

高等教育使谁获益更多？

——2003 - 2015 年中国高等教育异质性回报模式演变

郭 冉 周 皓

提要: 本文主要讨论不同群体获得高等教育回报的模式演变。在积极选择假设和消极选择假设两种理论框架的基础上,结合中国现实,本文提出了“倒U形”回报模式的设想。分析结果表明,从2003年到2015年,我国高等教育回报模式从“波动—中底层激励型”向“稳定—中层激励型”转变,为“倒U形选择假设”提供了实证支持。这种教育回报模式的演变过程说明社会中间阶层已成为高等教育最主要的受益人,表明中国社会变迁过程中制度性结构的变化有益于巩固新兴中产阶级并维持社会结构稳定。

关键词: 高等教育回报 选择性 异质性 反事实分析框架 回报模式

一、问题的提出

教育回报研究是探究人力资本形成、评估相关政策乃至衡量收入不平等、测度阶层变迁的重要途径。改革开放四十年以来,我国经济以9% - 10%的年均速度持续高速增长。恢复高考以来高等教育的不断发展,市场改革的持续深化,人力资本投资的日益增加,都为我国经济发展储备了充足的人力资源。这不仅促进了宏观经济的持续增长,提高了个人在劳动力市场的竞争力,也使得人力资本回报不断提升。

教育的社会功能与阶层结构变迁/维系有着千丝万缕的联系,教育在社会流动中的作用也与社会结构的开放性密切关联(方长春、风笑天,2018)。关注高等教育回报问题,有助于了解不同人群中教育回报的变动趋势和教育回报模式的差异,把握它同社会分层之间的关系;同时也可以反映劳动力配置水平,评估相关教育政策,进而为教育体制改革和宏观经济政策制定、调整提供有力依据。

因此,教育回报研究是非常复杂的问题域。首先,其复杂性体现为教育收益在不同人群、阶层之间的分配差异以及相应的社会分层效应。其研究范围不仅包括家庭背景、城乡属性、不同教育层次(纵向分流)

等的影响,还包括高等教育扩招、重点学校招生(横向分流)等政策性“干预”对教育回报的影响。第二,教育回报的研究范式和方法多种多样,不同学科对于方法的选择也各异其趣,因此有必要针对不同估计方法之间的优势和局限进行讨论。第三,教育回报研究还包括针对高等教育的回报模式的探讨,但总体而言,既往研究对连续的、长时段变化的关注稍显不足。因此,本文需要解决的问题就是:进入21世纪之后的十几年间,伴随着教育制度的巨变和市场机制的长足发展,高等教育回报模式如何发生变化,这种动态变化又使得哪个阶层获得的回报更多,进而影响社会分层?

二、文献综述

(一)高等教育与社会分层

高等教育是影响社会流动和社会分层的重要机制。它不仅是个人的知识学习和社会文化的教化过程,也是确保每个人平等地接受教育、享受教育服务并实现有序的社会流动的过程(唐俊超,2015)。更重要的是,教育是重要的阶层再生产途径,它通过调节收入、职业流动来完成这个过程(李培林、朱迪,2015)。因此,教育平等一直以来都是各个国家和社会努力达到的目标。然而教育不平等却始终存在(Wu,2010;刘精明,2008),它存在于社会内部(城乡、民族、性别、地区、贫富)不同的群体之间,也存在于国家之间,还会随着时期、政策、经济形势等变动加剧/减轻(叶华、吴晓刚,2011;张兆曙、陈奇,2013)。其作用机制也在“社会/文化—经济/竞争”“先赋—自致”“市场化—再分配”等理论的演绎下变得更为复杂。

布迪厄对文化资本和教育场域的研究清晰地揭示了高等教育获得对阶级生产及维持的作用。高等教育不仅是一种获得更高阶层地位的途径,对于巩固阶层秩序、文化和符号再生产也起到了至关重要的作用。资本和权力更为集中的名牌大学在这种制度化的仪式中所起的作用更甚,并通过权力与社会结构的动态运作实现阶层和不平等的再生产(布尔迪厄,2004)。人力资本理论则更为直接地强调了经济因素对于获得高等教育的作用,同时强调了高等教育的获得可以进一步增强竞争能力,收获更高的经济回报。对教育的直接投资可以产生“投资

→回报→更多投资→更多回报”的良性循环(Becker,1964)。

在有关“市场转型”的诸多理论与争论中,教育毫无疑问处于中心位置。倪志伟的“市场机会假说”认为,从再分配制度转向市场经济的过程中,市场部门会分享更多的权力,人力资本也会更加受市场青睐,获得较高回报(Nee,1989)。魏昂德(Andrew Walder)也在论著中提到,中国城市中存在二元的职业路径,但大学学历无论对于高级专业技术地位还是高级行政管理地位的获得都有举足轻重的直接影响(Walder et al.,2000)。尽管关于“市场转型”的讨论存在诸多分歧,各方仍然在教育的作用上达成了共识:随着市场转型的推进,教育的价值会有更高的体现,人力资本的重要性将得到大幅度提升;接受更好的教育,尤其是高等教育,会带来更多经济回报或者获得更高的政治地位;受过高等教育的群体更易于向上流动到较高的职业地位和社会阶层(Nee,1989;郑辉、李路路,2009)。

先赋和自致因素则是另一个主要的研究视角。先赋性的家庭背景、城乡差异、性别等因素以及个体能力、素质等自致性因素都是影响教育机会及结果分配的重要变量(Blau & Duncan,1968; Lucas,2001)。其中,相对风险回避假设(Relative Risk Aversion, RRA)较多强调先赋因素:社会经济地位较高的家庭会利用手中的各种资源对子代可能的地位下降风险进行规避,最终这种不平等将影响到阶层的流动(侯利明,2015)。通常情况下,处于相对劣势的个体/群体在结果分配过程中会强化其不利地位,形成固化的状态。而“工业化假设”则强调随着工业化/现代化的推进,先赋因素影响逐渐降低,自致因素逐渐发挥主导作用,教育不平等程度也逐渐下降(李春玲,2014)。

除此之外,对某些政策因素的评估也是教育不平等研究的重要议题,最典型的则是教育扩招这一政策。几个主要的理论模型,如“最大维持不平等假说”(MMI)、“有效维持不平等假说”(EMI)等也都针对资源的稀释与分配给出了不同解释。已有研究发现,教育扩招之后,家庭背景、区域差异等影响被放大,高等教育机会的分配变得更加不平等(刘精明,2014;叶晓阳、丁延庆,2015)。

关于教育不平等还有非常丰富的理论视角和研究取向,篇幅所限,在此不能一一列举。但总体而言,高等教育与社会分层是一种正向相关的关系,教育成就和阶层地位具有很强的相关性,且二者通过一定的中介机制相连接。本文暂且搁置“谁先谁后”这种因果机制的探讨和

