

中图分类号：

密级：公开

学科分类号：

论文编号：182214001

山东财经大学  
硕士学位论文  
(专业学位)

老年人保健消费影响因素分析

作者姓名：高丽

学科专业：应用统计硕士

指导教师：张艳丽（副教授）

培养学院：统计学院

二〇二〇年六月六日

# **Analysis of Influencing Factors on Elderly Health Care Consumption**

A Dissertation Submitted for the Degree of Master

**Candidate: Gao Li**

**Supervisor: Prof. Zhang Yanli**

School of Statistics

Shandong University of Finance and Economics

中图分类号：

密级：公开

学科分类号：

论文编号：182214001

## 硕士学位论文

# 老年人保健消费影响因素分析

作者姓名：高丽

申请学位级别：专业学位硕士

指导教师姓名：张艳丽

职 称：副教授

学 科 专 业：应用统计

研 究 方 向：计量经济

学 习 时 间：自 2018 年 9 月 1 日 起 至 2020 年 6 月 30 日 止

学位授予单位：山东财经大学

学位授予日期：2020 年 6 月



## 摘 要

随着经济的快速发展和居民生活水平的逐步提升,人均寿命延长,老龄人口比例增加,我国的老龄化时代已到来,老年人成为当前社会不可忽视的重要群体。由于老年人特殊的生理及心理状态,在其全部消费中,保健消费占据十分重要的位置。本文旨在研究老年人保健消费影响因素,对相关文献资料和研究成果进行整理和归纳,基于前人研究,经过充分分析与思考,构建老年人保健消费影响因素宏微观指标体系,于“中国健康与养老追踪调查”数据库和国家统计年鉴获取相关数据,对变量进行说明与描述性统计分析。观察数据结构可知,80%的受访老人保健消费支出金额为零,针对老年人保健消费金额零膨胀数据特征,基于两部模型原理将老年人保健消费行为分为保健消费决策与保健消费支出两部分相互独立的过程,分别探究其影响因素。同时考虑个人层面微观因素与地区层面宏观因素,建立分层线性模型探究其对老年人保健消费的影响。并构建广义部分线性可加模型深入探究老年人保健消费微观影响因素,更加全面、精准地测度了部分微观变量对老年人保健消费的非线性影响,得到了更具现实意义的结论。

通过构建分层线性模型与广义部分线性可加模型,得出人均地区生产总值、医疗保健类居民消费价格指数 2 个宏观因素以及年龄、性别、婚姻状况、受教育年限及其与人均地区生产总值交互项、退休状况、户口类型、自评健康、孤独感、子女实物支持 9 个微观因素对老年人保健消费决策存在显著线性影响;慢性病患情、抑郁程度、认知能力、家庭负债、子女看望次数、社交情况 6 个微观因素对保健消费决策存在显著非线性影响。居民人均可支配收入宏观因素以及年龄、工作状况、养老保险、户口类型、抑郁程度、家庭金融资产 6 个微观因素对老年人保健消费支出金额存在显著线性影响;受教育年限、子女实物支持、子女看望次数对保健消费支出存在显著非线性影响。本文的研究拓展了老年群体相关问题的研究领域,为老年人保健方面的研究提供了分析视角与实证材料;帮助老年人充分认识自身状况,对保健品与保健服务树立理性消费观;倡导子女多关心陪伴父母,给予老年人一定情感慰藉;指导保健品市场有针对性地对保健品与保健服务进行推广;呼吁社会有关部门关注老年人生理及心理需求,做好保健品理性消费观的宣传与普及,并加大保健品市场监管力度,促进保健消费市场的平稳健康发展。

**关键词：**老年人保健消费；影响因素；两部模型；分层线性模型；广义部分线性可加模型

## Abstract

With the economic development and the improvement of residents' living standards, the average life expectancy has been extended, and the proportion of the elderly population has increased. The era of aging has arrived in our country, and the seniors have become an important group that cannot be ignored in the current society. Health care consumption occupies a very important position in all consumption of the elderly because of their special physiological and psychological state. The purpose of this article is to study the influencing factors of elderly health care consumption. We summarize relevant literature and research results, and construct the macro and micro index system for influencing factors of elderly health care consumption after full analysis and thinking. We obtain relevant data from "China Health and Retirement Longitudinal Study" database and "National Bureau of Statistics", and perform a descriptive statistical analysis of the variables. The data structure shows that 80% of the elderly health care consumption expenditure is zero. Against the characteristics of the zero-inflation data, based on the two-part model principle, the health care consumption of the elderly is divided into two separate processes of health care consumption decision-making and health consumption expenditure, and their influencing factors are explored separately. Micro-factors at the individual level and macro-factors at the regional level are considered at the same time, and a hierarchical linear model is established to explore its impact on elderly health care consumption. A generalized partially linear additive model is constructed to further explore the micro-factors of the elderly care health consumption, and a more comprehensive and more realistic non-linear conclusion is obtained.

By constructing the hierarchical linear model and the generalized partially linear additive model we know that the macro-factors of GDP per capita, consumer price index for health care, as well as age, gender, marital status, years of education, retirement status, household registration type, self-assessed health, loneliness, and in-kind support of children have significant linear effects on health care consumption decisions for the elderly. Chronic conditions, degree of depression, cognitive ability, household debt, number of child visits, and social status have significant non-linear effects on health care consumption

decisions. The macro-factors of residents' per capita disposable income, as well as age, work status, pension insurance, household registration type, degree of depression and household financial assets have significant linear effects on health care expenditure for the elderly. Years of education, in-kind support of children and number of child visits have significant non-linear effects on health care expenditure. This article expands the research field about elderly issues, and provides analytical perspectives for research on elderly health. It helps the elderly fully understand their own conditions, and establish a rational consumption for health products and services. Children should pay more attention to accompany their parents and give the elderly some emotional comfort. And it guide the health care product market to promote products and services in a targeted manner. This article calls on relevant social departments to pay attention to the physical and psychological needs of the elderly, do a good job in propagating and popularizing the rational consumption concept of health products, and strengthen the supervision of the health care product market to promote the stable and healthy development of the health care consumption market.

**Key words:** Elderly Health Care Consumption; Influencing Factors; Two-part Model; Hierarchical Linear Model; Generalized Partially Linear Additive Model



第 1 章 绪论.....	1
1.1 研究背景及意义.....	1
1.1.1 研究背景.....	1
1.1.2 研究意义.....	2
1.2 国内外文献综述.....	2
1.2.1 老年人保健消费微观影响因素文献综述.....	2
1.2.2 老年人保健消费宏观影响因素文献综述.....	3
1.2.3 老年人保健消费影响因素研究方法文献综述.....	4
1.2.4 总体评价.....	6
1.3 研究方法与主要内容.....	6
1.3.1 研究方法.....	6
1.3.2 技术路线.....	6
1.3.3 主要研究内容.....	7
1.4 创新点.....	7
第 2 章 变量选取与描述性统计分析.....	8
2.1 数据来源及样本说明.....	8
2.2 指标选取及变量说明.....	9
2.2.1 被解释变量界定及说明.....	9
2.2.2 微观解释变量选取及说明.....	10
2.2.3 宏观解释变量选取及说明.....	14
2.3 老年人保健消费及其影响因素描述性统计分析.....	15
2.3.1 老年人保健消费情况描述性统计分析.....	15
2.3.2 老年人保健消费影响因素描述性统计分析.....	16
2.4 本章小结.....	19
第 3 章 老年人保健消费决策影响因素实证分析.....	20
3.1 相关模型介绍.....	20
3.1.1 两部模型思想.....	20
3.1.2 分层线性模型相关理论.....	21
3.1.3 广义部分线性可加模型相关理论.....	24
3.2 基于分层线性模型的影响因素分析.....	26
3.3 基于广义部分线性可加模型的影响因素分析.....	32

---

---

3.4 本章小结.....	39
<b>第4章 老年人保健消费支出影响因素实证分析.....</b>	<b>40</b>
4.1 基于分层线性模型的影响因素分析.....	40
4.2 基于广义部分线性可加模型的影响因素分析.....	45
4.3 老年人保健消费决策与支出影响因素对比.....	50
4.4 本章小结.....	51
<b>第5章 结论与展望.....</b>	<b>52</b>
5.1 结论与建议.....	52
5.1.1 研究结论.....	52
5.1.2 建议.....	55
5.2 研究展望.....	56
参考文献.....	57
致    谢.....	60

## 第1章 绪论

### 1.1 研究背景及意义

#### 1.1.1 研究背景

随着我国经济的快速发展，医疗卫生支出增加，科研及人才培养体系逐步健全，医疗水平提高，人均寿命延长，老龄人口比例不断攀升。目前全国60岁以上老年人口比例已达17.88%，山东省老年人口比例已达21.40%，国际上通常把60岁以上人口占总人口比例达到10%作为国家或地区进入老龄化的标准。我国已处于老龄化社会阶段，在人口老龄化如此严峻的背景下，老年人问题应得到足够重视。

老年人随着年龄的增长，身体机能逐渐衰退，机体调节与控制能力降低，器官功能下降，感知能力下降，自身免疫力下降，使得老年人患各种慢性病的概率提升，生活不能自理的人数增加。另外，由于老年人大多已离开工作岗位，家庭环境与社会环境均发生改变，或家庭成员的帮助和照顾不足，使老年人感到无助、缺少关爱。老年人特殊的生理机制和心理状态促使他们更多的关注保健品来改善机体功能。我国老年人蕴藏着较大的消费潜力，Karma指出，60岁以上消费者的消费能力大于中青年消费者，徐丽萍等也通过定量分析得出老年人消费是家庭人均消费的1.9倍<sup>[1-2]</sup>。在老年人的全部消费中，保健消费占据十分重要的位置，且存在较严重的分化现象，对老年人保健消费行为进行研究很有必要<sup>[3-5]</sup>。

保健品消费包括人们对保健品、保健器具与保健服务等方面的消费。与医疗消费相比，保健品消费相对成本更低、周期更长。我国老年人的保健品消费呈现三个突出特征。第一，由于老年人特殊的生理和心理特点，成为保健品市场主要消费者，老年群体对于保健产品和服务的消费需求日益旺盛。第二，相当多的老年人对保健消费相对盲目，保健品市场鱼龙混杂，老年人的认知与鉴别能力有限，对保健讲座及保健品销售人员过度信任，平时生活节俭的老年人，花费大部分甚至全部积蓄购买保健品的情况屡见不鲜。第三，保健品销售商大力宣传保健品是子女亲人对老年人情感关怀的体现，赋予了保健品更多功能，保健消费不仅满足了老年人的生理健康需求，一定程度上也带来了情感慰藉。

综上所述，我国老年人保健消费不断增长，这背后可能有宏观经济的改善、国家对医疗卫生的重视，也可能有老年人认知及健康水平的影响，或由于老年人缺乏情感

慰藉等。因此，探究老年人保健品消费影响因素变得尤为重要。

### 1.1.2 研究意义

#### 1. 理论意义

老年人由于特殊的生理和心理特点，对保健产品和服务的消费需求日益旺盛。但保健品市场鱼龙混杂，老年人的认知与鉴别能力有限，对保健消费相对盲目。目前专门针对老年人保健消费的研究较少，已有文献多为定性分析，多从微观视角出发，研究其线性关系。本文采用定量分析的方法，同时考虑个人微观层面与地区宏观层面，从线性与非线性角度探究老年人保健消费决策与保健消费支出影响因素，不仅为保健品领域的研究提供分析依据，同样也扩展了老年群体相关问题的研究领域。通过对老年群体维度的保健品购买问题的剖析，旨在能够为老年人保健和养生方面的研究提供可借鉴的分析视角和实证材料。

#### 2. 现实意义

随着我国人口老龄化日趋严峻，老年人群体作为社会弱势群体，在生活中会遇到很多方面的问题，需要社会各界予以关爱和保护。老年人的身心健康影响社会和谐与稳定，国家繁荣与发展同样需要老年人的共同参与。因此，社会各界应采取积极的引导措施来保障老年人的健康生活。另外，保健品行业迅速发展，影响范围广，涉及种类多，但随之而来的问题层出不穷，保健品市场急需规范和管理。本文旨在研究老年人保健消费现状及其影响因素，深入剖析老年人的保健消费行为，引导老年人树立正确消费观，对保健品及保健服务进行理性消费，鼓励老年人健康生活；倡导子女多关心陪伴父母，给予老年人一定情感慰藉；呼吁社会及有关部门关注老年人的生理及心理需求。另外，本文也将给保健品企业产品推广提供相关理论依据，促进保健消费市场的平稳健康发展。

## 1.2 国内外文献综述

### 1.2.1 老年人保健消费微观影响因素文献综述

从微观视角探究医疗保健消费问题的文献中，其影响因素大致可归结为社会人口学特征（如性别、年龄、受教育水平等）、经济状况、生理健康与心理健康、代际支持以及医疗保险与社会保障拥有情况等。

在社会人口学特征及经济状况对医疗保健消费影响的研究中，宋璐和左冬梅从性

别的视角研究农村老年人医疗保健支出状况及其影响因素，Tobit 模型回归结果显示，农村老年人医疗保健支出的影响因素存在明显的性别差异，男性老人的保健支出更多取决于自身的潜在因素，而女性老人的保健支出则更多依赖于家庭和子女支持等诱发因素<sup>[6]</sup>。魏瑾瑞和张睿凌指出年龄节点是影响医疗保健消费的重要因素<sup>[7]</sup>。基于中国营养和健康调查数据，韩雪梅和王增福发现居民受教育水平正向影响医疗保健消费支出，且与其他影响因素相比影响最大，并根据研究结果提出了相关对策建议<sup>[8]</sup>。Kukk 研究发现家庭债务对包括医疗保健的多项消费存在抑制作用<sup>[9]</sup>。

在生理健康及代际支持对医疗保健消费影响的研究中，周晶和韩央迪等研究了照料孙子女的经历对农村老年人生理健康的影响，并且讨论了成年子女提供的代际支持在其影响路径中的作用，研究发现，相比于未曾照料过孙子女的农村老年人，提供持续性隔代照料的农村老年人的自陈健康状况和日常生活自理能力更好，同时，情感支持在照料经历对其自陈健康状况的影响路径中发挥了一定的中介作用，即能否从其成年子女处获得情感支持影响了他们的自陈健康状况<sup>[10]</sup>。

在医疗保险对医疗保健消费影响的研究中，Kim 和 Richardson 对老年群体医疗保健消费支出的影响因素进行了探究，认为医疗保险与慢性病对医疗保健支出存在显著影响。研究发现，患有慢性病且有医疗保险的老年人的医疗保健支出明显低于没有医疗保险的老年人<sup>[11]</sup>。Bíró 研究发现补充性商业健康保险能够增加老年人医疗卫生服务需求<sup>[12]</sup>。臧文斌和刘国恩等利用中国城镇居民基本医疗保险入户调查九个城市的的面板数据，实证分析了城镇居民基本医疗保险对城镇家庭消费的影响，研究发现，参保家庭的医疗保健支出与未参保家庭无显著差异，而非医疗消费大约比未参保家庭高 13%<sup>[13]</sup>。

### 1.2.2 老年人保健消费宏观影响因素文献综述

国内外学者关于医疗保健消费影响因素的宏观层面研究，大多集中在居民人均收入、人口结构、医疗保险政策、政府对卫生事业的投入、医疗保健供给、医疗价格、环境质量等角度。

在居民人均收入及人口结构对医疗保健消费影响的研究中，田侃和瞿华利用 1990-2011 年间中国中、东、西部面板数据，探究了三大地区城乡居民收入与医疗保健消费支出之间的关系，发现中、东、西部人均可支配收入与医疗保健支出存在显著正相关关系，且中西部地区相关系数高于东部地区<sup>[14]</sup>。王学义和张冲分析了我国人均

收入变化和人口年龄结构对居民医疗保健支出的影响,研究发现,我国老年抚养系数对居民医疗保健消费支出有正影响,即老年抚养系数的上升会带动居民人均医疗保健消费支出的增加,而少儿抚养系数对医疗保健支出不存在显著影响<sup>[15]</sup>。Cantarero 整理西班牙各地区相关数据,对居民医疗消费支出进行探究,得出年龄结构对居民医疗支出水平存在重要影响<sup>[16]</sup>。Bech 发现年龄结构变化只会在短期内引起医疗消费的变化,长期来看人口老龄化对医疗消费的增长并不显著<sup>[17]</sup>。

在医疗保险政策、政府对卫生事业的投入、医疗保健供给以及医疗价格对医疗保健消费影响的研究中,刘国恩和蔡春光等得出国家医保政策改变了老年人医疗需求的结论<sup>[18]</sup>。Chou 探究了政府预算赤字对中国医疗保健支出产生的重大长期影响,为中国医疗融资政策的制定提供了理论依据<sup>[19]</sup>。聂聆和李斌用面板数据探究了我国江苏、浙江与广东三个省份的城镇居民医疗保健消费情况,发现三省城镇居民医疗保健支出与政府公共卫生事业支出存在显著负相关关系。同时发现由于健康服务和产品特别是非医疗性健康服务和产品供给不足,近年来城镇居民医疗保健支出占人均消费支出的比例以及医疗保健边际消费倾向均偏低,表明中、高收入人群的医疗保健消费仍有较大发展空间<sup>[20]</sup>。Molla 对孟加拉样本进行研究后,同样得出政府医疗支出与居民医疗保健消费支出呈负相关<sup>[21]</sup>。聂裕以我国 31 个省份的城镇居民为研究对象,得出医疗价格的上涨会降低医疗保健消费的支出<sup>[22]</sup>。

在环境质量对医疗保健消费影响的研究中,徐冬林和陈永伟分析了环境质量与城镇居民医疗保健消费之间的长期关系,研究结果表明环境恶化导致居民健康状况变差,从而使得医疗保健消费增加<sup>[23]</sup>。叶小青和李先玲同样研究了环境质量对医疗保健支出的影响,得出了相同结论<sup>[24]</sup>。Yang Jing 发现空气污染对中部地区医疗保健消费支出的影响最大,东部地区次之,而对西部地区影响不显著<sup>[25]</sup>。

### 1.2.3 老年人保健消费影响因素研究方法文献综述

国内外学者针对医疗保健消费影响因素问题采用定性分析或定量分析进行研究,定量研究中多用多元线性模型、方差分析、因子分析、结构方程、空间回归模型、两部模型、Tobit 模型等方法。

在对医疗保健消费影响因素进行定性分析的研究中,谈煜鸿基于文献回顾,采用描述性统计方法,概括和总结了老年人保健品消费水平、认知、态度、行为等方面的特点,分析并探讨了影响老年人保健品消费的诸如中医养生文化、经济收入状况和受

教育程度等因素,为未来的深入研究提供思路<sup>[26]</sup>。厉杰通过对选取的老年人进行访谈调查,采用定性分析方法,得出影响老年人保健品购买行为的因素主要包括个体因素和社会因素<sup>[27]</sup>。

在对医疗保健消费影响因素进行定量分析的研究中,刘斯敏和曾学文对南宁市消费者购买中药保健品的影响因素进行抽样调查,采用南宁市 462 名消费者的调查数据,将中药保健品的消费金额分层,设定为有序因变量,建立多元定序 Logit 模型,分析各因素对中药保健品消费额的影响程度<sup>[28]</sup>。石贝贝通过采用单因素方差分析和多元线性回归分析方法对全国老年人口及分城乡老年人口医疗保健消费决定因素进行了研究<sup>[29]</sup>。王菲同样采用了方差分析方法对老年人医疗消费行为展开研究<sup>[30]</sup>。肖波涛以成都市保健品市场消费者态度为研究对象,构建态度评价指标体系,采用主成分分析,在损失少量信息的条件下,对多个因素进行降维,得到简化后的公共因子,然后采用聚类分析将样本划分为具有不同结构特征的群组,完成对保健品市场消费群体的市场细分<sup>[31]</sup>。陈茗和刘素青采用了与主成分分析法相近的因子分析法探究影响老年人消费规律的主要因素<sup>[32]</sup>。王沛立和李恩平采用弹性网方法对居民医疗负担影响因素进行了分析<sup>[33]</sup>。

