

网络游戏付费群体消费行为多维解构

——基于PEST、Pearson相关检验、GLM与K-Means、列联交叉模型

江宇欣, 方慧琳, 芮招彪*, 江 庭

巢湖学院数学与大数据学院, 安徽 合肥

收稿日期: 2026年8月17日; 录用日期: 2026年9月9日; 发布日期: 2026年9月18日

摘 要

在当今社会信息技术不断发展的情况下, 网络游戏上的不同年龄段群众也逐渐增多。数字文娱产业快速扩张, 网络游戏虚拟道具付费成为青年主流消费场景, 圈层攀比催生非理性高额氪金问题。为系统拆解年龄、性别、职业、可支配收入多重人口特征下的付费异质性, 本文构建“宏观环境-线性关联-回归识别-无监督分层-分类交互”多层实证分析框架。依托线上问卷获取有效样本542份, 融合PEST行业环境分析、Pearson相关检验、广义线性模型(GLM)、K-Means聚类、列联卡方五类统计工具开展量化研究。研究结果显示: 第一, 职业、月可支配收入与游戏总付费仅存在微弱不显著负相关, 游戏付费属于偏好驱动的非刚性消费, 收入约束效应极弱; 第二, 职业显著区分付费游戏品类, 职场人群更倾向休闲益智类轻度付费内容; 第三, 付费能力与年龄呈非线性正向关系, 30岁以上男性为重度付费核心群体; 第四, 性别存在稳定消费偏好分野, 男性偏向MOBA、射击竞技类游戏, 女性更青睐剧情养成、休闲解谜类内容。本文研究丰富了数字消费多元统计建模范式, 为游戏厂商精细化分层运营、监管机构开展青少年理性消费引导提供量化实证支撑。

关键词

网络游戏付费行为, PEST分析, Pearson相关检验, 广义线性模型, K-Means聚类, 列联卡方分析

Multi-Dimensional Deconstruction of Customer Behavior among Online Game Players

—Based on PEST, Pearson Correlation, GLM, K-Means and Contingency Cross Models

Yuxin Jiang, Huilin Fang, Zhaobiao Rui*, Ting Jiang

School of Mathematics and Big Data, Chaohu University, Hefei Anhui

*通讯作者。

文章引用: 江宇欣, 方慧琳, 芮招彪, 江庭. 网络游戏付费群体消费行为多维解构[J]. 统计学与应用, 2026, 15(9): 171-180. DOI: 10.12677/sa.2026.159216

Abstract

With the continuous advancement of information technology and the rapid expansion of the digital cultural and entertainment industry, online games have attracted players of all age groups. The digital entertainment industry is expanding rapidly, and purchasing virtual items in online games has become a mainstream consumption activity among young people. Peer comparison within gaming communities has led to the prevalence of irrational high-value in-game spending. To systematically dissect the heterogeneity of consumption behaviors shaped by demographic factors including age, gender, occupation and monthly disposable income, this paper constructs a multi-layer empirical analytical framework consisting of macro-environment analysis, linear correlation test, regression identification, unsupervised clustering and categorical interaction analysis. Based on 542 valid online questionnaire samples, this paper adopts five statistical methods for quantitative research: PEST analysis, Pearson correlation test, Generalized Linear Model (GLM), K-Means clustering and contingency chi-square test. The empirical results demonstrate four key conclusions: First, occupation and monthly disposable income only have weak and insignificant negative correlations with total in-game expenditure, indicating that game consumption is preference-driven non-essential spending barely restricted by income levels. Second, occupation significantly differentiates players' preferred game genres for payment; office workers are more inclined to make low-cost purchases on casual puzzle games. Third, there is a non-linear positive correlation between age and consumption capacity, and males over 30 form the core group of high-spending players. Fourth, obvious gender gaps exist in gaming preferences: males favor MOBA and shooting competitive games, while females prefer story-based simulation and casual puzzle games. This study expands the multivariate statistical modeling paradigm for digital consumption, and provides quantitative empirical support for game enterprises to carry out refined segmented operations as well as regulatory authorities to guide adolescents to form rational consumption concepts.

