

大学生婚育意愿调查

周瑞民 于越 詹祺元 林奕文

January 6, 2023

摘要

随着社会变迁和代际更替，中国的婚育模式已发生显著的变化，年轻人的婚育观念显现出一系列的新变化与新特征。大学生作为青年群体的中坚力量，是未来十年内婚育行为的核心主体，因此有必要研究大学生的婚育意愿和观点及其可能的形成原因。

本文根据问卷调查获得的一手数据，分析了大学生婚育意愿的现状，并按照性别分层抽样，通过对比分析和聚类、回归等方式研究该现状的差异性及影响机制。对于两性差异，本文通过卡方检验等手段分析了恋爱状态、结婚意愿、促进/抑制结婚的因素、生育观念、提高生育的政策、受外界影响程度等婚育观的异同。对于互联网参与的影响，本文定性和定量地讨论了亚文化圈层、APP 使用等因素与婚育意愿的相关性，并通过对比和统计检验等形式探讨了部分可能的原因。

目录

1	绪论	6
1.1	研究背景与意义	6
1.1.1	研究背景	6
1.1.2	研究意义	6
1.2	研究内容与框架	6
1.3	研究方法	7
1.3.1	文献研究法	7
1.3.2	词云图	8
1.3.3	问卷调查法	8
1.3.4	实证分析法	8
1.3.5	信效度分析法	8
1.4	文献综述	9
2	基于线上数据的网络调查分析	9
2.1	全国人口结果与婚育现状	9
2.2	基于互联网的言论调查分析	11
3	调查方案设计	12
3.1	调查对象与调查方法	12
3.2	调查方案与问卷设计	13
3.2.1	调查方案	13
3.2.2	问卷设计	13
4	问卷调查情况	15
4.1	问卷有效性	15
4.1.1	样本数据信度检验	15
4.1.2	样本数据效度检验	15
4.2	受访者基本信息	16
5	大学生婚育意愿的现状与分析	18
5.1	婚姻意愿及观点现状	18
5.1.1	恋爱状态及次数	19
5.1.2	结婚意愿	19
5.1.3	理想结婚年龄	20
5.1.4	促进结婚的因素	20
5.1.5	抑制结婚的因素	21
5.1.6	对婚姻行为的认同度	21

5.2	生育意愿与观点现状	21
5.2.1	理想及意愿子女数	22
5.2.2	生育带来的价值	23
5.2.3	抑制生育的因素	23
5.2.4	提高生育意愿的政策	24
6	婚育意愿的性别差异	25
6.1	研究方法介绍	25
6.2	结婚意愿	26
6.3	恋爱状态	26
6.4	促进/抑制结婚的因素	27
6.5	理想/意愿子女数	28
6.6	提高/抑制生育的政策	29
6.7	婚育观念受何外界因素影响	30
7	互联网参与与婚育意愿	30
7.1	总体方法框架介绍	30
7.2	数据预处理	32
7.3	聚类分析	33
7.4	回归分析	38
7.4.1	数据预处理	39
7.4.2	分层抽样说明	39
7.4.3	圈层影响	40
7.4.4	APP 使用影响	42
7.4.5	其他外部因素	44
7.5	原因探究	45
7.5.1	数据预处理	45
7.5.2	统计检验设计	45
7.5.3	检验结果和解读	46
8	报告总结	47

插图索引

1	研究框架	7
2	中国、美国、日本的生育率	10
3	中国、美国、日本的人口增长率	10
4	中国人口结构数据	10
5	总获赞数排名前五的提问内容	12
6	提问内容的词云图	12
7	调查问卷的结构设计	14
8	受访者性别分布	16
9	受访者年龄分布	17
10	受访者人数大于 10 的省份	17
11	受访者户口类型分布	17
12	受访者受教育程度分布	17
13	受访者独生情况分布	18
14	受访者月消费额度分布	18
15	大学生恋爱状态	19
16	大学生恋爱次数	19
17	结婚意愿	19
18	理想结婚年龄	20
19	促进结婚的因素	20
20	抑制结婚的因素	21
21	对婚姻行为的认同度	22
22	理想及意愿子女数	22
23	理想儿子及女儿数	22
24	生育带来的价值	23
25	抑制生育的前六个因素	24
26	促进生育的前六个因素	24
27	结婚意愿的性别差异	26
28	恋爱状态的性别差异	27
29	促进结婚因素的性别差异	27
30	抑制结婚因素的性别差异	28
31	理想子女数的性别差异	28
32	意愿子女数的性别差异	28
33	提高生育的政策	29
34	阻碍生育的因素	29
35	婚育观受外界因素的影响	30

36	总体分析框架图	31
37	$K = 4$ 与 $K = 5$ 时各聚类样本量占比	34
38	$K = 4$ 时不同聚类的男女比例饼图	35
39	$K = 4$ 时不同聚类的婚育观念柱状图	37
40	上网时长深入分析——对结婚观的抑制作用	38
41	上网时长深入分析——对生育观的抑制作用	38
42	动漫圈层深入分析——对结婚观的抑制作用	41
43	动漫圈层深入分析——对生育观的抑制作用	41
44	小红书 APP 深入分析——对结婚观的影响	43
45	小红书 APP 深入分析——对生育观的影响	43

表格索引

1	2016-2020 结婚及对应出生数据	11
2	调查方案	13
3	信度检验结果	15
4	效度检验结果	16
5	性别差异卡方检验结果	26
6	$K = 4$ 情形下不同聚类的各特征贡献表	34
7	$K = 5$ 情形下不同聚类的各特征贡献表	34
8	圈层参与度对婚育意愿的回归贡献度	40
9	APP 对婚育意愿的回归贡献度	42
10	外部影响对婚育意愿的回归贡献度	44

1 绪论

1.1 研究背景与意义

1.1.1 研究背景

生育观作为个体观念在生育问题上的集中体现，根植于社会的文化结构中，深受时代变化的影响。随着社会的变迁和代际的更替，中国的婚育模式已发生显著的变化，特别是在年轻人群体中，其生育观念显现出一系列的新变化与新特征。大学生作为青年群体的中坚力量，是未来十年内婚育行为的核心主体。第七次全国人口普查数据显示，中国 18-22 岁的青年群体中，受教育程度在大学专科及以上者占比已达 52.20%。同时，作为成长于互联网技术高速发展时期的“Z 世代”，他们的婚育观念也有其独特性。大学生的婚育观念不仅影响其个体的婚姻、家庭选择以及人生发展路径，更将深远地影响到中国的生育水平、人口结构、代际养老支持体系等。通过与身边朋友交流、浏览互联网社区等途径，我们注意到大量年轻人或玩笑或认真地表达过独身生活的愿望，以及自己产生此种想法的原因、在此过程中与传统价值观的碰撞等。

今年 4 月，中国人民大学人口与发展研究中心人口与发展研究中心公布了《中国大学生婚育观报告》，基于分层多阶段 PPS 抽样组织实施了有全国代表性的 2021 年中国大学生婚育观调查^[1]。调查归纳了当前大学生婚育观的五大特征：对大学生而言，婚姻和生育仍然没有明显的去制度化，但其价值基础已发生改变；巨大的婚育阻力而非观念转变是大学生低婚育意愿的主要原因；婚育观呈现出显著的性别和社会经济地位分化；互联网与圈层文化和婚恋观之间形成互构。我们对其中涉及生育观念的部分结论产生浓厚的兴趣，希望进一步探究其变化轨迹、中介机制和背后的家庭及社会因素等。

1.1.2 研究意义

我国历经了快速的社会变革，在公私领域分离、市场经济转型不断深入的同时，面临着后疫情时代逆全球化、人口老龄化和少子化的挑战。我国老龄人口总量大，老龄化速度在加快，且面临“未富先老”的难题——富裕程度还未赶上世界平均水平，社会保障及养老体系远未健全。面对尚不确切的人口走势，了解青年婚育观念至关重要。一方面，年轻人的生育观是预测未来生育水平变动的基础；另一方面，对婚育观的探索也是理解行为转变背后机制的重要途径。大学生作为青年群体的中坚力量，是未来十年内婚育行为的核心主体，其婚育态度引发学界乃至社会的广泛关注。大学生的婚育观念不仅影响其个体的婚姻、家庭选择以及人生发展路径，更将深远地影响到中国的生育水平、人口结构、代际养老支持体系等。

1.2 研究内容与框架

本文首先分析全国总体婚育情况以及年轻人在互联网上发表的关于婚育相关的言论。通过分析这些内容，并结合生活经验，设计了全面丰富的调查问卷。根据调查数据，

我们首先阐述了大学生婚育意愿的现状，为了进一步了解产生该现状的原因，我们从三个角度进行了分析。第一，以性别为基准，对所有的数据进行横向分析，研究不同性别之间的婚育观差异，得出性别所导向的婚育观结论；第二，以网络媒体、APP 为基准，研究不同互联网圈层对于婚育观的影响；第三，对于已经采集到的问卷对象基于其基本信息、互联网圈层等进行聚类分析，将调查问卷对象的特征集成因子，基于这些因子对数据进行纵向分析，以挖掘相关的婚育观规律。最后希望得到多方面的影响当代青年人的婚育观的因素，并提出有效建议。

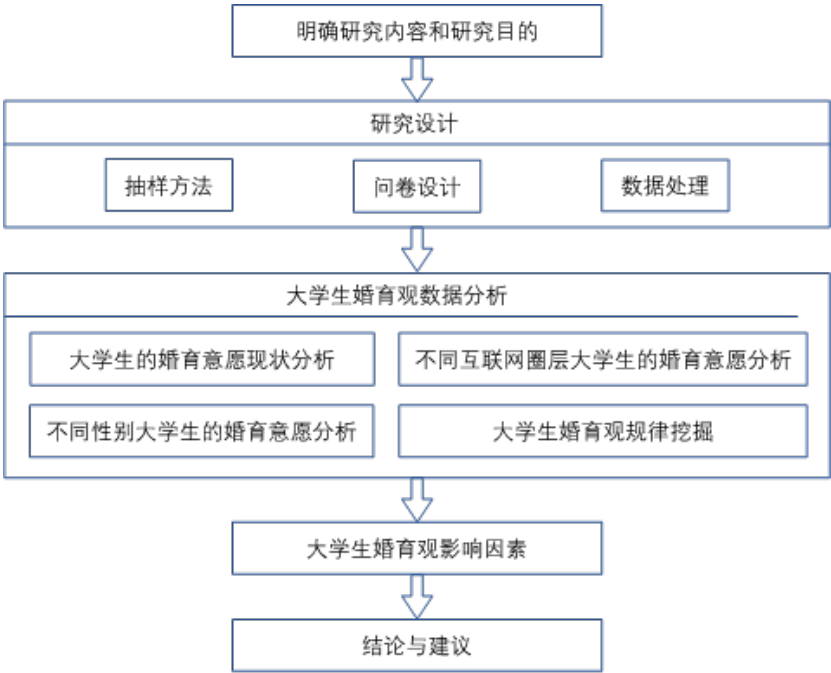


图 1: 研究框架

1.3 研究方法

本文运用了文献研究法、词云图、问卷调查法、实证分析法以及信效度分析法等方法对大学生的婚育意愿以及互联网圈层等方面展开研究。

1.3.1 文献研究法

文献研究法是根据一定的研究目的或课题，通过调查文献来获取资料，从而全面地、正确地了解掌握所要研究问题的一种方法。全面的文献分析是研究的基础，因而文献研究法被广泛运用于各种学科。通过对婚育意愿问题的学术论文和研究期刊进行阅读分析，了解了人口学、婚姻去制度化、人口老龄化的现有理论，理清了与研究主题有关的观点，找到了本研究的研究方向，为研究模型的提出、问卷逻辑设计奠定了基础。本文从性别和互联网圈层两个主要视角进行切入，研究大学生的婚育观情况和影响因素。

1.3.2 词云图

词云图，由词汇组成图形，也称“文字云”，是基于网络文本的爬取，对于出现频率较高的“关键词”予以视觉的突出，图形是最直观的呈现形式，词云图更是帮助使用者从大量的文本信息过滤出最值得重视的信息。本研究的词云图分析体现在互联网平台上年轻人分享有关婚育意愿的信息研究，利用该方法直接获取年轻人在观看及讨论婚育相关话题时的内容，以了解当代年轻人对于婚育的观点。

1.3.3 问卷调查法

问卷调查法是调查研究中采用最频繁的一种基本方法，它是研究者用统一严格设计的问卷，通过书面语言与被调查者进行交流，来搜集研究对象关于感兴趣问题或现象的信息和资料的方法。了解问卷调查法的特点、种类，掌握问卷编制的技术和实施问卷调查的基本要求是相当重要的。问卷调查法的基本工具是统一严格设计的问卷，调查过程是调查者向被调查者发放问卷，由被调查者书面回答来完成的。本文的问卷经过小组讨论和邀请同学试填问卷，删除或修改不适宜问项、排列问项次序、确定问项的表达方式以及初级量表的设置。通过对收集上来的数据进行简单处理，使之能够用于描述性统计分析、对比分析和聚类分析中。

1.3.4 实证分析法

实证研究法是科学实践研究的一种特殊形式。其依据现有的科学理论和实践的需要提出设计，本文是在问卷调查的条件下，通过有目的调查和分析，根据相关统计软件对收集到的数据进行统计分析，验证研究模型。在清洗问卷数据后，得到样本的基本数据资料，描述大学生的婚育意愿现状。进一步使用 SPSS 软件完成对量表的信度和效度的验证性分析。最后使用 python，对婚育意愿、性别、互联网圈层等维度进行对比分析和聚类分析。

1.3.5 信效度分析法

信度是衡量问卷结果稳定性与可靠性的重要指标，反映了问卷调查过程中对误差控制能力的大小。问卷的信度水平可用信度系数来衡量，信度系数越高即表示该测验的结果越一致、稳定与可靠。本研究在检验指标的信度时，主要采用的是 Cronbach's Alpha 系数，通过计算指标的 Cronbach's Alpha 值来判断指标信度是否通过检验。效度就是正确性程度，即样本在多大程度上反映了所要考查内容的真实情况。效度越高，即表示样本越可信。效度主要包括内容效度和结构效度，内容效度和结构效度可分别从定性和定量两方面分析。

1.4 文献综述

关于婚育观的概念，目前学界多从其内涵与外延进行讨论。狭义而言，婚育观是指在一定社会环境下，人们对于恋爱、两性关系、婚姻家庭及生育问题的基本看法与根本态度，受个体的人生观和价值观的影响^[2]。广义而言，在前述的基础上进一步囊括受访者的婚育意愿，即个体在适婚育龄期对婚育决策的具体意向，包括结婚意愿、理想子女数、意愿子女数与子女性别偏好等。本调研采用婚育观的广义概念。在婚姻与生育领域，已有不少经典理论用以刻画婚育观念和行为的变动趋势及其驱动机制，而第二次人口转变理论是解释当今世界逐渐扩散开来的低婚育意愿的主流框架。该理论的核心观点是：个体主义的兴起与观念的变革，使得婚姻对个体的吸引力下降，生育与婚姻逐步解绑，动摇了婚育的制度性基础^[3]。在该理论的预期下，青年的结婚意愿和结婚率都会下降，而同居率、离婚率与非婚生育率则会上升。在情感价值与自我认同成为个体的核心需求下，结婚和生育也只是生活方式的众多选项之一。中国很早就开展了关于青年婚育态度的调查，综合不同的研究有以下一些初步结论：中国青年择偶标准日趋多元化^[4]；理想结婚年龄集中在 26-30 岁；观念更开放^[5]，但仍重视婚恋的忠诚度^[6]；多数青年的意愿子女数小于两个孩子，且开始偏好女儿^[7]；住房压力近年来逐渐成为压抑青年生育意愿的主要动因^[8]，但抑制生育意愿的因素存在一定性别差异^[9]。

2 基于线上数据的网络调查分析

2.1 全国人口结果与婚育现状

根据世界银行的人口数据，中国的生育率从 1970 年代之前的 6 左右，降至 1990 年的 2 左右，再降至 2010 年后的 1.5 左右，之后虽然略有回升，但 2020 年我国生育率仅 1.70，低于维持人口总量不变的世代更替水平 2.1。在 1990 年之后，中国的生育率甚至低于美国，接近于日本的生育率。从人口增长率来看，2021 年我国人口增长率仅为 0.09，人口负增长将成为未来一段时间内的重要问题。问卷调查法是调查研究中采用最频繁的一种基本方法，它是研究者用统一严格设计的问卷，通过书面语言与被调查者进行交流，来搜集研究对象关于感兴趣问题或现象的信息和资料的方法。了解问卷调查法的特点、种类，掌握问卷编制的技术和实施问卷调查的基本要求是相当重要的。问卷调查法的基本工具是统一严格设计的问卷，调查过程是调查者向被调查者发放问卷，由被调查者书面回答来完成的。本文的问卷经过小组讨论和邀请同学试填问卷，删除或修改不适宜问项、排列问项次序、确定问项的表达方式以及初级量表的设置。通过对收集上来的数据进行简单处理，使之能够用于描述性统计分析、对比分析和聚类分析中。

根据第七次全国人口普查数据，全国人口中，0—14 岁人口为 25338 万人，占 17.95%；15—59 岁人口为 89438 万人，占 63.35%；60 岁及以上人口为 26402 万人，占 18.70%，其中 65 岁及以上人口为 19064 万人，占 13.50%。与 2010 年第六次全国人口普查相比，0—14 岁人口的比重上升 1.35 个百分点，15—59 岁人口的比重下降 6.79 个百分点，60