邓敏和杨莉等利用计划行为理论和结构方程模型,从行为信念和主观规范等方面研究老年人医疗消费行为<sup>[34]</sup>。冯桂平和刘爽强调不同问题的不同变化过程,基于跨理论模型变化阶段与变化过程理论,将医疗保健消费行为与跨理论模型相结合,通过建立健康行为变化阶段与老年人医疗保健品消费行为影响因素之间的联系,分析其行为发生和变化的动态过程中的微观因素,提出一种可介入的变化阶段框架以及对应的干预方法,从而实现对老年人医疗保健品消费行为的有效管理<sup>[35]</sup>。戴平生和李芳芳利用误差空间自回归模型,对 2009 年我国 31 个省域城乡居民医疗保健消费支出的影响因素进行了实证分析<sup>[36]</sup>。

苏春红和李齐云等与叶春辉和封进等针对很多样本医疗支出为零的情况,采用两部模型将消费者行为分为相互独立的两个阶段:第一阶段,消费者决定是否进行医疗消费;第二阶段,在决定参与医疗消费的前提下,确定医疗支出具体金额。该方法将以上两个阶段视为有先后顺序的过程,由两部分模型计算医疗支出偏效应,避免了零膨胀数据导致的样本选择偏误问题<sup>[37-38]</sup>。李涛和徐翔等与胡宏伟和张小燕等则针对限值因变量采用了 Tobit 模型,探究了老年人医疗保健支出的影响因素<sup>[39-40]</sup>。

### 1.2.4 总体评价

通过对国内外文献的查阅及整理,发现已有文献均从宏观或微观单一角度对保健消费进行探究,多为定性研究或线性分析,且很少专门针对老年人这一特殊群体。事实上,保健消费有可能同时受微观因素与宏观因素的影响,且影响机制不一定只是简单的线性关系,也可能是更为复杂的非线性关系。因此,本文同时引入个人层面微观因素与地区层面宏观因素,构建分层线性模型,并进一步建立广义部分线性可加模型,深入探究微观因素对老年人保健消费的非线性影响。

## 1.3 研究方法 with 主要内容

### 1.3.1 研究方法

1. 文献研究法。本文对相关的文献资料 and 研究成果进行整理和归纳,通过分析前人研究开拓写作思路。

2. 描述性统计分析法。对老年人保健消费及其影响因素进行描述性统计分析。

3. 两部模型思想。鉴于存在相当一部分老年人保健消费金额为零,因此采用两部模型思想将老年人保健消费分为保健消费决策与保健消费支出两部分相互独立的过程,分别展开研究。

4. 实证分析法。从“中国健康与养老追踪调查”数据库获取微观数据,从国家统计局年鉴获取宏观数据,构建分层线性模型与广义部分线性可加模型对老年人保健消费决策与保健消费支出影响因素进行实证分析。

### 1.3.2 技术路线

本文技术路线如图 1-1 所示。首先明确研究目的与研究内容,研究关于老年人保健消费问题的已有文献与相关理论;然后建立指标体系并进行数据处理,对相关变量做描述性统计分析;最后对老年人保健消费决策与保健消费支出两方面的影响因素进行实证分析,得出相关结论与建议。



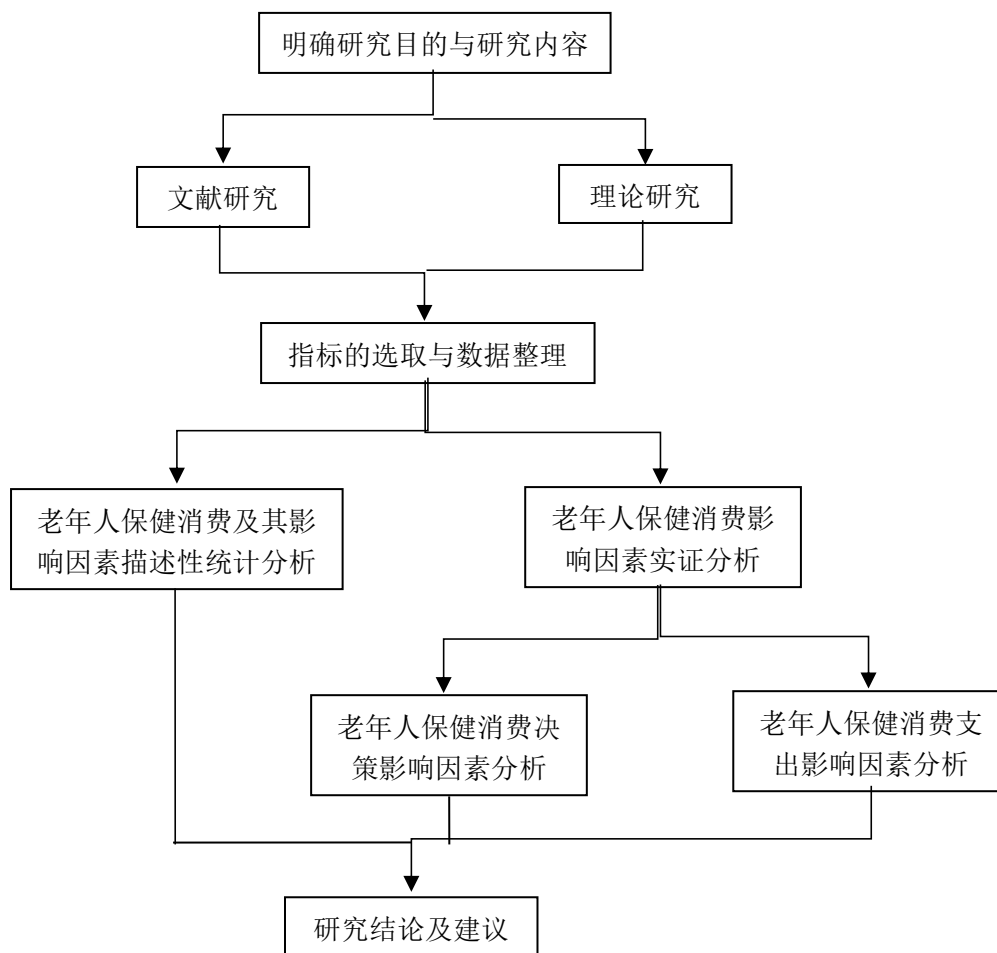


图 1-1 技术路线图

### 1.3.3 主要研究内容

本文首先建立老年人保健消费影响因素宏微观指标体系，并对变量进行说明与描述性统计分析。然后基于两部模型原理将老年人保健消费行为分为保健消费决策与保健消费支出两部分相互独立的过程，同时考虑个人层面微观因素与地区层面宏观因素，建立分层线性模型探究其对老年人保健消费的影响。最后从非参数角度建立广义部分线性可加模型进一步深入探究微观因素的非线性影响。

### 1.4 创新点

1. 本文同时从个人微观层面与地区宏观层面两个角度探究老年人保健消费影响因素。
2. 在建立线性模型的基础上，进一步建立广义部分线性可加模型，从非参数角度探究微观因素对老年人保健消费的非线性影响，得出更深层、更准确的结论。

## 第 2 章 变量选取与描述性统计分析

本章内容将说明实证部分所用数据来源，构建宏微观指标体系，并对被解释变量与解释变量进行描述性统计分析，为实证分析奠定基础。

### 2.1 数据来源及样本说明

本文所用个人层面的微观数据来源于“中国健康与养老追踪调查”(CHARLS)，该项目由北京大学负责开展，由国家自然科学基金委资助，以我国 45 岁及以上中老年人作为调查对象，旨在整理一套关于中老年人生活的高质量微观数据。CHARLS 于 2011 年开展全国基线调查，采用多阶段抽样的方法，定期追访并不断增加新的样本，调查范围涵盖了全国除海南、西藏、宁夏的 28 个省份或直辖市的 150 个县中的 450 个社区(村)。至 2015 年追访时，总计收集了 2.3 万名受访者信息。该项目委员会成员来自中国社会科学院、北京大学、中国人民大学、斯坦福大学、哈佛大学和和兰德公司等国内外知名高校及研究机构。问卷设计在充分考虑我国实际的基础上参考了国际调查经验，包括美国健康与退休调查(HRS)、英国老年追踪调查(ELSA)以及欧洲的健康、老年与退休调查(SHARE)。CHARLS 问卷的信度和效度均经过有效检验，问卷应答率高且数据可靠，被广泛应用于各领域学术研究，深受学者认可。

本文采用了“中国健康与养老追踪调查”最新一轮，即 2015 年数据进行实证分析。CHARLS 的样本容量很大，但并非所有样本数据都合格，需根据研究目的筛选有效样本。由于我国《老年人权益保障法》将老年群体的年龄起点标准规定为 60 周岁，故本文首先筛选了 CHARLS 受访时年龄在 60 周岁以上的样本，并对保健消费问题回答为“不知道”、“记不清”、“拒绝回答”的无效样本进行删除。然后由问卷信息的各个数据集匹配整理出本文需要的微观层面指标数据，对于少数出现矛盾的数据，通过人工判断进行相应修正，如某人回答无保健消费支出却回答了保健消费具体金额，则将其修正为有保健消费支出。最后，删除数据缺失严重的样本，对于极少数缺失数据，采用适当方法进行插补，例如对数值型变量“受教育年限”在有保健消费与无保健消费下分别采用均值插补，对“自评健康”分类型变量在有保健消费与无保健消费下分别采用众数插补。微观缺失数据仅占全部数据的 1%左右，经过合理插补，几乎不会对分析结果造成偏差。

除此之外，本文同时考虑宏观指标对老年人保健消费的影响，地区层面的宏观数

据来源于 2015 年中国统计年鉴分省年度数据。根据受访老人个人 ID 与 PSU 数据集，匹配出各样本所属省份，将样本所涉及的 28 个省份分别赋值为 1-28，并整理相关地区变量。

经筛选后最终样本量为 4041，其中保健消费不为零的有 788 例，约占全部样本的 20%。保健消费支出为零的有 3253 例，约占全部样本的 80%。即最终用于保健消费决策模型的样本量为 4041，用于保健消费支出模型的样本量为 788。

## 2.2 指标选取及变量说明

### 2.2.1 被解释变量界定及说明

#### 1. “保健消费”基本概念及界定

保健消费包括食用维生素和补品等保健品、使用保健设备以及享受保健服务所花费的金额，不包括医疗费用。

其中保健品即保健食品，指具有特定保健功能的食品，适宜于特定人群食用，具有机体调节功能，具有特定的功效，且不以治疗疾病为目的，不可代替药品。保健食品与一般食品不同，一般食品不强调或不具备特定功能，无特定的食用人群范围，可不限食用。保健食品与药品不同，药品的生产能力与技术条件均需严格审查，经多年临床观察及有关部门鉴定批准后才可投入市场，对治疗疾病有一定疗效，并配有详细且严谨的使用说明、适应症、注意事项、不良反应等；而保健品生产标准较低，没有疾病治疗作用，且说明书不如药品详细、严格。

保健设备包括普通健身设备、保健用按摩器、具有保健功能的服饰及用品等，对疾病或损伤具有预防、缓解等作用。保健设备与医疗设备不同，保健设备为强身健体防病用，而医疗设备为诊断与治疗疾病用。

保健服务消费包括由健康养生咨询、保健按摩、推拿、足疗、经络养生、减压疏导等项目产生的消费，多侧重于身体健康调理与疾病预防。

#### 2. 被解释变量说明

本文探究老年人保健消费的影响因素，因此以 60 周岁以上受访老人过去一个月的保健消费总额为响应变量。

由于受访老人中有相当一部分无保健消费支出，即保健消费变量取值为零，为保证模型准确性，采用两部模型思想，将老年人保健消费行为分解为独立的两部分，分

别定义被解释变量:

第一部分为保健消费决策行为,将发生保健消费记为1,无保健消费记为0,定义0-1二分类变量作为被解释变量。

第二部分为保健消费支出行为,即在决定进行保健消费的条件下,实际发生的保健消费金额,根据两部模型原理,将该金额的对数定义为被解释变量。

本文将分别研究老年人保健消费决策与保健消费支出的影响因素。

## 2.2.2 微观解释变量选取及说明

通过对已有文献的研究、总结与拓展,将老年人保健消费的微观影响因素归纳为健康状况、认知能力、社会人口学特征、财富状况、代际支持和社会互动几个方面。

### 1. 健康状况

健康状况是影响老年人保健消费的重要因素,保健品以提升人体某种机能为主要功能,保健消费通常被看作健康状况的引致消费,一般来说,健康状况越差,由健康问题引致的保健消费可能就越多。由于老年人闲暇时间较多,子女作为中青年群体,大多生活压力较大且生活节奏越来越快,无法给予父母足够的陪伴和关心。因此我们在关注老年人生理健康的同时,心理健康不可忽视,应考虑由于孤独感引发的保健消费。本文从两个方面构建健康状况指标,一是老年人生理健康状况,以自评健康和慢性病患病情况度量;二是老年人心理健康状况,以抑郁程度和孤独感度量。

(1) 自评健康。自评健康是受访老人对健康状况的自我评价,考虑到受访者对自己的健康状况没有足够的认知,CHARLS问卷在“健康状况”部分分别于首尾两次询问自评健康状况,本文选取了经过一系列健康访谈后对自己的健康状况有更深刻认知的第二次自评健康数据。若受访者回答自己的健康状况为“极好”、“很好”和“好”,则认为自评健康较好,赋值为1;若受访者回答“一般”、“不好”和“很不好”,则认为自评健康较差,赋值为0,由此构建0-1虚拟变量。

(2) 慢性病患情。问卷调查了受访者高血压、高血脂、糖尿病、恶性肿瘤、慢性肺部疾患、肝脏疾病、肾脏疾病、心脏病、消化系统疾病、中风、哮喘、关节炎或风湿病、情感及精神方面问题、与记忆相关疾病共14种慢性病患病情况,将受访者患病种类数作为慢性病患情变量。

(3) 抑郁程度。CHARLS问卷使用了流调中心抑郁量表(CES-D)的九个问题,包括“感到情绪低落”、“感到害怕”、“因一些小事而烦恼”、“觉得无法继续生

活”四个负面情绪问题，“睡眠不好”、“在做事时很难集中精力”、“做任何事都很费劲”三个生理症状问题以及“感到愉快”和“对未来充满希望”两个积极情绪问题。每个问题有四个选项，分别为“很少或没有”、“不太多”、“有时或一半时间”、“大多数时间”，对于两个积极情绪问题四个选项分别记为4分、3分、2分、1分，其他问题分别赋值为1分、2分、3分、4分，将九个问题得分相加，构造抑郁程度变量，变量取值越高，受访者抑郁程度越严重。

(4) 孤独感。根据问卷中“是否感到孤独”这一问题，将回答“很少或根本没有”选项的记为0，其他选项记为1，构造0-1虚拟变量。

## 2. 认知能力

老年人保健消费也可能会受认知能力影响。销售人员误导老年人花费大金额进行不必要保健消费的报道层出不穷，因此较高的认知能力有助于老年人辨别信息真伪，可能对保健消费存在反向影响。同时，较高的认知能力也可能帮助老年人选择适合自己的保健品或保健服务，从而对保健消费产生正向影响。

根据 CHARLS 问卷中的简易智能精神量表 (MMSE) 构建认知能力指标。其中包括对定向力、计算力、记忆力、复写能力四个方面的考察，每个方面总分为5分。定向力包括五个小问题，即令受访者说出当前的季节、年份、月份、日期以及星期几，每答对一个记1分。计算力的考察则要求受访老人计算100减7，连续减五次，每答对一次记1分。记忆力分为即时记忆力与延时记忆力，访员读出十个词，令受访者回忆，并间隔一段时间后令受访者再次回忆，每记起一个词记0.5分，取即时记忆与延时记忆平均分作为记忆力得分。复写能力的考察则通过向受访者展示一张两个五角星重叠的图片，若受访老人画出了该图片，则复写能力记5分。将定向力、计算力、记忆力与复写能力得分相加，作为认知能力变量，变量取值越大，说明受访老人认知能力越强。

## 3. 社会人口学特征

可能对老年人保健消费产生影响的社会人口学特征包括年龄、性别、婚姻状况、受教育年限、工作状况、退休状况、养老保险和户口类型。

(1) 年龄。以受访年份2015减受访者出生年份所得值作为年龄变量。由于只研究60岁以上老年人保健消费情况，因此年龄变量最小值为60。

(2) 性别。将女性赋值为1，男性赋值为0，构建0-1虚拟变量。

(3) 婚姻状况。问卷中婚姻状态有7个选项，分别为已婚且与配偶一同居住、

已婚但暂不与配偶一起居住、长期分居（不再作为配偶共同生活）、离异、丧偶、从未结婚、同居。本文将已婚且与配偶一同居住、已婚但暂不与配偶一起居住、同居三项视为“有配偶”，赋值为1；将长期分居、离异、丧偶、从未结婚四项视为“无配偶”，赋值为0，构造0-1虚拟变量。

(4) 受教育年限。根据问卷中对受访者“最高受教育水平”的调查，赋值为相应受教育年限，如小学毕业赋值为6，初中毕业赋值为9。对回答“未读完小学”的赋值为具体的读书年数，对具体年数不清楚的受访者统一将其受教育年限赋值为3。

(5) 工作状况。将工作状况分为工作与不工作两类，从事农业生产经营活动（包括种地、养鱼、养牲畜等）、个体经营、挣工资劳动或无工资为家庭经营活动帮工等均算是工作，义务志愿劳动与家务劳动不包括在内，工作赋值为1，不工作则赋值为0，构建0-1虚拟变量。

(6) 退休状况。将已办理退休手续记为1，反之记为0，构造0-1虚拟变量。

(7) 养老保险。我国养老保险种类包括政府机关与企事业单位养老保险、城乡居民养老保险、城镇居民养老保险、新型农村养老保险、失地农民养老保险、商业养老保险以及其他养老保险。若受访老人参加了或正在领取以上任一种养老保险，则赋值为1，否则赋值为0，构建0-1虚拟变量。

(8) 户口类型。户口类型变量以是否非农户口来表示，非农户口记为1，农业户口记为0，定义为0-1虚拟变量。

#### 4. 财富状况

由于多数保健品及保健服务价格较高，财富状况也是影响老年人保健消费的一个重要方面。本文选取个人年均收入、家庭金融资产、家庭负债三个指标作为反映老年人财富状况的变量。

(1) 个人年均收入。CHARLS 问卷对受访者各项收入进行了详细询问，本文个人年均收入指标以过去一年受访老人工资性收入（包括奖金）、养老金以及各项补助（包括失业补助、高龄老人养老补助、伤残补助、社会捐助等）之和计算。其中养老金包括由政府机关和企事业单位养老保险、城乡居民养老保险、城镇居民养老保险、新型农村养老保险、失地农民养老保险、商业养老保险以及其他养老保险所得到的养老金。考虑到我国农村居民的实际情况，要量化农村居民个体对家庭收入的贡献存在一定困难。因此，针对没有个人工资的农户，本文将其家庭农业年收入（包括种植业、林业、畜牧业以及渔业收入）除以参与该农业活动的家庭成员数，得到其农业人均收

