

# 贫困家庭高中阶段教育选择: 基于行为经济学的视角<sup>\*</sup>

李芳华 肖 寒

内容提要: 本文使用河南省某样本县 2012—2019 年全部初中毕业生的中考成绩和高中录取数据, 基于普通高中第一轮录取的分数线可能对学生家庭造成的关于“考上大学”主观概率的冲击, 以及 2016 年针对建档立卡户的普通高中学费减免政策这一准自然实验, 采用“断点-差分法”和“双重差分法”研究后发现: 面对同等的来自“考上大学”概率的负面冲击, 贫困家庭的学生有更大的概率选择职业高中; 当贫困家庭普通高中学费减免(即普通高中的净收益增加)时, 这一概率差大幅下降。在排除了贴现因子差异、收入水平上升、教育观念突变与信息不足四种可能的解释后, 本文充分证明, 造成贫困家庭与一般家庭高中选择差异的原因在于贫困家庭更高的风险厌恶程度。本研究厘清了高中阶段教育决策的微观机制, 为研究低收入群体的风险偏好提供了来自中国的证据, 对于普职协调发展和加强技能型人才培养具有鲜明政策启示。

关键词: 贫困 教育选择 职业高中

## 一、引 言

高中阶段教育是我国国民教育体系的重要组成部分, 在人才培养中起着承上启下的关键作用。<sup>①</sup>我国高中阶段教育主要包括普通高中和职业高中(包括职业中专、技工学校等)两个选择, 普通高中的教育是高等教育的准备(李中国和彭李, 2015), 是培养复合型创新型人才的基础; 而职业教育以培养技能为主, 可有效解决劳动力短缺问题和缓解青少年就业的不稳定性(欧阳忠明和王江雁, 2022)。如何引导不同能力、不同偏好的青少年选择合适的教育类型对经济发展、人力资本积累具有关键的作用(Falk et al., 2020)。我国一直采用“分数为主, 个人选择为辅”的高中录取机制, 虽然家庭有一定的选择空间, 但中考分数是决定子女高中阶段去向的主要依据。当前, 高中阶段的普职分流政策加重了家庭的教育焦虑, 导致在义务教育阶段的过度投资。同时, 我国在经济转型期间对职业技能型人才的需求持续上升。在此背景下, 2022 年 4 月我国修订《职业教育法》明确提出“普职协调发展”的指导思想, 对中等职业教育定位进行了调整, 未来有望对高中阶段分流方式进行持续的改革。因此, 现阶段厘清农村家庭高中阶段入学选择的微观机制, 对于政府进一步推动教育改革具有重要意义。

家庭教育决策是经济学的经典议题之一, 其中低收入家庭的教育选择问题对阻断贫困代际传递、提高社会流动性具有重要意义。学界在这一领域已经得到较为一致的结论, 即低收入家庭的教

<sup>\*</sup> 李芳华, 澳大利亚新南威尔士大学经济学院, 邮政编码: 2052, 电子信箱: fanghua.li@unsw.edu.au; 肖寒(通讯作者), 中国人民大学应用经济学院, 邮政编码: 100872, 电子信箱: xiaohanlinda@ruc.edu.cn。作者感谢匿名审稿专家的宝贵建议, 但文责自负。

<sup>①</sup> “教育这十年”“1+1”系列发布会<sup>②</sup>《介绍党的十八大以来普通高中教育改革成效》, <http://www.scio.gov.cn/xwfbh/gbwxfbh/xwfbh/jyb/Document/1726593/1726593.htm>。

育投资显著低于高收入家庭(邹薇和郑浩,2014),家庭收入的增加可以显著提高教育投资水平(Shapiro & Tambashe,2001)。但是,收入影响教育投资的机制目前尚不明确。传统的经济学解释有两种,其一是低收入家庭的教育理念落后导致对子女教育投入的不足。部分贫困率较高的发展中国家与发达国家存在教育理念上的差别,例如,“读书无用论”流行于我国广大农村地区(李涛和邬志辉,2015;蒋承等,2018;黄政,2021)。其二是低收入家庭面临紧的预算约束导致教育投资的决策与高收入家庭不同(Becker & Tomes,1985;Lucas & Kerr,2013;李力行和周广肃,2015;王伟同等,2021)。家庭资金或信贷约束对子女的教育投入具有显著的负向影响(Becker & Tomes,1979;Shultz,1993;李菁等,2002)。因此,改善家庭预算约束的政策如财税减免、低保项目实施、个人所得税改革以及住房财富的增加均对农村家庭的教育投资产生显著的正向影响(陈永伟等,2014;Gao et al.,2014;刘利利和刘洪愧,2020;任伟聪和梁若冰,2022)。

近年来产生了有别于传统经济理论的另一种行为经济学理论,<sup>①</sup>即低收入家庭由于在效用函数基础参数上与一般家庭不同,因此在包括教育投资的各种决策中存在非理性行为。Banerjee(2006)和Haushofer & Fehr(2014)认为,贫困个体所处的不利条件会引发心理变化,例如,折现率更高(impatient),风险厌恶程度更高等,导致其行为具有不经济性、可能引发“行为型”贫困陷阱,降低对教育、健康等有长期收益但存在一定风险的投资。以上不同机制对应的政策含义明显不同,因此对低收入家庭教育投资机制的识别不仅具有重要的学术价值,还能为提高政策工具的有效性提供可靠的参考。但是机制识别,特别是风险厌恶、折现率等参数的识别在现实中并不容易,目前的实证结果多是来源于实验室研究或随机实验(Ashraf et al.,2006;Fehr et al.,2022)。

中国高中阶段的学校选择为验证贫困家庭的偏好参数与一般家庭是否存在显著差异提供了一个绝佳的现实设定。我国高中阶段教育主要包括普通高中和职业高中(包括职业中专、技工学校等)两个选择,<sup>②</sup>在这两种类型的高中之间,遵循“分数为主,个人选择为辅”的分配机制。这两个选择在成本和风险程度上有很大不同。普通高中作为高等教育的准备阶段,毕业后可以参加普通高考,以一定的比例考入大学。而普通高中毕业、职高毕业和大专以上学历毕业的回报率存在明显的差距。根据目前文献的研究结果,三者的关系为:农村职业高中的回报率显著高于普通高中(颜敏,2012;侯风云,2004;胡咏梅和陈纯槿,2013;祁占勇和何佑石,2022),但显著低于大专以上学历(陆万军和张彬斌,2021)。<sup>③</sup>因此,从理论上来看,学生在职业高中与普通高中之间的选择,可以看作在“无风险中等收益”与“风险高收益”之间进行的最优化决策,可通过不同家庭的选择识别家庭风险偏好的不同。在此思路下,本文基于我国某县全部中考学生的个体成绩数据和高中选择数据,并匹配建档立卡贫困户信息,借助中考录取中的两轮录取机制构建识别策略。具体而言,我们利用第一轮“正取分数线”对上大学主观概率形成的冲击,使用“断点-差分”方法检验此分数线附近的贫困和一般家庭学生在冲击下的选择是否存在显著的差异,以验证不同家庭在教育选择上是否存在显著的风险偏好差异。在此基础上,进一步利用2016年教育精准扶贫的准自然实验作为外生冲击,识别了教育价格对贫困和一般家庭学生的异质性影响。

<sup>①</sup> 传统经济理论以Schultz(1976)为代表,不考虑偏好参数的差异,认为贫困家庭与其他家庭同样理性且有效率,选择完全由预算约束导致。

<sup>②</sup> 基于我国当前的发展阶段和教育资源供给结构,根据教育部发布的《2020年全国教育事业统计公报》,2015年我国高中阶段教育毛入学率为87.0%,在1500万左右的招生名额中普通高中招生数不足900万(约占60%),中等职业教育占比40%以上。

<sup>③</sup> 以陆万军和张彬斌(2021)为例,中等职业教育年工资收入约比初中教育高24.3%,普通高中教育比初中教育高13.5%,大学专科和本科分别比初中教育高50.6%和84.1%。

从实证角度看,在高中阶段教育选择这一问题上,基于发达国家的实证研究结果对我国普职分流的借鉴意义有限,<sup>①</sup>而使用国内数据针对我国高中教育选择的研究受数据限制较大。由于我国中考的分配机制中成绩是主要决定性因素,因此在对高中教育选择进行实证研究时,需要学生层面的考试数据以及高中阶段的去向数据。同时,理想的个体数据应来自同一地区以排除社会偏好、地域特征等不可观测因素的影响。本文的边际贡献体现在以下三个方面:研究数据上,本文匹配了来自教育系统的学生中考成绩数据、来自普通高中学籍系统的高中入学数据和来自职业高中学籍系统的职业高中入学数据,得到了同一县内学生中考表现和初中毕业后的真实教育选择信息,解决了前文所述的数据问题。研究方法上,本文采用断点-差分设计,以高中录取机制中的正取分数线作为断点,利用断点两侧对考上大学主观概率的影响,检验不同家庭学生的风险偏好是否存在差异,并利用2016年建档立卡贫困高中生免除学费的外生政策冲击构造准自然实验,研究政策影响。断点-差分设计的优势在于将学生样本限制在第一轮正取分数线上下这一教育选择空间最大的学生群体,与我国当前高中录取“分数为主,个人选择为辅”这一特殊选择机制相符。研究内容上,本文研究了贫困和非贫困家庭在高中选择方面的影响因素,并且探讨了两类家庭教育选择产生差异的机制,为行为经济学提供新的实证证据,丰富了关于我国高中阶段选择的实证研究。<sup>②</sup>

本文余下部分安排如下:第二部分介绍我国高中阶段的入学分配机制,并使用一个简单的风险资产选择模型得到可检验假说;第三部分介绍了本文所使用的数据来源和回归策略;第四部分为实证结果分析;第五部分排除了贴现因子、收入水平和帮扶人政策对高中教育决策的影响,并进一步讨论关于性别的异质性结果;第六部分为结论和启示。

## 二、背景与理论框架

### (一) 高中阶段入学选择与教育扶贫政策

我国目前初中毕业生通过中考进入不同类型高中的分配并非按照分数进行“一刀切”,而是遵循“以中考分数为主,以个人选择为辅”的分配机制,其中考生与学校的双向选择环节给予了学生和家长一定的选择空间。具体而言,中考评分结束之后普通高中首先进行录取,录取过程基本遵循波士顿机制(Boston mechanism)。录取过程分两轮,第一轮是各地区各学校按照学位数目划定分数线,并且根据学生具体的高中志愿选择进行匹配,此分数线为“正取分数线”。在第一轮结束后,若学校还有剩余学位,则会降低招生分数线,仅划定一个相对较低的“控制分数线”,<sup>③</sup>并对有意愿的学生进行第二轮录取。而职业高中的招生环节在普通高中之后进行,录取所有有意愿的初中毕业生,录满为止。<sup>④</sup>录取环节存在的选择空间为研究不同家庭对高中的选择差异提供了机会。

