

高考复读有用吗？

——基于断点回归的因果推断^{*}

张 燕 丁延庆 李虹辉 李 伟 马玉洁 叶晓阳^①

摘要 高考复读现象在中国非常普遍,每年平均有10%—20%的高考学生选择复读。然而现有文献里缺少对高考复读的严格定量研究。本文使用来自宁夏的高考数据,发现理科、男性、农村、高考总分较低的学生更倾向于选择复读;高考复读收益具有异质性,高分学生复读不易有效。我们利用高考录取批次线的外生条件,采用断点回归设计估计高考复读的因果效果,发现二本线下学生更倾向于选择复读,且能取得积极效果。我们的实证发现与来自一线教师的观察是一致的。本文对高考复读的文献、个人决策以及政策讨论都提供了一定参考。

关键词 中国高考; 复读; 教育收益; 断点回归

作者简介 张 燕 / 宁夏教育考试院教师 (银川 750002)

丁延庆 / 北京大学教育学院副教授 (北京 100871)

李虹辉 / 北京大学教育学院硕士研究生 (北京 100871)

李 伟 / 宁夏大学政法学院教授 (银川 750021)

马玉洁 / 宁夏大学政法学院助理教师 (银川 750021)

叶晓阳 / 密歇根大学教育学院博士研究生 (密歇根州 48109)

一、引言

高考复读是学生在参加过一次高考后,选择再接受一次高三教育,再次或多次参加高考的行为。高考复读现象在我国长期存在。复读生包含的不仅有“被动选择”的落榜生,还有大量为了追求更好院校就读机会而“主动选择”的学生。

^{*} 本文系国家自然科学基金2016年面上项目“‘教育致贫’与‘教育脱贫’:现状及相关政策评估——基于国家贫困人口建档立卡数据的研究”(项目编号:71613013)、教育部人文社会科学重点研究基地北京大学教育经济研究所2016年度重大项目“经济新常态下的教育扶贫与教育公平研究”(项目编号:16JJD880001)和宁夏自治区重大科技攻关项目“宁夏教育与扶贫数据库建设及相关政策评估研究”的阶段性成果。

^① 注:共同一作,排名不分先后。我们感谢戴冰清、耿未、杨晋、伍银多的共同努力使得本项目得以顺利进行,陈霜叶、陈晓宇、李锋亮、马莉萍和岳昌君诸师,以及两位匿名评审的宝贵意见。文责自负。

近年来复读群体日益庞大,已经成为非常普遍的社会现象,这不再是学生个人问题,已经发展成为一个重要的公共政策议题。尤其是,1999 年高校扩招以后,学生们不再满足于“有学上”,纷纷追求“上好学”,高考复读率持续升高。2002 年 2 月教育部发文,禁止公办中学占用学校正常的教育资源举办高中毕业复读班,2007 年 8 月 4 日,时任教育部部长周济参加“落实中职国家资助政策及 2007 年招生任务中部片区座谈会”时再次明确表态,从 2008 年起,全国所有公办普通高中禁止开办复读班,并认为这是利用国家资源谋取利益,是对教育公平的挑战。

然而高考复读的现象并没有因此冷却,公办高中通过开设分部,学生们通过涌向私立高考复读班等办法,绕开政策,高考复读班仍在继续,并且愈演愈烈。仅就近几年的情况来说,2015 年陕西省高考往届生 5.1 万人,占高考报名总数的 14.97%;2015 年湖南省高考往届生 4.5 万人,占高考报名总数的 11.5%;2016 年湖北省高考往届生 2.98 万人,占高考报名总数的 8.24%,可见高考复读现象之普遍。

高考复读现象愈演愈烈,而教育资源作为准公共物品,具有稀缺性,是否应该让有限的教育资源,尤其是优质公立高中的教育资源,是否首先满足应届生的受教育需求呢?然而适龄青年,为了更好地发展,希望第二次受教育的权利又是否应该被剥夺呢?尽管复读生在某种程度上对应届生是不公平的,但若将复读班一刀切,对于复读生来说又是否公平?要回答上述重要的教育政策问题,一个基本的出发点是,理解是哪些人选择高考复读?高考复读是否能够提高个人能力或高考分数;如果能,是对所有人都有正面影响还是有高考收益的异质性?在高考复读如此普遍的情况下,既有文献里仍然缺少对高考复读学生群体的严谨实证研究。本文对高考复读的相关文献和政策讨论进行了极大补充。

本文使用宁夏回族自治区 2015 年和 2016 年的高考数据研究复读现象。宁夏位居我国西北,历史上为各族人民杂居之地,新中国建立后,宁夏是我国唯一的省级回族自治区。据 2015 年统计,宁夏全区人口 667.88 万人,其中城镇人口 368.90 万人,占总人口的 55.23%;回族人口 240.74 万人,占总人口的 36.05%,当年高考人数 64 378 人。

宁夏高考复读十分普遍,2014—2016 年,宁夏高考复读率连续三年超过 18%,且以农村学生为主。2014 年高考报名人数为 64 378 人,其中复读生 12020 人。宁夏回族自治区 2015 年高考报名人数为 67707 人,其中复读生 12415 人,2014 年应届生 11204 人,2013 年应届生 988 人。宁夏回族自治区 2016 年高考报名人数为 69 119 人,其中复读生 12 652 人,2015 年应届生 11 692 人,2014 年应届生 841 人,2013 年应届生 119 人。

我们主要使用定量方法分析来自宁夏的高考数据,并使用定性研究焦点访谈的方法,对两名有高三经验的教育工作者的访谈,试图对本文实证结论背后的原因进行探讨。我们首先使用 OLS 和 logit 模型,预测不同群体的高考复读决策,发现理科、男性、农村、高考总分较低的学生更倾向于选择复读;复读行为在民族和年龄上并没有太大差异。

其次,我们比较同一届应届参加高考的学生中,选择直接进入大学(或者进入劳动力市场)的学生和选择复读一年的学生,最终参与高考录取的成绩的差异;即2015届应届生中,直接参与录取学生的2015年高考成绩,和选择复读学生的2016年高考成绩的差异。为使得不同年份成绩可比,我们使用标准分(即等位分)。我们按照2015年应届生在当年的高考表现,分成上本科一本线、本科二本线、本科三本线及专科线四个群体,在控制了考生个人特征、户籍所在地区及高中毕业学校的固定效应后,第一次高考分数在一本线上的复读效果显著为负;其他几类学生均经历了高考成绩的增长。