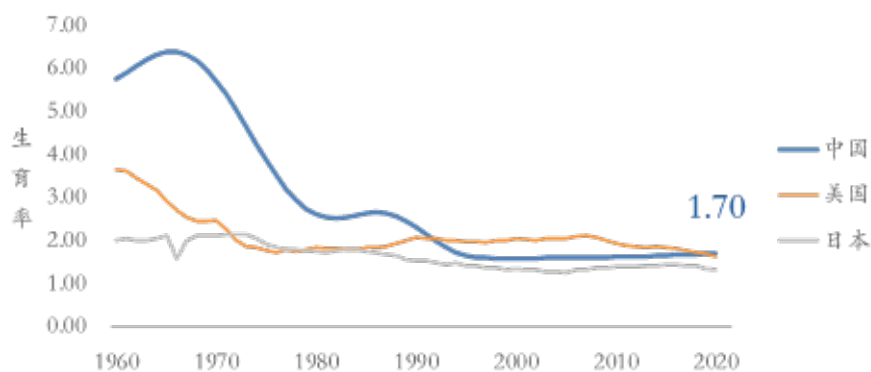


图 2: 中国、美国、日本的生育率

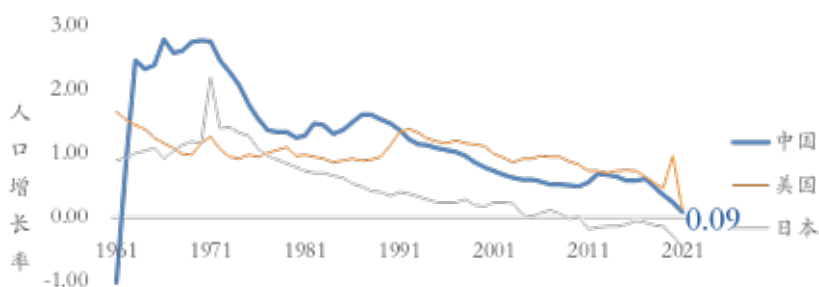


图 3: 中国、美国、日本的人口增长率

岁及以上人口的比重上升 5.44 个百分点，65 岁及以上人口的比重上升 4.63 个百分点。据联合国测算，我国人口将在 2030 年达到 14.6 亿的峰值后趋于下降，2050 年降为 14 亿，预计老年人口占 25.6%，劳动人口占 59.8%，少儿人口占 14.1%，人口结构将从长方形变成倒金字塔形。可以看出，我国少儿人口比重回升，生育政策调整取得了积极成效。但与此同时，人口老龄化程度进一步加深，未来一段时期将持续面临人口长期均衡发展的压力。

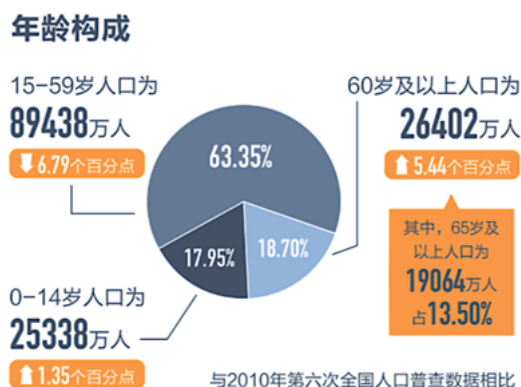


图 4: 中国人口结构数据

随着经济社会的持续发展，尤其是工业化、现代化带来的人口生育观念转变等方面的影响，低生育已经成为大多数发达国家普遍面临的问题，也将成为我国面临的现实问

题。根据民政局公布数据分析，2020 年全国结婚登记数相比 2016 年数据减少 329 万对，为 1986 年以来最低。而与出生人口更密切相关的初婚人数，从 2016 年的 1913 万人持续下降至 2020 年 1229 万人，下降幅度达 35.8%。年轻人数量下降，适婚人群总量减少，是结婚人数和结婚率下降的一个关键因素。根据国家统计局数据，我国 80 后、90 后、00 后人口分别为 2.23 亿人、2.1 亿和 1.63 亿，整体呈不断下降趋势。近 8 年 20 岁-29 岁对应出生人口总体下降 11%。

从表1中可以看出中国 2016-2020 年结婚及对应出生数据随着经济社会的持续发展，尤其是工业化、现代化带来的人口生育观念转变等方面的影响，低生育已经成为大多数发达国家普遍面临的问题，也将成为我国面临的现实问题。根据民政局公布数据分析，2020 年全国结婚登记数相比 2016 年数据减少 329 万对，为 1986 年以来最低。而与出生人口更密切相关的初婚人数，从 2016 年的 1913 万人持续下降至 2020 年 1229 万人，下降幅度达 35.8%。年轻人数量下降，适婚人群总量减少，是结婚人数和结婚率下降的一个关键因素。根据国家统计局数据，我国 80 后、90 后、00 后人口分别为 2.23 亿人、2.1 亿和 1.63 亿，整体呈不断下降趋势。近 8 年 20 岁-29 岁对应出生人口总体下降 11%。

表 1: 2016-2020 结婚及对应出生数据

年份	结婚登记（万对）	初婚人数（万人）	20-29 岁对应年份 出生人数（万人）	对应出生年份
2016	1143	1913	22569	1987-1996
2017	1063	1746	22069	1988-1997
2018	1014	1599	21547	1989-1998
2019	927	1399	20967	1990-1999
2020	814	1229	20347	1991-2000
近五年下降总比例	-28.8%	-35.8%	-9.8%	

2.2 基于互联网的言论调查分析

言论调查部分，本报告收集知乎 # 生育意愿话题下获赞数大于 1000 的回答及其对应的提问，共 235 条数据，用于对当代年轻人的整体生育意愿进行基本的把握，并初步探索影响生育意愿的因素。根据总获赞数进行排序，排名前五的提问内容、总获赞数以及对应的回答数量如下图5：

知乎用户观察到了当前生育意愿下降的问题，并对当前的生育政策和女性生育歧视的问题表达了不满。同时也存在一些用户，虽然目前不愿意生育，但担心无人养老的问题。

我们在剔除停用词后，选取词频前 15 名，绘制词云图如下图6：

可以发现知乎用户关于生育意愿的提问主要集中在女性上，并且会关心最佳生育年龄、生育政策等内容，但整体上表达了不愿意进行生育的观点。因此我们将生育政策、女性权益以及“养儿防老”等观点考虑进我们的问卷中。



图 5: 总获赞数排名前五的提问内容



图 6: 提问内容的词云图

3 调查方案设计

3.1 调查对象与调查方法

本文的调查对象即为全国大学生，其中包括专科生，本科生，硕士生和博士生。因此，《大学生婚育意愿调查》的调查对象纳入标准为：（1）中国某大学在校学生（2）自愿接受调查。本文采用网络问卷的方式进行调查。为了获得更多数据以进行分析，本文并未设置除在校学生以外的限制。

3.2 调查方案与问卷设计

3.2.1 调查方案

调查方案基于时间线进行设计，本次调查研究以调查问卷为主要研究工具，问卷是基于文献研究和前期搜集数据的基础上自行编制而成，具体调研安排如下表2所示。

表 2: 调查方案

阶段	时间	任务
第一阶段	2022 年 9 月 15 日至 2022 年 9 月 29 日	讨论、确定题目、查阅相关资料
第二阶段	2022 年 9 月 29 日至 2022 年 10 月 20 日	确定调查大纲、爬虫
第三阶段	2022 年 10 月 20 日至 2022 年 11 月 3 日	问卷设计、预发放、修改
第四阶段	2022 年 11 月 3 日至 2022 年 12 月 8 日	问卷发放、数据分析与处理
第五阶段	2022 年 12 月 8 日至 2022 年 1 月 6 日	报告撰写

3.2.2 问卷设计

1. 问卷类型

调查问卷以电子问卷的形式，让受访者进行自填问卷为主。调查小组将调查问卷以二维码、网页链接等电子问卷发放方式发放给受访者，受访者填写完毕后数据自动存储在问卷星中，便于分析整理。

2. 结构设计

调查问卷的结构设计分为六个部分：第一部分为受访者的婚姻意愿，包括受访者的恋爱状态、恋爱次数、结婚意愿，题型为单选题。第二部分为受访者的婚姻观点，包括受访者的理想结婚年龄、认为促进结婚的因素、认为抑制结婚的因素、对于婚前同居、婚前性行为、不婚、丁克、同性恋、闪婚、非婚非育、堕胎、开放性关系的婚姻行为的认同程度，题型为单选题、多选题与矩阵量表题。第三部分为受访者的生育意愿，包括意愿子女数、理想子女数、理想儿子数、理想女儿数，题型为矩阵量表题。第四部分为受访者的生育观点，包括认为生育带来的价值、认为阻碍生育的因素、认为可以提高生育意愿的政策，题型为排序题和多选题。第五部分为受访者的网络社交情况，包括受访者的休闲上网时间、对网络话题的参与程度、对 APP 的使用频率、是否分享婚育观点、是否在网上秀恩爱、未来是否会晒娃、认为影响婚育观念的因素，题型为单选题和矩阵量表题。第六部分为受访者的基本信息，包括受访者的性别、年龄、地区、户口类型、职业、最高学历、是否独生子女、消费水平，题型为单选题和填空题。问卷结构如下图7所示：

3. 内容设计

主题划分	调查主题	题型	调查内容
第一部分	婚姻意愿	单选题	受访者的恋爱状态、恋爱次数、结婚意愿
第二部分	婚姻观点	单选题、多选题与矩阵量表题	受访者的理想结婚年龄、认为促进结婚的因素、认为抑制结婚的因素、对于婚前同居、婚前性行为、不婚、丁克、同性恋、闪婚、非婚非育、堕胎、开放性关系的婚姻行为的认同程度
第三部分	生育意愿	矩阵量表题	受访者的意愿子女数、理想子女数、理想儿子数、理想女儿数
第四部分	生育观点	排序题与多选题	受访者认为生育带来的价值、认为阻碍生育的因素、认为可以提高生育意愿的政策
第五部分	网络社交情况	单选题与矩阵量表题	受访者的休闲上网时间、对网络话题的参与程度、对 APP 的使用频率、是否分享婚育观点、是否在网上秀恩爱、未来是否会晒娃、认为影响婚育观念的因素
第六部分	基本信息	单项选择题与填空题	受访者的性别、年龄、地区、户口类型、职业、最高学历、是否独生子女、消费水平

图 7: 调查问卷的结构设计

基本信息的作用主要在于为所有被调对象基于背景而分类。在大学生中，我们主要基于性别、户口类型、最高学历进行分类，希望基于问卷返回的结果，得出最终大学生们的不同婚育观与学生背景之间的关系，探索二者之间存在的相关性。这一部分的另一作用在于对调查对象分层。这样我们就可以基于不同的层次，利用统计调查的相关理论给出相对科学而准确的分析，并且使得我们的调查有据可循。在调查内容方面，我们先从最主要的入手。显然，在这部分最主要的是学生是否会结婚和生育。我们基于学生的恋爱背景调查他们的结婚意愿，基于结婚意愿希望得到他们关于婚姻因素、婚姻政策以及一些婚姻行为的看法，作为当代大学生对于结婚观念的输出。基于学生是否愿意生育，愿意生几胎的调查结果，将他们分类，进而挖掘他们对于生育价值、生育政策、生育因素的不同看法，形成看法与结果之间的相关性链条，为探索真正有用的、与社会现状相符合的生育政策及措施扫除一定的障碍。最后，我们对大学生们的网络社交与婚育观之间的关系很感兴趣。这是因为婚育观的形成势必与个人以及外界因素二者都息息相关。我们希望基于这些外界因素，来观察是哪些声音在影响着当代大学生们的观念，这种影响是积极的还是消极的，哪些渠道可以帮助大学生们在婚育观的问题上发声。我们在查阅相关文献时发现，有基于调查问卷的研究曾发现了生育意愿与互联网圈层的参与度相关的现象，例如参与动漫、网文、“追星”更多的青年群体与相对更

低的婚育意愿存在相关，参与健身、电竞更多的青年群体与相对更高的婚育意愿存在相关等。我们希望利用这次问卷调查获取一些数据，来验证这个结果并尝试探究出现这种情况原因。

4 问卷调查情况

4.1 问卷有效性

在调查问卷发放过程中，仅要求全日制学生受访者进行答卷，在一定程度上保证了问卷的准确性。本次问卷调查共计回收 545 份问卷：其中有效问卷为 455 份，构成本次问卷调查的样本数据；无效问卷为 90 份，无效问卷主要排除非全日制大学生的受访者所填问卷以及其他原因所导致的无效问卷。有效问卷的比例是 83.49%，其回收数量满足了问卷调查最初的设计目标，保证了问卷的准确性并减小了误差。

4.1.1 样本数据信度检验

信度（reliability）衡量的是测量结果的可靠性与一致性。问卷设计质量的信度检验，指的是对问卷测量结果准确性的分析，即对设计的问卷在多次重复使用下得到的数据结果的可靠性的检验。在实际应用中，信度检验多以相关系数表示，常用的方法有：重测信度，复本信度，折半信度，克朗巴哈信度，评分者信度等。本研究采用了 Cronbach's Alpha 信度系数值来检验各量表问题的信度，信度检验结果如下表3所示：

表 3: 信度检验结果

变量	Cronbach's Alpha 信度系数	项数
对婚前行为的认同程度	0.851	9
对网络话题的参与程度	0.639	7
对 APP 的使用频率	0.542	8
社会行为对婚育观的影响	0.787	7

对婚前行为的认同程度的 Cronbach's Alpha 值为 0.851，对网络话题的参与程度的 Cronbach's Alpha 值为 0.639，对 APP 的使用频率的 Cronbach's Alpha 值为 0.542，社会行为对婚育观的 Cronbach's Alpha 值为 0.787，影响的所有变量的信度均在 0.5 以上，样本信度适中，证明样本数据可信度高，可进行深入分析和研究。

4.1.2 样本数据效度检验

效度（validity）即有效性，是指测量工具或手段能够准确测出所需测量的事物的程度。效度是指根据所测量到的结果反映所想要考察内容的程度，测量结果与要考察的内容越吻合，效度越高；反之，效度越低。效度分为三种类型：内容效度、准则效度和结构

效度。本研究利用前人已研究过的成熟量表，因此具有良好的内容效度和准则效度，下面对样本的结构效度进行检验。首先对样本数据进行 KMO 值检验和 Bartlett 球体检验的近似卡方值显著性检验。KMO 值应大于 0.6，显著性应小于 0.01 才能符合要求，表示该变量内维度相关性较强。检验结果如下表4所示：

表 4: 效度检验结果

变量	KMO	Sig
对婚前行为的认同程度	0.789	0.000
对网络话题的参与程度	0.639	0.000
对 APP 的使用频率	0.646	0.000
社会行为对婚育观的影响	0.812	0.000

对婚前行为的认同程度的 KMO 值为 0.789，检验显著；对网络话题的参与程度的 KMO 值为 0.639，检验显著；对 APP 的使用频率的 KMO 值为 0.646，检验显著；社会行为对婚育观的影响的 KMO 值为 0.812，检验显著。所有变量的 KMO 值都在 0.6 之上，检验显著值都小于 0.001，满足因子分析的要求，且效度检验结果令人较满意。

4.2 受访者基本信息

调查问卷的最后一部分为基本信息，在基本信息部分调查了受访者的性别、年龄、出生省份、户口类型、职业、受教育程度、独生情况、月消费额度共 8 个方面，调查问卷基本信息如下：

1. 性别。本次问卷在受访者性别方面，男性占比 61.65%，女性占比 38.35%。性别不平衡的情况需要在之后的分析中进行调整。

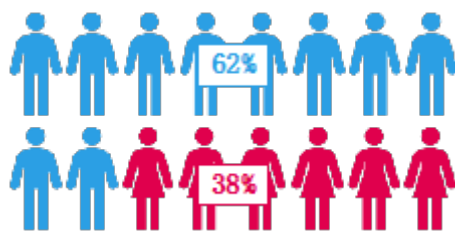


图 8: 受访者性别分布

2. 年龄。在受访者年龄方面，集中分布在 18-24 岁，年龄状况符合研究要求。
3. 出生省份。在受访者地区来源方面，全国各地均有所涉及，其中来自江苏、河北、福建的最多，分别占比 18.35%、13.03% 和 11.38%。
4. 户口类型。在受访者户口方面，类型为农村的受访者占比 25.14%，类型为城镇的受访者占比 74.86%。