判断,只在宏观上呈现既有研究对这对关系的理论探讨。图1归纳了几种类型的研究取向。

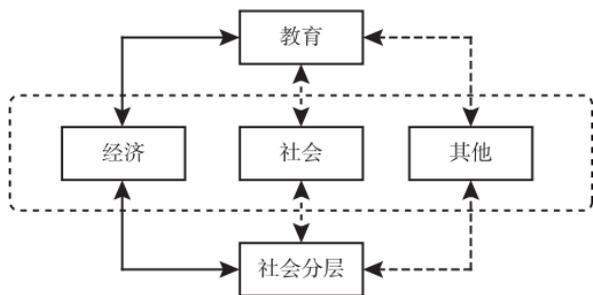


图1 教育与社会分层之间的关系

根据分析策略和研究设计,本文倾向于采取“社会/文化—经济/竞争”这一理论路径进行分析。第一,从整体来看,高等教育通过经济、社会以及其他诸多因素与社会分层产生关联。人为界定“因果关系”容易失之武断,故而各个因素与教育和社会分层都是通过双向箭头来表示相互的影响。第二,由于部分经济学研究主要强调经济因素在这对关系中的作用,因此经济因素与二者之间的连线为实线。第三,与经济因素不同,社会因素和其他因素(如文化、符号认知等)与二者的连线为虚线。由于社会学研究除了关注社会和其他因素之外,同样也关注经济因素的影响,因此经济、社会和其他因素都被囊括进大的虚线框。在经验研究中,这两个研究取向分别以积极选择假说和消极选择假说为代表。

(二) 积极选择假说与消极选择假说的争论

从实证研究的角度考察教育对于社会分层的影响,一个直接途径就是估计教育在不同阶层、群体中的边际回报及其模式。已有研究对教育回报的异质性模式的讨论大致可以归纳为两个竞争性的研究路径和假说:“理性—行为”模式及其对应的积极选择假说(positive selection hypothesis)、“理性选择+社会规范”模式及其对应的消极选择假说(negative selection hypothesis)。

积极选择假说深受人力资本理论影响,它从理性选择和理性行为出发,认为只有当高等教育可以给个人带来更多净收益的时候,人力资本能够带来更大的价值,人们才会选择接受高等教育(Manski, 1990)。

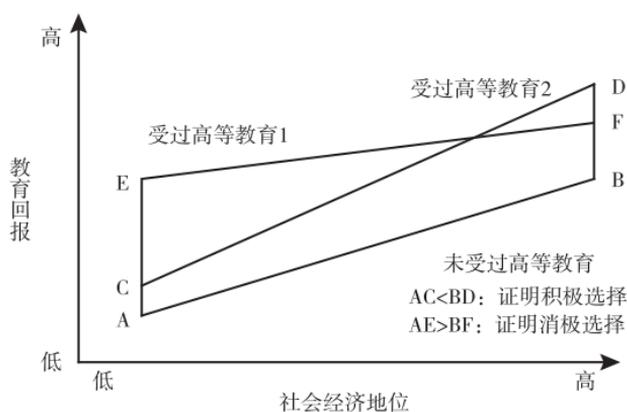


图2 教育回报的两种假说

回报率越高,进入大学的可能性越高。但更好的高等教育需要更多的投资和成本,而较低社会经济地位的家庭难以负担如此高昂的教育费用。因此,家庭经济条件更好的个体更容易进入高等教育,反之则不易进入。不同社会阶层在教育回报方面也存在差异,表现为家庭社会经济地位越高的个体越容易获得更高的教育回报。积极选择假说表明,进入大学时存在一定的选择性,可能造成选择性偏误。积极选择假说站在“成本—收益”的角度指出,接受高等教育的前提在于其收益高于成本,并由此断定经济因素发挥了主导作用,教育回报随着阶层地位上升而提高,如图2所示, $AC < BD$,平均回报率随之提高。从这个思路可以推导出经济因素直接导致了教育回报在不同阶层间的不平等分配,而非经济因素,如社会规范、文化以及个体、家庭结构特征等因素往往被忽视。

消极选择假说则强调,让人们做出选择不仅是经济动因,其原因更应该放到社会结构和规范的框架中来考虑。人们接受高等教育的行为不仅是基于经济因素的理性选择,也受到社会规范和家庭、文化等多重因素的制约和影响(Blau & Duncan,1968)。这种影响在自古以来就重视教育的中国更加突出:接受更多的教育不仅意味着更高的经济回报,也意味着取得更高的社会地位、社会的普遍尊重,进而维持阶层的再生产。社会结构的复杂性和社会情境的多元性也表明,不同社会背景的群体接受高等教育的影响机制是不同的,影响因素也是多元的(Beattie,2002; Lucas,2001)。较高阶层的个体接受教育的可能性比较高,不仅由于其家庭经济地位较高,可以获得的物质资源更为丰富,更

是因为其家庭拥有更丰富的“文化资本”,例如教育经历、文化品位等等(许多多,2017)。这种文化资本通过教育及其背后的阶层再生产制度对下一代进行社会化,使之获得更高阶层的通行证。因此,较高阶层的个体接受高等教育更多出于文化期待而非经济诱因,其经济回报并不会比其家庭地位带来的回报更多。相比之下,较低阶层的个体由于家庭经济地位较低,从家庭地位得到的物质激励也更少。因此,他们接受高等教育尽管也是出于文化的期待,但获得的经济激励的相对量会比先赋家庭经济地位较好的个体更多(Brand & Xie, 2010)。如图 2 中所示, $AE > BF$, 社会出身较低群体接受高等教育获得的收入增量大于较高群体的增量,这个增量随着社会阶层的提高而减小。

需要注意的是,高等教育的经济回报是指控制了其他因素之后的“净”效应。无论积极选择假说还是消极选择假说,都认为接受高等教育之后经济回报的增量不同,但较低阶层获得的教育回报在绝对总量上可能一直低于较高阶层。因此,较低阶层通过教育实现社会阶层向上流动仍然存在较大的阻力。另外,两种假设强调的教育回报增量也都呈线性增、减的关系,这种理想状态也有待进一步探讨。

(三)评述

尽管已有研究在诸多方面对高等教育回报率及其模式进行了丰富的讨论,但在某些方面仍然存在不足。在数据方面,由于不同调查数据在研究目标、抽样方案上存在差异,导致不同数据间缺乏兼容性。在方法方面,尽管国内已有不少研究利用各种统计方法(如工具变量、倾向得分等)来估计教育回报率,但却较少关注教育回报的模式及其问题。在研究视角方面,较少有研究去讨论这种模式是如何随时间变化的。积极选择假设认为较高阶层的人群相对获益更多,而消极选择假设则认为较低阶层的人群相对获益较多。但要注意的是,这两种竞争性假说都是以社会结构相对稳定的成熟社会为基础的。考虑到以往研究所使用的数据及其对应的时代背景和社会结构变迁,这两种假设对正处于转型进程中的中国社会是否适用尚有待讨论。

基于此,本文希望通过多轮调查数据来估计教育回报率的变化过程,同时对教育回报模式在近年来的变化进行描述,以说明高等教育在我国社会发展进程中的作用及未来的方向。

三、研究假设：“倒U形”回报模式假设

尽管高等教育不是社会分层的唯一路径,但毫无疑问,不同的教育选择模式对应着不同的获益阶层,也指向了不同的社会分层模式。积极选择假说和消极选择假说在不同研究中都得到过实证结果的支持,并长期存在争论(Brand & Xie,2010),其中既有理论之争,亦有对研究方法的商榷。

就当下中国社会而言,中产阶级是否已经兴起,是否已经形成中产阶级主导的“橄榄型”社会结构,仍然有待观察。但不可否认的是,改革开放以来,随着国民受教育程度的上升以及人均收入的大幅增加,以各种不同口径测量的“中产阶级”人数和比例都在不断增长,新的中国社会阶层结构正在逐步构建(李培林、张翼,2008)。一方面,中产阶级通过外显的行为方式,如旨趣和品位来重新标定阶层界限;另一方面,他们也通过阶层文化的社会化过程、代际传承、教育和教养方式等来维持阶层再生产。政策导向上对于“橄榄型”社会结构的普遍推崇,加上快速发展的经济等诸多利好因素,使得中产阶级从社会转变进程中获得了较多的回报,在发展中也会获得相当大的优势。不难推断,这种优势也同样体现在教育回报上。

