

阈值依赖的教育扩张 与教育机会不平等*

——以 A 省某年高考数据为例

李代

提要:本文基于 A 省某年高考成绩的总体数据,采用计算方法来模拟扩大招生对优势与劣势地区间升学机会分配的影响。研究发现,假设教育扩张是外生的,不与人的行动发生互动,扩张对不平等的影响是先升后降:始于高分段的扩张会导致教育不平等程度上升,精英高校尤为明显;而始于低分段的扩张则会减少教育不平等。由此,本文提出了“阈值依赖不平等”理论。

关键词:教育不平等 地区差异 教育扩张

一、研究背景

高等教育扩张,或者说大学扩招,是改革开放以来中国的一件大事。由于教育与个人未来的职业收入、社会地位等因素密切相关,教育扩张带来的教育机会分配的变化直接影响着劳动力市场供求的变化,更关系着千万人社会流动的前景。从城乡差异的角度来看,教育扩张到底有没有增进教育平等呢?

刘精明(2006)采用 GSS2003 数据,认为与 1978 - 1997 年相比,1998 年开始的大学扩招导致本科教育不平等加大。李春玲(2010)指出,根据国家统计局 2005 年 1% 人口抽样调查数据,与 20 世纪 90 年代初相比,扩招后高等教育机会的不平等不降反升。李春玲(2014)进一步对 2006、2008 和 2011 年全国抽样调查数据的整合数据进行研究,从 1940 年以来,高等教育的城乡差距在扩大。与中国教育扩张与教育不平等相关的经验研究还有很多(李春玲,2003;刘精明,2008;李

* 作者感谢张涵、程思薇、周飞舟对本文的重要建议与贡献。

煜,2006;吴愈晓,2013a, 2013b),发现的趋势比较相近,这里不一一赘述。

现有研究的理论和数据方面仍有局限。其理论往往借鉴西方,但对其暗含的社会条件与中国的情况契合与否关注不多;数据则多采用全国的大范围抽样调查数据,但由于中国高考以省为单位进行,招生是各校下到各省分别进行,这些数据反映的情况与实际的升学存在层次的差异。高考分数和招生名额的地域分布差异巨大,抽样调查数据中即使有高考分数这个变量,不按省份区别对待也是不合理的,而这就带来了自由度的问题。何况,受访人是否愿意或能够准确地报告高考卷面分数也是一个问题。

有感于此,本文希望在这两方面有所推进。本文关心的问题是中國高等教育扩张对各群体间教育不平等有何种影响。本文研究的人群并非全体适龄国民,而是报名参加高考的人。本文采用的数据是近十年中某年A省的高考卷面分数——也就是俗称“裸分”的几乎全部数据。^①

本文的贡献在于两点。首先,本文所用数据是录取过程中主要的依据,以此来勾勒升学的情况,损失的信息要远远小于社会分层研究常用的调查数据,其十万级的样本规模已经非常接近总体。作为省级数据,尽管还是没有各个招生院校的信息,但在阐释上犯“生态谬误”的风险已经有所降低。其次,本文利用这一数据建构的理论不但更加贴近中国升学机制,而且大可加以延伸,用于研究其他社会语境相同的问题。本文以为,对一省一年之数据详加分析,虽然并非社会分层与流动研究中常见的范式,但廓清了很多此前的理论中没能分离出来的机制,对理解问题有很大的益处。下文将简单介绍该领域常见的理论,并在其基础上展开分析。

二、理论综述

不少社会学研究者有一种假设,就是随着国家工业化的推进、教育的扩张,不同背景的人群获得教育机会的概率会越来越接近,教育机会

^① 本文得到授权使用该数据进行学术研究,但按国家保密政策规定,在此不透露具体来源,下文也不会报告准确的人数、分数和分数线。但是这并不影响本文分析的展开。

在社会中的分布更加平等(Treiman, 1970)。但是经验研究表明,在很多工业化社会里,教育不平等的程度并没有明显降低。例如,沙维特和布洛斯菲尔德(Shavit & Blossfeld, 1993)发现,在其研究涉及的13个工业化国家里,除瑞典与荷兰两个国家之外,其他国家社会背景与教育程度之间的相关性并未下降。当然也有教育不平等程度下降的例子,例如布林和乔森(Breen & Jonsson, 2005)的综述文章报告了在德国(Henz & Maas, 1995; Jonsson et al., 1996; Müller & Haun, 1994)、法国(Vallet, 2004)、意大利(Shavit & Westerbeek, 1998)、挪威(Lindbakk, 1998)等国家社会背景与教育程度间的相关性出现了下降。

为了描述某些国家出现的教育扩张并未减少教育不平等的现象,较早提出的模式是“最大化维持不平等”(Maximally Maintained Inequality, 简称 MMI, 参见 Raftery & Hout, 1993)。它主要包括三方面内容:一是其他条件不变,教育扩张反映的是人口的增长及其构成变化所导致的需求变化,这种情况下,教育不平等程度不会改变。二是如果扩张快于需求的增长,全体的升学率都会增长,但教育不平等程度不会改变。三是如果优势群体在某一程度的教育上达到饱和,而教育扩张不能通过其他形式完成,才会惠及弱势群体,教育不平等程度下降。作者在文中声明,这只是对若干经验事实的描述,作者进一步对这个模式给出了一个基于理性选择的解释(rational-choice explanation)。

由于 MMI 并没有考虑教育资源的差异性,在此基础上学者进一步提出了“有效维持不平等”(Effectively Maintained Inequality, 简称 EMI, 参见 Lucas, 2001)模式,认为优势群体会努力寻求优势,如果量上的差别大就寻求量上的优势,如果质上的差别大就寻求质上的优势。也就是说,如果某个层级的教育稀缺,例如大学本科,那么优势群体就努力寻求这一层级的教育;而如果大学本科对所有人不那么稀缺了,优势群体就努力寻求其中的优质教育,如精英大学;当然更理想的是二者兼得,多多益善。这一理论可以在不同水平的大学学生群体构成的分化(Brint & Karabel, 1989; Karabel, 2005; Ayalon & Shavit, 2004)现象中得到支持。

上述模式和理论解释力也有局限,不能说明某些国家的情况。有鉴于此,学者们又化用理性选择理论,提出了一些新的解释(Jonsson et al., 1996; Ballarino et al., 2009; De Graaf et al., 2000, 等等)。

然而,正如布林(Breen, 2010)所指出的,教育扩张对教育不平等的

影响取决于具体的实施政策和社会经济条件,不同国家的情况不同。欧美社会发生的教育扩张面临的社会条件与实施方式都与中国不同。中国这个以高考分数为主、优惠政策为辅、按省统一考试、省校配额招生的制度,最明显的特征在于录取的资格分数维度比较单一,而且是外显可见的。社会学家关心的社会经济地位等因素在中国皆是通过上述机制发挥作用的,简单地说就是主要“体现在分数上”。欧美社会不一定那么依赖标准化考试分数,因此社会学家关心的因素可以通过其他渠道发挥作用。例如,美国的大学招生不但关心 SAT、ACT 等标准化考试成绩,还关注学生的平时成绩、课外活动、个人陈述所展现的独特经验与品格,等等,而学生在这些方面取得的成果都在一定程度上受其社会经济背景的影响。

中国与欧美社会的第二点差异在于,中国大学扩招主要的推力是外生的政策力量,与欧美社会不尽相同。从计划经济时代延续下来的高考招生制度市场化的程度不那么高,大学招生通过配额的方式确定规模,这与西方社会不大一样。当然,通过国家政策推动教育扩张,中国绝对不是孤例,只是不同国家的实施形式不同。例如,以色列在 20 世纪 90 年代进行的教育改革也是国家主导、政策推动的 (Ayalon & Shavit, 2004), 和中国的高校扩招有一定的可比性。