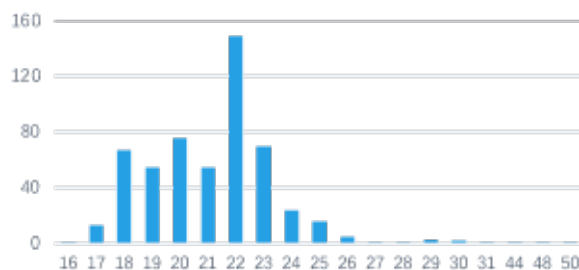


图 9: 受访者年龄分布

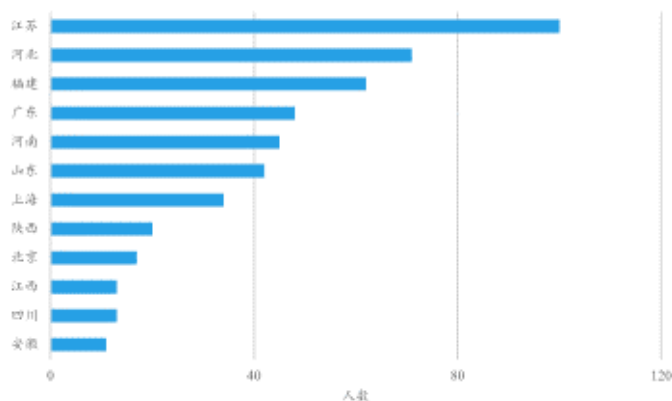


图 10: 受访者人数大于 10 的省份

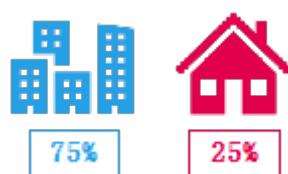


图 11: 受访者户口类型分布

5. 职业。在受访者职业方面，职业为学生的受访者占比 83.49%；职业为企事业单位职工的受访者占比 8.07%；职业为专业人士的受访者占比 3.67%，其中包括教师、律师、医护人员等；职业为其他的受访者占比 4.77%。
6. 受教育程度。在受访者受教育程度方面，受教育程度为专科的受访者占比 3.12%；受教育程度为本科的受访者占比 66.24%；受教育程度为硕士的受访者占比 28.99%；受教育程度为博士及以上的受访者占比 1.65%。其中，受访者的受教育程度主要为本科和硕士，占比 95.23%。



图 12: 受访者受教育程度分布

7. 独生情况。在受访者独生情况方面，不是独生子女的受访者占比 37.80%；是独生子女的受访者占比 62.2%。

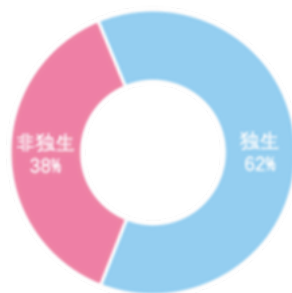


图 13: 受访者独生情况分布

8. 月消费额度。在受访者月消费区间方面，月消费额为 1000 元以下的调查者占比 7.34%；月消费额为 1000-3000 的受访者占比 73.03%；月消费额为 3000-5000 的受访者占比 14.13%；月消费额为 5000 以上的受访者占比 5.51%。

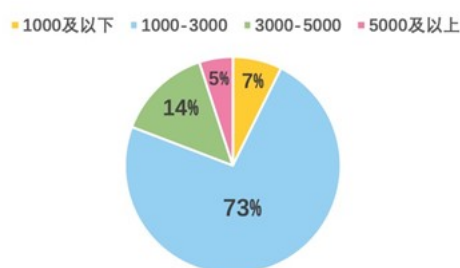


图 14: 受访者月消费额度分布

综合调查问卷基本信息可知，调查问卷的回收满足初始方案的设计，各调查指标百分比满足最初制定的可接受标准，因此数据可做进一步统计、推断和分析。

5 大学生婚育意愿的现状与分析

为了解现阶段大学生婚育意愿的现状，调查小组对受访者进行了一系列调查。旨在通过分析数据背后的规律，探究导致婚育意愿下降的原因及影响婚育意愿的因素。

5.1 婚姻意愿及观点现状

就大学生婚姻意愿现状而言，通过概述大学生的结婚意愿、认为促进结婚的因素、认为抑制结婚的因素的调查数据，分析大学生的婚姻观及其存在的规律。

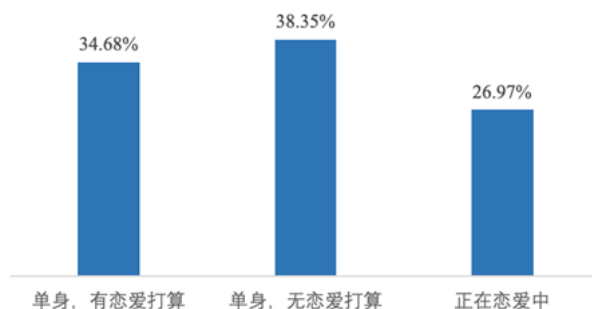


图 15: 大学生恋爱状态

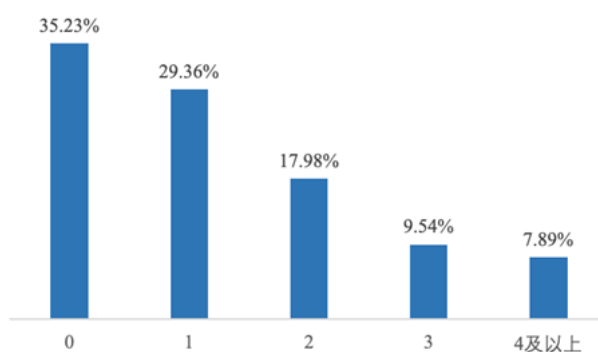


图 16: 大学生恋爱次数

5.1.1 恋爱状态及次数

大学生的恋爱状态及恋爱次数的调查结果如上图15与16所示。从恋爱状态来看，34.68% 的大学生目前为单身且有恋爱打算；38.35% 的大学生目前单身且无恋爱打算；26.97% 的大学生正在恋爱中。从恋爱次数来看，35.23% 的大学生没有谈过恋爱，29.36% 的大学生谈过 1 次恋爱，17.98% 的大学生谈过 2 次恋爱，9.54% 的大学生谈过 3 次恋爱，7.89% 的大学生谈过 4 次及以上恋爱。结果表明，大部分的大学生有较强的恋爱意愿，并且至少都谈过 1 次恋爱。

5.1.2 结婚意愿

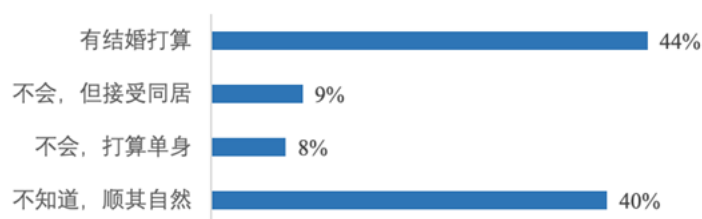


图 17: 结婚意愿

大学生的结婚意愿的调查结果如上图17所示。44% 的大学生有结婚打算；9% 的大学生没有结婚意愿但接受同居；8% 的大学生没有结婚意愿且打算单身；40% 的大学生

没有明确的意愿。结果表明，超过半数的大学生没有明确的结婚意愿。

5.1.3 理想结婚年龄

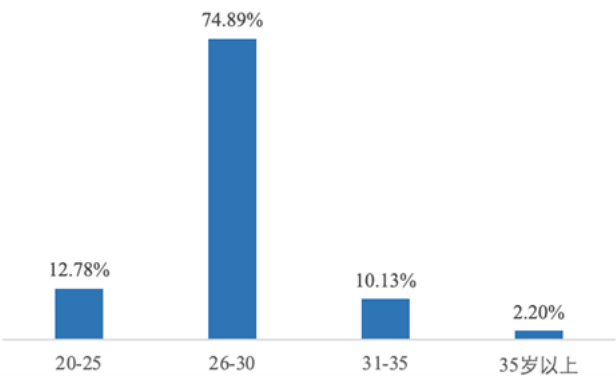


图 18: 理想结婚年龄

有结婚意愿的大学生的理想结婚年龄的调查结果如上图18所示。12.78% 的大学生希望在 20-25 岁结婚；74.89% 的大学生希望在 26-30 岁结婚；10.13% 的大学生希望在 31-35 岁结婚；2.20% 的大学生希望在 35 岁结婚。结果表明，大部分大学生希望在 26-30 岁这个年龄段结婚。

5.1.4 促进结婚的因素

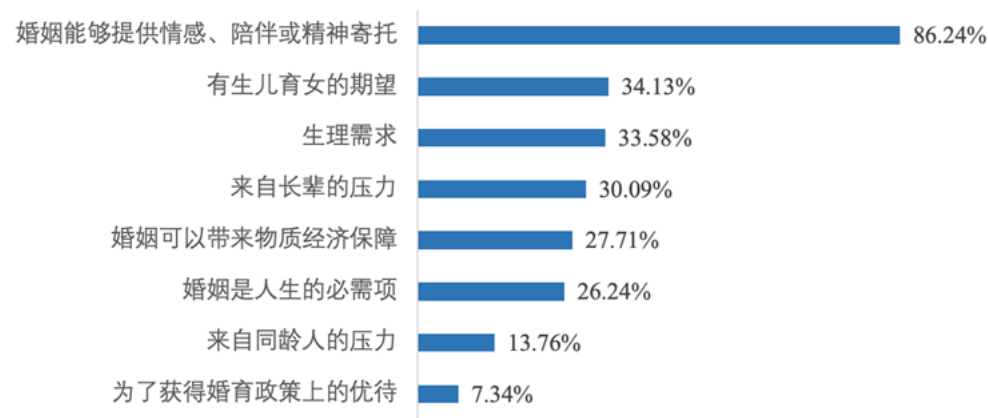


图 19: 促进结婚的因素

大学生认为促进结婚的因素的调查结果如上图所示。86.24% 的大学生认为婚姻能够提供情感、陪伴或精神寄托；30% 以上的大学生认为结婚可以达成他们生儿育女的期望值、满足生理需求，且来自长辈的压力是重要的影响因素；20% 以上的大学生认为婚姻可以带来物质经济保障，并且是人生的必需项；认为来自同龄人的压力和为了获得婚育政策上的优待是促进因素的大学生分别占 13.76% 和 7.34%。结果表明，绝大多数的大学生都认同婚姻带来的精神价值，但长辈的压力和物质保障也有一部分的影响。

5.1.5 抑制结婚的因素

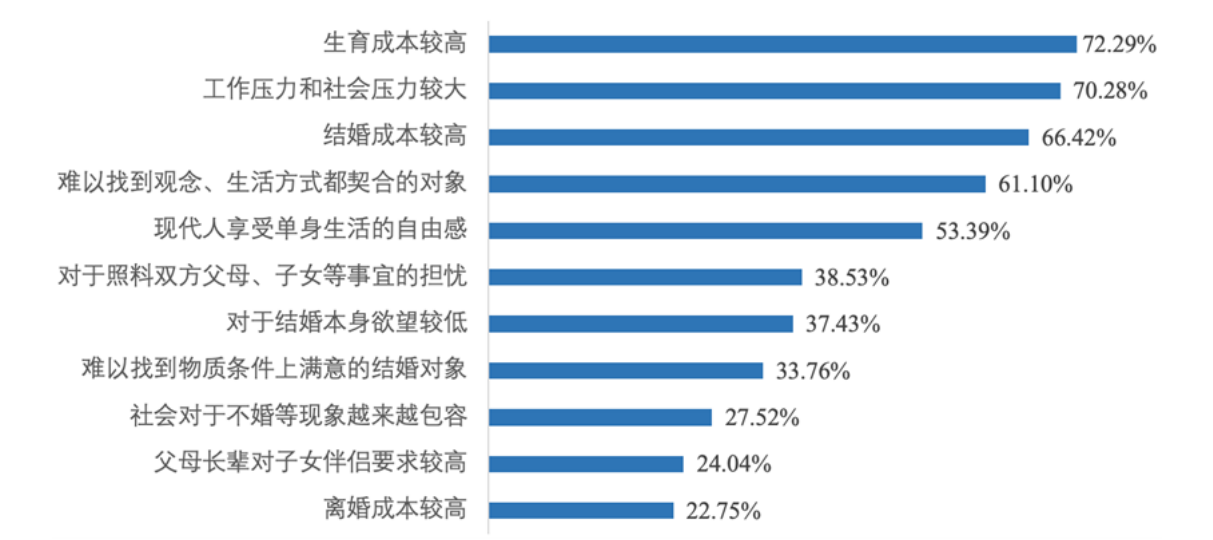


图 20: 抑制结婚的因素

大学生认为抑制结婚的因素的调查结果如上图所示。70% 左右的大学生认为较高的生育和结婚成本、工作和社会压力是导致他们不愿结婚的原因；61.10% 的大学生认为自己难以找到观念、生活方式都契合的生活对象，53.39% 的大学生享受单身生活的自由感，这些观念上的因素都使得他们或不愿或难以结婚；超过 30% 的大学生考虑到照顾双方父母和子女等事宜、结婚欲望不高、难以找到物质条件上满意的结婚对象而倾向于不结婚；认为由于社会对于不婚等现象越来越包容、长辈对子女伴侣的要求较高、离婚成本较高导致不愿结婚的大学生分别占比 27.52%、24.04% 和 22.75%。结果表明，婚育成本和生活成本较高是抑制结婚意愿的最主要的原因，而对于自己和伴侣的精神上的追求也使得大学生没有那么强的意愿结婚。

5.1.6 对婚姻行为的认同度

大学生对各种婚姻行为的认同度的调查结果如下图21所示。超过半数的大学生都认同或较认同婚前同居、婚前性行为、不婚和丁克行为，大部分的大学生不认同或不太认同闪婚、非婚非育和开放性关系行为，而对于堕胎和同性恋行为，大学生们的观点并不一致。结果表明，大部分大学生的婚姻观念已经比较开放，但对于一些不太符合社会风俗的行为较为保守。

5.2 生育意愿与观点现状

就大学生生育意愿现状而言，通过概述大学生的生育意愿、比较意愿生育和不愿生育的大学生对于生育带来的价值、抑制生育的因素、促进生育的政策的不同，分析大学生的生育观及其存在的规律。

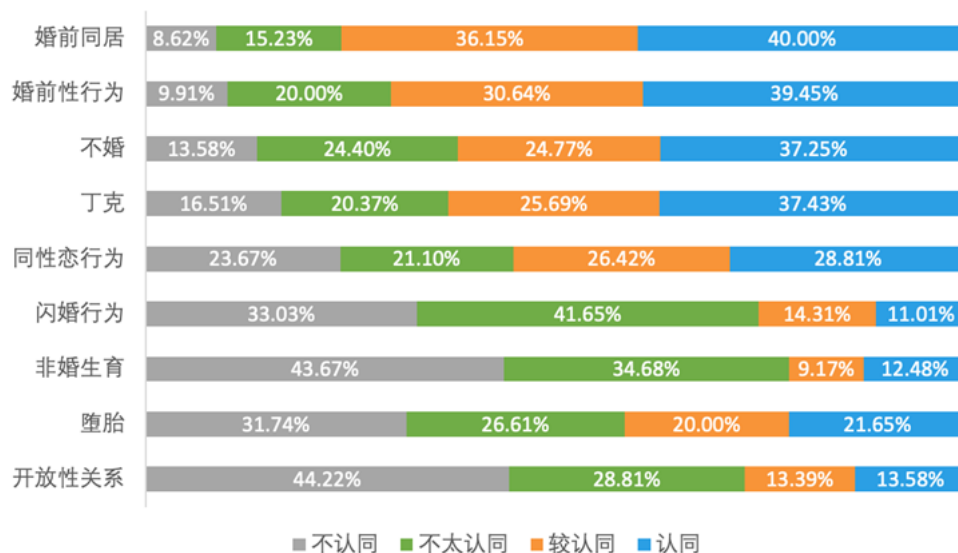


图 21: 对婚姻行为的认同度

5.2.1 理想及意愿子女数

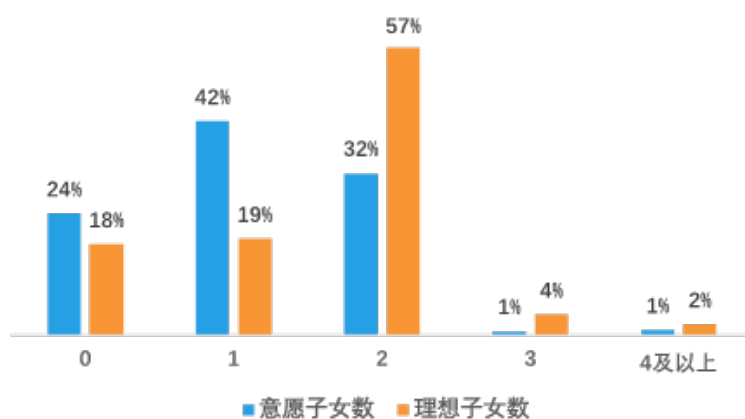


图 22: 理想及意愿子女数

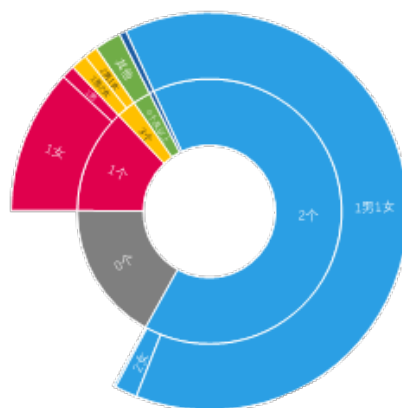


图 23: 理想儿子及女儿数

大学生的生育意愿的调查结果如上图22与23所示。从意愿子女数来看，24% 的大学生不愿生育，42% 的大学生意愿生育 1 胎，32% 的大学生意愿生育 2 胎，2% 的大学生意愿生育 3 胎及以上。从理想子女数来看，18% 的大学生不希望生育，19% 的大学生希望生育 1 胎，57% 的大学生希望生育 2 胎，6% 的大学生希望生育 3 胎及以上。从理想儿子及女儿数来看，62.94% 的大学生希望生育一儿一女，11.74% 的大学生希望生育一女，1% 左右的大学生希望生育一儿、两女、两儿一女或一儿两女。结果表明，大部分大学生意愿生育 1 胎，而希望生育 2 胎，说明由于各种现实原因，大部分大学生的生育意愿下降。同时，大部分大学生都期望生育一男一女，一部分希望生育一女，这可能是因为受到“龙凤呈祥”、“女儿是贴心小棉袄”等社会观念的影响。