然而,高考复读决策时内生的,OLS的估计值存在偏误。我们利用高考录取批次线的外生条件,使用断点回归方法,对高考复读的收益进行因果推断。我们发现,学生第一次高考成绩略低于二本线的,显著地更倾向于选择复读;然而,其他批次线附近的学生并没有显著不同。我们使用参数和非参数方法,发现二本线下的学生选择复读的,最终参与高考录取的成绩(即被大学录取的资格),得到了显著提高。我们的实证发现与来自教学一线的教师观察是一致的。

本文属于高考复读领域第一篇严格的因果推断论文,也是罕见的使用大规模行政数据进行定量研究的论文。我们的研究发现,丰富了中国高考复读现象的文献,为个人交易选择提供一定参考;更重要的,我们的研究可以参与政府制定关于高考复读更加有效、可行的政策的讨论;尤其是,我们发现高考复读确实对一部分学生群体带来显著的收益,这可能很好地解释为什么禁止高考复读的政策,最终的结果是“上有政策,下有对策”,屡禁不止。我们的政策启示是,如果在高考复读存在正向收益,那么我们的政策制定,不是要取消这种收益,而是应该指向复读的个人和公共成本,使得学生个人和公共政策实现更好的成本-收益权衡取舍。

二、相关研究

(一) 国内研究

国内关于高考复读现象的研究始于上世纪八十年代,但1986年—1998年之间,相关国内研究不过100余篇。而1998年我国高校扩张以后,高考复读生群体迅速增加,因此1999年开始,相关研究开始显著增加,尤其是2007年8月,时任教育部部长周济表示禁止公办普通高中开办复读班以来,相关评论和研究大量增加。

1. 对复读生群体的关注

对于高考复读生来说,他们比往届生承受了多一年高压力、大强度的“高三学习”,他们个体在这一年中的状态、生活质量成为了学者们关注的重点之一,吴颢通过对280名高考复读生的调查,认为高考复读生在生理领域上与大学生无明显差异,心理、社会、环境领域则优于大学生,具体来说,主动复读者优于被动复读学生,男生在很多方面略强于女生。^[1]

高考复读生的心理状态是学者研究的重要领域,包括自我效能、心理弹性等要素,以及考前心理状态等。程丽提到,高考复读生的心理弹性水平较之所有参加高考的学生来说,处于中等偏下水平,但善于学习的平均分是最高的。^[2]另有研究也发现,复读生相比应届生更需要社会支持,以增加其心理资本。^[3]

2. 对复读现象的探究

学生为什么要选择复读呢?高考复读在本质上是什么现象?有没有解决措施?对于“高考复读热”现象,较多学者对其中的原因进行了探究并提出了解决措施。

首先是客观上,也是最主要原因,即高等教育普及化以后,人们迫于找工作、出类拔萃的种种需求,不得不寻求更高的学历,相关的理论解释包括:社会不合理的人才消费观,^[4]和人力资本的信号理论,“文凭至上”观念不断在就业中体现,甚至将就业压力提前至高中时期。^[5]

教育不公平问题和高考制度的缺陷也是学生复读的重要原因。例如地区之间教育资源不均衡问题。有学者认为,在已经弱化“为地区培养人才”这一观念的今天,高考录取线不统一的政策凸显了教育不公平问题,除北京、上海等资源丰富之地外的学生纷纷为自己争取多一次的高考机会。^[6]高考录取制度的缺陷还体现在填报志愿的方式上,不过我国已经在近两年陆续在多省份实行了知分填报志愿制度和平行志愿制度,有效减少了因为志愿填报不当而导致的未被录取的考生数量。^[7]随着近年来新高考机制的改革,高考科目的设置更加多元化,高考成绩、平时因为制度缺陷导致的复读生数量会降低。学者们认为必须进一步完善高考制度,包括录取投档方式的转变,加强对志愿填报的指导,甚至增加多次考试录取的机会,完善自主招生制度,改变“一考定终身”的现状,^[8]也是解决高考复读热的重要举措。

从主观上来说,很多学生存在工具性复读或者习惯性复读现象,^[9]这与我国重视教育的传统,一味追求名校,鼓吹名校对于就业、生活,甚至是人生的改变,而忽视技术实用型的大专院校教育等观念也是息息相关的。^[10]除此之外,还可能存在随波逐流性复读,即受到外部环境的影响,从众跟风复读。或符号性复读,即为了追逐快乐感和愉悦感,向社会传递“自己能行”而复读的行为。^[11]因此,从舆论上加以引导,正确疏导社会各阶层的“名校”心理情结,使学生根据自身实际情况择校,也成为重要的举措之一。^[12]

3. 复读行为对相关主体的影响

高考复读现象的兴起有利有弊,其对于社会各个主体的影响是普遍而广泛的。高考复读本身是考生追求再教育权利的体现,然而病态的高考复读热,恐怕对于学生、学校、社会都不是一件好事。

对于中学来说,多收一个复读学生,通过种种学费、杂费、住宿费、课本费等复读费用,首先可以获得经济上的利益,这是利益驱动。同时,学校还可能因为招收优秀的复读生,提高升学率、重本率,为学校赚得宝贵的声誉,这是名

誉驱动。同时学校也有可能因此提高自身教学质量,这是复读的积极影响之一。^[13]

对于学生来说,一部分高考复读的学生可能多付出一年的学杂费用,就能够一举“飞上金枝头”,成为“人生赢家”。然而也有学生作出不明智的选择,浪费了时间与金钱,最后徒劳无功。还应该注意的,学生复读一年的时间成本,对于贫困家庭的学生来说,少读一年书,早读一年专科学校,就可以早一年赚钱。^[14]

对于高考录取来说,每一届考生都要与上一届考生同场竞争,在有限的招收名额里,一些成功的复读生挤占应届生的资源,又将成为应届生的榜样,不断延续,导致复读现象越来越普遍,^[15] 每一年都有大量复读生,即马太效应。

对于社会来说,一部分学生不复读就会走上工作的道路,因此复读在一定程度上可以舒缓社会矛盾,促进高考经济及相关行业的快速增长。同时,复读行为也间接地为国内高校输送了更多的优秀生源。^[16] 但是,高考复读这种“片面追求升学率”,“唯分数论”的情况,与国内正在大力推行的素质教育是背道而驰的,不少学者将这种现象表述为“教育改革上的拦路虎”,“置教育改革于不尴不尬之地”。^[17] 同时,也有学者指出,高考复读现象实际上是供求失衡引起的,尤其是一部分高分复读学生,其实质是对教育资源的浪费。^[18]

(二) 国外研究

与国内相关研究相比,国外对于高考复读的学术性研究明显较少,很大一部分原因是国外不存在中国式的高考,也不存在这样的主动复读,但是这并不意味着,国外就没有复读现象存在。