<sup>①</sup> 对职业教育和普通教育选择的实证研究主要来自德国,德国目前已形成较完善的“双元制”职业教育体系。虽然大量研究已经发现,德国低收入家庭的学生在教育选择上处于明显弱势,在控制了学生的成绩表现后高社会地位背景的儿童仍然比低社会地位背景的学生更有可能进入学术教育轨道(Falk et al. 2020),但由于德国的职业教育和普通教育基本免费,对教育价格如何影响教育类型选择的实证研究较为稀缺。此外,美国公立和私立中小学选择与我国高中选择有一定相似性。不同的是,我国职业教育和普通教育的差异不仅体现在教学质量方面,更体现在职业发展期望、上学年限期望等方面,因此,基于美国“公立-私立”中小学选择的研究结论(Epple & Romano, 1998)不能直接推广到我国高中阶段的教育选择。

<sup>②</sup> 目前的研究主要分为两种:一是根据个体高中阶段去向的数据进行描述性和相关性分析(方长春和风笑天, 2005; 陈胜祥, 2011; 沈有禄, 2017; 周洁和张俊, 2019; 石伟平和李鹏, 2021),缺乏学生层面的成绩数据以排除学生能力的差异,也无法排除地域因素的影响,难以进行细致的因果识别;二是以抽样调查的形式采集学生高中阶段教育选择意愿(苏丽锋等, 2016; 苏丽锋等, 2017),但是由于缺乏学生在高中阶段的实际去向数据,无法识别家庭的教育选择决策。

<sup>③</sup> 以样本县为例,每年的一轮正取分数线在450分以上,而二轮的分数线仅在350分左右。

<sup>④</sup> 多个省份包括河南省,中考同时也是初中结业考试,是获得初中学历证书的必要条件。因此,虽然很多职业高中不存在竞争性,但需要初中学历证书,即中考成绩。

不同类型的高中对应不同的教育成本。自2009年起,农村的家庭经济困难学生进入职业高中已经不需要缴纳学费,2012年后进一步扩大免学费范围,覆盖了所有农村(含乡镇)学生以及涉农专业学生。<sup>①</sup>到2013年,我国所有职业高中学生已经实现全部免费。此外,2006年起农村户籍的职高学生还享受助学金政策。<sup>②</sup>相比之下,普通高中的助学金政策开始得更晚,<sup>③</sup>并且只覆盖小部分家庭经济困难的学生。<sup>④</sup>对于农村地区的学生来说,学费与助学金共同形成了两种高中类型之间的价格差。如果学生来自农村的经济困难家庭,那么在2015年选择上高中比选择上职高的成本高2423元/年,<sup>⑤</sup>约占2015年农村人均纯收入的22%。<sup>⑥</sup>

2016年为了落实《中共中央国务院关于打赢脱贫攻坚战的决定》的要求,来自建档立卡贫困家庭的学生享受以下两个优惠措施:第一是对建档立卡家庭的普通高中学生免除学费和住宿费,第二是将所有来自建档立卡贫困户的学生补助级别提升到最高的1250元/学期。简而言之,2016年仅针对建档立卡家庭的精准扶贫教育政策构成了关于普通高中价格的准自然实验。

不同类型的高中也对应着不同的教育回报。普通高中作为高等教育的准备阶段,在毕业后可以参加普通高考,以一定的比例<sup>⑦</sup>考入大学;而普通高中毕业、职高毕业和大专以上学历毕业的回报率存在明显的差距。因此,高中阶段的选择是家庭根据成本与期望收益综合决策的结果。从实际情况看,我们发现,对于不同中考成绩的学生而言,选择进入高中的概率存在很大差异。以本文使用的样本县为例,2016年以后中考成绩在全县前40%(也就是普通高中学位数\*80%)的学生,<sup>⑧</sup>有98.94%的学生选择普通高中。正取分数线(即一轮录取分数线)一般由该地区的高中学位数决定。样本县的正取分数线在50%左右,而中考排名在全县40%—50%以内的学生中,只有53.6%的学生最终选择进入普通高中,即并非所有超过正取分数线的学生都选择了普通高中。对于这些学生而言,上大学虽然可以获得高回报率,但由于存在考不上大学的风险,且高中毕业的预期工资率低于职业高中毕业生,家庭需要在无风险的职业高中与有风险的普通高中之间进行选择。

## (二) 高中阶段入学选择模型

考虑以下家庭决策模型:家庭在职业高中( $S=V$ )与普通高中( $S=HS$ )之间选择,以最大化折现总净收益的效用。假设不同教育水平的工资水平由市场外生给定,初中毕业、职高、普通高中和大学对应的工资水平为 $\{w^M, w^V, w^H, w^C\}$ 。根据文献,满足 $w^M < w^H < w^V < w^C$ ;且各类教育的成本由政府确定(外生),即职高、普高和大学的折现总成本向量为 $\{P^V, P^H, P^C\}$ 。为方便阐述,我们假设

① 职业高中对城市学生免学费比农村学生更晚,自2012年开始针对城市涉农专业学生和家庭经济困难学生免学费。

② 补助标准为每生每学年1000元,2007年这一资助标准提高为每生每年1500元。

③ 2010年,国务院发布《财政部、教育部关于建立普通高中家庭经济困难学生国家资助制度的意见》,明确提出建立健全普通高中国家家庭经济困难学生资助政策体系,建立了国家助学金制度。详见:[http://www.gov.cn/gongbao/content/2011/content\\_1786418.htm](http://www.gov.cn/gongbao/content/2011/content_1786418.htm)。

④ 高中助学金的比例较低,根据政策规定资助面约占全国普通高中在校生总数的20%,其中东部地区为10%,中部地区为20%,西部地区为30%。建档立卡家庭一般被纳入家庭经济困难学生中,可以享受资助。

⑤  $2423 \text{元} = \text{职高和普高学费之差} + \text{职高和普高助学金之差}$ 。根据教育统计年鉴,2015年高中平均学费为1923元/学年,职高平均学费为631元/学年,差距为1292元/学年。职高助学金为2000元/学年,普高助学金为1500元/学年,差距为500元/学年。

⑥ 2015年农村居民人均纯收入为10772元。

⑦ 高中升学率存在很强的省际差异,根据样本所在省招生办公室公布的数据,2019年普通高考录取率为78.58%,即参加高考的学生中有21%的高考生只能获得高中学历。此外,职业高中毕业生也有一定比例可以升入高等院校。根据北京大学中国教育财政科学研究所2020年全国中职毕业生抽样调查,约有10%升入本科院校。但这一情况并不改变现有研究得到的普通高中、职业高中和大专以上学历毕业生的工资回报率关系。

⑧ 样本县所在河南省的高考录取率为78.58%,样本县的高中录取率为50%,因此可计算得,前 $50\% * 78.58\% = 40\%$ 的学生考上大学的概率接近1。

存在借贷市场以平滑各期消费,即不存在单期预算约束。

基于上述的定义,家庭  $i$  在第 0 期选择职业高中 ( $S = V$ ) 的折现总净收益可以表示为折现净收入与折现净教育支出之差:<sup>①</sup>

$$R_{i,0}^V(\beta_i, w^V, P^V) = \sum_{t=3}^T \beta_t^t w_t^V - P^V$$

若家庭  $i$  在第 0 期选择普通高中 ( $S = HS$ ), 则学生将来可能仅高中毕业,也可能考上大学,这两种可能性分别对应折现总净收益  $R_{i,0}^{HS}(\beta_i, w^H, P^H)$  与  $R_{i,0}^C(\beta_i, w^C, P^C, P^H)$ , 可分别表示为:  $R_{i,0}^{HS}(\beta_i, w^H, P^H) = \sum_{t=3}^T \beta_t^t w_t^H - P^H$  与  $R_{i,0}^C(\beta_i, w^C, P^C, P^H) = \sum_{t=7}^T \beta_t^t w_t^C - (P^H + P^C)$ 。

假设效用函数  $U_i(\cdot)$  为冯·诺依曼效用函数,即满足  $U'(\cdot) > 0$ ,  $U''(\cdot) < 0$ 。对于普通高中选项,由于存在不确定性,在家庭判断上大学的概率为  $Pr_i^C$  时,期望效用为:

$$U_{i,0}^{s=HS} = (1 - Pr_i^C) U_i [R_{i,0}^H(\beta_i, w^{HS}, P^H)] + Pr_i^C U_i [R_{i,0}^C(\beta_i, w^C, P^C, P^H)]$$

假设个体之间不存在贴现率的差异,即  $\beta_i = \beta$ , 后续我们会对此假设进行检验。家庭在  $\{V, HS\}$  之间选择,可以由图 1 表示。假设一类人的效用函数  $U$  为图 1 中黑色实线,职业高中折现净收益  $R^V$  对应的效用为点  $a$ , 该点的风险厌恶程度为  $A(R^V) = -U''(R^V)/U'(R^V)$ , 则与该无风险收益效用相等的风险组合是与点  $a$  平行的点  $b$ 。也就是说,家庭效用最大化的最优策略为,当上大学的概率  $Pr_i^C$  使得对应的  $U_{i,0}^{s=HS}$  在点  $b$  右侧时选择上高中,反之则选择上职业高中。两种选择的比例如图 1 所示。若风险厌恶程度升高,即效用函数由实线  $U$  上升为虚线  $U'$ , 则无风险收益对应的效用上升为点  $a'$ , 最优决策的分界线向右平移至点  $b'$ , 选择普通高中的比例降低。换言之,在同样的上大学概率  $Pr^C$  下,风险厌恶程度高的个体,更倾向选择职业高中。由此可以得到第一个假说:

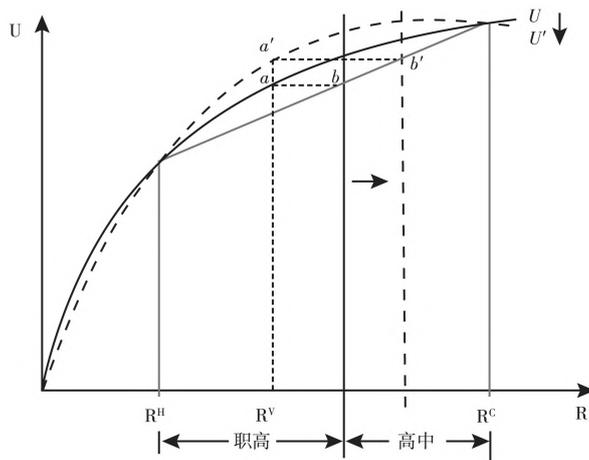


图 1 高中入学类型选择

假说 1: 假设贴现因子不存在个体差异,且不存在单期预算约束,则在相同的上大学概率  $Pr^C$  下,风险厌恶程度更高的家庭更倾向于选择无风险收益(职高)。换言之,对  $Pr^C$  下降更敏感(即决策  $S$  对  $Pr^C$  的弹性更大)的家庭,风险厌恶程度更高。