入，再加上养老金及各项补助，即得到农村人口个人年均收入。

(2) 家庭金融资产。将受访老人家庭现金数、金融机构存款数、现有债券面值、持有股票当前价值以及现持有的基金当前价值之和记为家庭金融资产变量。

(3) 家庭负债。CHARLS 问卷分别询问了未还清贷款总额、信用卡所欠金额、欠其他家庭或单位钱数。如果受访者不愿回答或者忘记了，则分级展开提问，确定欠款区间，并以上限和下限的均值作为该受访者该项欠款金额。将以上三项欠款相加，则得到家庭负债变量数据，若无债务，则变量取值为 0。

## 5. 代际支持

子女给老年人提供的物质及精神支持也可能影响其保健消费水平，其中物质支持以子女给老人的货币支持和子女给老人的实物支持两个指标度量，精神支持以子女一年中看望次数度量。另外，加入老年人是否照顾孙辈虚拟变量，子女看望与照顾孙辈都可以在一定程度上给予老年人精神寄托，从而反向影响其由于缺乏精神慰藉而产生的保健消费。

(1) 子女货币支持。将过去一年从各个子女处获得的货币总额记为子女货币支持变量。若受访者拒绝回答或不知道，则分级询问该问题，并赋值为上限和下限的均值。

(2) 子女实物支持。将过去一年从各个子女处获得的食物、衣物以及其他物品总额记为子女实物支持变量。同样的，若受访者拒绝回答或不知道，则分级询问该问题，并赋值为上限和下限的均值。

(3) 子女看望次数。首先将问卷中子女看望频率问题转换为一年内看望次数，例如每天看望则记为 365 次，每周一次则记为 52 次，每半月看望则记为 24 次。同住子女按照每日看望（即一年看望 365 次）计算。将所有子女一年内看望次数相加，得到最终的“子女看望次数”变量。

(4) 是否照顾孙辈。过去一年中花时间照顾了孙子女记为 1，否则记为 0，构造 0-1 虚拟变量。

## 6. 社会互动

社会互动也可能通过情景互动和内生互动影响老年人保健消费。情景互动是一种单向作用，个体的决策受到群体成员特征的影响，但不能反作用于群体成员特征。内生互动即个体和群体成员之间互为影响。老年人保健消费决策及金额也可能会受互动群体的影响，同时，社交活动也可以排解老年人负面情绪，有利于心理健康，故加入

社交情况变量。

社交情况。根据问卷中对老年人过去一个月打麻将、跳舞、参加培训课程等 11 项社交活动的调查，将“未参加”赋值为 0 分，“不经常”赋值为 1 分，“差不多每周参加”赋值为 2 分，“差不多每天参加”赋值为 3 分，将 11 项得分加总得到社交情况变量。取值越高，说明该受访老人社交活动越丰富。

### 2.2.3 宏观解释变量选取及说明

老年人保健消费不仅受个人微观层面因素的影响，还可能受地区宏观层面因素的影响。将宏观影响因素归结为地区经济、政府卫生支出、保健消费价格以及保健供给四个方面。

#### 1. 地区经济

一般来说保健品及保健服务定价较高，地区经济将影响个人消费能力与消费观念，因此引入地区经济类宏观指标，以各省份人均地区生产总值与居民人均可支配收入度量。人均地区生产总值与居民人均可支配收入单位为“元”，由于数值较大故对其进行对数处理纳入模型。

#### 2. 政府卫生支出

政府卫生支出方面以地方财政医疗保健支出度量，政府对医疗保健的公共支出一定程度上减轻了公民个人的医疗保健消费负担，因此理论上对老年人保健消费实际支出存在一定影响。

#### 3. 保健消费价格

根据经济学一般规律，商品价格影响购买支出，保健消费价格也会影响老年人保健消费，以医疗保健类居民消费价格指数度量。医疗保健类居民消费价格指数是与上年相比医疗保健类消费价格变动趋势和程度的相对数。

#### 4. 保健供给

保健供给以地区医疗卫生机构数和卫生人员数度量，保健供给程度越高，该地区保健市场发展越好，老年人保健消费往往越高。医疗卫生机构数是指从卫生行政部门取得许可证，或从管理部门取得登记证书的工作单位数量，包括基层医疗卫生机构、专业公共卫生机构、医院及其他医疗卫生机构。卫生人员数指在医疗卫生机构工作的职工人数，包括乡村医生和卫生员、卫生技术人员、管理人员和工勤人员等。

综上所述，总结本文解释变量如表 2-1 所示：



表 2-1 老年人保健消费宏微观解释变量

	指标类别	指标名称	数据类型	符号表示
个人微观指标	健康状况	自评健康	0-1 变量	$X_1$
		慢性病患病	连续型变量	$X_2$
		抑郁程度	连续型变量	$X_3$
		孤独感	0-1 变量	$X_4$
	认知能力	认知能力	连续型变量	$X_5$
	社会人口学特征	年龄	连续型变量	$X_6$
		性别	0-1 变量	$X_7$
		婚姻状况	0-1 变量	$X_8$
		受教育年限	连续型变量	$X_9$
		工作状况	0-1 变量	$X_{10}$
		退休状况	0-1 变量	$X_{11}$
		养老保险	0-1 变量	$X_{12}$
		户口类型	0-1 变量	$X_{13}$
	财富状况	个人年收入	连续型变量	$X_{14}$
		家庭金融资产	连续型变量	$X_{15}$
		家庭负债	连续型变量	$X_{16}$
	代际支持	子女货币支持	连续型变量	$X_{17}$
		子女实物支持	连续型变量	$X_{18}$
		子女看望次数	连续型变量	$X_{19}$
		是否照顾孙辈	0-1 变量	$X_{20}$
	社会互动	社交情况	连续型变量	$X_{21}$
地区宏观指标	地区经济	人均地区生产总值	连续型变量	$Z_1$
		居民人均可支配收入	连续型变量	$Z_2$
	政府卫生支出	地方财政医疗保健支出	连续型变量	$Z_3$
	保健消费价格	医疗保健类居民消费价格指数	连续型变量	$Z_4$
	保健供给	医疗卫生机构数	连续型变量	$Z_5$
		卫生人员数	连续型变量	$Z_6$

## 2.3 老年人保健消费及其影响因素描述性统计分析

### 2.3.1 老年人保健消费情况描述性统计分析

本文实证部分所用样本共 4041 例，其中，3253 位受访老人无保健消费支出，约占全部样本的 80%，788 位受访老人存在保健消费支出，约占全部样本的 20%。老年人保健消费决策直观分布情况如图 2-1 所示：

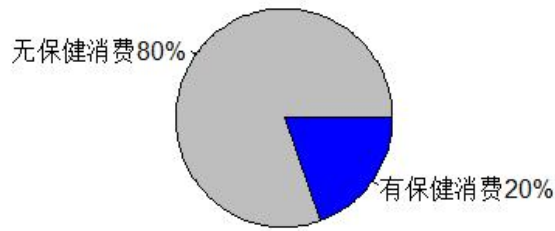


图 2-1 保健消费决策样本分布图

对于存在保健消费的老年人，其保健消费金额分布情况如表 2-2 及图 2-2 所示：

表 2-2 保健消费支出样本分布表

保健消费金额区间	(0, 100)	[100,500)	[500,1000)	[1000,5000)	5000 及以上	总计
样本数量	338	278	73	83	16	788
样本占比	42.89%	35.28%	9.27%	10.53%	2.03%	100%

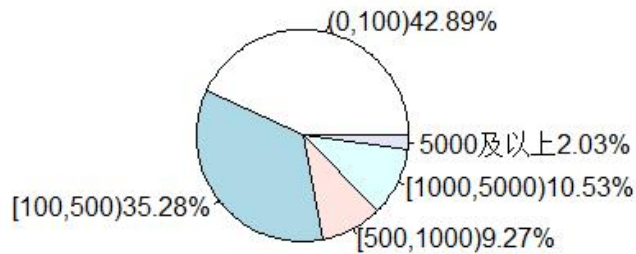


图 2-2 保健消费支出样本分布图

表 2-2 及图 2-2 显示了保健消费支出样本分布情况，在存在保健消费支出的 788 例样本中，每月消费金额在 0-100 元区间内的有 338 位老人，占比最大，而月消费 5000 及以上的仅为 16 人，占比最小。观察表 2-2 及图 2-2 可以看出，在有保健支出的样本中，将近 80% 的老年人月保健消费在 500 元以内，仅有少数老年人存在过高保健消费情况。

### 2.3.2 老年人保健消费影响因素描述性统计分析

#### 1. 0-1 型微观变量描述性统计分析

微观影响因素指标体系中，有自评健康、孤独感、性别、婚姻状况等 9 个 0-1 型虚拟变量，对 0-1 型微观变量进行描述性统计，统计结果如表 2-3 所示：

表 2-3 0-1 型微观影响因素描述性统计

指标名称	变量水平	样本数	有保健消费者比例	保健消费支出均值
自评健康	较好	1090	13.21%	537.30
	较差	2951	21.82%	648.67
孤独感	孤独	1092	25.00%	434.92
	不孤独	2949	17.46%	622.71

表 2-3 0-1 型微观影响因素描述性统计 (续表)

指标名称	变量水平	样本数	有保健消费者比例	保健消费支出均值
性别	女性	1975	22.38%	552.76
	男性	2066	16.75%	563.90
婚姻状况	有配偶	3100	19.61%	556.48
	无配偶	941	19.13%	561.63
工作状况	工作	2099	15.05%	288.91
	不工作	1942	24.30%	737.58
退休状况	已退休	680	35.44%	962.89
	未退休	3361	16.27%	379.11
养老保险	有养老保险	2894	19.28%	752.46
	无养老保险	1147	20.05%	477.36
户口类型	农业	3069	15.25%	259.66
	非农业	972	32.92%	993.47
是否照顾孙辈	是	1601	18.24%	558.37
	否	2440	20.33%	557.23

表 2-3 展示了各微观虚拟变量两水平样本数, 并经计算得到了各水平下有保健消费者比例以及存在保健消费老年人的保健消费支出金额均值。自评健康较好的受访者中, 有 13.21% 的老年人存在保健消费, 且保健消费支出均值为 537.30; 自评健康较差的老年人中有 21.82% 存在保健消费, 且保健支出均值为 648.67。可以看出不论是保健消费概率还是支出金额, 健康较差的受访者均高于健康状况较好的受访者。同样地, 由上表还能看出孤独者较不孤独者、女性老年人较男性老年人、不工作者较工作者、已退休者较未退休者、非农村老年人较农村老年人, 保健消费概率更高; 而婚姻状况、是否有养老保险与是否照顾孙辈就保健消费概率来看, 未见明显差异, 需由下文构建模型分析。另外, 自评健康较差的受访者较较好的受访者、不孤独的老年人较孤独的老年人、不工作者较工作者、已退休老年人较未退休老年人、有养老保险者较无养老保险者、非农村老年人较农村老年人, 平均保健消费支出金额更高; 而不同性别、婚姻状况、是否照顾孙辈就消费支出均值来看无明显差异。变量间具体作用关系将在下文构建模型进一步研究。

## 2. 连续型微观变量描述性统计分析

慢性病患病、抑郁程度、认知能力、年龄等 12 个连续型微观影响因素的描述性统计如表 2-4 所示:

表 2-4 连续型微观影响因素描述性统计

指标名称	度量单位	最小值	最大值	总体均值	有保健消费者变量均值	无保健消费者变量均值
慢性病患情	种	0.00	11.00	1.76	2.40	1.61
抑郁程度	分	1.00	36.00	15.95	16.61	15.80
认知能力	分	0.00	20.00	9.93	11.00	9.66
年龄	岁	60.00	101.00	68.73	69.38	68.57
受教育年限	年	0.00	19.00	4.49	5.99	4.13
个人年均收入	元	0.00	614400.00	8387.78	12483.08	7395.74
家庭金融资产	元	0.00	8001000.00	17624.48	32606.82	13995.19
家庭负债	元	0.00	6000000.00	7001.42	9810.53	6320.95
子女货币支持	元	0.00	328000.00	4131.34	3969.40	4170.57
子女实物支持	元	0.00	1000000.00	3210.57	5211.56	2725.85
子女看望次数	次	0.00	2555.00	363.95	319.61	374.69
社交情况	分	0.00	16.00	1.65	2.34	1.48

表 2-4 列出了指标体系中 12 个连续型微观变量的度量单位、最小值、最大值、总体均值以及有保健消费者与无保健消费者的变量均值。分析表中数据可知,有保健消费老年人的慢性病患情、抑郁程度、认知能力、年龄、受教育年限、个人年均收入、家庭金融资产、子女实物支持以及社交情况等变量均值高于无保健消费老年人相应变量均值;而子女货币支持与子女看望次数等变量均值表现为无保健消费老年人高于有保健消费老年人。由此可初步看出各变量对老年人保健消费的影响方向,但具体作用关系需根据下文模型得到。

### 3. 宏观变量描述性统计分析

本文实证部分所用老年人样本来自我国除海南、西藏、宁夏之外的 28 个省级行政区,因此收集该 28 省的相关数据,对人均地区生产总值、居民人均可支配收入等 6 个宏观变量的描述性统计如表 2-5 所示:

表 2-5 宏观影响因素描述性统计

指标名称	度量单位	最小值	最大值	平均值	标准差
人均地区生产总值	元	26165.00	107960.00	54606.29	24000.44
居民人均可支配收入	元	13466.59	49867.17	22525.61	9210.28
地方财政医疗保健支出	亿元	99.43	918.36	415.40	186.71
医疗保健类居民消费价格指数	百分比	100.60	106.60	102.59	1.30
医疗卫生机构数	个	5016.00	80109.00	34549.29	21687.45
卫生人员数	万人	4.84	85.57	37.61	20.51

表 2-5 列出了人均地区生产总值、居民人均可支配收入、地方财政医疗保健支出、医疗保健类居民消费价格指数、医疗卫生机构数以及卫生人员数 6 个宏观影响因素的度量单位、最小值、最大值、平均值以及标准差，初步展示了宏观变量的数据分布情况。

## 2.4 本章小结

本章首先在 2.1 节介绍了宏微观数据来源以及样本筛选过程。2.2 节则对被解释变量进行了界定及说明，分析了老年人保健消费影响因素并从微观与宏观两个方面选取了解释变量，最终确定了健康状况、认知能力、社会人口学特征、财富状况、代际支持与社会互动 6 个方面共 21 个人微观层面影响因素，以及地区经济、政府支出、保健消费价格与保健供给 4 个方面共 6 个地区宏观层面影响因素。在构建了指标体系后，2.3 节对老年人保健消费及其影响因素进行了描述性统计分析，其中，老年人保健消费情况被分为保健消费决策与保健消费支出两部分进行分析，影响因素被分为 0-1 型微观变量、连续型微观变量与宏观变量三部分进行描述性统计分析。总的来说，本章主要进行了变量选取与描述性统计分析工作，为下一章实证分析奠定基础。

## 第3章 老年人保健消费决策影响因素实证分析

由上一章描述性统计分析可知，老年人保健消费金额存在大量零值，对于零膨胀数据，根据两部模型思想，将老年人保健消费行为分为保健消费决策与保健消费支出两个相互独立的过程进行研究。本章将分析老年人保健消费决策影响因素，首先对相关模型进行介绍，然后采用分层线性模型同时引入宏观变量与微观变量，并构建广义部分线性可加模型探究微观变量间的非线性关系。

### 3.1 相关模型介绍

#### 3.1.1 两部模型思想

##### 1. 模型适用范围及优势

两部模型是专门处理零膨胀数据的模型。零膨胀数据相比于普通数据类型较为复杂，分为零膨胀连续数据和零膨胀计数数据。零膨胀连续型数据即被解释变量的取值范围不是传统模型下的 $(-\infty, +\infty)$ ，而是 $[0, +\infty)$ ，且通常以超过50%的概率取零值，大于零的部分为连续型分布，例如家庭教育支出、家庭医疗保健支出、家庭保险支出等都属于零膨胀连续型数据。若大于零的部分为离散可数数据，即为零膨胀计数数据，如保险索赔次数等。

对零膨胀数据建模存在一定困难，由于大量零值的存在不符合普通模型的建模条件，且传统模型不能充分利用零值所隐含的信息，因此模型参数的估计容易产生偏移，导致样本选择误差，使模型不再有效。针对这类问题，两部模型被提出并得到广泛应用。两部模型对数据的产生机制不做过强的前提假设，结构较为灵活，将零膨胀类问题分为两个相互独立的过程进行研究。如本文探究的老年人保健消费问题，两部模型即将老年人保健消费分解为保健消费决策与保健消费支出两个过程独立研究，避免大量零值导致的样本选择偏误，且充分利用零值信息。

##### 2. 模型形式及估计方法

以本文所研究的老年人保健消费问题为例，两部模型将老年人保健消费行为分为以下两个阶段：

第一阶段为保健消费决策阶段，即老年人决定是否参与保健消费，此时决策行为符合二项离散分布，可使用Logit模型进行拟合。模型形式为公式(3.1)所示：

$$\log\left(\frac{P_i}{1-P_i}\right) = \beta_1 X_i + u_{1i} \quad (3.1)$$

模型(3.1)中 $P_i$ 表示第 $i$ 个受访老人参与保健消费的概率, $X_i$ 表示影响受访者参与保健消费的因素, $u_{1i}$ 为随机扰动项。Logit模型一般采用极大似然估计法进行系数估计。

第二阶段为保健消费支出阶段,即老年人在参与保健消费的前提下,决定保健消费支出金额的大小。在实证分析中,一般采用对数线性模型处理因变量存在偏度的问题,模型形式为公式(3.2)所示:

$$\log(Y_j | P_j = 1) = \beta_2 X_j + u_{2j} \quad (3.2)$$

模型(3.2)中 $Y_j$ 表示参与保健消费的第 $j$ 个老年人保健支出金额, $X_j$ 表示影响受访者保健消费支出的因素, $u_{2j}$ 为随机扰动项,该模型作为传统对数线性模型,采用普通最小二乘法进行估计。

以上两个阶段共同构成了两部模型,两部模型将两个阶段的行为视为有先后顺序且相互独立的过程。

### 3.1.2 分层线性模型相关理论

#### 1. 模型适用范围及优势

分层线性模型主要用来处理嵌套结构数据。在社会科学类问题的研究中,变量间关系往往存在多层次的嵌套结构,如本文研究的老年人保健消费问题,保健消费不仅受性别等低层次个人微观因素的影响,同时也会受省域经济发展水平等地区宏观因素的影响。考虑到个体嵌套于地区这一分层数据特征,属于同一省域的不同个体的地区层面变量取值相同,由一个人的地区层面变量值可以获知该地区其他受访者的这一变量取值,这就违背了传统模型中解释变量观测独立性假定。

分层模型则解决了这一问题,分层模型假定每个高层单位都存在各自不同的回归模型,即不同的截距和斜率,而高层单位也是抽样的,因而认为不同的截距和斜率是从全部高层单位的截距和斜率中抽取的,是总体高层单位截距和斜率的随机样本。研究者可以通过分层模型探究微观因素及宏观因素对结局测量的影响,检验宏观因素是如何调节微观因素对结局测量的效应。且通过建立分层线性模型,可将结局变量的变

化分解为组内变异与组间变异，从而分析结局测量在低层次与高层次的相对变异情况。另外，分层模型不需要各个样本相互独立，所以可以修正由于观察数据的非独立性而引起的参数标准误的估计偏差。

## 2. 模型形式及估计方法

分层线性模型的建立需分为两个阶段，在第一阶段构建零模型，检验数据是否存在明显组间异质性，是否有必要构建分层模型；若数据适合建立分层模型，则进行第二阶段即完整模型的构建，逐步确定最终模型形式。分层模型可以为两水平、三水平或更多水平结构，下面以两层模型为例介绍模型基本框架：