1. 国外复读现象简介

在国外也存在一种“复读现象”,即因为成绩太差导致的留级,这与中国的复读现象有一部分相通之处,但总体来说,他们的重修是为了迫使学生得到其必须的学术或社会技能,实现培养目标。这个制度的运用广泛程度在上世纪初开始波动上升,在1970年代达到顶峰后又逐渐下降,又在上世纪90年代开始被广泛使用。^[19]

该复读制度在北美、拉丁美洲、西欧、中亚,以及撒哈拉以南的非洲都被使用,并且学者发现,越是不发达的国家,该制度总是导致了更低的升学率和更高的辍学率。如北美、西欧和中亚地区的复读率约为1%—2%,升学率在86%以上,而拉丁美洲、北非等使用该复读政策的地区复读率高达6%—9%,升学率在62%—73%之间。^[20]

尽管如此,学者也发现美国的复读现象比较普遍并且正在增多,有数据显示,7%的美国学生会被迫重修数学课,而代数课的重修率更是高达15%,更加值得注意的是,美国很多州政府正在提高数学课在总成绩中的占比及数学成绩的合格要求,因而预期需要重修数学课的学生比例会继续增多。^[21]

同时学者还发现,黑人小孩的留级率是白人小孩的几乎两倍之多,并且这个数字还在增加。^[22] 另外有研究发现,能够顺利升学的高中生与留级的高中生

人数比例之差在不同人种之间也有区别:白人小孩相差 13%,黑人小孩相差 11%而拉丁裔相差 16%。^[23]

2. 对复读现象影响及其原因的研究

由于国外留级现象属于被动留级,而非我国高考复读主动留级现象,因此复读生在心理上所承受的压力和日后表现,受到了十分广泛的关注,并且以此向政府提出有关该政策的建议。然而这个话题一直充满争议,各个学者利用不同的样本也有不同的研究结论。但关于复读政策负面影响的研究大多围绕更低的学业成就、辍学、低升学率进行。

Jacob & Lefgren 的研究发现,低年级的学生留级对升高中影响不大,但八年级的学生留级则很可能降低学生的高中升学率。^[24] Hill 提到只有在重修生份额在 5—10% 的情况下,才会导致第一次修课的学生挂科率增加;并且认为这种挂科率的增加不完全是重修的学生带来的同伴效应。^[25]

Manacorda 使用 1996—1997 年 7—9 年级的学生数据,并一直跟踪到 2000—2001 年,最后发现,在 4—5 年以后,挂科都会导致持续的挂科和更低的学业成就。^[26]但也有学者指出,尽管复读政策具有激励效应,但低年级的复读行为并不会导致学习情况的提高,^[27]甚至可能因为社会压力及低自尊等心理情况而导致辍学率升高。^[28]

一些学者试图探究留级、学校背景与辍学之间的关系。有学者认为,在学校留级会使学生感到沮丧,因而降低自尊心和自信心,进而导致留级现象发生。^[29]另一些学者发现,如果学生在学习期间更多参与学校活动,则辍学的概率会显著降低。^[30]类似地,如果学生与老师、同学等与学校相关的关系越少,则越容易导致辍学。^[31]

但也有学者关注到了复读政策的积极影响。Manacorda 在研究中指出,尽管挂科可能给学生造成负面影响,但复读政策对学生努力学习的激励是可能是无法估计的,这也可能是该政策仍然被使用的最大原因。^[32] Jacob 也在研究中指出,复读政策对于学生来说,仿佛是学习懈怠的遏制剂,尽管这可能导致学习的周期更长。^[33]

值得注意的是,与国内研究不同的是,国外学者多采用因果推断策略,将着眼点放在了留级生自身学习效果、对非留级生学习情况的影响等,进行实证研究,而非对留级现象背后的制度探讨、客观分析。因此,对于美国留级现象的研究,尤其是研究方法,可以为我们提供参考。

三、理论框架

高考复读是典型的教育选择。人力资本理论是用以解释教育选择的最经典理论,其基本框架为,个人通过权衡一项教育决策(如,高考复读)的成本与收益的差别,进而做出最终的选择。通过高考复读,学生可以预期通过多一年的学习和复习,第二年的高考成绩得到提高,进而能够进入更好大学

(收益)。与此形成对比的是,高考复读所付出的成本包括一年时间的机会成本,复读产生的直接费用(如学费等)。如果考生预期收益大于成本,将会选择复读。

除了人力资本理论以外,筛选理论认为,由于就业市场存在不确定性和信息不对称,雇主无法深入识别求职者的能力,只有通过标签和信号进行筛选,因此出现过度教育、文凭主义等现状。由于高考复读很大程度上只是多花一年时间进行考试准备(test preparatin),并不是学习新的知识和技能。考生在做出复读决策时,更看重的是通过复读,获取更高的分数,进入更好的大学,获得更高的文凭。

上述两种理论都是建立在个人理性选择及信息充分的基础上。然而,在现实中,考生的高考复读决策并不完全是理性的。一方面,他们可能并不清楚高考复读的收益,更可能的是高估复读的收益,因而更倾向于选择复读;另一方面,认知偏差,诸如从众心理、小数法则等,使得考生对自我定位及复读效果的错误认识。本文的一个重要目的,即是对高考复读的效果进行科学评估,为考生的决策提供参考。

四、数据和样本

我们主要考察 2015 年参加高考的宁夏学生。针对复读生,我们使用学生的身份证号(在保密机房里完成,保证数据不泄露)匹配至 2016 年的高考信息。我们将参加 2015 年高考的应届生分为两个群体,群体一为 2015 年高考后未选择复读的 2015 年应届生:即除去 2014 年(或更早)应届、2015 年复读的学生,及 2015 年应届、2016 年复读的学生,或即只参加过 2015 年一届高考的学生;群体二为 2015 年应届,并选择在 2016 年复读的学生,或即至少同时参加过 2015、2016 年高考的学生。

2015 年共有 63 207 名参加高考的学生,除去 11 881 名在 2015 年就非应届生身份的学生,剩余 2015 年应届生 51 326 名,其中群体一包含学生 40 144 名,占 78.21%;群体二包含学生 12 037 名,占 21.79%。文科应届生 19 046 名,占 37.11%;理科应届生 32 280 名,占 62.89%。男生 23 920 名,占 46.60%;女生 27 406 名,占 53.40%。城镇学生 23 113 名,占 45.03%;农村学生 28 213 名,占 54.97%。汉族学生 35 565 名,占 69.29%;回族学生 15 184 名,占 29.58%。平均年龄 20.00 岁。

未选择复读的应届生共有文科生 15 513 名,理科生 24 631 名。男生 18 763 名,女生 21 381 名。城镇学生 19 384 名,占 48.29%,农村学生 20 760 名,占 51.71%。汉族学生 27 423 名,占 68.31%,回族学生 12 205 名,占 30.40%。平均年龄 19.99 岁。

