

生育问题网络关注度的时空演变及影响因素研究

石莹莹, 崔炳伟

北方工业大学理学院, 北京

收稿日期: 2023年7月11日; 录用日期: 2023年8月1日; 发布日期: 2023年8月15日

摘要

生育问题作为当前社会亟需解决的问题, 国家政府出台了一系列相应政策措施, 但效果并不理想。本文在研究过程中发现生育网络关注度与生育率显著正相关, 生育网络关注度在一定程度上可以反映出生育率的变化规律。故本文利用百度指数平台的搜索指数和社会经济数据, 综合运用探索性空间数据分析、FE面板数据模型等研究方法, 探究生育问题的网络关注度时空演变特征及影响因素, 深度挖掘生育问题影响因素本质, 并对生育问题的改善提出切实有效的建议。

关键词

网络关注度, 空间自相关, 传统Hausman检验

A Study on the Spatial and Temporal Evolution of Online Attention to Fertility Issues and the Influencing Factors

Yingying Shi, Bingwei Cui

College of Science, Northern University of Technology, Beijing

Received: Jul. 11th, 2023; accepted: Aug. 1st, 2023; published: Aug. 15th, 2023

Abstract

As an urgent social problem, the national government has introduced a series of corresponding policies and measures, but the effect is not satisfactory. In the course of this paper, we found that fertility network attention is significantly and positively correlated with fertility rate, and fertility

network attention can reflect the changing pattern of fertility rate to a certain extent. Therefore, this paper uses the search index and socio-economic data of Baidu index platform, integrates exploratory spatial data analysis, FE panel data model and other research methods to explore the spatial and temporal evolution characteristics and influencing factors of the network concern of fertility issue, deeply explores the essence of the influencing factors of fertility issue, and puts forward practical and effective suggestions for the improvement of fertility issue.

Keywords

Internet Attention, Spatial Autocorrelation, Traditional Hausman Test

Copyright © 2023 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

国家统计局在 2023 年 1 月 17 日公布最新人口数据中, 2022 年我国人口出生率为 6.77‰, 人口自然增长率-0.60‰, 这是我国自 1960 年后人口自然增长率首次呈现负数。生育问题成为我国关注的重点, 我国也相继出台了鼓励生育二孩、三孩等相关政策应对生育问题, 并采取了多方面措施鼓励生育, 但其效果甚微。另一方面, 数字化时代中互联网的磅礴发展与人们的生活息息相关。2023 年 3 月 2 日, CNNIC 官网第 51 次互联网发展报告显示: 截至 2022 年 12 月, 我国网民规模达 10.67 亿, 较 2021 年 12 月增长 3549 万, 互联网普及率达 75.6%。互联网的迅猛发展中, 每个人都既可以是传播者, 也可以是受众, 逐渐形成了“网络社会”。

在网络社会中, 人们足不出户也能知晓天下事, 在这样的环境下, 大众对于关注的话题, 通常会在网络上进行搜索以及发表自己的观点, 人们也对关心的社会问题有极大的参与感。过高的网络关注度对很多现实问题具有推动作用。网络关注度是了解民众意愿的重要途径, 也成为解决实际问题的重要参考。近年来, 生育问题便是广受关注的重点之一, 与生育问题相关的话题也频频出现在各大社交软件的热门词里, 并引起网络上的广泛关注与传播, 民众从不同角度对生育问题进行搜索与意见发表。

梳理生育问题的网络关注情况, 了解民众对生育问题的实际看法, 研究网络关注度与生育之间的关系能够丰富生育问题的学术研究。对人口增长的政策落实和提升民众满意度有重要的参考意义。本文将主要探讨对生育问题网络关注度的时空演变过程是怎样的? 具有哪些重要的影响因素?

1.2. 文献综述

穆光宗等人[1] (2021)发现中国早在 2000 年就已步入内生性低生育阶段, 政策性低生育阶段远比人们想象和预料得短促。刘伟杰等人[2] (2021)系统梳理了学界关于青年婚姻观与生育观的研究成果, 分析了其中的薄弱环节, 根据不同地域、不同职业、不同文化、不同户口下的不同群体提出针对性的解决措施。景鹏等人[3] (2023)发现当前经济社会条件下, 伴随生育支持力度的加大, 合理调整生育支持结构可以实现生育水平提升与经济持续增长的“双赢”。田悦芳和孔广用[4] (2022)从纪录片《奇妙的蛋生》中发现其在女性视角凝视下的主题愿景是拥有平等的机会, 增强女生生育问题的“可见性”。

在关于网络与生育的相关文献中, 刘璐[5] (2022)借助心理学和传播学等理论, 把适龄未婚育的女性

群体“恐育”问题作为切入点,探讨了社交媒体环境下应激心理对女性“恐育”传播行为的作用和影响,发现社交媒体环境中“恐育”行为的传播是女性群体对生育风险问题的应激反应,进而造成集体避险式情绪表达,且此传播行为在一定程度上引发生育风险的过度防范。乔嘉如[6] (2019)发现网络社交媒体的使用增强了女性对生育风险的“在场感”,网络媒体使用在生育风险对生育意愿上的影响起到放大和缩减的双重作用。高颀源[7] (2022)爬取社交平台的数据进行数据挖掘,发现生育舆情近十年总体呈现负面生育态度,其中二孩政策对生育态度有积极影响,生育舆情中限制网民生育态度的决策因素主要是:经济压力、保障压力、养育压力和身体因素。黄君洁和邱婷[8] (2022)通过问卷调查发现在社交网络上浏览亲子信息会对个人的生育意愿产生影响。

在搜索与生育问题相关的话题时发现,生育问题与很多疾病挂钩。廖晓妹等人[9] (2020)采用单因素和多因素 Logistic 回归模型评估社会人口统计学因素、肿瘤因素与生育相关问题关注度之间的关系,发现治疗前未生育的患者对生育相关问题关注度高。周旭平等[10] (2023)通过单因素和多因素 Logistic 回归模型,发现在三孩生育政策推行 1 年期间,产后出血的产前高危因素呈现新的分布特点。董翠丽等人[11] (2023)通过方差分析、t 检验、Pearson 相关分析及多元线性回归分析了育龄期女性乳腺癌患者生育忧虑现状及其影响因素,为临床改善女性癌症患者的生育忧虑水平提供依据。朱艺等人[12] (2023)采用 Harman 单因子检验法进行共同方法偏差检验,探讨了产后抑郁症状和育儿自我效能之间的现状及作用路径,分析了影响产后抑郁症状和育儿自我效能的人口学因素。

