

生育机会成本及异质性

——基于机器学习算法的发现

袁 益, 张 力

(复旦大学 人口研究所, 上海 200433)

摘 要: 生育机会成本对生育意愿和生育行为的重要性已被广泛认可, 但量化研究仍处于探索阶段。把生育机会成本视为一种“反事实”的生育行为的收入代价, 运用机器学习算法测量女性的生育机会成本并考察生育机会成本的异质性, 通过回归模型识别生育机会成本对生育意愿和生育数量的边际效应。结果显示, 生育男孩的机会成本低于女孩; 收入水平越高、受教育程度越高的女性, 生育绝对机会成本和相对机会成本越低, 这种异质性与能否获得生育保险密切相关; 职业替换门槛低的女性群体, 生育保险的覆盖率更低, 生育机会成本更高。进一步的研究发现, 生育绝对机会成本上升1单位, 女性生育二孩和多孩意愿分别下降1.6%和1.5%, 实际生育数量下降1.1%。研究发现对认识生育率发展趋势具有参考价值。

关键词: 生育机会成本; 随机森林算法; 低生育率; 生育行为; 生育意愿

中图分类号: C924.24 **文献标识码:** A **文章编号:** 1000-4149 (2021) 06-0040-14

DOI: 10.3969/j.issn.1000-4149.2021.00.047

Childbearing Opportunity Costs and Its Heterogeneity: Findings Drawn from Machine Learning Method

YUAN Yi, ZHANG Li

(Institute of Population Studies, Fudan University, Shanghai 200433, China)

Abstract: The significance of childbearing opportunity cost (COC) on fertility preference and fertility behaviors has been recognized widely, but the quantitative analysis is still under-developed. In this paper, fertility opportunity cost is regarded as a counterfactual

收稿日期: 2021-05-12; 修订日期: 2021-08-04

基金项目: 国家社会科学基金项目“老年贫困与养老待遇充分性保障研究”(19BRK019)。

作者简介: 袁益, 复旦大学人口研究所博士研究生; 张力, 复旦大学人口研究所教授。

income cost of fertility behavior, and the machine learning algorithm is used to measure women's fertility opportunity cost and investigate the heterogeneity of fertility opportunity cost. Through regression model, the marginal effect of fertility opportunity cost on fertility intention and fertility quantity is identified. The results show that the COC for male births is lower than that for female births. The women with higher education income levels have lower absolute and relative opportunity costs of childbearing. This heterogeneity is closely related to the access to maternity insurance. Low-income female laborers are largely outside the coverage of maternity insurance and have higher COC. The further analysis reveals that one unit of increasement in absolute COC will lead to the intentions to have two or more children decrease by 1.6 percent and 1.5 percent, respectively, and the actual number of births decreases by 1.1 percent. The findings make good reference for understanding the future fertility trend.

Keywords: childbearing opportunity costs; random forest; low fertility; fertility behavior; fertility preference

一、引言

自 20 世纪 80 年代计划生育政策被确定为基本国策以来,我国生育率在更替水平下持续走低,逐步放开的政策限制也未能显著提升育龄人群的生育意愿和生育潜力^[1-6]。国家统计局相关调查和第七次全国人口普查结果显示,育龄妇女的意愿生育子女数仅为 1.8,育龄妇女总和生育率已低至 1.3^①。这种状况与生育成本高导致低生育意愿和生育数量的观点受到社会的普遍认同。

生育成本指养育子女需要投入的物质和非物质资源的价值,是影响生育决策的重要因素。生育成本包括直接成本和机会成本,前者指子女抚养的各种家庭开支,后者指女性选择生育而需要放弃其他机会的损失。相对于比较直观的生育直接成本,生育机会成本定义易,量化难,专门的测度成果较少,至今尚未形成权威的测量方法^[7-8]。但是,生育机会成本是生育成本不可忽视的部分。缺乏生育机会成本的考察,生育成本的分析将有失偏颇。

生育机会成本测量复杂的一个重要原因是涉及对已育女性“反事实”生育(不生育或少生育子女)状态下收入的预测,而构造“反事实状态”需要控制其他变量计算出生育女性与未育女性的收入差值。运用传统计量方法进行的收入预测对样本偏误和变量遗漏等有较高要求,预测误差较大。基于此,本文的研究重点是生育机会成本的测度及其对生育意愿和生育水平的影响。本文尝试运用机器学习算法测量女性因生育而导致的绝对机会成本和相对机会成本并考察生育机会成本的异质性,还将通过回归模型识别生育机会成本对生育意愿和实际生育数量的边际效应。研究结果显示,生育机会成本对二孩和多孩的生育意愿以及实际生育数量都有较为显著的影响,但由于异质性的存在,对生育率的抑制作用低于很多研究的

① 国务院新闻办公室第七次全国人口普查主要数据结果发布会,2021 年 5 月 11 日。

预期。

本文在已有研究的基础上对生育机会成本的分析进行了三方面的拓展：第一，运用在生育经济学领域还较少应用的机器学习算法估算了生育机会成本，在方法上丰富了生育机会成本测度的研究；第二，区分生育绝对机会成本和生育相对机会成本，有助于更充分地理解收入约束下生育机会成本可承受的程度；第三，考察了生育机会成本的异质性，有助于更深入地认识生育机会成本对不同收入水平、不同职业和不同教育水平的女性群体的不同作用。研究结果对从生育成本角度思考生育率发展趋势具有参考价值。但受限于数据，本文未能考察生育机会成本的动态变化。

二、文献综述

生育理论的经典文献中，生育决策涉及利他、利己的生育预期效用和生育成本的考虑。前者关乎生育的收益，后者关乎生育的负担^[9-10]。随着社会经济发展，生育预期效用越发不稳定，生育成本越来越高，很多学者认为成本约束成为主导生育决策的因素^[7,11-12]。

生育成本有多种类型。罗丽艳把生育成本分为家庭成员承担的内部成本和由全社会承担的外部成本^[13]。徐安琪认为，除了一些直接费用外，生育成本还应包含因生养孩子而损失的职业升迁或收入增加的机会（即“母职惩罚”）^[14]。谭雪萍把经济代价、机会代价和心理代价都包含在生育成本内^[15]。概括而言，生育成本既涵盖显性支出，也包括隐性损失。只要在生育问题上存在选择和取舍，生育机会成本就有可能存在。