选择复读的文科应届生 3 533 名,复读的理科应届生 7 649 名,文科复读率 18.55%,理科复读率 23.70%。复读的男生 5 157 名,女生 6 025 名,男生复读率

21.56% ,女生复读率 21.98%。城市学生 3 729 名 ,农村学生 7 453 名 ,城市学生复读率 16.13% ,农村学生复读率 26.42%。汉族学生 8 142 名 ,回族学生 2 979 名 ,汉族学生复读率 22.89% ,回族学生复读率 19.62%。复读生平均年龄 20.05 岁。

表 1 中比较了 15 年高考后升学或工作的应届生群体和 15 年高考后选择复读的应届生群体。理科、男性、农村、高考总分较低的学生更倾向于选择复读;复读行为在民族和年龄上并没有太大差异。

表 1 2015 年选择复读与未复读应届生的个人信息等差异

		未复读群体	复读群体
文科学生	数 量	15 513	3 533
	比 例	38.64%	28.71%
男性学生	数 量	18 763	5 157
	比 例	36.56%	42.84%
城镇学生	数 量	19 384	3 729
	比 例	48.29%	30.98%
汉族学生	数 量	27 423	8 142
	比 例	68.31%	67.64%
平均年龄		19.99	20.05
2015 年高考总分		372.06	336.90

由于文理科不可比 ,需要单独分析 ,为了节省篇幅 ,本文的分析集中于理科学生 ,文科学生的情况类似。表 2 比较了理科复读生在 15 年(应届)和 16 年(复读)的成绩差异。各科成绩及总成绩均已转化为均值为 0、标准差为 1 的标准分。结果显示 ,理科复读生每一科均有显著的分数提高 ,尤其是综合科目和数学科目 ,经过复读一年的大量练习 ,成绩得到较大提高。然而 ,这只是描述统计结果 ,成绩的差异不一定完全是由于复读带来的 ,还应当考虑其他影响因素的作用 ,需要进一步的深入分析。

表 2 理科复读生 2015、2016 年各科高考成绩表

	15 年均值	16 年均值	均值差值	相关系数
语文	-0.140	-0.064	0.077	0.359
数学	-0.259	0.011	0.364	0.721
英语	-0.270	0.001	0.271	0.767
综合	-0.341	0.010	0.436	0.739
总分	-0.316	0.060	0.377	0.775

五、实证研究结果

(一) 高考复读决策的影响因素

我们使用回归分析描述高考复读决策的影响因素,基本的回归模型如下:

$$Y(\text{Repeated} = 1)_{ij} = \beta_0 + \beta_1 X_{ij} + \delta_j + \varepsilon_{ij} \quad (1)$$

其中,因变量 $Y(\text{Repeated} = 1)_{ij}$ 是一个两分类的结果变量,选择复读取值为 1,不选择复读取值为 0; X_{ij} 包括个人特征、家庭特征、行为经济学特征等; δ_j 是毕业高中固定效应,用以控制地区 and 高中差异; ε_{ij} 是不包含在模型当中的随机扰动项。我们将标准误差聚集在学校层面。

在估计模型的选择上,本文将分别进行比较线性模型(OLS)和非线性模型(Logit model)并比较其差别。这是因为针对二元分类的结果变量,两类模型在中间人群的估计结果比较接近,非线性模型在极值拟合更好。由于事先并不知道复读学生的分布情况,使用两类模型的比较将有助于得到更稳健的结果。

在该模型中,因变量 repeat 为是否复读这一两分类的结果变量,选择复读赋值为 1,选择不复读赋值为 0;自变量依次放入性别变量,为两分类变量,其中男生赋值为 0,女生赋值为 1;文理科变量,为两分类变量,其中文科生赋值为 0,理科生赋值为 1;民族代码变量,为三分类变量,分为汉族、回族和其他少数民族;城乡代码变量,为两分类变量,其中城镇学生赋值为 0,农村学生赋值为 1;年龄变量为连续变量,为截至 2016 年 12 月 31 日,该考生的年龄,取整数;2015 年高考总分变量,为连续变量,为该生在 2015 年的高考总成绩。

表 3 汇报了 Logit 模型的边际效应和 OLS 模型的估计值,两者非常一致。以 Logit 模型为例,回归结果显示,在其他条件不变的情况下,女生比男生复读的概率高 2.39 个百分点;理科生比文科生复读概率高 4.52 个百分点;回族学生复读概率比汉族学生要低 4.78 个百分点;农村学生复读概率比城市学生要高 10.08 个百分点;学生年龄每增加一岁,复读概率就降低 1.22 个百分点;学生高考总分每增加一分,复读概率就降低 0.04 个百分点。并且,这些回归结果在 0.01 的水平上显著。

表 3 2015 年应届生样本 Logit 和 OLS 模型比较结果

	Logit (边际效应)	OLS
女 性	0.023 9*** (0.003 7)	0.024 0*** (0.003 8)
理 科	0.045 2*** (0.003 8)	0.045 8*** (0.003 8)
回 族	-0.047 8*** (0.003 8)	-0.049 5*** (0.003 9)

续 表

	Logit (边际效应)	OLS
农 村	0.100 8 *** (0.038 1)	0.101 6 *** (0.003 9)
年 龄	-0.012 2 *** (0.001 7)	-0.012 4 *** (0.001 8)
高考总分	-0.000 4 *** (0.000 2)	-0.000 4 *** (0.000 0)
样本总量	51 326	51 326

注: *** 表示在 1% 的水平上显著, ** 表示在 5% 的水平上显著, * 表示在 10% 的水平上显著。

(二) 高考复读的收益: OLS 估计

进一步看,本文估计高考复读对高考成绩和录取结果的影响。如前所述,我们仍然应当比较同一年应届生的不同复读决策。本文使用 2015 年所有理科学生的样本,由于文理科的考试科目及分数线不同,本文仅比较应届理科学生中决定复读的学生在 2016 年的成绩,及应届生中选择升学或工作的学生 2015 年的成绩。因此,结果变量是非复读生应届高考成绩和复读生下一年的高考成绩;该结果变量用以衡量学生最后一次参与大学录取的资格。为保证结果的可比性,本文使用标准化后的高考成绩(即考生的相对排位或者等值分)。基本回归模型如下,我们控制学校的固定效应。

$$Y_{ij} = \beta_0 + \beta_1 \text{Repeated}_{ij} + \beta_2 X_{ij} + \delta_j + \varepsilon_{ij} \quad (2)$$