当高中的教育成本降低时,高中毕业与大学毕业的折现净收益上升。此时风险组合的两种可能结果所对应的折现收益皆上升,在同等条件下比无风险收益效用更高。在风险厌恶程度不变的情况下(对应效用函数  $U$  的形状不变),选择风险组合(即上高中)的比例上升,由此可以得到第二

<sup>①</sup> 增加“从职高升入大学”是一种更复杂的职高教育选择,但是由于职高和大学的回报率均高于初中毕业和普通高中毕业,因此这种情况下的平均回报率  $w^V$  仍然满足目前模型中的  $w^M < w^H < w^V < w^C$  设定。

个假说:

假说 2: 假设贴现因子不存在个体差异,且不存在单期预算约束,则在相同的上大学概率  $Pr^C$  下,当上高中的成本下降时,选择风险组合(即上高中)的比例显著提高。

本文通过比较贫困学生与一般家庭学生在高中入学选择时的差异,利用正取分数线构造的关于考上大学主观概率( $Pr^C$ )的断点以及教育扶贫政策所构成的准自然实验,验证以上两个假说,以证明贫困家庭与一般家庭的风险厌恶程度存在显著差异。

### 三、数据与回归策略

#### (一) 数据和变量

##### 1. 研究区域的代表性

本文使用的数据来自我国河南省某样本县,我们使用 2016 年《中国县域统计年鉴》和人口普查数据中各类经济指标、教育指标与样本县进行对比,结果如表 1 所示。对比发现对于大部分经济指标,样本县与全国层面统计指标的中位数或均值非常接近,说明样本县的经济发展水平、收入以及储蓄水平处于全国平均;教育方面,样本县人口的平均受教育年限与全国人口的均值和中位数持平,普通中学和小学在校学生数的人口比例略高于全国均值但在河南省中位数附近,表明样本县的教育资源和在校生情况分别具有全国代表性和全省代表性。由于本文主要研究贫困和非贫困家庭的差异,所以着重比较样本县的贫困家庭是否具有全国代表性,使用中国家庭金融调查 2015 年数据(CHFS2015)中农村贫困家庭与样本县的贫困家庭进行对比后发现,样本县的贫困家庭收入水平、家庭结构、供养负担、劳动人口受教育情况和年龄分布均非常接近全国抽样调查的均值或中位数,说明样本县的贫困人口也具有良好的全国代表性。

同时,样本县的教育政策也与全国政策一致。从 2009 年起,样本县所有职业高中一、二年级的学生均可享受 1000 元/学期的助学金,并且免除学费。根据国家精准扶贫的政策,从 2016 年秋季学期开始,对所有来自建档立卡贫困户的学生免除学费(850 元人民币/学期)和住宿费(300 元人民币/学期),并享受补助 1250 元/学期。<sup>①</sup>

表 1 样本县各项指标与全国对比

指标	全国 均值	全国 中位数	河南 均值	河南 中位数	样本县
Panel A: 样本县的经济指标					
人均地区生产总值(单位:万元)	4.45	3.38	3.81	3.14	3.85
第二产业占比(单位:%)	42.94	43.84	49.33	48.26	42.40
规模以上工业企业(单位:个)	127	62	161	142	63
人均储蓄余额(单位:万元)	2.38	1.90	1.71	1.56	1.95
人均储蓄余额/人均可支配收入	1.35	1.19	1.05	0.96	1.23
Panel B: 样本县的教育指标					
平均受教育年限(单位:年)(2020年)	8.57	8.63	8.77	8.61	8.82
(普通中学+小学在校学生数)/人口	0.11	0.11	0.13	0.13	0.14

<sup>①</sup> 按照国家政策,2016 年前样本县对于家庭困难的学生的补助分为三档,分别是 750 元/学期、1000 元/学期和 1250 元/学期。根据河南省的“雨露计划”补助政策,2016 年及以后,职业高中对建档立卡贫困学生的助学金额度同样为 1250 元/学期,与普通高中相同。

续表 1

指标	全国 均值	全国 中位数	河南 均值	河南 中位数	样本县
Panel C: 样本县贫困家庭和个人特征					
人均纯收入(单位: 元)	5082	2500			3554
抚养比	0.77	0.5			0.68
家庭规模(单位: 人)	4	4			3.3
土地面积(单位: 亩)	2	2			2
性别比例	0.53	1			0.55
劳动力人口受教育年限(单位: 年)	7.53	9			7.72
劳动力人口平均年龄(单位: 岁)	39.46	40			39.78

## 2. 数据描述

本文主要使用了三套数据。第一套数据为样本县初中毕业生的中考信息,数据来源于样本县教育局,时间跨度为2014—2019年。这一数据库记录了全县所有参加中考的学生姓名、性别、出生日期、身份证号等个人信息,中考总分和各科分数等成绩信息,以及初中所在学校、班级编号、是否参加中考等信息。第二套数据为样本县初中毕业生的高中选择数据库,由普通高中学生信息和职高学生信息构成。其中,普通高中的学生数据覆盖2012—2019年全县高中学生的姓名、性别、出生日期、身份证号等个人信息和所在学校名称等学业信息。职高学生数据还包括了所学专业信息、学费减免和奖学金发放情况。第三套数据是2014—2019年样本县建档立卡贫困户数据库,来自样本县脱贫攻坚指挥部,记录了建档立卡贫困户的家庭成员信息和家庭收入信息。此外,还记录了精准扶贫期间的具体帮扶政策,包括此户对应的扶贫第一书记和帮扶人的姓名、性别、年龄、受教育程度和帮扶人的工作单位。<sup>①</sup> 本文将以上不同来源的学生信息及建档立卡贫困户数据通过身份证号进行1:1匹配,构建包含学生基本信息、中考成绩、高中去向信息以及家庭信息的学生层面数据库。本文使用的数据优势在于,由于此套行政数据包含了县域内全部初中毕业生的信息,极大地避免了抽样调查数据中由于农村家庭迁移、子女随迁等导致的样本选择问题。

学生和家庭层面主要变量的描述性统计见表2。其中Panel A展示了所有初中毕业生的相关变量描述性统计。2012—2019年样本县平均每年参加中考的人数为4711人,期间共有37688个观测值。初中毕业生参加中考的比例(*Participate*)约为95%,表明样本的选择性问题非常小。Panel B为建档立卡贫困家庭学生的相关变量。样本中贫困生共有5063个观测值,占全部学生的13.4%。贫困家庭人均收入(*perincome*)的平均值为5783元,2015年农村居民人均纯收入为10772元,贫困人口人均收入仅为2015年农村人均纯收入的53.68%,建档立卡贫困户的收入远低于农村家庭的平均水平。<sup>②</sup>

<sup>①</sup> 脱贫攻坚期间针对建档立卡贫困户建立扶贫第一书记制度和帮扶人制度,制度细节与实施情况参见Zhang et al. (2021)、汪宗金等(2021)、杨晓婷等(2022)。

<sup>②</sup> 值得说明的是,全县所有初中毕业生都可以选择到县外就读普高或者职高,但这种情况相对较少,高中阶段留在县内上学的比例高达96%以上,并且2014—2019年间不存在明显的变化。根据全国学籍管理系统,2017年样本县建档立卡人口中在县外上高中或者职高的比例为3.76%。中国流动人口动态监测调查数据显示,2017年全部县外流动人口当中,5—9岁群体占比为7.6%,而15—19岁群体占比为4.0%。这一组数据说明,在子女进入高中阶段后,随迁比例大幅下降,与样本县的情况一致。

表 2 数据描述性统计

变量名称	变量含义	年份	观测值	均值	标准差
Panel A: 参加中考的学生					
<i>Score</i>	中考成绩(分)	2012—2019	37688	407.88	118.04
<i>Age</i>	学生参加中考时的年龄(岁)	2012—2019	38685	15.627	0.677
<i>Female</i>	性别虚拟变量,女生取值为 1	2012—2019	40068	0.458	0.498
<i>Poor</i>	学生来自建档立卡贫困户取值为 1	2012—2019	40068	0.130	0.336
<i>Participate</i>	初中毕业生参加中考的比例	2014—2019	31058	0.950	0.219
<i>school_dum</i>	学生进入普高或职高取值为 1,不上学为 0	2012—2019	40068	0.911	0.285
<i>HS_dum</i>	学生进入普高取值为 1,职高取 0	2012—2019	36493	0.495	0.500
<i>HS</i>	学生高中选择变量,取值为 0,1,2	2012—2019	40068	1.362	0.640
Panel B: 参加中考的贫困学生					
<i>n5</i>	家庭中 5 岁以下人口	2012—2019	5063	0.077	0.298
<i>n10</i>	家庭中 5—10 岁人口	2012—2019	5063	0.237	0.483
<i>n15</i>	家庭中 10—15 岁人口	2012—2019	5063	0.594	0.648
<i>dependent</i>	抚养比	2012—2019	5063	0.897	0.904
<i>edu_max</i>	成年家庭成员的最高受教育年限(年)	2012—2019	5033	3.081	0.792
<i>perincome</i>	家庭人均纯收入(元)	2012—2019	5063	5783.28	4602.80
<i>Inc_rank</i>	家庭人均收入在全村贫困户中的排名	2012—2019	5058	0.108	0.070
<i>Eduhelper</i>	帮扶人来自教育系统则取值为 1	2012—2019	5200	0.159	0.366

(二) 识别策略及回归方程设定

根据本文的高中选择模型,验证上述假说的关键,在于获得一个其他条件保持不变情况下的关于考上大学主观概率( $Pr^C$ )的外生冲击。其中,保证家庭收入等条件相同对于分离风险厌恶程度是至关重要的。由于职业高中普遍存在学费减免现象,即 $P^V=0$ ,相比之下高中学费相对贫困家庭收入较高。正如王浩名和岳希明(2019)指出的,贫困家庭的低收入是产生信贷约束的主要掣肘,因此如果贫困家庭存在紧预算约束,则选择职业高中可能是由于没有办法借贷,无法支付高中学费所导致。为了获取其他条件不变情况下关于考上大学主观概率( $Pr^C$ )的外生冲击,我们聚焦高中录取过程中的正取分数线。正取分数线由学位数决定,学生无法预先估计或干预,因此可以认为正取分数线两侧的学生在各方面(能力、家庭条件等)皆可比。此外,由于正取分数线仅是一轮录取的标准,且超过正取分数线的学生有相当比例选择不上高中,这使得正取分数线左右两侧的学生都可以选择上普通高中或职业高中,即学生面对同样的选择集( $S = \{0, V, HS\}$ )。但是,低于正取分数线作为一个负面的信号,会极大的降低学生的信心以及考上大学的主观概率( $Pr^C$ ),<sup>①</sup>因此可以认为正取分数线左右两侧存在 $\Delta Pr^C > 0$ 。