Keywords

Online Game Payment Behavior, PEST Analysis, Pearson Correlation Test, Generalized Linear Model (GLM), K-Means Clustering, Contingency Chi-Square Analysis

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

1.1. 研究背景

据《2024 年中国游戏产业报告》¹，国内网络游戏市场全年实际销售收入达 3257.83 亿元。Z 世代成为游戏核心用户，虚拟道具逐步演变为社交身份符号，圈层攀比心理催生大量超前、非理性付费行为，未成年人过度氪金、网络超前消费等议题持续受到监管部门与学界关注。外部产业环境同步发生结构性调整：国家持续收紧游戏版号发放、落地分级防充值政策；居民收入分层重塑用户付费预算；AI 个性化推送、云跨端技术持续放大精准付费转化。单一统计模型无法兼顾宏观行业约束与微观用户分层差异，现有研究缺少多模型融合一体化分析框架，因此本文融合多类统计方法，以此为切入点，多维解构网络

¹https://www.nppa.gov.cn/xxfb/ywdt/202412/t20241216_877445.html.

游戏付费人群消费特征。

1.2. 文献综述

现有网络游戏消费研究分为三大分支：

(1) 付费动机心理研究：Hsu & Lu (2004) 基于 TAM 模型证实社交沉浸、竞技成为线上游戏付费核心驱动[1]；国内马晨(2022)指出大学生易受同辈攀比诱导，大量资金投入竞技外观道具[2]。

(2) 用户分群统计研究：Kim (2019) 仅采用单一 K-Means 划分氪金玩家，缺少回归模型验证变量因果效应，无法解释偏好形成机制。

(3) 消费计量建模研究：国内现有文献多单独使用线性回归或简单交叉表，存在两处研究缺口[2] [3] [4]：未结合 PEST 宏观行业框架；未同步拆解年龄、性别、职业三重人口异质性，缺乏多模型交叉稳健验证。本文弥补上述空白，融合宏观环境分析、线性关联、广义回归、无监督聚类、列联交互五类方法，构建完整数字消费量化分析体系。

1.3. 研究目的与研究意义

1.3.1. 研究目的

本文的研究目的包括以下几个方面：(1) 借助 PEST 框架系统梳理网络游戏付费市场政策、经济、社会、技术外部约束，并据此推演出四个核心研究维度(4 Hypotheses)，与后续 Pearson 相关、GLM、K-Means、列联卡方四类统计模型形成一一对应的假设检验关系；(2) 量化检验职业、月可支配收入与游戏累计总付费的线性相关强度与显著性；(3) 识别不同职业人群对游戏付费品类的差异化影响效应；(4) 通过 K-Means 无监督聚类完成付费用户分层，刻画四类群体付费规模、人口画像；(5) 结合列联卡方检验拆解年龄、性别、职业与游戏偏好的交互关系，提出分层落地优化方案。

1.3.2. 理论意义

本文构建“宏观环境 - 线性关联 - 回归识别 - 无监督分层 - 分类交互”多层统计分析范式，拓展数字文娱消费实证方法体系，为同类线上虚拟商品消费研究提供可复制建模思路。

1.3.3. 实践意义

为游戏企业提供分年龄、性别、职业精细化运营策略；为教育、文旅监管部门提供青少年理性消费治理的数据支撑。

1.4. 研究创新点

框架创新：构建“PEST 研判→理论推演→实证假设→结果对话”全链条闭环范式，突破宏观分析与微观验证脱节的传统模式。

方法融合创新：将宏观 PEST 行业分析与四类微观统计模型结合，线性、非线性分析互补，突破单一模型局限。

视角创新：同步考察年龄、性别、职业三重人口异质性，区分付费金额、付费品类双重维度差异。

实证稳健创新：采用 Pearson、GLM、K-Means、列联卡方多模型交叉印证，统一实证结论，提升可信度。

1.5. 研究逻辑

全文分析链条：PEST 宏观行业研判→引出 4H 理论假设→问卷数据预处理与信效度检验→Pearson 线性相关性检验→广义线性模型识别品类影响因素→K-Means 用户付费分层→列联卡方分析人群偏好交互

