

# 高等教育质量收益的估计:基于倾向指数匹配法的研究

刘泽云 邱牧远

**【摘要】**文章基于2005年中国城镇居民调查数据,估计接受高质量高等教育对劳动者工资的影响。研究以“211”大学作为高质量大学的标准,并使用倾向指数匹配法克服传统的OLS估计方法的缺陷。研究发现,个体能力、家庭背景、高中教育质量及大学入学机会都是影响个体能否进入高质量大学的因素。同时,匹配法的估计结果表明,在中国高等教育质量具有较高的收益,对毕业于“211”大学的劳动者而言,上“211”大学的平均收益可能使其年工资增长了37.5%。

**【关键词】**教育质量 教育收益率 匹配法 倾向指数 “211”大学

**【作者】**刘泽云 北京师范大学经济与工商管理学院,副教授;邱牧远 北京大学经济学院,博士研究生。

大量研究表明,受教育年限越多的劳动者工资收入更高(Psacharopoulos等,2004)。但随着研究的深入,越来越多的学者认识到在分析人力资本投资的作用时仅仅考虑教育数量是不够的,必须同时考虑教育质量。

在中国,尽管关于教育数量的收益率的研究非常多,并且对计量方法进行了总结和探索,但对于教育质量的关注程度却远远不够。最近的一些研究估算了中国中学教育质量对劳动者受教育状况的影响(Park等,2010)和对地区劳动生产率的影响(张海峰等,2010),而对教育质量和劳动者收入的关系则鲜有论及<sup>①</sup>。因此本文试图考察劳动者所接受高等教育的质量对其工资收入的影响。中国人口众多,加之受传统文化的影响,对高等教育的需求非常旺盛。而高质量大学的数量相对较少,导致升入高质量大学的竞争非常激烈。一般而言,能够通过高考的激烈竞争进入高质量大学的都是禀赋较高的学生,这些能力较高的学生即便不进入好大学,也能在劳动力市场上获得较高的收入,因此他们能否从接受高质量的高

<sup>①</sup> 任兆璋、范闽(2005)应用招聘网站的简历信息估算了劳动者就读于“211”高校的额外收益,但样本的随机性和代表性值得怀疑,而且样本量偏小,模型设定也存在较大问题。

等教育中获得较高的经济收益是值得怀疑的<sup>①</sup>。本文根据国家统计局的城镇居民住户调查数据,以“211”大学作为中国高质量大学的标准,运用倾向指数匹配法对中国高等教育质量的经济收益进行定量估计。

## 一、高等教育质量的衡量

教育数量一般用受教育年限或受教育程度衡量,国内外研究者对此均无异议。但如何衡量教育质量却并非易事,相关研究也未达成共识。在国外关于高等教育质量经济收益的研究中,通常使用两种方法衡量高等教育质量。一种是使用某类投入指标(如教师平均工资、生师比、教师学历、学费、学生平均入学成绩)衡量高校质量,或基于若干类投入指标生成高校的教育质量指数(Black等,2004,2006;Long,2008)。另一种是采用高等学校的评级作为识别高校教育质量的标准(Brand等,2003;Liu等,2010;Long,2010)。在中国,高等学校的投入数据几乎无法获得,于是我们借鉴美国学者用大学排名来衡量高等教育质量的思路。然而,中国目前还没有一个被广泛认可的高校排名。因此我们根据教育部确定的“211”大学来识别高等教育的质量。也就是说,如果个体所上的是“211”大学,就认为其接受了高质量的高等教育。

中国政府早在1954年便确立了重点建设一部分高等学校的政策,当年确定了6所全国重点大学,此后陆续增加重点大学的数目,截至1981年,共确定96所全国重点大学,这些学校在师资队伍、科学研究、教学管理、校园建设等方面都得到了各级政府的大力支持<sup>②</sup>。1993年,中共中央和国务院发布了《中国教育改革和发展纲要》,提出在21世纪重点建设100所左右的高等学校和一批重点学科及专业的建设工程,简称“211”工程,而被列入这一工程的学校便被称为“211”大学。1995年,“211”工程正式启动。“九五”和“十五”期间,“211”工程一期和二期分别涵盖了99所和107所高校,目前正在实施“211”工程三期。根据教育部公布的最新名单,目前全国共有112所“211”大学<sup>③</sup>。

在中国,“211”大学是各级政府重点建设的大学 and 高中毕业生首先选择的大学,也是劳动力市场上雇主做出雇佣决策的重要参考依据,甚至国外大学在招收中国留学生时也会考虑学生是否毕业于“211”大学。因此,我们认为“211”大学可以作为高质量大学的代理变量<sup>④</sup>。