在该方程中,因变量 Y 是未选择复读的学生 2015 年的成绩标准分及选择复读的学生 2016 年的成绩标准分,下文称为高考收益。 i 代表学生个体, j 代表学校。因此, Y_{ij} 是学校 j 的学生 i 的高考成绩。在该模型中,控制了个人和家庭特征以及毕业学校固定效应之后, β_1 即是复读的收益的 OLS 估计值。

表 4 2015 年应届理科生高考复读收益回归结果

变 量	模型一	模型二	模型三	模型四
复读决策	0.052 (0.049)	0.120 ** (0.053)	0.143 *** (0.054)	0.326 *** (0.039)
女 性		0.159 *** (0.017)	0.163 *** (0.017)	0.127 *** (0.016)
汉 族		-0.172 * (0.089)	-0.140 * (0.073)	-0.002 (0.056)
回 族		0.007 (0.106)	0.065 (0.080)	0.185 *** (0.057)
农 村		-0.179 *** (0.069)	-0.113 *** (0.032)	0.069 *** (0.023)

续 表

变 量	模型一	模型二	模型三	模型四
年 龄		-0.236 *** (0.012)	-0.233 *** (0.011)	-0.152 *** (0.011)
R-Squared	0.005	0.086	0.102	0.439
样本量	29 113	29 113	29 113	29 113

注：*** 表示在 1%的水平上显著，** 表示在 5%的水平上显著，* 表示在 10%的水平上显著。

表 4 的结果显示,在不控制其他因素的情况下,复读学生的最终参与录取的高考成绩(复读后)比未复读学生的最终参与录取的高考成绩(应届)高 0.052 个标准差,但统计上并不显著;显然这个估计值存在遗漏变量带来的偏差。第 2 列控制了学生的个人特征,复读学生的高考成绩高 0.12 个标准差(在 0.05 的水平上显著)。第 3 列控制了考生户口所在地区的固定效应,在其他条件不变的情况下,复读能使学生高考收益提高 0.143 个标准差(在 0.01 的水平上显著)。第 4 列进一步控制了考生高中毕业学校的固定效应,在其他条件不变的情况下,复读能使学生高考收益提高 0.326 个标准差(在 0.01 的水平上显著)。

在表 5 中,我们按照 2015 年应届生在当年的高考表现,分成上本科一本线,本科二本线,本科三本线,及专科线四个群体,在控制了考生个人特征、户籍所在地区及高中毕业学校的固定效应,考察高考复读收益的异质性。结果显示,第一次高考分数在一本线上的复读效果显著为负;其他几类学生均经历了高考成绩的增长。

表 5 各批次线上理科生高考复读收益回归结果比较

变 量	一 本	二 本	三 本	专 科
复读决策	-0.793 *** (0.036)	0.500 *** (0.011)	0.778 *** (0.010)	0.765 *** (0.039)
女 性	0.015 (0.015)	-0.002 *** (0.004)	0.040 *** (0.008)	0.103 *** (0.012)
汉 族	-0.126 ** (0.058)	-0.056 * (0.024)	-0.015 (0.036)	0.073 (0.055)
回 族	-0.090 (0.061)	-0.063 ** (0.025)	0.051 (0.037)	0.170 *** (0.057)
农 村	0.053 *** (0.019)	0.005 (0.005)	-0.005 (0.011)	0.045 ** (0.019)
年 龄	-0.085 *** (0.011)	-0.005 *** (0.002)	-0.021 *** (0.004)	-0.018 *** (0.006)
R-Squared	0.556	0.590	0.566	0.355
样本量	12 667	2 856	7 530	6 063

注：*** 表示在 1%的水平上显著，** 表示在 5%的水平上显著，* 表示在 10%的水平上显著。

(三) 高考复读的收益: 断点回归设计

表 4 和表 5 使用 OLS 估计高考复读的效果。然而复读决策很有可能是内生的,即选择复读和选择不复读的学生可能存在不可观测的、系统性的差异。回归模型(2)中的 ε_{ij} 中包含不可观测的因素,可能同时影响了复读和高考成绩,进而导致 β_1 的估计是有偏的。因此,我们进一步采用断点回归设计来对复读收益进行因果推断。

我们利用高考录取批次线的外生条件。举例而言,两个高考总分相差 1 分的学生,平均意义上应该是相似的;其中高 1 分的学生刚好上了一本录取线,而另一位学生没有上一本线因此选择复读。在这种情况下,两个相似学生是否复读的决策只受到了一本录取线的影响,与其他内生的、不可观测的变量无关,我们因而能够估计出高考复读对高考成绩因果效应。

我们以理科一本线 445,二本线 416,三本线 316,专科线 200 分作为断点,以 200 分作为观察窗口,观察断点上下考生复读率的变化。从图 1 可以明显看出,一本线、三本线及专科线附近的复读率都比较连续,只有二本线 416 分附近存在明显断点,为了更好地观察,图 2 将窗口设为 50 分。如图所示,在二本线下五分的复读率约为 30%,而复读线上五分的复读率仅为 10%。而二本线下 50 分(三本线附近)的学生,复读率超过 50%。

图 3 和图 4 展示了,在二本线 416 分附近,考生不存在年龄、城乡分布的不连续。我们还检验了其他学生层面的变量,在断点附近都是连续的。因此,如果最终成绩在二本线上下出现差异,只可能是因为二本线导致学生的复读差异带来的。