为验证关于主观概率( $Pr^C$ )的外生冲击是否对不同家庭的高中选择产生异质性影响,我们

<sup>①</sup> 心理学研究显示,学生对于自己成绩的自信程度影响对结果的预期(Williams, 2010)、课程和任务的选择(Bandura, 2001)以及最终结果本身(Pajares & Schunk, 2005)。根据 Bandura et al. (1999) 的理论模型以及实证研究(Schöber et al. 2018),考试成绩是影响自信程度的最重要因素,尤其是重大的事件及考试(Tram & Cole 2000)。

使用“断点-差分法”(RD-DID, 或称“difference-in-discontinuities”)进行分析。主要的结果变量为是否上普通高中( $HS\_dum$ )这一虚拟变量,取值为1表示上普通高中,取值为0代表上职业高中。

由于不同年份的正取分数线根据当年招生学位数量有所浮动,我们将学生标准化成绩与当年标准化正取分数线作差,得到中心-标准化成绩作为驱动变量,具体做法如下所示:

$$S_{it}^* = \bar{S}_{it} - \bar{S}_t^c \quad (1)$$

其中 $\bar{S}_{it}$ 为标准化中考分数, $\bar{S}_t^c$ 为第 $t$ 年的标准化正取分数线, $S_{it}^*$ 为中心化后的中考成绩,是本文的驱动变量。如果学生成绩高于正取分数线,则 $S_{it}^*$ 为正,反之为负。基于驱动变量 $S_{it}^*$ 构建虚拟变量 $D_i$ :学生 $i$ 的中考成绩高于正取分数线则 $D_i$ 取值为1,低于正取分数线则取值为0。

首先对贫困家庭与一般家庭学生分别构造断点回归方程(RD)如下:

$$y_i = \gamma_1 D_i + \gamma_2 f(D_i) + \gamma_3 X_i + \varepsilon_i \quad (2)$$

其中 $y$ 为本文关注的结果变量,为学生的高中选择,如前文定义; $D_i$ 为处理变量; $f(D_i)$ 为 $D_i$ 的多项式函数。 $X_i$ 为考生个人层面的控制变量,包括考生参加中考时的年龄、性别。估计样本的范围为 $-h < D_i < h$ ,其中 $h$ 为选择的带宽。参考Calonico et al. (2019)求最优带宽的方法,利用Epanechnikov核函数构造局部多项式。由于高中选择大多集中在正取录取分数线附近,因此我们采用低于重点高中录取分数线全部学生作为主样本进行回归。 $\gamma_1$ 是核心待估参数,根据前文假说,如果使用贫困学生样本估计得到的 $\gamma_1 > 0$ ,意味着在正取分数线处存在断点。我们注意到以上精确断点回归估计存在潜在干扰:本文根据普通高中第一轮的正取规则计算得到理论上的正取分数线,<sup>②</sup>可能存在测量误差从而导致断点回归估计不准确。为了排除以上干扰,我们参考Almond & Doyle (2011)、Barreca et al. (2016)和王天宇等(2021)使用甜甜圈断点回归估计方法(donut RDD),即将正取分数线上下的观测值去掉后再使用方程(2)估计。

为了直接比较贫困学生相对非贫困学生在断点处的选择差异,我们进一步采用Grembi et al. (2016)提出的“断点-差分法”法(RD-DID),<sup>③</sup>即分别对贫困学生样本和非贫困学生样本进行断点回归后,得到各自的最优带宽 $h^{poor}$ 与 $h^{non-poor}$ ,以其平均值为带宽( $h^* = \frac{1}{2}(h^{poor} + h^{non-poor})$ ),对带宽内样本( $-h^* < S^* < h^*$ )进行以下回归:

$$y_i = \tau_0 + \tau_1 S_{it}^* + D_i(\gamma_0 + \gamma_1 S_{it}^*) + Poor_i[\alpha_0 + \alpha_1 S_{it}^* + D_i(\beta_0 + \beta_1 S_{it}^*)] + \beta_3 X_i + \varepsilon_i \quad (3)$$

其中 $Poor$ 为是否贫困的虚拟变量,如果来自建档立卡贫困家庭则取值为1,否则取值为0; $D$ 为前文定义的处理变量, $S^*$ 为驱动变量, $X_i$ 为考生和家庭层面的控制变量。 $\beta_0$ 即为RD-DID估计的处理效应,即交互项 $D \times Poor$ 的回归系数。

我们使用此方法分别对2016年以前( $0 = P^V < P^V$ )和2016年之后( $0 = P^V = P^V$ )的数据进行回归,分别验证假说1与假说2。

① 由于中考的考试难度在不同年份之间有所波动,为了使考生的实际成绩具有可比性,首先对2014—2019年学生的中考成绩进行标准化处理生成标准化分数。参照王军鹏等(2020)的做法,处理方式如下: $\bar{S}_{it} = \frac{S_{it} - \bar{S}_t}{\sigma_t}$ 。其中,下角标 $i$ 代表第 $i$ 个学生, $t$ 代表学生参加中考的年份, $S_{it}$ 为学生 $i$ 的实际中考分数, $\bar{S}_t$ 为考试年份 $t$ 所有考生的平均分数, $\sigma_t$ 是第 $t$ 年中考分数的标准差, $\bar{S}_{it}$ 即为 $t$ 年考试的学生 $i$ 的标准化成绩。

② 计算方法为根据本地区计划招生的学位数量确定正取分数线。部分地区可能出现计划学位中包含择校名额的情况,造成估计的正取分数线不等于真实分数线。

③ 根据Hsu & Shen (2019),当两个子样本的样本量差距较大时,分组断点回归估计得到的系数不完全具有可比性,可能导致错误的统计推断结果。

## 四、基本回归结果

### (一) 断点回归设计的适用性检验

断点回归设计的应用前提是驱动变量在断点前后不能被完全操纵。理论上,正取分数线由学校在中高考成绩公布后确定的,因此录取分数线不会被考生个人操纵。同时我们使用本文数据库对这一前提进行检验,具体采用 Cattaneo et al. (2018) 提出的方法对学生在正取分数线左右分布的连续性进行检验。<sup>①</sup> 连续性检验显示,在正取分数线两边的差值仅为 0.033,对应的  $t$  统计量为 0.67,显示不存在录取分数线的操纵问题。

其次,断点回归设计需要满足的另一个条件是控制变量在断点前后不存在跳跃。作为平衡性检验,我们分别以各个控制变量(包括是否来自贫困家庭、性别、年龄、家庭收入以及家庭结构)作为被解释变量进行断点回归,<sup>②</sup>结果均不显著,表明学生特征在正取分数线处没有出现显著的断点,证明了断点回归的有效性。

### (二) 普职教育价格差为正时的断点-差分回归结果

图 2a 给出了普职教育价格差为正时贫困学生在高中选择上出现断点的证据。可以观察到,来自建档立卡贫困家庭的学生在正取分数线处的高中选择存在明显的“跳跃”,位于分数线以下的贫困学生选择普通高中的概率出现了显著下跌。非贫困家庭的学生在分数线处也出现选择高中概率下降的情况,但幅度相对较小。值得注意的是,无论是贫困学生还是非贫困学生,录取分数线右侧散点的纵坐标并没有达到 1,说明不是所有分数超过正取分数线的非贫困学生都选择了普通高中。

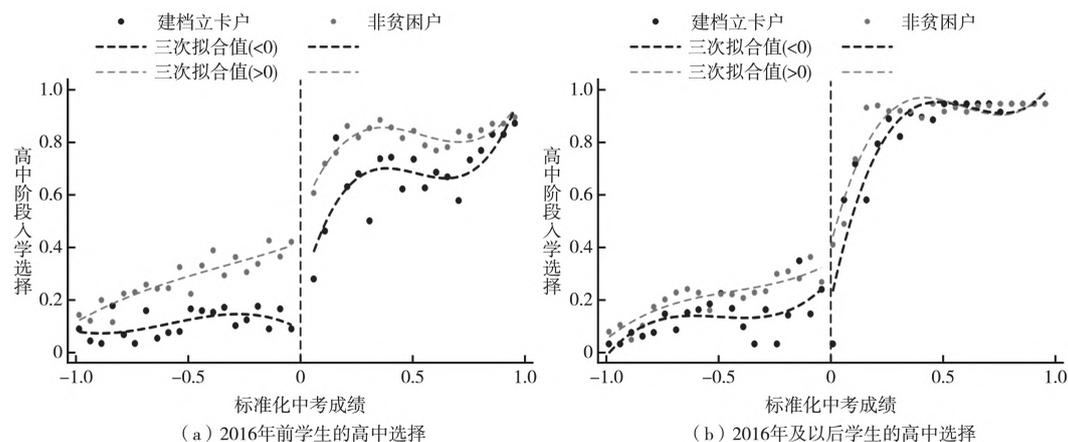


图 2 政策实施前后学生的高中选择

注:横轴为驱动变量  $S^*$ ,纵轴为学生对应的高中选择(0=上职高,1=上普高),黑色和灰色散点分别代表来自建档立卡贫困家庭的学生和非贫困学生,虚线为多项式拟合曲线。

我们首先对假说 1 进行验证,没有超过正取分数线可以看作是关于上大学概率的负面冲击,在两侧学生个人特征及家庭经济状况没有显著差异的前提下,若贫困家庭对这一负面信号的反应更大,即可以验证贫困家庭风险程度更高。首先对贫困家庭和一般家庭分别回归。表 3 报告了这一变量的甜甜圈 RD 估计结果,并且分别报告了常规方法、误差修正方法和稳健方法下的估计值。<sup>③</sup>

① 由于篇幅原因,连续性检验结果未予报告,感兴趣者可以向作者索要。

② 由于篇幅原因,平衡性检验未予报告,感兴趣的读者可以向作者索要。

③ 常规方法指直接进行断点回归的估计结果,误差修正是在常规方法的基础上对带宽边缘的样本进行优化处理,稳健方法则是使用稳健标准误的估计结果。

其中第(1)列与(3)列分别展示贫困学生与非贫困学生的基准回归结果。贫困学生在断点处存在选择上显著的跳跃,而非贫困学生在断点处不存在显著跳跃。这与图2a所展示的一致,即相对于非贫困家庭,正取分数线左侧(即收到上大学概率的负面信号)的贫困家庭学生更倾向于选择无风险职业高中。由于两侧家庭完全可比,贫困家庭在断点处的选择差异排除了紧预算约束的可能。

作为稳健性,进一步将样本扩展为全部参加中考的学生,并将被解释变量扩展为 $HS$ ,包括三种取值( $HS=0$ 为就业, $HS=1$ 为职高, $HS=2$ 为普高)。甜甜圈RD结果如第(2)列和第(4)列所示,贫困样本的估计系仍然显著为正,非贫困学生的估计系数不显著,与前文的估计结果一致。

表3 断点回归基准结果: 职业高中 vs 普通高中

	(1)	(2)	(3)	(4)
	贫困学生		非贫困学生	
	$HS\_dum$	$HS$	$HS\_dum$	$HS$
常规法	0.995** (0.400)	0.530** (0.243)	0.062 (0.056)	0.043 (0.062)
误差修正法	1.006** (0.400)	0.537** (0.243)	0.042 (0.056)	0.021 (0.062)
稳健法	1.006** (0.489)	0.537* (0.290)	0.042 (0.064)	0.021 (0.071)
带宽	0.203	0.283	0.214	0.250
观测值	1449	1773	8801	11324