### (1) 零模型

零模型又称为空模型、截距模型或无条件均值模型，该模型与单因素随机效应方差分析的原理相同，第一层与第二层模型中均不引入解释变量，只将模型分解为由个体差异造成的部分和由组群差异造成的部分，由此判断样本数据是否适合采用分层线性模型。零模型形式由公式 (3.3) 所示：

$$\begin{aligned} \text{第一层模型: } g(y_{ij}) &= \beta_{0j} + e_{ij} & \text{Var}(e_{ij}) &= \sigma^2 \\ \text{第二层模型: } \beta_{0j} &= \gamma_{00} + u_{0j} & \text{Var}(u_{0j}) &= \sigma_{u0}^2 \\ \text{组合模型: } g(y_{ij}) &= \gamma_{00} + u_{0j} + e_{ij} \end{aligned} \quad (3.3)$$

其中  $g(\cdot)$  为连接函数，依据因变量不同分布形式可选取单位函数、Logit 函数、Log 函数等作为连接函数，为叙述方便，下文以单位连接函数为例对零模型进行介绍。

第一层模型表达式中， $\beta_{0j}$  为第  $j$  组的因变量均值， $e_{ij}$  为围绕该均值的随机个体变异，表示组内个体之间的差异。第二层模型表达式中， $\gamma_{00}$  为总截距，表示因变量  $y_{ij}$  的各组总平均值， $u_{0j}$  则表示第  $j$  组的因变量均值与总平均值之间的差异，为组均值之间的变异。

上式中  $\sigma_{u0}^2$  表示组间方差， $\sigma^2$  表示组内方差，组内相关系数 ICC (Intra-Class Correlation Coefficient) 定义为组间方差与总方差之比，反映组间异质性与组内同质性。ICC 取值范围在 0 到 1 之间，越趋于零，组间变异越小，组群效应越弱；越趋于 1，组间变异越大，组群效应越强。一般来讲，若  $ICC = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma^2} > 0.059$ ，或组间方



差  $\sigma_{u_0}^2$  显著不为零, 则认为通过零模型检验, 数据存在组群效应, 适合构建分层线性模型。

## (2) 完整模型

若数据通过了零模型检验, 则说明样本数据存在组间差异, 有必要构建完整分层线性模型进行分析。完整模型包含了两个不同层次的解释变量, 分别以  $x$ 、 $z$  和  $w$  表示, 模型结构如公式 (3.4) 所示:

$$\begin{aligned}
 \text{第一层模型: } g(y_{ij}) &= \beta_{0j} + \sum_{p=1}^P \alpha_p x_{pij} + \sum_{q=1}^Q \beta_{qj} z_{qij} + e_{ij} \\
 \text{第二层模型: } \beta_{0j} &= \gamma_{00} + \sum_{m=1}^M \gamma_{0m} w_{mj} + u_{0j} \\
 \beta_{1j} &= \gamma_{10} + \sum_{m=1}^M \gamma_{1m} w_{mj} + u_{1j} \\
 &\dots\dots\dots \\
 \beta_{Qj} &= \gamma_{Q0} + \sum_{m=1}^M \gamma_{Qm} w_{mj} + u_{Qj} \\
 \text{组合模型: } g(y_{ij}) &= \gamma_{00} + \sum_{m=1}^M \gamma_{0m} w_{mj} + \sum_{p=1}^P \alpha_p x_{pij} + \sum_{q=1}^Q \gamma_{q0} z_{qij} + \sum_{q=1}^Q \sum_{m=1}^M \gamma_{qm} w_{mj} z_{qij} \\
 &\quad + (u_{0j} + \sum_{q=1}^Q z_{qij} u_{qj} + e_{ij})
 \end{aligned} \tag{3.4}$$

第一层模型截距  $\beta_{0j}$  为随机截距,  $P$  个水平 1 解释变量  $x_{pij}$  为固定效应,  $Q$  个水平 1 解释变量  $z_{qij}$  具有随机效应。每个水平 1 随机回归系数 ( $\beta_{0j}, \beta_{qj}, q=1, 2, \dots, Q$ ) 作为  $M$  个水平 2 解释变量  $w_{mj}$  的线性函数, 即具有随机效应的水平 1 解释变量系数受水平 2 模型中环境变量的影响, 由此构成了  $Q+1$  个方程。

多层模型的估计方法有最大似然法、限制性最大似然法、迭代广义最小二乘法、限制性迭代广义最小二乘法、经验贝叶斯估计法、自助法等。本文则采用最常用的限制性最大似然法, 以全残差为基础, 迭代过程中首先产生参数估计初始值, 一次迭代后产生的参数估计值作为下一次迭代的初始值, 迭代过程循环多次, 直至达到收敛标准, 所有参数估计值趋于稳定, 该过程借助多层次数据处理专用软件 HLM 对分层线性模型进行估计。

### 3.1.3 广义部分线性可加模型相关理论

#### 1. 模型适用范围及优势

广义部分线性可加模型是对标准线性模型的推广,每个变量用一个非线性函数替换,同时保持了模型的整体可加性。相比于传统线性模型,广义部分线性可加模型具有多方面优点。首先,广义部分线性可加模型适用范围更广,对因变量数据分布无过多限制,在无法确定响应变量分布情况或其不服从正态分布时,广义部分线性可加模型依然适用,对于二分类因变量、计数型因变量、正态分布连续型因变量及非正态分布连续型因变量,分别有与其对应的连接函数构建可加模型。其次,与传统线性模型的驱动机制不同的是,部分线性可加模型为数据驱动,数据决定了响应变量与解释变量之间的本质关系,不仅能得出各变量间的线性关系,还能描绘不同形式的非线性影响,不需为每个自变量分别手动设置不同的变形方式,就可以对每个变量建立非线性函数,得出更为灵活多变的回归模型。另外,相比于非参数模型,半参数部分线性可加模型可以避免变量个数过多时带来的维数灾难问题。且模型具有可加性,可以在保持其他变量不变的情况下探究某一自变量对因变量的影响,保留了可加结构的解释性,为分别检验各变量的影响效应提供了基础。这些优点使得广义部分线性可加模型对数据的拟合程度比传统线性模型高,模型解释意义更为丰富,且模型预测更为精准。部分线性可加模型即在可加模型探究非线性关系时加入线性部分,同时考虑线性影响因素与非线性影响因素,因此,对于可能存在非线性关系且未知具体关系形式的数据来说,建立广义部分线性可加模型是较好的处理方式。

#### 2. 模型形式及估计方法

广义部分线性可加模型的模型形式为:

$$g(E(Y | X_1, X_2, \dots, X_{p+r})) = \alpha + \sum_{i=1}^p \beta_i X_i + \sum_{j=1}^r s_j(X_{p+j}) \quad (3.5)$$

其中,  $\alpha$  为常数项,  $\sum_{i=1}^p \beta_i X_i$  为模型线性部分,  $\sum_{j=1}^r s_j(X_{p+j})$  为模型的非参数部分,  $s_j(X_{p+j})$  为各非参数变量的平滑函数。  $g(\cdot)$  为连接函数,由于部分线性可加模型较为灵活,对被解释变量的形式没有具体规定,对于被解释变量的不同分布形式,应采用不同的模型连接函数。因变量分布与连接函数选取的对应关系如表 3-1 所示:

表 3-1 因变量分布与其对应的连接函数

因变量分布	连接函数
二项分布	Logit 连接: $g(\mu) = \log(\mu/(1-\mu))$
负二项分布	倒数连接: $g(\mu) = 1/\mu$
正态分布	单位连接: $g(\mu) = \mu$
Gamma 分布	Log 连接: $g(\mu) = \log(\mu)$
Poisson 分布	Log 连接: $g(\mu) = \log(\mu)$

目前估计部分线性可加模型的方法主要有三种,第一种是向后拟合算法(Backfitting Algorithm),第二种为局部拟差分方法(Local Quasi-Differencing Approach),第三种为边际积分算法(Marginal Integration Method)。考虑到向后拟合算法的巧妙迭代与计算简便等优势,本文实证分析中将采用该方法进行模型估计,下面对向后拟合算法做简要介绍。

采用向后拟合算法对可加模型进行估计时,将模型线性部分看作一个特殊的非参数函数,记  $f_0(x_p) = \alpha + \sum_{i=1}^p \beta_i X_i$ ,则部分线性可加模型的估计转变为对

$g(E(y|x_1, x_2, \dots, x_{p+r})) = f_0(x_p) + \sum_{j=1}^r s_j(x_{p+j})$  中函数  $f_0(\cdot)$  与各个  $s_j(\cdot)$  的估计。在估计

其中一个  $s_k$  时,假定除了  $s_k$  以外的其他  $r-1$  个  $s_j(\cdot)$  以及  $f_0(\cdot)$  均已知,定义偏残差为

$u_k = g(E(y|x_1, x_2, \dots, x_{p+r})) - f_0(x_p) - \sum_{j=1, j \neq k}^r s_j(x_{p+j})$ ,通过最小化该偏残差得到估计函数

$\hat{s}_k$ ,循环该过程得到  $r$  个分量  $s_1, s_2, \dots, s_r$  的估计值,记为  $\hat{s}_1, \hat{s}_2, \dots, \hat{s}_r$ 。同理,在估计

$f_0(x_p)$  时,假定其余  $s_j(\cdot)$  均固定,得到最优估计值  $\hat{\beta}_i (i=1, 2, \dots, p)$ 。在具体实现时,首先

初始化函数  $\hat{f}_0^0, \hat{s}_1^0, \hat{s}_2^0, \dots, \hat{s}_r^0$ ,假定  $\hat{s}_2^0, \dots, \hat{s}_r^0$  以及  $\hat{f}_0^0$  固定,通过上述方法,

得到偏残差最小时的估计函数  $\hat{s}_1^1$ ,同样的步骤得到  $\hat{f}_0^1, \hat{s}_1^1, \hat{s}_2^1, \dots, \hat{s}_r^1$ ,不断

进行上述迭代过程,直到残差  $RSS = \sum_{j=1}^r [g(E(y|x_1, x_2, \dots, x_{p+r})) - f_0(x_p) - \sum_{j=1}^r s_j(x_{p+j})]^2$

达到事先设定的收敛标准,从而得到部分线性可加模型线性部分与非参数部分的最终估计结果。

### 3.2 基于分层线性模型的影响因素分析

为从宏微观两方面探究老年人保健消费决策影响因素,引入健康状况、认知能力、社会人口学特征、财富状况、代际支持和社会互动六个方面共 21 个微观指标,以及地区经济、价格水平、政府卫生支出和保健供给四个方面共 6 个地区宏观指标,并以地区编号作为两层数据的连接变量。将个人微观指标作为第一层、地区宏观指标作为第二层变量,以 0-1 变量“是否有保健消费”作为模型响应变量,以 Logit 函数作为连接函数,构建分层 Logistic 模型。

由于对于老年人保健消费问题的个人微观层面及地区宏观层面解释变量较多,并需要甄别变量的固定效应和随机效应以及考虑跨层交互效应,因此在构建分层模型时,不能将以上各方面一次性全部纳入模型,否则将使得模型混乱且准确性差,应逐步的、有条不紊的将上述变量纳入模型。为此,本文在查阅了大量文献后,最终按照以下步骤构建分层线性模型:

(1) 首先不放入任何解释变量,建立零模型,检验层-2 残差项是否显著不为零,并计算组内相关系数 ICC 值,若层-2 残差项通过了卡方检验,且 ICC 值大于 5.9%,则说明不同地区老年人保健消费存在显著差异,这时建立分层线性模型才有理论与实际上的支持;

(2) 将地区层面宏观变量引入模型层-2,筛选出显著变量,以解释地区间差异;

(3) 将个人层面微观变量引入模型层-1,筛选出显著变量,以解释组内差异,此时假定层-1 变量系数为固定效应,即不随地区变化而变化,在地区层面没有差异;

(4) 对层-1 变量系数是否存在随机效应进行检验,并对存在随机效应的变量探究其与层-2 变量是否存在交互效应;

(5) 综合上述步骤,得到最终完整模型,并对结果进行解释说明。

#### 1. 零模型

首先令层-1 与层-2 均不包含任何解释变量,建立零模型,以检验不同地区老年人保健消费决策是否存在显著差异,该检验的理论及效果与单因素方差分析一致。即检验层-2 的残差项是否显著不为 0,若通过该检验,层-2 残差项显著不为 0,则说明各个地区间的老年人保健消费决策存在显著差异,有必要建立分层模型。具体检验模型为公式 (3.6) 与 (3.7) 所示:

层-1 模型:

$$\begin{aligned} Prob(Y = 1 | \beta) &= P \\ \log[P/(1-P)] &= \beta_0 \end{aligned} \quad (3.6)$$

层-2 模型:

$$\beta_0 = -1.347 + u_0 \quad (3.7)$$

零模型检验结果如表 3-2 所示:

表 3-2 老年人保健消费决策零模型检验结果

随机效应	标准差	方差	自由度	卡方值	p 值
$u_0$	0.463	0.214	27	126.981	0.000

由表 3-2 可知, 来自地区层面的组间方差为 0.214, 随机效应的卡方检验统计量为 126.981,  $p$  值为 0.000, 在 1% 的显著性水平下通过了卡方检验, 说明层-2 残差项显著不为 0, 即不同地区的老年人保健消费决策存在显著差异。组内相关系数  $ICC = 0.214 / [0.214 + (\pi^2 / 3)] = 6.11\% > 5.9\%$ , 说明老年人保健消费决策差异有 6.11% 是由地区差异引起的, 从零模型检验  $p$  值角度与组内相关系数  $ICC$  角度均说明数据存在相当程度的组间异质性与组内同质性, 有必要建立分层模型, 同时考虑个人层面与地区层面因素对老年人保健消费决策的影响。

## 2. 用地区变量解释组间差异

由零模型得出的组内相关系数及随机效应的显著性可知, 老年人保健消费在不同地区间存在显著差异, 因此接下来引入地区层面的变量去描述这种差异。第一层模型形式与零模型相同, 在第二层模型中引入上述地区经济、政府卫生支出、保健消费价格及保健供给四个方面共 6 个宏观变量 (表示为  $Z_1, \dots, Z_6$ )。以 10% 的显著性水平作为判定标准, 寻找对被解释变量有显著影响的宏观变量, 为了避免遗漏重要变量, 保留  $p$  值接近显著性水平的变量, 得到的模型表示为公式 (3.8) 与 (3.9) 所示:

层-1 模型:

$$\begin{aligned} Prob(Y = 1 | \beta) &= P \\ \log[P/(1-P)] &= \beta_0 \end{aligned} \quad (3.8)$$

层-2 模型:

$$\beta_0 = 3.834 + 0.010 \log Z_1 - 0.046 Z_4 + u_0 \quad (3.9)$$

运行结果如表 3-3 及表 3-4 所示:

表 3-3 加入宏观影响因素的模型回归结果

变量	系数	标准误	t 值	自由度	p 值
$\log(\text{人均地区生产总值})$	0.010 ***	0.003	3.760	25	0.001
医疗保健类居民消费价格指数	-0.046 ***	0.013	-3.424	25	0.002

附注 1: \*\*、\*、\*分别表示当显著性水平为 1%、5%、10%时显著, 下同。

表 3-4 加入宏观影响因素后模型组间方差估计

随机效应	标准差	方差	自由度	卡方统计量	p 值
$u_0$	0.327	0.107	25	77.416	0.000

由表 3-3 统计结果可知, 在宏观层面, 人均地区生产总值与医疗保健类居民消费价格指数两个指标在 5% 的显著性水平下对老年人保健消费决策分别存在正向与负向的影响。

由表 3-4 可知, 零模型组间方差为 0.214, 引入地区层面宏观影响因素后, 模型组间方差估计值减小为 0.107, 表明新加入的人均地区生产总值与医疗保健类居民消费价格指数宏观变量对于组间差异存在一定解释能力, 约 50% 的地区老年人保健消费决策差异能够被人均地区生产总值与医疗保健类居民消费价格指数说明。

### 3. 纳入层-1 解释变量

在上述模型的基础上考虑个人层面微观因素对老年人保健消费的影响。本部分考虑受访老人健康状况、认知能力、社会人口学特征、财富状况、代际支持和社会互动六个方面共 21 个微观变量 (表示为  $X_1, \dots, X_{21}$ ), 为使建模过程清晰有逻辑, 暂不考虑变量随机效应 (将在下文第 4 个建模步骤中考虑), 当作固定效应引入模型。共得到 12 个微观显著变量, 模型表示为:

层-1 模型:

$$\begin{aligned} \text{Prob}(Y=1|\beta) &= P \\ \log[P/(1-P)] &= \beta_0 - 0.462X_1 + 0.187X_2 + 0.565X_4 + 0.016X_5 \\ &\quad + 0.028X_6 + 0.548X_7 + 0.069X_9 + 0.350X_{11} \\ &\quad + 0.323X_{13} + 0.001X_{18} - 0.001X_{19} + 0.091X_{21} \end{aligned} \quad (3.10)$$

层-2 模型:

$$\beta_0 = -1.386 + 0.013 \log Z_1 - 0.030Z_4 + u_0 \quad (3.11)$$

模型估计结果如表 3-5 所示:

表 3-5 加入宏微观影响因素的模型回归结果

变量	系数	标准误	t 值	自由度	p 值
log(人均地区生产总值)	0.013 *	0.007	1.966	25	0.060
医疗保健类居民消费价格指数	-0.030 *	0.015	-1.999	25	0.056
自评健康	-0.462 ***	0.079	-5.796	4026	0.000
慢性病患情	0.187 ***	0.024	7.714	4026	0.000
孤独感	0.565 ***	0.123	4.579	4026	0.000
认知能力	0.016 *	0.008	1.938	4026	0.052
年龄	0.028 ***	0.007	4.293	4026	0.000
性别	0.548 ***	0.095	5.763	4026	0.000
受教育年限	0.069 ***	0.012	5.661	4026	0.000
退休状况	0.350 ***	0.077	4.575	4026	0.000
户口类型	0.323 ***	0.097	3.326	4026	0.001
子女实物支持	0.001 ***	0.000	2.662	4026	0.008
子女看望次数	-0.001 ***	0.000	-3.898	4026	0.000
社交情况	0.091 ***	0.015	5.917	4026	0.000

由表 3-5 可以看出,在个人微观层面上,自评健康、慢性病患情、孤独感、年龄、性别、受教育年限、退休状况、户口类型、子女实物支持、子女看望次数及社交状况在 1% 的显著性水平下显著影响老年人保健消费决策,认知能力则在 10% 的显著性水平下显著影响保健消费决策。

#### 4. 层-1 随机斜率检验及跨层交互效应评估

本部分内容的主要目的是探究上一步骤模型中得出的年龄、性别、认知状况、社交情况等 12 个显著微观变量对老年人保健消费决策的影响是否随地区的变化而变化。将以上 12 个变量的系数设为随机性,即第二层模型中加入残差项,若残差项显著不为零,则说明该微观变量对老年人保健消费决策的影响存在地区差异。得到各项残差中只有“受教育年限”变量的残差显著不为零,检验结果如表 3-6 所示:

表 3-6 “受教育年限”变量随机效应检验

随机效应	标准差	方差	自由度	卡方统计量	p 值
$u_9$	0.030	0.001	27	42.052	0.074

表 3-6 表明,“受教育年限”变量的残差项显著,卡方检验对应的 p 值为 0.074,即在 10% 的显著性水平下残差项显著不为零,存在随机效应,受教育年限对老年人保健消费决策的影响存在地区差异性。