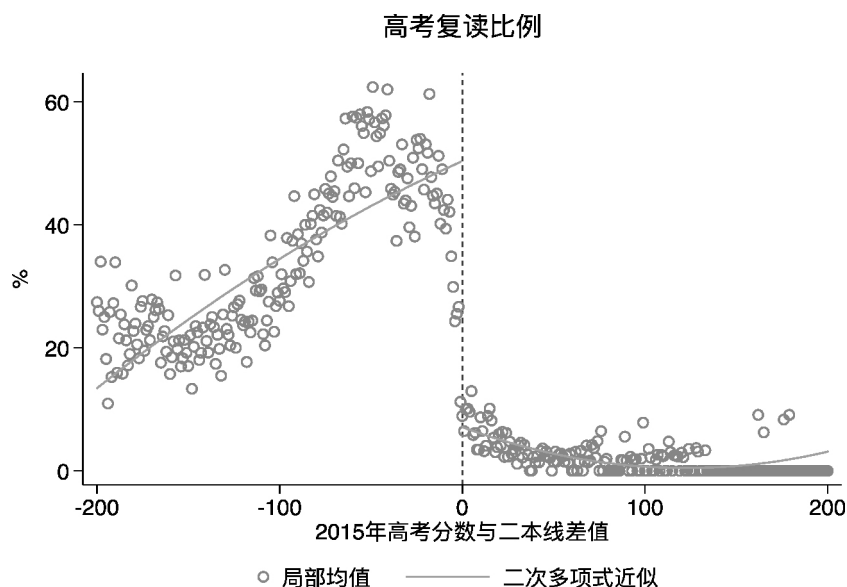


图 1 二本线附近理科生复读率

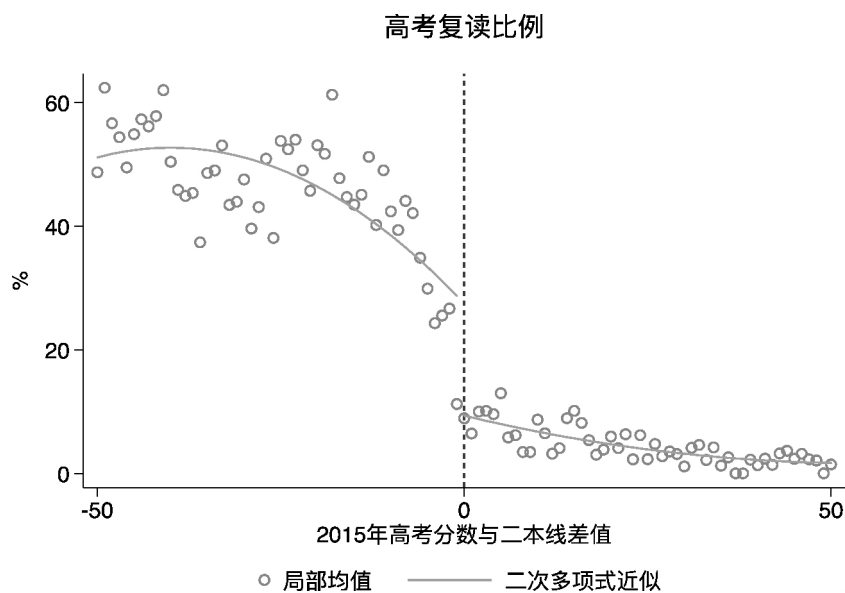


图2 二本线附近理科生复读率(小窗口)

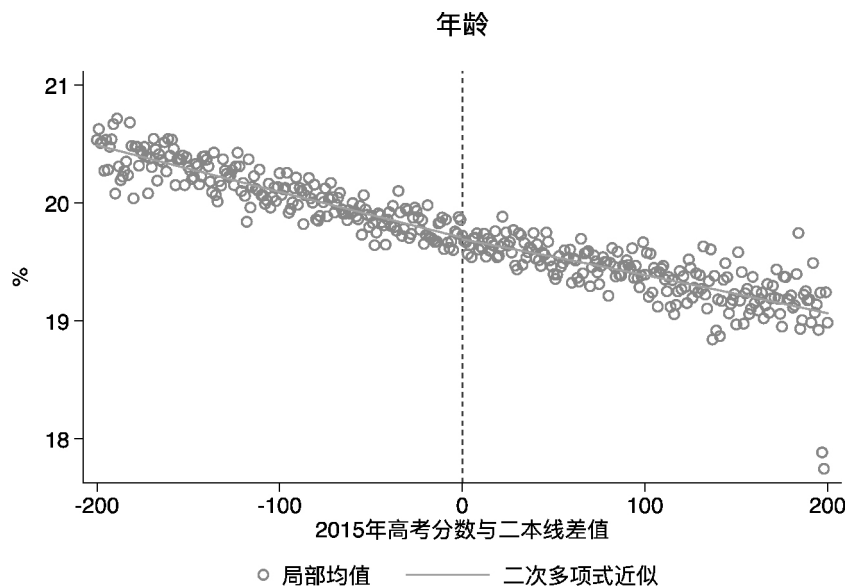


图3 二本线附近理科生年龄分布

图5展示了学生高考分数在二本线附近的变化。考生的最终高考分数衡量方式为:2015年末选择复读的应届生,其收益为他在2015年高考成绩的标准分;2015年选择复读的应届生,其收益为他在2016年高考成绩的标准分。这个分数衡量的是学生最终参与大学录取的基础。总体而言,学生的最终高考分数与2015年的初始高考分数正向相关(对于2015年已经升学的学生而言,这个相关系数是1)。在二本线附近出现了下降,由于二本线左侧的学生选择复读的比例(平均30%)高于右侧的学生(平均10%),这个成绩的变化可以