值得注意的是,随着社会结构趋于稳定,阶层流动渠道也逐渐收紧。教育对社会分层,特别是对较低阶层群体向上流动的促进作用正在逐步减小。较低阶层本身就难以接受更好的教育,^①教育回报对其激励作用有限,不利的家庭出身也进一步制约了他们的职业发展和经济收入。反观较高阶层,由于接受高等教育更多出于文化期待而非力图“谋生”,因此他们的经济回报总量可能不会减少,但由教育带来的“净”回报可能“减少”。这种现象有两个可能的原因:第一,高等教育带来的相对回报被中产阶级超过;第二,教育回报带来的物质激励远远小于家庭出身等社会经济条件带来的物质激励。其数量关系见图3。

因此,本文推测是否存在第三种模式,即图3阴影部分中AG—

^① 尽管当前中国的高等教育选择仍然以能力为主(刘精明,2014),但不可否认,大学入学以前的能力培养越来越依赖家庭的社会经济地位。先赋地位在高等教育获得的过程中仍然扮演着举足轻重的作用。

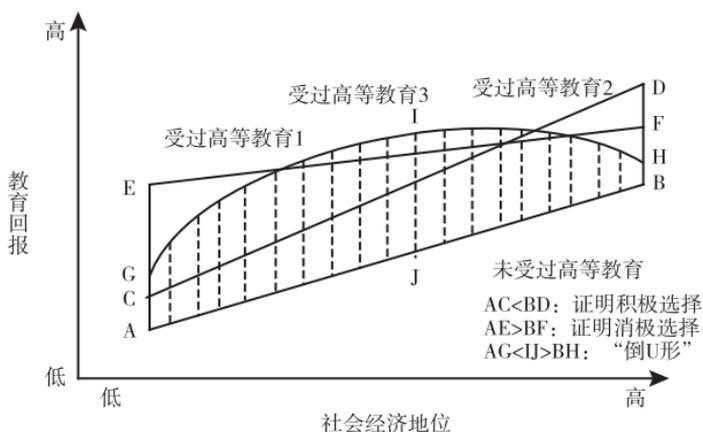


图3 教育回报的“倒U形”回报模式

IJ—BH的“倒U形”回报模式:从接受高等教育的可能性来看,具有中等可能性的人的教育回报相对高于其两侧(进入大学的可能性较高或较低)的人。这种教育回报模式隐含的意义在于:中产阶级可能是高等教育最大的受益者,从高等教育中获得更高的回报和激励最大。在这种假设的视角下,底层和高层的较低回报则体现出更为复杂的影响机制:除了教育本身的作用之外,家庭的社会经济地位、社会资源、社会结构等诸多因素都会对教育回报模式产生至关重要的影响,这也将涉及更多关于教育回报模式和社会分层模式的讨论。这种教育回报模式的影响已超过本文的讨论范围,此处不再过多着墨。

四、数据与方法

(一) 数据和样本选择

本文选用由中国人民大学进行的全国综合社会调查(CGSS)2003—2015年的九期数据。CGSS数据关注社会变迁的趋势,探讨了多个具有科学意义和现实意义的研究议题,是全国性、综合性、连续性的调查项目。该数据采用分层抽样设计方案,全面收集了社会、社区、家庭和个人多个层次的数据,具有较好的覆盖性和较强的推断能力。^①

^① CGSS全面介绍可参见其网站(<http://www.chinagss.org/>),最终访问时间为2019年1月。

由于本文主要关注不同阶层高等教育回报的问题,因此将研究样本限定为正在从事非农职业(有工资收入)的城镇居民。根据被访者是否读过大学,进一步将样本控制在“高中”和“大学”两组。本文主要考察教育回报模式的年代变化趋势,因此并未将数据汇总为统一的数据集,而是对这十年中的九期数据分别加以分析。

(二)方法

1. 同质性假定

当前国内有关教育回报的研究主要是基于明瑟方程(李实、丁赛,2003),其理论基础是同质性假设(即教育的获得及其回报对每个人都是相同的)。

$$Y_i = \alpha + \beta D_i + \gamma EXP_i + \theta EXP_i^2 + \tau \epsilon_i \quad (1a)$$

$$Y_i = \alpha + \beta D_i + \sum \gamma_i X_i + \epsilon_i \quad (1b)$$

其中,因变量 Y_i 为个人月收入自然对数, α 为截距项,自变量中 D_i 为是否接受大学教育, EXP_i 和 EXP_i^2 分别为工作年限和工作经验的平方, ϵ_i 为误差项。实际研究中,如公式(1b)所示,会加入许多影响因素 X_i ,以表示控制影响个人收入的各种因素以获得教育回报的净效应。它假定单位教育回报率为常数,认为现实中个体教育回报率的差别主要取决于随机因素。

对“同质性假设”的批评主要集中于两点:第一,常规 OLS 方法存在内生性问题,无法获得教育回报的无偏估计值(张巍巍、李雪松,2014),其中最有可能的就是遗漏个人能力变量 A_i ,并包含于误差项 ϵ_i 中,这导致 $Cov(A_i, \epsilon_i) \neq 0$,进而 $E(D_i | \epsilon_i) \neq 0$ 。由于存在相反处理结果的数据“缺失”,不可能同时获得 Y_{0i} 和 Y_{1i} 的结果,不能使用传统的 OLS 及 IV 等估计方法,否则会产生估计偏误。第二,有学者认为同质性假设下不能区分不同阶段的教育程度对于教育回报的影响,也不能克服个体选择性对估计结果的干扰。尽管部分研究希望通过设置代理变量、固定效应模型和工具变量等方法来解决内生性问题,但相关研究仍然无法完全解决第二个挑战,即不同阶段教育程度对收益的影响以及个体教育回报率的异质性问题(简必希、宁光杰,2013;刘泽云,2009)。

2. 异质性假定

异质性假设模型是为解决同质性假设存在的问题而提出的,通常

与因果推论相关联。异质性可以分为两种类型(Xie, 2013):实验前的异质性(pre-treatment heterogeneity)^①与实验效应的异质性(heterogeneous treatment effect)。前者指实验组与控制组两组人群可能在本质上就有差异,后者则是指实验组中的每个个体在对待相同的一个实验或干预时的反应各不相同。以高等教育获得为例,实验前的异质性是指上大学的群体与不上大学的群体可能存在本质上的差异(不管这种差异的来源是什么),从而导致OLS等传统估计方法产生较大的系统性偏误;而实验效应的异质性则是指获得高等教育(上大学)所能带来的教育回报对每个人而言都可能是不同的。

Heckman的样本选择模型常用于解决实验前异质性问题:标准模型通过引入选择方程来描述样本的选择行为,继而假定收入方程和选择方程的不可观测变量服从联合正态分布,最后通过计算求得逆米尔斯比率来消除选择偏差。基于Heckman模型,部分研究放宽了其过强的假设,采用了半参数和非参数方法进行估计(Ichimura, 1991)。这也推动了更多半参及非参估计方法及其他配套估计方法的出现,部分研究还使用“局部线性回归”或者“局部工具变量回归”方法(Imbens & Angrist, 1994)、边际实验效应(MTEs)(Heckman & Vytlacil, 1999)来估算匹配估计量;而社会学研究中则通常使用倾向得分匹配方法。