本文在此提出欧美社会与中国社会的差异,并不是要倡导“中国特殊论”。恰恰相反,本文认为前述理论提出时都暗含着各自社会面对的具体条件,包括其经济人口因素、社会文化和升学制度的安排等,所以将其拿来研究中国的情况,必须更加审慎地思考这些条件是否与中国的情况相合。以 MMI 为例,本文认为实际上中国的现实条件与 MMI 并不相符,因此不会出现优势群体饱和之后才轮到弱势群体受益的局面。

MMI 提出之始,研究者希望以此建立一个研究的基线,用于未来经验研究的检验 (Raftery & Hout, 1993), 而这也正是后来不少研究所做的工作。它们把不平等维持或增加的迹象看作 MMI 模式的证据。然而本文以为,MMI 模式只是可能导致不平等维持或增加的解释之一;要理解教育扩张与教育不平等的关系,进行检验的基线和建构研究的框架在逻辑上看还有其他选择,那就是本文提出的“阈值依赖不平等”(Threshold Dependent Inequality, 简称 TDI) 理论。本文以为,先排除人行动的因素,单纯看教育扩张的效应,得到的情况宜作为基线;在

此基础之上合理地增添人的行动与政策的互动,叠加政策和社会条件,才能理解教育扩张在具体社会中的作用,才能理解现实中政策与政策改革的效果。研究不能毕其功于一役,本文完成的工作主要是第一步,同时也给出了进一步研究的设想。根据本文主体部分的分析,可以看到 MMI 是特定条件下的一种情况,并不具备普遍性。

为简化分析,本文先做了两个假设:假设分数线与人没有互动,假设分数线的移动外生于系统。在此前提下,本文的理论认为,如果两个群体存在差异,那么从分数线由最高到最低的移动,也就是教育扩张的整个过程会使得群体不平等程度先增后降。也就是说,始于高分段的教育扩张会导致不平等程度上升,对精英高校而言尤为明显,而始于低分段的教育扩张则会减少教育不平等。

如何得到这一结论呢?本文的理论与 MMI 和 EMI 的关系具体来说是怎样的呢?不妨先说明本文的研究设计,展开对经验数据的分析。

三、研究设计

本文所用数据是近十年中某年 A 省的高考卷面分数。对此数据需要加以详细介绍,以明确其优势与局限。

首先,所谓卷面分数,俗称“裸分”,指的是考试过程中考生答题、判卷者给分所形成的客观分数,在此数据中文、理科均介于 0 分到 750 分之间。这里需要说明两点:第一,我国近年来一般不对高考分数加以非线性处理,不像美国的 SAT 或 ACT 等考试那样进行标准化变换。第二,卷面分不含我国各院校在招生过程中给予考生的优惠政策加分。诸如少数民族、竞赛表彰、大学自主招生等因素对实际招生结果有非常大的作用,但并不能在这一数据中得到体现。

其次,本文所用的数据并非 A 省该年全部考生的数据。出于保密的需要,在此不便透露具体数字,但本文所用数据的规模大概为 A 省总报考人数的九成左右。未包含其中的两类学生,一是小语种考生,也就是外语不考英语而考俄语、日语、德语、法语、西班牙语等语言的考生;二是以少数民族语言答卷的考生,他们不使用统一的语文试卷,而是使用把汉语作为第二语言来考察的汉语言试卷,其他学科试卷命题也翻译成少数民族语言。了解高考的人都知道,这两类考生的分数与其他

考生很可能有系统性差异,但在招生时都是有效分数。所以说,本文所用数据缺少小语种考生和部分少数民族考生的信息,有一定的偏差。

最后,这一数据只有考生地区和分数的变量,而没有任何个人属性、家庭背景以及实际录取结果的变量。这意味着很多社会学家关心的因素,例如性别、家庭背景、民族身份等无法考察。但由于根据数据可以推知地级市及其下属区县一级行政区划的信息,本文可以近似地考察城区与非城区学生的差异。此外,由于没有录取结果,本文只能使用公开报道中的大学录取分数线,假定超过某一分数线的人一定得到录取,而实际情况当然不是如此。^① 高考先有名额,后有录取,录取完成才知分数线,所以这种近似虽然合理,但毕竟存在缺陷。

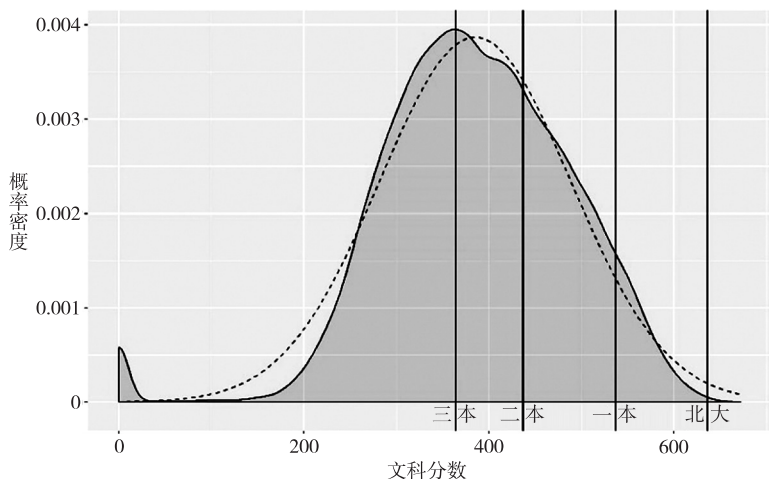
上文说明了数据的缺陷,但与优点两相权衡,本文认为优点更为突出。首先,由于数据规模较大,可达十万的级别,尤其是在右尾上几乎完整的数据让我们对研究的发现更有信心。其次,作为省级数据,本文可谓是“解剖麻雀”的方法,该省并非直辖市,其工农产业并不存在非常独特的情况,对于全国多数省份的情况有一定的代表性。由于招生名额的分配和操作都在一省之内进行,全国范围的抽样调查得到的笼统结果其实对于任何一个面对高考的学生来说都不够精确,因为其中隐含的各省分数分布、大学名额差异的权重未加控制。更重要的是,如若本文的研究结果能促进我国掌握大量数据的有关机构加强与学界合作,或自己更细致地展开分析,有益于未来公共政策的设计与实施,则善莫大焉。

回到本文的研究设计上,用一省一年的截面数据推论教育扩张和教育不平等升降这两个变化过程的关系,应如何进行? 本文进行了一系列简单的计算,以模拟考分不变而分数线移动对录取结果的影响。

本文对自变量采取了近似的办法,把县级市、区并为城区,县视为非城区,将样本划分为两个群体,城区与非城区考生比例在 2:1 左右。这种做法可比拟于以学籍所在地划分城乡,类似梁晨等(2012)的办法。这种做法当然不够准确,既然不以户口为界,每个群体内部都可能