→多模型综合验证→结果与 PEST 对话讨论→结论与分层对策。

2. 网络游戏行业 PEST 宏观环境分析

2.1. 政策环境

国家新闻出版署常态化收紧游戏版号发放,《未成年人网络保护条例》²严格限制未成年人单次充值上限与每日游戏时长;政策重点扶持国风、科普、教育向游戏,对高强度氪金、攀比诱导类游戏实施内容整改,从制度层面约束非理性付费行为。

2.2. 经济环境

本次有效样本 542 份中,18~25 岁学生占比 80.81%,月可支配资金集中在 1000~2000 元,月游戏付费以 1~200 元小额支出为主;职场人群收入更高,但衣食住行现实支出挤压游戏预算,市场付费呈现两极分化:少量中年群体大额氪金,多数青年仅小额投入。

2.3. 社会环境

Z 世代将游戏作为线下社交延伸载体,限定皮肤、稀有角色成为圈层身份标识;女性玩家规模持续扩张,恋爱养成、休闲解谜类付费潜力释放;心理成就感、角色美化、竞技战力提升为三大核心付费动机,同辈攀比是非理性消费关键诱因。

2.4. 技术环境

云游戏、多端互通拓宽用户覆盖边界;AI 用户画像算法实现道具、礼包精准推送,提升付费转化效率;大数据运营助力厂商分层营销,但算法诱导充值加剧青少年超前消费风险。

2.5. PEST 综合耦合研判

政策形成外部付费约束,收入水平决定付费预算上限,社交心理是付费内生驱动,大数据技术兼具产业赋能与消费诱导双重作用,四类外部因素共同塑造当前分层、差异化网络游戏付费市场。

3. 数据、问卷设计与研究假设

3.1. 数据来源与样本描述

2025 年 3~4 月采用腾讯问卷线上线下同步发放,累计投放 706 份,回收 548 份,剔除作答时长不足 30 秒、全部相同、关键变量缺失无效问卷 6 份,最终有效样本 542 份,样本回收率 76.77%。见表 1,描述样本人口分布三线表。

Table 1. Three line table of sample population distribution
表 1. 样本人口分布三线表

特征	分类	样本数量	占比(%)
年龄	18 岁以下	48	8.86
	18~25 岁	408	75.28
	25~30 岁	62	11.44
	30~60 岁	24	4.43

²https://www.gov.cn/zhengce/zhengceku/202310/content_6911289.htm

续表

性别	男	222	40.96
	女	320	59.04
职业	学生	438	80.81
	企事业单位	52	9.59
	教师/商业服务等	52	9.60
游戏参与	游玩	336	62.00
	不游玩	206	38.00

数据来源：本次调查问卷整理。

3.2. 问卷设计与信效度检验[3]

问卷共 11 道封闭客观题，分为人口特征、可支配收入、游戏参与、月度/累计付费、付费动机、偏好游戏类型六大模块。选取职业、月可支配资金、总付费三项核心题测算 Cronbach’s $\alpha = 0.627 > 0.6$ ，达到社科问卷可接受信度标准；题项逻辑统一，具备基础结构效度。数据编码修正：删除原总消费 5000 元以上 1.5 异常编码，全部付费区间统一整数标准化编码，无消费样本赋值 0，规避数值逻辑错误。

3.3. 变量定义

如表 2 所示，本研究构建多维度的变量体系进行分析。在自变量方面，选取了年龄、性别、职业及月可支配收入等人口统计学特征，这些分类变量将用于 Pearson 相关性分析、广义线性模型(GLM)及列联分析。在因变量设定上，分为两类：一是“游戏累计总付费”，作为标准化连续编码变量，建立 Pearson 相关；二是“偏好付费游戏类型”，作为无序多分类变量，通过广义线性模型和卡方检验来识别差异。此外，为了进一步研究，提取了标准化年龄、性别、月付费及总付费作为聚类特征，通过无量纲标准化处理，采用 K-Means 聚类算法进行多角度分析构建。