关于生育问题的影响因素研究中,陈钰晓等人[13] (2023)根据 CFPS 数据,采用基准分析、异质性分析、稳健性分析发现家庭养老照护会显著降低育龄女性生育二孩的概率。彭争呈等人[14] (2023)基于 2015 年全国 1% 人口抽样调查数据,研究发现单独二孩政策并未显著降低政策目标女性的劳动参与,但要实现女性生育与就业的协调,需要有可替代的家庭和公共儿童照护资源做支撑。陆温婷[15] (2023)发现城市化是最好的“避孕药”,养老观念在城市化与生育意愿的关系中起调节作用,“自我养老”观念能够缓解城市化对生育意愿的负向影响。姚懿轩和王百玲[16] (2019)发现“全面二孩”政策下农村育龄女性生育中存在的问题包括家庭经济状况及照看二孩的时间和精力问题、女性自身和胎儿健康状况的问题、就业与自身发展问题、农村生育保险制度及托育服务不完善。

通过对生育问题相关文献的查阅,发现关于生育问题及其影响因素已经有较为丰硕的研究成果,关于网络关注对生育问题的影响,主要集中在女性的生育意愿上,没有将网络关注度和整体生育问题之间的关系进行深入挖掘,本文将重点考察网络关注度对生育问题的影响上。

2. 数据来源与研究方法

2.1. 数据收集与处理

百度作为全球领先的中文搜索引擎,每天响应来自 100 余个国家和地区的数十亿次搜索请求,是网民获取中文信息的最主要入口,也是国人最为熟识的信息检索方式。百度指数是百度旗下的服务平台,主要分享百度网民的行为数据。平台的搜索指数以海量网民在百度的搜索量为数据基础,以关键词为统计对象。庞大的用户基数和国民认可度都使得百度指数平台的搜索指数可作为网民关注度的数据化体现。本文选择百度指数作为数据来源。

本文选用“直接取词法”和“范围取词法”[17]选取关键词,具体步骤如下:以“生育”作为直接关键词,利用百度指数需求图谱及站长工具(<http://mtool.chinaz.com/>)中的关键词挖掘功能对“生育”相关关键词进行补充,并在百度指数上检测收录情况。通过关键词筛选,最终选取“生育 + 生育津贴 + 生育保险 + 生育政策”为组合关键词。利用 python 工具在百度指数平台分别爬取 2013 年 1 月~2023 年 3 月

全国尺度及我国的 31 个省区市省级尺度的逐日关键词搜索指数, 作为研究生育网络关注度的源数据。(注: 百度指数平台的搜索指数并不是网民实际搜索频次, 而是根据各个关键词在百度网页搜索中搜索频次的加权得到的指数, 在本文中单位记为: 点。)

为确保数据的有效性和科学性, 将日均生育网络关注度与生育率进行 Pearson 相关性分析。生育率数据可从《中国人口和就业统计年鉴》中得到, 但由于数据周期为前年 11 月至次年 10 月, 与其他数据的数据周期不同, 且 2021 年、2022 年数据暂无, 故本文对 2013~2022 年我国生育率进行测算, 根据生育率计算公式:

$$\text{生育率} = \frac{\text{出生人数}}{\text{平均育龄妇女人数}}$$

式中出生人数数据由国家统计局得到, 平均育龄妇女人数由年末总人数乘以育龄女性比例, 数据均来自 2014~2022 年《中国统计年鉴》。

日均生育网络关注度及生育率数据见表 1, Pearson 相关分析结果见表 2, 相关系数为 0.814, 且 $p = 0.004 < 0.01$, 说明日均生育网络关注度与生育率显著正相关, 生育网络关注度在一定程度上可以反映出生育率的变化规律, 具有较高的合理性和可信度。

Table 1. Average daily fertility network attention and fertility rate
表 1. 日均生育网络关注度及生育率

年份	生育网络关注度	生育率
2013	2946.934	44.00308
2014	3947.321	45.83238
2015	3884.373	45.62217
2016	4182.699	52.66679
2017	4204.836	50.00265
2018	4204.06	44.33221
2019	3909.518	43.32883
2020	2992.842	37.23855
2021	3025.233	33.47307
2022	2931.896	30.51681

Table 2. Pearson correlation analysis
表 2. Pearson 相关性分析

		生育网络关注度	生育率
生育网络关注度	相关系数	1	
	<i>p</i> 值		
生育率	相关系数	0.814**	1
	<i>p</i> 值	0.004	

* $p < 0.05$; ** $p < 0.01$.

2.2. 研究方法

2.2.1. 季节性集中指数

季节集中指数可用于分析生育网络关注度在一年 12 个月内分布的集中程度, 其计算公式如下:

$$R = \sqrt{\frac{\sum_{i=1}^{12} (X_i - 8.33)^2}{12}}$$

其中, X_i 为每月生育网络关注度与全年生育网络关注度总数的比值。季节性集中指数值越大, 表示生育网络关注度在每月的分布越集中, 季节性变化越大; 相反, 该值越趋近于 0, 表示生育网络关注度在每月的分布越均匀, 季节性变化越小。

2.2.2. 变异系数

变异系数主要用于比较不同省区间的生育网络关注度的离散程度。变异系数越大, 说明不同省区间的生育网络关注度差异越大; 反之, 变异系数越小, 说明不同省区间的生育网络关注度差异越小。变异系数计算公式如下:

$$C_v = \frac{1}{\bar{X}} \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}}$$

其中, C_v 为变异系数, X_i 为 i 地区的网络关注度, \bar{X} 为 X_i 的平均值。

2.2.3. 全局空间自相关

全局空间自相关用于反应整个研究区域中所有要素之间的显著性、平均关联程度和空间分布模式[18]。本文采用 Global Moran's I 统计量反应各省区生育网络关注度空间关联程度的总体特征, Global Moran's I 统计量具体计算公式如下:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{X})(x_j - \bar{X})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij} \sum_{i=1}^n (x_i - \bar{X})^2}$$

其中, n 为总省区数, x_i , x_j 分别为省区 i 和省区 j 的生育网络关注度, \bar{X} 为全部省区生育网络关注度的平均值, w_{ij} 为空间权重矩阵 W 的 (i, j) 元素, 空间权重矩阵 W 选用基于邻接关系的 K-Nearst Neighbor 矩阵, K 取常用值 4。式中 I 值的显著性采用 Z 值检验:

$$Z = \frac{I - E(I)}{\sqrt{Var(I)}}$$

其中, $E(I)$ 表示 Moran's I 的期望值, $Var(I)$ 表示 Moran's I 的方差。在正态分布下, 若 $1.645 \leq |Z(I)| \leq 1.96$, $p < 0.1$, 表示区域间观测值在 10% 的显著性水平上存在空间相关性; 若 $1.96 \leq |Z(I)| \leq 2.58$, $p < 0.05$, 表示区域间观测值在 5% 的显著性水平上存在空间相关性; $|Z(I)| \geq 2.58$, $p < 0.01$, 表示区域间观测值在 1% 的显著性水平上存在空间相关性。Global Moran's I 的取值范围为 $[-1, 1]$, 在通过显著性检验的前提下, 若 $I > 0$, 说明所研究区域存在空间正相关, 且值越接近于 1, 空间正相关性越强, 研究对象的值呈聚集分布; 若 $I < 0$, 说明所研究区域存在空间负相关, 且值越接近于 -1, 空间负相关性越强, 研究对象的值呈离散(互斥)分布。