女性因生育而引发的潜在损失可以发生在家庭内外的多个场域，但如果以女性初婚和初育平均年龄推算，多数女性的生育行为发生在进入劳动市场以后。基于此，对生育机会成本的讨论，不少是围绕女性职业发展和收入增长展开。有研究指出，生育导致女性职业中断^[16]。市场化条件下，女性很难平衡母亲身份与工作的冲突。李芬的研究发现，社会普遍忧虑女性工作、母职兼顾会影响其工作表现（即职场中的“瘢痕效应”）^[17]。循此逻辑，对生育机会成本的测度，多数是经济损失的测量。例如，讨论二孩政策遇冷时，朱兰把生育机会成本理解为女性为提高生育数量而放弃在计划生育政策下获得的社会经济福利^[18]。刘金菊通过回归模型估计出分年龄段的女性劳动参与率，据此编制女性工作预期寿命表，估算女性因生育而造成的工作年限损失，再把损失年限按照市场收入水平换算为收入损失^[19]。刘金菊“职业中断年数”的研究思路，对生育机会成本的估算很有启发意义，但技术层面上，其估计结果的可靠性受到回归模型的样本分布的前提假设、因果识别的模糊以及变量之间的内生性的影响。

国内目前还缺乏严谨的实证分析证实生育造成的职业间断会影响到女性收入水平。很多研究指出，生育会对女性的职业发展和收入产生负面影响^[18,20-22]，但事实并非绝对，也存在与该预期相悖的情况。阿圭罗（Agüero）和马克斯（Marks）使用6个拉丁美洲国家的数据发现，孩子数量对女性劳动参与率没有影响，如果将数据扩充到26个国家，基于全样本和中等收入国家的子样本可以得出生育对女性劳动参与率基本没影响，基于低收入国家和35岁以下女性的子样本则得出生育会降低女性劳动参与率^[23]。艾达（Adda）等人的研究发

现,20岁时,已生育女性的收入比未生育女性高5%,但到40岁时,未育女性的收入比已育女性高10%,工作中断导致一段时间没有收入来源是主要的生育成本^[24]。

国内研究方面,贾男等基于CHNS 2010数据发现,对于不同生育类型的女性,生育对收入的负面影响程度不同。对于外生生育类型(非意愿生育)的女性,生育当年的工资率(单位时间的工资水平)下降17.6%,并且生育只对当年工资有影响,前后几年的工资几乎都不受影响;而内生生育类型(理性生育)的女性,工资的下降在统计上不显著^[25]。智联招聘在2017年发布的《2016年职场妈妈生存状况调查报告》显示,认同生育后薪资待遇会出现下降的人数明显高于认为薪酬会上升的人数;但同时也有近1/4女性相信,产后重返职场对她们的收入水平没有影响,也有人生育后工资提升了^①。《2020年职场妈妈生存状况调查报告》则显示,职场已婚未育女性比已婚已育女性在升职加薪上面临更多的障碍,职业收入对家庭总收入的贡献比重比已婚已育女性更低^②。上述研究和调查反映,生育对女性收入水平的影响有可能受到职业差异、社会保障差异等多重异质性的制约。因此,分析生育机会成本时,需要考察不同人群生育机会成本的异质性。

三、方法与数据

1. 测算思路

生育机会成本虽然可以理解为女性因生育而导致的经济、受教育机会和工作晋升等多个维度的损失,但从量化的角度看,这些损失最终都可以体现在收入的变化上。因此,本文把生育机会成本视为一种“反事实”的生育行为的收入代价,即如果一个女性已经生育了 n 个子女,相比于其本人只生育 $n-1$ 个子女,收入水平的下降程度。假设女性生育 n 个子女后的年收入为 y_n ,该女性仅生育 $n-1$ 个子女的年收入为 y_{n-1} 。当 $y_{n-1} > y_n$ 时,生育发生了机会成本;当 $y_{n-1} < y_n$ 时,生育并未导致收入损失。

如何获得 y_{n-1} 是测度生育机会成本的最大挑战,原因是我们只能观测到事实(一个女性生育了 n 个子女后的年收入 y_n)而无法同时观测到反事实(该女性如果少生育一个子女后的年收入 y_{n-1})。生育机会成本的测度需要估计反事实状态的结果 y_{n-1} 。考虑到机器学习算法在预测“反事实的结果”方面有较强的实现能力^[26-27],本文的解决思路如下:首先基于微观数据构建训练模型,利用机器学习算法增强特征变量(解释变量)对响应变量(被解释变量)的预测能力;然后依据拟合优度(Rsquared)、平均绝对离差(MAE)和均方误差的开平方(RMSE)三个指标的参数结果,选择对响应变量预测能力最强的算法估算 y_{n-1} 。

2. 机器学习算法选择

机器学习算法众多,需要择优选取。首先,使用中山大学2016年中国劳动力动态调查(CLDS 2016)数据测试包含决策树、装袋法、随机森林、线性回归、梯度提升、支持向量机和人工神经网络七种机器学习算法的预测能力,把拟合优度作为评估的基准指标,选取出拟合优度最大的两种算法;其次,在拟合优度最大的两种方法中比较平均绝对离差,选取平

① 参见:https://www.sohu.com/a/73840140_119737
② 参见:<https://new.qq.com/rain/a/20200519AOIPFD00>