注: \*、\*\*和\*\*\*分别表示变量在10%、5%和1%水平上显著,括号内为标准误。以下各表同。

接下来,我们使用RD-DID方法,直观比较贫困和非贫困群体之间选择的区别并将回归结果报告于表4 Panel A,①使用的回归带宽为0.2665。②可以看到,不论使用哪一个样本,交互项 $D \times Poor$ 的系数均显著为正,即贫困家庭学生对上大学概率的负面冲击更敏感。这与表3的分样本回归结果一致。为了检验甜甜圈RD的稳健性,我们保留了分数线附近的观测值进行标准的精确断点回归估计,并在Panel B报告了相应回归设定的传统精确RD估计系数,结果显示传统断点回归中的估计值与甜甜圈断点回归的结果非常相近。

表4 RD-DID 回归结果

	(1)	(2)
	$HS\_dum$	$HS$
Panel A: 甜甜圈 RD-DID 回归结果		
$D \times Poor$	1.587** (0.722)	1.575* (0.839)
观测值	1801	1875
$R^2$	0.242	0.193

① 由于篇幅原因,此处仅展示关键系数。如有需要,可向作者索要完整回归结果。

② 根据RD-DID的回归设计,此处选择的带宽 $h=1/2*(0.283+0.250)$ 。

续表 4

	(1)	(2)
	<i>HS_dum</i>	<i>HS</i>
Panel B: 精确 RD-DID 回归结果		
<i>D × Poor</i>	1.735 <sup>***</sup> (0.651)	1.778 <sup>**</sup> (0.765)
观测值	1945	2027
R <sup>2</sup>	0.231	0.184
Panel C: 数值模拟结果		
Coefficient	0.933 <sup>***</sup>	0.487 <sup>***</sup>
95% 置信区间	[-0.252 0.228]	[-0.284 0.263]
Coefficient(左侧)	-0.899 <sup>***</sup>	-0.474 <sup>***</sup>
95% 置信区间	[-0.182 0.197]	[-0.176 0.231]
Coefficient(右侧)	0.03	0.012
95% 置信区间	[-0.182 0.197]	[-0.176 0.231]

Bertrand et al. (2004) 和 Rosenbaum(2001) 推荐在非完全随机设定处理组的情况下,使用数值模拟方法计算标准误。因此,为了更稳健地验证贫困与非贫困学生之间差距的存在,以及分别研究这两个群体在断点两侧的选择差异,我们进行了数值模拟得到两组差值的安慰剂效应分布,并计算处理效应的置信区间。<sup>①</sup> 表 4 的 Panel C 第(1) — (2) 行展示了两组断点系数之差  $\Delta\gamma_1$  的估计值和数值模拟生成的置信区间,结论与前面的结果一致。

使用数值模拟的另一个优势是能够捕捉到断点回归两侧考生选择的不对称情况,为比较断点两侧的回归结果提供了可行性。表 4 的 Panel C 第(3) — (6) 行报告了断点左右两侧的估计系数之差  $\Delta\gamma_1^-$ 、 $\Delta\gamma_1^+$  和相应的置信区间。估计结果均显示断点左侧的回归系数之差显著为负,表明低于分数线的贫困学生选择普通高中的概率显著低于非贫困学生。断点右侧回归系数不显著异于 0,表明当且仅当受到负面冲击时,不同家庭的选择才会出现差异,而成绩高于录取分数线时,没有显著的区别。断点左侧和右侧的非对称性证明了贫困学生和非贫困学生高中选择的差异主要来自录取分数线以下的样本,进一步证明了风险厌恶在高中选择中的重要作用。<sup>②</sup>

综上所述,本节稳健地证明在正取分数线附近,贫困家庭的学生在风险升高时,更倾向于选择无风险的职高,表明贫困家庭的风险厌恶程度更高,假说 1 得到验证。

### (三) 普职教育价格差消失时的断点 - 差分回归结果

下面我们利用 2016 年精准扶贫教育政策的准自然实验,检验普职教育价格差为 0 时的贫困户

① 数值模拟具体步骤如下:(1) 按照据贫困学生占中考总人数的比例在样本中随机抽取 13% 作为“贫困家庭”,将他们定义为“伪处理个体”。(2) 使用伪处理个体分别对贫困和非贫困两个样本进行断点回归估计,并得到两组系数之差。(3) 重复前两个步骤 2000 次,得到两组差值的安慰剂效应分布,进一步计算得到处理效应的置信区间。由于篇幅原因,未在正文中报告结果,感兴趣的读者可以向作者索要。

② 一个潜在的样本选择问题是,如果贫困家庭比非贫困家庭的孩子有更高的概率放弃中考,那么上述回归可能有偏。根据全国中小学学籍管理系统的初中毕业信息,我们计算得到 2014 年初中毕业生中没有参加中考的比例为 14.41%,2015 年这一比例为 10.75%,占比均较低。进一步使用这两个年份数据分别用变量 *Poor* 对是否弃考进行 Probit 回归,得到的估计系数分别为 0.24% (p 值 = 0.851) 与 -0.45% (p 值 = 0.577),在经济意义和统计上均不显著,证明不存在样本选择问题。

行为变化。这一政策对贫困户来说是外生的,因为贫困户的认定时间为2014年(李芳华等,2020;黄薇和曹杨,2022),且认定标准由政府决定,与贫困户子女中考的时间无关。我们也使用平衡性检验验证在价格差改善前后建档立卡户的家庭特征是否有显著区别:分别将表示政策实施前后的虚拟变量 *After16* 回归到不同的家庭特征变量上,估计系数均不显著,表明2016年前后贫困户的家庭特征没有显著性差异。<sup>①</sup>图2(b)展示了2016年及以后贫困学生的高中选择情况。与图2(a)相比,来自建档立卡贫困家庭的学生在正取分数线处的高中选择的断点已经消失,与非贫困家庭的学生在分数线附近的选择差异不明显。我们对2016—2019年参加中考的贫困和非贫困学生分别进行了甜甜圈RD回归估计,结果显示2016年及以后的贫困学生在分数线附近不存在高中选择的显著断点,与非贫困学生的高中选择没有统计上的显著差异。<sup>②</sup>

2016年及以后的RD-DID结果如表5 Panel A第(1)列所示,交互项  $D \times Poor$  的回归系数不显著。为了进一步验证这一变化是否在普职教育价格差变为0后立即发生,我们将样本限制为价格差为0的第一年(即2016年)进行同样的回归。如第(2)列所示,交互项系数同样不显著,表明2016年贫困生和非贫困生在职高和普高之间的选择在分数线处均不存在显著断点。这说明消除教育价格差异后,贫困家庭的高中选择立刻产生了变化,与非贫困家庭的选择不存在显著差异。参照上文 Panel B 报告了使用数值模拟得到的估计系数和置信区间。可以看到2016—2019年样本的交互项估计系数为-0.022,置信区间表明系数在统计上不显著。同样,断点左侧和右侧的回归系数均不显著。将样本扩展为全部参加中考的学生得到与以上一致的估计结果,证明回归结果十分稳健。<sup>③</sup>

以上结果表明,当普职教育价格之差下降至0,选择高中的净收益显著上升时,在正取分数线以下的贫困学生选择高中的概率较之前显著提高,且与非贫困学生的选择没有差异,假说2得到验证。

表5 2016年及以后RD-DID回归结果

	(1)	(2)
	<i>HS_dum</i>	
	2016—2019年	2016年
Panel A: RD-DID 结果		
$D \times Poor$	0.480 (1.294)	-0.214 (0.711)
观测值	1628	2243
带宽	0.188	0.261
R <sup>2</sup>	0.295	0.369
Panel B: 数值模拟结果		
Dif in Coefficient	-0.022	0.228
95% 置信区间	[-0.282 0.204]	[-0.426 0.444]
Dif in Coefficient(左侧)	-0.052	-0.2
95% 置信区间	[-0.171 0.230]	[-0.228 0.335]
Dif in Coefficient(右侧)	-0.074	-0.028
95% 置信区间	[-0.171 0.230]	[-0.228 0.335]

① 由于篇幅原因,未在正文中报告这一回归结果,感兴趣的读者可以向作者索要。

② 由于篇幅原因,未在正文中报告2016年及以后的甜甜圈RD结果,感兴趣的读者可以向作者索要。

③ 由于篇幅原因,未在正文中报告这一回归结果,感兴趣的读者可以向作者索要。

(四) 双重差分回归结果

上述断点回归检验了中考成绩在正取分数线附近的学生的普通高中选择,估计了局部平均处理效应(LATE)。我们还将样本范围扩展到全部存在高考风险的学生,即考试成绩低于“高中学位数×78.58%”的学生<sup>①</sup>,以非贫困学生作为控制组,采用双重差分法检验不同教育价格差下,贫困学生相对非贫困学生的选择差异。由于本文的被解释变量为分类变量,采用logit模型进行回归,设定如下:

$$Y_{it} = \alpha_0 + \alpha_1 Poor_i + S_i^* (\alpha_2 + \alpha_3 Poor_i) + After_{it} [\alpha_4 + \alpha_5 Poor_i + \alpha_6 S_i^* \times Poor_i] + \varepsilon_i \quad (4)$$

其中,被解释变量、虚拟变量  $Poor$ 、标准化中考分数  $S_i^*$  与前文定义相同。变量  $After$  表示政策实施时间的虚拟变量,若参加中考的年份在2016年及以后,取值为1,反之则为0。我们直接使用  $S_i^*$  控制家庭对学生考上大学的主观概率,因此变量  $Poor$  的系数  $\alpha_1$  和交互项  $S^* \times Poor$  的系数  $\alpha_3$  表示2016年之前贫困学生与非贫困学生在上大学的主观概率可比情况下的初始差异,交互项  $Poor \times After$  的系数  $\alpha_5$  和  $S^* \times Poor \times After$  的系数  $\alpha_6$  为我们所关注的DID系数,表示普职教育价格差变动对贫困家庭学生选择的影响。

DID结果显示,变量  $Poor$  显著为负,说明贫困学生比非贫困学生在普职教育价格差为正时,选择普通高中的概率更低。 $Poor \times After$  系数为正,显示教育价格差为零减小了不同家庭的选择差异,三重交互项  $S^* \times Poor \times After$  的系数显著为正,表明随着标准化分数(上大学概率)的上升,贫困学生选择高中的改善幅度更大。<sup>②</sup>

为了得到政策对高中选择差异改善的程度,我们根据DID回归结果的系数计算了贫困和非贫困学生对高中选择的边际效应(图3)。图3(a)显示2016年前贫困学生和非贫困学生上高中的概率之差在整个区间显著为负,正取分数线处贫困学生的边际效应为-8.27%。图3(b)中2016年及以后贫困学生和非贫困学生之间上高中的概率之差出现大幅下降。当学生标准化成绩超过0后两组学生之间的差距急剧减小,当标准化分数超过0.3时已经不存在显著差距。同时观察到,当学生分数显著低于正取分数线,高中选择的差距仍然显著,但是绝对值较2016年之前有所减小(-2.87%)。