进一步要考虑地区层面的宏观变量人均地区生产总值和医疗保健类居民消费价格指数是否对受教育年限和保健消费决策的关系存在影响效应,即添加宏观变量与受教育年限变量的交互项,探究其是否显著。发现人均地区生产总值对数与受教育年限

变量交互项显著，即人均地区生产总值对受教育年限和保健决策的关系存在影响效应。且系数值为负，即人均地区生产总值弱化了受教育年限与保健决策的关系，人均地区生产总值越低，受教育年限对老年人保健消费的影响越显著。从现实意义来讲，即越贫困的地区，受教育程度的差异越能引起老年人保健消费概率的差异。

5. 完整模型

综合以上内容，整合层-1 微观变量与层-2 宏观变量以及宏微观变量交互项，得到最终模型 (3.12) 与 (3.13)。

模型形式：

层-1 模型：

$$\begin{aligned}
 & Prob(Y = 1 | \beta) = P \\
 & \log[P / (1 - P)] = \beta_0 - 0.461X_1 + 0.186X_2 + 0.562X_4 + 0.016X_5 \\
 & \quad + 0.028X_6 + 0.552X_7 + 0.661X_9 + 0.363X_{11} \\
 & \quad + 0.327X_{13} + 0.001X_{18} - 0.001X_{19} + 0.089X_{21}
 \end{aligned} \tag{3.12}$$

层-2 模型：

$$\begin{aligned}
 \beta_0 &= -1.555 + 0.012 \log Z_1 - 0.029Z_4 + u_0 \\
 \beta_9 &= 0.661 - 0.126 \log Z_1 + u_9
 \end{aligned} \tag{3.13}$$

模型最终回归结果如表 3-7 所示：

表 3-7 最终分层线性模型回归结果

变量	系数	比数比	标准误	t 值	自由度	p 值
log(人均地区生产总值)	0.012 *	1.012	0.006	1.948	25	0.062
医疗保健类居民消费价格指数	-0.029 **	0.971	0.014	-2.050	25	0.048
自评健康	-0.461 ***	0.631	0.078	-5.877	4025	0.000
慢性病患者	0.186 ***	1.204	0.025	7.555	4025	0.000
孤独感	0.562 ***	1.754	0.123	4.564	4025	0.000
认知能力	0.016 *	1.016	0.008	1.947	4025	0.051
年龄	0.028 ***	1.028	0.006	4.347	4025	0.000
性别	0.552 ***	1.737	0.096	5.778	4025	0.000
受教育年限	0.661 ***	1.937	0.192	3.442	4025	0.001
受教育年限*log(人均地区生产总值)	-0.126 ***	0.882	0.040	-3.132	4025	0.002
退休状况	0.363 ***	1.438	0.072	5.076	4025	0.000
户口类型	0.327 ***	1.387	0.099	3.314	4025	0.001
子女实物支持	0.001 ***	1.001	0.000	2.654	4025	0.008
子女看望次数	-0.001 ***	0.999	0.000	-3.984	4025	0.000
社交情况	0.089 ***	1.093	0.015	5.786	4025	0.000

由表 3-7 可知，老年人保健消费决策同时受地区层面宏观因素与个人层面微观因



素的影响。

宏观方面,地区生产总值对老年人保健消费决策在 10%的显著性水平下有正向影响,地区生产总值越高,地区经济发展越好,居民物质条件相对较高,且消费观念较为积极,因而老年人更容易存在保健消费行为;除此之外,在 5%的显著性水平下,医疗保健类居民消费价格指数对老年人保健消费决策存在显著负向影响,价格指数越高,保健消费概率越低,这也符合经济学基本原理。

微观方面,老年人保健消费决策受健康状况、认知能力、社会人口学特征、代际支持以及社会互动五个方面共 12 个变量的影响。

从健康状况角度,慢性病患者数越多的老年人由于生理需要,保健消费概率越大,慢性病患者数每增加 1 种,保健消费概率提高至原来的 1.204 倍;自评健康较好的老年人保健消费概率是健康较差老年人的 0.631;孤独感越强的老年人,越希望与其他人进行交流,保健消费能提供社会互动的渠道,老年人通过购买保健品可以获得更多的交流机会、交流话题和交流人群,也可以获得保健销售人员的情感慰藉,因而统计结果显示孤独的老年人是不孤独老年人保健消费概率的 1.754 倍。

从认知能力角度,认知能力越强的老年人,越能理性且正确的分析判断自己的身体状况,且结合医嘱选择适合自己的保健产品及保健服务,因此存在保健消费的概率较认知能力低的老年人更高。

从社会人口学特征角度,年龄越大,老年人身体机能越差,购买保健品的概率越高;女性老年人的保健消费概率是男性的 1.737 倍;已退休的老年人由于经济上的宽裕与时间上的充裕保健消费概率显著高于未退休老年人;城市老年人由于消费观念与经济水平的不同是农村老年人保健消费概率的 1.387 倍;受教育水平越高的老年人越容易存在保健消费,且受教育年限对老年人保健消费决策的影响存在地区差异性,人均地区生产总值弱化了受教育年限与保健决策的关系,人均地区生产总值越低,受教育年限对老年人保健消费概率的正向影响越显著,人均地区生产总值越高,受教育年限对老年人保健消费决策的影响越弱,从现实意义来讲,即越贫困的地区,受教育程度的差异越能引起保健消费决策的差异。

从代际支持角度,子女实物支持在 1%的显著性水平下正向影响保健消费概率,由于子女实物支持一定程度上降低了老年人经济负担,相对提高老年人物质条件,从而提高老年人保健品购买概率;子女看望能给予老年人情感上的慰藉,降低其孤独感,提高心理健康水平,降低了保健品需求,因此子女看望次数多的老人较子女看望次数

少的老人更不易产生保健消费行为。

从社会互动角度，一方面，购买保健品与保健服务的老年人在与销售人员的社会互动中得到亲情服务，受到尊重与关注；另一方面，老年人在与其他购买者的社会互动中实现了信息交流与消费愉悦分享。这些原因均会导致社会互动对老年人保健消费概率的正向影响，统计结果显示，老年人社交情况得分每增加 1 分，保健消费概率平均提高至原来的 1.093 倍。

总的来说，多层线性模型从宏观角度与微观角度全面探究了老年人保健消费决策影响因素，较传统单一线性模型传递了更多信息，对问题的剖析更加全面。

### 3.3 基于广义部分线性可加模型的影响因素分析

观察分层线性模型可知，老年人保健消费决策存在个人差异与地区差异，同时受微观因素与宏观因素的影响。由零模型可知，老年人保健消费决策的差异有 6.11% 是由地区差异引起的，其余 93.89% 的差异均由个人差异引起，且分层模型共得到 12 个显著微观变量与 2 个显著宏观变量，可以看出保健消费决策主要受微观因素的影响。前文分析了微观因素对老年人保健消费决策的线性影响，但线性假定不一定能真实反映其作用关系，实际的经济运行过程由于其复杂性与波动性，往往会表现出非线性特征，相关微观因素与保健消费决策间有可能存在复杂非线性关系，因此接下来引入半参数方法，控制地区因素，采用广义部分线性可加模型对老年人保健消费决策非线性微观影响因素进行探究。

引入前文健康状况、认知能力、社会人口学特征、财富状况、代际支持和社会互动六个方面共 21 个微观指标，由于自变量既有性别、工作状况、是否照顾孙辈等虚拟变量，又有年龄、个人年收入、子女实物支持等连续型变量，因此首先将虚拟变量作为可加模型的线性部分，将连续型变量作为非参数部分，构建部分线性可加模型。在模型构建之前，首先进行共曲线性检验，可加模型共曲线性将影响模型准确度，只有各非参数项的光滑函数之间不存在共曲线性才可以构建半参数可加模型。

#### 1. 共曲线性检验

在线性模型构建过程中，如果某变量与其他变量线性相关，则变量之间存在共线性，这将对模型的参数估计与准确度产生不利影响。同理，在可加模型中，如果某变量的光滑函数与其他非参数项的拟合值高度相关时，则说明模型中存在共曲线性。当共曲线性严重时，被估计的光滑函数变得不稳定，且可能出现同一组数据所得可加模

型的解不唯一的情况。

本文采用了 R 软件中用于共曲线性检验的 `estimate` 数值,该数值基于的计算理念为,模型中光滑函数  $s_i(X_i)$  所解释的变化空间  $a$  可以分解为两部分,一是位于其他光滑函数  $g(X)$  所解释的变化空间中的部分  $a_1$ ,二是不位于其他光滑函数所解释的变化空间部分  $a_2$ 。如果  $a_1$  占据了  $s_i(X_i)$  所解释的变化空间  $a$  的很大一部分,则存在共曲线性问题。`estimate` 数值是  $g(X)$  的基础 F-范数的平方除以  $s_i(X_i)$  基础 F-范数,衡量的是光滑函数  $s_i(X_i)$  在多大程度上被其他光滑函数  $g(X)$  解释,若该数值大于 0.5,则说明两光滑函数间存在共曲线性。

每个平滑项之间的成对一致性度量结果如表 3-8 所示:

表 3-8 部分线性可加模型共曲线性检验结果

	$s_2(X_2)$	$s_3(X_3)$	$s_5(X_5)$	$s_6(X_6)$	$s_9(X_9)$	$s_{14}(X_{14})$
$s_2(X_2)$	1.000	0.040	0.003	0.008	0.009	0.003
$s_3(X_3)$	0.044	1.000	0.047	0.006	0.025	0.017
$s_5(X_5)$	0.002	0.032	1.000	0.083	0.294	0.068
$s_6(X_6)$	0.046	0.004	0.090	1.000	0.024	0.003
$s_9(X_9)$	0.011	0.021	0.355	0.026	1.000	0.115
$s_{14}(X_{14})$	0.004	0.022	0.103	0.008	0.152	1.000
$s_{15}(X_{15})$	0.006	0.019	0.092	0.009	0.114	0.129
$s_{16}(X_{16})$	0.003	0.003	0.008	0.017	0.010	0.009
$s_{17}(X_{17})$	0.002	0.004	0.007	0.004	0.013	0.012
$s_{18}(X_{18})$	0.005	0.003	0.008	0.007	0.020	0.029
$s_{19}(X_{19})$	0.008	0.007	0.008	0.030	0.012	0.005
$s_{21}(X_{21})$	0.004	0.014	0.065	0.013	0.061	0.039
	$s_{15}(X_{15})$	$s_{16}(X_{16})$	$s_{17}(X_{17})$	$s_{18}(X_{18})$	$s_{19}(X_{19})$	$s_{21}(X_{21})$
$s_2(X_2)$	0.003	0.001	0.002	0.002	0.004	0.005
$s_3(X_3)$	0.008	0.003	0.001	0.003	0.004	0.023
$s_5(X_5)$	0.034	0.003	0.004	0.004	0.006	0.070
$s_6(X_6)$	0.003	0.001	0.002	0.002	0.018	0.006
$s_9(X_9)$	0.091	0.021	0.004	0.005	0.006	0.064
$s_{14}(X_{14})$	0.165	0.001	0.002	0.003	0.008	0.059
$s_{15}(X_{15})$	1.000	0.001	0.013	0.004	0.005	0.048
$s_{16}(X_{16})$	0.005	1.000	0.001	0.262	0.003	0.004
$s_{17}(X_{17})$	0.081	0.008	1.000	0.018	0.004	0.007
$s_{18}(X_{18})$	0.059	0.002	0.057	1.000	0.004	0.015
$s_{19}(X_{19})$	0.004	0.006	0.003	0.001	1.000	0.005
$s_{21}(X_{21})$	0.020	0.002	0.004	0.006	0.004	1.000

由表 3-8 各平滑项之间的一致性度量结果可知,各连续型自变量光滑函数之间的共曲线性 estimate 数值均小于 0.5,因此变量间不存在共曲线性,可以构建部分线性可加模型。

## 2. 模型初步构建

初步构建部分线性可加模型探究老年人保健消费决策的影响因素,由于被解释变量“老年人保健消费决策”为 0-1 二分类变量,因此选择 Logit 函数作为连接函数,将自评健康、孤独感、性别、婚姻状况、工作状况、退休状况、养老保险、户口类型、是否照顾孙辈 9 个虚拟变量引入模型线性部分,将慢性病患情、抑郁程度、认知能力、年龄、受教育年限、个人年收入、家庭金融资产、家庭负债、子女货币支持、子女实物支持、子女看望次数、社交情况 12 个连续型变量作为非参数部分,加入地区控制变量,构建广义部分线性可加模型。保留在 10%显著性水平下显著的变量,得到的模型为公式 (3.14),估计结果如表 3-9 与表 3-10 所示:

$$\begin{aligned} \log[P/(1-P)] = & 0.366 + s_2(X_2) + s_3(X_3) + s_5(X_5) + s_6(X_6) + s_9(X_9) \\ & + s_{16}(X_{16}) + s_{18}(X_{18}) + s_{19}(X_{19}) + s_{21}(X_{21}) - 0.310X_1 \\ & + 0.516X_4 + 0.612X_7 + 0.294X_8 + 0.475X_{11} + 0.299X_{13} \end{aligned} \quad (3.14)$$

表 3-9 广义部分线性可加模型线性部分初步回归结果

变量	系数	标准误	z 统计量	p 值
自评健康	-0.310 **	0.136	-2.274	0.023
孤独感	0.516 ***	0.130	3.960	0.000
性别	0.612 ***	0.116	5.296	0.000
婚姻状况	0.294 **	0.138	2.129	0.033
退休状况	0.475 ***	0.165	2.875	0.004
户口类型	0.299 *	0.155	1.929	0.054

表 3-10 广义部分线性可加模型非参数部分初步回归结果

变量	参考自由度	p 值	变量	参考自由度	p 值
s(慢性病患情)	6.178	0.000	s(家庭负债)	2.313	0.054
s(抑郁程度)	2.099	0.095	s(子女实物支持)	1.001	0.063
s(认知能力)	7.071	0.000	s(子女看望次数)	5.653	0.000
s(年龄)	1.000	0.000	s(社交情况)	3.597	0.000
s(受教育年限)	1.001	0.044			

由表 3-9 与表 3-10 可知,初步构建的半参数可加模型的线性部分得到 6 个显著变量,非参数部分得到 9 个显著光滑函数。观察非参数部分回归结果,发现年龄、受教育年限、子女实物支持三个变量的自由度接近于 1。模型自由度表示模型中不受限

制的待估计变量的个数,如果广义部分线性可加模型中某光滑函数的自由度接近于1(在0.8与1.2之间),表示该光滑函数中只有一个待估计参数,即为一元一次函数,没有必要将其作为模型非参数项,因此应将年龄、受教育年限、子女实物支持三个连续型变量从非参数部分删除,添加到模型线性部分,得到调整后的最终模型。

### 3. 最终模型

将自由度接近1的连续型变量由非参数部分调整到线性部分后,得到最终广义部分线性可加模型形式为公式(3.15),估计结果如表3-11与表3-12所示:

$$\begin{aligned} \log[P/(1-P)] = & -6.010 + s_2(X_2) + s_3(X_3) + s_5(X_5) + s_{16}(X_{16}) + s_{19}(X_{19}) \\ & + s_{21}(X_{21}) - 0.319X_1 + 0.512X_4 + 0.039X_6 + 0.602X_7 \\ & + 0.289X_8 + 0.870X_9 + 0.462X_{11} + 0.294X_{13} + 0.001X_{18} \end{aligned} \quad (3.15)$$

表3-11 广义部分线性可加模型线性部分最终回归结果

变量	系数	标准误	z 统计量	p 值
自评健康	-0.319 **	0.136	-2.346	0.019
孤独感	0.512 ***	0.130	3.928	0.000
年龄	0.039 ***	0.009	4.504	0.000
性别	0.602 ***	0.115	5.235	0.000
婚姻状况	0.289 **	0.138	2.090	0.037
受教育年限	0.870 **	0.369	2.361	0.018
退休状况	0.462 ***	0.165	2.798	0.005
户口类型	0.294 *	0.155	1.898	0.058
子女实物支持	0.001 *	0.000	1.797	0.061

表3-12 广义部分线性可加模型非参数部分最终回归结果

变量	估计自由度	p 值	变量	估计自由度	p 值
s(慢性病患情)	6.368	0.000	s(家庭负债)	2.314	0.072
s(抑郁程度)	2.052	0.099	s(子女看望次数)	5.649	0.000
s(认知能力)	7.028	0.000	s(社交情况)	3.597	0.000

由最终构建的广义部分线性可加模型,共得到15个变量对老年人保健消费决策存在显著影响,其中9个变量为线性影响,6个变量为非参数部分的非线性影响,非线性影响可由偏残差图3-1直观显示:

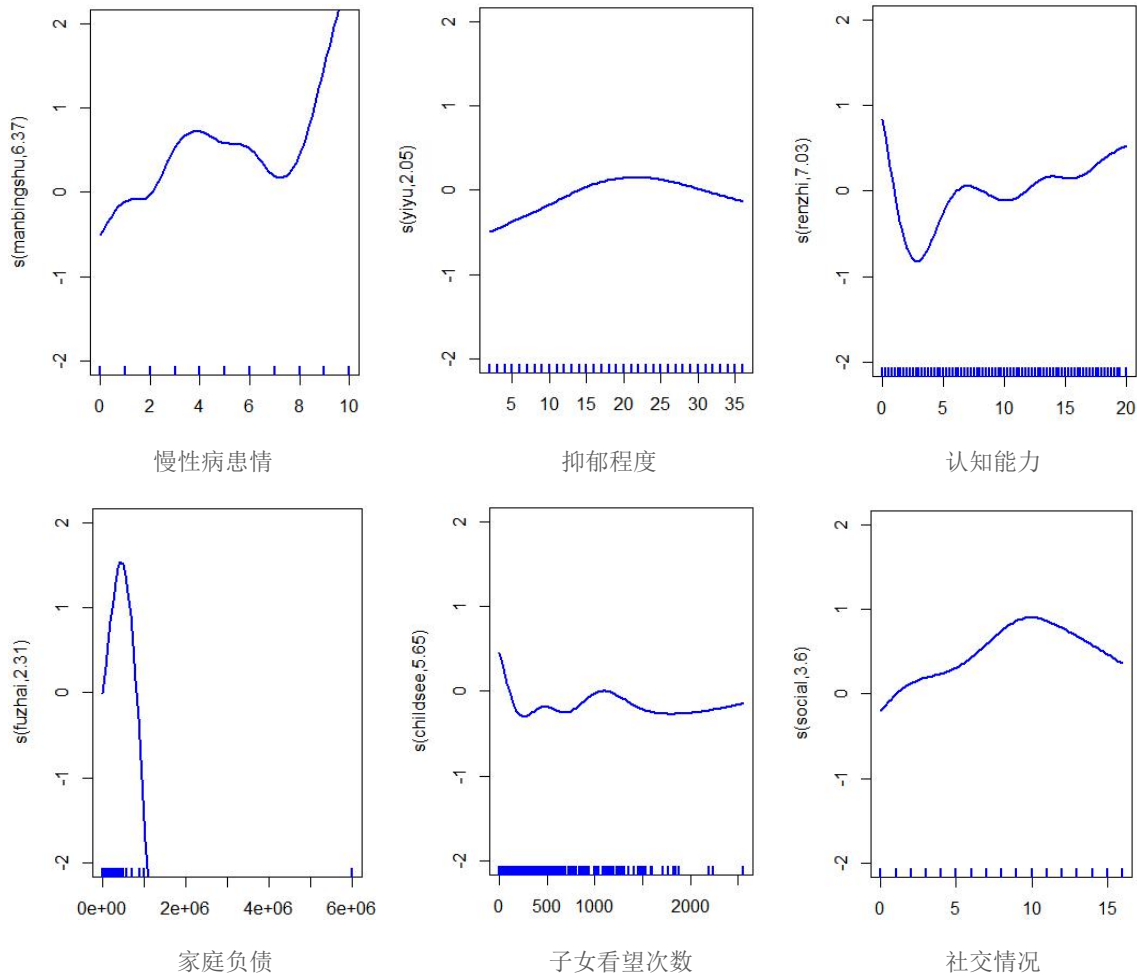


图 3-1 老年人保健消费决策非线性影响因素偏残差图

对于老年人保健消费决策影响因素的探究，广义部分线性可加模型与分层线性模型的统计结果对比如表 3-13 所示：

表 3-13 可加模型与分层模型所得影响因素对比

指标名称	分层线性模型系数符号	广义部分线性可加模型系数符号	整合结果
$\log(\text{人均地区生产总值})$	正	—	正
医疗保健类居民消费价格指数	负	—	负
自评健康	负	负	负
孤独感	正	正	正
年龄	正	正	正
性别	正	正	正
婚姻状况	—	正	正
受教育年限	正	正	正
受教育年限* $\log(\text{人均地区生产总值})$	负	—	负
退休状况	正	正	正
户口类型	正	正	正
子女实物支持	正	正	正