均绝对离差最小的算法作为测度方法，均方误差的开平方作为辅助评价指标。图 1 是三个指标的中位数数值和 95% 的置信区间。

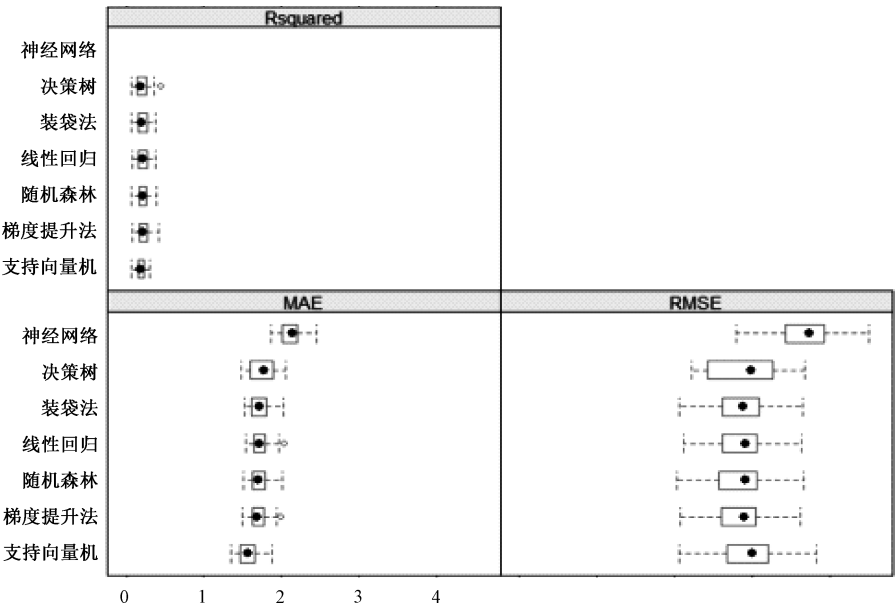


图 1 不同机器学习算法的预测效果评价

图 1 显示，拟合优度最高的两种算法分别是线性回归、随机森林；平均绝对离差最小的是支持向量机、梯度提升和随机森林；均方误差的开平方最小的是装袋法、梯度提升、线性回归和随机森林。如果考虑均方误差的开平方的均值，最小的是梯度提升和随机森林。综合而言，随机森林是最优算法，针对本文的研究问题，能够有效解决特征变量之间的相关性和过拟合等问题，提高预测精度。

3. 基于随机森林算法的生育机会成本测算

随机森林算法是一种基于决策树的集成学习算法。决策树是一种单一学习方法，被称为“基学习器”（base learner），多棵决策树组合成“森林”则构成了集成学习方法中的强学习器（strong learner），能够有效提升算法的预测能力。令随机森林模型：

$$y_n = f^*(\cdot) = \sum_{b=1}^B f^b(n, x, x_i, t_i, T, \lambda) / B$$

其中， x_i 和 t_i 分别是随机森林算法中按照“最小化残差平方和”原则得到的每棵决策树在第 i 层分裂时的特征变量和对应的分裂临界值。为解决过拟合问题，定义每棵决策树的终节点数目 T 为复杂性（通过“成本复杂性修枝”得到）， λ 为使用交叉验证法得到的对复杂性 T 的惩罚力度；响应变量为女性当前已生育 n 个子女的年收入为 y_n ， n 为核心特征变量， x 为其余特征变量； B 为随机森林中包含的决策树的数目。

通过算法训练和验证得到 x_i 、 t_i 、 T 和 λ 等参数结果后，令所有样本女性已生育子女数为 $n-1$ ，得到已经生育了 n 个子女（事实生育）的女性，假如生育 $n-1$ 个子女（反事实生育）的年收入为 y_{n-1} ，则生育的绝对机会成本可通过 $y_{n-1} - y_n$ 得到，相对机会成本可通过

$(y_{n-1} - y_n)/y_{n-1}$ 得到。生育相对机会成本可以结合生活负担水平衡量生育机会成本可以被承受的程度。

4. 数据来源和变量选择

分析数据来自 CLDS 2016。CLDS 2016 覆盖全国 29 个省、自治区和直辖市的 14226 个家庭和 21086 个个体。本文将研究对象限定为至少生育过一个子女、所有响应变量和特征变量均无缺失值的初婚女性，据此共得到 2679 个有效分析样本。其中，生育过男孩的样本有 2038 个，生育过女孩的样本为 1628 个。

响应变量为样本女性的全年收入，核心特征变量为生育行为，包括生育子女数量、生育男孩数量、生育女孩数量，其他特征变量包括了个人层面和家庭层面影响个人收入的重要因素（见表 1）。因素的重要性基于节点不纯度（IncNodePurity）标准确定^①。

表 1 变量赋值与描述性统计

变量	赋值方法	均值	最小值	最大值
收入	实际填写数（单位：万元）	2.85	0.08	20
生育子女数量	实际填写数（单位：人）	1.72	1	6
生育男孩数量	实际填写数（单位：人）	0.92	0	4
生育女孩数量	实际填写数（单位：人）	0.78	0	5
个人特征				
年龄	2016 - 出生年份	40.54	20	51
年龄的平方	$(2016 - \text{出生年份})^2$	1698.81	400	2601
户口类型	农业户口或现在居民户口但以前是农业户口 = 0，非农户口或现在居民户口但以前是非农户口 = 1	0.20	0	1
户口所在地	本村/居委会 = 1，本乡镇（街道）其他村居委会 = 2，本县（县级市、区）其他乡镇街道 = 3，本县区以外 = 4	1.50	1	4
受教育程度	初中及以下 = 1，普高、职高、中专和技校 = 2，大专和大学本科 = 3，硕博 = 4	1.44	1	4
健康水平	非常健康 = 1，健康 = 2，一般 = 3，比较不健康 = 4，非常不健康 = 5	3.71	1	5
就业单位类型	共有党政机关或人民团体或事业单位、国有企业、民营或私营企业、务农、自由职业者等 11 种类型，设定 10 个虚拟变量	7.79	1	11
居住地类型	城市 = 1，农村 = 0	0.37	0	1
居住地区	东部 = 1，中部 = 2，西部 = 3	1.87	1	3
家庭特征				
父亲户口类型	农业户口或现在居民户口但以前是农业户口 = 0，非农户口或现在居民户口但以前是非农户口 = 1	0.17	0	1
父亲受教育程度	初中及以下 = 1，普高、职高、中专和技校 = 2，大专和大学本科 = 3，硕博 = 4	1.16	1	3
母亲户口类型	农业户口或现在居民户口但以前是农业户口 = 0，非农户口或现在居民户口但以前是非农户口 = 1	0.14	0	1
母亲受教育程度	初中及以下 = 1，普高、职高、中专和技校 = 2，大专和大学本科 = 3，硕博 = 4	1.07	1	3
兄弟姐妹数量	实际填写数（单位：人）	3.09	0	14