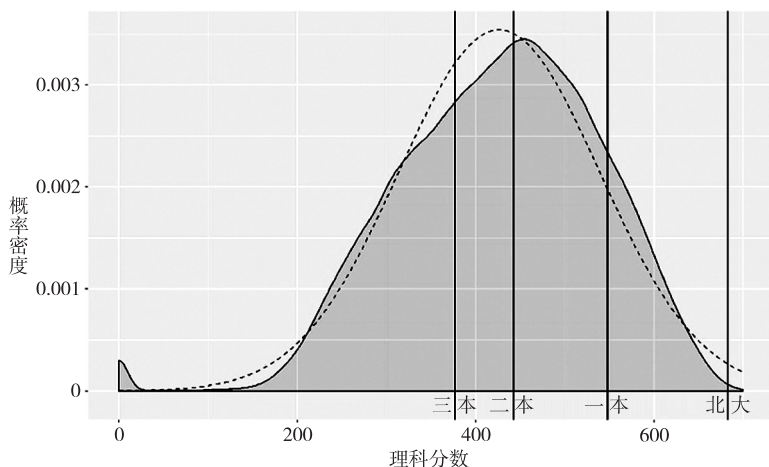
① 以该年理科成绩为例,按照本文的方式估计的高于 A 省三本线的考生人数为 7 万多,而该年计划录取人数为近 9 万,前者是后者的 83%。前文已经说过本文所用的数据中不包含小语种考生和采用民族语言答卷的考生,也许两个数字的差异可以由此得到部分解释。如果实际上高于分数线的考生不一定得到录取,那么本文估计的人数应该高于实际录取人数,而目前的状况恰恰相反,这说明本文的近似方法与实际情况相差可能不大。

包含城乡居民,因此其结果不能直接与以往研究进行比较。但是这一方面是数据所限,另一方面,鉴于城乡教育和地区教育都存在差异,以后者为例举一反三,对辨明前者的机制有很大价值。



注:深色阴影为城区,浅色阴影为非城区。

图1 A省某年文科高考考生分数分布

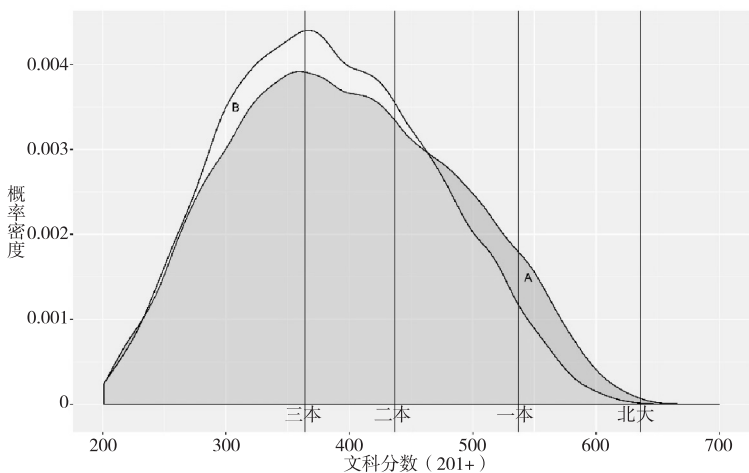


注:深色阴影为城区,浅色阴影为非城区。

图2 A省某年理科高考考生分数分布

在此基础之上,首先分文、理科绘制全体学生分数的概率密度分布,并叠加以该年 A 省一本、二本、三本大学文、理科录取分数线。为了进一步考察精英大学的情况,在图中又加上了北京大学的文、理科录取分数线(见图 1、图 2)。由于概率密度曲线下方面积或者说其定积分就是比例,那么处于分数线右侧、曲线下方的面积也就是升学率,而这个面积与分数线左侧面积之比,就是该群体升学的比数(odds)。可以看到,数据有很多 0 值,也就是完全白卷或考试作弊者,后续分析中不妨舍去;而且,0 值之外两个分布的形状都类似虚线所勾勒的正态曲线。我们还注意到,北大、一本、二本、三本大学的录取分数线处于分布的不同位置,因此有不同的意涵。这一点对后文的理论阐述有非常重要的意义。

这样,把两个群体的分数线都描画出来,不就可以看到不同群体升学比例的差距了吗?本文制图时截取了 201 分及以上的结果(见图 3、图 4),^①深色阴影者为城区,浅淡者为非城区。

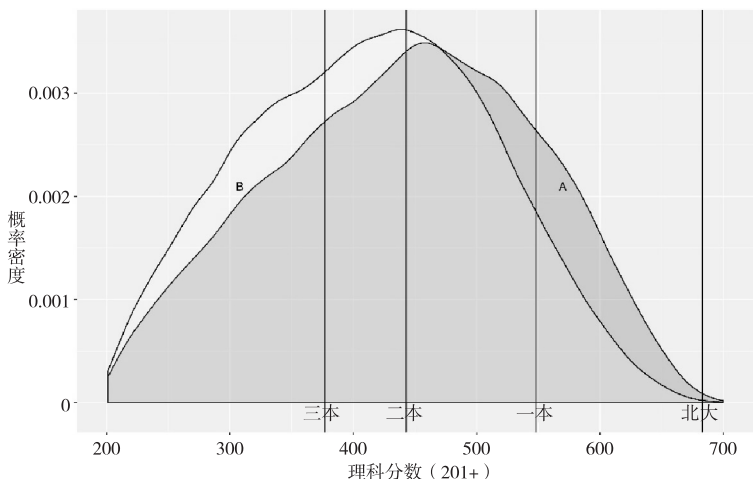


注:深色阴影为城区,浅色阴影为非城区。

图 3 A 省某年高考文科城区与非城区考生分数分布

如果我们的视角微观到每一分的层次上,那么对于横轴上的每一个具体分数,其对应的曲线高度便可以近似地理解成取得这个分数的

^① 虽然制图做了处理,但计算时并没有舍去除前面提及的 0 分考生之外的低分。以下皆同。



注：深色阴影为城区，浅色阴影为非城区。

图4 A省某年高考理科城区与非城区考生分数分布

城区学生和非城区学生各自在城区与非城区学生总数中所占的比例。曲线高者说明比例较高，也就是占有优势。所以，两条曲线的交点，也就是微观层面上城、乡录取多寡的相互关系发生转变的点，在交点右侧城区学生录取比例较高，在交点左侧非城区学生录取比例较高。

如果我们的视角转换为累积所得的学生升学率，也就是概率密度分布曲线下面、录取分数线右边的积分面积，那么A部分，也就是右方城区的概率密度分布高于非城区的面积，就是教育不平等拉大的绝对值；而B部分，左方非城区概率密度分布高于城区的面积，就是教育不平等缩减的绝对值。本文着眼于累积所得的升学率的层面，因此下面的阐述也在这个层面展开。

四、教育扩张与教育不平等

图3与图4虽然直观，但因为是静态的，并不能说明不平等变化的过程。教育扩张——也就是一省文理科招生名额之增加，体现在此框架下也就是录取分数线自右向左的移动。那么一分一分穷举式地移动分数线，不就可以计算并观察分数线右侧面积的变化——也就是教育

机会不平等的变化了吗?

为了简化分析,上述观察需要两个假设:假设分数线与人没有互动,假设分数线的移动外生于系统。也就是说这两张图中的概率密度分布不会因分数线的移动而改变。这当然不是现实情况,但也是比较合理的。为什么这么说呢?因为前面提到中国的教育扩张是政府主导的,相对而言,扩招带来分数线的移动的确更类似一个外生的冲击。此外,高中三年苦读之后才参加考试,而大学录取的名额是高考之前不久才确定的。很难想象有大批学生会因招生名额发生变化而临时决定不参加高考,或者未读高中的人前来参加考试。因此,招生名额变化与选择是否就学的决策有三年之差,而如果说在上高中之前因为预计了三年后招生名额的变化而改变上学与否的决策,恐怕也不是实情。最极端的例子是在开始扩招的第一年,受惠的学生已经历高中三年的学习,其数量和学习成绩早已确定了,不会因为高校扩招而发生太大的改变,不管名额增加导致分数线如何移动,高考成绩分布都应该不大会受影响。因此,从短期看扩招可以看作一个外生冲击,与高考成绩分布彼此独立,这是一个比较合理的假设。

每给定一个分数线,文理科概率密度分布曲线就分别对应着两个右侧的面积,也就是在文科、理科生中模拟的城区与非城区学生各自的升学率。通过罗列升学率得到经验累积分布函数(Empirical Cumulated Distribution Function, 简称 ECDF),叠加之后两相对比可以更加清楚地展示升学率差异的变化过程(见图5、图6)。^①