2.2.4. 局部空间自相关

局部空间自相关用来识别不同空间单元上可能存在的空间关联模式, 反应要素的空间聚集和分异特

征[19]。本文采用 Local Moran's I 统计量来探索局部空间聚类分布特征, Local Moran's I 统计量具体计算公式如下:

$$I_i = \frac{n(x_i - \bar{X}) \sum_{j \neq i} w_{ij} (x_j - \bar{X})}{\sum_{i=1}^n (x_i - \bar{X})^2}$$

其中, n 为总省区数, x_i 为省区 i 的生育网络关注度, \bar{X} 为全部省区生育网络关注度的平均值, w_{ij} 为空间权重矩阵 W 的 (i, j) 元素。Local Moran's I 指数检验方式与 Global Moran's I 指数相同。通过 Local Moran's I 指数可以区分具体集聚模式, 包括 H-H 集聚(高观测值地区被高观测值地区包围)、L-L 集聚(低观测值地区被低观测值地区包围)、L-H 集聚(低观测值地区被高观测值地区包围)和 H-L 集聚(高观测值地区被低观测值地区包围)。

3. 生育网络关注度的时空演变特征

3.1. 时间演变特征

3.1.1. 年际变化

根据 2013~2023 年全国尺度的生育网络关注度分别计算各年份平均关注度及每月的平均关注度。2013~2023 年我国生育网络关注度年际变化趋势图如图 1 所示, 2013~2014 年我国生育网络关注度呈现上升趋势, 生育网络关注度由 2013 年的 2946.93 点上升到 2014 年的 3947.32 点, 年际增长率达到 33.95%; 2014~2019 年我国生育网络关注度呈现平稳态势, 其中 2017 年关注度达到峰值 4204.8 点; 生育网络关注度在 2019~2020 年出现大幅下降, 由 2019 年的 3909.52 点降低到 2020 年的 2992.84 点; 生育网络关注度在 2020~2023 年呈现平稳态势。由于 2020 年生育网络关注度的大幅下降, 2020~2022 年的生育网络关注度在近 10 年间处于低位, 这与新冠疫情持续时间重合, 表明新冠疫情对我国生育网络关注度产生了明显的负面影响。根据 2023 年 1~3 月的数据, 网络关注度的变化率为 8.39%, 有“抬头”倾向, 这可能与新冠疫情结束有关。

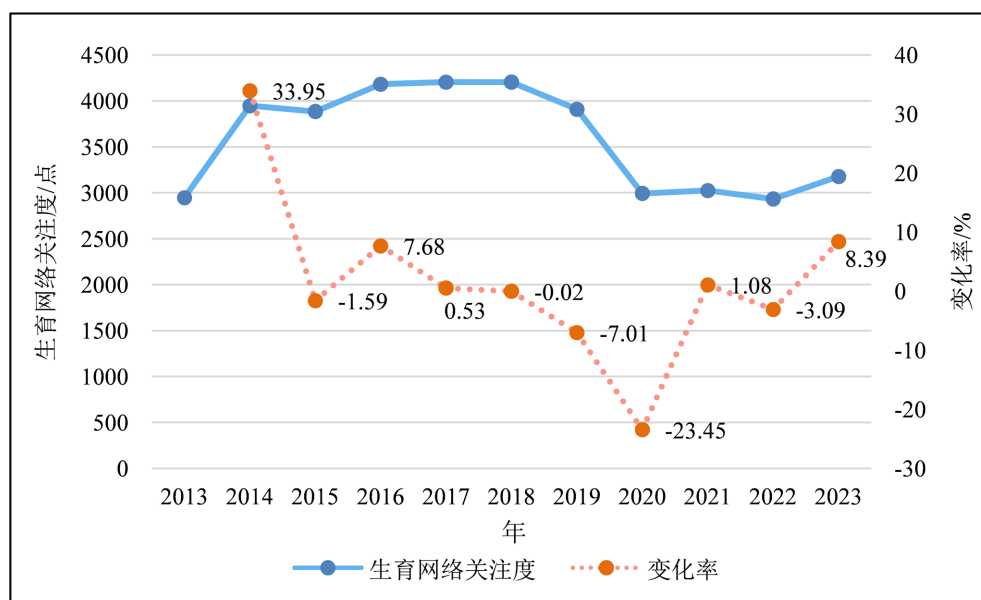


Figure 1. Inter-annual change trend chart of China's fertility network attention from 2013 to 2023
图 1. 2013~2023 年我国生育网络关注度年际变化趋势图

3.1.2. 月际变化

为研究我国生育网络关注度是否根据季节变化而变化, 进一步对月际变化进行分析。2013~2022 年我国生育网络关注度逐年月际变化趋势图如图 2 所示。

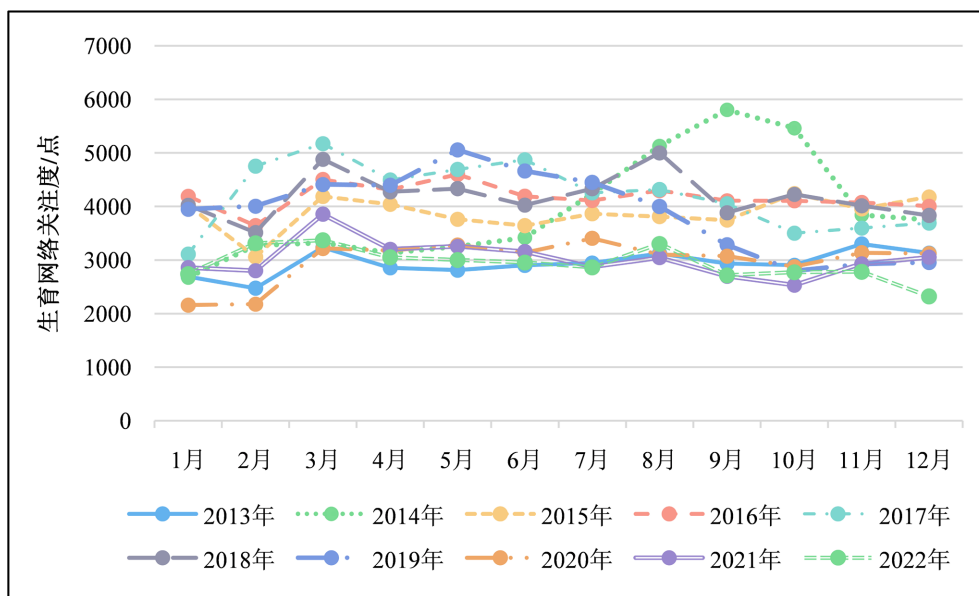


Figure 2. Inter-month change trend chart of China's fertility network attention from 2013 to 2023
图 2. 2013~2023 年我国生育网络关注度月际变化趋势图