Table 2. Variable definition
表 2. 变量定义

变量类型	变量名称	变量说明	建模用途
自变量	年龄、性别、职业、月可支配收入	分类人口特征	Pearson、GLM、列联分析
因变量 1	游戏累计总付费	标准化连续编码	Pearson 相关
因变量 2	偏好付费游戏类型	无序多分类变量	广义线性模型、卡方检验
聚类特征	标准化年龄、性别、月付费、总付费	无量纲标准化	K-Means 聚类

3.4. 研究假设：基于 PEST-4H 理论推理(配套消费分层理论)

H1 (政策 - 职业假说)：职业身份显著区分用户的付费游戏品类选择(细分市场理论)。政策层面对未成年人实施充值限额与时长管控，使得学生群体与职场人群面临差异化的制度约束环境——学生更倾向于选择 MOBA、竞技等强社交对抗品类进行投入；而职场人群不被干预，且受碎片化休闲需求驱动，更偏好休闲益智类轻度付费内容。该假说由广义线性模型(GLM)承担主要验证，列联卡方提供佐证。

H2 (经济 - 收入假说)：不同于收入消费理论，职业、月可支配收入与游戏累计总付费之间不存在显著线性相关关系。经济环境维度显示，游戏虚拟道具属于偏好驱动的非刚性消费，低收入学生群体受社交攀比心理驱动仍有小额高频充值，高收入职场人群则因现实开支削减游戏预算，二者对冲削弱收入

约束效应。该假说由 Pearson 相关检验承担验证。

H3 (社会 - 性别假说): 年龄、性别与游戏付费品类之间存在显著交互关联(社会认同理论)。社会环境维度下, 传统性别角色社会化过程塑造了差异化偏好——男性在成长中更易被竞技对抗、射击类游戏吸引, 以获取成就感与社会认同; 女性则更倾向剧情养成、休闲解谜类内容, 以满足情感陪伴与审美表达需求。该假说由列联表卡方检验承担主要验证。

H4 (技术 - 年龄假说): 年龄、性别显著区分付费金额, 高年龄男性付费层级更高(人口分层消费理论)其中 30 岁以上男性为重度付费核心群体。技术环境维度下, AI 精准推送与跨端技术随用户年龄增长持续强化消费习惯, 形成“年龄增长→消费能力跃升→算法强化推荐→付费再升级”的正反馈循环; 而性别层面, 男性在竞技排行中的投入意愿亦高于女性。该假说由 K-Means 聚类完成用户分层验证。

4. 基于 Pearson 相关的付费线性关联检验

4.1. Pearson 模型原理

Pearson 相关系数用于衡量两组标准化变量线性关联强弱, 取值区间 $[-1, 1]$, 绝对值越大线性关系越强, 其计算公式为[4]:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}, \quad (1)$$

$p < 0.05$ 代表变量线性相关性具备统计学显著性。其中, X_i 和 Y_i 为样本值, $\bar{X}$ 和 $\bar{Y}$ 为均值。

4.2. 实证相关结果

- (1) 职业与游戏总付费: $r = -0.025$, $p = 0.746 > 0.05$, 微弱负相关并且不具备统计显著性;
- (2) 月可支配收入与游戏总付费: $r = -0.021$, $p = 0.785 > 0.05$, 微弱负相关且不具备统计显著性。

4.3. 经济含义解读

网络游戏虚拟道具不属于衣食住行刚性消费, 属于兴趣偏好驱动的可选消费, 收入水平对付费规模约束作用极弱; 低收入学生群体受圈层攀比小额充值, 高收入职场人群受现实开支挤压减少游戏投入, 因此呈现微弱负相关, 假设 H2 成立。

4.4. 模型局限性

Pearson 仅捕捉线性关系, 无法刻画年龄、性别带来的非线性付费分层差异, 因此进一步采用广义线性模型、K-Means 聚类开展深度分析。

5. 广义线性模型(GLM)付费品类影响识别

5.1. 模型选择依据

因变量“偏好付费游戏类型”为无序多分类离散变量, 普通线性回归正态、方差齐性假设无法满足; 广义线性模型通过链接函数适配离散型被解释变量。卡方预检验显示月可支配收入 $p = 0.944$ 无显著解释力, 仅保留职业为核心自变量纳入模型。