表 3-13 可加模型与分层模型所得影响因素对比 (续表)

指标名称	分层线性模型系数符号	广义部分线性可加模型系数符号	整合结果
慢性病患情	正	非线性	非线性
抑郁程度	—	非线性	非线性
认知能力	正	非线性	非线性
家庭负债	—	非线性	非线性
子女看望次数	负	非线性	非线性
社交情况	正	非线性	非线性

附注 1: “—”表示影响不显著。

由表 3-13 可知,对于老年人保健消费决策影响因素的研究,分层线性模型从宏观两个方面进行分析,得到人均地区生产总值和医疗保健类居民消费价格指数 2 个宏观变量、自评健康和孤独感等 12 个微观变量以及受教育年限与人均地区生产总值交互项对保健消费概率存在显著影响。广义部分线性可加模型则进一步发现了婚姻状况对老年人保健消费决策的线性影响以及慢性病患情、抑郁程度、认知能力、家庭负债、子女看望次数和社交情况的非线性影响,其他微观影响因素及影响方向与分层线性模型均相同。

对于广义部分线性可加模型得出的慢性病患情、抑郁程度、认知能力、家庭负债、子女看望次数及社交情况 6 个非线性影响因素,观察图 3-1,图形横坐标表示相应自变量的取值,纵坐标表示偏残差,即排除其他变量影响后所得到的由相应自变量造成的老年人保健消费概率的差异,反映了某个自变量对因变量的影响效果。

慢性病患情对老年人保健消费概率的影响为随着慢性病患者种类数的增多先正向、后负向、再正向地影响因变量。这可能是由于在慢性病患情较轻微时,无需太多药物治疗,依靠保健品就可以改善身体机能,因此在慢性病患者数为 0-4 种时,对保健消费概率为正向影响;当慢性病患者数增加到 5-7 种时,患病数较多,身体机能变差,老年人主要依靠效果更为显著的药物来控制病情,不再过多依靠效果较慢的保健品,因此该阶段慢性病患情与保健消费概率呈负相关关系;而当慢性病种数增加至 8-11 种时,老人身体机能则变得格外差,需要同时依靠药品和保健品控制病情,因此慢性病患情与保健消费概率再次呈现正向相关关系。从实际意义角度,半参数模型结果相比于线性模型的简单正向相关关系更具解释力。

抑郁程度对老年人保健消费概率的影响曲线先上升后下降,即随着抑郁程度的加重,对保健消费概率的影响先正而后负向,抑郁程度得分为 23 时,保健消费概率达

到峰值附近。这可能是由于当抑郁程度较轻时，一般不需采取药物治疗，只需要保健品即可缓解，所以老年人会更倾向于购买保健品；而抑郁程度加重时，易导致身心健康状况变差，服用保健品的效果不明显，更多的需要依靠药物及心理干扰进行治疗，因此保健品消费概率降低。对于线性模型无法捕捉的变量间的非线性关系，广义部分线性可加模型可做出合理解释。

认知能力对老年人保健消费决策的影响整体上呈现先负向后正向的关系，认知能力得分为 2.5 时，保健消费概率达到最低值附近。老年人认知水平很低时，一般是健康状况或智力水平存在问题，认知能力越差，问题越严重，基于生理需要，对保健品的需求越高，因此此时的认知能力与保健消费概率为反向关系；当认知水平在正常范围内时，认知水平越高，分析能力越强，越有可能根据自己的实际需要购买保健品，老年人保健消费概率随之提高。相比于线性模型得出的认知能力与老年人保健消费决策之间的简单正向相关关系，广义部分线性可加模型所得经济意义更加丰富。

家庭负债对老年人保健消费的影响关系为先正向后负向，当负债较少时，不足以影响生活水平，可能会由于超前消费的家庭理念增大保健消费概率；而家庭负债较多时，经济压力与思想压力随之加大，将会减少非必需开支，作用效果较慢的保健品消费概率也会降低，呈现负向相关关系。对于家庭负债与老年人保健消费概率之间的复杂关系，线性模型未能得出，而广义部分线性可加模型做出了精准描述。

子女看望次数与老年人保健消费决策的关系为随着子女看望次数的增多，老年人保健消费概率先下降后不变。这可能是由于老年人容易因为孤独感而从保健品与保健品销售人员处获得情感慰藉，从而购买保健品与保健服务，而适当的子女陪伴可以减轻老年人孤独感，削弱老年人对保健品的需求与依赖，一定程度上降低保健消费概率；而当子女看望次数达到一定程度后，老年人孤独感几乎为零，这时再增加子女陪伴，边际效用递减，几乎不再对老年人保健消费概率产生影响。相比于线性模型得出的简单负向相关关系，部分线性可加模型所得到的经济意义更加丰富且接近现实。

随着社交频率的增加，老年人保健消费概率表现为先上升后下降。社交情况得分小于 10 时，社交频率越大，可能会由于老年人之间的相互宣传与相互影响以及老年人与保健销售人员的接触而使得老年人更容易存在保健消费；而当社交情况得分大于 10 时，一方面，过多的社交活动会消耗老年人很多时间和精力，从而无暇考虑购买保健品，另一方面，能参与如此多的社交活动的老年人一般身体素质较好，无需服用保健品，这两个方面都会使得社交频率对老年人保健消费概率产生削弱作用。相比于



线性模型得出的简单正向相关关系,广义部分线性可加模型对社交情况与老年人保健消费决策之间关系的描述更切合实际。

总的来说,针对老年人保健消费决策影响因素问题的探究,广义部分线性可加模型对比于前文分层线性模型,进一步发现了婚姻状况的线性影响,以及慢性病患情、抑郁程度、认知能力、家庭负债、子女看望次数、社交情况6个因素的非线性影响机制,得出更具解释意义的统计结论。

### 3.4 本章小结

本章对老年人保健消费决策影响因素进行了实证分析,3.1节简要介绍了实证部分所涉及的两部模型思想、分层线性模型理论以及广义部分线性可加模型理论。3.2节同时引入个人层面微观指标与地区层面宏观指标,构建分层线性模型,得到了地区生产总值与医疗保健类居民消费价格指数2个宏观影响因素与自评健康、孤独感、年龄、性别、受教育年限、退休状况、户口类型、子女实物支持、慢性病患情、认知能力、子女看望次数和社交情况12个微观影响因素,以及受教育年限与人均地区生产总值交互项的影响作用。为对微观影响因素进行深入分析,3.3节构建广义部分线性可加模型,进一步发现了婚姻状况对老年人保健消费决策的线性影响以及慢性病患情、抑郁程度、认知能力、家庭负债、子女看望次数与社交情况的非线性影响关系,所得结果有较强的解释意义。

## 第4章 老年人保健消费支出影响因素实证分析

根据两部模型的思想,在建立保健消费决策模型之后,分离存在保健消费即保健消费金额不为零的样本,共788例,并使用该部分样本建立保健消费支出模型,进一步探究对于存在保健消费的受访老人,影响其具体保健消费金额的因素。

### 4.1 基于分层线性模型的影响因素分析

对于老年人保健消费支出问题,首先同时考虑宏观因素与微观因素,构建分层线性模型。解释变量包括健康状况、认知能力、社会人口学特征、财富状况、代际支持和社会互动六个方面共21个微观指标,以及地区经济、价格水平、政府医疗保健支出和保健供给四个方面共6个地区宏观指标,并以地区编号作为两层数据的连接变量。将个人微观指标作为第一层变量,地区宏观指标作为第二层变量,以保健消费金额作为该模型的响应变量,构建分层线性模型。

#### 1. 正态性检验

为确定分层模型连接函数,首先对存在保健消费的老年人的保健消费支出金额数据做Q-Q图进行正态性检验,得到图4-1。

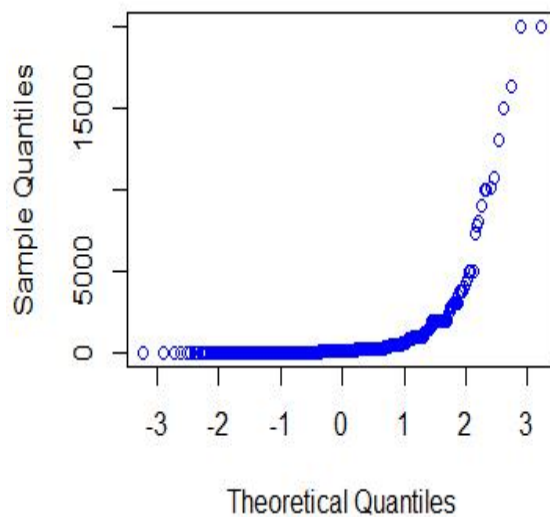


图4-1 老年人保健消费支出Q-Q图

由图4-1可以看出,Q-Q图上的点并没有大致分布在一条直线上,说明数据不服从正态分布,存在保健消费的样本中,其保健消费金额分布呈现一定程度偏态。因此取对数减弱分布偏态性,并检验对数处理后的保健支出数据是否服从正态分布。

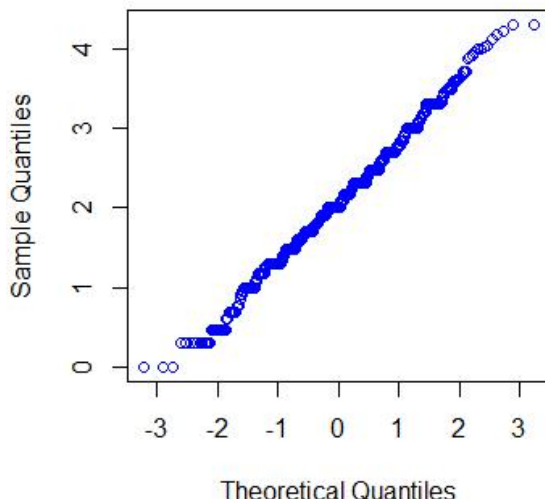


图 4-2 老年人保健消费支出对数 Q-Q 图

观察图 4-2，发现 Q-Q 图上的点近似分布在一条直线附近，说明对数处理后的老年人保健支出数据近似服从正态分布。因此可以以保健支出对数为因变量，构建分层线性模型。

### 2. 零模型

首先令层-1 与层-2 均不包含任何解释变量，建立零模型，以检验不同地区老年人保健消费支出是否存在显著差异。该检验的理论及效果与单因素方差分析一致，即检验层-2 的残差项是否显著不为 0，目的是评估组内同质性及组间异质性。若通过该检验则说明老年人保健消费支出金额存在显著地区差异，单个人层面微观因素不足以全面说明影响机制，有必要引入地区层面宏观因素，建立分层线性模型。具体检验模型为公式 (4.1) 与 (4.2)：

层-1 模型：

$$\log Y = \beta_0 + r \tag{4.1}$$

层-2 模型：

$$\beta_0 = 2.09 + u_0 \tag{4.2}$$

零模型检验结果如表 4-1 所示：

表 4-1 老年人保健消费支出零模型检验结果

随机效应	标准差	方差	自由度	卡方值	p 值
$u_0$	0.188	0.035	27	59.697	0.000
$r$	0.741	0.549			

由表 4-1 可知，来自地区层面的方差为 0.035，随机效应  $u_0$  的卡方检验统计量为

59.697,  $p$  值为 0.000, 在 1% 的显著性水平下通过了卡方检验, 说明层-2 的残差项显著不为 0。组内相关系数  $ICC = u_0 / (u_0 + r) = 0.035 / (0.035 + 0.549) = 6.02\% > 5.9\%$ , 说明老年人保健消费支出的差异有 6.02% 是由地区差异引起的。从零模型检验  $p$  值角度与组内相关系数  $ICC$  角度均说明样本数据存在明显组间异质性或组内同质性, 有必要建立分层线性模型, 同时考虑个人层面与地区层面因素对老年人保健消费支出金额的影响, 更加全面的分析相关问题。

### 3. 用地区变量解释组间差异

由零模型得出的组内相关系数及随机效应的显著性可知, 老年人保健消费在不同地区间存在显著差异, 因此引入地区层面宏观变量。第一层模型形式与零模型相同, 在第二层模型中引入人均地区生产总值、居民人均可支配收入、地方财政医疗保健支出、医疗保健类居民消费价格指数、医疗卫生机构数以及卫生人员数 6 个宏观变量(表示为  $Z_1, \dots, Z_6$ )。以 10% 的显著性水平为判定标准, 保留对响应变量有显著影响的解释变量。得到以下模型:

层-1 模型:

$$\log Y = \beta_0 + r \quad (4.3)$$

层-2 模型:

$$\beta_0 = -0.361 + 0.567 \log Z_2 + u_0 \quad (4.4)$$

运行结果如表 4-2:

表 4-2 加入宏观影响因素的模型回归结果

变量	系数	标准误	t 值	自由度	p 值
$\log$ (居民人均可支配收入)	0.567 *	0.327	1.734	26	0.094

表 4-3 加入宏观变量后模型组内方差与组间方差估计

随机效应	标准差	方差	自由度	卡方统计量	p 值
$u_0$	0.151	0.023	26	57.269	0.001
$r$	0.741	0.549			

由表 4-2 所示的模型统计结果表明, 在宏观层面, 居民人均可支配收入对老年人保健消费支出金额在 10% 的显著性水平下存在显著正向影响, 即地区居民人均可支配收入越高, 该地区老年人保健消费支出金额就越高。

由表 4-3 可知, 纳入宏观影响因素后的模型组内方差为 0.549, 与空模型的组内

方差估计值相同，而组间方差却由空模型的 0.035 较大幅度减小为 0.023。这表明新加入的居民人均可支配收入宏观变量不能对组内变异进行说明，而对组间变异存在一定解释能力，约 34.29% 的地区间老年人保健消费金额差异能够被居民人均可支配收入解释。

#### 4. 纳入层-1 解释变量

引入地区宏观变量后，考虑保健消费支出同时还受个人因素影响，接下来在上述模型的基础上加入个人层面微观因素。将受访老人健康状况、认知能力、社会人口学特征、财富状况、代际支持和社会互动六个方面共 21 个微观变量（表示为  $X_1, \dots, X_{21}$ ）引入第一层模型，最终共得到 7 个微观显著变量。

模型表示为公式（4.5）与（4.6）：

层-1 模型：

$$\log Y = \beta_0 + 0.011X_3 + 0.007X_6 - 0.109X_{10} + 0.114X_{12} + 0.256X_{13} + 0.034 \log X_{15} + 0.001X_{18} + r \quad (4.5)$$

层-2 模型：

$$\beta_0 = -0.053 + 0.405 \log Z_2 + u_0 \quad (4.6)$$

回归结果由表 4-4 所示：

表 4-4 加入宏微观影响因素的模型回归结果

变量	系数	标准误	t 值	自由度	p 值
$\log$ （居民人均可支配收入）	0.405 **	0.203	1.997	26	0.046
抑郁程度	0.011 ***	0.004	2.809	779	0.006
年龄	0.007 **	0.004	1.989	779	0.047
工作状况	-0.109 **	0.054	-2.026	779	0.043
养老保险	0.114 *	0.063	1.825	779	0.068
户口类型	0.256 ***	0.071	3.610	779	0.001
$\log$ （家庭金融资产）	0.034 *	0.019	1.773	779	0.076
子女实物支持	0.001 ***	0.000	4.498	779	0.000

表 4-5 加入宏微观变量后模型组内方差与组间方差估计

随机效应	标准差	方差	自由度	卡方统计量	p 值
$u_0$	0.110	0.012	26	42.004	0.024
$r$	0.623	0.388			

由表 4-4 可知，在个人微观层面上，存在 7 个因素对老年人保健消费支出金额存在显著影响。在 1% 的显著性水平下，抑郁程度与子女实物支持对老年人保健支出金额存在正向影响，非农村老人保健支出显著高于农村老人；在 5% 的显著性水平下，

年龄对保健金额存在正向影响，有工作老年人保健支出显著低于无工作老年人；在10%的显著性水平下，家庭金融资产越高，受访老人保健支出越高，有养老保险者的保健支出显著高于无养老保险者。

由表4-5可知，纳入层-1变量后，与只纳入层-2变量相比，组内方差估计量由0.549减小为0.388，表明有29.38%的老年人保健消费支出个体水平变异能够被层-1微观变量解释说明。组间方差估计量由0.023减小为0.012，表明将层-1变量引入模型后，不仅可以解释层-1变异，而且可以解释层-2变异。说明层-1微观变量既可在个体水平，也可在群组水平上影响老年人保健消费支出金额，因为该组微观变量的值可能既存在组内变异，也存在组间变异。

#### 5. 层-1随机斜率检验及跨层交互效应评估

本部分内容的主要目的是探究上一步骤模型中得出的年龄、户口类型、工作状况、养老保险、抑郁程度、家庭金融资产以及子女实物支持7个显著微观变量对老年人保健消费支出的影响是否随地区的变化而变化。将以上7个变量的系数设为随机性，即模型中加入残差项，若残差项显著不为零，则说明该变量对因变量的影响存在地区差异。

对上述7个微观变量进行卡方检验，对应p值均大于0.1，即残差项均与零没有显著差异，不存在随机效应，各微观变量对老年人保健消费支出的影响不存在地区差异。因此将该7个变量均作为固定效应引入模型，不再考虑宏微观变量交互效应。

#### 6. 完整模型

综合以上内容，整合层-1微观变量与层-2宏观变量，得到最终层-1模型与层-2模型的模型形式如上文公式(4.5)与(4.6)所示。回归结果如上文中表4-4所示。

由表4-4分层线性模型估计结果可知，老年人保健消费支出金额同时受地区层面宏观因素与个人层面微观因素的影响。

宏观方面，居民人均可支配收入对老年人保健消费支出存在正向影响，即地区居民人均可支配收入越高，居民的购买能力越强，该地区老年人保健消费支出金额就越高，居民人均可支配收入每增加1%，老年人保健支出金额平均增加0.4%。

微观方面，老年人保健消费支出金额受健康状况、社会人口学特征、财富状况以及代际支持四个方面共7个变量的影响。

从健康状况角度，生理健康问题可能更多导致老年人医疗消费的增高而不是保健消费的增高，因而模型统计结果显示生理健康状况与老年人保健消费支出金额并无显

著联系；而由抑郁程度所度量的心理健康状况对老年人保健消费金额存在显著正向影响，即抑郁程度越高，心理健康状况越差，越能引发老年人的保健消费需求，保健消费支出也就越高，抑郁程度得分每增加1分，老年人保健消费金额平均增加1.1%。