3. 统计模型

根据反事实分析框架,本文首先将样本范围控制在高中和大学(及以上)两种教育水平。令 $D_i = 1$ 表示大学学历, $D_i = 0$ 表示高中教育。大多数情况下,教育回报与受教育水平并不呈线性关系,这也使得传统收入对数对于回归系数的估计方法产生较大偏差。进入大学“与否”这个选择过程遵循下述规则:

$$D_i = \begin{cases} 0, D_i^* < 0 \\ 1, D_i^* \geq 0 \end{cases} \quad (2)$$

$$\begin{aligned} D_i^* &= P_i(X_i) + F(u) \\ &= p + \epsilon_{Di} \end{aligned}$$

① 内生性偏差的来源既有实验前异质性,也有包括测量误差等在内的其他各种可能的原因。严格地讲,实验前异质性(或选择性偏差, selection bias)并不等价于内生性,而只是引起内生性的一种原因。但社会科学研究一般都会忽略测量误差等原因,因此可以将实验前异质性等价于内生性。

$$\begin{aligned}
 &= W_i E_i X_i + (1 - W_i) S_i X_i + F(u) \\
 &= W_i \lambda'_1 X_i + (1 - W_i) \lambda'_2 X_i + F(u) \\
 &= \lambda' X_i + F(u)
 \end{aligned} \tag{3}$$

$$p = P_i(X_i | D_i = 1) = \lambda' X_i \tag{4}$$

其中 D_i^* 为潜变量,代表大学入学的净收益, X_i 是可观测到的变量向量, E_i 为经济因素(是协变量 $\lambda'_1 X_i$ 的线性函数), S_i 为非经济因素(是协变量 $\lambda'_2 X_i$ 的线性函数), W_i ($0 \leq W_i \leq 1$) 为经济因素所占权重。 ϵ_{Di} 表示个体在选择过程中未被观察到的异质性, $F(u)$ 表示潜变量 ϵ_{Di} 的分布函数。 $p = P_i(X_i | D_i = 1)$ 表示接受大学教育的概率,并受到一组协变量的作用。在本文中,采用 *logit* 模型来估计不同年份个体大学入学概率,即预测倾向值得分 p 。

第二步,解决实验效应的异质性,由于存在两种选择结果:

$$\begin{cases} Y_{0i} = \gamma_0 X_i + \epsilon_{0i} & (\text{当 } D_i = 0 \text{ 时}) \\ Y_{1i} = \gamma_1 X_i + \epsilon_{1i} & (\text{当 } D_i = 1 \text{ 时}) \end{cases} \tag{5}$$

可知:

$$\begin{cases} E(\epsilon_{0i} | X_i) = 0 & (\text{当 } D_i = 0 \text{ 时}) \\ E(\epsilon_{1i} | X_i) = 0 & (\text{当 } D_i = 1 \text{ 时}) \end{cases} \tag{6}$$

进一步推导可得:

$$Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i} \tag{7}$$

$$= [(\gamma_1 - \gamma_0) X_i] D_i + \gamma_0 X_i + [\epsilon_{0i} + (\epsilon_{1i} - \epsilon_{0i}) D_i] \tag{7a}$$

$$= [(\gamma_1 - \gamma_0) X_i + (\epsilon_{1i} - \epsilon_{0i})] D_i + \gamma_0 X_i + \epsilon_{0i} \tag{7b}$$

$$= \beta_i D_i + \gamma_0 X_i + \epsilon_{0i} \tag{7c}$$

其中,

$$\beta_i = (\gamma_1 - \gamma_0) X_i + (\epsilon_{1i} - \epsilon_{0i}) \tag{8}$$

β_i 表示个体的异质性教育回报,是服从某种分布的随机变量:(1) 当 $\gamma_1 \neq \gamma_0$ 时,表明存在观测到的异质性 $(\gamma_1 - \gamma_0) X_i$; (2) 当 $\epsilon_{1i} \neq \epsilon_{0i}$ 时,存在未被观测到的异质性 $(\epsilon_{1i} - \epsilon_{0i})$ 。不难得知 $E(\epsilon_{1i} - \epsilon_{0i} | X_i) = 0$, 因此,在给定 X_i 的条件下, β_i 的均值为:

$$\beta = E(\beta_i | X_i) = E(\gamma_1 - \gamma_0) X_i \tag{9}$$

第三步,在异质性假设的条件下,根据公式(1b),我们认为 α 为实验前异质性, β 为实验效应异质性,可以得到统计模型:

$$Y_i = \alpha + \beta D_i + \sum \gamma_i X_i + \epsilon_i \quad (10)$$

为讨论实验效应的异质性问题,本文拟采用 HTE 模型(Heterogeneous Treatment Effect),将 β 进行分解。该模型将倾向值分成若干层次(block),采用半参数或者非参数方法,分别估计各层的 ATT,这些层次类似于分层线性模型中的层-2 分析单位(Brand & Xie,2010)。该方法仍然遵循倾向得分匹配的“可忽略假设”,可以在同一研究中使用多种参数(Zhou & Xie,2016),同时避免单纯倾向得分匹配可能带来的不平衡问题(King & Nielsen,2016),适用于 LIV 方法并非最佳的情况。

HTE 模型有三种估计方法(Xie et al.,2007)。本文采用基于多项式回归的“平滑差分”(Smoothing Difference Methods)方法,通过差分来消除实验前异质性的影响。首先,选择进入模型的自变量向量集,采用 logit(或 probit)模型估计倾向值;其次,基于不同的倾向值得分分别对控制组和实验组进行非参数回归,根据匹配的带宽(bandwidth)使用局部多项式回归;再次,在不同水平的倾向得分上获得控制组和实验组的非参数多项式回归拟合线条,以获得异质性实验效应在不同倾向值上的分布。最后,利用非参数方法分别对每年的数据估计不同倾向得分群体(潜在的)高等教育收益,并归纳教育回报的模式。

(三)变量选择

根据已有研究,本文筛选受过高中及以上教育、在职且从事非农工作的城市居民作为分析样本,而农业工作、无工作、退休等样本在本文中暂时不加考虑。这样既能保证样本框的一致性,也可以保留比较的可能性。从不同年份的样本变异程度、应答情况和缺失值、模型设置等几个方面考虑,各年份样本、变量情况会有些许不同。

主要的研究变量为月收入 and “是否读过大学”。考虑到通货膨胀因素,研究使用的因变量为月收入的对数^①以及“是否读过大学”。月

① 在教育回报研究中,通常的做法是选用月工资或者周薪、时薪等指标。由于 CGSS 数据中时薪和周薪等测量指标缺失,本文使用月收入作为因变量。CGSS 数据在 2003-2006 年测量了月工资收入,在 2008-2015 年测量了年工资收入和总收入,本研究在对测量指标进行操作化处理时,将年工资收入除以 12,得到其月工资收入。

收入以 2003 年为基准,进行不变价格处理;对收入取自然对数,这样可以使收入具有可比性,并消除其潜在的异方差性等问题。是否读过大学则是以“高中”参照组,“大学”为干预组。

自变量的选取主要分成三部分:控制变量、个体层面变量和家庭层面变量。控制变量^①主要包括:性别、婚姻状况、党员身份、工作所处部门性质、是否拥有房屋产权等。此外,控制变量还包括住房面积和年龄两个连续变量。