2013~2022 年间逐年月际变化整体呈现平稳趋势, 存在年内变化差异, 生育网络关注度在每年 3 月均出现不同程度的升高, 在每年 4 月出现不同程度的降低; 除 2014 年, 其余年份的生育网络关注度在每年 9 月也出现不同程度的升高。2013~2022 年关注度的月度峰值出现在 3~11 月, 无明显特征; 2013~2018 年关注度的月度谷值均出现于 1 月和 2 月, 且每年谷值月份与当年春节假期所在月份完全重合, 可以说明春节假期期间人们的生育网络关注度有所下降。为进一步分析我国生育网络关注度的月际分布集中特征, 应用季节性集中指数(表 3)。

Table 3. Seasonal concentration index
表 3. 季节性集中指数

年份	2013 年	2014 年	2015 年	2016 年	2017 年	2018 年	2019 年	2020 年	2021 年	2022 年
季节性集中指数	0.6294	2.0391	0.6605	0.4669	1.1952	0.8005	1.5312	1.0750	0.8905	0.8212

2013~2022 年的季节性集中指数最大值为 2014 年的 2.0319, 其余年份的季节性集中指数均小于 2, 说明一年中生育网络关注度分布较为均匀。

3.2. 空间演变特征

3.2.1. 时空格局变化

根据 2013~2023 年我国 31 个省区市省级尺度的生育网络关注度计算分省尺度各年份平均关注度及各年份变异系数, 变异系数数据趋势见表 4 及图 3。

Table 4. Coefficient of variation
表 4. 变异系数

年份	2013年	2014年	2015年	2016年	2017年	2018年	2019年	2020年	2021年	2022年
变异系数(%)	0.6621	0.4879	0.4869	0.5069	0.5659	0.5533	0.5840	0.5130	0.5006	0.5115

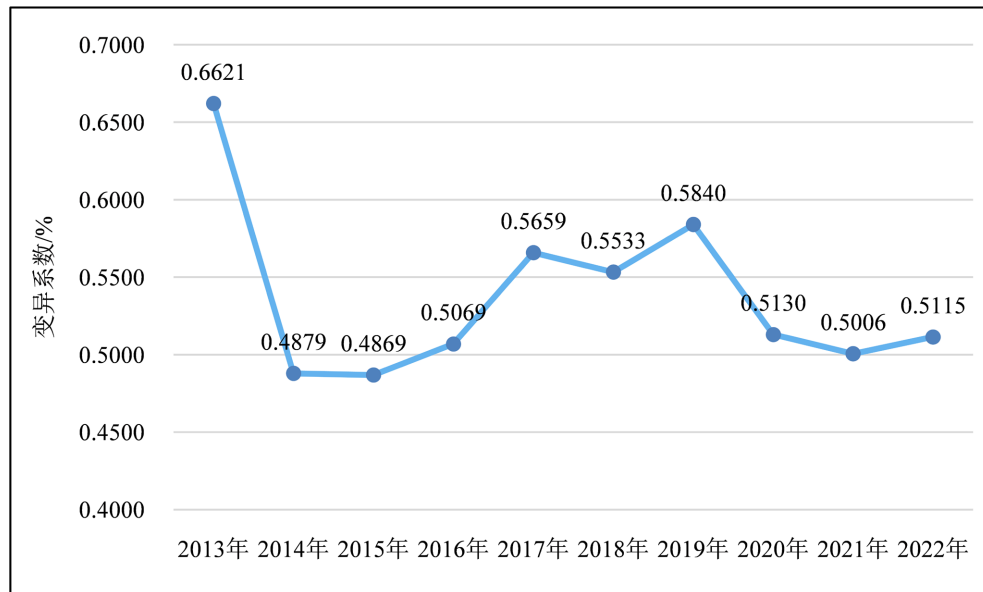
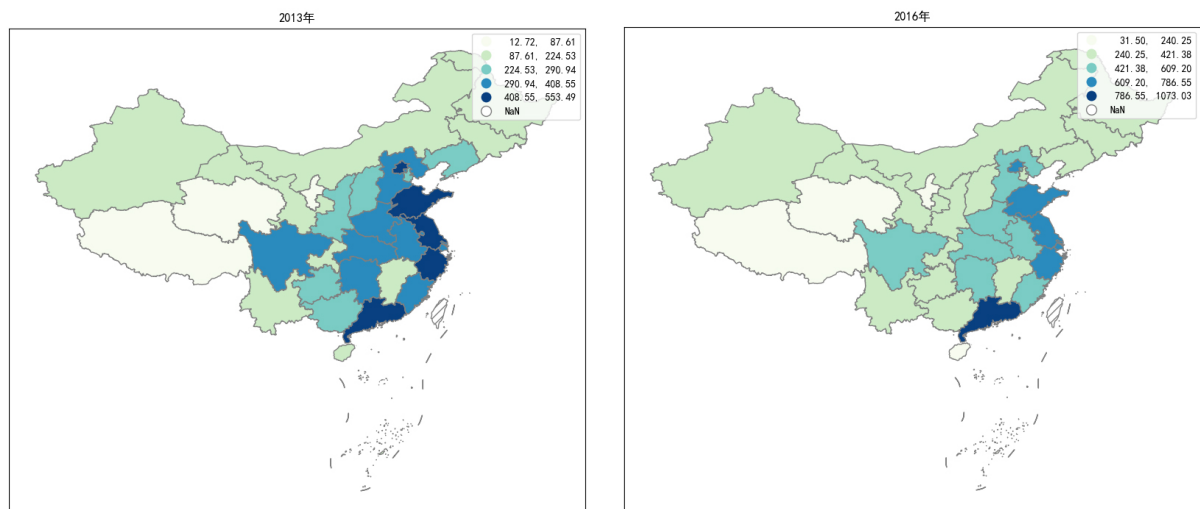


Figure 3. Coefficient of variation
图 3. 变异系数

如上图所示,2013~2022年我国31个省区市的生育网络关注度的变异系数呈现“下降-上升-下降”趋势,总体呈现缩小趋势,2020~2022年变异系数呈稳定状态,表明地区之间的生育网络关注度差异在缩小。