从社会人口学特征角度，年龄越大，由于身体机能越差等原因，保健消费支出金额越高；无工作的老年人由于有足够的时间关注保健品与保健服务，因此保健消费支出金额较有工作的老年人平均高出10.9%；有养老保险的老年人由于一般物质条件更为宽裕，经济状况预期较为乐观，相比无养老保险的老年人平均保健消费支出金额高11.4%；非农户口老人由于经济状况与消费观念的不同比农村户口老人保健消费支出金额平均高25.6%。

从财富状况角度，由家庭现金数、金融机构存款数、现有债券面值、持有股票当前价值以及现持有的基金当前价值共同构成的家庭金融资产对数值对老年人保健消费支出金额存在显著正向影响，即老年人家庭经济状况越好，保健消费支出金额也就越高。

从代际支持角度，子女实物支持一定程度上降低了老年人经济负担，相对提高老年人物质条件，因而子女实物支持对老年人保健消费支出金额在1%的显著性水平下存在显著正向影响。子女实物支持价值量每增加1元，老年人保健消费支出金额平均增加0.1%。

总的来说，经上述步骤建立的分层线性模型从宏微观两个方面探究老年人保健消费支出金额的影响因素，较普通线性模型得出了更丰富的结论，对实际问题的剖析更加全面。

## 4.2 基于广义部分线性可加模型的影响因素分析

观察前文分层线性模型中零模型可知，老年人保健消费支出的差异有6.02%是由地区差异引起的，其余93.98%的差异均由个人差异引起，且分层线性模型最终得到7个显著微观变量与1个显著宏观变量，可以看出老年人保健消费支出的影响因素主要来源于个人微观层面。线性假定不一定能真实反映其作用关系，实际的经济运行过程由于其复杂性与波动性，往往会表现出非线性特征，相关微观因素与保健消费支出金额间有可能存在复杂非线性关系。因此接下来引入半参数方法，控制地区因素，采用广义部分线性可加模型对老年人保健消费支出非线性微观影响因素进行探究。

同样引入健康状况、认知能力、社会人口学特征、财富状况、代际支持和社会互

动六个方面共 21 个微观指标，将虚拟变量作为线性部分，将连续型变量作为非参数部分，构建部分线性可加模型。

### 1. 共曲线性检验

与老年人保健消费决策模型相同，在建立可加模型之前，首先进行共曲线性检验，若共曲线性评估值大于 0.5，则说明两光滑项之间存在共曲线性，不能直接建立可加模型。经检验，每个平滑项之间的成对一致性度量结果如表 4-6 所示：

表 4-6 部分线性可加模型共曲线性检验结果

	$s_2(X_2)$	$s_3(X_3)$	$s_5(X_5)$	$s_6(X_6)$	$s_9(X_9)$	$s_{14}(X_{14})$
$s_2(X_2)$	1.000	0.048	0.023	0.020	0.009	0.007
$s_3(X_3)$	0.036	1.000	0.143	0.088	0.040	0.023
$s_5(X_5)$	0.018	0.076	1.000	0.168	0.209	0.068
$s_6(X_6)$	0.016	0.014	0.195	1.000	0.056	0.016
$s_9(X_9)$	0.007	0.050	0.369	0.074	1.000	0.108
$s_{14}(X_{14})$	0.014	0.055	0.147	0.031	0.223	1.000
$s_{15}(X_{15})$	0.012	0.041	0.190	0.023	0.142	0.050
$s_{16}(X_{16})$	0.011	0.019	0.016	0.026	0.014	0.014
$s_{17}(X_{17})$	0.030	0.024	0.018	0.022	0.026	0.024
$s_{18}(X_{18})$	0.015	0.013	0.015	0.032	0.011	0.014
$s_{19}(X_{19})$	0.012	0.013	0.057	0.057	0.052	0.023
$s_{21}(X_{21})$	0.011	0.049	0.149	0.053	0.061	0.018
	$s_{15}(X_{15})$	$s_{16}(X_{16})$	$s_{17}(X_{17})$	$s_{18}(X_{18})$	$s_{19}(X_{19})$	$s_{21}(X_{21})$
$s_2(X_2)$	0.021	0.027	0.015	0.009	0.008	0.017
$s_3(X_3)$	0.042	0.001	0.005	0.014	0.005	0.115
$s_5(X_5)$	0.138	0.011	0.013	0.019	0.005	0.156
$s_6(X_6)$	0.001	0.010	0.014	0.026	0.026	0.036
$s_9(X_9)$	0.168	0.012	0.018	0.025	0.007	0.108
$s_{14}(X_{14})$	0.154	0.020	0.013	0.026	0.004	0.090
$s_{15}(X_{15})$	1.000	0.004	0.007	0.028	0.007	0.126
$s_{16}(X_{16})$	0.005	1.000	0.004	0.243	0.005	0.042
$s_{17}(X_{17})$	0.030	0.006	1.000	0.061	0.028	0.021
$s_{18}(X_{18})$	0.091	0.035	0.140	1.000	0.004	0.048
$s_{19}(X_{19})$	0.045	0.013	0.025	0.040	1.000	0.030
$s_{21}(X_{21})$	0.079	0.013	0.013	0.046	0.025	1.000

由表 4-6 平滑项之间的一致性度量结果可知，各非参数项光滑函数之间的共曲线性数值均小于 0.5，即引入模型的连续型变量光滑函数间不存在共曲线形性，可以直接构建部分线性可加模型。



## 2. 模型初步构建

为探究老年人保健消费支出金额的微观影响因素，以对数函数作为连接函数，将自评健康、孤独感、性别等 9 个虚拟变量引入模型线性部分，将年龄、认知能力、个人年收入、子女看望次数等 12 个连续型变量引入模型非参数部分，加入地区控制变量，构建广义部分线性可加模型。保留在 10% 的显著性水平下显著的变量，剔除不显著变量，初步构建部分线性可加模型的模型形式为公式 (4.7) 所示，回归结果如表 4-7 与表 4-8 所示：

$$\log Y = 2.810 + s_3(X_3) + s_6(X_6) + s_9(X_9) + s_{15}(\log X_{15}) + s_{18}(X_{18}) + s_{19}(X_{19}) - 0.238X_{10} + 0.163X_{12} + 0.175X_{13} \quad (4.7)$$

表 4-7 广义部分线性可加模型线性部分初步回归结果

变量	系数	标准误	t 值	p 值
工作状态	-0.238 ***	0.070	-3.385	0.001
养老保险	0.163 **	0.065	2.495	0.013
户口类型	0.175 **	0.072	2.428	0.016

表 4-8 广义部分线性可加模型非参数部分初步回归结果

变量	估计自由度	p 值	变量	估计自由度	p 值
s(抑郁程度)	1.000	0.001	s(log 家庭金融资产)	1.000	0.043
s(年龄)	1.000	0.048	s(子女实物支持)	1.932	0.010
s(受教育年限)	3.967	0.047	s(子女看望次数)	3.827	0.064

由表 4-7 与表 4-8 可知，初步构建的半参数可加模型共得到 9 个显著变量，其中线性部分得到 3 个显著影响因素，非线性部分得到 6 个显著影响因素。进一步观察估计结果可以发现，抑郁程度、年龄、家庭金融资产三个变量的自由度为 1，这意味着对于其光滑函数来说，不受限制的待估计参数个数为 1，即为一元一次函数，因此这三个因素与老年人保健消费支出为线性关系，没有必要将其放在非参数部分进行研究，应将抑郁程度、年龄、家庭金融资产三个变量从非参数部分删除，添加到模型线性部分，得到调整后的最终模型。

## 3. 最终模型

将自由度为 1 的连续型变量抑郁程度、年龄、家庭金融资产由非参数部分调整到线性部分后，重新拟合模型，得到最终的部分线性可加模型形式为公式 (4.8) 所示，回归结果如表 4-9 与表 4-10 所示：

$$\log Y = 2.422 + s_9(X_9) + s_{18}(X_{18}) + s_{19}(X_{19}) + 0.021X_3 + 0.168X_6 - 0.232X_{10} + 0.171X_{12} + 0.157X_{13} + 0.052 \log X_{15} \quad (4.8)$$

表 4-9 广义部分线性可加模型线性部分最终回归结果

变量	系数	标准误	t 值	p 值
抑郁程度	0.021 ***	0.006	3.657	0.000
年龄	0.168 ***	0.063	2.677	0.008
工作状况	-0.232 ***	0.070	-3.296	0.001
养老保险	0.171 ***	0.065	2.623	0.009
户口类型	0.157 **	0.074	2.119	0.035
log(家庭金融资产)	0.052 **	0.023	2.251	0.025

表 4-10 广义部分线性可加模型非参数部分最终回归结果

变量	估计自由度	p 值
s(受教育年限)	3.974	0.020
s(子女实物支持)	1.854	0.024
s(子女看望次数)	3.678	0.087

由表 4-9 与表 4-10 可知，共得到 9 个变量对老年人保健消费支出金额有着显著影响，其中抑郁程度、年龄、工作状况、养老保险、户口类型与家庭金融资产 6 个变量对因变量为线性影响，受教育年限、子女实物支持与子女看望次数对因变量存在显著非线性影响，对于非线性影响，由图 4-3 偏残差图进行直观展示：

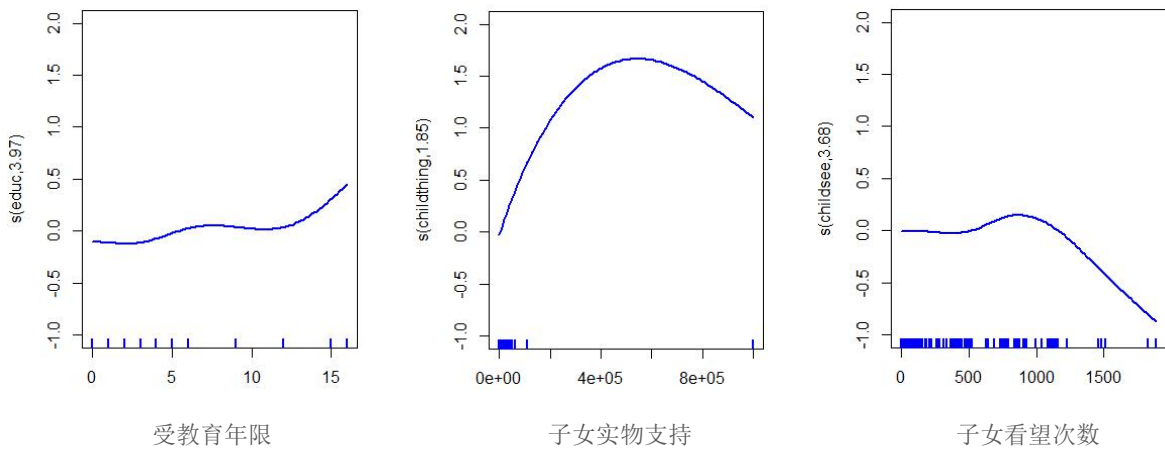


图 4-3 老年人保健消费支出非线性影响因素偏残差图

对于老年人保健消费支出金额影响因素的探究，广义部分线性可加模型与分层线性模型的统计结果对比如表 4-11 所示：

表 4-11 可加模型与分层模型所得影响因素对比

指标名称	分层线性模型系数符号	部分线性可加模型系数符号	整合结果
log(居民人均可支配收入)	正	—	正
抑郁程度	正	正	正
年龄	正	正	正
工作状况	负	负	负

表 4-11 可加模型与分层模型所得影响因素对比 (续表)

指标名称	分层线性模型系数符号	部分线性可加模型系数符号	整合结果
养老保险	正	正	正
户口类型	正	正	正
$\log$ (家庭金融资产)	正	正	正
受教育年限	—	非线性	非线性
子女实物支持	正	非线性	非线性
子女看望次数	—	非线性	非线性

附注 1: “—” 表示影响不显著。

由表 4-11 可知, 对于老年人保健消费支出影响因素的研究, 分层线性模型从宏观微观两个方面进行分析, 得到居民人均可支配收入显著宏观变量以及年龄等微观影响因素。广义部分线性可加模型进一步发现了受教育年限、子女实物支持与子女看望次数的非线性影响, 其他微观影响因素及影响方向均与分层线性模型相同。

对老年人保健消费支出金额存在显著影响的非参数部分指标为受教育年限、子女实物支持以及子女看望次数。观察图 4-3, 其横坐标表示相应解释变量的取值, 纵坐标表示排除其他变量影响后所得到的由该解释变量造成的老年人保健消费支出金额的差异, 即偏残差, 反映了某个解释变量对老年人保健消费支出的作用效果。

受教育年限变量对老年人保健消费支出存在正向影响, 但影响作用由小变大, 其残差图在受教育年限取值为 0-12 时较为平缓, 取值大于 12 即存在高等教育经历时曲线较为陡峭。当受教育年限较低时, 不同受教育水平的个人综合素质与消费观念差距不大, 如六年受教育年限的小学学历与九年受教育年限的初中学历并不会造成太大的个人差异, 因此对老年人保健消费支出的影响较小; 而当受教育水平较高时, 教育年限的差异就会引起个人消费能力与消费观念较大的不同, 保健消费支出金额也就存在较大差异。对于线性模型无法度量的受教育年限与老年人保健消费支出之间的非线性关系, 广义部分线性可加模型可做出合理解释。

子女实物支持对老年人保健消费支出的影响曲线先上升后下降, 即随着子女实物支持价值量的增多, 对老年人保健消费支出金额的影响先正而后负向。这可能是由于子女对老年人衣物和食物等的实物支持可以使老人减少该部分开支, 降低经济负担, 结余下的钱就可能会购买更多的保健产品与保健服务, 保健消费金额也就增多; 而当子女一年内实物支持价值量达到 50 万元甚至更多时, 这其中很可能已经包含了保健品类的实物支持, 而不需老年人自己过多购买, 因此老年人保健消费支出金额随之下降。相比于线性模型得出的简单正向相关关系, 广义部分线性可加模型对子女实物支

持与老年人保健消费支出之间的非线性关系描述更具解释意义。

子女看望次数较少时,对老年人保健支出并无明显影响,而当一年内子女看望次数超过 1000 时,很可能意味着老人与子女同住,这时可能由于老年人从子女处获得情感慰藉,会把更多的精力放在子女和大家庭上,从而忽视保健品的购买;也可能由于同住子女负担了老年人多数保健消费,因此老年人保健消费支出金额下降。对于线性模型无法度量的子女看望次数与老年人保健消费支出之间的非线性关系,广义部分线性可加模型可做出合理解释。

总的来说,针对老年人保健消费支出微观影响因素的研究,部分线性可加模型对比于前文分层线性模型,进一步发现了受教育年限、子女实物支持与子女看望次数 3 个因素对老年人保健消费支出的非线性影响作用,得到了更具现实意义的统计结果。

为检验模型稳健性,根据 CHARLS 相关数据构造老年人日常生活自理能力变量与工具性日常生活活动能力变量,替换原模型中度量健康状况的自评健康与慢性病患病变量,所得结论与原模型一致。此外,分别去掉退休状况变量、用居住地类型变量替换户口类型变量构建模型,并分别对城镇老年人与乡村老年人子样本进行回归,得到的结果均与原模型几乎完全一致,一定程度上验证了模型稳健性。

### 4.3 老年人保健消费决策与支出影响因素对比

根据所构建的分层线性模型与广义部分线性可加模型结果,对比老年人保健消费决策与保健消费支出影响因素,得到表 4-12:

表 4-12 老年人保健消费决策与支出影响因素比较

指标类别	指标名称	保健消费决策模型系数符号	保健消费支出模型系数符号
宏观因素	$\log$ (人均地区生产总值)	正	—
	$\log$ (居民人均可支配收入)	—	正
	医疗保健类居民消费价格指数	负	—
健康状况	自评健康	负	—
	慢性病患者情	非线性	—
	抑郁程度	非线性	正
	孤独感	正	—
认知能力	认知能力	非线性	—
社会人口学特征	年龄	正	正
	性别	正	—
	婚姻状况	正	—
	受教育年限	正	非线性

表 4-12 老年人保健消费决策与支出影响因素比较 (续表)

指标类别	指标名称	保健消费决策模型系数符号	保健消费支出模型系数符号
社会人口学特征	受教育年限*log(人均地区生产总值)	负	—
	工作状况	—	负
	养老保险	—	正
	退休状况	正	—
	户口类型	正	正
财富状况	log(家庭金融资产)	—	正
	家庭负债	非线性	—
代际支持	子女实物支持	正	非线性
	子女看望次数	非线性	非线性
社会互动	社交情况	非线性	—

附注 1: “—”表示影响不显著。

由表 4-12 可知,老年人保健消费决策与保健消费支出金额均受抑郁程度、年龄、受教育年限、户口类型、子女实物支持与子女看望次数的影响。其中年龄、户口类型、子女看望次数对老年人保健消费决策与保健消费支出的影响机制相同;抑郁程度对老年人保健消费概率为先正向后负向的非线性影响,而对保健消费金额为正向线性影响;受教育年限与子女实物支持变量对保健消费决策为正向影响,而对保健消费金额为更复杂的非线性影响。除此之外,人均地区生产总值、医疗保健类居民消费价格指数、自评健康、孤独感、性别、婚姻状况、退休状况对老年人保健消费决策存在线性影响,慢性病患情、认知能力、家庭负债、社交情况对消费决策存在非线性影响,而对老年人保健消费金额无显著影响;居民人均可支配收入、工作状况、养老保险和家庭金融资产对老年人保健消费金额存在线性影响,而对保健消费决策无显著影响。

#### 4.4 本章小结

本章对老年人保健消费支出影响因素进行了实证分析。4.1 节同时引入个人层面微观指标与地区层面宏观指标,构建分层线性模型,得到了居民人均可支配收入宏观影响因素与抑郁程度、年龄、工作状况、养老保险、户口类型、家庭金融资产、子女实物支持 7 个微观影响因素。为对微观影响因素进行深入分析,4.2 节构建广义部分线性可加模型,进一步发现了受教育年限、子女实物支持与子女看望次数对老年人保健消费支出的非线性影响关系,所得结果有较强的解释意义。4.3 节对老年人保健消费决策与保健消费支出影响因素进行了比较,直观展示了老年人保健消费决策与保健消费支出影响因素的相同点与不同点。

## 第5章 结论与展望

### 5.1 结论与建议

#### 5.1.1 研究结论

老年人由于特殊的生理机制与心理状态,使得他们更加倾向于通过医疗保健手段改善身体机能,保健品也因此受到越来越多老年人的关注,老年人保健消费问题一度成为社会热点。本文主要对老年人保健消费影响因素进行探究,首先建立老年人保健消费影响因素宏微观指标体系,确定了健康状况、认知能力、社会人口学特征、财富状况、代际支持和社会互动6个方面共21个微观指标与地区经济、政府卫生支出、保健消费价格与保健供给4个方面共6个宏观指标。然后针对老年人保健消费金额零膨胀数据特征,基于两部模型原理将老年人保健消费行为分为保健消费决策与保健消费支出两个相互独立的过程。同时考虑个人层面微观因素与地区层面宏观因素,建立分层线性模型研究变量间线性关系,并进一步建立广义部分线性可加模型,从非参数角度深入探究微观因素对保健消费的非线性影响,得到的研究结论如下:

1. 对于老年人保健消费决策,共得到17个影响因素:

(1) 从宏观角度,得到人均地区生产总值与医疗保健类居民消费价格指数2个线性影响因素。人均地区生产总值对老年人保健消费概率存在正向影响,地区经济发展越好,居民生活水平相对较高,消费观念较为积极,更容易存在保健消费行为;医疗保健类居民消费价格指数越高,保健消费概率越低,符合经济学基本原理。

