

地方政府评价与育龄群体二孩生育决策 —— 基于CFPS数据的实证分析

[摘 要] 2016年1月1日，“全面二孩”生育政策在全国范围内正式实施。通过聚焦育龄群体对“全面二孩”政策的预期生育行为，利用2014、2016、2018年中国家庭追踪调查（CFPS）数据，采用混合截面双重差分（DID）和倾向得分匹配（PSM）相结合的方法，估计了对地方政府的评价如何影响育龄群体对“全面二孩”政策的预期生育反应。实证结果发现育龄群体会对未来生育政策做出预期反应，育龄群体对政府的信任是一种策略性信任，对地方政府评价越高越会对未来生育政策做出积极反应从而进行预期生育行为。异质性分析结果显示，18至29岁和东中部地区育龄群体的预期行为受到对地方政府评价的影响更甚。政府如果希望未来政策更快更好达到效果，就不能忽视政策颁布前与群众就政策相关信息进行沟通、宣传与预期引导，且应该重视长期政策信誉建设。

[关键词] 生育意愿；生育行为；全面二孩政策；公众预期；人口政策；政府信誉

[中图分类号] C913.7 [文献标识码] A [文章编号] 1002-8129 (2022) 11-0053-15

一、引言

在我国，生育行为具有一定的特殊性，其在很大程度上受到生育政策的影响。计划生育政策作为我国一项基本国策，一直以来都是公众生育行为决策的风向标，其对中国人口问题和经济发展的积极作用值得肯定，但也带来了一系列诸如人口老龄化、少子化等负面问题。为此，政府对计划生育政策进行了部分调整，分别在2013、2014、2016年颁布实施“双独二孩”“单独二孩”“全面二孩”政策，但政策效果并不十分理想（乔晓春，2015；陈友华，孙永健，2021）[1][2]。

学术界在既有计划生育政策背景下围绕着生育决策的影响因素进行了深入的研究探讨，但研究多集中于个人特征（如年龄、健康状况、收入、受教育状况、婚姻状况等）（陈建新，王莉君，2021；邓敏，陈宝璘，2020；王一帆，罗淳，2022）[3][4][5]、家庭社会特征（如父母支持、社会养老保险、社会医疗保险、房价等）（李婉鑫，杨小军，杨雪燕，2021；阮荣平，焦万慧，郑风田，2021；孟颖颖，王静，张孝栋，2020）[6][7][8]、文化观念特征（如性别平等、“养儿防老”、工作—家庭冲突等）

（胡荣，林彬彬，2020；王国军，高立飞，2021）[9][10]，在一定程度上忽视了宏观层面政府与公众在生育决策形成过程中的政策信息互动。要全面分析公众生育行为决策，就必须把握生育政策本身在其中起到的作用，不能忽视公众在政策未明确颁布阶段会如何做出预期和行为反应。

公众对政府的信任是政府用来制定和执行具有约束力的政策所依赖的政治资源，政府信任可以降低交易成本，即政府与公众就生育政策达成一致的成本，信任能增加政策被接受的可能性（Chanley, Rudolph, Rahn, 2001）[11]，从而，政府可以花费更少的努力让公众响应生育政策。现有研究认为对地方政府的评价在一定程度上代表了群众对地方政府的信任程度，尽管许多因素都可能降低公众对政府的信任程度，但最终都可以反映到政府绩效上，政治信任来源于政府绩效以及信任主体对这种绩效的评价（Michler, Richard, 2001；吴华根，2008）[12][13]。

理性预期理论表明个人能够有效地利用自己已掌握的信息形成预期（D'Haultfoeuille, Gaillac, Mauriel, 2021）[14]，而政策颁布之前既已形成的对政府的信任程度，正是个人据此形成政策预期的重要信息。具体到计划生育政策变化对人们生育行为的影响方面，2015年12月21日上午，十二届全国人大常委会第十八次会议初次审议了《中华人民共和国人口与计划生育法修正案（草案）》，草案提出，“全面二孩”政策自2016年1月1日起施行。在此之前，相关部门曾放出未具体到细节的“全面二孩”政策相关消息。面对一项已经释放模糊消息但未明确实施的政策，公众将做出何种反应，以及何种因素影响公众预期，正是我们关心的问题。

本文以2016年1月1日正式实施的“全面二孩”政策为切入点，根据在二孩政策明确颁布之前有迹可循的政策消息“放风”时间点，将“全面二孩”政策具体颁布时点之前的消息“放风”阶段看作政策模糊阶段，将“全面二孩”政策颁布之后视为政策明确阶段，研究发现对政府评价高的个人会在政策模糊阶段对政府未来政策产生积极预期，而对地方政府评价低的人会对政府未来政策产生消极预期，而这种预期会影响决策行为，对于全面二孩政策而言，其直接表现为是否提前为生育二孩进行准备，并在政策正式颁布到10个月孕期结束之前存在生育行为。本文利用2014年、2016年、2018年的家庭追踪调查（CFPS）数据，首先采用Probit模型检验育龄群体对生育政策是否存在预期行为，即在政策正式颁布实施之后到政策生效不满10个月期间生育二孩。实证发现其预期行为并不显著，研究发现该结果的产生可能是因为平均效应掩盖了异质性，于是将育龄群体分为高政府评价组和低政府评价组，实证分析发现高政府评价组的二孩生育预期行为显著，而低政府评价组的二孩生育预期行为不显著。随后，采用混合截面双重差分法（DID）实证分析对地方政府评价如何影响育龄群体对生育政策的预期行为，这也是本研究主要关注的问题，研究发现育龄群体对地方政府评价越高越会对未来生育政策做出积极反应，

从而进行预期生育行为。

本文可能有以下几点贡献和创新：首先，将预期理论引入到政府生育政策，对预期理论进行了拓展；其次，从地方政府与公众互动视角，对二孩生育决策影响因素进行了实证分析；最后，已有实证研究多用生育意愿来衡量生育水平，忽视了生育意愿并不一定会转化为生育行为（张冲，李想，2020）[15]，且将“无政策条件下理想子女数”和“期望生育子女数”作为被解释变量忽视了中国严格执行计划生育政策这一背景，并不能很好衡量真实生育水平（风笑天，2017）[16]，本研究将实际二孩生育行为作为被解释变量，更准确地衡量二孩生育决策。

二、政策背景与理论假设

（一）政策相关背景

2021年5月31日，中共中央政治局召开会议，会议指出：“进一步优化生育政策，实施一对夫妻可以生育三个子女政策及配套支持措施。”同年8月，全国人大常委会会议表决通过了关于修改人口与计划生育法的决定，修改后的人口计生法规定，国家提倡适龄婚育、优生优育，一对夫妻可以生育三个子女。