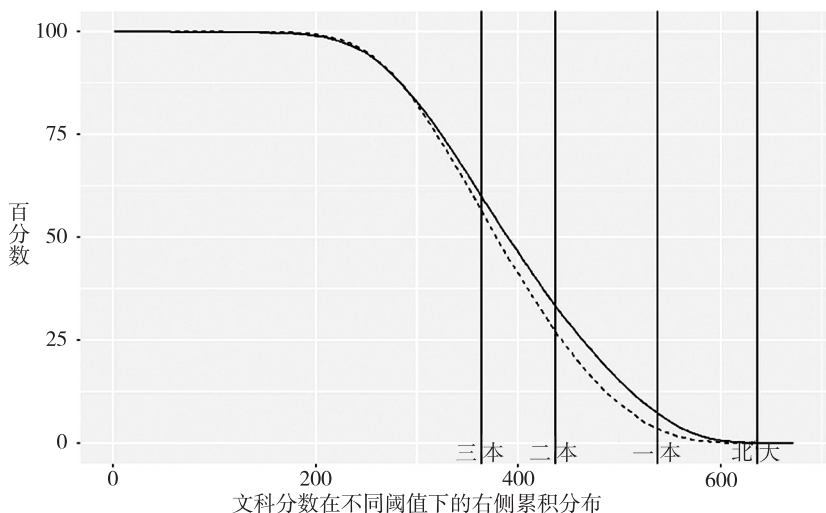
根据一般报告教育不平等的惯例,我们可以想象至少三种基本的报告差异变化的形式。首先是百分比的差值的变化,也就是“城区升学率比非城区高若干百分点”的形式。这可以通过简单地对两条经验累积分布函数作差得到(见图7、图8)。为了帮我们看清右尾处的情况,这里截取了301分及以上的部分,下同。

其次,另一种常见的报告形式是升学率的倍数变化,也就是“城区升学率是非城区的若干倍(百分之多少)”的形式。这可以通过简单地以两条经验累积分布函数求商得到(见图9、图10)。

最后,因为处理这类问题的一个常见模型是logistic回归,模型系数直接报告的是比数比(odds ratio)的对数,我们也可以利用经验累积分布

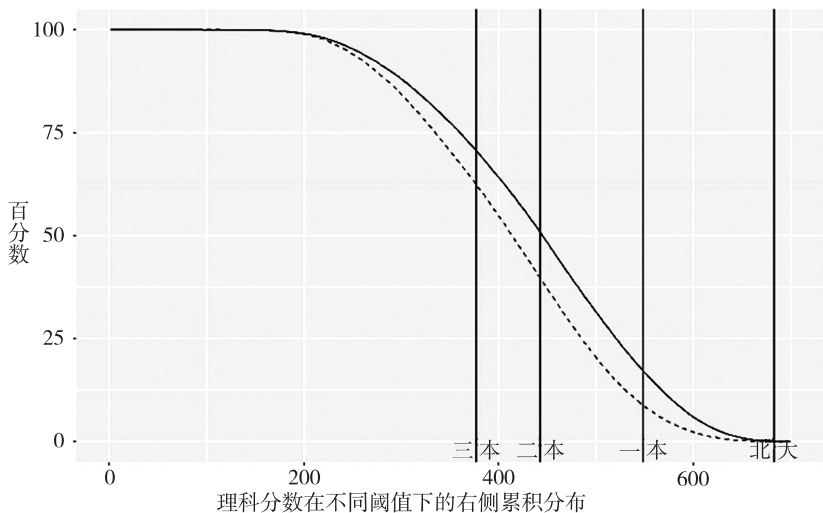
^① 实际上这里报告的是 $1 - \text{ecdf}(x)$, 因为直观报告的是分数线阈值右侧的面积。