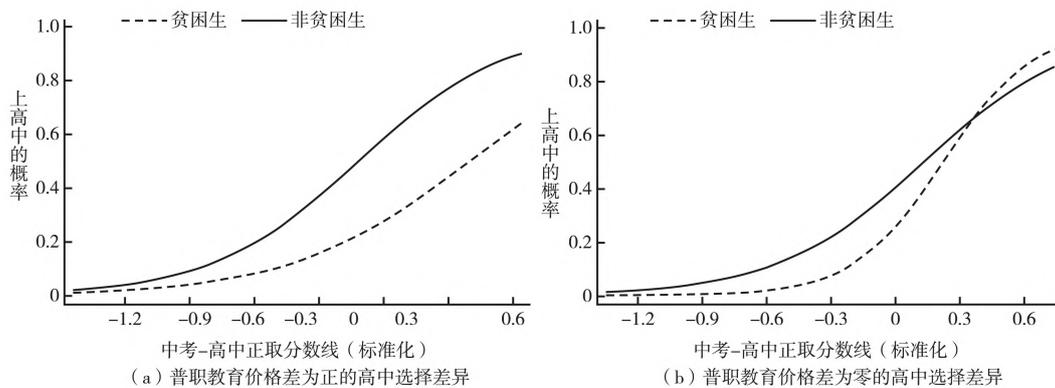


图3 普职教育价格差变化的高中选择差异

① 样本县在河南省,研究区间内,河南的高考录取率为78.58%。

② 由于篇幅原因,DID结果在此不做展示。作为稳健性,我们使用全部中考考生、全部低于正取分数线的学生重复上述DID回归,得到相似的结果。以上结果如有需要,可向作者索要。

以上双重差分回归结果再一次验证了假说 1, 即贫困家庭的风险厌恶程度高于非贫困家庭。当样本从正取分数线附近扩展到全部风险区间后, 普高与职高价格差的降低普遍提高了贫困学生选择普通高中的概率。虽然断点 - 差分法证明在正取分数线附近的贫困生与非贫困生无显著差异, 但是即使在价格差为零时, 风险较高的贫困生(即上大学的概率  $Pr^C$  较低) 选择普通高中的概率仍然显著低于非贫困生, 这间接地证明了贫困家庭的风险厌恶程度的确高于非贫困家庭。

## 五、进一步讨论

本节首先使用不继续上学和职业高中之间的选择, 排除“贴现因子”这一可能干扰结果解读的机制, 然后使用收入变动与帮扶人制度排除其他影响贫困与非贫困家庭行为差异的可能机制。最后, 扩展上一部分的结果, 进一步探讨贫困家庭在男女教育投入上的异质性。

### (一) 可能的机制 1: 贴现因子

在本文的理论模型中, 得到两个假说的关键假设是两类家庭的贴现率相同, 否则无法仅通过家庭在相同上大学概率下的不同选择, 唯一识别风险厌恶程度的差异。<sup>①</sup> 验证这一假设的思路如下: 当家庭面临子女不上学 ( $S=0$ ) 与无风险职业高中 ( $S=V$ ) 两个选择时, 不上学的折现净收益可以表示为:  $R_{i,0}^0(\beta_i, \mu^M) = \sum_{t=0}^T \beta_i^t w_i^M$ , 职业高中的折现净收益如前文所述为:  $R_{i,0}^V(\beta_i, \mu^V, P^V) = \sum_{t=3}^T \beta_i^t w_i^V - P^V$ 。家庭的最优决策是在两者之间选择折现净收益最高的选项。由于  $w^M, \mu^V, P^V$  均为外生给定, 因此两者折现净收益之差完全由贴现因子  $\beta_i$  决定。由此可知, 在相同的成绩下, 不同家庭  $i$  在不上学与职高之间的选择, 可唯一识别贴现因子  $\beta$ ,<sup>②</sup> 即在相同成绩下若更倾向于选择不上学, 则该家庭的贴现因子更低(即贴现率更高); 反之, 则不存在贴现因子的个体差异。综上所述, 通过比较贫困和非贫困家庭在不上学 ( $S=0$ ) 与无风险职业高中 ( $S=V$ ) 间的选择是否有显著差异, 判断是否存在贴现因子的差异。

我们定义虚拟变量 ( $occ\_dum$ ): 取值为 1 表示上职业高中, 取值为 0 表示不继续上学; 将方程 (3) 的被解释变量替换为  $occ\_dum$  进行同样的断点回归和 RD-DID 回归, 断点回归结果显示贫困和非贫困样本的估计系数均不显著, 可见贫困家庭与非贫困家庭在断点处选择相同, 都不存在显著跳跃, 且其行为与非贫困学生没有区别。<sup>③</sup> 表 6 展示了 RD-DID 结果, 使用数值模拟方法得到估计系数和置信区间, 结果不显著, 即贫困学生和非贫困学生在断点处的选择不存在差异。<sup>④</sup> 作为稳健性, 我们将被解释变量替换为代表是否继续上学的虚拟变量  $school\_dum$ , 取值为 0 表示不继续上学, 取值为 1 表示继续上学(职高或普高), 然后进行回归, 结果如第 (2) 列所示。三种方法估计得到的交互项系数均不显著, 稳健地验证了贴现因子相同这一重要假设。

① 如果贴现率下降, 则高中毕业与大学毕业的折现净收益同时下降, 且后者下降幅度更大。对应图 1 中 X 轴上的  $R^H$  和  $R^C$  同时向左平移, 即使效用函数保持不变, 上高中的概率也可能下降。由于篇幅原因, 如果需要具体的推导, 可向作者索要。

② 不同家庭对不同学历的预期回报率可能存在差异, 即可能存在  $\varepsilon_i$ , 家庭层面的预期工资率可以表示为  $w_i^m(1+r^s) - \varepsilon_i$ 。 $\varepsilon_i$  可以理解文献中所探讨的教育观念, 例如上学无用论或者女性上学无用论等(李涛和邬志辉, 2015; 蒋承等, 2018; Jayachandran 2015; Choi & Hwang 2015; 杜丽群和王欢 2021; 黄政 2021)。从建模的角度而言, 设定回报率家庭差异  $\varepsilon_i$  与贴现因子差异等同, 即对任意  $\varepsilon_i$ , 存在  $\beta_i$ , 使得  $w_i^m(1+r^s) - \varepsilon_i = \beta_i w_i^m(1+r^s)$ 。因此从简化模型的角度, 我们用贴现因子涵盖所有的教育观念差异。

③ 由于篇幅原因, 未在正文中展示断点回归基准结果。如有需要, 可向作者索要完整回归结果。

④ 本文还进行了甜甜圈 RD 估计和精确 RD 估计, 估计系数均不显著。由于篇幅原因, 未在正文中展示结果。如有需要, 可向作者索要完整回归结果。

表6 贴现因子检验的 RD-DID 数值模拟结果

	(1)	(2)
	<i>occ_dum</i>	<i>school_dum</i>
Coefficient	-0.0197	0.0079
95% 置信区间	[-0.082 0.067]	[-0.197 0.162]
Coefficient(左侧)	0.051	0.015
95% 置信区间	[-0.058 0.054]	[-0.146 0.101]
Coefficient(右侧)	0.031	0.023
95% 置信区间	[-0.058 0.054]	[-0.146 0.101]

注: 根据 RD-DID 的回归设计, 此处选择的带宽  $h = 1/2 * (0.398 + 0.508) = 0.453$ 。

(二) 可能的机制 2: 收入增长

可能导致上述结果的机制之二为, 2016 年贫困生所属的建档立卡户家庭与非贫困家庭的收入差距消失。但实际上, 样本县 2018 年正式脱贫, 2016 年建档立卡户的人均家庭年收入为 5258 元, 远低于当年样本县所在省农村家庭人均可支配收入 18443 元。<sup>①</sup> 此外, 2015 年贫困户的平均家庭人均收入比 2014 年增加 30.66%, 这一增幅与 2015—2016 年的收入增幅非常相近。<sup>②</sup> 如果高中选择差异的消失是由收入提高造成的, 应观察到贫困学生 2015 年的高中选择比 2014 年显著改善。因此, 我们使用前文所述的 RD-DID 方法对 2014 年和 2015 年的贫困学生样本进行以下估计:

$$y_i = \tau_0 + \tau_1 S_i^* + D_i(\gamma_0 + \gamma_1 S_i^*) + After15_{it} [\alpha_0 + \alpha_1 S_i^* + D_i(\beta_0 + \beta_1 S_i^*)] + \beta_3 X_i + \varepsilon_i \quad (5)$$

其中  $D, S^*, X$  均为前文定义的变量,  $After15$  代表 2015 年的虚拟变量, 学生参加高考的年份为 2015 年取值为 1, 否则取值为 0。回归使用的带宽等于 2014 年考生和 2015 年考生分别做断点回归估计使用的带宽平均值。 $\beta_0$  为 RD-DID 估计的处理效应, 即交互项  $D \times After15$  的回归系数。表 7 第(1)列和第(2)列使用不同样本得到的  $D \times After15$  交互项的估计系数均不显著, 表明 2014 年和 2015 年的高中选择没有出现显著差异, 证明收入上升并非是导致贫困家庭行为改变的机制。我们进一步使用双重差分估计, 选取 2014—2015 年作为样本区间, 通过将 2015 年作为虚拟的政策实施时点进行回归估计, 回归结果显示年份虚拟变量  $After15$  与贫困状态变量  $Poor$  的交互项系数均不显著,<sup>③</sup> 再一次证明收入并非产生基准结果的机制。

表7 收入增长与教育观念改变的 RD-DID 回归结果

	(1)	(2)	(3)	(4)
	<i>HS_dum</i>	<i>HS</i>	<i>HS_dum</i>	<i>HS</i>
$D \times After15$	-0.220 (1.411)	-0.184 (1.601)		
$D \times Eduhelper$			0.168 (1.245)	-0.129 (1.302)

① 数据来源: <http://www.henan.gov.cn/2017/01-26/369749.html>。

② 2014 年贫困人口人均纯收入为 2955 元, 2015 年为 3861 元, 2016 年为 5193 元。2014—2015 年收入增幅为 30.66% =  $(3861 - 2955) / 2955$ ; 2015—2016 年收入增幅为 34.50% =  $(5193 - 3861) / 3861$ 。

③ 由于篇幅原因, 双重差分结果未予报告, 感兴趣可以向作者索要。

续表 7

	(1)	(2)	(3)	(4)
	<i>HS_dum</i>	<i>HS</i>	<i>HS_dum</i>	<i>HS</i>
控制变量	是	是	是	是
观测值	180	182	502	512
R <sup>2</sup>	0.346	0.305	0.348	0.338

