一步分析表 3、4,可以得到结论。

大学生的年级 (X_2) 在 Probit 回归分析中模型一和模型二中均通过显著性检验,影响系数为正。在方差分析中对消费支出也有显著影响。此结论与曾本均 (2008)^[3] 的研究结论相一致。大学生总体消费支出水平随着年级的提高而呈现逐年上升趋势。这是因为大一新生刚进入大学,尚没有形成固定的朋友圈,学生的交际消费尚不明显。而随着一到两年的相处,大学生基本形成了自己的交际圈,交际消费也会逐步上升。

生源地 (X_3) 和家庭月收入状况 (X_6) 在 Probit 回归分析中模型一和模型二及在方差分析中模型一和模型二对消费支出都有显著影响。印证了梁前德 (2010)^[4] 和赵凯 (2002)^[5] 的研究结论。由于多数大学生的经济尚未独立,其资金来源主要从家庭获取,家庭的收入状况对大学生的消费水平起到决定性的作用。由于城乡二元经济结构的存在,农村经济明显落后于城市经济,农村家庭的收入水平一般也显著落后于城镇家庭。因此,农村家庭的大学生消费水平一般也低于城镇家庭的大学生。

从 Probit 模型和方差分析模型看是否有记账习惯 (X_7) 对于大学生消费都有显著影响。反应了大学生

的理财意识和理财能力。有记账习惯的大学生可以清楚的知道自己的每天花销的数额和花销方式。因此,他们节约消费的意愿也更强。因此,他们的消费水平整体上要低于那些没有记账习惯的大学生。

Probit 模型和方差分析模型都可以看出大学生的性别(X_1)和专业类型(X_4)对大学生消费水平影响不显著。这主要是因为当前无论是男生还是女生,无论是理工科大学生还是文科大学生,其消费均呈现出超前消费、盲目消费和攀比消费特点。因此大学生的消费水平在专业类型上并没有显著的差异性。家庭兄弟姐妹个数(X_5)对大学生消费水平的影响在模型一中显著而在模型二影响变为不显著。这是因为,当前随着中国经济的快速发展,居民的收入水平均显著提高,家庭也有能力供养多个大学生的消费。

进一步分析表 5、6,可以得到具体哪些水平之间有影响。从表 5 年级(X_2)的均值多重 t 检验,看出大一和大二,大一和大三,大三和大四之间消费的差异性很明显,进一步说明了大二的消费水平相对于其他年级的消费较低。表 6 是不同家庭收入(X_6)的均值多重 t 检验。结果显示:家庭收入很低、较低、正常、较高、很高之间对于消费几乎全部都有差异性,这进一步说明了家庭收入的多少对于消费多少是一个最显著影响因素。

4 结论性评述

正确引导大学生合理消费对树立大学生科学消费观念,培养其勤俭节约意识具有重要意义。在 512 份安徽大学大学生调查问卷的基础上,运用有序 Probit 模型和方差分析方法探析大学生的个人特质对大学生消费水平的影响。结果表明:大学生所在的年级、生源地、家庭收入情况以及大学生是否有理财意识和能力对大学生的消费水平具有显著影响,而大学生的性别和专业类型在大学生消费上没有明显差异,大学生兄弟姐妹个数在 Probit 模型一中有影响显著,而在 Probit 优化模型中系数变为不显著,表明随着中国经济的快速发展,居民的收入水平均显著提高,家庭有能力供养多个子女。特别的,在从均值多重 t 检验得到了家庭的收入多少是大学生消费多少的最显著性因素。

基于以上分析,大学生要培养科学的消费观念,首先应遵守量入为出原则,大学生应尽量避免随大流、盲目攀比高消费,应注意减少随意性消费,抵御奢侈品的诱惑。其次应把握消费的统筹规划,即根据自己实际成长发展需要安排合适的消费结构,并对自己的消费情况定时进行评估。再次,购物时遵从经济适用原则。在购物时,多向有利于提高自己身心健康的方面投资,并注重投资效益。

参考文献:

- [1] 郭奉元. 影响大学生消费的个人因素分析[J]. 市场与人口分析,1998(3):33-36
- [2] 曾本君. 大学生消费行为与个人价值观关系的实证研究[J]. 宜宾学院学报,2008(4):122-124
- [3] LI S, LI Y. An Exploration of the Psychological Factors Influencing College Students Consumption of Mobile Phone in West China [J]. International Journal of Business and Management,2010,28: 132-138
- [4] ZHAO Y. Influence of Consumption Psychology on Users' Purchasing Decision and Marketing Countermeasures [J]. Science of Social Psychology,2004,21:36-40
- [5] 梁前德. 父亲从业状况与大学生消费的实证分析—基于武汉地区 2662 名大学生消费的调查[J]. 教育与经济,2010(4):23-30
- [6] CLAEYS C, SWINNEN A, VANDEN A P. Consumers Means-end Chains for "think" and "feel" Products [J]. International Journal of Research in Marketing,1995(3):13-20
- [7] 张志辉. 当代大学生消费的特征及趋势[J]. 中国青年研究,2002(5):33-36
- [8] 王玉枝,邢春峰,常广平. 分析交互作用方法的应用与比较[J]. 重庆工商大学学报:自然科学版,2010,27(1):36-38
- [9] 胡雪梅. 浅谈基于模拟的计量经济学方法[J]. 重庆工商大学学报:自然科学版,2014,31(2):74-76
- [10] 周丽. 影响大学生消费的因素分析[J]. 统计观察,2007(2):100-101