进行计算并加以报告,也就是先分别计算两条函数某处取值与其补数的比得到两个比数,接着求商得到比数比,再取对数(结果见图 11、图 12)。



注:城区为实线,非城区为虚线。

图 5 A 省某年高考文科不同分数线上城区与非城区升学率的变化



注:城区为实线,非城区为虚线。

图 6 A 省某年高考理科不同分数线上城区与非城区升学率的变化

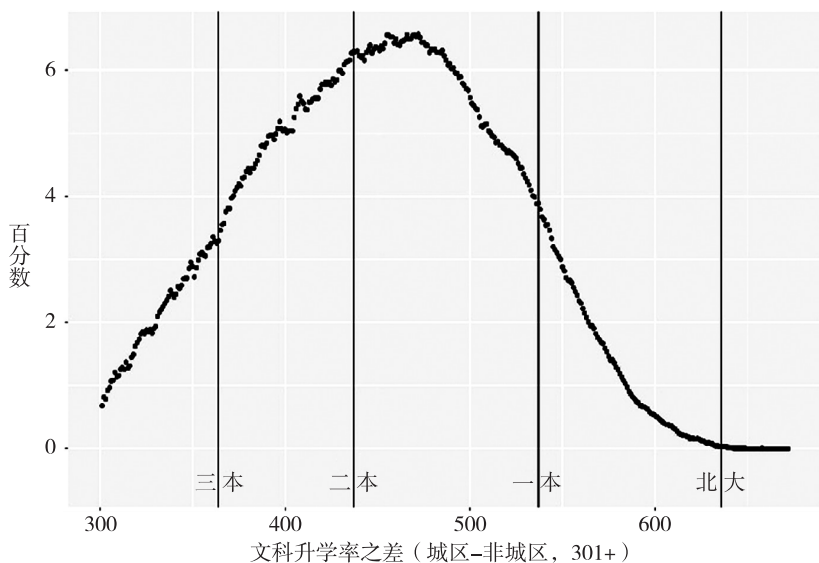


图7 A省某年高考文科不同分数线上升学率之差的变化

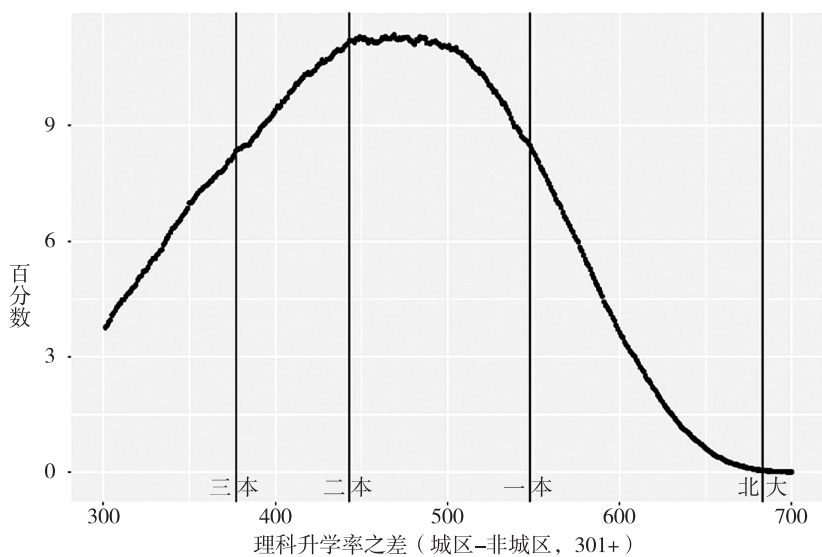


图8 A省某年高考理科不同分数线上升学率之差的变化

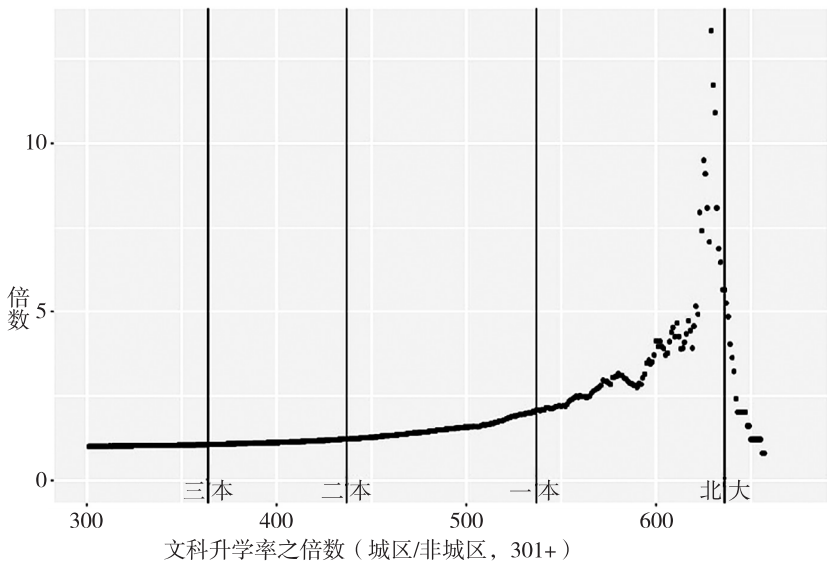


图9 某年A省高考文科不同分数线上升学率之倍数的变化

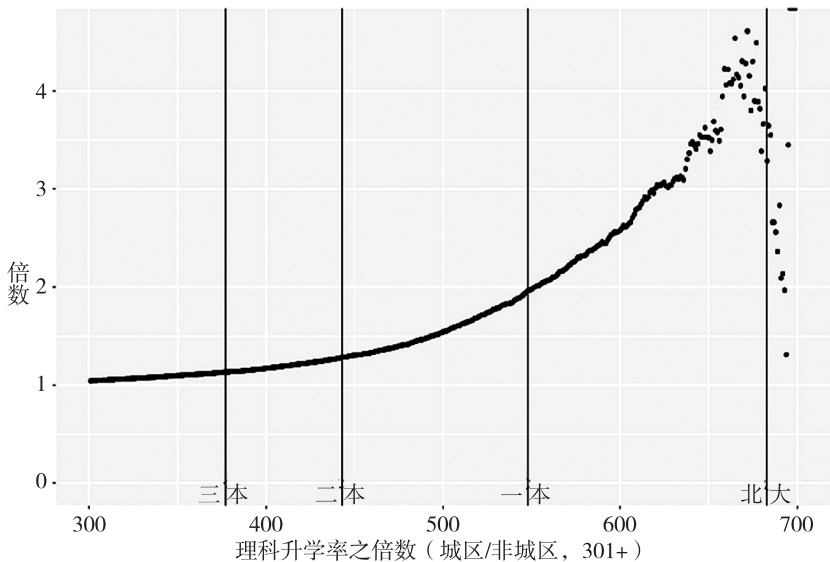


图10 某年A省高考理科不同分数线上升学率之倍数的变化

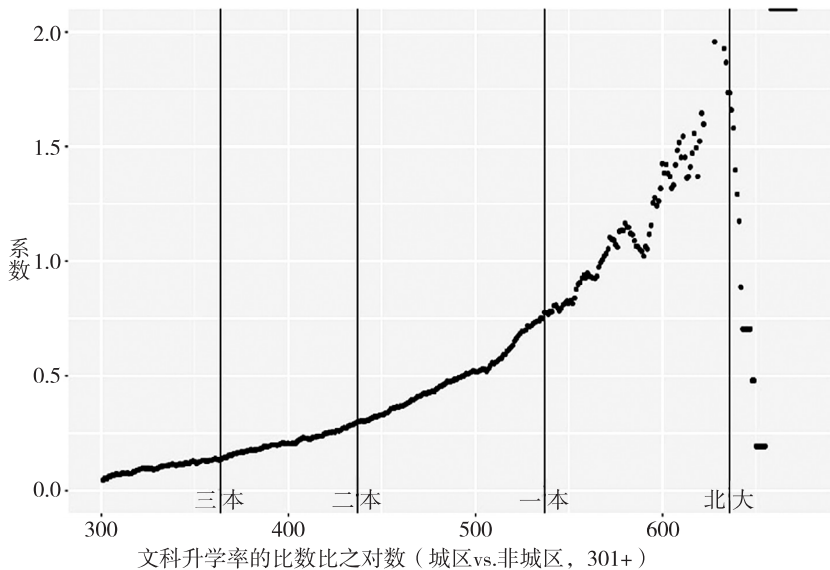


图 11 A 省某年高考文科不同分数线上升学率之比数比对数的变化

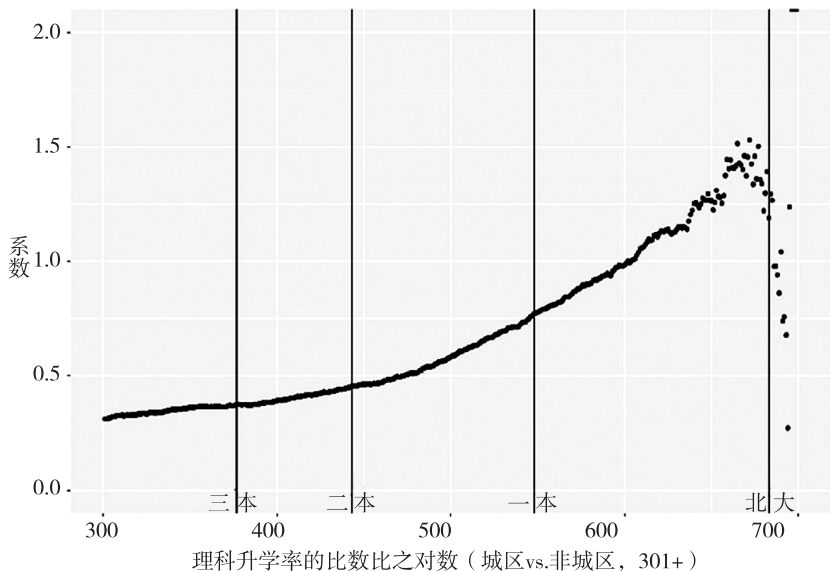


图 12 A 省某年高考文科不同分数线上升学率之比数比对数的变化

这三种不同的报告形式,背后的内容是相同的,只是不同的形式有不同的阐释。

首先,可以看到,在按照我们近似的方法划分的城区与非城区之间,高考分数的分布均类似正态分布,但彼此又有一定的差异(见图3、图4)。这一差异主要体现在均值上,方差的差异相对较小。可以想见,如以社会学家关心的其他变量为标准来划分群体,这种分布类似正态、均值不等而方差相近的情况应是常见的。例如,性别之间可能有差异,城乡之间可能有差异,不同民族之间可能有差异,等等。这里绝无孰优孰劣的判断,这种差异的存在只是一种客观事实,必须具体情况具体分析。

其次,我们看到文理科的分布存在差异,这一差异不但体现为分数的概率密度曲线峰值位置有别,还尤其体现在两条概率密度分布曲线的交点与其各自峰值的关系上(见图3、图4)。但是二者又有共性,体现在交点与各类本科院校录取分数线的关系上。如果考虑差值,在交点右侧城区的分布曲线高于非城区,因此城乡不平等是拉大的,而左侧正好相反。我们看到二本与三本分数线在交点左侧,而一本和北大分数线在其右侧。假如扩招之后四条分数线都在交点右侧,那就说明对于一本大学包括北大,扩招到现在教育的不平等程度还是在拉大,而事实上二本大学和三本大学的教育不平等程度在拉大之后已经过了峰值,现在开始降低了。^①