### (三) 可能机制 3: 教育观念突变和信息差消失

另一个可能的影响机制为 2016 年之前贫困家庭的教育观念落后,对子女接受教育持有消极或者负面的态度,或者因为信息不足而错误估计对不同教育选择的期望收益,或者低估上大学的概率,而 2016 年这一观念突然转变或信息不足的情况突然消失。

对于这一机制,我们利用对建档立卡贫困户实施的“帮扶责任人”制度来进行检验。“帮扶责任人”是 2014 年启动的一项通过为贫困家庭分配政府部门工作人员帮助他们尽快脱贫的政策。如果以上观察到的结果来自于教育观念的突变或信息不足的改善,那么会存在因为帮扶人专业优势导致的结果差异。具体而言,如果帮扶人来自教育系统(包括教育部门和普通高中),则他们更有可能使用自己的工作经验和信息优势对贫困家庭的教育观念产生影响,帮助贫困父母为孩子的高中选择进行决策。因此,可以通过检验帮扶人是否存在异质性影响,检验这一可能的机制。

首先对贫困户对应的帮扶人进行平衡性检验,检验显示帮扶人来自教育系统与贫困户的家庭特征无明显相关性。<sup>①</sup> 然后使用 RD-DID 方法对贫困学生家庭进行以下估计:

$$y_i = \tau_0 + \tau_1 S_i^* + D_i(\gamma_0 + \gamma_1 S_i^*) + Eduhelper_i[\alpha_0 + \alpha_1 S_i^* + D_i(\beta_0 + \beta_3 S_i^*)] + \beta_3 X_i + \varepsilon_i \quad (6)$$

其中  $X$  均为前文定义的变量。 $D S^* Eduhelper$  表示帮扶人背景的虚拟变量,若来自教育系该取值为 1,否则取值为 0。方程使用的带宽等于有教育帮扶和无教育帮扶的两个学生组分别做断点回归估计使用的带宽平均值。表 7 第(3)列和第(4)列使用不同被解释变量得到的  $D \times Eduhelper$  交互项的估计系数均不显著,表明不同帮扶人对高中选择没有显著差异。我们还将变量  $Eduhelper$  和交互项  $S^* \times Eduhelper$  作为核心解释变量进行回归,结果表明二者的系数均不显著,<sup>②</sup>说明  $Eduhelper$  对各分数段贫困生的高中选择均没有显著影响,即家庭教育观念并不是导致贫困家庭行为改变的机制。

### (四) 性别异质性

贫困家庭中普遍存在重男轻女现象(杜丽群和王欢 2021),面临严重预算约束的家庭对女孩的轻视直接体现在教育投资方面(Jayachandran 2015),其所在地区的文化、社会偏好和社会规范也导致家庭对女孩的教育投资少于男孩。为了检验贫困女生是否在高中选择上比贫困男生面临更大的劣势,首先使用 RD-DID 对于正取分数线附近的样本进行检验。<sup>③</sup> RD-DID 的回归方程如下所示:

$$y_i = \tau_0 + \tau_1 S_i^* + D_i(\gamma_0 + \gamma_1 S_i^*) + Female_i[\alpha_0 + \alpha_1 S_i^* + D_i(\beta_0 + \beta_1 S_i^*)] + \beta_3 X_i + \varepsilon_i \quad (7)$$

其中  $D_i S^* X$  均为前文定义的变量,  $Female$  为考生的性别虚拟变量。 $\beta_0$  为 RD-DID 估计的处理效应,即交互项  $D \times Female$  的回归系数。使用的带宽等于女生和男生分别做断点回归估计使用

<sup>①</sup> Zhang et al. (2021) 对帮扶人制度进行了详细介绍和评估,并从机制上说明了分配的随机性。由于篇幅原因,未予报告,感兴趣的读者可以向作者索要。

<sup>②</sup> 由于篇幅原因,交互项回归未予报告,感兴趣的读者可以向作者索要。

<sup>③</sup> 我们直接使用全样本回归得到的带宽,其原因在于,断点回归需要依靠大样本渐进理论进行统计推断,对样本量有较高要求。将贫困生按性别分为子样本分别估计会导致样本量大幅减少,降低了断点回归估计的准确性。

的带宽平均值。虽然性别虚拟变量的系数不显著,但是我们发现将贫困生按照性别分组,分别估计会导致样本数量大幅减少,由于断点回归需要依靠大样本渐进理论进行统计推断,这一操作降低了断点回归估计的准确性。<sup>①</sup>

因此,本部分沿用上文的思路对贫困学生使用多项 logit 和 logit 模型将性别虚拟变量 *Female*、中考成绩  $S^*$  以及与性别的交互项  $S^* \times Female$  回归到学生的高中选择上。为了直观展示普职教育价格差的消失对不同性别学生的异质性影响,图4报告了2016年前和2016年及以后的男生、女生高中选择的边际效应。<sup>②</sup>可以看到,2016年以前,当标准化分数高于-0.3时,男性和女性的边际效应出现了巨大的差异,且随着分数的增加而增大。这意味着相同上大学概率下,贫困家庭更倾向于投资男生而非女生。但是,当普职教育价格差在2016年消失以后,男性和女性的边际效应差异基本消失,即教育价格差的消失对贫困女生的福利改善更大。

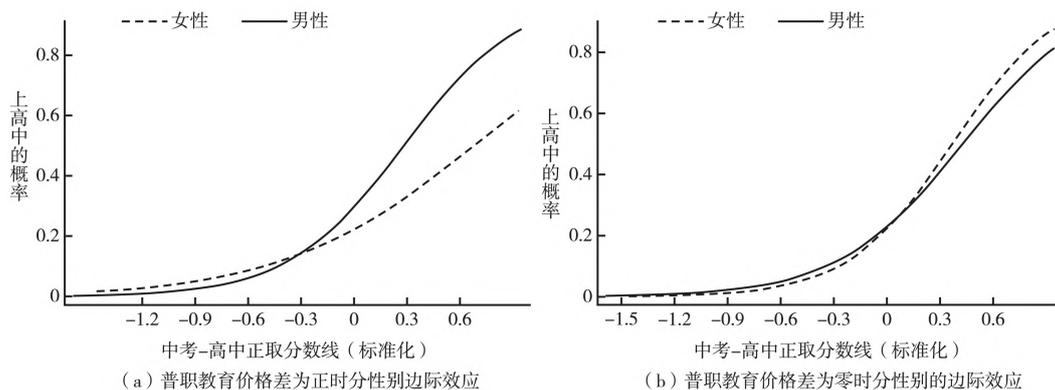


图4 不同普职教育价格差下性别的边际效应

## 六、结论和启示

中国高中阶段的学校选择为验证贫困家庭的偏好参数与一般家庭是否存在显著差异提供了一个绝佳的现实设定。本文基于2014—2019年初中毕业生的高中选择和2016年针对建档立卡户的准自然实验,使用断点-差分法和双重差分法得到以下发现:第一,在职业高中价格低于普通高中价格时,如果受到关于上大学概率的负面冲击,在其他条件不变的情况下,贫困家庭显著降低选择普通高中(高风险高回报)的概率,而一般家庭没有显著变化;第二,贫困家庭学生“职高 vs 普高”的选择显著地受到教育价格的影响,在普职教育价格差降低为0时,上述选择差异大幅降低,且在风险较低的分段(正取分数线附近),这一差异几乎完全消失。以上两个结果共同显示贫困家庭的风险厌恶程度高于一般家庭。另一方面,在职高教育成本为零的前提下,贫困生在“不上学 vs 上职高”的选择上没有差距,这一现象不论在正取分数线附近还是在全分数段区间皆成立。这证明低收入家庭的贴现因子与一般家庭没有显著区别。我们进一步排除了收入上涨机制、教育观念突变机制和信息不足消失机制,稳健地证明了上述结论。

本文的发现为贫困家庭风险偏好差异提供了来自实验室外的现实证据,厘清了农村家庭教育高中阶段入学选择的微观机制,对于政府进一步推动“普职协调发展”具有重要意义。本文的模型显示,职高与普通高中之间的选择可以看作无风险中等收益与风险高收益之间权衡的最优化决策,政府可以通过影响模型中的重要参数,调节学生在职高与普通高中之间的选择,包括不同学历的教

① 由于篇幅原因, RD-DID 回归未予报告,感兴趣的读者可以向作者索要。

② 由于篇幅原因,此处只展示边际效应,多项 logit 和 logit 回归结果未予报告,感兴趣的读者可以向作者索要。

育价格向量、教育回报率向量以及高考风险系数。在其他参数不变的情况下,不同教育选择的价格是调节低收入家庭教育决策的有力政策工具。但是,随着脱贫攻坚的结束和低收入群体收入的上升,未来价格工具的作用空间可能逐渐减小,因此政府可以考虑除价格工具外的其他工具。若政府希望提高选择职业高中的比例以填补国家技能型人才的缺口,可以通过进一步增加职业高中的选择空间、扩大升学空间和就业途径,提高职业教育的回报率,或调节高中生升入大学的概率,即提高普通高中的升学风险,以提高职业教育作为无风险收益的吸引力。