注：引入年龄平方的变量，目的是平滑年龄与收入的倒“U”型关系。

① 节点不纯度指树模型中特征变量在节点分裂后可以使分裂后的两个子节点内部“纯度”提高的程度，可通过构造节点不纯度函数计算得到。

四、实证结果

1. 生育对收入水平的影响

图 2 显示的是影响收入水平的特征变量的重要性排序。居首的是受教育程度，即人力资本是影响收入水平的最为关键的因素。其次是年龄平方和年龄，反映出工作经验的积累对收入水平的重要性。值得注意的是，就业单位类型对收入水平的影响很小，可能的原因是单位招聘时已经综合考虑了申请人的受教育程度等个人特征，单位类型能够解释收入水平的信息已经被包含在其他变量之中。在全部 28 个特征变量中，生育子女数量的重要性排在前三分之一（第 8 位），说明生育行为对女性收入水平有一定影响。统计参数的符号显示，这种影响为负，即生育子女数量越多，女性年收入越低。基于图 2 得出的重要信息是，生育行为产生的对女性收入的“惩罚效应”不容忽视，但其并非是影响女性收入水平的最重要因素。这个结果也说明，对不同的女性群体，机会成本对生育影响存在异质性。

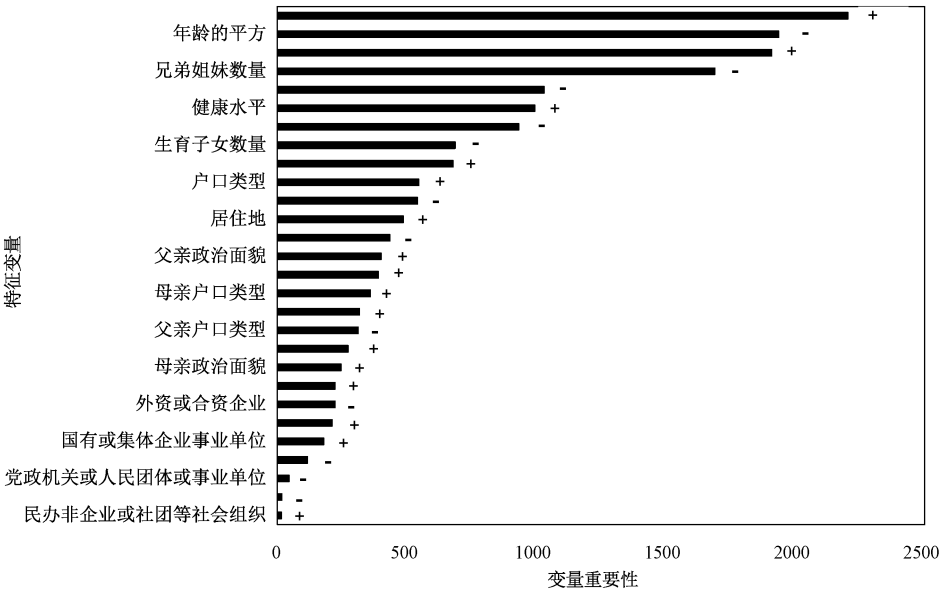


图 2 收入水平影响因素的重要性排序

说明：+和-表示特征变量对响应变量（收入）作用的方向。

2. 生育机会成本测算结果

图 3（a）展示了生育的绝对机会成本的估算结果，从中可得出一些基于目前资料文本难以直接讨论，但值得后续研究的发现：第一，一些女性会因生育而导致年收入损失，这已被大量的研究所证实；但也有女性因生育而导致年收入增加，可能的原因是部分职场母亲在生育后责任感增强，雇主看中更稳重、更有责任感的职场母亲，对其给予升职加薪^[28]。第二，相对于“0”值，生育绝对机会成本明显右偏，结合密度（落在某个收入区间的样本数的比重）分析，可以得出对大多数女性而言，生育绝对机会成本为正值。换言之，大部分初婚女性因为生育导致了年收入的下降，昭示生育惩罚效应和职业瘢痕效应的存在。第三，生育女孩的绝对机会成本与生育子女的绝对机会成本在总体分布上接近一致，而当生育绝对

机会成本大于 0 时, 生育男孩的绝对机会成本曲线与生育女孩的绝对机会成本曲线有较明显的偏离。生育男孩的绝对机会成本在 0—1.1 万元之间的比重大于生育女孩, 但当生育绝对机会成本大于 1.1 万元时, 生育女孩的绝对机会成本比生育男孩更大。从均值上看, 生育男孩的绝对机会成本 (0.28 万) 小于生育女孩 (0.49 万)。这或与女孩在成年前有更高的死亡率有关, 因此需要母亲在家更多的关心和照料^[29-30]。第四, 三条曲线显示出明显的数据离散度, 说明生育绝对机会成本在个人之间存在很强的异质性。

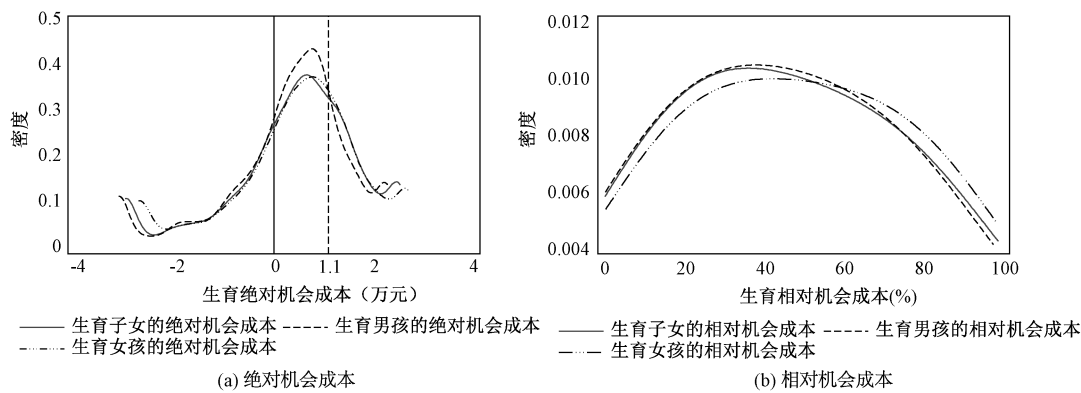


图 3 生育机会成本的性别差异

说明: 为便于图形展示, 图 (b) 不包括未受“收入惩罚”的样本, 即生育相对机会成本小于 0 的女性; 图 中线段为使用高斯核密度函数估计得到。

图 3 (b) 展示了生育的相对机会成本结果。相比于绝对机会成本, 相对机会成本更能反映预算约束条件下女性作生育决策时根据其收入状况对预期成本承担的意愿。图 3 (b) 显示, 有七成样本生育子女相对机会成本少于 50%, 有三成样本生育子女相对机会成本超过 50%, 高者甚至接近年收入。生育相对机会成本存在性别差异, 生育男孩相对机会成本超过 50% 的样本占生育男孩总样本的 29.69%, 生育女孩相对机会成本超过 50% 的样本占生育女孩总样本的 34.21%。生育男孩的相对机会成本在 0—60% 区间内的比例显著高于生育女孩的相对机会成本, 生育女孩的机会成本大于 60% 的比重要显著高于生育男孩。从均值上进行比较, 生育女孩的相对机会成本 (24.22%) 高于生育男孩 (18.77%), 这与生育的绝对机会成本一致。