5.2.2 生育带来的价值

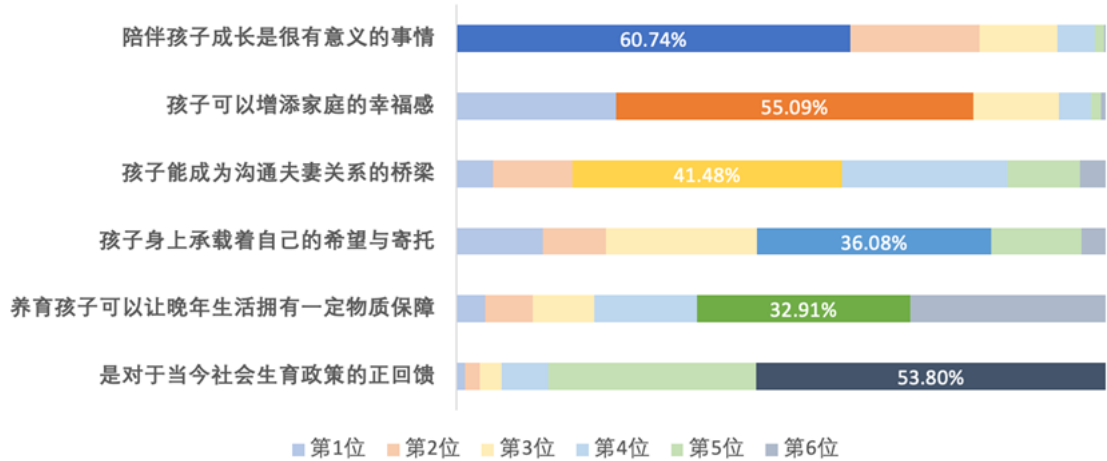


图 24: 生育带来的价值

大学生认为生育带来的价值的调查结果如上图24所示。60.74% 的大学生认为陪伴孩子成长很有意义是生育的最大价值；55.09% 的大学生认为孩子可以增添家庭幸福感是生育的极重要价值；41.48% 的大学生认为孩子能成为沟通夫妻关系的桥梁是生育的重要价值；36.08% 的大学生认为生育的一部分价值来自于孩子身上承载着自己的希望与寄托；32.91% 的大学生不太在意养育孩子可以让晚年生活拥有一定物质保障；53.80% 的大学生不认为生育是对于当今社会生育政策的正回馈。结果表明，大部分大学生更加在意孩子成长本身的意义以及孩子为家庭带来的正面影响，而不太在意孩子带来的物质保障和当今的社会生育政策。

5.2.3 抑制生育的因素

大学生认为抑制生育的因素的调查结果如图25所示。从意愿生育的大学生角度看，82.33% 的大学生认为抚养压力会明显降低他们的生育意愿；63.98% 的大学生认为居住环境压力也是降低生育意愿的重要原因；超过 50% 的大学生认为抑制生育的重要因

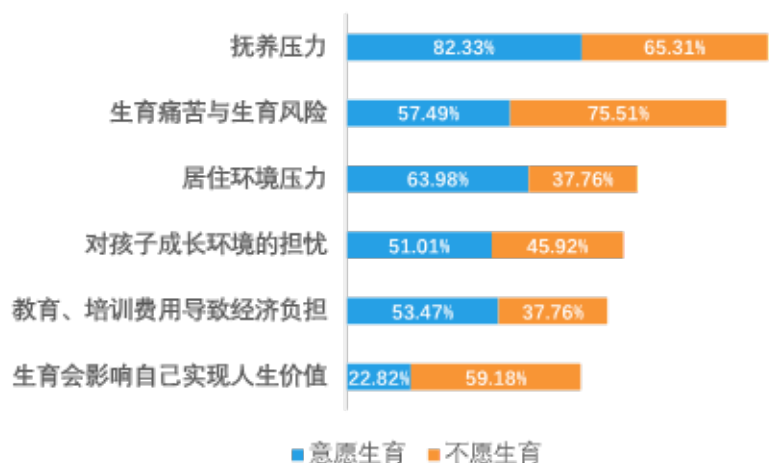


图 25: 抑制生育的前六个因素

素包括生育痛苦与生育风险、对孩子成长环境的担忧以及教育培训费用导致经济负担；22.82% 的大学生认为生育会影响自己实现人生价值而不愿生育。从不愿生育的大学生角度看，75.51% 的大学生认为生育痛苦与风险会明显降低他们的生育意愿，60% 左右的大学生认为抚养压力和影响人生价值实现也是降低生育意愿的重要原因；45.92% 的大学生担忧孩子的成长环境而不愿生育，超过 37% 的大学生认为抑制生育的因素还包括居住环境压力和教育培训费用带来的经济负担。结果表明，意愿生育的大学生更担心抚养和居住环境压力，不愿生育的大学生更担心生育痛苦与风险和抚养压力，说明生育痛苦与风险是影响生育意愿的重要因素，当前的妇幼医疗水平和医学知识普及的不足抑制了大学生的生育意愿。

5.2.4 提高生育意愿的政策

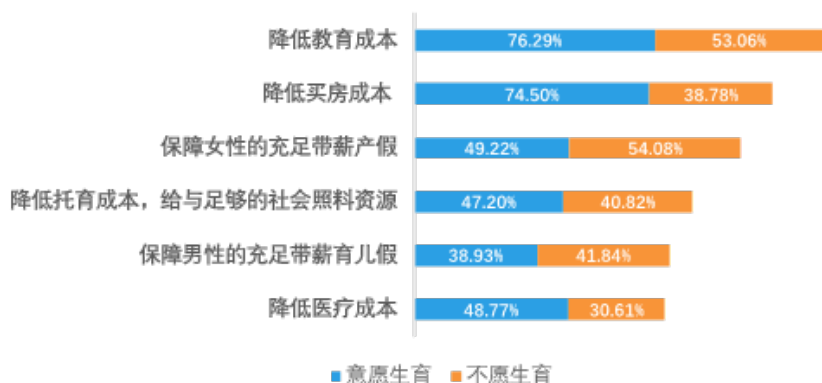


图 26: 促进生育的前六个因素

大学生认为促进生育的因素的调查结果如上图26所示。从意愿生育的大学生角度看，超过 70% 的大学生认为降低教育成本和买房成本能够明显提升他们的生育意愿；50% 左右的大学生认为促进生育的重要因素包括保障女性的充足带薪产假、降低托育成本和降低医疗成本；38.93% 的大学生认为保障男性的充足带薪育儿假也能一定程度促进生

育。从不愿生育的大学生的角度看，超过 50% 的大学生认为降低教育成本和保障女性的充足带薪产假能够明显提升他们的生育意愿；40% 左右的大学生认为促进生育的重要因素包括降低买房成本、降低托育成本和保障男性的充足带薪育儿假；30.61% 的大学生认为降低医疗成本也能一定程度促进生育。结果表明，意愿生育的大学生更在意教育和买房成本，不愿生育的大学生更在意教育成本和女性产假，说明女性产假是影响生育意愿的重要因素，当前的工作和社会歧视是导致生育意愿较低的重要原因。

6 婚育意愿的性别差异

在 2021 年澎湃网的一项有关婚育的调查发现，对于“Z 世代”来说女性的结婚意愿明显低于男性。女性表示“不结婚”和“不确定会不会结婚”的人数占 43.92%，比男性多出 19.29%。女性表现出更加强烈的婚姻犹疑态度。而在对恋爱态度的调查中，并未发现这种性别差异。这说明在婚姻这件事上，女性的态度更加谨慎。探究男性和女性在婚育方面的差异对我们的研究是很重要的，如果两性之间对于婚育方面存在较大的差异，那么相关的政策建议也不应该“一视同仁”。

6.1 研究方法介绍

在这一部分，我们主要采用对比分析的方法研究性别之间的差异。首先，为了使得研究结果更为直观，我们使用了分层抽样的方法，基于性别对问卷搜集到的数据进行了分层抽样。具体地，我们分别随机抽取了男性和女性各 200 个样本，共计 400 个样本。完成分层抽样以后，我们主要通过列联表分析以及卡方检验的方法对男女婚育观念的差异进行比较。主要分析的角度是：

1. 性别和结婚意愿
2. 性别和促进/抑制结婚的因素
3. 性别和理想/意愿子女数
4. 性别和提高生育的政策
5. 性别和婚育观念受何外界影响

下表5是卡方检验中，在 0.05 的显著性水平下，表现非常显著的变量。

从表5中可以整体看出，男性和女性大学生群体在婚育方面的观念存在非常大的差异。其中在意愿/理想子女数、阻碍/提高生育的因素、婚育观念受到外界因素的影响这几个方面表现出的差异最大。

表 5: 性别差异卡方检验结果

变量	χ^2 值	自由度	P 值
结婚意愿	13.77	3	0.003
恋爱状态	7.48	2	0.023
促进结婚的因素	24.59	7	0.001
抑制结婚的因素	38.87	10	0.000
意愿子女数	26.48	4	0.000
理想子女数	25.64	4	0.000
阻碍生育的因素	34.05	10	0.000
提高生育的因素	43.38	11	0.000
恋爱相亲育儿综艺	30.38	3	0.000
国家生育政策	16.15	3	0.001
专家教授发表看法	15.46	3	0.001

6.2 结婚意愿

从图27中可以看出，男女大学生在结婚意愿上存在较大的差异。其中，不愿结婚希望保持单身的男性和女性分别占整体的 4.5% 和 3.75%；不愿结婚但可以接受同居的男性和女性分别占 5.5% 和 4.5%。这两方面男性和女性并未表现出差异。在结婚意愿上，相较于男性，女性更倾向于顺其自然，在百分比差异上比男性高出 7.25%；而男性表现出了较强烈的结婚意愿，所抽取的样本中，有明确结婚打算的男大学生所占的比例比女大学生多了 9%。



图 27: 结婚意愿的性别差异

综上所述，在结婚意愿方面，男性的结婚意愿比女性更为强烈，女性对于结婚没有强烈的意愿，更倾向于顺其自然的发展。

6.3 恋爱状态

从图28中可以看出，男女大学生在恋爱状态上也存在一定的差异。其中，单身且没有恋爱打算的性别差异不大，仅有 1.25%。在恋爱中的女性和男性分别占比 11.75% 和 16.5%，女性比男性更多。单身且有恋爱打算的状态下，男性和女性表现出很大差异，此选项中，男性和女性分别占比 20.5% 和 14.5%，相差了 6%。

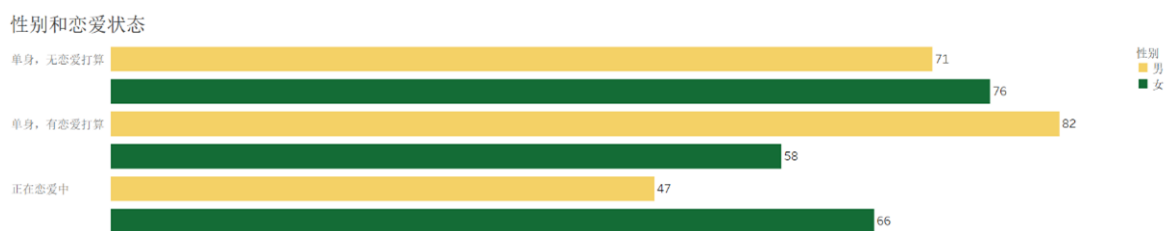


图 28: 恋爱状态的性别差异

因此，可以看出在恋爱状态方面，虽然目前男大学生恋爱经验远低于女大学生，但是男大学生的有较积极的恋爱打算，他们希望可以早日“脱单”。综合，考虑 2.2 中讨论的男女在结婚意愿方面的差异，可以看出对于恋爱的态度，女性并未表现出和其在结婚意愿一致的不乐观，可见女性对于婚姻是比较严谨的。

6.4 促进/抑制结婚的因素

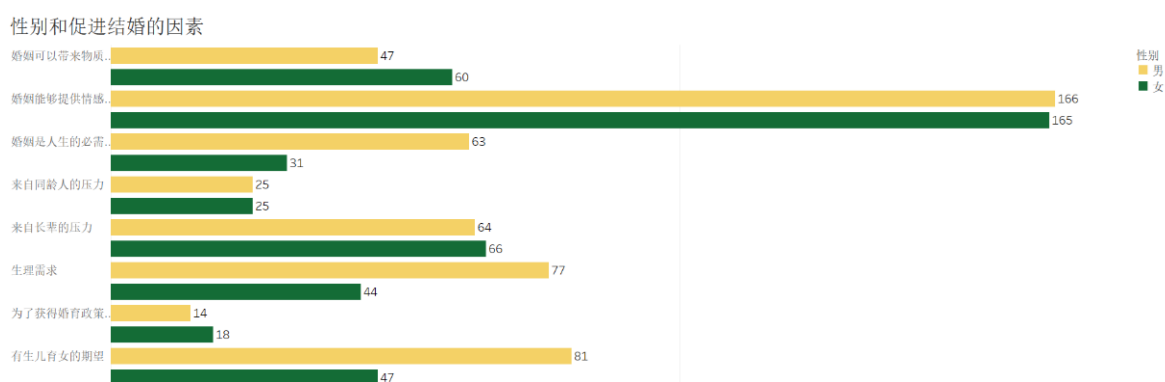


图 29: 促进结婚因素的性别差异

从图29中可以看出，男女大学生在对促进结婚的因素方面的看法是不同的。其中，认为婚姻是人生的必需项的男性占比比女性高了 3.22%；认为生儿育女的期望是促进结婚因素的男性比女性高了 3.42%；同时，男性更多地认为生理需求也是促进结婚的重要因素。二者一致认同的是婚姻能够提供情感、陪伴或精神寄托。在促进结婚因素的看法上，由此可以看出男性的婚恋观念更加的传统和保守。

从图30中可以看出，男女大学生在对抑制结婚的因素方面的看法同样大相径庭。其中，最为鲜明的几个特点为：第一，相较于男性，更多的女性认为对于结婚本身欲望的降低会抑制结婚。第二，女性也会更多地担忧结婚之后对于双方父母、子女等的照顾问题。此外，女性也认为现代人享受单身生活的心态也是一大抑制结婚的因素。而男性更关注的问题在于结婚的成本以及社会的压力问题，具体地，认为结婚成本高的男性占比 7.1%，女性占比 5.6%，二者相差 1.5%，认为工作和社会压力大的男性也明显高于女性。

总而言之，在抑制结婚的因素方面，男性更着眼于结婚成本、社会压力等经济问题，

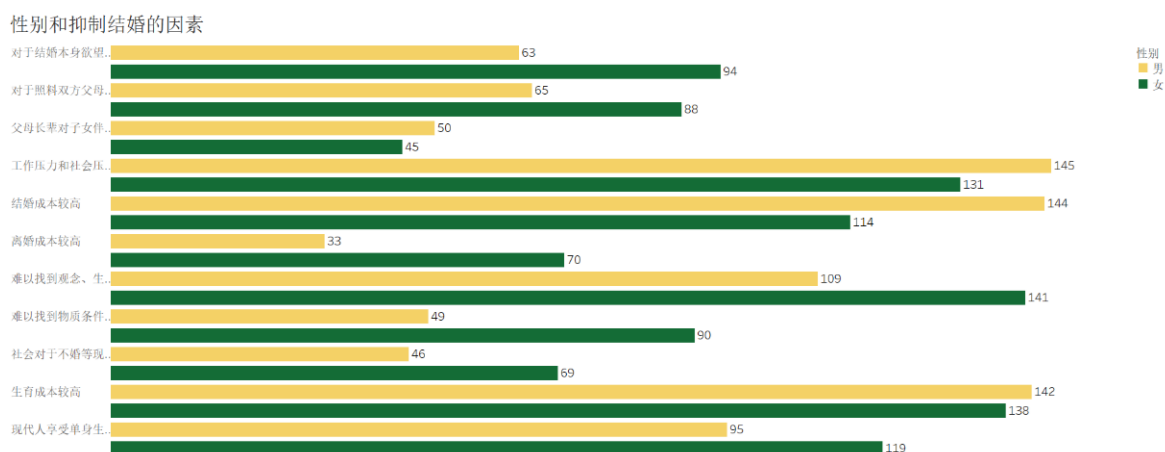


图 30: 抑制结婚因素的性别差异

而女性则更关注婚姻本身存在的家庭矛盾。

6.5 理想/意愿子女数

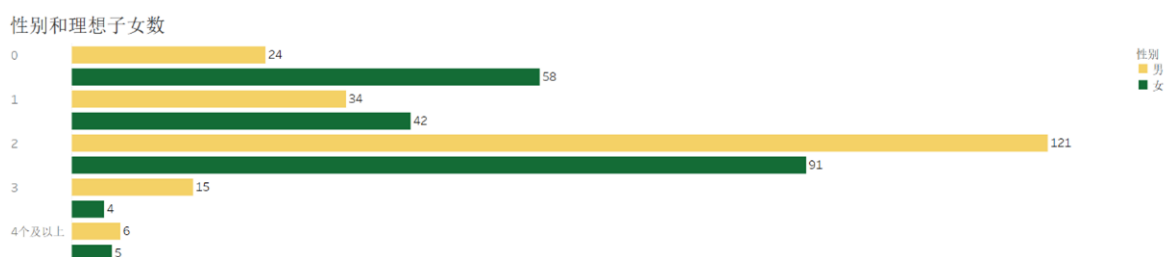


图 31: 理想子女数的性别差异

图31表示不同性别的理想子女数（不考虑任何因素，如经济实力等）。理想子女数为 0 的大学生中，男性和女性分别占整体的比例为 6% 和 14.5%；而选择理想子女数为 2 的男性和女性的占比为 30.25% 和 22.75%。可以发现，男性相较于女性更期望可以生儿育女。