根据国家统计局数据显示，2000年，我国60周岁以上老年人口占到总人口的10%，按照国际通行标准，我国人口年龄结构已进入老龄化阶段。一段时间以来，我国人口少子化和老龄化问题严重，为了解决这一问题，2011年11月、2013年11月各省份分别开始实施“双独二孩”“单独二孩”，但政策效果并不理想。此后，何时全面开放二孩及其具体政策开始引发热议，2015年3月国务院政府工作报告中，延续多年的“坚持计划生育基本国策不动摇”提法被删除，且李克强总理在记者招待会上回答关于人口政策提问时表示，“去年已经开始实施单独二孩政策，现在正在推进，也正在进行全面的评估”。同年5月，时任卫计委新闻发言人在卫计委例行发布会上表示，“单独二孩不是句号，往下走的目的就是要促进人口的长期均衡发展”。同年7月，据《第一财经日报》透露，包括卫计委在内的相关方面已经着手就全面放开二孩政策进行评估和推进，如果顺利，年内就会开始实施。之后，卫计委计划生育基层指导司在回应记者关于“何时全面放开二胎”的问题时表示，要积极做好进一步调整完善生育政策的研究论证工作，同时强调了“正在按中央的要求抓紧推进有关工作”。这些信息无疑透露着同一个信号：全面二孩政策即将来临。2015年10月，党的十八届五中全会明确提出“全面实施一对夫妇可以生育两个子女的政策”。随后，2015年12月21日，十二届全国人大常委会第十八次会议初次审议了《中华人民共

和国人口与计划生育法修正案（草案）》，草案提出，本修正案自2016年1月1日起施行。对于全面二孩的政策效果，一种观点认为，全面二孩政策符合预期、政策效果明显（原新，高瑗，2017）[17]。另一种观点认为，政府相关部门对全面二孩政策实施效果的乐观判断缺乏可靠的依据（风笑天，2020）[18]，政策效果并不理想（任远，2017）[19]。虽然在实施全面二孩政策之后，受到政策影响，2016年人口出生率达到自2001年以来最高，出生人数比2015年全年出生人数增加171万。但自2017年以来，我国出生人口数量已连续5年下降。理清在当下低生育率时代，人们的生育决策受何种因素影响、生育政策如何更好发挥作用，是政府部门急需解决的问题，也是学术界长期关注的问题。

（二）理论假设

个人如何形成对不确定的未来结果的信念对于理解个人决策至关重要。“计划行为理论”认为个人行为意向受到态度、主观规范和知觉行为控制的影响，笔者认为高地方政府评价增强了个人二孩生育行为的知觉行为控制，即对地方政府评价越高的人越认为自己能够掌握的生育信息和资源更多、受到的阻碍越少。生育行为发生后往往有许多不确定性，国务院提出完善生育休假制度、给予税收优惠、住房支持和推进教育公平等一系列措施，这些措施有利于减轻生育二孩的负担，在一定程度上减少了孩子出生后的经济、教育、社会不确定性，但地方政府如何落实尚不明确。公众对政府表现的预期会影响他们对政策的态度及行为决策，公众预期的形成依赖于对以往事实进行回顾。政府以往的工作业绩与之后的工作业绩有很强的正相关关系（Meier，O'Toole，2002）[20]，公众对地方政府评价越高，对地方政府以往工作成效越肯定，就会认为地方政府能够一如既往落实中央所制定的政策和做出的承诺，更倾向于认为政府在之后的生育政策执行工作中会表现出色，甚至在细则方面做得更好。

当然，人们的预期不一定是完全理性的，预期往往会受到信息约束，对于个人而言，他们很少有充足的信息和能力去评估一项中央政策是否合理以及对已有生活造成何种利弊影响，个人更可能会依靠政治信任来补偿信息约束（Hetherington，2006）[21]。缺乏政府信任会导致群众抵制或不响应政府所颁布的政策（Trudinger，Steckermeier，1998）[22]，低信任度更会创造出一种让政策更难实施的环境，同时公众的支持会影响到政策的实施（Hetherington，1998）[23]。反之，高信任度使得政府政策获得意外的支持。结果导向政府信任理论认为公众对政府政策的信任依赖于其短期政策结果对个人是否有利，过程导向政府信任理论认为，公众更关心政府政策执行过程是否公平（孟天广，2022）[24]。这两种导向的政府信任分别以自利结果和过程公平为信任条件，当公众对政府表现评价高时，他们就倾向于相信政府，这种信任是一种策略性信任，反过来又会成为公众行为与政府政策达成一致的前提条件（孟天广，2022；Levi，Sacks

，Tyler，2009）[24][25]。但在传统文化和信念的影响下，公众可能对政府无条件信任，出于对政府的信任，即使公众的利益和要求没有得到完全满足，即使这些决定的后果被认为是有风险或不利的，他们也倾向于支持政府颁布的政策、认同政府的决定（Gabriel，Trudinger，2011）[26]。从二孩生育政策来看，策略性政府信任表现为育龄群体对政府未来二孩生育政策进行预期，预期政策结果对自己有利便会进行二孩生育行为，无条件的政府信任表现为不管二孩生育政策是否有利于自己，政府明确颁布二孩政策之后，无条件响应政府号召进行二孩生育行为。根据国际知名公关咨询公司Edelman发布的“Edelman Trust Barometer 2021”显示，中国政府信任度排名第一，这项报告是在2021年10月19日至11月18日对全世界28个国家的3.3万多名受访者进行的实地调查。在低生育率时代，在“全面二孩”政策号召下，育龄群体无条件信任政府的直接表现就是进行二孩生育行为，且这种行为具有持续性，不因生育政策配套措施是否完善、政府表现是否优秀、政策结果是否有利于个人而改变。