## 参考文献

- 陈胜祥 2011 《农村中职免费政策失灵:表现、成因与对策——基于浙、赣、青三省的调查》,《教育科学》第5期。
- 陈永伟、顾佳峰、史宇鹏 2014 《住房财富、信贷约束与城镇家庭教育开支——来自 CFPS2010 数据的证据》,《经济研究》增1期。
- 杜丽群、王欢 2021 《家庭经济视角下人力资本理论研究进展》,《经济学动态》第5期。
- 方长春、风笑天 2005 《阶层差异与教育获得——一项关于教育分流的实证研究》,《清华大学教育研究》第5期。
- 胡咏梅、陈纯槿 2013 《农村职业教育投资回报率的变化:1989—2009年》,《教育与经济》第1期。
- 侯风云 2004 《中国农村人力资本收益率研究》,《经济研究》第12期。
- 蒋承、刘霄、戴君华、金文旺 2018: 当前农村高中教育的发展瓶颈与应对策略,《中国教育学刊》第1期。
- 李芳华、张阳阳、郑新业 2020 《精准扶贫政策效果评估——基于贫困人口微观追踪数据》,《经济研究》第8期。
- 李涛、邬志辉 2015 《“乡土中国”中的新“读书无用论”——基于社会分层视角下的雍村调查》,《探索与争鸣》第6期。
- 李中国、彭李 2015 《被裹挟的普通高中教育——基于近30年〈教育部工作要点〉的词频分析》,《湖南师范大学教育科学学报》第2期。
- 李菁、林毅夫、姚洋 2002 《信贷约束、土地和不发达地区农户子女教育投资》,《中国人口科学》第6期。
- 李力行、周广肃 2015 《家庭借贷约束、公共教育支出与社会流动性》,《经济学(季刊)》第1期。
- 刘利利、刘洪愧 2020 《个人所得税改革与家庭教育支出——兼论教育负担与教育差距》,《经济科学》第1期。
- 陆万军、张彬斌 2021 《大学扩招、就业挤压与中等职业教育收益变迁》,《经济学动态》第12期。
- 黄政 2021 《读书真的无用?——农村居民教育期望的形成机制研究》,《教育科学研究》第6期。
- 黄薇、曹杨 2022 《常态化精准扶贫政策的完善:反福利依赖的视角》,《经济研究》第4期。
- 欧阳忠明、王江雁 2022 《何以成为缓冲器:德国职业教育促进就业的变迁与坚守》,《河北师范大学学报(教育科学版)》第3期。
- 祁占勇、何佑石 2022 《我国中等职业教育城乡个体回报率变动研究——基于2008—2017年CGSS的实证分析》,《西南大学学报(社会科学版)》第2期。
- 任伟聪、梁若冰 2022 《农村家庭税费负担与人力资本积累——基于农村税费改革的经验分析》,《财政研究》第5期。
- 沈有禄 2017 《谁上中职?谁上普高?有何差异》,《中国职业技术教育》第11期。
- 苏丽锋、孙志军、李振宇 2016 《高中阶段教育选择影响因素研究——基于高中与中职在校学生的调查分析》,《清华大学教育研究》第4期。
- 苏丽锋、孙志军、李振宇 2017 《初中后教育选择意愿及影响因素研究——普高、中职还是不再读书》,《华中师范大学学报(人文社会科学版)》第5期。
- 汪崇金、杨亿、谷军健 2021 《第一书记驻村帮扶能提升乡村社会资本吗?——一项田野实验研究》,《财经研究》第3期。
- 王军鹏、张克中、鲁元平 2020 《近朱者赤:邻里环境与学生学习成绩》,《经济学(季刊)》第2期。
- 王天宇、高秋明、赵丽秋 2021 《医药分开改革中的供方行为:基于医保住院赔付记录的分析》,《世界经济》第3期。
- 王伟同、周洪成、张妍彦 2021 《看不见的家庭教育投资:子女升学压力与母亲收入损失》,《经济研究》第9期。
- 王浩名、岳希明 2019 《贫困家庭子女受教育程度决定因素研究进展》,《经济学动态》第11期。
- 颜敏 2012 《技能高中还是普通高中?——中国农村学生的教育选择》,《中国农村经济》第9期。
- 杨晓婷、陆镜名、刘奕辰、于滨铜 2020 《“资本下沉”赋能“资源释放”:第一书记带动贫困村脱贫的行动逻辑与高效机制》,《中国农村观察》第6期。
- 邹薇、郑浩 2014 《贫困家庭的孩子为什么不读书:风险、人力资本代际传递和贫困陷阱》,《经济学动态》第6期。
- Almond, D., and Doyle, J., 2011, “After Midnight: A Regression Discontinuity Design in Length of Postpartum Hospital Stays”, *American Economic Journal: Economic Policy*, 3(3), 1—34.

- Ashraf, N., Karlan, D., and Yin, W., 2006, "Tying Odysseus to the mast: Evidence from a Commitment Savings Product in the Philippines", *Quarterly Journal of Economics*, 121(2), 635—672.
- Bandura, A., Freeman, W. H., and Lightsey, R., 1999, *Self-efficacy: The Exercise of Control*, Macmillan.
- Banerjee, A. V., Benabou, R., and Mookherjee, D., 2006, *Understanding Poverty*, Oxford University Press.
- Barreca, A. I., Lindo J. M., and Waddell G. R., 2016, "Heaping-induced Bias in Regression-discontinuity Designs", *Economic Inquiry*, 54(1), 268—293.
- Becker, G. S., and Tomes, N., 1986, "Human Capital and the Rise and Fall of Families", *Journal of Labor Economics*, 4(3 part2), S1—S39.
- Becker, G. S., and Tomes, N., 1979, "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility", *Journal of Political Economy*, 87(6), 1153—1189.
- Bertrand, M., Duflo, E., and Mullainathan, S., 2004, "How much should We Trust Differences-in-differences Estimates?", *Quarterly Journal of Economics*, 119(1), 249—275.
- Calónico, S., Cattaneo, M. D., Farrell M. H., and Titiunik R., 2019, "Regression Discontinuity Designs Using Covariates", *Review of Economics and Statistics*, 101(3), 442—451.
- Cattaneo, M. D., Jansson, M., and Ma X., 2018, "Manipulation Testing Based on Density Discontinuity", *Stata Journal*, 18(1), 234—261.
- Choi, E. J., and Hwang, J., 2015, "Child Gender and Parental Inputs: No More Son Preference in Korea?", *American Economic Review*, 105(5), 638—43.
- Epple, D., and Romano, R. E., 1998, "Competition between Private and Public Schools, Vouchers, and Peer-group Effects", *American Economic Review*, 88(1), 33—62.
- Falk, A., Kosse, F., and Pinger, P., 2020, "Mentoring and Schooling Decisions: Causal Evidence", IZA Working paper, 13387.
- Fehr, D., Fink, G., and Jack, B. K., 2022, "Poor and Rational: Decision-making under Scarcity", *Journal of Political Economy*, 130(11), 2862—2897.
- Gao, Q., Zhai, F., Yang, S., and Li, S., 2014, "Does Welfare Enable Family Expenditures on Human Capital? Evidence from China", *World Development*, 64, 219—231.
- Grembi, V., Nannicini, T., and Troiano, U., 2016, "Do Fiscal Rules Matter?", *American Economic Journal: Applied Economics*, 8(3), 1—30.
- Haushofer, J., and Fehr, E., 2014, "On the Psychology of Poverty", *Science*, 344(6186), 862—867.
- Hsu, Y. C., and Shen, S., 2019, "Testing Treatment Effect Heterogeneity in Regression Discontinuity Designs", *Journal of Econometrics*, 208(2), 468—486.
- Jayachandran, S., 2014, "The Roots of Gender Inequality in Developing Countries", NBER Working Paper, 20380.
- Lucas, R. E. B., and Kerr, S. P., 2013, "Intergenerational Income Immobility in Finland: Contrasting Roles for Parental Earnings and Family Income", *Journal of Population Economics*, 26(3), 1057—1094.
- Pajares, F., and Schunk, D., 2005, *Competence Perceptions and Academic Functioning*, Handbook of Competence and Motivation, Guilford Press.
- Rosenbaum, P. R., 2001, "Effects Attributable to Treatment: Inference in Experiments and Observational Studies with a Discrete Pivot", *Biometrika*, 88(1), 219—231.
- Schöber, C., Schütte, K., Köller, O., McElvany, N., Gebauer, M. M., 2018, "Reciprocal Effects between Self - efficacy and Achievement in Mathematics and Reading", *Learning and Individual Differences*, 63, 1—11.
- Schultz, T. W., 1964, *Transforming Traditional Agriculture*, New Haven, Yale University Press.
- Schutz, T. P., 1993, "Investment in the Schooling and Health of Women and Man", *Journal of Human Resources*, 28(4), 694—734.
- Shapiro, D., and Tambashe, B. O., 2001, "Gender, Poverty, Family Structure, and Investments in Children's Education in Kinshasa, Congo", *Economics of Education Review*, 20(4), 359—375.
- Tram, J. M., and Cole, D. A., 2000, "Self-perceived Competence and the Relation between Life Events and Depressive Symptoms in Adolescence: Mediator or Moderator?", *Journal of Abnormal Psychology*, 109(4), 753.
- Williams, D. M., 2010, "Outcome Expectancy and Self-efficacy: Theoretical Implications of an Unresolved Contradiction", *Personality and Social Psychology Review*, 14(4), 417—425.
- Zhang, Y., Zheng, X., and Lunyu, X., 2021, "How Do Poverty Alleviation Coordinators Help the Impoverished in Rural China? — Evidence from the Chinese Poor Population Tracking Dataset", *China Economic Review*, 69, 101686.

## A Study of Educational Choices at the High School Level for Poor Families: A Behavioral Economics Perspective

LI Fanghua<sup>a</sup> and XIAO Han<sup>b</sup>

( a: The University of New South Wales; b: School of Applied Economics , Renmin University of China)

**Summary:** High school education is an important part of China's national educational system and plays a key role in talents cultivation. China's high school education mainly includes two options , namely , regular high school and vocational high school. How to guide adolescents with different abilities and preferences to choose the proper type of high school is crucial for economic development and human capital accumulation. China has been adopting the score-based high school admission mechanism supplemented by personal choice , and though families have a certain degree of choice , the score of the secondary school entrance examination is the main basis for deciding what kind of high school children will finally go. Currently , the diversion policy of regular and vocational high school has aggravated the educational anxiety of families and led to overinvestment in compulsory education. Meanwhile , demand for vocational labors continues to rise during China's economic transformation. Therefore , it is important for the government to further promote the education reform by clarifying the micro mechanism of school choice at the high school level for families.

In order to verify whether there are “behavioral” differences between poor and non-poor families and to fill the research gap at the high school education in China , this paper utilizes the education price variation due to the free tuition and student subsidies policy for vocational high school students before 2016 and the natural experiment of educational policy of pairing-off assistance in targeted poverty alleviation started in 2016 as exogenous shocks; and it uses the RD-DID method to identify the heterogeneous effect of education investment between poor and non-poor students. This paper adopts the individual score data and high school choice data of all secondary school students in a county in China , combined with the information of identified poor households. Furthermore , we examine whether high school choice is influenced by family poverty for students in the entire score distribution using the difference-in-differences method. We also use gender information to test whether there is gender heterogeneity in educational decisions of poor families , and discuss the effect of educational perceptions on high school choice using data from targeted poverty alleviation.

This paper has two main findings. Firstly , faced with an equal negative shock to the probability of “being enrolled by universities” , students from poor families are more likely to choose vocational high schools; and this difference in probability decreases significantly when tuition is waived for poor families' regular high schools ( i. e. , the net benefit of regular high schools increases). Secondly , after eliminating four possible explanations , namely , discount factor differences , rising income levels , sudden changes in educational perceptions , and insufficient information , this paper provides sufficient evidence that the difference in high school choice between poor and non-poor families is due to the higher risk aversion of poor families.

The findings of this paper have two policy implications. Firstly , holding all other parameters constant , the education price is a powerful policy tool in order to manage educational decisions of low-income households , and this paper demonstrates that the risk aversion has a significant effect on low-income households by driving them to prefer risk-free vocational high schools. Secondly , in order to increase the proportion of vocational high school choices to fill the gap of national skill-based talents , Chinese government can increase the rate of return of vocational education by offering more choices of vocational high schools , expanding the space for further education and career paths. It can also regulate the probability of high school students being admitted into universities by increasing the risk of further education of students in regular high schools to increase the attractiveness of vocational education as a risk-free option.

**Keywords:** Poverty; Educational Choices; Vocational High School

**JEL Classification:** D01 , J24 , O12

( 责任编辑: 王利娜) ( 校对: 曹 帅)