3. 生育“收入惩罚”和生育保险

表 2 对比了女性生育后 (事实生育) 和生育前 (“反事实”生育) 的年收入。从生育后 (事实生育的女性) 的收入分布来看, 随着收入的增加, 收入较高的女性在女性群体中的比例不断减少, 年收入 5 万元以下的女性占了 85% 以上。与此相比, 年收入在 5 万元以上的 “反事实” 生育的女性比例没有显著增加, 说明生育行为对收入的影响主要集中在年收入 5 万元以下的女性群体。生育后收入小于 1.5 万元的女性比重明显比生育前更大, 相应地, 生育后收入在 1.5 万—5 万元之间的女性比重比生育前更小, 反映生育导致部分年收入在 1.5 万—5 万元之间的女性的工资下滑到 1.5 万元以下。

表 2 不同群体女性的收入比较				个,%
年收入区间	事实生育的女性		“反事实”生育的女性	
	样本数	比重	样本数	比重
1.5 万元以下	983	36.69	324	12.09
1.5 万—5 万元	1393	52.00	2031	75.81
5 万元以上	303	11.31	324	12.09
合计	2679	100	2679	100

图 4 以绝对生育成本为例，进一步呈现了“反事实”生育的女性年收入和生育机会成本的关系。散点分布的拟合线斜率为负，说明少生育一个子女时的年收入越高，生育绝对机会成本越低。以年收入 5 万元为分界线，就平均水平而言，年收入 5 万元以下的初婚女性生育的绝对机会成本大于 0；年收入 5 万元以上的初婚女性生育绝对机会成本小于 0。

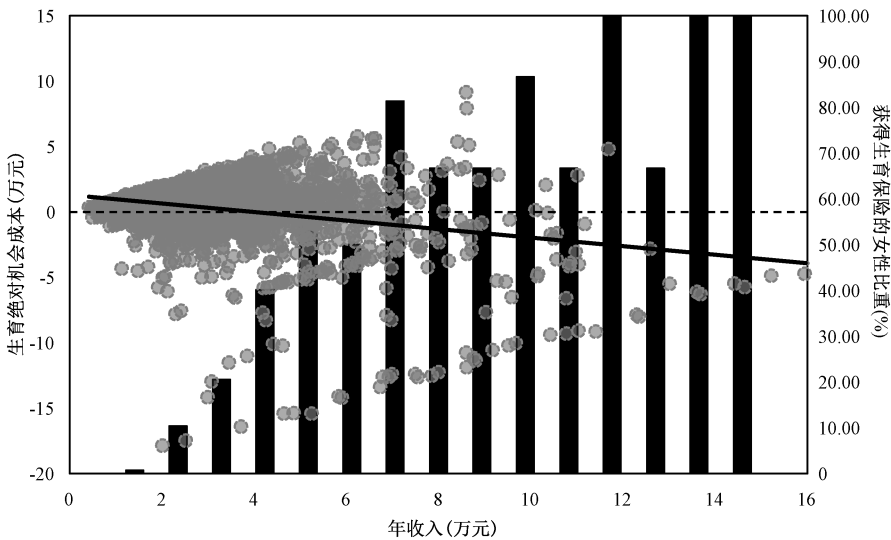


图 4 收入水平、生育绝对机会成本和生育保险

说明：生育绝对机会成本的散点和拟合线对应左侧坐标轴，条形图（生育保险覆盖率）对应右侧坐标轴。收入区间获得生育保险的女性比重=收入区间获得生育保险的女性人数/收入区间女性总人数。

为何低收入者会比高收入者承担更多的生育绝对机会成本，可以从生育保险制度的视角去理解。作为一种基本社会保障制度，生育保险保障妇女劳动者不会因为生育行为而被解雇，在因怀孕和分娩暂时中断劳动时，享受生育医疗待遇和生育津贴。低收入女性从事的行业和职业所需的人力资本积淀和替换门槛相对较低，雇主为控制劳动力成本，可能不会严格落实生育保险制度。如图 4 所示，收入越低的女性中，获得生育保险的比例越小，生育绝对机会成本越高。在 0—1 万元年收入区间的女性中，没有一人获得生育保险；1 万—2 万元年收入区间获得生育保险的女性仅占 1 万—2 万元年收入女性的 0.7%。与此相反，年收入在 13 万—14 万和 14 万—15 万元之间的女性，全部获得生育保险。可见，低收入女性承担更多生育机会成本，与缺乏生育保险福利有关。

我国生育保险的覆盖面仍然偏低。国家统计局公布的数据显示，2010 年全国生育保险

参保人数占总就业人口比例为 16.2%, 占城镇就业人口的比例为 35.6%; 2019 年, 两者分别提高到 27.6%和 48.4%, 远低于城镇职工基本养老保险和城镇职工基本医疗保险的参保率^[31]。由于生育保险费用由用人单位承担, 个人不用缴纳, 所以数量庞大的农民工和其他非正规就业的劳动人口, 几乎没有被生育保险覆盖。

不同教育程度和不同就业单位的女性生育机会成本差异也证实了生育保险的作用。从图 5 可见, 就平均水平而言, 教育程度低的女性, 生育的绝对机会成本和相对机会成本都很高。生育绝对机会成本和相对机会成本最低的是获得硕士和博士学位的女性。这部分女性群体大概率从事高收入工作, 职业门槛更高, 生育保险覆盖率也更高; 教育程度越低的女性, 从事低收入工作的可能性越大, 职业门槛越低, 社会保障越不完善, 面对越高的生育机会成本。