综上所述，笔者认为全面二孩生育政策的实际影响可能在不同群体间存在差异，因此本文提出以下实证假说以验证我们的理论猜测：

假说1：育龄群体会对生育政策变化的实际效果进行预期，因此生育行为反应并不一定完全符合政策设计目标。

假说2：全面二孩政策实施之后，育龄群体的二孩生育决策可能取决于其对地方政府的评价，高评价个体对政策的响应程度更高，低评价个体的响应程度则较低。

三、实证设计

（一）数据来源及样本特征

本文数据来源于2014、2016、2018年中国家庭追踪调查（CFPS）数据。CFPS样本覆盖25个省、自治区、直辖市，目标样本规模为16000户，调查对象包含样本家户中的全部家庭成员，CFPS调查问卷共有社区问卷、家庭问卷、成人问卷和少儿问卷四种主体问卷类型，是一项全国性、大规模、多学科的社会跟踪调查项目。CFPS家庭数据包含了受访者的子女数量以及子女的出生日期、性别等数据，CFPS成人数据中则包含了受访者的基本信息和政治态度等数据。本文利用三次连续的CFPS数据得到了两组混合截面数据：其一为在2018年家庭数据与2016年成人数据中同时出现的样本（实验组），其二为在2016年家庭数据和2014年个人数据中同时出现且未在2018年家庭数据中出现的样本（对照组）。经

数据处理之后，样本规模为3495人，其中实验组2986人，对照组509人，样本覆盖全国21个省（青海省和台湾地区数据缺失）、4个直辖市、5个自治区。

在2011年11月和2013年12月，我国分别实施“双独二孩政策”和“单独二孩政策”，为了剔除“双独”和“单独”生育政策影响、聚焦在二孩政策出台前后育龄人群的二孩生育行为，本文选取CFPS2014、2016和2018年的数据进行观察，在实验组中取所有一孩出生在2016年1月前的受访者数据，而在对照组中，则保留了所有一孩出生在2015年1月之前的受访者数据。“全面二孩”政策具体实施颁布时间为2016年1月1日，孕期为280天，取9个月，将2016年1月1日至2016年9月30日进行二孩生育行为看作育龄群体对该政策预期产生的行为决策。此外，对于来自同一个家庭的受访者，利用stata随机保留在数据中第一位出现且符合要求的受访者。从生育行为角度，对数据进行清洗和处理，选取了年龄为18至49岁的育龄人群，剔除了解释变量和控制变量的缺失值。

（二）相关变量解释

1. 被解释变量。本文被解释变量为二孩生育行为，在所观察时间段内生育二孩，该变量赋值为1，未生育则为0。目前研究主要将“无政策约束下理想子女数”和“期望子女数”作为被解释变量，相较于生育意愿，实际生育行为更能真实有效反映生育决策的影响因素。本文将成人数据中孩子1、孩子2、孩子3出生年月按大小进行排序得出不同年份的二孩生育情况。
2. 核心解释变量。核心解释变量为对地方政府的评价和是否受到二孩生育政策冲击的交互项，其中，对地方政府的评价来自于调查中“您对去年本县/县级市/区政府工作的总体评价”这一问题，受访者被要求对本县市政府在上一年工作由好到坏给出从1到4四个等级的评价，“有很大成绩”为1、“有一定成绩”为2、“没有多大成绩”为3、“没有成绩”为4。此外，若受访者认为当地政府的工作“比之前更糟了”，则政府评价变量会被记为5，因“比之前更糟了”与之前4个评价等级口径不一致，在基本回归中先删除此回答，放在稳健性部分再去讨论。
3. 主要控制变量。借鉴以往关于生育意愿及政府信任的研究，本文将控制变量按照个人特征、家庭社会特征、地区特征、政府特征5个层面分类设定。个人特征包括主观收入、年龄、性别、最高学历；家庭社会特征包括是否参加养老保险、是否参加医疗保险、一孩性别；地区特征主要是根据我国现行

地域划分标准将数据样本中30个省（直辖市、自治区，青海省、台湾地区数据缺乏）划分为东、中、西部地区以及居住地城乡属性纳入回归模型；政府特征包括中国腐败问题严重程度、中国住房问题严重程度、中国社会保障问题严重程度、中国教育问题严重程度、中国医疗问题严重程度、到政府办事是否受到拖延推诿、与政府干部冲突、是否遭遇政府不合理收费。具体赋值与描述性统计如表1所示。

(表1)		描述性统计			
变量	变量描述	样本数	平均值	最小值	最大值
对地方政府评价	对去年“本县/县级市/区政府工作的总体评价”打分，详细赋值见上文核心解释部分	3495	2.558	0	4
主观收入	受访者对个人收入打分，5分最高	3495	2.464	1	5
年龄	对照组和实验组年龄分别以2014年和2016年为基准	3495	37.18	18	49
最高学历	未上学为1，小学为2，初中为3，高中为4，大学专科为5，大学本科为6，研究生及以上为7	3495	3.294	1	7
性别	女性赋值为0，男性赋值为1	3495	0.586	0	1
健康状况	1为非常健康，2为很健康，3为比较健康，4为一般，5为不健康	3495	2.779	1	5
养老保险	1为参加养老保险，0为没有	3495	0.680	0	1
一孩性别	女性赋值为0，男性赋值为1	3495	0.583	0	1
医疗保险	1为有医疗保险，0为没有	3495	0.914	0	1
城乡虚拟变量	1为城镇户籍，0为农村户籍	3495	0.608	0	1
东部地区	1为属于东部地区，0为不属于	3495	0.470	0	1
中部地区	1为属于中部地区，0为不属于	3495	0.296	0	1
西部地区	1为属于西部地区，0为不属于	3495	0.223	0	1
政府拖延程度	1为曾经遭遇过，3为曾经听说过，5为既没听说且没遭遇过	3450	4.298	1	5
与政府干部发生冲突	同上	3450	4.832	1	5
政府不合理收费	同上	3450	4.667	1	5
中国住房评价	共分为0到10分，0代表该问题不存在，10代表该问题非常严重。分数越高代表受访者认为该问题越严重	3450	6.329	0	10
中国社保评价	同上	3450	6.293	0	10
中国腐败问题评价	同上	3450	7.056	0	10
中国教育问题评价	同上	3450	6.540	0	10
中国医疗问题评价	同上	3450	6.586	0	10

（三）模型设定和结果分析

2016年1月1日，全面二孩政策在全国范围内正式实施，本研究将其视为一项准自然实验，对假说1进行检验，考虑到被解释变量是虚拟变量，本文选择probit模型进行回归分析：

$$y=\alpha_1+\alpha_2treat+X_1+\varepsilon_1 \tag{1}$$

其中被解释变量y为是否生育二孩的虚拟变量，treat表示对照组和实验组的虚拟变量，对照组和实验组具体解释参考本文数据来源部分。对于来自对照组的样本，值恒为0，而对于实验组样本，其值为1。根据平均孕期280天，对于实验组样本，观察其在2016年1月1日至2016年9月30日的二孩生育行为，若生育二孩则y为1，未生育二孩则y为0，对于对照组样本，若其在2015年1月1日至2015年9月30日生育二孩，则y赋值为1，若未生育二孩，则y赋值为0。X1为一系列控制变量，其中包括个人基本特征和家庭社会特征，个人基本特征控制变量主要包括主观收入、年龄、最高学历、性别、健康状况，家庭社会特征包括一孩性别、是否参加养老保险、是否参加医疗保险，除此之外还控制了对县市政府的评价。