个体信息变量主要包括:家庭规模、14 岁时户籍所在地和大学入学队列。大学入学队列根据重大历史事件和历史发展进程,以参加高考和(可能的)高校入学时间来划分组别,共分为四组。^②

父母信息变量共四个,包括父母受教育程度和父母职业得分。这些变量均为子女 14 岁时父母的信息,且均为连续变量。根据惯例,将不同教育程度转换为不同受教育年限:文盲 = 0,私塾/小学 = 6,初中 = 9,中专/普通高中/职业高中 = 12,大专 = 14,大学本科 = 16,研究生及以上 = 19。父母职业得分则根据 ISCO - 88 标准,转换为百分制。

五、分析结果

(一)基于倾向得分匹配的教育回报率估计^③

为解决实验前异质性问题,本文将采用 logit 模型(选择模型)来估计每个案例进入大学、获得高等教育的概率,即倾向得分值。具体操作中,对这九期数据逐个计算,以呈现不同时期高等教育入学的具体情况(见表 1)。

① 这些变量大多数为当前状态的测量,且可随时间变化,并不能直接“影响”个体的大学入学。之所以将其纳入模型,一是为了控制不可观测的异质性,二是为了将其作为可能遗失变量的代理变量,包括其个人能力、禀赋、先赋特质等。

② 1966 年为第一个时间节点,以 1966 年以前参加高考人群为参照组,分别设定 1967 - 1982 年组、1983 - 1998 年组以及 1999 年之后参加高考组别。严格地讲,1966 - 1982 年应该分为两个阶段。第一个阶段是 1966 - 1976 年,此时正常的高考被取消,但仍然保留了部分推荐读大学(工农兵学员)的名额。由于样本中在这个时间段就读大学并完成学业的人数较少(<1%),且数据中也没有涉及工农兵学员的选项,故与后者合并在一起。第二个阶段为 1977 - 1982 年,即恢复高考的前五年,此时由于大量文革期间中断学业的考生参加考试,出现了“堆积”现象。因此,在入学队列上将这两个群体归为同一组。

③ 篇幅所限,倾向得分各层匹配结果未附于正文中,如有需要可以联系作者。

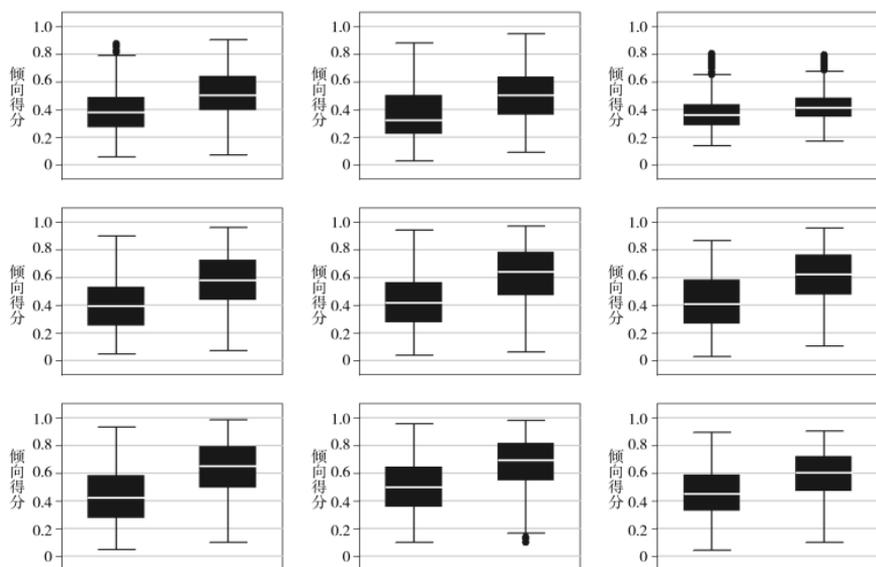
表 1 倾向得分估计模型(选择模型)

	2003	2005	2006	2008	2010	2011	2012	2013	2015
性别(女=0)	0.086 (0.111)	-0.054 (0.009)	0.025 (0.116)	-0.055 (0.128)	-0.118 (0.103)	-0.372* (0.156)	-0.264* (0.116)	-0.306** (0.105)	
年龄	-0.020** (0.007)	1.250*** (0.178)	-0.040** (0.014)	-0.024 (0.016)	1.714*** (0.126)	-0.041* (0.020)	-0.032* (0.013)		
党员身份 (非党员=0)	1.403*** (0.132)	0.191 (0.155)	1.326*** (0.170)	1.432*** (0.175)	0.563*** (0.136)	1.177*** (0.214)	1.568*** (0.140)	1.037*** (0.164)	
国有部门 (私有部门=0)						0.972*** (0.157)		0.606*** (0.119)	
高考队列			0.008 (0.191)	0.300 (0.216)	0.590*** (0.075)	0.275 (0.266)	0.327 (0.179)	0.713*** (0.202)	0.462*** (0.078)
14岁户口(农村=0)	-0.116 (0.120)	1.313 (0.685)	-0.183 (0.131)	-0.249 (0.138)	0.094 (0.106)	0.137 (0.161)	0.186 (0.116)	0.307* (0.125)	0.261* (0.112)
家庭规模	-0.194*** (0.051)	-0.087 (0.066)	-0.027 (0.033)	-0.136* (0.060)	-0.208*** (0.043)	-0.166** (0.063)	-0.206*** (0.046)	-0.153** (0.052)	-0.089* (0.042)
父亲职业得分	0.001 (0.004)	0.028*** (0.004)	0.007* (0.004)		0.007* (0.003)	0.006 (0.005)	0.013** (0.004)	0.006 (0.004)	0.010** (0.003)
父亲教育程度	0.078*** (0.018)			0.051* (0.021)	0.109*** (0.015)	0.056* (0.023)	0.082*** (0.017)	0.063** (0.020)	0.104*** (0.016)
母亲教育程度	0.049** (0.016)			0.085*** (0.020)			0.045** (0.017)	0.058** (0.018)	
样本量	1614	1076	1382	1265	2023	859	1959	1564	1724
LL	-1002.068	-649.976	-882.327	-759.721	-1188.305	-512.652	-1137.477	-908.407	-1073.252
Pseudo R ²	0.097	0.114	0.050	0.133	0.150	0.138	0.159	0.137	0.101

注:(1)单元格括号内为标准误。(2) $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ 。(3)考虑到模型的基本设定和变量缺失值、倾向得分匹配平衡性等因素,进入选择模型的变量在不同年份会有不同;高考队列在此作为连续变量;2005年数据中没有14岁户口信息,故用现户口替代;2006年职业地位变量为被访者18岁时指标;2012年父亲职业地位缺失值较多,以母亲职业地位替换。(4)篇幅所限,常数项和控制变量中住房和婚姻信息没有放进表格。

1. 匹配后实验组和控制组共同支持区间

为方便比较不同群体之间的差异,本文将上述计算得到的倾向得分分值划分为十层。通过变量的筛选和模型设置来平衡模型并完成倾向值打分,使层内的实验组和控制组在主要特征上基本相同,进而为后续的各种匹配方法提供基础。对 CGSS 2003-2013 的九期数据分别进行倾向得分匹配,得到历年两个组别的共同支持区间(详见图 4)。可以看出,九期数据中实验组与控制组有足够的共同支持区间。