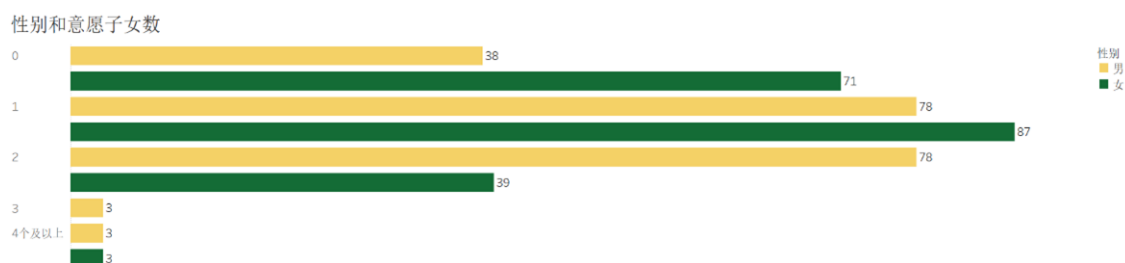


图 32: 意愿子女数的性别差异

图32表示不同性别的意愿子女数（考虑现实因素），明显看出，考虑到现实因素后，相较于理想子女数，无论是男性还是女性，选择不生育和生育 1 个孩子的比例有所增加，

生育 2 个以上孩子的比例明显降低。经济、社会、工作等方面的压力对于生育意愿产生了一定的抑制作用。此外，在意愿子女数上，男性仍然表现出更强的生儿育女意愿。

6.6 提高/抑制生育的政策

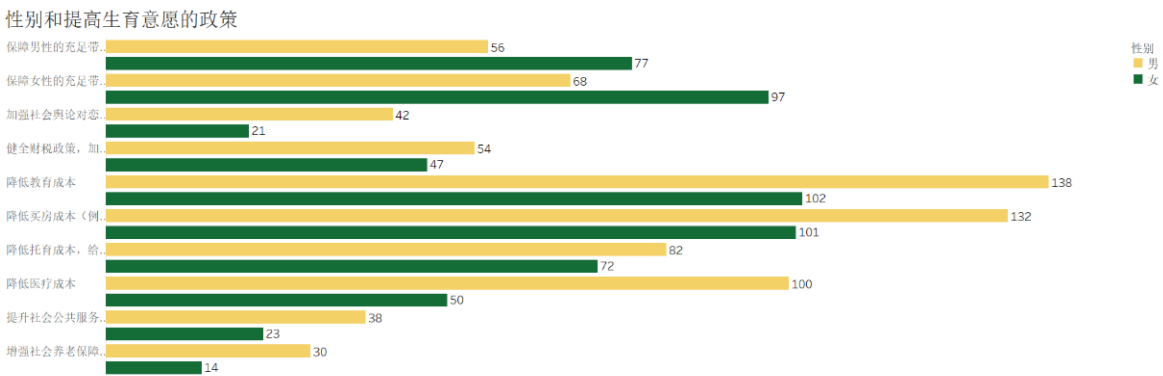


图 33: 提高生育的政策

图33表示不同性别大学生对如何提高生育的看法，其中，男性相较于女性更看重社会可以降低一些教育、买房、医疗等方面的成本，男性更愿意从经济等客观角度出发思考如何提高生育意愿。认为保障双方的充足带薪产假可以有效提高生育的男性和女性占整体的比例分别为 9.23% 和 12.95%，除去生育的成本问题，女性更关注一些带薪产假之类的人文因素。

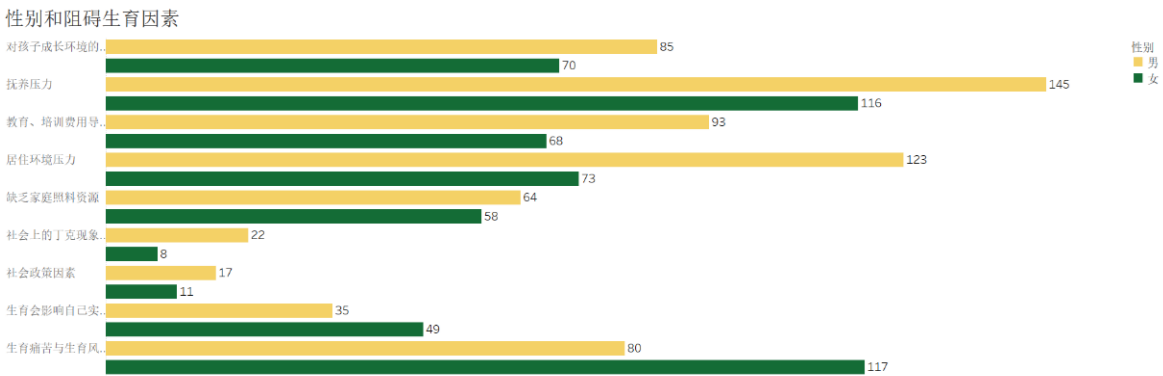


图 34: 阻碍生育的因素

图34表示不同性别大学生对有哪些因素阻碍会生育的看法。这一点和两性对提高生育政策的看法是较为一致的。具体表现为，男性更看重抚养压力、居住环境压力等经济问题，而女性更关注生育痛苦和职业发展等人文要素。

6.7 婚育观念受何外界因素影响

下图35表示不同性别大学生的婚育观受恋爱育儿综艺、国家生育政策、专家教授看法等外界因素的影响程度。可以看出，外界信息对男性婚育观的影响更小，大部分男性完全不会受到这些隐私影响，女性更喜欢关注婚育相关的外界信息。

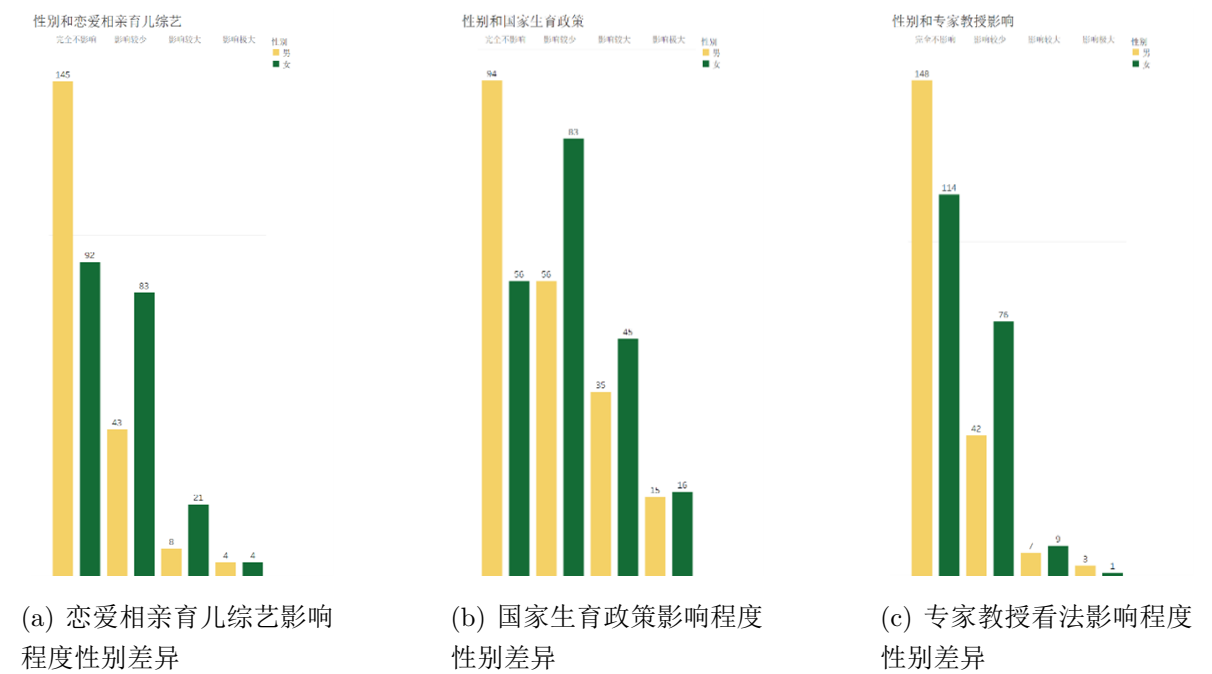


图 35: 婚育观受外界因素的影响

7 互联网参与与婚育意愿

为了得知各种互联网参与对婚育意愿会产生什么样的影响，我们组在调查问卷中特意划出了一大版块，设计了一系列矩阵量表题，获取了当代青年人们的互联网参与信息。在这一部分，我们将互联网参与分为“互联网圈层参与”、“APP 使用频率”以及“其他互联网外部影响”三个部分，着重于这三个部分对于婚育意愿可能产生的影响。

7.1 总体方法框架介绍

在这一部分，我们主要采用机器学习方法对数据进行针对性地分析。分析内容主要被分为两部分——一是对被调查者总体进行定性分析的部分，核心是将被调查者依据其圈层参与情况与 APP 使用情况进行非监督分类，以此对他们的婚育观进行类型上的比较；二是对被调查者进行定量分析的部分，核心是将被调查者的圈层参与情况与 APP 使用情况进行量化，进而通过数值表示体现不同的圈层、APP 以及其他外部因素对于当代青年人们的影响情况。首先给出关于总体方法的框架图供各位参考，如下图36所示。

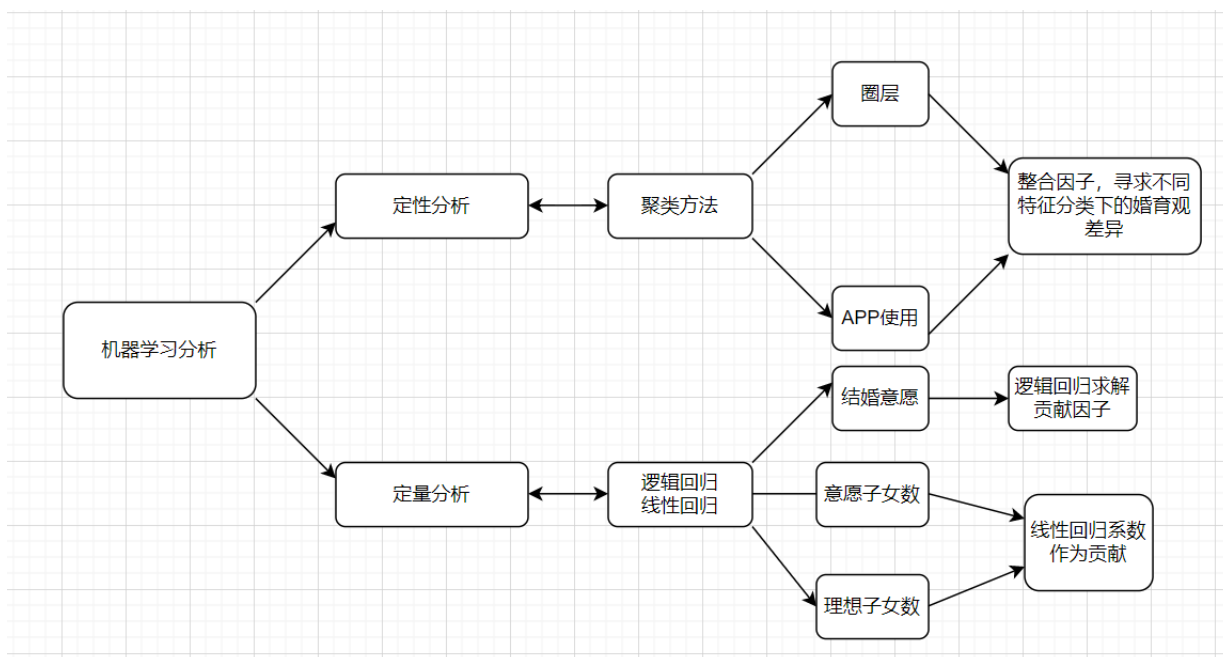


图 36: 总体分析框架图

在定性分析方面，我们主要采用聚类这一典型的无监督方法。针对已获得的休闲上网时长、圈层参与程度以及 APP 使用情况信息，对所有的样本进行聚类分析，起到整合因子的效果，将其自然聚为几个大类，接着统计这几个大类上样本的结婚意愿、生育意愿以及恋爱次数信息，最后对这些信息进行可视化比较，用于寻求不同特征分类下的婚育观差异。

而在定量分析方面，我们主要采用了逻辑回归与线性回归的模型。在机器学习当中，线性回归常常运用于回归问题，而逻辑回归则运用于分类问题。在这里分别运用到这两个模型，和我们对于结婚意愿以及生育意愿这两个问题的机器学习定义有关。我们将结婚意愿这方面的问题给主要作为一个分类问题进行看待，因而最终得到的各特征的影响程度数值是作用于一个分类模型的，故对于各个特征的贡献因子计算，我们采用逻辑回归的方法；对于生育意愿而言，我们衡量一个当代年轻人生育意愿采用的是他们的理想子女数以及意愿子女数这两个数值，因此这个问题更多的是一个偏向回归方面的问题，故我们采用线性回归模型计算各个特征对于生育意愿的影响因子。综上，在这一部分，我们主要以各个特征的贡献因子数值作为载体，从而对当代青年人各个圈层的参与程度、各个 APP 的使用程度以及其他的外部因素对于婚育意愿的影响程度进行分析，最终得出相关结论。

最后需要再次强调，在这一部分尽管我们使用了一些可解释性较强的线性回归方法，但是人的婚育观念本身还是受到各种方面以及各种因素的影响的，互联网的参与程度可能仅仅是其中的一小部分，因而在这里我们得到的影响因子与贡献因子结果体现的是相关关系而非因果关系。我们对“互联网行为影响婚育观念”这一可能的议题有一定的研究兴趣，进而进行了相关研究，希望可以通过当代青年人的互联网行为探究这些行

为对他们的相关性影响，进而提出一些可能的建议。

7.2 数据预处理

接着，我们将介绍一下这部分得到的调查数据以及我们对于这部分数据的预处理工作。数据预处理是为了便于后续的机器学习分析。因为我们得到的数据更多是被调查者们对于圈层参与以及 APP 使用情况的文字叙述，因而将数据的这些特征转为数值类型是十分必要的。

我们得到的数据是来自 545 位被调查者的调查问卷，其中男性 336 人，女性 209 人。由于我们的调查问卷发放主要以小组成员在 QQ 群和微信群等网络媒介发放为主，而小组成员的男生要多于女生，因而相比之下问卷所发的群内男性要略多于女性。这在一定程度上会涉及样本不平衡问题，这一问题我们后续会通过分层抽样的方法尽可能缓和这一问题。

这部分的数据特征主要是各圈层的参与程度、各 APP 的使用频率以及被调查者所受到的外部因素影响程度大小。而数据标签主要是可以反映被调查者结婚观的结婚意愿、恋爱次数以及可以反映被调查者生育观的理想子女数、意愿子女数等等。

首先是对于数据标签的处理。对于“恋爱次数”这一标签，我们在调查问卷当中设置的选项为 0 次、1 次、2 次、3 次与 4 次及以上。为了后续研究的方便，我们将恋爱次数“4 次及以上”的被调查者的恋爱次数统一记为 4 次。其他的“理想子女数”、“意愿子女数”、“理想生儿子数”、“理想生女儿数”同样都有“4 个及以上”这一选项，我们也统一记为 4 次。

在这里需要说明一下“理想子女数”与“意愿子女数”之间的区别。“理想子女数”指的是不考虑任何物质、环境条件下，被调查者想要的子女数；而“意愿子女数”是在考虑各方面物质、环境及其他条件的情况下，被调查者想要的子女数。总体而言，意愿子女数小于等于理想子女数才是比较合理的。

对于结婚意愿这一标签，我们给出的问卷选项依据结婚意愿从低到高依次为“不会，打算单身”、“不会但接受同居”、“不知道，顺其自然”以及“有结婚打算”。因此，我们采用 python 的顺序编码器 OrdinalEncoder，依次以 0、1、2、3 分别对这 4 个选项进行顺序编码，0 代表“不会，打算单身”而 3 代表“有结婚打算”。这样，就完成了对于数据标签的数值化编码。

其次是对上网时长这一特征进行数值化编码。上网时长这一问题我们设置的选项依次为“2 小时及以下”、“2-4 小时”、“4-6 小时”、“6-8 小时”以及“8 小时及以上”。我们通过取选项“中位数”的方法进行数值化编码，即分别用 1、3、5、7、9 依次代表选择这些选项的被调查者们的每日休闲上网时长。对于相关性分析而言，用中位数代表相关的选项区间是合理的。

接着是对于各个互联网圈层的参与程度的特征预处理。在这里，我们选择了“动漫”、“综艺”、“体育/电竞”、“明星”、“科技/数码”、“网文”以及“游戏”这 7 个互联网重要圈层进行研究。在这里，我们采用了矩阵量表题的形式，对于每个圈层，都设置

了“不参与”、“参与较少”、“参与较多”以及“参与程度极高”4个选项。我们依次以0、1、2、3分别对这4个选项进行顺序编码，0代表“不参与”而3代表“参与程度极高”。通过顺序编码刻画各圈层的参与程度在一定程度上是比较合理的。