本研究主要采用probit模型对本文假说1展开实证回归。模型（1）回归结果如表2所示，从第（1）（2）列可以看出，尽管有一定正向促进作用，育龄群体并未通过采取二孩生育行为对政策产生显著反应。这说明公众对政策实际效果的预期不足，进而产生响应不足的结果。更进一步，将公众预期看作是一致地过分简单化了个人预期，本文认为对二孩生育政策预期的平均效应掩盖了异质性，个人预期具有异质性，这种异质性体现为不同个体对同一政策预期的不一致（张晓晶，曾一巳，邵兴宇，2021）[27]。预期产生的前提是社会变化会对公众利益产生影响且该社会变化能够被公众察觉（江世银，2017）[28]。公众对政府评价越高，政治参与度会越高，从而更容易关注到政府颁布的政策。参考Ashraf、Bau、Nunn、Voena（2020）[29]的成果，本研究扩展了模型（1），将对地方政府评价高的育龄群体和对地方政府评价低的育龄群体分组进行分析，估计高地方政府评价组和低地方政府评价组对全面二孩政策的预期是否有差异，结果如第（3）（4）列所示，高地方政府评价显著提高了二孩生育行为，对于低地方政府评价育龄群体而言，没有显著差异。这与假说2的结果相符。

(表2) 不同政府评价群体二孩生育行为影响因素回归结果

变量	整体	整体	高政府评价群体	低政府评价群体
	(1)	(2)	(3)	(4)
treat	0.187 (0.118)	0.167 (0.118)	0.532*** (0.203)	-0.115 (0.148)
主观收入	0.009 (0.044)	0.005 (0.044)	0.015 (0.061)	0.028 (0.061)
年龄	-0.056*** (0.006)	-0.057*** (0.006)	-0.063*** (0.009)	-0.048*** (0.009)
最高学历	-0.015 (0.034)	-0.021 (0.034)	-0.020 (0.044)	-0.043 (0.054)
性别	0.183** (0.085)	0.185** (0.086)	0.048 (0.114)	0.359*** (0.129)
健康状况	-0.049 (0.038)	-0.045 (0.038)	-0.089* (0.053)	-0.009 (0.055)
养老保险	-0.084 (0.087)	-0.097 (0.088)	0.088 (0.127)	-0.252** (0.123)
一孩性别	-0.254*** (0.079)	-0.256*** (0.079)	-0.247** (0.107)	-0.294** (0.116)
医疗保险	0.552*** (0.202)	0.556*** (0.202)	0.398 (0.270)	0.780** (0.312)
城乡虚拟变量	-0.070 (0.086)	-0.065 (0.086)	-0.093 (0.117)	-0.025 (0.125)
县市政府评价		-0.082* (0.047)	-0.001 (0.071)	-0.098 (0.101)
常数项	-0.189 (0.325)	0.080 (0.361)	0.009 (0.492)	-0.240 (0.592)
观测样本数	3601	3601	1921	1786

注：括号内为稳健标准误，*、**、***分别表示在10%、5%和1%水平上显著。

不过高评价组与低评价组个体也可能在生育行为上本来就存在系统性差异，因此本研究接下来通过更严格的因果识别方法进一步检验假说2，这也是本研究主要关注的假说，搞清楚育龄群体对地方政府评价如何影响二孩生育预期决策。本文借鉴Duflo(2001)[30]、李明和郑礼明(2021)[31]的估计方法，建立了以下混合截面双重差分模型：

$$y_{it} = \beta_1 + \beta_2 \text{treat}_{it} + \beta_3 \text{evaluation}_{it} + \beta_4 \text{treat}_{it} * \text{evaluation}_{it} + X_{it} + D_{it} + \varepsilon_{it}$$

(2)

其中， i 为18至49岁的育龄个体，跟前文一样，被解释变量 y_{it} 表示第 i 个育龄个体是否生育二孩的虚拟

变量，下标 t 表示是否处于实验组。 $treatit$ 表示对照组和实验组的虚拟变量，对于来自对照组的样本，值恒为0，而对于实验组样本，其值为1。对于对照组样本，若其在2015年1月1日至2015年9月30日生育二孩，则 yit 赋值为1，若未生育二孩，则 yit 赋值为0。相似地，对于实验组样本，观察其在2016年1月1日至2016年9月30日期间的二孩生育行为，若生育二孩则 yit 为1，未生育二孩则 yit 为0。 $evaluation$ 为政府评价变量，数据来源于CFPS2014（对照组）和CFPS2016（实验组）的成人调查数据。 Xit 是一系列控制变量， Dit 为地区固定效应，本文按照经济发展情况，将数据中各个省级行政区划分为东部、中部、西部地区。本研究重点关注交互项系数 β_4 ，该系数衡量了育龄群体对地方政府的评价多大程度影响育龄群体的二孩生育行为。

本文的模型设计不同于传统的双重差分模型研究政策实施前后实验组和对照组二孩生育行为差异，而是研究实验组和对照组在不同生育政策下二孩生育行为差距是否受公众对地方政府评价影响，以及全面二孩生育政策模糊阶段和明确阶段二孩生育行为是否受公众对地方政府评价影响的比较。

结果如表3（1）列所示，回归结果表明，在控制住对照组和实验组的情况下，其与对地方政府评价的交互项系数 β_4 仍然显著为负，说明对地方政府评价显著影响育龄群体二孩生育预期决策行为，对地方政府评价越高，越会对地方政府配套措施及表现产生积极预期从而提前进行二孩生育行为。表3（2）列是增加东部、中部、西部地区固定效应的结果，回归结果与基本回归结果相似。