注:(1)从上排左起分别为2003、2005、2006年,第二排左起分别为2008、2010、2011年,第三排左起分别为2012、2013、2015年。(2)每个单元格的箱线图中,左侧为控制组(高中学历),右侧为实验组(大学及以上学历)。

图4 倾向得分匹配后控制组和实验组共同支持区间(箱线图)

2. 匹配估计量和倾向得分加权回归

在上述分层匹配以后,本文利用近邻匹配、马氏距离法、分层匹配、核匹配这四种基于半参数和非参数的估计方法,估计得到各年份的平均实验效应(ATT,即教育回报率)。^①有关结果请见表2。以2003年

① 高等教育回报率的估计通常根据明瑟方程的设定而来, $Y_i = \alpha + \beta D_i + \sum \gamma_i X_i + \epsilon_i$,其中 β 为是否就读大学的教育系数。由于通常假定大学教育为四年制,所以教育回报率就等于 $\beta/4$ 。例如, $\beta=0.414$,则高等教育回报率为10.325%。倾向得分的匹配估计量等同于 β 值,其计算回报率的过程与之相同。

的教育回报率为例,采用近邻匹配法、马氏距离法、分层匹配和核匹配的估计结果分别是 0.414、0.393、0.408 和 0.399,比较接近。

为了在同一个标准之下比较教育回报,本文进一步使用了倾向加权回归方法。倾向加权回归与上述匹配估计方法的区别在于:匹配估计量是利用不同的方法根据样本间距离进行匹配及估计,而倾向得分加权回归则是利用相应的公式将倾向得分值转化为权重,采用普通最小二乘法估计平均实验效应。权重的计算为:

$$\begin{cases} ATE_w = 1/(1 - ps) & (\text{当 } D_i = 0 \text{ 时}) \\ ATE_w = 1/ps & (\text{当 } D_i = 1 \text{ 时}) \end{cases} \quad (2)$$

表 2 通过自助法 (Bootstrap) 以及倾向得分加权得到的各种估计量

年份	匹配估计量				倾向得分加权
	近邻匹配	马氏距离	分层匹配	核匹配	
2003	0.414	0.393	0.408	0.399	0.375
	8.214	9.469	9.813	11.771	10.255
2005	0.304	0.249	0.329	0.329	0.332
	4.982	4.166	7.524	7.541	8.166
2006	0.362	0.269	0.343	0.341	0.351
	5.723	5.039	9.051	9.030	8.481
2008	0.409	0.432	0.409	0.408	0.394
	6.450	9.204	7.641	9.419	8.843
2010	0.535	0.456	0.499	0.523	0.521
	8.703	7.541	10.049	12.851	11.448
2011	0.365	0.386	0.419	0.406	0.437
	3.699	6.253	6.270	6.565	6.350
2012	0.460	0.418	0.403	0.431	0.409
	5.693	9.592	7.393	9.417	9.325
2013	0.368	0.303	0.381	0.398	0.377
	6.543	7.529	7.012	8.608	8.472
2015	0.424	0.448	0.425	0.445	0.439
	6.888	8.824	8.483	11.145	10.376

注:单元格内上下两行数字分别为教育回报估计量和 t 值。

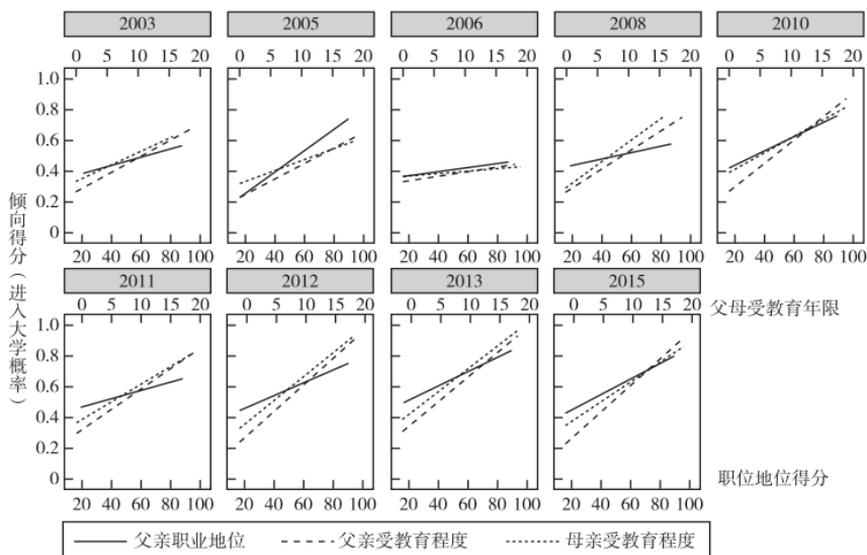
(二) HTE“平滑差分”模型和“倒 U 形”模式的验证

1. 社会经济地位与高等教育获得概率间的关系

以往的研究表明,表现家庭社会经济地位(如父母受教育程度、职业得分等)的先赋特征对高等教育入学有直接的积极影响(Blau &

Duncan, 1968)。这些变量的系数也均呈现显著的影响。

为了直观地展示这种影响,在控制其他相关变量的前提下,分别选择父母的受教育年限和父亲职业地位为代表性指标,以描述家庭经济地位与倾向得分之间的关系(Brand & Xie, 2010)。图5显示,家庭的社会经济地位越高,获得高等教育的概率就越大,而且历年来这种正相关关系基本稳定。



注:图中左侧纵坐标表示倾向得分,即获得高等教育的概率;图形底部为主横坐标,表示父亲的职业得分,而图形顶部次横坐标为父母的受教育年限。

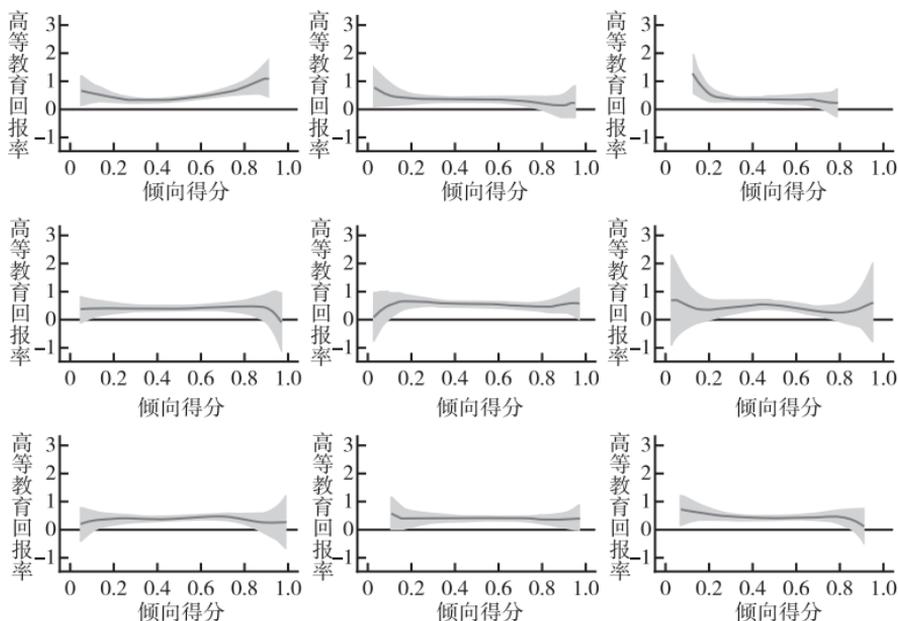
图5 中国高等教育入学概率与社会经济地位的关系

2. HTE 模型估计结果

不同年份高等教育回报率的异质性估计结果可见图6。其中横坐标为倾向得分,表示接受高等教育的概率。纵坐标为高等教育回报率,虚线表示每个倾向得分值所对应的高等教育回报率的均值。阴影部分表示置信区间,一旦置信区间包含0值,则表明在95%的置信度下可以接受原假设($H_0: ATE = 0$),即高等教育回报率等于0。否则倾向得分值对应的教育回报率显著不等于0。