5.2. 整体 Omnibus 显著性检验

似然比卡方 = 31.602, $df = 16$, $p = 0.011 < 0.05$, 模型整体统计显著, 证明至少一类职业对游戏付费

品类存在显著影响。模型 $AIC = 436.21$, $BIC = 479.58$, 拟合效果良好。

5.3. 参数估计与经济解读

休闲益智类游戏系数 1.318, $p < 0.001$, 在 1%水平下显著正向; Rougelike 类系数 0.568, $p = 0.09$, 10%水平边缘显著。说明职场人群碎片化休闲需求更强, 选择休闲益智付费概率显著高于学生群体; 其余竞技类游戏参数均无统计显著性, 假设 H1 成立。

5.4. GLM 核心结论

职业是区分付费游戏赛道关键人口变量, 学生群体偏向 MOBA、重度竞技付费, 职场人群偏好碎片化轻度休闲游戏。

6. 基于 K-Means 聚类的付费用户分层异质性

6.1. 最优聚类 K 值确定

选取标准化年龄、性别、月度付费、累计总付费 4 项特征, 采用肘部法则计算不同 K 值组内平方误差 SSE, 曲线拐点对应 $K = 4$; 聚类轮廓系数 0.71, 组内相似度高、组间区分度良好, 聚类效果可靠。

6.2. 四类付费用户画像

如表 3 所示, 基于聚类分析结果, 将付费用户划分为四类群体(Cluster0 至 Cluster3), 并从年龄区间、男女比例、累计总付费、月均付费四个维度进行了特征刻画, 并赋予了相应的群体标签: 学生轻度付费群体(18~25 岁, 付费低)、职场中度付费群体(25~30 岁, 付费中等)、中年核心付费群体(30~40 岁, 男性为主, 付费较高)以及高收入重度付费群体(40~60 岁, 男性主导, 付费极高)。

Table 3. Four types of paid user profiles
表 3. 四类付费用户画像

聚类编号	年龄区间	男女占比	累计总付费	月均付费	群体标签
Cluster0	18~25 岁	男 60%, 女 40%	<2000 元	<200 元	学生轻度付费群体
Cluster1	25~30 岁	男 70%, 女 30%	2000~5000 元	200~500 元	职场中度付费群体
Cluster2	30~40 岁	男 85%, 女 15%	5000~10,000 元	500~10,000 元	中年核心付费群体
Cluster3	40~60 岁	男 90%, 女 10%	>10,000 元	>1000 元	高收入重度付费群体

6.3. 聚类规律总结

- (1) 付费能力与年龄呈非线性正向关系, 30 岁之后付费金额明显提升;
- (2) 所有高付费聚类男性占比超 85%, 女性集中于低付费学生群体; 假设 H4 成立, 年龄、性别形成显著付费金额分层异质性。

7. 列联表与卡方检验：人口特征与游戏偏好交互

7.1. 列联分析原理

构建二维列联表, 卡方检验原假设为变量相互独立, $p < 0.05$ 则拒绝原假设, 变量存在显著交互关联。

7.2. 分维度检验结果

年龄: 18 岁以下男性 65%偏好 MOBA; 18~25 岁为付费主力; 25 岁以上样本转向轻度休闲。

性别维度卡方检验 $\chi^2 = 48.37$, $p < 0.001$, 性别偏好差异高度显著: 男性集中于 MOBA、射击竞技; 女性偏好女性向养成、休闲益智。

职业维度: 学生群体 MOBA 付费占比最高, 职场人群休闲付费占比更高, 与 GLM 模型结论高度吻合; 假设 H3 成立, 年龄、性别、职业共同交互影响付费游戏品类选择。

8. 多模型交叉稳健验证

线性层面: Pearson 证明收入无法单独解释付费高低, 聚类模型补充年龄、性别带来的非线性分层特征。

偏好识别层面: GLM 模型识别职业整体影响, 列联卡方进一步细化不同年龄、性别下的细分偏好差异。

四类模型实证结论相互支撑、彼此印证, 形成完整用户分层消费画像, 实证结果具备稳健性。

9. 实证发现与 PEST 宏观环境的对话讨论

9.1. 政策规制与职业分流(对话 H1)