(表3) 政府评价对二孩生育预期行为的影响

变量名称	(1)	(2)	(3)	(4)
treat	0.095*** (0.033)	0.093*** (0.033)	0.094*** (0.033)	0.094*** (0.033)
县市政府评价	0.016 (0.011)	0.017 (0.011)	0.017 (0.011)	0.018 (0.011)
treat*县市政府评价	-0.031*** (0.012)	-0.031*** (0.012)	-0.032*** (0.012)	-0.032*** (0.012)
主观收入	-0.001 (0.004)	-0.001 (0.004)	-0.002 (0.004)	-0.002 (0.004)
年龄	-0.005*** (0.000)	-0.004*** (0.000)	-0.005*** (0.000)	-0.005*** (0.000)
最高学历	-0.004 (0.003)	-0.003 (0.003)	-0.002 (0.003)	-0.002 (0.003)
性别	0.013* (0.007)	0.011 (0.007)	0.011 (0.007)	0.011 (0.007)
健康状况	-0.004 (0.003)	-0.004 (0.003)	-0.004 (0.003)	-0.003 (0.003)
养老保险	-0.010 (0.008)	-0.010 (0.008)	-0.008 (0.008)	-0.009 (0.008)
一孩性别	-0.025*** (0.007)	-0.025*** (0.007)	-0.024*** (0.007)	-0.024*** (0.007)
医疗保险	0.036*** (0.013)	0.033*** (0.013)	0.037*** (0.013)	0.037*** (0.013)
城乡虚拟变量	-0.009 (0.008)	-0.008 (0.008)	-0.007 (0.008)	-0.007 (0.008)
东部地区		-0.134*** (0.034)	-0.139*** (0.034)	-0.139*** (0.034)
中部地区		-0.129*** (0.034)	-0.133*** (0.034)	-0.132*** (0.034)
西部地区		-0.117*** (0.034)	-0.120*** (0.035)	-0.119*** (0.035)
中国腐败评价			-0.001 (0.002)	-0.001 (0.002)

(续表3) 政府评价对二孩生育预期行为的影响

变量名称	(1)	(2)	(3)	(4)
中国住房评价			-0.003 [*] (0.002)	-0.003 [*] (0.002)
中国社保评价			0.004 ^{**} (0.002)	0.004 ^{**} (0.002)
中国教育评价			-0.002 (0.002)	-0.002 (0.002)
中国医疗评价			-0.000 (0.002)	-0.000 (0.002)
政府拖延程度				0.003 (0.003)
与政府干部发生冲突				0.003 (0.005)
政府不合理收费				-0.003 (0.004)
常数项	0.173 ^{***} (0.041)	0.295 ^{***} (0.052)	0.312 ^{***} (0.054)	0.294 ^{***} (0.060)
观测样本数	3495	3495	3450	3450

注：括号内为稳健标准误，*、**、***分别表示在10%、5%和1%水平上显著。

在中国，中央政府和地方政府的关系较为复杂，虽然地方政府在执行中央政策时具有一定的灵活性、自主性，但地方政府始终是中央政府颁布政策的实施者，地方政府在政策落实过程中一定程度上代表着中央政府，地方政府的某些行为可能会引起公众对中央政府态度的改变，导致公众对中央政府的评价和对地方政府评价趋于一致。为了排除研究对象二孩生育行为受到对中央政府评价问题的干扰，本研究在回归模型（2）的基础上增加了中国腐败问题严重程度、中国住房问题严重程度、中国社会保障问题严重程度、中国教育问题严重程度、中国医疗问题严重程度等控制变量，结果如表3（3）列所示，其结果与模型（2）回归结果相似，该结果在一定程度上排除了研究对象不同的生育行为是因为对中央政府态度不同的结果。考虑到其他政府行为在一定程度上同时对本文研究解释变量和被解释变量造成影响，本文将数据中政府办事是否受到拖延推诿、与政府干部冲突、是否遭遇政府不合理收费控制在内，结果如表3（4）列所示，交互项系数显著为负，在一定程度上排除了内生性干扰。

（四）稳健性讨论

1. 倾向得分匹配。本文选取2018年家庭数据与2016年成人数据中同时出现的样本作为实验组，2016年家庭数据和2014年个人数据中同时出现且未在2018年家庭数据中出现的样本为对照组，由于实验组和对照组不同期，单纯使用DID模型没有考虑两组数据可观测特征是否可比较。因此本文在混合截面双重差分估计模型（DID）的基础上，进一步将DID与倾向得分匹配（PSM）相结合，解决了实验组和对照组不同期而受到限制的共同趋势检验问题，控制了实验组和对照组样本选择差异，同时较好地解决了模型内生性问题。