再次,如果我们考虑的不是差值,而是倍数或者比数比的关系,那么能看到什么呢?从图9、图10可见,一、二、三本分数线都在关系发生转变的拐点以左,因此对应的教育不平等下降了,继续扩招会进一步降低教育不平等。精英大学处于分布的极右端,情况很不连续,而且其大趋势是扩招导致的教育不平等还要拉大。这说明以仅采用精英大学的数据,例如北京大学和苏州大学学生的数据(梁晨等,2012)或者“首都大学生成长追踪调查”(吴晓刚,2016)这样的数据所做的研究来跟采用大范围抽样数据的研究结果进行比较需要格外注意。^②

① 当然三本的学生应是二本线与三本线之间的面积,二本的学生应是二本线与一本线之间的面积,这里为了行文方便有所省略。

② 由于是我们采用的几乎是总体数据,所以不存在一个“样本量过小”的问题。现实生活中高分段分数上人数的分布就是比较稀疏的,并不是在每一个分数上都有不少学生,有些分数对应的可能仅是一两个学生,甚至没有学生。

这里非常值得注意的是,用差、倍数还是比数比来衡量教育不平等,教育扩张与教育不平等关系由增转衰的拐点所在位置不一,这说明即便背后的现象完全一样,采用不同的报告形式也可能表述得完全不同。选择哪一个拐点、哪一个趋势进行报告,取决于研究者想探究的问题和数据的条件。报告差和倍数的描述统计方法在一般新闻报道和政府报告中比较多见,而报告比数比的模型化研究在社会学的学术研究中比较多见。前者更加直观易懂,后者要更加严谨可靠,不同的工具各有利弊。如果对测量不平等所用的工具不加注意,是不是会出现背后的社会事实一样而只是因算法不同导致结论分歧呢?本文对三种测量方法得到的结果都进行了报告并且比较了差别,就是希望能在这一点上有一个清晰的呈现。

最后,前文讲的实际上是教育不平等随着分数线的左移而变化的连续趋势。^①但是既然分数线的变动定然表现为一个个截面上不连续的跳动,而不论用差、倍数还是比数比的 \log 来描画不平等,抽象而得的一般变化趋势都有一个比照基点,那么在实际的经验研究中也必须对自己的参照点在哪里非常小心。下面我们举几个例子来加以说明。

假如只能找到相隔10年的两个截面数据,那么若不小心使用,很有可能两个截面背后分数的分布不变,但各自处于不平等峰值的两侧。如果只进行这样简单的线性描述,其结论当然不能说明问题。即便两个数据处于峰值的同一侧亦须留意。以直观的差来说,假设某两国分数分布完全一致,升学机制也完全一样,在某国作为比照基点的截面数据对应的位置相当于这里的二本线,那么教育扩张导致其移动到三本线的位置,我们会观察到教育不平等下降;而另一国基点位置在一本线的位置,教育扩张导致其移动到二本线的位置,我们会观察到教育不平等上升。这样看上去似乎有了国别差异,若误以不同理论加以阐释,岂不成了缘木求鱼?经验研究当对此加以注意。

同样,这个道理也说明,谈论教育扩张时一定要明确指的是哪一层次的教育,是高等教育还是义务教育,是高质量的高等教育还是低质量的高等教育,等等。对于近乎普及的义务教育面对的基点,其分数线合理的估计位置应在峰值以左,那么我推测教育扩张从一开始就使教育

① 如果用数学的语言来打比方,也就是不平等函数的一阶导数的情况。

平等程度有所提升。高质量的高等教育很可能是稀缺的,分数线位置在峰值以右,那么教育扩张导致不平等增大也就不难理解了。

五、进一步研究的设想

至此本文的主体分析已基本完成,也得到了相应的结论。正如前文所说,这是一个比较简化的框架,因为我们做了两个假设,即完全排除分数线的移动对社会经济和人的行动的影响,假设分数线的移动完全是外生因素所致。这一方面是由于数据所限,如果能获得更多变量和更多年份的数据,情况当有所改善;另一方面也是为了分析问题的起点清晰简洁,为后续研究提供比较的基线。

正如前文所说,这种假设从短期来看是比较合理的,而且我们模拟的情况正体现了教育扩张对教育机会分配最直接的效果。

但是从长期来看,教育扩张会从多个方面间接地影响人的行动,分数线的移动也就不能说是完全外生的。例如,媒体上有一种意见认为高等教育的快速扩招导致高等教育学历在劳动力市场上贬值,反而引起人们对“大学生找不到工作”的担忧,于是不少学生干脆不读高中而直接去打工。如果这种情况确实成立,对高考分数分布当然就会产生影响。我们还要看到,这种间接效应往往是众多因素共同作用的结果,教育扩张只是其中一环,将其总效应全归结其上,恐怕言过其实了。采取常见的社会调查数据进行的研究,往往无法区分教育扩张的直接效应和间接效应,而这正是本文提供的框架可能有所贡献的地方。在下一步的工作中,运用更丰富翔实的数据,逐步放松本研究提出的假设,可能会让我们对教育扩张的政策效果有更清晰的认识。

此外,不少经验研究受数据限制,也无法有效地从总效应中甄别出其他会对教育不平等产生影响的因素。我国力行教育扩张的政策正发生在转型期社会大变革的背景之下,影响人们就学意愿、就学机会和就学能力的因素很多,如果只比较扩张政策执行前后不同群体教育机会的分配多寡,恐怕失之笼统。本文提供的框架可以帮助研究者结构化地理解繁杂因素的作用。总的来讲,前面提供的框架告诉我们,教育不平等发生变化无非三种情况:一是录取人数与总量关系的变化。简单而言,就是由于存在群体分布的差异,假如报考人数不变而录取人数变

化,分数线便会发生移动,从而改变教育不平等的状况。这也正是本文主要分析的情况。二是群体规模相对关系的变化。如要保持录取人数与总量的关系不变,优势群体相对增加,则分数线向右移动,反之则向左移动,导致教育不平等状况发生相应变化。理解这一点很容易,想象一下如果优势群体从一万人减少到一百人,而录取的总人数不变,这个缺口肯定需要从劣势群体中最靠前的部分补取,也就意味着分数线要向左移动。^①三是分数分布的变化。假如分数分布在不同群体之间发生了变化,不论是均值差、方差差异还是分布形状的其他变化,都有可能给给定分数线下不平等关系的改变。

总之,导致教育机会不平等发生变化的机制在逻辑上必定通过这三种渠道发挥作用。其中只有第一个渠道可以确定是教育扩张的效果。当然,需要重申,从长期来看教育扩张可能会通过与其他因素的互动间接导致后两个方面的变化,这就需要以进一步的数据和研究来进行探讨。

六、阈值依赖不平等

本文用以说明教育扩张与教育不平等关系的研究框架可以称作“阈值依赖不平等”理论。概括来说,假如候选人的资格超过某一遴选的阈值才能获得教育机会,那么遴选的阈值始于高分段的教育扩张,则表现为教育不平等程度的上升;始于低分段的扩张,则表现为教育不平等程度的下降。也就是说,教育扩张与教育不平等的正向或负向的关系取决于升学之时遴选的阈值是多少,所以本文有了“阈值依赖”这个表达。在中国高考制度的语境下,“遴选的阈值”就是指大学录取分数线。

首先需要说明,教育扩张与不平等关系发生转变的这一拐点阈值

① 通过积分计算也容易说明这一点,因为优势群体的概率密度分布曲线不变,其积分面积所对应的实际人数减少,则劣势群体积分面积对应的实际人数增多。设 a 、 b 分别表示分数线右侧优势、劣势群体的积分面积, X 、 Y 分别表示优势、劣势群体的总人数,则录取人数 $n = aX + bY$ 。若优势群体减少 d 人,由于总人数不变,此时劣势群体增加 d 人,于是如果分数线不变,录取人数 $m = a(X - d) + b(Y + d) = n - (a - b)d < n$,所以要维持录取人数与总人数的关系,分数线需向左移。