然后是针对各APP的使用频率的特征预处理。在这里，我们主要选择了一系列以“互联网发声”以及“互联网交流讨论”为主体的主流APP，包括“微博”、“知乎”、“豆瓣”、“虎扑”、“百度贴吧”、“小红书”、“B站”以及“短视频”类APP。对于每个APP，我们设置了“完全不使用”、“使用较少”、“使用较多”以及“使用频率极高”4个选项。同样地，依次以0、1、2、3分别对这4个选项进行顺序编码，0代表“完全不使用”而3代表“使用频率极高”。

这样，我们就把这部分得到的所有数据都进行了数值化编码，最后的数据类型都被转化为int或是float类型。这为后续的机器学习建模做了一个较好的铺垫。

7.3 聚类分析

首先对获得的数据进行聚类分析，探究不同聚类下的被调查者的婚育观差异。在聚类方法方面，我们选择了较为经典的K-means算法。K-means算法的计算速度较快，算法实现不困难且算法原理简单、可解释性较好，这是我们选择这一方法的主要原因。

在聚类超参数 K 的选择上，我们主要对 $K = 4$ 以及 $K = 5$ 都进行了尝试。一方面，我们认为这一问题当中涉及到特征还是比较多的，如果选定聚类类数 $K < 4$ ，很可能无法较为准确地依据特征对样本进行合理地聚类划分。我们都知道，聚类的目标是尽可能使得类内距离小而类间距离大，而在特征较多的情况下，如果聚类超参数 $K < 4$ ，对于这一问题的划分很有可能不够充分。而如果选定聚类类数 $K > 5$ ，鉴于我们获取的被调查者样本数也并没有特别多，如果聚类类数过多，一方面在样本不是特别多的情况下某些聚类的划分不具有代表性，这会给结果分析带来困难，另一方面很有可能导致某一类的样本数较少，而事实上这一样本数较少的聚类很有可能本身应属于另一聚类，这会使得聚类结果的可解释性变差。打个不恰当的比方，后者情况类似于树模型当中的节点样本数很少的子树，理论上需要剪枝操作删去。因此，我们主要对超参数 $K = 4$ 以及 $K = 5$ 进行尝试。

分别以 $K = 4$ 与 $K = 5$ 进行聚类后得到的各聚类样本量占比分别如下图37所示：

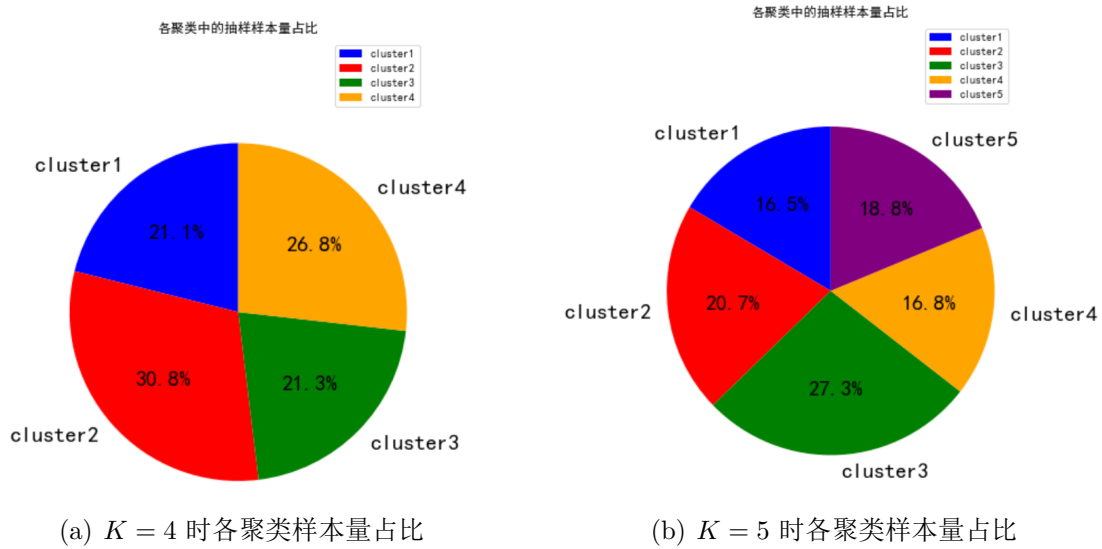


图 37: $K = 4$ 与 $K = 5$ 时各聚类样本量占比

从图中可以看到，选择 $K = 4$ 与 $K = 5$ 的各聚类样本量占比都是相对比较平衡的。这一部分我们用上了全部的 545 个样本，这是为了尽可能在聚类过程中用上所有能用上的信息。除此之外，因为聚类是无监督算法，并不需要对样本量的平衡性有非常严格的要求，因此我们用上了全部样本。

从聚类样本量占比的角度，貌似选择 $K = 4$ 与 $K = 5$ 都可以进行聚类的划分。为了更进一步选择出对于该问题而言最优的聚类参数，我们对于这两种情况计算各个聚类中各个特征的均值，希望通过比较均值寻找出类内距离较小而类间距离较大的一种聚类模型。 $K = 4$ 与 $K = 5$ 情形下的各个聚类中各个特征的均值分别如下表6与7所示。在表中各个特征的最大值用红色字体标出，而最小值用蓝色字体标出。

表 6: $K = 4$ 情形下不同聚类的各特征贡献表

	上网时长	动漫	综艺	体育/电竞	明星	科技/数码	网文	游戏	微博	知乎	豆瓣	虎扑	贴吧	小红书	B 站	短视频
cluster1	2.22	0.63	0.28	0.94	0.31	0.98	0.32	0.97	0.75	1.11	0.40	0.29	0.30	0.76	2.10	0.52
cluster2	3.67	1.76	0.61	1.65	0.43	1.55	1.02	2.20	0.75	1.36	0.29	0.22	1.07	0.43	2.63	0.82
cluster3	8.09	1.40	0.78	1.54	0.72	1.34	1.09	1.93	1.27	1.28	0.58	0.43	0.87	1.10	2.33	1.09
cluster4	3.77	0.77	1.36	0.90	1.34	0.90	1.07	1.03	1.99	1.37	0.86	0.24	0.62	1.98	2.11	1.34

表 7: $K = 5$ 情形下不同聚类的各特征贡献表

	上网时长	动漫	综艺	体育/电竞	明星	科技/数码	网文	游戏	微博	知乎	豆瓣	虎扑	贴吧	小红书	B 站	短视频
cluster1	4.52	1.06	0.80	1.23	0.83	1.21	1.08	1.59	1.50	1.50	0.53	0.18	0.82	1.26	2.21	1.08
cluster2	4.28	1.10	0.93	1.42	0.86	1.27	0.86	1.65	1.34	1.17	0.54	0.42	0.72	1.17	2.34	1.14
cluster3	4.21	1.14	0.72	1.27	0.56	1.15	0.72	1.47	1.12	1.28	0.57	0.32	0.81	1.04	2.29	0.93
cluster4	4.13	1.31	0.94	1.25	0.78	1.12	0.93	1.39	1.25	1.36	0.64	0.22	0.78	1.07	2.31	0.70
cluster5	3.96	1.15	0.71	1.36	0.68	1.25	0.88	1.64	1.31	1.39	0.48	0.35	0.67	1.12	2.37	0.91

比较两种不同超参数情形下的各聚类的特征，我们可以明显地发现， $K = 4$ 情形下不同的聚类之间的类间差异要明显大于 $K = 5$ 的情形。以“上网时长”这一特征进行

举例，在 $K = 4$ 情形下，休闲上网时长最长的第三类聚类的平均日休闲上网时长达到了 8.09 小时，而休闲上网时长最短的第一类聚类仅仅只有 2.22 小时，极差为 5.87 小时；相比之下， $K = 5$ 时“上网时长”这一特征的极差仅仅为 0.56 小时，可以看到两种情形下“上网时长”这一特征类间差距确实不在一个量级。事实上将表6与7的其他特征依次进行横向对比，会发现 $K = 4$ 情形下的其他特征的总体极差也是要大于 $K = 5$ 情形的。因此对比两种不同参数 K 值下的聚类特征，我们最后选择了类间差距更大的 $K = 4$ 作为 K-means 聚类方法的超参数，并以此建立聚类模型进行结果分析。

在 $K = 4$ 情形下，4 个不同的聚类上的男女比例图分别如下图38所示。

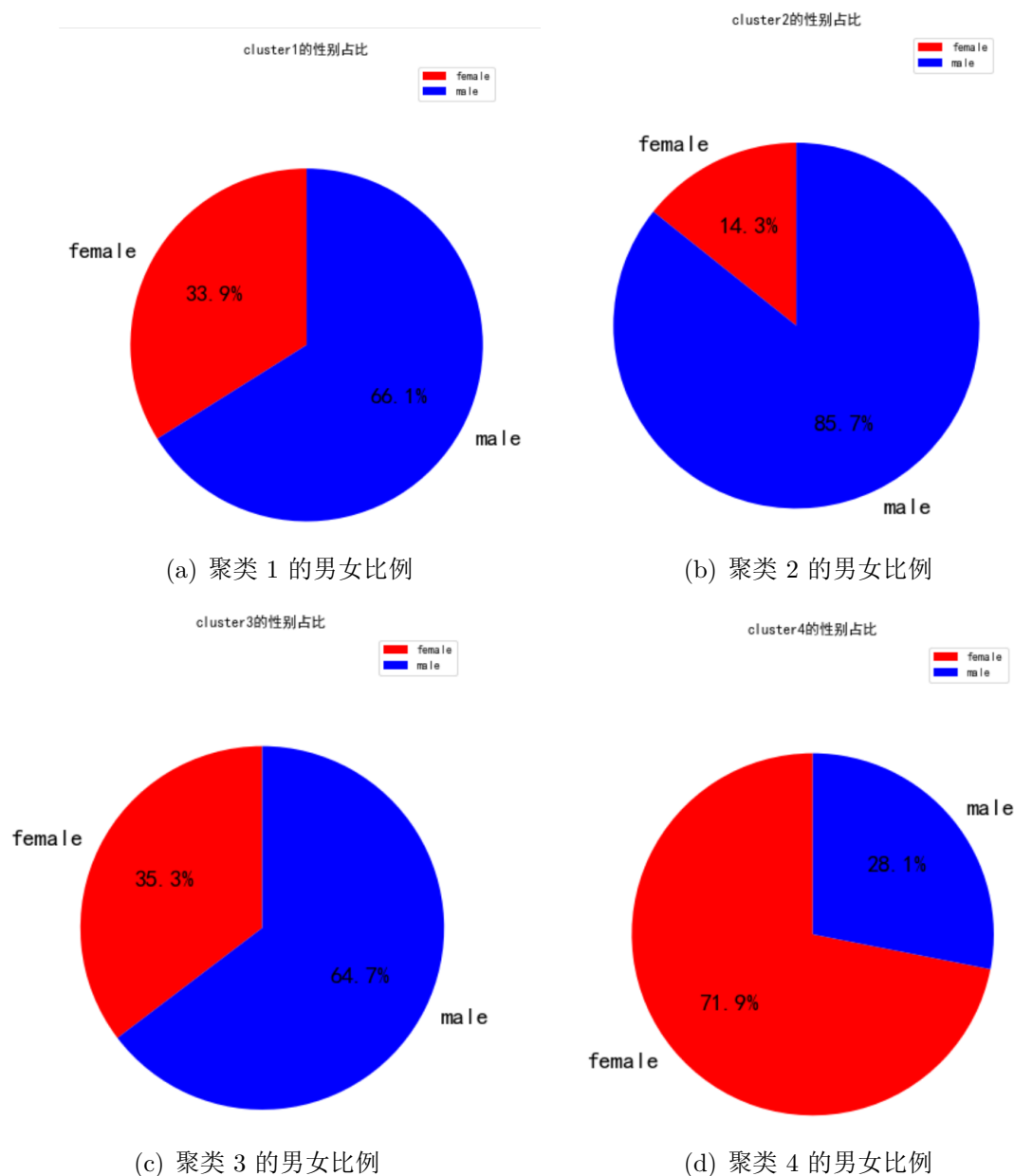


图 38: $K = 4$ 时不同聚类的男女比例饼图

横向对比 4 个比例饼图，可以发现 4 个聚类之间的性别差异还是比较明显的。当

然，这和我们获取的全部样本量当中男性样本要多于女性样本有一定的关系，但是这一因素不是决定性的。可以看到聚类 1、聚类 2 和聚类 3 的男性比例要明显多于女性，尤其是聚类 2，基本上男女比例呈 6:1 的趋势；而聚类 4 当中的这一趋势被明显地颠倒，女性比例要显著多于男性，男女比例呈 1:3 的趋势。要分析不同聚类之间男女比例差异的原因，需要重新审视表格6，从不同聚类的特征表现上寻找。

我们将不同聚类反映的特征均值以及男女比例结合起来，对 4 个不同的聚类进行了被调查者画像分析，由此显现出各个聚类在圈层参与度与 APP 使用上的偏好差异。分析结果如下：

1. 聚类 1 是平均休闲上网时间最短的群体，几乎对于所有的圈层基本上都不怎么偏好，对于大多数 APP 的使用频率也较低。这一类人属于互联网轻度参与群体，其生活受到互联网的影响也较小。
2. 聚类 2 是偏好圈层为动漫、体育/电竞、科技/数码与游戏，偏好 APP 为百度贴吧与 B 站的群体。该聚类中的人群追星较少，对豆瓣、微博、虎扑和小红书的使用较少，且该聚类以男性群体主导，是典型的“宅圈文化（ACG 宅男）”群体。
3. 聚类 3 是平均休闲上网时间最长的群体，对于各个圈层几乎都有所涉猎，对于网文圈层较为关注。APP 方面对于虎扑和 B 站使用较多，没有使用频率较低的 APP。可以认为聚类 3 是互联网重度用户群体。
4. 聚类 4 对于明星和综艺较为偏好，是较为典型的追星群体，而对于科技/数码与体育/电竞的关注度较低。APP 方面微博、知乎、豆瓣、小红书和短视频类 APP 使用频率均较高，且该聚类以女性全体为主，是较为典型的“追星综艺”型群体。

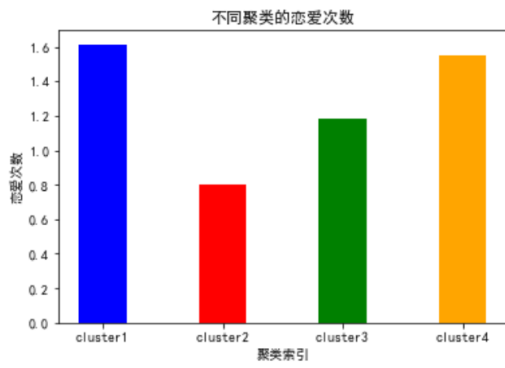
在分析完各个群体的特征后，我们依次给出各个聚类的婚育意愿结果图，以柱状图的形式给出。所有柱状图的纵轴是不同聚类的婚育观标签的均值。我们以柱状图的形式进行展示，是为了更好地起到聚类之间的对比效果。结果柱状图如下图39所示。

横向对比各个图像，我们可以得到这样一些结论：

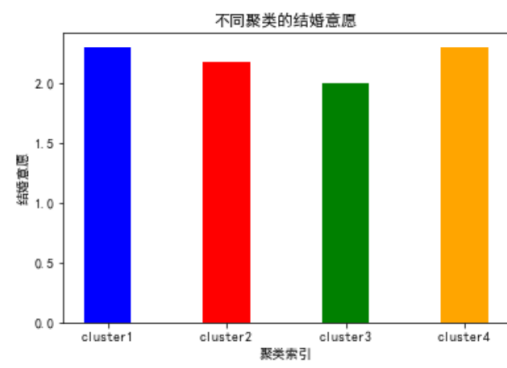
聚类 1 是休闲上网时间最短的群体，且各圈层参与度均不高，属于互联网轻度用户，而聚类 3 是休闲上网时间最长的群体，属于互联网重度用户。相比之下，聚类 1 的各项婚育观数值几乎都达到了 4 个聚类中的最高，而聚类 3 的各项婚育观数值基本上都是 4 个聚类中的最低（恋爱次数除外）。这在一定程度上反映随着休闲上网时间的增长与互联网参与度的增高，年轻人们的婚育意愿会出现下降趋势。

为进一步验证这一结论，我们绘制了不同上网时长区间下结婚意愿和生育意愿的百分比堆积条形图。图40和图41直观展示出，随着上网时长的增加，明确表示有结婚意愿的比例（图40中的深绿色部分）在减少，而表示不想要孩子的（图41中的黄色部分）越来越多，结论对男性和女性调查者均成立。

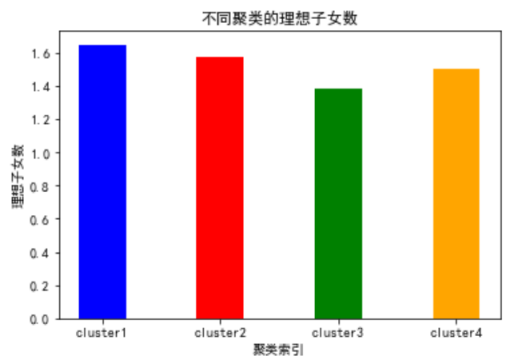
聚类 2 群体的恋爱次数较少，但是仍然具有较高的婚育意愿，说明尽管宅圈文化群体在恋爱方面较为萌新，但是仍然对于结婚生育抱有较高的向往和期望。