(2) 从健康状况角度,得到自评健康、孤独感2个线性影响因素以及慢性病患者、抑郁程度2个非线性影响因素。自评健康较好的老年人保健消费概率是健康较差老年人的0.631;孤独感越强的老年人,越希望通过购买保健品获得更多的交流机会、交流话题和交流人群,孤独的老年人是不孤独老年人保健消费概率的1.754倍。随着慢性病患者种类数的增多,对老年人保健消费概率表现为先正向、后负向、再正向的非线性影响,这可能是因为当慢性病患者数为0-4种时,身体状况不会太差,无需过多药物治疗,多依靠保健品进行调理,因此慢性病患者对保健消费概率表现为正向影响;当慢性病患者数增加到5-7种时,身体机能变差,老年人主要依靠效果更为显著的药物来控制病情,不再过多依靠效果较慢的保健品,因此该阶段慢性病患者与保健消费概率呈负相关关系;而当慢性病种数增加至8-11种时,老人身体机能则变得格

外差,需要同时依靠药品和保健品控制病情,因此慢性病患情与保健消费概率再次呈正向相关。抑郁程度对老年人保健消费概率的非线性影响曲线先上升后下降,即随着抑郁程度的加重,对保健消费概率的影响先正向后负向,这可能是由于当抑郁程度较轻时,一般不需采取药物治疗,只需要保健品即可缓解,所以老年人会更倾向于购买保健品;而抑郁程度加重时,易导致身心健康状况变差,服用保健品的效果不明显,更多的需要依靠药物及心理干扰进行治疗,因此保健品消费概率降低。

(3)从认知能力角度,得到1个非线性影响因素。随着认知能力的提升,对老年人保健消费概率表现为先负向后正向的非线性影响,认知能力得分为2.5时,保健消费概率达到最低值。老年人认知能力得分低于2.5时,一般健康状况或智力水平存在问题,认知能力越差,问题越严重,基于生理需要,对保健品的需求越高,因此此时的认知能力与保健消费概率为反向关系;当认知水平在正常范围内时,认知水平越高,分析能力越强,越有可能根据自己的实际需要购买保健品,老年人保健消费概率随之提高。

(4)从社会人口学特征角度,得到年龄、性别、婚姻状况、退休状况、户口类型、受教育年限及其与人均地区生产总值交互项6个线性影响因素。年龄越大,老年人身体机能相对越差,越容易产生保健消费;由于男性与女性消费习惯不同,使得女性老年人保健消费概率是男性的1.737倍;已婚、已退休老年人保健消费概率分别高于未婚、未退休老年人;城市老年人由于消费观念与经济水平的不同是农村老年人保健消费概率的1.387倍;受教育年限越高,保健消费概率越大,且人均地区生产总值弱化了受教育年限与保健决策的关系,越贫困的地区,受教育程度的差异越能引起保健消费决策的差异。

(5)从财富状况角度,得到家庭负债1个非线性影响因素。家庭负债对老年人保健消费概率表现为先正向后负向的非线性影响,当负债较少时,不足以影响生活水平,可能会由于超前消费的家庭消费理念增大保健消费概率;而家庭负债较多时,经济压力与思想压力随之加大,将会减少非必需开支,作用效果较慢的保健品消费概率也会降低,呈现负向相关关系。

(6)从代际支持角度,得到子女实物支持1个线性影响因素与子女看望次数1个非线性影响因素。子女实物支持价值量对老年人保健消费概率存在显著正向影响,由于子女实物支持一定程度上降低了老年人经济负担,相对提高老年人物质条件,从而提高老年人保健品购买概率。子女看望次数与老年人保健消费决策的关系为随着子

女看望次数的增多,老年人保健消费概率先下降后不变,这可能是由于老年人希望从保健品与保健品销售人员处获得情感慰藉,而适当的子女陪伴可以减轻老年人孤独感,削弱老年人对保健品的需求与依赖,一定程度上降低保健消费概率;而当子女看望次数达到一定程度后,子女看望效用边际递减,几乎不再对老年人保健消费概率产生影响。

(7) 从社会互动角度,得到社交情况 1 个非线性影响因素。随着社交频率的增加,对老年人保健消费概率表现为先上升后下降的非线性影响,社交情况得分小于 10 时,社交频率越大,可能会由于老年人之间的相互宣传与相互影响以及老年人与保健销售人员的接触而使得老年人更容易存在保健消费;当社交情况得分大于 10 时,一方面,过多的社交活动会消耗老年人很多时间和精力,从而无暇考虑购买保健品,另一方面,能参与如此多的社交活动的老年人一般身体素质较好,无需服用保健品,这两个方面都会使得社交频率对老年人保健消费概率产生削弱作用。

2. 对于老年人保健消费支出金额,共得到 10 个影响因素:

(1) 从宏观角度,得到居民人均可支配收入 1 个线性影响因素。居民人均可支配收入越高,居民的购买能力越强,保健消费金额越高,居民人均可支配收入每增加 1%,老年人保健支出金额平均增加 0.4%。

(2) 从健康状况角度,得到抑郁程度 1 个线性影响因素。抑郁程度越高,心理健康状况越差,越能引发老年人的保健消费需求,保健消费支出也就越高,抑郁程度得分每增加 1 分,老年人保健消费金额平均增加 1.1%。

(3) 从社会人口学特征角度,得到年龄、工作状况、养老保险、户口类型 4 个线性影响因素与受教育年限 1 个非线性影响因素。年龄越大,保健消费支出越高;无工作的老年人由于有较多闲暇时间关注保健产品与保健服务,其保健消费支出金额较有工作的老年人平均高出 10.9%;有养老保险的老年人由于收入预期较为乐观,保健消费支出比无养老保险的老年人平均高 11.4%;非农村老年人由于经济状况较好且消费观念不同,保健消费支出较农村老年人平均高 25.6%。受教育年限变量对老年人保健消费支出存在正向影响,但影响作用由小变大,当受教育年限较低时,不同受教育水平的个人综合素质与消费观念差距不大,如六年受教育年限的小学学历与九年受教育年限的初中学历并不会造成太大的个人差异,因此对老年人保健消费支出的影响较小;而当受教育水平较高时,教育年限的差异就会引起个人消费能力与消费观念较大



的不同,保健消费支出金额也就存在较大差异,受教育水平越高,越能根据自身需要选择适合的保健产品与保健服务,保健消费支出金额越高。

(4) 从财富状况角度,得到家庭金融资产 1 个线性影响因素。由家庭现金数、金融机构存款数、现有债券面值、持有股票当前价值以及现持有的基金当前价值共同构成的家庭金融资产对数值对老年人保健消费支出金额存在正向显著影响,即老年人家庭物质基础越牢固,保健消费支出金额越高。

(5) 从代际支持角度,得到子女实物支持与子女看望次数 2 个非线性影响因素。随着子女实物支持价值量的增多,对老年人保健消费支出金额表现为先正而后负向的非线性影响,这可能是由于子女对老年人衣物和食物等的实物支持可以使老人减少该部分开支,降低经济负担,结余下的钱就可能会购买更多的保健产品与保健服务,保健消费金额也就增多;而当子女一年内实物支持价值量达到 50 万元甚至更多时,这其中很可能已经包含了保健品类的实物支持,而不需老年人自己过多购买,因此老年人保健消费支出金额随之下降。子女看望次数对老年人保健消费金额同样存在非线性影响,子女看望次数较少时,对老年人保健支出并无明显影响;而当一年内子女看望次数超过 1000 时,很可能意味着老人与子女同住,这时老年人可能会把更多的精力放在子女和大家庭上,从而忽视保健品的购买,也可能由于同住子女负担了老年人多数保健消费,因此老年人保健消费支出金额下降。

### 5.1.2 建议

本文深入剖析了老年人保健消费现状及影响因素,拓展了老年群体相关问题的研究领域,同时为保健品领域的研究提供分析依据,为老年人保健方面的研究提供了分析视角与实证材料。结合本文研究结论,提出如下建议:

1. 针对老年人自身来说,老年人受教育程度普遍较低,且主动学习意识淡薄,但社会是不断发展进步的,这便造成老年人辨别能力较弱,对此,老年人应学习养生保健常识,提高明辨是非的能力,充分认识自身健康状况、机体需求以及经济能力,理性购买消费品。同时,老年人应积极参与社交活动,在社会互动中充分发挥自身优势,开阔个人视野,实现个人价值,在与人交流中也降低了内心孤独感,得到精神慰藉,分散对保健产品与保健服务的注意力。另外,老年人应合理锻炼身体,科学养生,通过多种途径提高身体机能,而不是单一依靠药物与保健产品。

2. 针对老年人家庭来说, 老年人多数离开工作岗位, 家庭社会环境发生改变, 孤独感增强, 这时子女在为父母提供经济支持及生活照料的基础上, 还应多关心陪伴父母, 加强沟通和交流, 给予老年人一定情感慰藉, 引导其树立正确的观念, 以避免老人盲目不合理的保健消费。

3. 针对保健品企业来说, 不能虚假宣传保健品功能, 夸大保健产品效果, 人为诱导老年人进行保健消费, 应遵循诚信营销的理念, 严格把控保健品质量, 合理制定保健品价格。另外, 保健品部门应充分分析地区宏观现状与老年人个体特征, 有针对性地进行保健品与保健服务的推广。例如可以把经济发达地区的已退休、自评健康较差、孤独感较强的女性老年人作为保健品重点推广对象, 销售成功率将会增大; 把无工作、有养老保险、家庭金融资产较高的老年人发展为稳定客户, 其保健消费金额往往较大, 保健品销售额将会快速提升。

4. 针对政府来说, 有关部门应关注老年人生理及心理需求, 做好保健品理性消费观的宣传与普及, 积极倡导老年人树立正确的健康观, 科学养生。另外, 政府应严格把控对保健品市场的监管, 加大对媒体广告制作和宣传的检查力度, 健全监管机制, 严格追究保健市场的违法违规行为, 促进保健消费市场的平稳健康发展。

根据研究所得的结论, 本文分别从老年人自身、老年人家庭、保健品企业以及政府四个角度给出相关建议。由于本文实证部分均采用 2015 年数据, 数据存在时效性, 因此以上建议均针对于 2015 年的情况。

## 5.2 研究展望

1. 由于本文涉及众多个人微观指标, 对比各微观数据库后, 确定采用“中国健康与养老追踪调查”(CHARLS)数据库中的数据。但目前 CHARLS 中的数据仅更新至 2015 年, 因此本文无法获取最新数据进行研究, 接下来将关注相关数据更新情况, 及时对该研究进行补充。

2. 受本人能力所限, 对于老年人保健消费影响因素问题, 并未同时引入宏观变量与微观变量构建分层半参数模型探究其非线性影响, 接下来我将继续就该问题进行学习与探究, 以期得出更全面的结论。

## 参考文献

- [1] Karma H. The Grey Market[J]. *Economists*, 2016(04): 09-15
- [2] 徐丽萍, 王小林, 郭平, 等. 中国老年人生活成本和标准消费系数测算[J]. *人口与发展*, 2011, 17(03): 26-34
- [3] Banerjee S. Expenditure Patterns of Older Americans[J]. *Employee Benefit Research Institute*, 2012(368): 75-87
- [4] 茅锐, 徐建炜. 人口转型、消费结构差异和产业发展[J]. *人口研究*, 2014, 38(03): 89-103
- [5] 王俊, 龚强, 王威. “老龄健康”的经济学研究[J]. *经济研究*, 2012, 47(01): 134-150
- [6] 宋璐, 左冬梅. 农村老年人医疗支出及其影响因素的性别差异: 以巢湖地区为例[J]. *中国农村经济*, 2010(05): 74-85
- [7] 魏瑾瑞, 张睿凌. 老龄化、老年家庭消费与补偿消费行为[J]. *统计研究*, 2019, 36(10): 87-99
- [8] 韩雪梅, 王增福. 城市居民医疗保健消费的省际差异及影响因素分析[J]. *甘肃社会科学*, 2015(02): 183-186
- [9] Kukk M. How did Household Indebtedness Hamper Consumption During the Recession? Evidence from Micro Data[J]. *Journal of Comparative Economics*, 2016, 44(3): 764-786
- [10] 周晶, 韩央迪, Mao Weiyu. 照料孙子女的经历对农村老年人生理健康的影响[J]. *中国农村经济*, 2016(07): 81-96
- [11] Kim J, Richardson V. The Impact of Poverty, Chronic Illnesses, and Health Insurance Status on Out-of-Pocket Health Care Expenditures in Later Life[J]. *Social Work in Health Care*, 2014, 53 (10): 932-949
- [12] Bíró A. Supplementary Private Health Insurance and Health Care Utilization of People Aged 50+[J]. *Empirical Economics*, 2014, 46(2): 501-524
- [13] 臧文斌, 刘国恩, 徐菲, 等. 中国城镇居民基本医疗保险对家庭消费的影响[J]. *经济研究*, 2012, 47(07): 75-85
- [14] 田侃, 瞿华. 我国城乡居民医疗保健服务消费实证研究[J]. *财政研究*, 2015(07): 46-50

- [15] 王学义, 张冲. 中国人口年龄结构与居民医疗保健消费[J]. 统计研究, 2013, 30(03): 59-63
- [16] Cantarero P.D, Santiago L.P. Decomposing the Determinants of Health Care Expenditure: the Case of Spain[J]. The European Journal of Health Economics, 2012, 13(1): 19-27
- [17] Bech M, Christiansen T, Khoman E, etc. Ageing and Health Care Expenditure in EU-15[J]. The European Journal of Health Economics, 2011, 12(5): 469-478
- [18] 刘国恩, 蔡春光, 李林. 中国老人医疗保障与医疗服务需求的实证分析[J]. 经济研究, 2011, 46(03): 95-107
- [19] Chou Win-Lin. Explaining China's Regional Health Expenditures Using LM-type Unit Root Tests[J]. Health Economics, 2007, 26(4): 682-698
- [20] 聂聆, 李斌. 城镇居民医疗保健消费研究——基于粤苏浙三省的相关情况分析[J]. 广东商学院学报, 2012, 25(06): 64-70
- [21] Molla A.A, Chi Chunhui, Mondaca A.L.N. Predictors of High Out-of-pocket Healthcare Expenditure: an Analysis Using Bangladesh Household Income and Expenditure Survey[J]. BMC Health Services Research, 2017, 17(1): 152-192
- [22] 聂裕. 城镇居民医疗保健消费的区域性差异分析[J]. 时代金融, 2018(15): 181-184
- [23] 徐冬林, 陈永伟. 环境质量对中国城镇居民健康支出的影响[J]. 中国人口资源与环境, 2010, 20(04): 159-164
- [24] 叶小青, 李先玲. 环境质量、居民收入与医疗健康支出关系研究——基于截面相关的省际面板数据模型[J]. 统计与信息论坛, 2013, 28(12): 42-46
- [25] Yang Jing, Zhang Bing. Air Pollution and Healthcare Expenditure: Implication for the Benefit of Air Pollution Control in China[J]. Environment International, 2018, 39(5): 443-455
- [26] 谈煜鸿. 老龄化背景下我国老年人保健品消费研究述评[J]. 经济论坛, 2016(06): 108-111
- [27] 厉杰. 对老年人购买保健品行为的市场调查统计分析[J]. 商场现代化, 2018(19): 12-13

- [28] 刘斯敏, 曾学文. 基于多元定序 Logit 模型的中药保健品消费影响因素研究——以南宁市为例[J]. 中国市场, 2017(27): 91-95
- [29] 石贝贝. 我国城乡老年人口消费的实证研究——兼论“退休—消费之谜” [J]. 人口研究, 2017, 41(03): 53-64
- [30] 王菲. 我国城市老年人消费行为的实证研究[J]. 人口与发展, 2015, 21(03): 101-112
- [31] 肖波涛. 基于消费者态度的成都市保健品市场细分研究[D]. 成都: 西南财经大学, 2014
- [32] 陈茗, 刘素青. 老年人消费类型分析——二、三线城市个案研究[J]. 老龄科学研究, 2015, 3(01): 11-18
- [33] 王沛立, 李恩平. 我国居民医疗负担及其影响因素分析——基于弹性网方法的实证研究[J]. 数学的实践与认识, 2019, 49(14): 97-107
- [34] 邓敏, 杨莉, 陈娜. 医养结合下老年人医疗消费行为影响因素分析——以南京市为例[J]. 中国卫生政策研究, 2017, 10(01): 52-57
- [35] 冯桂平, 刘爽. 基于跨理论模型的老年人医疗保健品消费行为研究[J]. 消费经济, 2018, 34(04): 74-81
- [36] 戴平生, 李芳芳. 基于误差空间模型城乡居民医疗保健消费影响因素的实证分析[J]. 中国卫生统计, 2012, 29(04): 514-519
- [37] 苏春红, 李齐云, 王大海. 基本医疗保险对医疗消费的影响——基于 CHNS 微观调查数据[J]. 经济与管理研究, 2013(10): 23-30
- [38] 叶春辉, 封进, 王晓润. 收入、受教育水平和医疗消费: 基于农户微观数据的分析[J]. 中国农村经济, 2008(08): 16-24
- [39] 李涛, 徐翔, 张旭妍. 孤独与消费——来自中国老年人保健消费的经验发现[J]. 经济研究, 2018, 53(01): 124-137
- [40] 胡宏伟, 张小燕, 郭牧琦. 老年人医疗保健支出水平及其影响因素分析——慢性病高发背景下的老年人医疗保健制度改革[J]. 人口与经济, 2012(01): 97-104

## 致 谢

时光荏苒，一转眼又到了毕业季。两年前第一次走进山东财经大学参加复试的场景还历历在目。在山财，我认识了很多知识渊博、和蔼可亲的老师；结交了很多志趣相投、乐观优秀的朋友；参加了许多丰富多彩、妙趣横生的活动。能成为山东财经大学的一员让我倍感荣幸。在这段学习生活中有太多的事值得我铭记，亦有太多的人值得我感谢。

硕士生涯即将结束之际，首先想感谢的是我的导师张艳丽老师。入学以来，张老师经常给我们推荐专业类书籍与文献，帮助我们不断积累理论知识。对于毕业论文，无论是论文选题、研究思路确定还是论文写作与修改阶段，张老师都给予了我很多帮助，每当我在论文写作过程中遇到问题时，她总能给我提供解决思路，使我豁然开朗。在生活中，张老师也教会了我许多为人处世的道理，对即将踏入社会的我受用无穷。感谢张老师两年来的督促和指导，在今后的生活和工作中，我会以张艳丽老师为榜样，激励自己不断进步。

同时，我要感谢孔令涛老师和戴洪帅老师，研究生一年级阶段每周一次的学习小组讨论使我获益匪浅，锻炼了我学术研究、语言表达等多方面的能力，感谢老师们的陪伴与辛苦付出。感谢所有统计学院老师们，为我们传授各门课程知识精髓，为我们的学术研究打下坚实的专业理论基础。在此，向老师们致以最崇高的敬意和最诚挚的感谢，愿老师工作顺利。

其次，感谢同师门的王明宣师兄和李丹华师姐，在学习与工作中给了我极大的帮助与指导。感谢我的同学们，感谢你们这两年的陪伴，仍记得上课时一起探讨问题，期末考试时一起紧张的备考，假期时一起出游。特别是与我一起学习、生活的室友们，谢谢你们对我的论文提出的建议以及生活上给予我的帮助，在你们身上我看到了很多值得我学习的闪光点。希望同学们都顺利毕业，未来前程似锦。

另外，感谢我的家人，感谢他们一直以来对我无条件的支持，让我衣食无忧，安心学习，给了我他们能给予的最好的生活。我会努力学习，努力工作，努力生活，让他们不再操劳。

最后，特别感谢为我评审和答辩的老师们的悉心阅读和批评指正！