表4是对照组和实验组主要控制变量描述性统计差异表，可以看出实验组和对照组对县市政府评价、

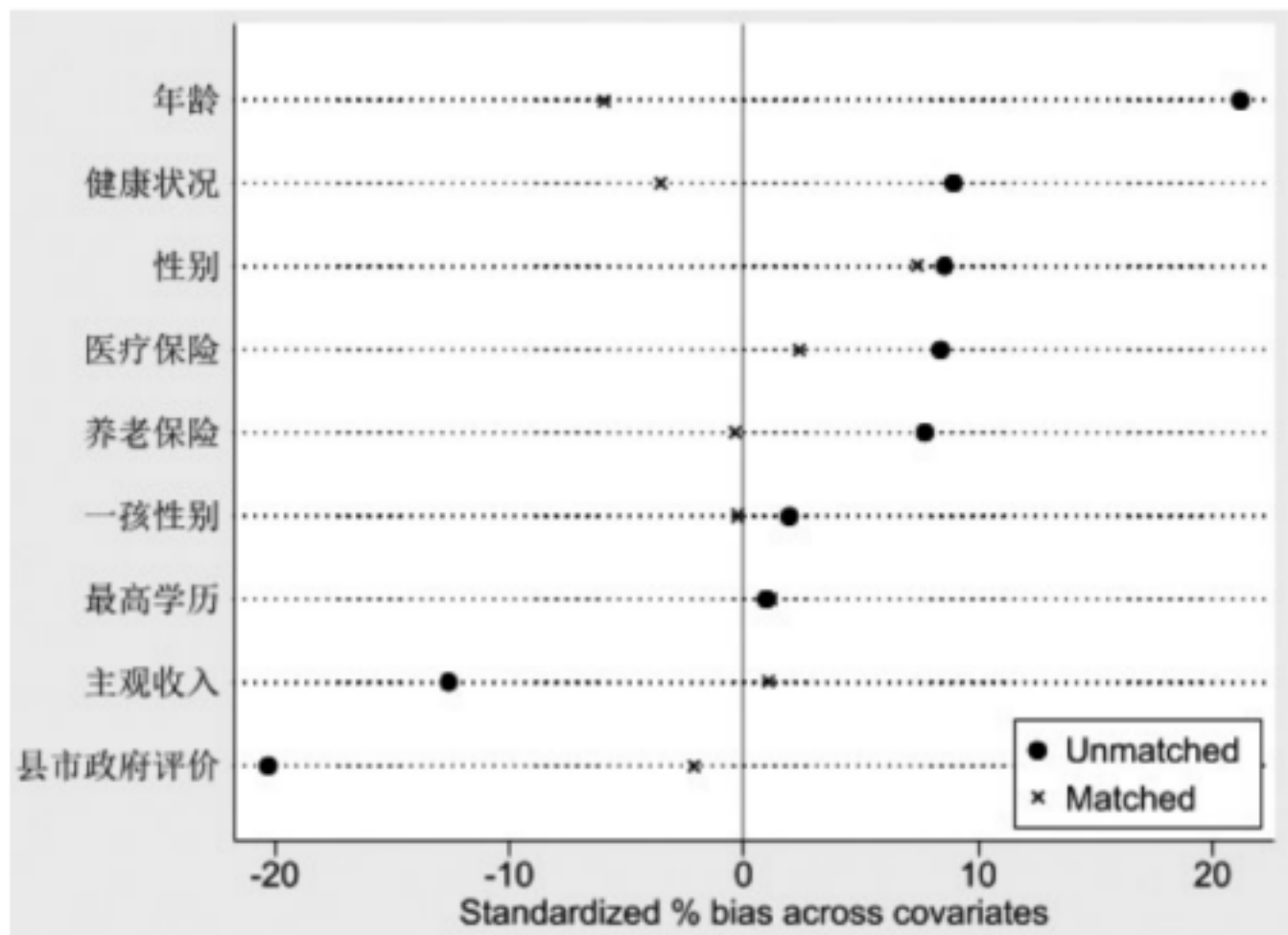
主观收入和年龄这几个变量在1 %的显著性水平上存在差异，性别、健康状况、养老保险和医疗保险在10 %的显著性水平上存在差异，因此本文选择对县市政府评价、主观收入、年龄、性别、健康状况、养老保险和医疗保险来进行样本匹配。表5为利用logit模型来估计倾向得分的结果，结果显示对县市政府评价、主观收入、年龄均在1 %的显著性水平上有差异，性别和医疗保险在5 %的显著性水平上有差异，虽然健康和养老保险的差异不显著，但随着经济社会发展和医疗卫生水平的提高，人们的健康问题普遍得到改善、养老保险的覆盖率也是逐步增加，对treat确有影响，因此选择这7个变量进行分析是可靠的。进一步对得分倾向匹配进行平衡性检验，结果如图1所示，所有匹配后的可观测变量标准偏差绝对值都小于10%，可证明本文所选取的可观测变量和匹配方法可靠且合适，DID-PSM回归结果可见表6第（3），交互项系数显著为负，说明模型（2）基准回归结果是稳健的。

(表4) 政府评价对二孩生育预期行为的影响

变量名称	实验组均值	对照组均值	平均值差值	p 值
县市政府评价	2.537	2.681	-0.144	0.000
主观收入	2.447	2.585	-0.138	0.001
年龄	37.38	35.66	1.718	0.000
最高学历	3.307	3.303	0.003	0.480
性别	0.593	0.553	0.040	0.047
健康状况	2.787	2.683	0.104	0.026
养老保险	0.689	0.651	0.038	0.046
一孩性别	0.583	0.573	0.010	0.331
医疗保险	0.919	0.890	0.029	0.027

(表5) logit模型回归结果

变量名称	估计系数	标准误	P 值
县市政府评价	-0.230	0.062	0.000
主观收入	-0.170	0.052	0.001
年龄	0.025	0.007	0.000
性别	0.232	0.099	0.019
健康状况	0.046	0.045	0.308
养老保险	0.005	0.110	0.964
医疗保险	0.332	0.166	0.045



(图1) 平衡性检验

2. 安慰剂检验。为了检验模型（2）基本回归结果是否可靠，本研究还通过将政策发生时间前置于2015年1月1日、后置于2017年1月1日，将其视为虚拟的全面二孩政策颁布时间，进行安慰剂检验，如果交互项系数显著，则说明安慰剂检验不通过。回归结果如表6第（1）（2）列所示，交互项系数不显著，说明本文基准回归结果是可靠的。

3. 敏感性检验和其他一些稳健性检验。在前文的分析中，观察对照组2015年1月1日至2015年9月30日二孩生育行为与实验组的二孩生育行为差异是否受到对地方政府评价影响，取1月至9月是为了与实验组月份统一，排除季节影响因素干扰。现在，为了检验模型（2）基准回归结果是否会因为遗漏了2015年10月1日至2015年12月31日这段日期而造成估计误差，将对照组二孩生育行为观察段设置为2015年4月1日至2015年12月31日，其结果如表7中第（1）列所示，交互项系数显著为负，与模型（2）基准回归结果一致。

(表7) 敏感性与稳健性检验结果

变量名称	对照组时间段	实验组时间段	县市政府评价取值	敏感性检验
	(1)	(2)	(3)	(4)
treat	0.080** (0.033)	0.067** (0.028)	0.078*** (0.030)	0.069** (0.029)
县市政府评价	0.012 (0.011)	0.013 (0.009)	0.015 (0.009)	0.007 (0.010)
treat*县市政府 评价	-0.026** (0.012)	-0.023** (0.010)	-0.025** (0.010)	-0.022** (0.010)
是否控制 X_{it}	是	是	是	是
是否控制 D_{it}	是	是	是	是
是否控制 I_{it}	是	是	是	是
常数项	0.313*** (0.060)	0.195*** (0.051)	0.294*** (0.057)	0.290*** (0.049)
观测样本数	3449	3459	3556	5454

注:括号内为稳健标准误,*,**、***分别表示在10%、5%和1%水平上显著。

其次,根据有迹可循的“全面二孩”生育政策颁布消息“放风”时间线,2015年7月官方对于实施“全面二孩”明确持有一个肯定的态度,并表示工作正在推进,以该“放风”消息为时间节点,保持对照组不变,观察实验组2016年4月1日至2016年9月30日二孩生育行为与对照组差异是否受到对地方政府评价影响。其结果如表7第(2)列所示,交互项系数显著为负,与模型(2)基准回归结果一致。

最后,在基准回归中,研究认为对县市政府评价数据中“比之前更糟糕了”和其他回答不是一个回答口径,因此将其去掉,现在为了检验数据稳健,将其加上,对模型(2)进行回归,其结果如表7第(3)列所示,与模型(2)基准回归结果保持一致。