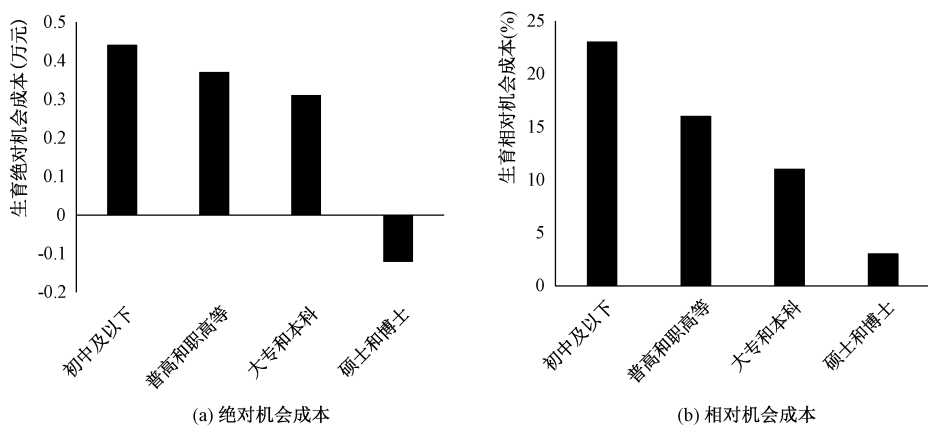


图 5 不同教育程度女性生育机会成本的差异

图 6 是按法人性质区分的用人单位的生育机会成本。相比于图 6 (b) 的用人单位, 图 6 (a) 的用人单位中不因为生育而减少了收入的女性比例更高, 因生育产生绝对机会成本的女性比例更低。公有制和私有制的用人单位女性生育绝对机会成本集中在 0.3 万—0.6 万元, 但自治和社会组织、自雇群体的女性生育绝对机会成本集中在 1 万—1.6 万元。生育绝对机会成本的单位差异与现有研究发现的生育保险参保率的职业人群差异类似^[32-33]。

图 6 (c) 和 (d) 显示, 生育的相对机会成本在不同单位类型分布并非和图 6 (a) 和 (b) 完全一致。私有制单位女性生育相对机会成本稍高于公有制单位女性 (图 6 (c)), 公有制单位女性生育的相对机会成本更为明显地集中在 0—10% 区间, 而私有制集中于 10%—20% 区间。自治和社会组织女性生育相对机会成本集中在 0—10% 的区间内, 并且在 0 上下有对称分布的特性; 但自雇群体的生育相对机会成本却更集中在 50%—80% 区间内 (图 6 (d))。相对于其他单位类型, 自雇群体一方面没有生育保险给予权益保障, 另一方面收入较低, 使得生育相对机会成本更高。

4. 生育机会成本对生育意愿和生育行为的影响

生育机会成本对生育意愿和生育行为的影响可以通过回归模型来考察。模型选择考虑了

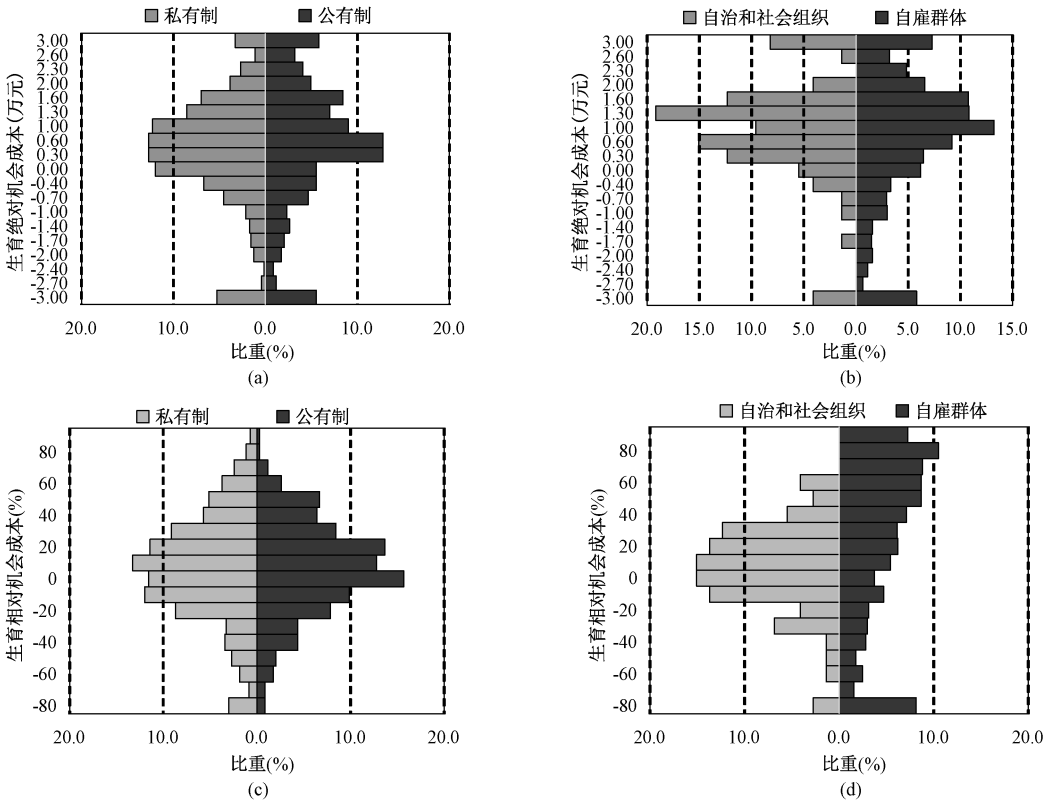


图 6 不同就业单位女性的生育机会成本对比

说明：公有制单位包括：党政机关或人民团体或军队、国有或集体事业单位、国有企业、集体企业四类；私有制单位包括：民营或私营企业、外资或合资企业；自治或社会组织包括：村居委会等自治组织和民办非企业或社团等社会组织；自雇群体包括：个体工商户、自由职业者、务农。

因变量类型：因生育意愿为二分类变量，采用 Probit 回归进行参数估计；实际生育数量为计数变量，采用泊松回归进行参数估计，使用负二项回归作稳健性检验。考虑到不同地区生育文化、经济发展等差异，回归中通过加入虚拟变量控制了地区固定效应，解决个体因地区异质性而产生的参数估计的不一致，提高参数估计的无偏性。