如前所述,获得高等教育的概率与家庭社会经济地位呈正相关。因此,可以将倾向得分解读为家庭社会经济地位:不同倾向得分的人群可能获得的教育回报率等同于不同家庭经济地位的人群可能获得的教育回报率。这也契合了积极选择假说和消极选择假说中关于不同社

会经济地位人群的高等教育回报率存在差异的讨论。由图 6 可以看出,从 2003 年到 2015 年,高等教育回报经历了两个阶段的演变。



注:从上排左起分别为 2003、2005、2006 年,第二排左起分别为 2008、2010、2011 年,第三排左起分别为 2012、2013、2015 年。

图 6 2003 - 2015 年高等教育回报的异质性模式

2003 - 2006 年表现为“波动—中低层激励型”。2003 年的回报率大体随倾向得分值升高而升高,即阶层越高,教育回报率越高,但底层的回报率已经有向上抬升的趋势。2005 年,从均值来看,倾向得分值处于较低部分的回报率继续提升。但倾向得分值较低和较高部分的教育回报的区间包含 0 值,这个结果可能与匹配过程中两组的内部方差较大有关。2006 年的估计结果进一步向较低层次倾斜,教育回报率随倾向得分值的提高而逐步下降,即倾向得分值越低,教育回报率越高,倾向得分值最高部分教育回报率不显著。三轮调查的分析结果表明,教育回报率逐渐从较高的倾向得分群体向得分较低群体转移。对这一结果可以这样理解:由于 2003 年之前的大学毕业生仍然保留了扩招前的部分“精英”特征,接受高等教育仍然可以获得非常可观的回报;而在 2005 - 2006 年期间,教育扩招已推行了数年并对劳动力市场产生了一定的影响,教育回报的天平就不断向中层和底层倾斜。在这一阶段,高等教育扩展在促进教育公平的层面上取得了积极效果。

2008 - 2015 年体现为“稳定—中层激励型”。自 2008 年开始,高等教育回报模式趋于稳定,倾向得分值较低与较高的群体的教育回报率基本不显著(即阴影部分包括 0 值),只有倾向得分值处于中间的群体有显著不为 0 的教育回报率。换言之,从高等教育中受益最多的是倾向得分居中的群体,也就是社会地位上的中间阶层或中产阶层群体;而对于社会经济地位较低或较高的群体,接受高等教育并不能带来显著的激励作用。可以推测,在实施教育制度改革较长时间之后,劳动力市场已经形成一套稳定的、有利于中间阶层的社会分流机制。

可以想见,因家庭经济地位较低而不太可能接受高等教育的人群在就业市场中依旧处于不利地位。有效维持不平等假设(EMI)认为,相比逐渐“饱和”的大学教育机会,重点大学,也即优质教育资源仍然是稀缺的,较低社会阶层群体仍然难以接受重点大学的教育。即使接受了高等教育,其在劳动力市场中的不利地位也会使他们无法获得相对较高的教育回报,从而使高等教育失去了对这部分人的激励作用。另一方面,社会经济地位相对较高、最有可能接受高等教育的人群有各种丰富的资源和较好的先赋条件,在就业市场中占据着有利地位。因此,即使他们不接受高等教育,同样也可能获得较高的收入。来自家庭和先赋条件的优势掩盖了高等教育带来的正向激励。因此,在这个阶段,高等教育真正的激励作用发生在中间阶层身上。他们不具备能轻易获得高等教育就学资格的优势地位,但又不至于被甩在分层的末端。他们需要通过一定的努力和竞争获得高等教育资格,并在就业市场取得更高收益,实现社会阶层进一步的向上流动。

这些结果表明,经历了高等教育扩展初期的普及化和平等化趋势之后,高等教育的回报模式逐渐稳定下来,形成了对中产阶层最有利却不利于社会底层的模式。这个结果在某种程度上检验了本文提出的“倒 U 形回报模式”假设。

(三) 稳健性检验

高等教育的重要性不仅在于提供知识和工作技能,也在于提供一个开放的、公平竞争的平台(许多多,2017)。对于传统上在工作中处于弱势地位的女性更是如此。高等教育的扩展使女性获得了更多的教育机会,也大幅提高了女性的人力资本,近年来年轻女性接受大学教育的比例已经超过了男性(贺光烨,2018)。因此,尽管仍然存在针对女

性的就业歧视,但女性获得的工作机会仍然得到了大幅的增长。

为此,本文分别针对女性和男性两个样本子群体进行模型检验。^①从女性样本来看,2003 - 2006 年回报模式仍然表现为“波动—中低层激励型”。其中 2003 年体现为积极选择的形态,2005 年为过渡形态,2006 年体现为消极选择的形态。2008 - 2015 年这六期数据中,除 2012 年以外,其他年份均体现出中间阶层的实际教育回报高于两端的趋势,验证了“稳定—中层激励型”的趋势。男性样本的波动形态与女性大致相似,但时间上更早进入倒 U 形的形态。

六、结论和讨论

(一) 结论

本文利用 CGSS 2003 - 2015 年的九期数据,使用 PSM、HTE 模型等多种方法解决忽略重要变量和样本选择性等内生性问题,估计了中国 2003 - 2015 年高等教育回报率的历时演变过程,并估计了不同时期的高等教育回报模式及其演变过程。从教育回报模式来看,2003 - 2006 年间表现为“波动—中低层激励型”,2008 年之后则呈现为“倒 U 形”回报模式,即高等教育回报率只对倾向得分值处于中间的人群(对应于中产阶级)显著,但对两侧(较高与较低倾向得分)的人群则不显著。

从教育回报模式来看,2003 - 2006 年,倾向得分处于中部或底部的群体潜在教育回报逐渐升高,即“波动—中底层激励型”。但如果从 1999 年高等教育扩招的时间来看,2003 年前后正是扩招的学生加入劳动力市场的阶段。而 2003 年以后的几年里中层及底层群体教育回报率提高,也正是高等教育扩招带来的正向的平等化效应。2008 年之后,高等教育回报模式发生了转化,并逐步趋于稳定。高等教育给最低和最高阶层的经济激励甚微,而中间阶层成为教育回报更高的人群。高等教育回报模式呈现出“稳定—中层激励型”态势,即“倒 U 形”回报模式。

从高等教育入学机会与高等教育回报率的角度来看,个人能力和

^① 篇幅所限,稳健性分析的图形略去,如有需要可联系作者。

家庭出身同时显著地影响高等教育入学机会(刘精明,2014)。但个人能力本质上仍然受自身家庭社会经济地位影响。即家庭出身影响着个人能力,能力又影响入学机会,进而影响个体的教育回报。从这个意义上讲,家庭出身不仅影响高等教育的入学机会,还影响着个体的高等教育回报率。对底层群体而言,获得优质大学教育资源的机会微乎其微,尽管可以就读于较差的本科或者专科,但这样所获得的经济回报也相当有限。这也符合“有效维持不平等假设”(EMI)。而处于较高阶层的人群由高等教育带来的“净”的经济回报看似较为有限,实则是被其家庭社会经济地位带来的其他形式的资本回报掩盖了,并因此导致这部分的经济回报“不显著”。而唯有对中间阶层的群体来说,是否接受高等教育决定了其能否进入收入相对较好的职业与行业,并带来显著正向的教育回报。