利用 Python 工具及 Pysal 开源库对我国31个省区市的生育网络关注度进行空间分析,利用自然断点法分为5个等级,时空演变特征如图4所示。



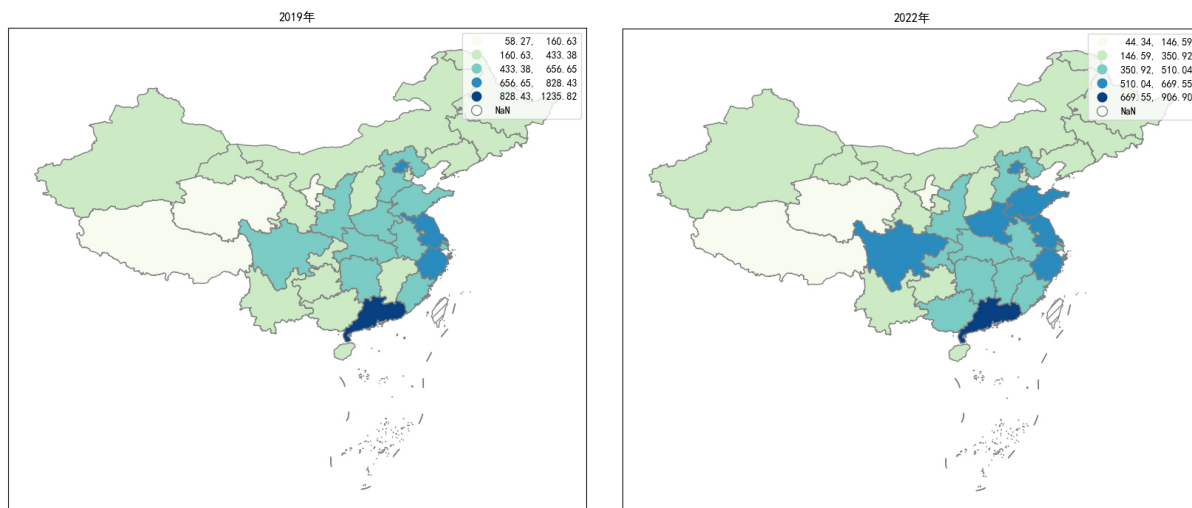


Figure 4. Attention distribution characteristics of fertility network in 31 provinces, autonomous regions and municipalities in 2013, 2016, 2019 and 2022

图 4. 2013 年、2016 年、2019 年、2022 年我国 31 个省区市的生育网络关注度分布特征

2013 年生育网络关注度最高省份包括广东、北京、山东、浙江和江苏, 较高省份包括上海、河南、河北、湖北、福建、湖南、安徽和四川, 最低省份包括宁夏、青海和西藏。2016 年生育网络关注度最高省份为广东, 较高省份包括北京、江苏、山东、浙江和上海, 最低省份包括海南、宁夏、青海和西藏。2019 年生育网络关注度最高省份为广东, 较高省份包括江苏、北京和浙江, 最低省份包括宁夏、青海和西藏。2022 年生育网络关注度最高省份为广东, 较高省份包括江苏、浙江、北京、山东、河南和四川, 最低省份包括宁夏、青海和西藏。2013~2022 年我国生育网络关注度第一等级城市由 5 个减少到 1 个, 第二等级城市由 8 个减少到 6 个, 第三等级由 6 个增加到 10 个, 第四等级由 9 个增加到 11 个, 第五等级城市保持在 3 个, 表明我国生育网络关注度在不断降低, 原先关注度较高的城市不能保持原有关注度。可以发现, 2013~2022 年我国不同地区生育网络关注度与地区生产总值排名有关, 网络关注度较高城市多集中于地区生产总值较高城市, 如广东、江苏、山东、浙江等, 网络关注度较低城市多集中于地区生产总值较低城市, 如宁夏、青海、西藏等。

此处引入“胡焕庸线”假说, “胡焕庸线”为我国地理学家胡焕庸在 1935 年提出的划分我国人口密度的对比线, 是我国人口发展水平和经济社会格局的分界线[20]。我国 31 个省区市的生育网络关注度也存在“胡焕庸线”现象, 网络关注度较高的城市集中在胡焕庸线以东, 呈现聚集趋势; 四川省为唯一处于胡焕庸线以西的较高生育网络关注度城市; 胡焕庸线以西其余省份均为关注度较低城市。

3.2.2. 全局空间特征

表 5 为 2013~2022 年我国 31 个省区市的 Global Moran's I 指数。P 值均小于 0.05, 表示在 5% 的显著性水平上存在空间相关性。可以看出, Moran's I 指数呈现下降趋势, 由 2013 年的 0.3726 下降到 2022 年的 0.2101, 表明我国 31 个省区市间的生育网络关注度存在空间正相关性, 且呈现减小趋势, 生育网络关注度在空间上表现为集聚分布, 具有较高(较低)生育网络关注度的省份与较高(较低)生育网络关注度的省份相对临近。

3.2.3. 局域空间特征

计算 2013~2022 年我国 31 个省区市的 Local Moran's I 指数, 结果见表 6。可以看出, 我国生育网络关注度具有空间自相关性的城市由 2013 年的 12 个降低到 2022 年的 6 个, 整体呈现下降趋势。2013 年

生育网络关注度有 38.71% 的省份呈现出空间集聚特征, 其中 H-H 型省份有 8 个, L-L 型省份有 4 个; 2016 年生育网络关注度有 25.81% 的省份呈现出空间集聚特征, 其中 H-H 型省份有 5 个, L-L 型省份有 3 个; 2019 年和 2022 年生育网络关注度均有 19.35% 的省份呈现出空间集聚特征, 其中 H-H 型省份有 4 个, L-L 型省份有 2 个。

Table 5. The Global Moran's I index

表 5. Global Moran's I 指数

年份	2013 年	2014 年	2015 年	2016 年	2017 年	2018 年	2019 年	2020 年	2021 年	2022 年
Moran's I	0.3726	0.3952	0.3565	0.2329	0.2474	0.2294	0.2075	0.2411	0.263	0.2101
p-value	0.001	0.001	0.002	0.007	0.009	0.011	0.017	0.014	0.008	0.013

Table 6. Cluster characteristics

表 6. 集聚特征

类型	2013 年	2016 年	2019 年	2022 年
H-H	天津、河北、上海、江苏、浙江、安徽、福建、山东	上海、江苏、安徽、福建、山东	上海、江苏、安徽、福建	上海、江苏、安徽、福建
L-L	陕西、青海、宁夏、西藏	陕西、宁夏、西藏	宁夏、西藏	宁夏、西藏
L-H				
H-L				

2013~2022 年间, 上海、江苏、安徽和福建这四个城市“常驻”H-H 类型省份, 表明这四个省份与相邻省份呈现出高生育网络关注度的集聚效应。四个城市均位于华东地区, 华东地区自然环境条件优越, 物产资源丰富, 是我国综合技术水平最高的经济区, 城市的经济活动等因素均对生育网络关注度产生影响。宁夏和西藏这两个城市“常驻”L-L 类型省份, 表明这两个省份与相邻省份呈现出低生育网络关注度的集聚效应。

4. 生育问题网络关注度的影响因素分析

目前我国的生育率水平出较低的状态上, 本节使用面板数据针对关于生育的网络关注度进行影响因素分析。如表 7 所示, 我们选取这些指标作为研究生育问题的网络关注度的影响因素分析。

Table 7. Selection and presentation of indicators

表 7. 指标选取及表示

指标	表示
人口自然增长率	gr
网络关注度	na
居民消费价格指数	pi
享受生育保险待遇人数	nb
地区生产总值	gdp

Continued

地区生产总值指数	igdp
全体居民人均可支配收入	incr
每十万人人口高等学校平均在校生数	avc
年末参加生育保险人数	yb
人均地区生产总值	avgdp
出生率	br
粗离婚率	rdr
教育经费	ef
结婚登记	m