GLM 显示, 职场人群休闲益智类付费概率显著高于学生, 而学生 MOBA 类付费占比最高。这一差异折射出政策层面的制度性筛选: 防沉迷机制对未成年人实施管控, 学生在受限窗口内将有限资源集中投入社交回报率最高的竞技品类以维系圈层认同; 职场人群不受此约束, 工作节奏反而驱动其偏好碎片化、低负荷的休闲内容。政策规制并非均质作用于所有群体, 而是以职业身份为中介, 间接引导了不同赛道付费流量的分流。

9.2. 收入约束失效的经济解释(对话 H2)

Pearson 显示月可支配收入与游戏总付费仅微弱负相关, H2 成立。这一反直觉发现印证了 PEST 经济维度的核心判断: 游戏虚拟道具属韦伯意义上的“地位性消费”——其效用源于社交圈层中的展示与比较功能。低收入学生受攀比驱动进行高频小额充值, 高收入职场人群则因现实开支而抵消消费冗余。两股力量对冲, 致使“收入 - 消费”线性关系整体失效。K-Means 进一步揭示, 30 岁后付费金额跃升, 说明生命周期阶段比当期收入更具分层解释力。

9.3. 性别偏好的社会再生产(对话 H3)

列联卡方显示性别与品类偏好高度显著, 男性集中于 MOBA 与射击类, 女性偏好养成与休闲解谜类。这并非天然差异, 而是性别角色期待的数字化投射——男性在传统社会化中被鼓励展现竞争性与支配性, MOBA 等竞技游戏提供展现场域; 女性则被引导发展共情与审美表达, 养成装扮类游戏精准对应这些偏好。聚类结果中, 女性在 30 岁以上高付费群体中占比不足 15%, 表明随年龄增长, 婚育、职场等现实挤压可能迫使女性提前退出高投入游戏场景, 而男性竞技消费惯习则持续被社会认同机制强化至中年。

9.4. 技术赋能的年龄累积效应(对话 H4)

K-Means 显示, Cluster2 (30~40 岁)与 Cluster3 (40~60 岁)男性占比分别达 85%和 90%, 累计付费超 5000 元乃至万元, 而 18~25 岁学生月均不足 200 元。这一陡峭的年龄 - 付费梯度, 不能简单归因于“年龄增长、收入增加”——更深层机制在于技术维度: AI 精准推送与云跨端技术构建了消费习惯的复利效应。30 岁以上核心用户经多年游戏参与和算法持续喂养, 偏好已被高度固化, 且跨端互通降低了多设备

持续投入的摩擦成本，形成“付费→精准推荐→更高额付费”的正反馈循环。技术环境由此成为年龄与付费非线性关系的放大器，而非单纯的工具性中介。

9.5. 多模型交叉验证的宏观一致性判断

四类统计模型从不同维度切入，其结论在 PEST 宏观框架下呈现高度内部一致性：政策对职业分流的制度筛选、经济对线性关联的整体对冲、社会对性别偏好的再生产塑造、技术对年龄梯度的非线性放大——四个方面的宏观力量分别经由职业身份、生命周期阶段、性别角色期待、算法固化四条微观传导路径，共同构建了当前网络游戏付费市场“学生小额高频 × 职场轻度休闲 × 中年男性重度氪金 × 女性偏好分化”的多元分层格局。这一宏观与微观的对话式验证，既增强了实证结论的解释力与外部效度，也为后续分层对策提供了精准的干预靶点。