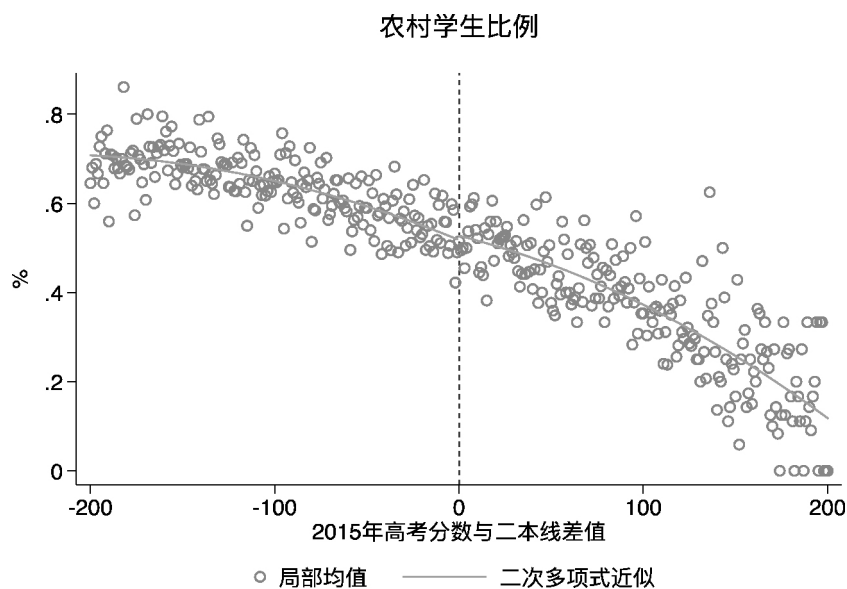


图 4 二本线附近理科生城乡分布情况

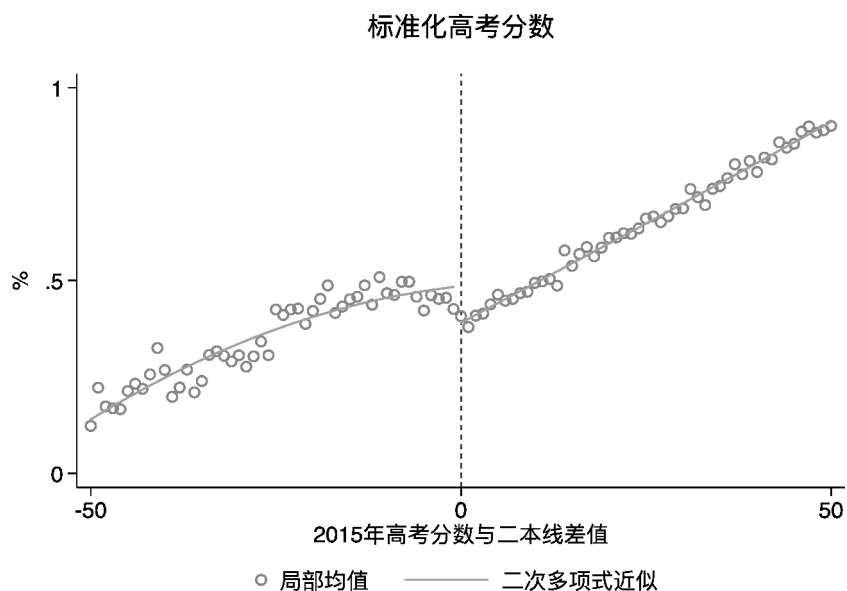


图 5 二本线附近理科生高考收益图(小窗口)

认为是复读带来的。我们同时检验了一本线、三本线及专科线附近的成绩变化,没有发现成绩的变化。

我们使用参数估计和非参数估计两种方法来准确地估计高考复读的收益。参数估计的方法不受限于样本量,但是对于较大的窗口或者高次项的估计较敏感且不精准;非参数估计的方法更加精准,但是对样本量有一定要求。我们使用不同的窗口大小以及一次项和二次项的方程形式。其中,第一阶段回归以考生复读率为结果变量,第二阶段回归以考生的高考收益为结果变量,

复读考生的高考收益为其 2016 年高考标准分,未复读考生的高考收益为其 2015 年高考标准分。

表 6 报告的估计结果。与图 5 一致,高考没有上二本线的考生,比考上二本线的考生,复读的概率高 10—20 个百分点。对这一部分学生,高考选择复读,带来 0.1—0.4 个标准差的成绩提升(参数估计大窗口的结果不准确;除此以外,估计结果非常一致)。

表 6 二本线附近理科生高考复读率及收益的断点回归结果

窗 口	5 分		10 分		20 分	30 分
	1	2	1	2	1	1
A. 参数估计						
第一阶段回归: 结果=复读率	-0.124***	-0.211***	-0.096***	-0.116***	-0.177***	0.255***
	(0.029)	(0.014)	(0.023)	(0.041)	(0.026)	(0.031)
第二阶段回归: 结果=高考收益	0.186***	0.189***	0.120***	0.118***	-0.106**	-0.351***
	(0.397)	(0.396)	(0.035)	(0.036)	(0.042)	(0.056)
观测值	1 049	1 049	2 018	2 018	4 097	6 101
B. 非参数估计 (rdrobust 命令)						
第一阶段回归: 结果=复读率	—	—	-0.164*	-0.19	-0.107**	-0.104***
			[-0.426, 0.031]	[-0.620, 0.233]	[-0.205, 0.010]	[0.175, 0.002]
第二阶段回归: 结果=高考收益	—	—	0.174	0.046	0.374**	0.401***
			[-0.539, 0.677]	[-1.09, 1.21]	[-0.011, 0.705]	[0.060, 0.675]
观测值	1 049	1 049	2 018	2 018	4 081	6 079

六、讨论和小结

对于什么样的学生会选择复读这一问题,2015 年所有高考应届生样本显示,女生、理科生、汉族学生、农村学生、年龄小的学生、高考成绩偏低的学生更可能选择复读。其中,根据断点回归结果,二本线下 10 分以内的学生复读率明显偏高,且有显著的正向收益。

为了更好地理解数据分析的结果,我们向四川某重点高中高三教师 A 及北京大学研究生 B 进行了访谈,前者有逾 20 年从教经验,教过历史、语文等科目,熟谙各个学科特点及学生特点;后者本科毕业后,于宁夏回族自治区、青海省支教一年有余,主要教语文科目,了解当地情况和学生及其家长想法。

首先,从复读学生群体的特征来看,两位老师认为学生及家长对教育改变

人生命运的期待、高考分数的差异(如农村学生显著低于城市学生,少数民族学生显著低于汉族学生)、以及高考志愿填报的信息、辅导差异是影响不同群体复读率差异的重要因素。根据我们的2016年问卷数据,49%的回族复读生是在填报志愿、获知录取结果后选择复读的(其余是在高考结束后或获知高考成绩后);汉族复读生约有28%是由于录取结果不理想选择复读。

总体来说,分数较低的学生更可能选择复读;就具体的分数段来看,无论是文科还是理科,二本线下及二本线上的学生存在复读率的跳跃。为此,A老师认为是学生及家长普遍认为上二本才算是上了大学的观念所致。B老师则提到,一般二本院校和一本院校学费较低,三本院校则学费显著较高,家长和学生很可能用一年时间复读,希望能考上二本,以规避高额学费。

其次,我们的实证结果显示,理科一本学生复读的收益为负(表5第1列)。当然,由于理科一本线附近并不存在断点,OLS估计值可能存在偏差。与此相比,应届成绩在三本以上、二本以下的学生复读收益最大(断点回归估计值)。然而,专科线上线下的考生往往基础过于薄弱,复读时也并不努力,所以成绩提高幅度反而比较小。

最后,在高考复读的作用机制上,两位老师提到了年龄以及自我感知的压力与成绩的关系。年龄越大的复读生,其复读效果相对较差。这与年纪小的学生心理素质好,记忆力好,学习能力强相关。在我们的数据中,有一小部分学生复读次数超过三次,其潜在的复读收益已经非常小了。再则,我们发现,自我压力大的学生,更有可能选择复读,但是实际复读效果并不好。因此,自我压力会给学生复读动力,但反而会成为学生提高学习成绩的障碍,复读生要合理对待自我压力,在复读过程中合理释放自身压力。

本文基于宁夏的高考数据,实证研究了高考复读的收益及其作用机制;高考录取批次线为断点回归的估计提供了可能,研究的结果与来自一线老师的判断基本一致。研究结果表明,高考复读的倾向、收益以及作用机制都存在显著的异质性。我们最重要的建议是:没有一个普适性的“高考复读有用”的答案;一旦基于自身偏好选择高考复读,投入和面对压力将是影响复读效果的重要因素。

限于学识和能力,本文还有诸多不足之处,留待后续更多学者的研究予以解答。首先,本文的假定都是基于考生能够在两次高考中正常发挥,忽略了偶然因素对于高考成绩的影响。因此,如果能够取得该生在模拟考试中的成绩,对于本文研究会很有助益。其次,还有一部分学生选择多次复读,对这一类学生的研究也非常重要。我们的研究结论基于宁夏的数据,对类似的低收入、民族地区应当具有相当的参考作用,然而对于其他地区(如发达地区)的复读现状与效果是否具有外部有效性,还需要更多来自其他地区的实证研究予以补充。