本节数据来源：其中，网络关注度为前文提到并爬取的数据；其余各项指标数据来源于国家统计局分地区年度数据库，由于人口自然增长率和出生率缺少 2020 年数据，本文采用算术平均计算 2021 年、2019 年、2018 年、2017 年数据获得，教育经费缺少 2021 年的数据，本文采用回归的方法进行预测获得。

根据现有模型等，本节构建面板回归模型针对生育网络关注度进行影响因素分析，具体模型结构如下：

$$na = \gamma_1 gr + \gamma_2 pi + \gamma_3 nb + \gamma_4 gdp + \gamma_5 igdp + \gamma_7 incr + \gamma_8 avc + \gamma_9 yb + \gamma_{10} avgdp + \gamma_{11} br + \gamma_{12} rdr + \gamma_{13} ef + \gamma_{14} m + \gamma_7 + \gamma_0 + u$$

4.1. 固定效应分析

本节使用 stata 软件针对以上涉及的所有的变量进行面板回归的固定效应进行分析，第一次结果如表 8 所示。从中我们可以看到最终结果当 P 值大于 0.0005 时，大于 F 值，但是我们可以看到三个指标地区生产总值、地区人均生产总值和结婚登记的 P 值较大，均大于 0.5。故需剔除这三个变量再次回归。我们得到结果如表 9 所示。同理从表 9 中我们不难看出最终结果当 P 值大于 0.0002 时，大于 F 值，但是我们可以看到居民消费价格指数、享受生育保险待遇人数和教育经费这三个指标的 P 值较大，均大于 0.1，因此我们同样需要剔除这三个变量再一次做面板回归。如此有了表 10。从表 10 中我们可以看到最终结果当 P 值大于 0.0000 时，大于 F 值，而且此时的各指标的 P 值都处于相对较小的数值，均小于 0.15，故选取剩余指标构建新的面板回归模型，具体如下显示：

$$na = 8.82gr + 1.64igdp + 3.38incr + 10.92avc - 4.58yb + 4.54br - 4.2rdr - 2.7$$

Table 8. Results of the first fixed-effects analysis

表 8. 第一次固定效应分析结果

na	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]
gr	17581.6300	2231.3000	7.88	0.000	13185.7300 21977.5400
pi	11274.6100	7290.1560	1.55	0.123	-3087.7970 25637.0200
nb	356.2004	357.2872	1.00	0.320	-347.6947 1060.0960
gdp	0.1417	0.6257	0.23	0.821	-1.0911 1.3745

Continued

igdp	4246.3190	2514.5720	1.69	0.821	-707.6635	9200.3020
incr	2.5318	2.0338	1.24	0.214	-1.4751	6.5387
avc	105.7789	10.4074	10.16	0.000	85.2751	126.2827
yb	-47.8174	20.7232	-2.31	0.022	-88.6444	-6.9905
avgdp	-0.1713	0.8498	-0.20	0.840	-1.8455	1.5029
br	13240.0200	4872.9450	2.72	0.007	3639.7790	22840.2500
rdr	-44197.9300	12337.3400	-3.58	0.000	-68503.8500	-19892.0200
ef	-0.0062	0.0040	-1.55	0.122	-0.0140	0.0017
m	-814.4031	1299.8800	-0.63	0.532	-3375.3100	1746.5030
_cons	-1743521.0000	833782.1000	-2.09	0.038	-3386163.0000	-100878.4000
sigma_u	95369.5110					
sigma_e	82667.3000					
rho	0.5710			(fraction of variance due to u_i)		
	F test that all u_i = 0: F (30, 235) = 2.22				Prob > F = 0.0005	

Table 9. Results of the second fixed-effects analysis

表 9. 第二次固定效应分析结果

na	Coef.	Std. Err.	t	P > t	[95% Conf.	Interval]
gr	17530.4400	2133.8100	8.22	0.000	13326.8700	21734.0000
pi	11039.8600	7237.1990	1.53	0.128	-3217.2880	25297.0100
nb	275.3949	328.7537	0.84	0.403	-372.2438	923.0336
igdp	4246.4580	2340.3490	1.81	0.071	-363.9871	8856.9030
incr	2.1577	0.6752	3.20	0.002	0.8275	3.4878
avc	103.8427	9.5268	10.90	0.000	85.0751	122.6103
yb	-43.3833	9.4381	-4.60	0.000	-61.9762	-24.7903
br	13184.0900	4710.6830	2.80	0.006	3904.1360	22464.0500
rdr	-43591.9700	12124.5400	-3.60	0.000	-67477.0900	-19706.8500
ef	-0.0043	0.0027	-1.62	0.105	-0.0095	0.0009
_cons	-1763430.0000	815350.6000	-2.16	0.032	-3369655.0000	-157204.0000
sigma_u	91577.4960					
sigma_e	82233.8340					
rho	0.5536			(fraction of variance due to u_i)		
	F test that all u_i = 0: F (30, 238) = 2.35				Prob > F = 0.0002	

Table 10. Results of the third fixed-effect analysis
表 10. 第三次固定效应分析结果

na	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]	
gr	17568.4700	1992.2710	8.82	0.000	13643.9800 21492.9600	
igdp	3791.6200	2316.0890	1.64	0.103	-770.7427 8353.9830	
igdp	2.2561	0.6667	3.38	0.001	0.9427 3.5695	
avc	104.0911	9.5356	10.92	0.000	85.3073 122.8748	
yb	-43.3045	9.4507	-4.58	0.000	-61.9210 -24.6881	
br	17849.3100	3930.6910	4.54	0.000	10106.4100 25592.2000	
rdr	-49454.1300	11778.8800	-4.20	0.000	-72656.8300 -26251.4300	
_cons	-671787.5000	248951.6000	-2.70	0.007	-1162186.0000 -181388.7000	
sigma_u	110516.3200					
sigma_e	82615.4080					
rho	0.6415				(fraction of variance due to u_i)	
F test that all u_i = 0: F (30, 241) = 3.71				Prob > F = 0.0000		

4.2. 随机效应分析

本节使用 stata 软件针对以上涉及的所有的变量进行面板回归的随机效应进行分析, 第一次结果如表 11 所示。从中我们可以看到四个指标居民消费价格指数、地区生产总值、地区人均生产总值和粗离婚率的 P 值较大, 均大于 0.1。

Table 11. Results of the first random-effects analysis
表 11. 第一次随机效应分析结果

na	Coef.	Std. Err.	z	P > z	[95% Conf. Interval]
gr	16592.8300	2327.5450	7.13	0.000	12030.9300 21154.7400
pi	11898.0500	7635.0740	1.56	0.119	-3066.4240 26862.5200
nb	744.6423	238.8449	3.12	0.002	276.5149 1212.7700
gdp	0.4309	0.6359	0.68	0.498	-0.8154 1.6773
igdp	4455.0550	2566.7560	1.74	0.083	-575.6945 9485.8040
incr	5.5662	2.0487	2.72	0.007	1.5508 9.5816
avc	102.0095	10.5732	9.65	0.000	81.2864 122.7326
yb	-55.2159	21.3967	-2.58	0.010	-97.1526 -13.2792
avgdp	-1.1967	0.8608	-1.39	0.164	-2.8838 0.4904