10. 核心结论与分层优化对策

10.1. 综合实证结论

宏观 PEST 层面：政策、经济、社会、技术四重外部因素共同塑造分层付费市场，社交攀比是非理性消费核心诱因。

线性关联：职业、月可支配收入与游戏总付费无显著线性关系，付费由个人兴趣偏好主导。

回归识别：职业显著区分付费赛道，学生偏爱竞技，职场偏爱休闲。

聚类分层：年龄越高付费能力越强，男性垄断重度氪金，八成用户为小额学生付费群体。

交互特征：性别形成稳定游戏偏好分野，男性偏向竞技，女性偏向养成休闲。

10.2. 分层落地对策

(1) 政府与社会监管层面

严格落实未成年人充值限额与防沉迷机制，中小学、社区联合开展理性消费科普；扶持国风、轻量教育类游戏，削减高强度氪金诱导设计。

(2) 游戏企业精细化运营

18~25 岁学生：推出低价小额竞技皮肤、二次元礼包，优化 MOBA 竞技内容；25~30 岁职场人群：碎片化休闲月卡、轻度道具，减少高强度充值活动；30 岁以上男性：典藏外观、长线会员养成体系；女性玩家：丰富剧情养成、休闲装扮类付费内容。

(3) 个体消费引导

玩家自主设置月度游戏消费上限，区分刚需娱乐与溢价虚拟道具，杜绝圈层攀比式超前充值。

11. 研究局限与未来展望

11.1. 研究局限

(1) 样本代表性与抽样偏差。本研究采用线上便利抽样，样本中 18~25 岁学生占比达 80.81%，而 30 岁以上群体仅占 4.43%，与第 53 次《中国互联网络发展状况统计报告》³中 30 岁以上网民占比约 45% 的结构存在显著偏差。学生群体过度代表、中老年在职样本严重不足，可能高估学生消费特征对整体结论的权重，削弱外部效度。未来应依据网民年龄与职业分布实施分层配额抽样，使各子群体样本量与总体比例匹配，以提升结论的泛化能力。

(2) 自我报告数据的偏差风险。问卷数据完全依赖受访者主观回忆，可能引入社会期望偏差(倾向美

³https://www.cac.gov.cn/2024-03/25/c_1713038218396702.htm

化或低估消费行为)与回忆偏差(对累计付费估算失准)。加之问卷未设置测谎题或交叉验证项,难以甄别随意作答或不实回答。后续可在问卷中嵌入重复题项、反向计分题或逻辑校验题(如“上月付费是否高于累计付费”等矛盾检测),以提升数据可靠性,并应在论文中明确讨论各类偏差的潜在影响方向。

(3) 横截面数据的因果解释局限。本文采用横截面设计,所有变量在同一时点采集,无法追踪用户付费行为随生命历程的动态变化(如从学生向职场过渡中的品类转换),也难以建立严格因果推断。变量层面,未纳入日均游戏时长、充值频次、活跃度等行为控制变量;模型层面,仅使用传统统计方法,未引入XGBoost等集成学习算法进行预测精度对比。

11.2. 未来研究方向

基于本文的研究,未来可以进一步研究的方向包括:(1) 调研拓展:开展季度追踪面板问卷,追踪同一批用户的付费行为动态变化,并严格按照网民人口结构实施分层抽样以增强代表,扩充中老年样本;(2) 模型升级:融合集成机器学习、玩家评论文本情感挖掘,丰富用户心理维度的解释路径;(3) 场景拓展:区分手游、端游、主机付费行为差异开展对比研究;(4) 数据质量管控:在问卷中嵌入测谎题与交叉验证题,同时结合用户后台消费日志数据对自我报告进行校准,降低主观偏差。

基金项目

本研究得到了巢湖学院大学生创新创业训练计划项目(编号:X202510380030)、学科质量提升工程项目(编号:XLZ202409)和科研启动项目(编号:KYQD-2023064)的资助。

参考文献

- [1] Hsu, C. and Lu, H. (2004) Why Do People Play On-Line Games? An Extended TAM with Social Influences and Flow Experience. *Information & Management*, **41**, 853-868. <https://doi.org/10.1016/j.im.2003.08.014>
- [2] 马晨. MMOG 时代下大学生网络游戏喜好的新特征、成因及应对措施[J]. 普洱学院学报, 2022, 38(5):78-83.
- [3] 张虎, 田茂峰. 信度分析在调查问卷设计中的应用[J]. 统计与决策, 2007(21): 25-27.
- [4] 贾俊平, 何晓群, 金勇. 统计学[M]. 第4版. 北京: 中国人民大学出版社, 2009.