在何处,无法通过先验的计算得出。前面的分析已经表明,这取决于两个群体分数分布的具体关系和形状,自然无法以函数来计算得出。但是本文的分析提供了一个寻找拐点阈值的通用步骤,因此它不是不可知的。

其次还要说明,尽管大趋势是一致的,但是用不同方式计算不平等,尽管背后的分布可能一致,得出的结果可能不大相同。因此拐点阈值在哪里,本身取决于研究者想采用什么方式来计算不平等。

尽管如此,大体的趋势是清楚的,对于以差衡量的不平等,拐点阈值在两条分布曲线的交点;对于以倍数或比数比来计算的不平等,拐点阈值在相当靠近右尾的位置。

那么 TDI 与前面提到的 MMI 和 EMI 之间是什么关系呢?现在可以进行简单的说明了。MMI 可以说是 TDI 框架在特定条件下的一种情况。

MMI 认为优势群体会尽可能满足自己的需求,在需求饱和之后教育扩张才可能开始惠及弱势群体。这里所谓“需求饱和”是一个定性的概念,并没有一定的数值,但至少意味着一个较大的百分比。假如给出一个拇指规则,令“需求饱和”的程度为“80%可以升学”,那么在我们的框架中理解这个问题,不就是说两个群体的资格分布相距之远,分数线需要向左移动,直到满足了优势群体 80% 的人之后才会发生关系的转变吗?至少在我们定义的城乡地区之间并没有这么大的差异,经验研究常用的变量(例如户口、性别、族裔、父辈状况等)界定出来的群体之间,会有这么大的认知能力与教育资源的差异吗?这需要今后研究的进一步考察。当然,如果改变区分优劣势群体的标准,有可能达到群体分数分布剧变的效果,这也恰恰说明了 MMI 不具有普遍性,而是有其特定的社会条件。

总之,依据 MMI 而认为中国教育不平等会扩大到一个较高的程度,甚至达到优势群体的“饱和”,可能不太符合实际。当年提出 MMI 的研究(Raftery & Hout, 1993)主要关心的是 20 世纪的英国与爱尔兰社会,可能其社会阶层的分化更加明显,而且不像中国这样主要根据标准化考试分数来定夺录取,因此资格分布差别大是有可能的。

另外,MMI 并不局限于高等教育入学这一个关口。卢卡斯(Lucas, 2001)把 MMI 的研究放在教育过渡(education transition)的议题的脉络中做了比较好的总结,并指出了教育过渡研究的一个重要发

现:随着教育阶段的提高,社会背景因素对升学的影响逐渐下降。在中国,不少学者同样发现,随着教育阶段的提升,城乡等因素对升学率的影响幅度下降,中国初升高的城乡差异大于高升本(李春玲,2014),与世界上其他国家情况类似(例如 Garnier & Raffalovich, 1984; Shavit & Blossfeld, 1993, 等等)。对此的一个合理解释是弱势群体中有更多的人被前面的遴选剔除出去,或者较早地选择不继续升学。因此当本文研究报名参加高考的人群时,优劣势群体的差异已经比较小了。如果把这个次第遴选的过程纳入 MMI 的含义当中,MMI 命题的含义就发生了改变,需将限定怎样的人群作为总体的讨论纳入其中。

本文把视野限定在高考升学这一关口,认为要探讨教育不平等需要明确限定的范围,例如讨论高考升学就应该只选取至少完成了高中教育的样本,而不应把所有样本汇聚到一起。限定人群不同,实际上本身也是研究者要阐释的机制的不同。在 TDI 的框架下,改变限定的人群,实际上是纳入了前面所提到的第二渠道的影响,也就是群体规模的相对关系对教育不平等的影响。举例来说,如果不把高考考生作为限定的人群,而是把全体同龄人口作为限定的人群,这就相当于在现有的分数分布左端增加极大比例的 0 值分布,对于计算获取教育的机会当然就有很大影响。如果不同群体的同龄人口中高考考生所占比例不一或年年有别,当然可以在很大程度上影响观察到的获取高等教育机会差异的分布。在这样的限定下,我们看到的教育机会的变化就还包含着人们是不是想获取高等教育的决策,而不只是教育扩张的直接结果。^①

总之,本文认为 MMI 内在的逻辑实质是 TDI,MMI 是 TDI 在特定条件下的一种情况。在中国的条件下,由于主要依照高考分数进行选拔,而群体之间的分数没有那么大的差距,本文相信应不至于出现优势群体需求饱和时教育扩张方才惠及劣势群体的局面,拐点阈值会出现得更早。因此,即使观察到教育不平等扩大的情况,也不能验证 MMI。

至于 EMI 理论,前面的分析很容易说明不同水平的大学学生群体出现分化现象的原因。观察四条分数线所在位置,进入精英大学者为精英分数线右侧之面积,进入一本非精英大学者为精英与一本线之间

^① 在此我并非想割裂二者之间的关系,上不上高中的决定和扩招之间当然可能有一定关系,前面也已经提到。只是在短期来看,这并非扩招的直接效应;长期来看,扩张政策对上高中的决定影响有多大目前也难以确定。

的面积,如此类推,则很容易看出为什么门槛高的精英大学与门槛低的其他大学日渐分化,或者门槛高的高等教育与门槛低的成人教育日渐分化,无非是在两条门槛间的那部分分布的差异使然。由此便可解释中国存在的教育分流现象(吴愈晓,2013a)。简而言之,EMI是TDI的一个方面。

七、结论与讨论

最后,回归到中国的教育扩张与教育不平等上,我们能做出何种推想呢?

在中国,录取阈值的变化始于高分段的教育扩张,直接效应是会先拉大教育不平等的程度,在经过一个拐点之后开始减小教育不平等的程度。而录取阈值的变化始于中低分段,教育扩张的直接效应是会减小教育不平等的程度。这一结果发生在完全客观的机制之下,完全依赖于标准化考试的分数,不平等的变化只是分数分布的客观结果。

正如前文提到的,这一观点主要说的是教育扩张的直接影响。但是长期来看,由于教育扩张通过第二与第三渠道对教育不平等产生间接影响,在经验研究中观察到的总效应可能会有不一样的答案。正是在这样的张力中,研究者得以不断地深入地挖掘影响教育不平等的复杂机制。间接效应在多大程度上可以归因于教育扩张,有多大程度是即使没有发生教育扩张也会发生的,可能可以通过进一步的考察来加以探讨。

本文当然也存在明显的缺陷与不足,可以在以后的研究中加以改进。本文对机制的解释高度抽象,对发生在高考制度中的种种变化,例如填报志愿时机与方法的改变、针对各类群体和表彰人群的优惠政策、部分高校推行的自主招生和校长实名推荐制等微观的机制完全没有涉及,而人们担心的往往正是学生家庭背景等因素可能会通过这些机制对不同群体产生不同的影响。本文无法对此做出有效的回应,一个主要的原因是数据年份有限、变量有限,对这些重要的问题难以进行考察。

本文一直在强调数据的不足,但是这种数据不足的性质不是有价值的无法收集或者收集的成本很高,而是在现有的高考报名制度体系之下,行政管理部门可能已经收集了相关的数据,只是未加整合与

公开。例如,由于中国存在学籍制度,考生的个人与家庭情况实际上可以和高考成绩、录取结果关联上;小语种考生与民族语言答卷考生的成绩也是已知的,只是没有和本文所用的数据放在一起;由于录取是一个统一进行的机械过程,各个高校的录取人数、录取分数实际上都是已知的,如果整合起来可以解决很多本文提到的不精确问题。^①即便是现在还没有收集的数据,如果价值较高,进行收集的成本也不是那么不可逾越,毕竟在报名高考时让考生额外填写一些信息并不是非常困难的事。总之,本文对于改进目前数据的缺陷是比较乐观的。与很多常见的大范围社会抽样调查不同的是,这里并不存在方法上或者成本上无法克服的问题。在我国政府政务电子化、透明化的大趋势之下,只要行政管理部门有意愿,完全可以使相关议题的研究可用的数据质量得到大幅提升。

本文的第二个疑问在于:本文的分析限于一省一年,将其推广至一般化的可能性有多大?