最后,在研究高考复读收益机制时,出于使用数据的局限,我们并不能对影响高考复读效果的因素进行更深入的探究,这需要通过更丰富的行政数据

和问卷调查数据予以补充;我们无法将自我的努力纳入自变量,因为存在潜在的内生性(后续的研究可以考虑随机试验设计);学生在复读一般会选择更好的学校,存在自选择偏差也无法观察到学校的增值效果。这些不足之处,需要更多的学者,收集更多数据,基于更丰富的理论和实证策略进行深入研究。总之,高考复读的个人决策和公共政策,都亟待更多更好的研究予以参考。

参考文献:

- [1] 吴颢.高考复读生存质量及其影响因素研究[J].中国健康心理学杂志,2010(10):1211-1213.
- [2] 程丽.高考复读心理弹性及其与学习动机的关系[D].开封:河南大学,2009.
- [3] 周哲.高考复读心理资本、社会支持与应对方式的关系研究[D].桂林:广西师范大学,2015.
- [4][10][16] 方光宝.高考复读热的原因及影响分析[J].科技信息,2009(13):28-29.
- [5] 杨高安.抚州市普通高考复读生复读行为调查研究[D].南昌:江西师范大学,2005.
- [6] 张克新,朱成科.关于我国高考复读现象之“繁荣”的思考[J].当代教育论坛:学科教育研究,2007(6):30-31.
- [7] 符太胜.从高考复读生群体的主体视角透视高考改革——问题及方向[J].现代教育管理,2010(6):22-25.
- [8] 潘俊峰.关于高考复读生增多的思考及对策[J].黑龙江教育学院学报,2004(5):18-19.
- [9][11] 陈益林,韦宁.复读现象剖析及对策探讨[J].继续教育研究,2008(8):136-138.
- [12] 邱纪香,康玉唐.高考复读现象解析[J].中国电力教育,2012(4):139-140.
- [13][15] 田虎.我国高考复读现象透视[J].教育发展研究,2007(11a):29-33.
- [14] 许祥云.家庭高等教育投资行为实证研究[D].武汉:武汉理工大学,2010.
- [17] 李莉莉.高考复读班透视[J].科学与文化,2007(10):4-6.
- [18] 刘福才.供求失衡:高考复读政策的实施困境[J].中国教师,2014(1):36-41+46.
- [19] Bali, V. A., Anagnostopoulos, D. & Roberts, R. Toward a Political Explanation of Grade Retention. [J] Educational Evaluation and Policy Analysis, 2005(2): 133-155.
- [20] Urquiola, M. & Calderon, V. Apples and Oranges: Educational Enrolment and Attainment across Countries in Latin America and the Caribbean [J]. International Journal of Educational Development, 2006(6): 572-590.
- [21] Reys, B. J., Dingman, S., Nevels, N. & Teuscher, D. High School Mathematics: State-Level Curriculum Standards and Graduation Requirements [J]. Center for the Study of Mathematics Curriculum, 2007: 13.
- [22] Moller, S., Stearns, E., Blau, J. R. & Kenneth, C. Land. Smooth and Rough Roads to Academic Achievement: Retention and Race/Class Disparities in High School [J]. Social Science Research, 2006(1): 157-180.
- [23] Stearns, E., Moller, S., Blau, J. & Potochnick, S. Staying Back and Dropping Out: the Relationship Between Grade Retention and School Dropout [J]. Sociology of Education, 2007(3): 210-240.
- [24] Jacob, B. A. & Lefgren, L. The Effect of Grade Retention on High School Completion. [J]. American Economic Association, 2009(3): 33-58.
- [25] Hill, A. J. Economics of Education Review [J]. Economics of Education Review, 2014(43): 91-105.
- [26][32] Manacorda, M. The Cost of Grade Retention [J]. Review of Economic and Statistics, 2012(2): 596-606.
- [27] McCoy, A. R. & Reynolds, A. J. Grade Retention and School Performance: An Extended Investigation

- [J]. *Journal of School Psychology* , 1999(3) : 273-298.
- [28] Jimerson , S. R. , Anderson , C. E. & Whipple , A. D. Winning the Battle and Losing the War: Examining the Relation between Grade Retention and Dropping Out of High School [J]. *Psychology in the Schools* , 2002(4) : 441-457.
- [29] Finn , J. D. Withdrawing from School [J]. *Review of Educational Research* , 1989(2) : 117-142.
- [30] Gottfredson , Denise C. , Carolyn M. Fink & Nanette Graham. Grade Retention and Problem Behavior. [J] *American Educational Research Journal* , 1994(4) : 761-784.
- [31] Jacob , B. Accountability , Incentives and Behavior: Evidence from School Reform in Chicago [J]. *Journal of public Economics* , 2005(5-6) : 761-796.
- [33] OECD. Competitiveness at School May Not Yield the Host Exam Results [EB/OL]. <http://www.economist.com/graphic-datial/2017/04/20/competitiveness-at-school-may-not-yield-the-host-exam-results>. 2017-4-20/2017-4-20.

Retaking College Entrance Exam: Evidence from a Regression Discontinuity Design

ZHANG Yan , DING Yanqing , LI Honghui ,

LI Wei , MA Yujie & YE Xiaoyang

(Ningxia Education Examination Authority , Yinchuan , 750002 , China ;

Graduate School of Education , Peking University , Beijing , 100871 , China ;

School of Political Science and Law , Ningxia University , Yinchuan , 750021 , China ;

School of Education , University of Michigan , State of Michigan , 48109 , United States)

Abstract: Retaking the Chinese College Entrance Exam (CEE) is prevalent in China that every year around 10 to 20 percent of high school graduates choose to retake CEE in the next year. However , there is a dearth of rigorous empirical work in the literature. This paper uses administrative data on students who took CEE in Ningxia and find that students in the STEM track , who are male , from rural families , and have lower CEE scores are more likely to retake CEE. Retaking CEE is heterogeneous that higher achieving students are less likely to benefit from it. Taking advantage of the exogenous admission tier cutoffs , we use Regression Discontinuity Design to estimate the causal effects of retaking CEE on eventual admission scores. Students with the first-time CEE score right below the second-tier cutoff are much more likely to retake CEE in the next year , and they largely benefit from retaking the CEE. Our empirical findings are consistent with the observations of high school teachers. This paper contributes to the related literature , individual decision , and education policy-making with strong empirical evidence.

Key words: Chinese College Entrance Exam (CEE) ; retaking CEE ; returns to education ; regression discontinuity

(责任校对: 李媛媛)