本文的数据以家庭为单位对其二孩生育行为进行研究,在选取的过程中,通过stata命令,同一家庭只保留第一位出现的基因组成员,将其作为家庭代表,为了剔除回归结果对样本数据具有敏感性这一影响,本文保留每组家庭中的所有个体,以个体数据来替代家庭数据进行回归分析。回归逐步控制住了 X_{it} 、 D_{it} 和 I_{it} ,含义与上文相同,结果如表7第(4)列所示,其结果是稳健的。

(五) 异质性讨论

考虑到不同年龄育龄群体的二孩生育行为可能存在异质性,本文将样本中各年龄阶段划分为18至29岁、30至39岁、40至49岁。结果如表8第(1)(2)(3)列所示,可以看出,18至29岁育龄群体全面

二孩政策预期管理下二孩生育行为受到对地方政府评价的影响十分显著，对地方政府的评价越高，越会选择生育二孩。

(表 8)		异质性分析回归结果				
变量名称	18-29 岁	30-39 岁	40-49 岁	东部地区	中部地区	西部地区
	(1)	(2)	(3)	(4)	(5)	(6)
treat*县市政府评价	-0.102*** (0.034)	-0.022 (0.025)	0.001 (0.006)	-0.036** (0.015)	-0.054** (0.021)	0.016 (0.033)
是否控制 X_{it}	是	是	是	是	是	是
是否控制 D_{it}	是	是	是			
是否控制 I_{it}	是	是	是	是	是	是
_cons	0.263 (0.215)	0.182 (0.139)	0.022 (0.041)	0.178*** (0.067)	0.016 (0.091)	0.367*** (0.121)
N	744.000	1229.000	1477.000	1618.000	1026.000	770.000

注：括号内为稳健标准误，*、**、***分别表示在 10%、5% 和 1% 水平上显著。

接下来再考虑二孩生育行为受到具体的生育环境影响，全面二孩政策实施存在空间异质性，本文将数据样本中30个省（直辖市、自治区）划分为东部地区、中部地区和西部地区，对模型（2）进行回归。结果如表8所示，第（4）（5）（6）列分别是东部地区、中部地区、西部地区回归结果，对于本文主要关注的交互项系数而言，从第（4）（5）列可以看出，东部地区 and 中部地区育龄群体地方政府评价对全面二孩生育政策预期管理有显著影响，从第（6）列可以看出，地方政府评价对西部地区育龄群体无显著影响。

四、结论与政策建议

中国作为世界人口大国，生育问题一直是我国政府和社会共同关注的问题。随着人口风险增加，原有的生育计划已经不足以适应当期的经济社会变化，因此，“三孩政策”随之而来。2021年8月20日，第十三届全国人大常委会第三十次会议表决通过了《全国人民代表大会常务委员会关于修改<中华人民共和国人口与计划生育法>的决定》，决定指出一对夫妻可以生育三个子女。生育计划不可能是一成不变的，政策应该随着人口结构的变化而不断调整，当原有的生育政策不能与当今的经济社会相适应时，政府相关部门理应对其做出改变以适应社会发展。政策改变或制度变革必然会触及一部分人的利益。在生育计划调整之后，如何使新的生育政策尽快达到预期效果就成为了一个重要问题。

本文利用2014、2016和2018年中国家庭追踪调查（CFPS）数据，采用Probit模型和混合截面双重差

分模型（DID）实证分析了对地方政府的评价如何影响育龄群体在政策模糊阶段对二孩生育政策的预期，进而影响二孩生育行为。研究发现育龄群体对生育政策预期不仅在对地方政府评价方面存在异质性，而且公众对地方政府评价越高越会对政府生育政策进行积极预期而进行二孩生育行为，这给我们的理论提供了一个很好的现实支撑。

通过实证分析发现育龄群体对政府的信任并不是完全无条件的信任，而是一种策略性信任，也就是说，政府能够通过采取一定的手段措施改变育龄群体对政府的信任程度进而使生育政策更快更好达到政策目标。一方面，政府应该加强政策颁布前与群众就政策相关信息的沟通、宣传与预期引导。首先，政府可以通过新闻发布会或一些官方媒体发布相关信息，拓宽与群众的沟通渠道；其次，政府可以对公众预期进行引导，社会在不断发展，人们获取信息的能力也在不断加强，政府可以利用其自身特殊地位来影响政策相关信息的发布和传播（龚维斌，2017）[32]，通过对政策内容、政策施行时间、政策实施强度和方式等消息的“放风”来引导公众预期。年龄异质性分析结果表明对地方政府的评价对18到29岁的育龄群体二孩生育预期行为影响更为显著，政府应加强对该年龄阶段育龄群体的积极政策预期引导，着重关注这部分群体生育的相关需求，提高其对政府的评价，以更好地推进政策达到预期效果。最后，鉴于预期产生的前提是社会变化会对公众利益产生影响且该社会变化能够被公众察觉，部分育龄群体可能并不关注“全面二孩”政策，从而造成本文在一定程度上低估了地方政府评价对育龄群体预期的影响，对于这部分群体，政府可以加大政策宣传力度，使其能够察觉到政策影响。

另一方面，地方政府应该重视长期政策信誉建设，对于政策预期引导不能过度夸大，西方政客面临竞选时往往会在这种“放风”信息中有意夸大政策影响或者对公众做出过度承诺来获得公众支持，但研究证明政客们在竞选时的承诺会提高公众预期需求，而这些需求在事后很难被满足，从而形成“期望差距”。公众会根据过往信息形成预期从而做出未来决策，这种“期望差距”会不利于后续相关政策的实施。政府想要缩小这种“期望差距”方法有二，一是压低公众的事先期望，二是增加公共服务供给和提高公共服务质量。但是压低公众事先期望必然会导致当期政策实施效果大打折扣，所以政府想要当期和后期政策实施达到预期效果，就要想办法增加政策相关服务供给、提高政策相关服务质量。新公共管理运动改变了传统公共管理模式下的政府与社会公众的关系，社会公众是政府提供的公共物品及服务的顾客，并对此进行评价（黎民，2011）[33]。地方政府提供的公共物品及服务质量越高，公众对地方政府评价就会越高。如此，便进入了一个良性循环。

[参考文献]