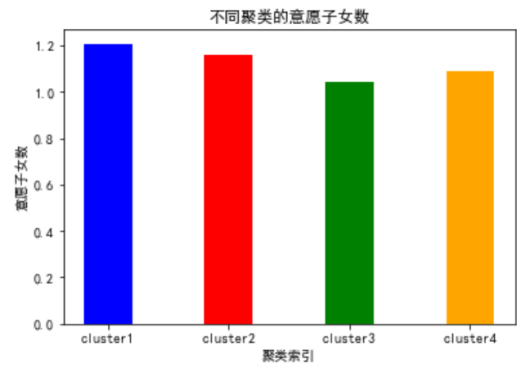
(a) 不同聚类的恋爱次数



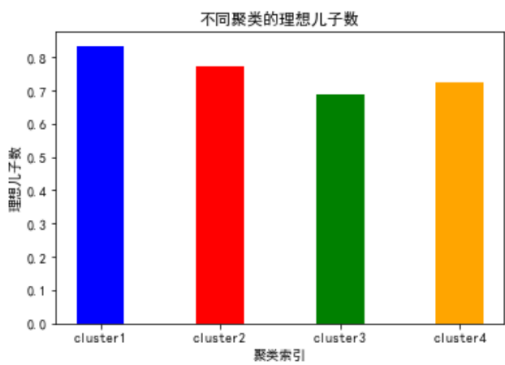
(b) 不同聚类的结婚意愿



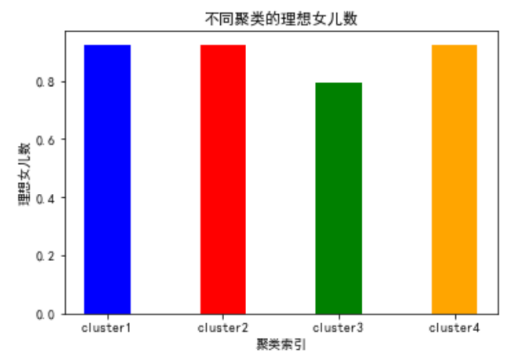
(c) 不同聚类的理想子女数



(d) 不同聚类的意愿子女数



(e) 不同聚类的理想生儿子数



(f) 不同聚类的理想生女儿数

图 39: $K = 4$ 时不同聚类的婚育观念柱状图

聚类 4 群体的恋爱次数和结婚意愿都较高，这在一定程度上说明关注明星、综艺，对于微博、小红书等 APP 使用频率较高的“追星”群体会有较强的结婚意愿。而相比之下，这类群体的生育意愿反而并没有特别强烈，各项数值没有聚类 2 的宅圈文化群体高。在该类群体上可以反映出结婚意愿与生育意愿并不是完全的正相关关系，在这里体现了一定的反差。

对比总体的各个聚类的理想生育儿子数以及理想生育女儿数，我们发现对于所有聚类群体而言，理想生育女儿数完全高于理想生育儿子数，这一定程度上说明上个世纪在生育方面较为频繁的“重男轻女”现象基本上已不复存在，当代青年人们还是更愿意生育女儿。

在生育儿子与生育女儿的对比上，我们发现聚类 2 群体和聚类 4 群体的差异最为明显，这两类群体对于生育女儿的倾向性明显更高。这也一定程度上反映了宅圈、动漫以及追星、综艺这些文化和圈层对于当代青年人的影响，也许是这些圈层中对于女性角色的刻画，加强了这些群体生育女儿的意愿，体现了一些正相关关系。

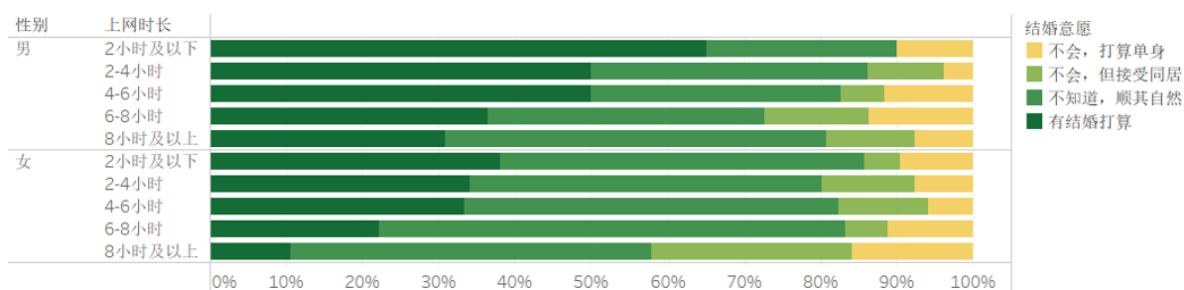


图 40: 上网时长深入分析——对结婚观的抑制作用

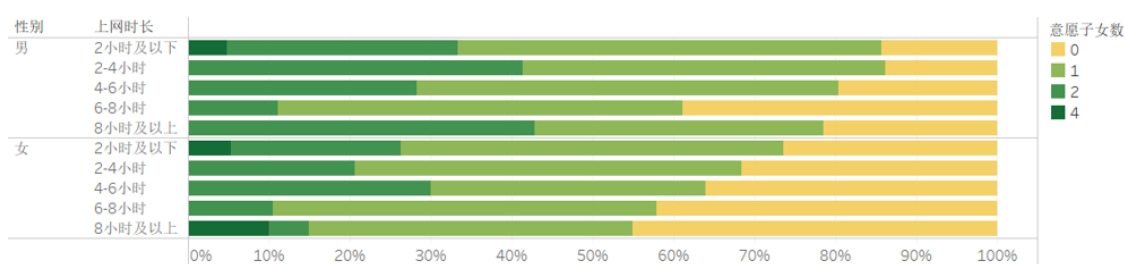


图 41: 上网时长深入分析——对生育观的抑制作用

通过聚类分析，我们得出了一些当代青年人的婚育观念与他们所参与的圈层以及使用的 APP 之间的定性关系，并得到了一些相关性的结论。除此之外，我们还对这二者之间的定量关系以及其他的外部因素对婚育观念的影响感兴趣，因此在下一部分，我们将采用回归分析的方法解决这一问题。

7.4 回归分析

接下来用回归分析方法探究被调查者的婚育观与他们所参与的圈层、使用的 APP 以及外部因素之间的定量关系。对于结婚观念的影响方面，我们采用逻辑回归模型进行

建模，而对于生育观念的影响方面，我们采用线性回归模型进行建模。这一点我们已经在方法框架介绍部分进行了原因解释，在这里不再赘述。

7.4.1 数据预处理

在数据预处理方面，这一部分和此前聚类分析的数据预处理有较多相似之处，其核心还是将文字描述类型的特征通过各种变换方式变为数值类型，并且通过 Python 当中的编码器进行顺序编码，体现出被调查者们在某一特征上的数值高低。而在这里，我们不仅仅需要研究圈层参与度与 APP 使用频率与婚育观念之间的关系，还需要研究其他各外部因素的影响，因此我们需要对各外部影响因素的特征预处理。

在这里，我们选择的主要是对婚育观可能有较大影响的一些外部影响因素，这些因素有互联网相关的，也有和互联网不相关的，主要包括“朋友圈秀恩爱晒娃”、“恋爱相亲育儿综艺”、“国家生育政策”、“专家教授发表看法”、“娱乐圈情感八卦”、“身边人经历”以及“相关艺术作品”。我们设置了“完全不影响”、“影响较小”、“影响较大”以及“影响极大”4个选项。和之前一样，我们以 0、1、2、3 分别对这 4 个选项进行顺序编码，0 代表“完全不影响”而 3 代表“影响极大”。

对于数据标签和其他特征的处理，方法和聚类分析中基本一致，故在这里不做赘述。

7.4.2 分层抽样说明

在之前进行聚类分析时，我们将所有的 545 个样本都进行了建模。这主要是由于我们此前做的是定性分析，需要比较各个聚类之间的婚育观念差异情况，因而我们用上了全部样本。

但是在这部分进行统计建模时，我们就需要考虑性别之间的样本不平衡问题。这一方面是因为在这里进行定量分析时，我们有意识地将男女性别分开进行分析，希望可以起到男女之间进行对比的效果，另一方面我们在计算对结婚观念的贡献因子时用到了逻辑回归模型，这是一个典型的有监督分类模型，因此需要考虑性别之间的样本不平衡问题，尽可能在样本平衡情况下计算各个特征及因素对婚育观念的影响贡献因子，力求做到参数估计的准确。

我们获得的调查问卷数据共为 545 份，其中 336 份为男性，209 份为女性。我们以男性和女性进行分层，各自抽取了 200 份样本，共计 400 份样本进行模型的建立。尽管采用分层抽样的方法，会损失 136 份男性问卷的信息，但是这是为了保证样本平衡的权宜之计。我们会将 400 份样本作为总体进行建模，同时再将样本的男女性别分开再进行一次建模，而每个特征的贡献因子值我们也会将总体与男女性别的建模结果放在一起展示，方便进行比对。

在这里，我们着重于分析各种因素对于结婚观念和生育观念带来的正向或负面影响，因而我们的标签选用“结婚意愿”、“理想子女数”与“意愿子女数”，就不对“恋爱次数”、“理想生儿子数”、“理想生女儿数”这些相对较为细致的标签进行分析。这样做可以在减小分析复杂度的同时，得出尽可能贴合我们调查主体的结论。

7.4.3 圈层影响

首先关注圈层参与度对于婚育观念的影响。在结婚观念方面，我们采用 sklearn 库当中的逻辑回归方法进行参数估计，设定参数为多分类器，采用随机梯度下降算法，计算每个特征对应系数的参数估计作为该特征对于结婚观念的贡献因子。在生育观念方面，我们采用 sklearn 库当中的线性回归模型进行建模，得到各个特征系数的参数估计。具体的圈层参与度对婚育意愿的回归贡献度如下表8所示。表中的回归贡献度若为正值表示该特征对于标签具有正相关贡献，相应地若为负值则表示该特征对于标签呈负相关贡献。表中的列的回归贡献度最大值我们用红色字体标出，而最小值我们用蓝色字体标出。

表 8: 圈层参与度对婚育意愿的回归贡献度

圈层	总体			男性			女性		
	结婚	理想子女	意愿子女	结婚	理想子女	意愿子女	结婚	理想子女	意愿子女
	logit	OLS	OLS	logit	OLS	OLS	logit	OLS	OLS
动漫	-0.201	-0.062	-0.073	0.117	-0.052	-0.001	-0.669	-0.071	-0.106
综艺	0.279	-0.030	-0.004	-0.227	-0.195	-0.164	1.085	0.136	0.166
体育/电竞	0.981	0.151	0.146	1.061	0.023	0.094	-0.610	0.139	0.205
明星	-0.031	-0.080	-0.012	1.753	0.160	0.057	-0.221	-0.089	-0.001
科技/数码	-0.196	0.105	0.135	-0.030	0.087	0.117	-0.487	0.079	0.040
网文	-0.122	-0.009	-0.021	0.569	-0.020	0.093	-0.545	-0.002	-0.086
游戏	-0.289	-0.098	-0.061	-0.870	-0.098	-0.118	0.653	-0.149	0.010

观察该表，我们可以得到如下相关性结论：

首先，各亚文化圈层的参与情况确实会影响当代年轻人的婚育观。从表中数据不难看出，各个亚文化圈层对于结婚观以及生育观都或多或少地具有一定的正向或负面影响，基本上没有贡献回归度绝对值为 0 的情况发生。

其次，关注体育和电竞可以增强婚育意愿，这对男女都较为适用；但关注动漫、网文和游戏会降低生育意愿；其中游戏对于男性，动漫对于女性的婚育意愿的抑制都起到了较为明显的作用。对于这一现象，我们给出了如下可能的解释：一方面，可能这和体育、电竞这两个圈层比较贴合竞技这一已发展多年的传统有关，因此关注这两个圈层的群体在婚育观方面也较为传统，他们会认为传宗接代的传统是应该去遵守的；另一方面，动漫和网文以及游戏更多反映了当代青年人们对于自己生活以及未来对象等多方面的预期，而关注这些圈层的群体们也许正是会将这一预期与真实生活进行比较，因而某种程度上降低了婚育意愿。

接着，追星对于男女之间的婚育观念的影响效应存在显著的性别差异。追星对于男性的婚育意愿有显著的促进作用，而会导致女性的婚育意愿存在不同程度的降低。我们认为，一方面对于女性而言，追星的原因更多是因为自己所追的明星比较完美，某种程

度上现实生活中所遇到的对象可能反而达不到她们的预期，因此追星对于女性的婚育意愿呈负相关关系；而另一方面对于男性而言，追星更多是他们对于现实生活的延续，某种程度上他们正是因为受到偶像的鼓舞，所以才希望在现实生活中努力找到符合他们偶像预期的人，因而追星更多成为了他们的动力来源，这就体现了追星对于男性的婚育意愿的正向作用。

为进一步探究以上相关性背后可能的原因，我们试图进行一些深入分析。我们注意到动漫、网文圈均有很强内容生产属性，这些内容大多是对现实生活的抽离或理想化，构建了具有某种反差性的平行空间，因此或许会削弱年轻人对现实空间的需求与期望。结合调查问卷中之前提到的抑制婚育观念的可能影响因素，我们分性别做出了一下的两个交叉分析图，如下图42与43所示。

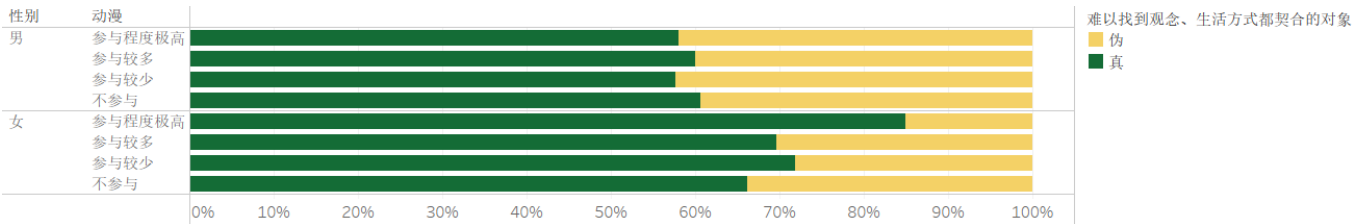


图 42: 动漫圈层深入分析——对结婚观的抑制作用

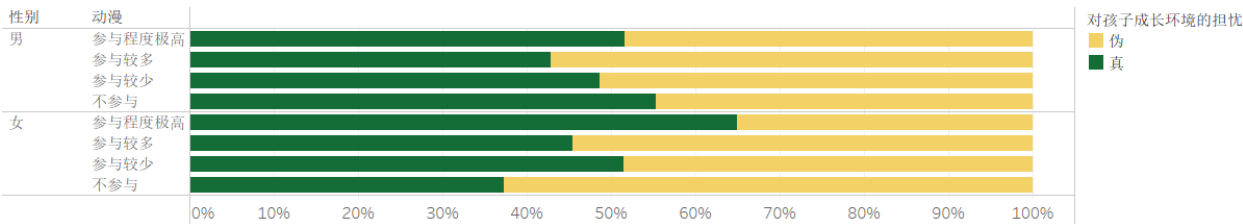


图 43: 动漫圈层深入分析——对生育观的抑制作用

从图上我们可以看出，动漫圈层会对结婚观念具有抑制作用的一大原因在于参与动漫亚文化圈层的群体“难以找到观念、生活方式都契合的对象”。在参与动漫圈层的群体中，女性受到这一因素的影响更为明显。正如上文所述，动漫更多反映了当代青年人们对于自己生活以及恋爱对象等多方面的预期。受到动漫这一圈层影响的群体更多的是受到了这种预期的影响，因而产生了对于“结婚”的抑制导向。而动漫圈层对于生育观念具有抑制作用的一大原因在于参与动漫圈层的群体对于未来孩子成长环境十分担忧。事实上这也很好理解，因为动漫很多时候是对于现实生活的刻画与反映，而一些现实场景例如校园一直是动漫中被刻画的一个重要主体，一些校园题材方面的热点，例如校园霸凌、人际关系等等都成为了动漫主体重点刻画的对象。因此参与动漫圈层的群体会对未来孩子的成长环境感到担忧，这是十分正常的现象。

需要指出的是，我们所进行的大量原因分析内容是基于日常生活经验等做出的判断，没有直接的调查数据支持，仅代表我们尝试对调查结果中发现的相关性所做出的解释。我们同样努力试图从调查数据中获取部分证据，以支持和解释互联网圈层参与度对

于婚育观念的影响，而非简单地诉诸主观的理论解释。以上的对比分析就是一个例子。事实上，在我们的问卷设计中包含许多对婚育问题看法的调查问题，对这些结果的解读可以是五花八门的。在这里我们的原则是尽可能采取直观的对比分析方法，仅展示其中相对显著的几个例子，一如前文中分析性别与婚育观的相关性时所采取的方法。这是因为，在所展示的相关性不代表因果性的前提下，进行对原因的探究性分析需要特别小心，这里的系统性错误对读者的误导可能是更深层的。我们的调查设计有十几个互联网圈层和几十个促进/抑制婚育相关因素的选项，从统计意义上来说，即使其背后没有任何关系，平均而言也会发现几个显著的联系，况且我们总是可以对其做出解读，这样的误导是我们希望尽可能避免的。因此，本节内我们在原因探究上使用的都是对比等直观方法，对于各圈层与影响婚育观原因的相关性分析将在下一节讲述。

从表格中我们还能看到其他一些圈层或多或少地对于婚育观念的正相关或是负相关关系，在这里囿于篇幅就不做赘述了。总之，亚文化圈层对于当代年轻人们的婚育观念的影响还是比较显著的。