实证结果显示（见表 3），生育绝对机会成本在 5% 显著性水平上对二孩和多孩生育意愿呈现负向影响，生育绝对机会成本上升 1 单位，女性生育二孩和多孩的意愿分别下降 1.6% 和 1.5%。生育绝对机会成本提高会通过生育意愿显著降低女性的实际生育数量，泊松回归结果表明，在 5% 显著性水平上抑制实际生育数量，生育绝对机会成本上升 1 单位，实际生育数量下降 1.1%；负二项回归结果证实了基于泊松回归得到的结论依然成立^①。

表 3 结果还显示，生育机会成本并非是推动生育率下降的唯一重要因素，除了机会成本的作用外，还有部分控制变量对生育意愿和实际生育数量产生显著的抑制影响。例如，常用来衡量生育观念的受访者的户口类型和受教育程度对实际生育数量有显著的负向影响，表明

① 限于篇幅，没有汇报生育相对机会成本影响生育意愿和生育行为的结果。

表3 生育机会成本对生育意愿和生育行为的影响

变量	二孩生育意愿 (Probit 回归)	多孩生育意愿 (Probit 回归)	实际生育数量 (泊松回归)	实际生育数量 (负二项回归)
生育绝对机会成本	-0.111 ** (0.051)	-0.106 ** (0.053)	-0.006 ** (0.003)	-0.006 ** (0.003)
生育绝对机会成本的边际效应	-0.016 ** (0.007)	-0.015 ** (0.007)	-0.011 ** (0.005)	-0.011 ** (0.005)
年龄	0.012 (0.228)	-0.018 (0.211)	0.07 *** (0.009)	0.07 *** (0.009)
年龄的平方	-0.003 (0.003)	-0.002 (0.003)	-0.001 *** (0.000)	-0.001 *** (0.000)
受教育程度	0.167 ** (0.085)	0.147 * (0.082)	-0.076 *** (0.013)	-0.076 *** (0.013)
户口类型	0.027 (0.215)	0.004 (0.205)	-0.104 *** (0.025)	-0.104 *** (0.025)
户口所在地	-0.043 (0.075)	-0.037 (0.07)	0.03 *** (0.011)	0.03 *** (0.011)
父亲受教育程度	-0.108 (0.139)	-0.062 (0.141)	0.014 (0.021)	0.014 (0.021)
父亲户口类型	0.657 (0.472)	0.797 * (0.442)	-0.098 *** (0.028)	-0.098 *** (0.028)
母亲受教育程度	-0.179 (0.187)	-0.238 (0.19)	0.042 (0.032)	0.042 (0.032)
母亲户口类型	-0.379 (0.455)	-0.488 (0.432)	0.007 (0.037)	0.007 (0.037)
兄弟姐妹数量	0.101 (0.067)	0.095 (0.068)	0.008 ** (0.004)	0.008 ** (0.004)
居住地类型	0.041 (0.149)	-0.038 (0.157)	-0.147 *** (0.027)	-0.147 *** (0.027)
地区固定效应	√	√	√	√
常数项	2.602 (3.623)	3.125 (3.407)	-1.171 *** (0.168)	-1.171 *** (0.168)
样本量	782	819	2618	2618
虚拟 R ²	0.487	0.485	—	0.056

注：1. 二孩生育意愿赋值方法：在已经生育一个子女的前提下，准备再生育一个子女，赋值为1；在已经生育一个子女前提下，不准备再生，或未生育子女前提下，不准备再生，赋值为0。多孩生育意愿赋值方法：在已经生育了一定数量子女的前提下，准备再生育子女，赋值为1；在已经生育一个子女前提下，不准备再生，或未生育子女前提下，不准备再生，赋值为0。2. ***、**、* 分别表示1%、5%和10%的显著性水平，括号内为省份层面的聚类稳健标准误。

随着人口城市化和教育扩张等社会经济多方面的发展，除生育机会成本之外的生育观念的变迁对生育率下降的影响。

五、结论与讨论

人口再生产所产生的机会成本被视为造成生育率持续降低的重要原因，但对生育机会成本的认识大多停于感性。基于全国性的微观调查数据和机器学习算法，本文的实证分析得出如下发现：①生育对女性收入水平有影响，会对女性造成“收入惩罚”；②不同收入水平和教育程度的女性的生育机会成本存在显著差别，收入越高、教育程度越高的女性，生育机会成本越低；③生育机会成本在就业市场中的“惩罚效应”，主要体现在职业替换门槛较低的

低收入妇女身上；④生育机会成本的收入人群差异与生育保险参保率的职业人群差异有关，收入越低的女性，生育保险参保率越低，生育的机会成本越高；⑤生育机会成本降低了生育二孩和多孩的意愿，也显著降低了实际生育数量；⑥生育机会成本对生育意愿和生育行为的影响显著，但并非唯一决定因素。

本文的研究表明，生育机会成本的高低取决于生育保障制度，与女性的职业、收入密切相关。以生育保险为代表的社会保障制度通过多主体的成本分摊，成为降低女性生育机会成本的重要工具。因此，完善生育保障制度，特别是提高生育保险对低收入女性的覆盖率或是应对中国当前低生育率的一个可行政策方向。

生育机会成本对生育决策的影响虽然重要，但在第二次人口转变的时代背景下，其对生育水平的抑制作用不宜过分夸大。如同世界上很多国家，中国已经开始进入由后物质主义价值观主导的第二次人口转变，低生育率是大势所趋^[34]。第二次人口转变理论认为，虽然经济成本对生育率的降低有影响，但在更大程度上，生育率的下降缘于观念的传播和扩散。生育不仅受到经济条件制约，而且还受到价值观变迁的冲击^[35]。尽管学界对第二次人口转变理论的普适性还存在质疑，但本文的实证分析在某种程度上也支持这一理论判断。在多重因素共同作用下，减少家庭层面上的生育机会成本，构建“生育友好型”的家庭政策虽有助于缓解，但无法从根本上“反转”难以避免的低生育率趋势。

第七次人口普查公报数据显示，2020年中国总人口比2010年增长了5.38%，年平均增长率为0.53%，比2000年到2010年的年平均增长率0.57%下降了0.04个百分点（注：国务院办公厅第七次全国人口普查主要数据结果发布会，2021年5月11日），由“少子化”和低生育意愿导致的人口增速放缓的潜伏危机受到多方关注。我国人口总量大概率在未来几年就会到达峰值，然后进入缓慢的负增长。人口因素对经济、社会、民生诸多方面的严峻挑战将在人口负增长时代显现。学者们也越来越多意识到仅依靠生育政策已经很难逆转生育水平下降的走向^[36-37]。面对微观家庭层面“少子化”成为常态的局面，如何构建与新人口形势相匹配的公共政策体系，应当成为应对低生育率挑战的逻辑起点。

参考文献：

[1] 姚引妹, 李芬, 尹文耀. 单独两孩政策实施中堆积夫妇及其生育释放分析 [J]. 人口研究, 2014 (4): 3-18.