首先,为了确保使用一年之数据得到的结果不是偶然的,本文利用该省前一年的高考数据重复了这里所做的研究,所得结果大体一致。这说明本文描述的短期效应并非某一年的特例,而是符合常情的。

其次,一般化的第二个问题在于,本文用以划分群体的自变量其实是一种近似,与此前社会学研究常见的社会人口变量大异其趣,因此无法直接对话。但是这一研究框架是可取的。假如我们承认不同群体之间存在优势与劣势的差异,假如我们承认所衡量的属性分布形状类似正态分布、均值不等而方差相仿,那么本文得到的结论大概也适用。而正态分布不正是大多数参数模型的前提假设之一吗?举一反三,根据本研究来推想其他变量的情况,也是合理的。

第三,高考分数不易得,社会人口变量易得。本文的方法对于没有高考分数的研究是否没有借鉴价值了呢?实不尽然。可以假设,以往经验研究关注的社会人口变量与分数之间存在着统计上的关系。对于没有高考分数这样单维度数据的情况,是不是可以利用合适层次、足够数量的录取数据,采用倾向值分析之类的方法,进行某种不精确的近似分析呢?如此不也可推展到其他国家的研究中去吗?

① 例如,不使用公开报道的一、二、三本录取分数线,因为作为极值,它们可能受极端值波动的影响。采取各自录取学生群体5%分位数的结果作为分数线可以使分析更加稳健。

行文至此,该如何评估我国的教育扩张政策呢?本文的分析尚不完善,这一问题本不宜回答,但有两点意见需要加以澄清。首先,教育平等绝非教育扩张的惟一目标。本文已经说明,对于稀缺的教育资源,要扩张就难免导致教育不平等,难道因噎废食,就不提高教育资源的供给了吗?是否应把结果平等当作惟一的追求目标,还需三思。其次,因为教育不平等程度随着教育扩张有升有降,所以不能抱着一个线性的假设,换句话说就是不可忽略比较的基点在哪里。我国的教育扩张发生于高等教育资源非常稀缺的历史条件下,可能是先造成了扩大教育不平等的结果。其他国家的教育扩张在某一历史时期可能发生在分数线已在峰值左侧的条件下,教育扩张促进教育平等亦是合情合理。两者背后机制即便全然一致,其表现出来的结果也可能截然不同。忽略这一点而简单加以比较,恐怕并不妥当。

参考文献:

- 李春玲,2003,《社会政治变迁与教育机会不平等》,《中国社会科学》第3期。
- ,2010,《高等教育扩张与教育机会不平等:高校扩招的平等化效应考查》,《社会学研究》第3期。
- ,2014,《教育不平等的年代变化趋势:对城乡教育机会不平等的再考察》,《社会学研究》第2期。
- 李煜,2006,《制度变迁与教育不平等的产生机制——中国城市子女的教育获得(1966-2003)》,《中国社会科学》第4期。
- 梁晨、李中清、张浩、李兰、阮丹青、康文林、杨善华,2012,《无声的革命:北京大学与苏州大学学生社会来源研究(1952-2002)》,《中国社会科学》第1期。
- 刘精明,2006,《高等教育扩展与入学机会差异:1978-2003》,《社会》第3期。
- ,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。
- 吴晓刚,2016,《中国当代的高等教育、精英形成与社会分层——来自“首都大学生成长追踪调查”的初步发现》,《社会》第3期。
- 吴愈晓,2013a,《中国城乡居民的教育机会不平等及其演变(1978-2008)》,《中国社会科学》第3期。
- ,2013b,《教育分流体制与中国的教育分层(1978-2008)》,《社会学研究》第4期。
- Ayalon, Hanna & Yossi Shavit 2004, "Educational Reforms and Inequalities in Israel: The MMI Hypothesis Revisited." *Sociology of Education* 77.
- Ballarino, Gabriele, Fabrizio Bernardi, Miguel Requena & Hans Schadee 2009, "Persistent Inequalities? Expansion of Education and Class Inequality in Italy and Spain." *European Sociological Review* 25.
- Breen, Richard 2010, "Educational Expansion and Social Mobility in the 20th Century." *Social Forces* 89.

- Breen, Richard & Jan O. Jonsson 2005, "Inequality of Opportunity in Comparative Perspective: Recent Research in Educational Attainment and Social Mobility." *Annual Review of Sociology* 31.
- Brint, Steven & Jerome Karabel 1989, *The Diverted Dream: Community Colleges and the Promise of Educational Opportunity in America, 1900 - 1985*. New York: Oxford University Press.
- De Graaf, Nan Dirk, Paul M. De Graaf & Gerbert Kraaykamp 2000, "Parental Cultural Capital and Educational Attainment in the Netherlands: A Refinement of the Cultural Capital Perspective." *Sociology of Education* 73.
- Garnier, Maurice A. & Lawrence E. Raffalovich 1984, "The Evolution of Equality of Educational Opportunities in France." *Sociology of Education* 57.
- Henz, Ursula & Ineke Maas 1995, "Chancengleichheit durch die Bildungsexpansion." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 47.
- Jonsson, Jan O., Colin Mills & Walter Müller 1996, "A Half Century of Increasing Educational Openness? Social Class, Gender and Educational Attainment in Sweden, Germany and Britain." In Robert Erikson & Jan O. Jonsson (eds.), *Can Education Be Equalized? The Swedish Case in Comparative Perspective*. Boulder: Westview.
- Karabel, Jerome 2005, *The Chosen: The Hidden History of Admission and Exclusion at Harvard, Yale, and Princeton*. Boston, MA: Houghton Mifflin Company.
- Lindbekk, Tore 1998, "The Education Backlash Hypothesis: The Norwegian Experience, 1960 - 92." *Acta Sociologica* 41.
- Lucas, Samuel R. 2001, "Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects." *American Journal of Sociology* 106.
- Müller, Walter & Dietmar Haun 1994, "Bildungsungleichheit im sozialen Wandel." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 46.
- Raftery, Adrian E. & Michael Hout 1993, "Maximally Maintained Inequality: Expansion, Reform and Opportunity in Irish Education, 1921 - 1975." *Sociology of Education* 66.
- Shavit, Yossi & Hans-Peter Blossfeld (eds.) 1993, *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Boulder, Colo.: Westview Press.
- Shavit, Yossi & Karin Westerbeek 1998, "Educational Stratification in Italy: Reforms, Expansion, and Equality of Opportunity." *European Sociological Review* 14.
- Treiman, Donald J. 1970, "Industrialization and Social Stratification." *Sociological Inquiry* 40.
- Vallet, Louis-André 2004, "The Dynamics of Inequality of Educational Opportunity in France: Change in the Association Between Social Background and Education in Thirteen Five-Year Birth Cohorts (1908 - 1972)." Paper presented to the ISA Research Committee on Social Stratification and Mobility Neuchatel, Switzerland. May 6 - 8.

作者单位:北京大学社会学系
责任编辑:杨 可