7.4.4 APP 使用影响

接着我们关注各种社交类型 APP 的使用频率对于当代年轻人们婚育观念的影响。我们采用的各种 APP 特征在前面已经介绍过，而建模与分析方法也与之前圈层参与度影响的一致。具体的 APP 使用频率对婚育意愿的回归贡献度如下表9所示。

表 9: APP 对婚育意愿的回归贡献度

APP	总体			男性			女性		
	结婚	理想子女	意愿子女	结婚	理想子女	意愿子女	结婚	理想子女	意愿子女
	logit	OLS	OLS	logit	OLS	OLS	logit	OLS	OLS
微博	-0.589	-0.151	-0.157	-0.345	-0.006	-0.069	-0.585	-0.134	-0.163
知乎	0.844	0.017	0.031	0.789	0.008	0.068	0.520	0.006	-0.040
豆瓣	0.221	-0.018	-0.014	0.159	-0.019	-0.095	0.043	-0.046	-0.034
虎扑	1.412	0.252	0.155	0.739	0.168	-0.006	0.847	0.297	0.350
百度贴吧	-0.437	-0.016	0.028	-0.684	-0.037	0.022	0.476	0.010	0.083
小红书	0.110	0.029	0.041	-0.063	-0.106	-0.008	0.827	0.263	0.198
B 站	0.237	0.084	0.059	-0.173	0.076	0.050	0.060	0.022	0.054
短视频	0.597	0.076	0.101	0.896	0.054	0.133	-0.100	0.053	0.030

观察该表，我们可以得出以下结论：

首先，从总体上看，APP 使用频率也会在一定程度上影响当代年轻人的婚育观。相对而言，这种影响对于结婚意愿的影响要大于生育意愿的影响。从表格9数据可以看出，结婚意愿一栏的回归贡献度绝对值要大于代表生育意愿的那两栏。

其次，基于总体而言，微博和百度贴吧对于婚育意愿的消极作用较为明显。而基于性别而言，微博对于女性，百度贴吧对于男性的婚育意愿的降低效果均十分明显。我们认为这一现象和这两个平台上本身具有的较强言论争议性有关。微博和百度贴吧这两个APP，总体而言其上探讨的话题和言论相对争议性较为明显，不同性别、不同群体的人们更容易形成言论上“针锋相对”的态势，各种立场的碰撞使得当代年轻人受到的影响也较为明显。除此之外，这也和百度贴吧的男性使用者相对更多，而微博的女性使用者相对更多有着一定的关系。这一点从之前聚类得到的结果表6也可以体现出来，以男性为主的聚类 2 群体贴吧特征贡献度为 1.07，而以女性为主的聚类 4 群体微博特征贡献度为 1.99，均为 4 个聚类中的最大值。我们认为这不同群体分别对这两个 APP 的性别偏好也一定程度上导致了这一现象的发生。

接着，我们发现使用虎扑频率较高的用户会有较高的婚育意愿。从表格上可以看出，虎扑一行对所有标签的贡献度基本都是正值。我们认为这一点可以和之前圈层参与度得出的“关注体育和电竞可以增强婚育意愿”的结论结合起来。因为虎扑上的社区和论坛，很多都是和体育以及电竞有关的话题。因此我们认为这里的虎扑 APP 对于婚育意愿的正相关关系与“虎扑 APP 与体育及电竞圈层联系紧密”的这一事实有一定的相关性。

除此之外，我们还发现很有意思的一点是，小红书对于男性和女性之间的婚育意愿影响效果差异最明显。小红书对于女性的婚育意愿有显著的促进作用，而对男性而言则有一定的抑制作用。我们认为这 and 不同性别对于小红书 APP 的使用偏好以及看法有一定的关系。小红书对于女性而言，其主题更偏向于“精致生活”，其上一些内容对女性而言可能更多反映生活的美好层面，这在一定程度上会对女性的婚育意愿有着促进效果。

开始时我们并不知道为什么小红书的使用会对男性的婚育意愿存在抑制作用，我们猜测也许是因为小红书上的一些内容相对“现充”，这可能会一定程度上给男性使用者带来一定的压力。于是我们对小红书 APP 的使用者们认为的可能抑制婚育意愿的因素进行了调查与深入研究，得到的相关结果如下图44与45所示。

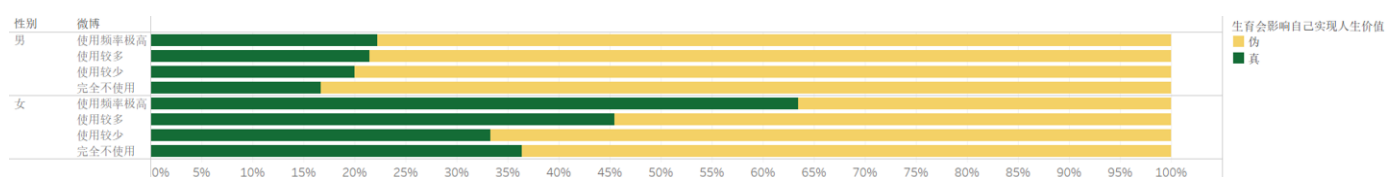


图 44: 小红书 APP 深入分析——对结婚观的影响

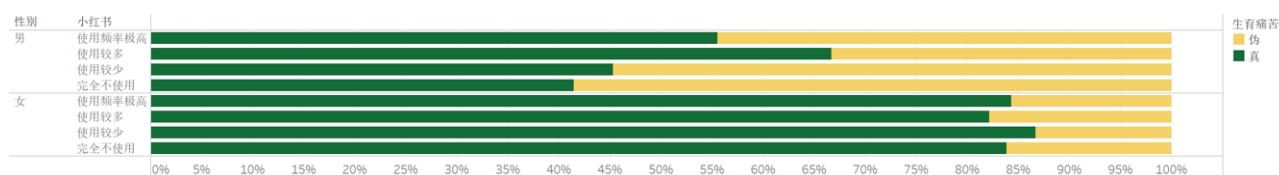


图 45: 小红书 APP 深入分析——对生育观的影响

从图上可以看出，小红书的女性使用者更多认为抑制当代年轻人婚育意愿的因素很多是“生育会影响自己实现人生价值”，这一点在男性使用者方面并不明显。但是在“生育痛苦”这一抑制因素上，无论是男性还是女性，小红书的使用者们都表示了一定程度的认同。因此我们认为，小红书会降低男性的婚育意愿是由于男性使用者们也意识到了“生育痛苦”有关，而不单单是这款 APP 可能给他们带来的生活压力。这一点也说明了对于“APP 使用频率对婚育观念造成的影响”这一议题，有必要进行深入的内在分析。

7.4.5 其他外部因素

最后我们关注其他外部因素对于当代年轻人们婚育观念的影响。这里的建模与分析方法也与之前的一致。具体的回归贡献度结果如下表10所示。

表 10: 外部影响对婚育意愿的回归贡献度

外部影响因子	总体			男性			女性		
	结婚	理想子女	意愿子女	结婚	理想子女	意愿子女	结婚	理想子女	意愿子女
	logit	OLS	OLS	logit	OLS	OLS	logit	OLS	OLS
朋友圈恩爱晒娃	2.046	0.345	0.245	1.535	0.206	0.233	0.961	0.454	0.153
恋爱育儿综艺	1.198	-0.026	0.001	1.192	-0.237	-0.130	1.408	0.204	0.128
国家生育政策	0.277	0.002	-0.003	0.083	0.142	0.210	0.581	-0.122	-0.123
专家教授看法	0.526	0.121	0.160	1.317	0.201	0.217	0.361	0.102	0.237
娱乐圈情感八卦	-0.879	-0.241	-0.070	0.149	-0.051	-0.061	-0.408	0.067	0.035
身边人经历	-0.362	-0.007	-0.045	0.630	0.004	-0.100	-1.001	0.067	0.035
相关艺术作品	-0.475	0.008	0.021	-0.832	0.002	0.007	-0.504	0.017	-0.017

观察表格10，我们可以得出以下结论：

首先，这些外部因素确实会一定程度上影响当代年轻人的婚育意愿。可以看到，表中很少出现回归贡献度绝对值接近于 0 的情况，这也在一定程度上反映了我们的这项研究是有意义的。

其次，基于总体而言，朋友圈秀恩爱晒娃对于当代年轻人的婚育意愿的促进作用最为显著；相反，娱乐圈情感八卦对于当代年轻人的婚育意愿的抑制作用最大。这一点还是比较符合实际的，当代年轻人们受到朋友圈当中秀恩爱内容的“耳濡目染”，确实会一定程度上激发他们对于结婚生育的美好向往。而近几年由于各种娱乐圈事件发生层出不穷，媒体相关报道比比皆是，这也变相导致了暴露于娱乐圈情感八卦当中的年轻人确实会抑制他们的婚育意愿。

接着，对于男性而言，他们更易受到朋友圈秀恩爱晒娃对于结婚生育的正向影响；而相关艺术作品对于男性的结婚意愿的消极作用最大。除此之外，专家教授发表看法也会对男性的婚育意愿起到一定的促进作用。相关的艺术作品可能更多反映了“生存压力”、“生活压力”等主题，这在一定程度上抑制了他们的结婚意愿。在专家教授发表看

法方面，可能是因为本身具有一定生育意愿的群体更关注这一问题，因而导致了这样的结果。

而对于女性而言，她们更易受到恋爱相亲育儿综艺节目对于结婚生育的正向影响；相反，身边人的经历会很大程度上抑制她们结婚的念头，而国家生育政策会抑制她们的生育意愿。在相亲育儿综艺节目方面，可能女性相对地更为感性一些，因此会受到这些节目对于她们婚育观念方面的鼓舞；也正因如此，身边人的经历对她们而言也是十分宝贵的，她们某种程度上也很容易受到这方面的影响。而国家生育政策，个人看来本身对于女性的影响就更大一些。女性的生育确实会影响生活和工作的方方面面，因此国家政策导向确实对她们会有很大的影响，这也可以反向激励国家政府提出真正有利于人民发展和女性群体的生育政策。

7.5 原因探究

上一节中提到，我们尝试基于经验和部分直观数据，对回归分析结果中互联网圈层与婚育意愿间的相关性做出了解读。对于调查问卷中涉及的十几个互联网圈层参与度，和数十个促进或抑制婚育的相关因素，我们也进行了统计检验，已试图发现一部分可能的婚育观影响机制。以下将具体展示这部分内容。

7.5.1 数据预处理

首先进行了分层抽样，与上一节的处理方式相同，男性和女性各自抽取了 200 份样本。在变量处理上有所不同，因变量方面选取了“请问您认为或您见到的促进结婚的因素有哪些”“请问您认为或您见到的抑制结婚的因素有哪些”“请问您见到的或认为的阻碍生育的因素有哪些”“请问您认为的可以提高生育意愿的政策有哪些”4 个问题，将这四个所选题按选项展开进行 one-hot 编码，共计 38 个选项作为 0-1 变量。对自变量即亚文化圈层参与度、APP 使用程度，同样作为有序分类变量进行编码，不同的是编码数值为 1、2、3、4，这是为了适应之后检验的需要。

7.5.2 统计检验设计

检验对象为二分类变量和有序分类变量间的两两相关性检验，适用 Cochran-Armitage 检验。设 y_i 为所关注的事件的发生频数， n_i 为各等级的样本量， x_i 为各等级的赋值，检验统计量为：

$$z = \frac{\sum_{i=1}^k y_i (x_i - \bar{x})}{\sqrt{\bar{p}\bar{q} \left[\sum_{i=1}^k n_i (x_i - \bar{x})^2 \right]}} \quad (1)$$

其中

$$N = \sum_{i=1}^k n_i \quad \bar{p} = \frac{1}{N} \sum_{i=1}^k y_i \quad \bar{q} = 1 - \bar{p} \quad \bar{x} = \frac{1}{N} \sum_{i=1}^k n_i x_i \quad (2)$$

对 z 统计量进行标准正态分布的双边检验求得 p 值。

两两检验时面临着比较严重的多重检验问题，检验的规模为 $38 \times 15 = 570$ ，因此以下仅展示 p 值小于 0.001 的部分。

7.5.3 检验结果和解读

对男性样本的检验结果发现，使用豆瓣频率与选择“难以找到观念、生活方式都契合的对象”作为抑制结婚的原因存在负相关， p 值为 0.000764；使用小红书频率与选择“婚姻可以带来物质经济保障”作为促进结婚的原因存在正相关， p 值为 0.000321。通常认为豆瓣和小红书都是女性参与度相对较高的社区，也许豆瓣男性用户更容易与女性群体精神上产生共鸣，从而没那么担心与另一半观念的契合，这也解释了回归分析中发现的男性使用豆瓣与结婚意愿正相关而与生育意愿负相关。而小红书用户对婚姻中物质保障的在意，可能与回归分析中发现的男性使用小红书与生育意愿负相关这一事实存在一定联系。

对女性样本的检验结果发现，使用抖音等短视频平台的频率与选择“父母长辈对子女伴侣要求较高”作为抑制婚姻的原因，以及选择“降低买房成本（例如对多孩家庭给予买房上的优惠）”作为促进生育的政策均存在正相关， p 值分别为 0.000917 和 0.000517。这一定程度上反应了短视频用户的关注点，也符合回归分析中发现的女性使用短视频平台与结婚意愿的负相关。

最后仍需强调的是，以上结果仍有可能稳健性不足，对某些小众圈层的参与者而言，在一些选项上选择的人数过少，从而对抽样结果比较敏感。主要原因还是时间和精力限制，男女各 200 人的样本对这样较大规模的相关性分析显得有所不足，容易得到虚假的结论。我们希望后续能够通过爬虫等方式，在各个互联网社区进一步检验调查数据中的发现，读者若有兴趣也可自行尝试。

8 报告总结

本报告以全国大学生为研究对象，通过问卷调查进行研究，信效度分析表示问卷结果具有良好的可参考性。通过对问卷调查得到的数据分析，总结大学生婚育意愿的现状为：超过半数没有明确的结婚意愿；理想结婚年龄集中在 26-30 岁；婚姻观念比较开放；意愿子女数比理想子女数平均少一个孩子；多数的理想子女情况为一男一女，且总体偏好女儿。大学生婚育观的现状为：在意婚姻和生育带来的精神价值，生活成本、女性生育痛苦与就业歧视是压抑大学生结婚和生育的主要动因。

通过对比男女大学生的婚育观，本文发现二者的婚育观念存在明显差异：第一，从婚育意愿或者恋爱意愿来看，相较于女性，男大学生往往表现得更加积极，具有更加强烈的结婚意愿。第二，谈到影响婚育的因素时，男性更多会考虑一些婚育成本、工作压力等经济因素，而女性倾向于一些诸如带薪产假、个人未来发展之类的人文因素。此外，女性更喜欢从恋爱育儿综艺、国家生育政策、专家教授看法等外界渠道接受一些有关婚育的信息。因此，在需制定提升婚育意愿的政策时，不能够忽略性别差异。

关于互联网参与与婚育意愿，本文主要针对“互联网圈层参与”、“APP 使用频率”以及“其他互联网外部影响”三个部分，着重于这三个部分对于当代年轻人婚育意愿可能产生的影响进行了研究。总体而言，婚育意愿与互联网使用时长呈现负相关，具体到各细分因素则各有不同影响，如与动漫/网文的参与度负相关、与体育/电竞的参与度正相关、与微博和百度贴吧的使用频率负相关等等。需要注意的是，以上都是基于相关性而非因果性质的推断。本文尝试探究了其可能的形成原因，发现了样本中互联网参与和关心的婚育议题存在相关性，例如豆瓣用户更在意和另一半精神上的契合，小红书用户更在意生育对自我人生价值实现的影响等等。总体而言，这三个部分确实会对当代年轻人的婚育意愿产生或多或少的影响，而如何进一步地实现社会结婚生育政策的稳定，缓解人口老龄化的严峻问题，还需要国家、社会和个人的共同努力。

参考文献

- [1] 李婷, 郑叶昕, and 闫誉腾. 中国的婚姻和生育去制度化了吗? *妇女研究论丛*, 2022. 1.1.1
- [2] 王飞. 当代青年的婚恋观及其影响因素分析——基于 17~ 34 岁年龄段的青年调查数据. *中国青年研究*, (7):73–76, 2015. 1.4
- [3] Ron Lesthaeghe. The second demographic transition: A concise overview of its development. *Proceedings of the National Academy of Sciences*, 111(51):18112–18115, 2014. 1.4
- [4] 贾志科 and 王思嘉. 当代青年婚恋研究述评与展望 (1985-2018). *青年探索*, (6):5–16, 2019. 1.4
- [5] 梁璐. 大学生对后现代婚姻形态的认同度研究——对中国青年政治学院 653 名大学生的问卷调查分析. *青年探索*, (1):21–25, 2013. 1.4
- [6] 姜永志. 当代大学生婚恋价值观研究现状分析与展望. *内蒙古师范大学学报: 哲学社会科学版*, 42(6):159–163, 2013. 1.4
- [7] 朱洪翠 and 刘冰. 新时代大学生生育观研究. *黑龙江高教研究*, 38(4):33–39, 2020. 1.4
- [8] 李宝礼 and 邵帅. 不安居, 则不乐育: 住房状况与青年人口生育意愿研究. *中国青年研究*, 3:53–62, 2022. 1.4
- [9] 风笑天 and 李芬. 再生一个? 城市一孩育龄人群的年龄结构与生育意愿. *思想战线*, 6(1):88, 2016. 1.4