[2] 石智雷, 杨云彦. 符合“单独二孩”政策家庭的生育意愿与生育行为 [J]. 人口研究, 2014 (5): 27-40.

[3] 翟振武, 陈佳鞠, 李龙. 中国出生人口的新变化与趋势 [J]. 人口研究, 2015 (2): 48-56.

[4] 张丽萍, 王广州. 中国育龄人群二孩生育意愿与生育计划研究 [J]. 人口与经济, 2015 (6): 43-51.

[5] 靳卫东, 宫杰婧, 毛中根. “二孩”生育政策“遇冷”: 理论分析及经验证据 [J]. 财贸经济, 2018 (4): 130-145.

[6] 陈友华, 孙永健. “三孩”生育新政: 缘起、预期效果与政策建议 [J]. 人口与社会, 2021 (3): 1-12.

[7] 郑真真, 李玉柱, 廖少宏. 低生育水平下的生育成本收益研究——来自江苏省的调查 [J]. 中国人口科学, 2009 (2): 93-102.

[8] 段继红, 苏华山, 张成. 生育成本对二孩生育意愿的影响 [J]. 当代财经, 2020 (1): 17-26.

[9] LEIBENSTEIN H. An interpretation of the economic theory of fertility: promising path or blind alley? [J]. Journal of Economic Literature, 1974, 12 (2): 457-479.

- [10] 加里·斯坦利·贝克尔. 家庭论 [M]. 北京: 商务印书馆, 1998: 151-153.
- [11] 王志章, 刘天元. 生育“二孩”基本成本测算及社会分摊机制研究 [J]. 人口学刊, 2017 (4): 17-29.
- [12] 茆长宝, 穆光宗. 国际视野下的中国人口少子化 [J]. 文化纵横, 2018 (4): 16-16.
- [13] 罗丽艳. 孩子成本效用的拓展分析及其对中国人口转变的解释 [J]. 人口研究, 2003 (2): 47-54.
- [14] 徐安琪. 孩子的经济成本: 转型期的结构变化和优化 [J]. 青年研究, 2004 (12): 1-8.
- [15] 谭雪萍. 成本—效用视角下的单独二胎生育意愿影响因素研究——基于徐州市单独家庭的调查 [J]. 南方人口, 2015 (2): 1-12.
- [16] 张霞, 茹雪. 中国职业女性生育困境原因探究——以“全面二孩”政策为背景 [J]. 贵州社会科学, 2016 (9): 150-154.
- [17] 李芬. 二孩政策下生育对女性职业发展的瘢痕效应——基于全国十二城市的实证研究 [J]. 河北学刊, 2018 (6): 193-199.
- [18] 朱兰. 生育政策、机会成本与生育需求——中国综合社会调查的微观证据 [J]. 西北人口, 2020 (2): 90-101.
- [19] 刘金菊. 中国城镇女性的生育代价有多大? [J]. 人口研究, 2020 (2): 33-43.
- [20] 宋健, 周宇香. 中国已婚妇女生育状况对就业的影响——兼论经济支持和照料支持的调节作用 [J]. 妇女研究论丛, 2015 (4): 16-23.
- [21] 张抗私, 谷晶双. 生育对女性就业的影响研究 [J]. 人口与经济, 2020 (5): 19-29.
- [22] 杨菊华, 卢瑞鹏. 性别红利: 理论意涵、基本特征与社会效应 [J]. 山东社会科学, 2021 (3): 51-59.
- [23] AGÜERO J M, MARKS M S. Motherhood and female labor force participation: evidence from infertility shocks [J]. The American Economic Review, 2008, 98 (2): 500-504.
- [24] ADDA J, DUSTMANN C, STEVENS K. The career costs of children [J]. Journal of Political Economy, 2017, 125 (2): 293-337.
- [25] 贾男, 甘犁, 张劼. 工资率、“生育陷阱”与不可观测类型 [J]. 经济研究, 2013 (5): 61-72.
- [26] 陈云松, 吴晓刚, 胡安宁, 贺光烨, 句国栋. 社会预测: 基于机器学习的研究新范式 [J]. 社会学研究, 2020 (3): 94-117.
- [27] 胡安宁, 吴晓刚, 陈云松. 处理效应异质性分析——机器学习方法带来的机遇与挑战 [J]. 社会学研究, 2021 (1): 91-114.
- [28] 祝红艳. 生育对女性职业发展有何影响——智联招聘特别调查 [J]. 职业, 2009 (16): 36-36.
- [29] AHMED F A. Gender difference in child mortality [J]. The Egyption Population and Family Planning Review, 1990, 24 (2): 60-79.
- [30] CHOWDHURY R, TANEJA S, MAZUMDER S, BHANDARI N, STRAND T A. Gender differences in infant survival: a secondary data analysis in rural North India [J]. BMJ Open, 2017, 7 (8): e014179.
- [31] 张盈华. 完善中国流动人口参保机制 [R]. 国际劳工组织中国和蒙古局, 2019.
- [32] 庄渝霞. 生育医疗待遇: 政策演变、人群差异及优化策略 [J]. 社会科学, 2017 (12): 79-89.
- [33] 盛亦男, 童玉芬. 生育政策调整对女性劳动力供需的影响研究 [J]. 北京社会科学, 2018 (12): 96-104.
- [34] 於嘉, 谢宇. 中国的第二次人口转变 [J]. 人口研究, 2019 (5): 3-16.
- [35] ZAIDI B, MORGAN S P. The second demographic transition: a review and appraisal [J]. Annual Review of Sociology, 2017, 43 (1): 473-492.
- [36] 陆杰华. 人口负增长时代: 特征、风险及其应对策略 [J]. 社会发展研究, 2019 (1): 21-32.
- [37] 张现苓, 翟振武, 陶涛. 中国人口负增长: 现状、未来与特征 [J]. 人口研究, 2020 (3): 3-20.

[责任编辑 刘爱华]