中间阶层作为高等教育回报率最高的群体,也将是当前社会结构中获益最多的群体。他们在维持社会再生产过程的同时仍然有一定的上升空间,从而在一定程度上保证了社会流动性,使社会结构趋于稳定和合理。

(二) 进一步讨论

教育回报是人力资本、社会流动和分层研究的重要议题。本文在两个方面继续深入:一是讨论了数据与估计方法的问题,如内生性问题、历年倾向得分估计模型中的变量设置问题、HTE模型本身的方法论等;二是讨论了我国当前高等教育回报率水平与模式,并对高等教育与社会不平等和社会分层的关系进行了深入的理论讨论。

必须注意的是,本文的研究方法仍然存在潜在的问题。一是数据的选取与使用问题。适合进行长时段的社会变迁研究的追踪数据相对较少,CGSS覆盖时间较长,但并非追踪数据,且更换过抽样框。因此,在进行因果推论时,采用CGSS这一混合截面数据需要有较强的前提假定,要求每期都有较好的代表性以及测量的稳健性。二是变量层面的问题,个体过往(如14岁时)信息的选择和采集存在“回溯性偏误”的可能。三是CGSS数据中缺少对高等教育质量的详细区分,如重点大学、专业等等,导致对不同阶层所获高等教育的质量、激励缺乏区分和进一步讨论。

针对“倒U形”回报模式本身,我们必须谨慎思考这一分层过程是

否逐步趋于稳定,是否强化了底层和上层的固化。这个问题的潜在含义在于:何种分层模式、何种社会结构是最优的?本文无意进行更多的价值判断。但从社会公平的角度考虑,教育能否成为社会底层有效的社会流动渠道,仍然需要在相关政策制定和评估时加强讨论,使教育真正成为塑造社会结构的积极因素。这既涉及基础教育的公平性问题,也涉及高等教育入学机会的公平性问题。本文对于高等教育回报模式的发现,特别是“倒U形”回报模式的发现仍然是初步的,尚需进一步的讨论与事实的检验。

参考文献:

- 布尔迪厄,皮埃尔,2004,《国家精英:名牌大学与群体精神》,杨亚平译,北京:商务印书馆。
- 方长春、风笑天,2018,《社会出身与教育获得——基于 CGSS 70 个年龄组数据的历史考察》,《社会学研究》第 2 期。
- 贺光桦,2018,《专业选择与初职获得的性别差异:基于“首都大学生成长追踪调查”的发现》,《社会》第 2 期。
- 侯利明,2015,《地位下降回避还是学历下降回避——教育不平等生成机制再探讨(1978 - 2006)》,《社会学研究》第 2 期。
- 简必希、宁光杰,2013,《教育异质性回报的对比研究》,《经济研究》第 2 期。
- 李春玲,2014,《教育不平等的年代变化趋势(1940 - 2010)——对城乡教育机会不平等的再考察》,《社会学研究》第 2 期。
- 李培林、张翼,2008,《中国中产阶级的规模、认同和社会态度》,《社会》第 2 期。
- 李培林、朱迪,2015,《努力形成橄榄型分配格局——基于 2006 - 2013 年中国社会状况调查数据的分析》,《中国社会科学》第 1 期。
- 李实、丁赛,2003,《中国城镇教育收益率的长期变动趋势》,《中国社会科学》第 6 期。
- 刘精明,2008,《中国基础教育领域中的机会不平等及其变化》,《中国社会科学》第 5 期。
- ,2014,《能力与出身:高等教育入学机会分配的机制分析》,《中国社会科学》第 8 期。
- 刘泽云,2009,《教育收益率估算中的几个方法问题》,《北京大学教育评论》第 1 期。
- 唐俊超,2015,《输在起跑线——再议中国社会的教育不平等(1978 - 2008)》,《社会学研究》第 3 期。
- 许多多,2017,《大学如何改变寒门学子命运:家庭贫困、非认知能力和初职收入》,《社会》第 4 期。
- 叶华、吴晓刚,2011,《生育率下降与中国男女教育的平等化趋势》,《社会学研究》第 5 期。
- 叶晓阳、丁延庆,2015,《扩张的中国高等教育:教育质量与社会分层》,《社会》第 3 期。
- 张巍巍、李雪松,2014,《中国高等教育异质性回报的变化:1992 - 2009——基于 MTE 方法的实证研究》,《首都经济贸易大学学报》第 3 期。
- 张兆曙、陈奇,2013,《高校扩招与高等教育机会的性别平等化——基于中国综合社会调查(CGSS2008)数据的实证分析》,《社会学研究》第 2 期。
- 郑辉、李路路,2009,《中国城市的精英代际转化与阶层再生产》,《社会学研究》第 6 期。

- Beattie, Irene R. 2002, "Are All 'Adolescent Econometricians' Created Equal? Racial, Class, and Gender Differences in College Enrollment." *Sociology of Education* 75(1).
- Becker, Gary S. 1964, *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Chicago: University of Chicago Press.
- Blau, Peter M. & Otis Duncan 1968, "The American Occupational Structure." *American Sociological Review* 33(2).
- Brand, Jennie E. & Yu Xie 2010, "Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." *American Sociological Review* 75(2).
- Heckman, James J. & Edward J. Vytlacil 1999, "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects." *Proceedings of the National Academy of Sciences* 96(8).
- Ichimura, Hidehiko 1991 "Semiparametric Least Squares (SLS) and Weighted SLS Estimation of Single-index Models." *Journal of Econometrics* 58(s1-2).
- Imbens, Guido W. & Joshua D. Angrist 1994, "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2).
- King, Gary & Richard Nielsen 2019, "Why Propensity Scores Should Not Be Used for Matching." *Political Analysis* 27(4).
- Lucas, Samuel R. 2001, "Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects." *American Journal of Sociology* 106(6).
- Manski, Charles F. 1990, "Nonparametric Bounds on Treatment Effects." *The American Economic Review* 80(2).
- Nee, Victor 1989, "A Theory of Market Transition: From Redistribution to Markets in State Socialism." *American Sociological Review* 54(5).
- Walder, Andrew G., Bobai Li & Donald J. Treiman 2000, "Politics and Life Chances in a State Socialist Regime: Dual Career Paths into the Urban Chinese Elite, 1949 to 1996." *American Sociological Review* 65(2).
- Wu, Xiaogang 2010, "Economic Transition, School Expansion and Educational Inequality in China, 1990-2000." *Research in Social Stratification and Mobility* 28(1).
- Xie, Yu 2013, "Population Heterogeneity and Causal Inference." *Proceedings of the National Academy of Sciences* 110(16).
- Xie, Yu, Jennie E. Brand & Ben Jann 2012, "Estimating Heterogeneous Treatment Effects with Observational Data." *Sociological Methodology* 42(1).
- Zhou, Xiang & Yu Xie 2016, "Propensity Score-based Methods Versus MTE-based Methods in Causal Inference Identification, Estimation, and Application." *Sociological Methods & Research* 45(1).

作者单位:中国社会科学院社会发展战略研究院(郭冉)
北京大学社会学系(周皓)

责任编辑:杨可