Continued

br	-7739.3500	2220.5480	-3.49	0.000	-12091.5400	-3387.1550
rdr	-6664.8150	6609.3850	-1.01	0.313	-19618.9700	6289.3430
ef	-0.0072	0.0015	-4.84	0.000	-0.0101	-0.0043
m	444.1482	346.3053	1.28	0.200	-234.5977	1122.8940
_cons	-1761440.0000	866699.8000	-2.03	0.042	-3460140.0000	1122.8940
sigma_u	0.0000					
sigma_e	82667.3000					
rho	0.0000				(fraction of variance due to u_i)	

故需剔除这四个变量再次回归。我们得到结果如表 12 所示。从表 12 中我们可以看到地区生产总值指数、粗离婚率和结婚登记这三个指标的 P 值较大, 均大于 0.3。

Table 12. A second random-effects analysis

表 12. 第二次随机效应分析

na	Coef.	Std. Err.	z	P> z	[95% Conf.	Interval]
gr	17085.6700	2125.8140	8.04	0.000	12919.1500	21252.1900
nb	704.7066	254.9057	2.76	0.006	205.1005	1204.3130
igdp	2322.3070	2360.7350	0.98	0.325	-2304.6490	6949.2620
incr	2.8874	0.6749	4.28	0.000	1.5646	4.2101
avc	96.8049	9.6626	10.02	0.000	77.8666	115.7433
yb	-46.6399	9.4500	-4.94	0.000	-65.1615	-28.1184
rdr	-1468.3100	6791.4940	-0.22	0.829	-14779.3900	11842.7700
ef	-0.0066	0.0015	-4.39	0.000	-0.0095	-0.0036
m	214.2210	356.7346	0.60	0.548	-484.9659	913.4079
_cons	-399493.2000	253496.7000	-1.58	0.115	-896337.6000	97351.1100
sigma_u	24197.5670					
sigma_e	83845.1140					
rho	0.0769				(fraction of variance due to u_i)	

因此, 我们同样需要剔除这三个变量再一次做面板回归。从表 13 中我们可以看到此时的各指标的 P 值都处于相对较小的数值, 均小于 0.01, 故选取剩余指标构建新的面板回归模型, 具体如下显示:

$$na = 8.82gr + 2.64nb + 4.32incr + 10.21avc - 5.01yb - 4.78ef - 5.12$$

Table 13. The third random-effect analysis
表 13. 第三次随机效应分析

na	Coef.	Std. Err.	z	P > z	[95% Conf.	Interval]
gr	17342.6000	1966.8230	8.82	0.000	13487.6900	21197.5000
nb	652.9630	247.3626	2.64	0.008	168.1412	1137.7850
incr	2.8157	0.6524	4.32	0.000	1.5370	4.0944
avc	97.6259	9.5607	10.21	0.000	78.8873	116.3646
yb	-46.4312	9.2736	-5.01	0.000	-64.6072	-28.2553
ef	-0.0061	0.0013	-4.78	0.000	-0.0085	-0.0036
_cons	-154243.1000	30112.5200	-5.12	0.000	-213262.5000	-95223.6300
sigma_u	25711.3890					
sigma_e	84607.1930					
rho	0.0845				(fraction of variance due to u_i)	

4.3. 传统 Hausman 检验

根检验显著性、相关联程度, 结果如表 14 和表 15 所示。

Table 14. Model significance and degree of correlation of fixed effects
表 14. 固定效应的模型显著性、相关联程度

na	Coef.	Std. Err.	t	P > t	[95% Conf.	Interval]
gr	17568.4700	1992.2710	8.82	0.000	13643.9800	21492.9600
igdp	3791.6200	2316.0890	1.64	0.103	-770.7427	8353.9830
incr	2.2561	0.6667	3.38	0.001	0.9427	3.5695
avc	104.0911	9.5356	10.92	0.000	85.3073	122.8748
yb	-43.3045	9.4507	-4.58	0.000	-61.9210	-24.6881
br	17849.3100	3930.6910	4.54	0.000	10106.4100	25592.2000
rdr	-49454.1300	11778.8800	-4.20	0.000	-72656.8300	-26251.4300
_cons	-671787.5000	248951.6000	-2.70	0.007	-1162186.0000	-181388.7000
sigma_u	110516.3200					
sigma_e	82615.4080					
rho	0.6415				(fraction of variance due to u_i)	
				F test that all u_i = 0: F (30, 241) = 3.71		Prob > F = 0.0000

Table 15. Model significance and correlation degree of random effects
表 15. 随机效应的模型显著性、相关程度

na	Coef.	Std. Err.	t	P > t	[95% Conf.	Interval]
gr	18546.3700	2037.8000	9.10	0.000	14532.2800	22560.4600
nb	586.7191	311.4113	1.88	0.061	-26.7036	1200.1420
incr	2.2459	0.6689	3.36	0.001	0.9282	3.5635
avc	103.2083	9.6229	10.73	0.000	84.2530	122.1636
yb	-44.8631	9.2635	-4.84	0.000	-63.1104	-26.6158
ef	-0.0083	0.0023	-3.68	0.000	-0.0128	-0.0039
_cons	-130056.4000	36448.8400	-3.57	0.000	-201853.9000	-58258.9200
sigma_u	48398.5610					
sigma_e	84607.1930					
rho	0.2465				(fraction of variance due to u_i)	
F test that all u_i = 0: F (30, 242) = 2.14				Prob > F = 0.0009		

由表 14 和表 15 得到固定效应的模型显著性和关联程度均优于随机效应, 因此本文选用固定效应模型进行分析。

综上所述, 对于生育问题的网络关注度的影响因素分析, 我们得出, 虽然影响因素很多, 但归根到底, 可以分为三个方面: 文化、政治和经济因素。本文重点研究的是文化和经济因素, 其中, 与人口自然增长率、地区生产总值指数人均可支配收入、每十万人中在校大学生的人数和出生率与网络关注度成正相关, 年末参加生育保险人数和粗离婚率与生育的网络关注度成负相关。由于目前我国的生育率在逐年降低, 而生育问题的网络关注度与其成正相关, 故我们需要正视生育问题, 增加对生育问题的网络关注度, 加强互联网的使用, 积极引导人们关注生育问题, 逐步提高生育率。

5. 主要结论与政策建议

5.1. 主要结论

通过对本研究的梳理, 发现生育网络关注度与生育率显著正相关, 生育网络关注度在一定程度上可以反映出生育率的变化规律。随后对生育网络关注度研究得出以下结论:

(一) 生育网络关注度的时空演变特征

1) 2013 年~2022 年生育网络关注度呈上升 - 平稳 - 下降趋势。2014 年开始上升且年度变化率最大, 可能与我国 2014 年的单独二胎政策实行有关。17 年达到峰值, 20 年受新冠疫情影响大幅下降, 20 年~22 年都处于近 10 年低位。23 年随着新冠疫情的结束, 生育网络关注度出现“抬头”倾向。生育网络关注度的每年月际变化整体呈现平稳趋势, 存在年内变化差异, 生育网络关注度一般在每年 3 月、9 月均出现不同程度的升高, 在 4 月出现不同程度的降低。而每年谷值月份与当年春节假期所在月份完全重合, 可以说明春节假期期间人们的生育网络关注度有所下降。

2) 2013~2022 年地区之间的生育网络关注度差异在缩小。我国生育网络关注度在不断降低, 原先关

关注度较高的城市不能保持原有关关注度。我国不同地区生育网络关注度与地区生产总值排名有关, 网络关注度较高城市多集中于地区生产总值较高城市, 如广东、江苏、山东、浙江等, 且只有广东始终在最高位, 在广东, 一方面潮汕等地的生育意愿比较高, 另一方面广东流入的人口最多, 流入人口以青壮年为主, 育龄年龄段的人口较多。网络关注度较低城市多集中于地区生产总值较低城市, 其中宁夏、青海、西藏始终位于最低位。

3) 2013~2022年, 我国31个省区市间的生育网络关注度存在较强的空间正相关性, Moran's I 指数呈现下降趋势, 空间集聚性较为减弱, 且空间差异程度呈强弱交叉状态出现。随着时间的推移, 形成以高值和低值集聚区为主的关联类型区, 其中, 华东地区这四个省份“上海、江苏、安徽和福建”与相邻省份呈现高生育率的集聚效应, 西北、西南地区呈现低生育率的集聚效应, 这可能与经济、社会、人口等各方面的因素有关。

(二) 生育问题网络关注度的影响因素

研究发现生育网络关注度的影响因素可以分为三个方面: 文化、政治和经济因素。本文重点研究的是文化和经济因素。

1) 人口自然增长率、地区生产总值指数人均可支配收入、每十万人中在校大学生的人数和出生率与网络关注度成正相关; 年末参加生育保险人数和粗离婚率与生育的网络关注度成负相关。

2) 生育养育孩子的经济成本、因疫情导致劳动就业面临的风险提升、新媒体快速传播低生育文化是影响我国生育率的主要因素。

5.2. 政策建议

基于以上结论, 本文给出如下建议:

(一) 重视生育政策实施过程中民众的声音。在实行政策的同时, 重视生育网络关注度的变化及时引导, 关注互联网上大众的意见反馈, 对政策做出及时调整, 以确保正确合适的政策实行对生育及生育网络关注度产生积极性的影响。

(二) 帮助家庭降低生育养育的经济成本。在就业、税收、医疗等方面制定有利于降低家庭生育养育成本的政策。完善生育保险制度, 保障养育期女性的经济收入, 同时, 政府应加大公共教育投资力度, 尤其是加大对婴幼儿和早期教育的财政投入。

(三) 营造生育友好型社会文化环境。通过有效宣传倡导, 比如: 政府通过运营自媒体、公众号、抖音等对婚恋、育儿等话题进行合理引导和讨论。同时, 加快家庭婚育文化建设, 积极引导年轻人婚育观点从重视自我向重视家庭转变。

参考文献

- [1] 穆光宗, 林进龙. 论生育友好型社会——内生性低生育阶段的风险与治理[J]. 探索与争鸣, 2021(7): 56-69+178.
- [2] 刘伟杰, 闫文辉. 当代青年婚姻观和生育观的研究现状与展望[J]. 山东农业工程学院学报, 2021, 38(1): 94-99.
- [3] 景鹏, 李桢, 王明璐. 生育支持、生育率与经济增长[J]. 经济科学, 2023(1): 169-186.
- [4] 田悦芳, 孔广用. 女性凝视下生育问题的“可见性”——基于纪录片《奇妙的蛋生》的文化反思[J]. 采写编, 2022(12): 158-159.
- [5] 刘璐. 生育风险应激: 社交媒体环境下的女性“恐育”行为研究[D]: [硕士学位论文]. 沈阳: 辽宁大学, 2022.
- [6] 乔嘉如. 网络媒体的风险放大机制[D]: [硕士学位论文]. 广州: 暨南大学, 2019.
- [7] 高鹿源. 网络舆情中的生育态度及其影响因素研究[D]: [硕士学位论文]. 杭州: 浙江大学, 2022.
- [8] 黄君洁, 邱婷. 社交网络中的亲子信息对生育意愿的影响研究[J]. 人口与社会, 2022, 38(6): 79-93.
- [9] 廖晓妹, 韩庆茹, 张摇, 马力. 极年轻乳腺癌患者生育问题关注度的影响因素及其预后分析[J]. 中华乳腺病杂志

- (电子版), 2020, 14(5): 285-290.
- [10] 周旭平, 张丽, 张百蕾, 王英, 吴嘉涵. 三孩生育政策下产后出血的产前高危因素分析[J]. 中国妇幼健康研究, 2023, 34(3): 21-28.
- [11] 董翠丽, 王红. 育龄期乳腺癌患者生育忧虑现状及其影响因素分析[J/OL]. 保健医学研究与实践: 1-13. <https://wap.cnki.net/touch/web/Journal/Article/GXBJ20230210001.html>, 2023-03-05.
- [12] 朱艺, 孙娜, 殷晓旭, 方鹏骞, 龚言红. 生育政策调整背景下我国女性产后抑郁症状与育儿自我效能双向关系的追踪研究——基于湖北省的实证调查[J]. 人口与发展, 2023, 29(1): 27-39.
- [13] 陈钰晓, 周魅. 家庭养老照护对育龄女性二孩生育行为的影响[J]. 湘潭大学学报(哲学社会科学版), 2023, 47(1): 52-57+148.
- [14] 彭争呈, 邹红, 宋泽. 多孩生育政策抑制了女性就业吗——基于单独二孩政策的研究[J]. 经济学动态, 2023(1): 89-105.
- [15] 陆温婷. 城市化、养老观念与育龄人口生育意愿分析[J]. 长江师范学院学报, 2023, 39(3): 9-18.
- [16] 姚懿轩, 王百玲. 全面二孩政策下农村育龄女性生育问题探析[J]. 农村经济与科技, 2019, 30(3): 269-270.
- [17] 孙毅, 吕本富. 网络搜索与经济行为相关性研究综述[J]. 管理评论, 2011, 23(7): 72-77.
- [18] Anselin, L. (1988) A Test for Spatial Autocorrelation in Seemingly Unrelated Regressions. *Economics Letters*, **28**, 335-341. [https://doi.org/10.1016/0165-1765\(88\)90009-2](https://doi.org/10.1016/0165-1765(88)90009-2)
- [19] Anselin, L. (1995) The Local Indicators of Spatial Association—LISA. *Geographical Analysis*, **27**, 93-115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- [20] 丁金宏, 何书金. 中国人口地理格局与城市化未来——纪念胡焕庸线发现 80 周年学术研讨会在上海举行[J]. 地理学报, 2015, 70(12): 1856.