- [1] 乔晓春.“单独二孩”,一项失误的政策[J].人口与发展,2015,(6).
- [2] 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).
- [3] 陈建新,王莉君.个人因素对不同阶段生育意愿的影响——基于CGSS2017数据的实证分析[J].调研世界,2021,(6).
- [4] 邓敏,陈宝璘.城乡、性别差异下一孩家庭二孩生育意愿的比较研究[J].人口与社会,2020,(4).
- [5] 王一帆,罗淳.促进还是抑制?
受教育水平对生育意愿的影响及内在机制分析[J].人口与发展,2021,(5).
- [6] 李婉鑫,杨小军,杨雪燕.儿童照料支持与二孩生育意愿——基于2017年全国生育状况抽样调查数据的实证分析[J].人口研究,2021,(5).
- [7] 阮荣平,焦万慧,郑风田.社会养老保障能削弱传统生育偏好吗? [J].社会,2021,(4).
- [8] 孟颖颖,王静,张孝栋.城乡居民医保整合对农民工城市融合的影响——
基于2017年全国流动人口动态监测调查数据[J].决策与信息,2020,(8).
- [9] 胡荣,林彬彬.性别平等观念与女性生育意愿[J].求索,2020,(4).
- [10] 王国军,高立飞.低生育意愿的一个解释:养儿防老向商业保险养老转变——基于CGSS2015数据的实证分析[J].兰州学刊,2021,(2).
- [11] Chanley, J., Rudolph, T. J., & Rahn, W. M..Public Trust in Government in the Reagan Year sand Beyond. InJ.R. Hibbing, &E. Theiss-Morse (Eds.),What is it About Government that Americans Dislike?[M].Cambridge:Cambridge University Press,2001.

- [12] Mishler W, Richard R. What Are the Origins of Political Trust? Testing Institutional and Cultural Theories in Post-Communist Societies[J]. Comparative Political Studies, Vol 34, No. 1, 2001.
- [13] 吴华根. 基于合作型政府信任关系的深圳市政府转型[J]. 深圳大学学报(人文社会科学版), 2008, (5).
- [14] D'Haultfoeuille X, Gaillac C, Maurel A. Rationalizing Rational Expectations: Characterization and Tests[J]. Quantitative Economics, 2020.
- [15] 张冲, 李想. 女性生育意愿与生育行为偏离的影响因素[J]. 中国卫生统计, 2020, (6).
- [16] 风笑天. 当代中国人的生育意愿: 我们实际上知道多少?[J]. 社会科学, 2017, (8).
- [17] 原新, 高瑗. 中国快速人口老龄化的战略机遇[J]. 河北学刊, 2017, (4).
- [18] 风笑天. “全面二孩”政策实施效果: 现有评价分析及其思考[J]. 东南大学学报(哲学社会科学版), 2020, (5).
- [19] 任远. 定位和定向: 我国生育政策未来[J]. 探索与争鸣, 2017, (7).
- [20] Meier K J, O'Toole L J. Public management and organizational performance : the effect of managerial quality[J]. Journal of Policy Analysis and Management, 2002, 21(4).
- [21] Hetherington M J. Why trust matters: Declining political trust and the demise of American liberalism[M]. Princeton: Princeton University Press, 2006.
- [22] Truidera E M, Steckermeier LC. Trusting and controlling? Political trust, information and acceptance of surveillance policies: The case of Germany[J]. Government Information Quarterly, 2017.
- [23] Hetherington, M. J. The political relevance of political trust[J]. The American Political Science Review, 1998.

- [24] 孟天广.福利制度的过程治理：再分配、政府质量与政治信任[J].行政论坛,2022,(1).
- [25] Levi M,Sacks A,Tyler T.Conceptualizing Legitimacy,Measuring Legitimizing Beliefs[J].American Behavioral Scientist,2009.
- [26] Gabriel O W,Trudinger EM.Embellishing Welfare State Reforms? Political Trust and the Support for Welfare State Reforms in Germany[J].Germany Politics,2011,20:2,pp 273-292.
- [27] 张晓晶,曾一已,邵兴宇.构建中国特色的预期管理[J].中共中央党校(国家行政学院)学报,2021,(3).
- [28] 江世银.社会预期管理论[M].北京:中国社会科学出版社,2017.
- [29] Ashraf N,Bau N,Nunn N,Voena A.Bride Price and Female Education[J].Journal of Political Economy,2020,vol. 128(2), pp 591-641.
- [30] Duflo E.Schooling and Labor Market Consequences of School Construction in Indonesia:Evidence from an Unusual Policy Experiment[J].American Economic Review,2001,Vol.91,No.4.
- [31] 李明,郑礼明.回不去的家乡?——教育公共品供给与人口回流的实证研究[J].金融研究,2021,(4).
- [32] 龚维斌.社会预期管理是社会治理的有效方式——评《社会预期管理论》[J].社会治理,2017,(3).
- [33] 黎民.公共管理学[M].北京:高等教育出版社,2011.

[责任编辑：汪智力 朱苗苗]

Local Government Evaluation and Decision-making of Two-child Birth for Childbearing Age Groups
—— An Empirical Analysis Based on CFPS Data

ZHANG Ping, XIAO Zhenni, WANG Weizhou

Abstract: On January 1, 2016, the universal two-child policy was officially implemented nationwide. By focusing on the expected fertility behavior of childbearing age groups on the ' universal two-child ' policy, using the data of China Family Panel Studies (CFPS) in 2014, 2016 and 2018, and using a combination of mixed-section difference-in-difference (DID) and propensity score matching (PSM), this paper estimates how the evaluation of local governments affects the expected fertility response of childbearing age groups to the ' universal two-child ' policy. The empirical results show that the childbearing age group will make an expected response to the future fertility policy. The trust of the childbearing age group to the government is a strategic trust. The higher the evaluation of the local government, the more it will respond positively to the future fertility policy and carry out the expected fertility behavior. The results of heterogeneity analysis show that the expected behavior of childbearing age groups aged 18 to 29 and in the eastern and central regions is more affected by the evaluation of local governments. If the government hopes that future policies will achieve better and faster results, it cannot ignore the communication, publicity and expectation guidance with the masses on policy-related information before the promulgation of the policy, and should attach importance to long-term policy credibility building.

Keywords: fertility desire ; reproductive behavior ; comprehensive two-child policy ; public expectations ; population policy ; government credibility

[收稿日期] 2022-07-30

[基金项目] 本文系2017年度国家社会科学基金项目“基于地方政府激励的统计虚报测度及校正机制研究”(编号:17BJL070)阶段性成果。

[作者简介]张平（1981-），男，陕西西安人，经济学博士，武汉大学社会保障研究中心副教授，主要从事公共经济学、发展经济学研究；肖真妮（1997-），女，湖南株洲人，武汉大学社会保障研究中心硕士研究生；王伟舟（1998-），男，山东济南人，伦敦大学学院斯拉夫东欧研究院硕士研究生。