① 李宏彬等(2008)基于中国双胞胎样本数据的研究表明教育收益率的OLS估计值是双胞胎样本估计值的3倍,他们认为这主要是因为在中国能力偏差对教育收益率的影响非常大,换言之,用OLS方法估计出来的教育收益率中绝大部分应归因于被忽略了的能力和背景的影响。

② <http://www.moe.gov.cn/edoas/website18/level3.jsp?tablename=1222139707228251&infoid=1223513711350102>。

③ [http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/moe\\_94/201002/82762.html](http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/moe_94/201002/82762.html)。

④ 一个可能存在的问题是,样本中包含了1995年之前上大学的劳动者,这部分人在接受高等教育时还未实施“211”工程。但原来的重点大学都成为了“211”大学,而少数“211”大学尽管过去不是重点大学,但都是办学时间长、办学质量较高的大学。因此,用“211”大学来表示高质量大学是可以接受的。

## 二、高等教育质量收益的估计方法

### (一) OLS 估计

在 20 世纪 90 年代之前,估算高等教育质量收益的常用方法是在明瑟收入方程中加入衡量高等教育质量的变量(Speakman 等,2006)。即:

$$Y_i = \gamma D_i + \beta X_i + U_i \quad (1)$$

$Y_i$  是工资收入的自然对数; $D_i$  是衡量高等教育质量的变量,如在本文中, $D_i=1$  表示个体上的是“211”大学, $D_i=0$  则反之; $X_i$  是一系列控制变量; $U_i$  是随机误差项。参数  $\gamma$  的 OLS 估计值即为高等教育质量的收益。

### (二) 匹配法

OLS 估计的本质是在控制可观测因素的情况下,比较接受高质量和低质量高等教育的个体的平均工资差异。这一方法存在两个缺陷(Black 等,2004):(1)假定对数工资与解释变量之间存在线性关系。(2)可能违反公共集假设,即在给定了解释变量的取值后,有时只存在接受高质量高等教育的个体,或者只存在接受低质量高等教育的个体,这时就无法比较两类个体的工资差异。针对上述情况,目前常用的一种方法是匹配法。这种方法放宽了被解释变量与解释变量存在线性关系的假设,并且有效地处理了公共集假设问题。

匹配法将是否上“211”大学视为一个处理:上“211”大学的群体是处理者;未上“211”大学的群体是非处理者。而上“211”大学对收入的影响就是处理效应。设  $Y$  为结果变量; $Y_{1i}$  为接受处理状态下的结果,即个体  $i$  上“211”大学的潜在收入; $Y_{0i}$  为未接受处理状态下的结果,即个体  $i$  未上“211”大学的潜在收入。

上“211”大学对个体收入的影响可通过多种处理效应体现。最常用的一种是处理者的平均处理效应(简称 ATT),即对于实际上了“211”大学的人而言,上“211”大学和假如不上“211”大学的潜在收入的平均差异。令  $D_i=1$  表示个体上了“211”大学,反之  $D_i=0$ 。则  $ATT = E(Y_{1i} | D_i=1) - E(Y_{0i} | D_i=1)$ 。如果在控制了决定上“211”大学的可观测特征  $X_i$  之后,是否上“211”大学与潜在收入  $Y_{0i}$  是独立的,即满足条件独立假设  $(Y_{0i} \perp D_i) | X_i$ ,那么,  $E(Y_{0i} | X_i, D_i=1) = E(Y_{0i} | X_i, D_i=0)$ 。从而:

$$\begin{aligned} ATT &= E(Y_{1i} | D_i=1) - E(Y_{0i} | D_i=1) \\ &= E\{[E(Y_{1i} | X_i, D_i=1)] | D_i=1\} - E\{[E(Y_{0i} | X_i, D_i=1)] | D_i=1\} \\ &= E\{[E(Y_{1i} | X_i, D_i=1)] | D_i=1\} - E\{[E(Y_{0i} | X_i, D_i=0)] | D_i=1\} \end{aligned}$$

在满足条件独立假设的情况下,基于可观测特征(也称控制变量) $X_i$ ,将样本分成若干个单元格。同一单元格内处理者与非处理者的收入差异仅仅归因于处理的结果,将收入差异基于处理者( $D_i=1$ )的数目进行平均就可以得到 ATT。

使用上述匹配法可能会违反公共集假设,即在控制了  $X_i$  之后,在一个存在处理者的单元格内,可能不存在非处理者的样本。倾向指数(简称 P-score)方法可以解决这一问题。倾

向指数定义为  $P(X_i) = \text{Prob}(D_i=1 | X_i)$ , 即在控制  $X_i$  的情况下, 参与处理  $D_i$  的概率。如果条件独立假设成立, 那么有  $(Y_{0i} \perp D_i) | P(X_i)$ , 即无论是控制  $X_i$  还是控制  $P(X_i)$ ,  $Y_{0i}$  与  $D_i$  都是独立的 (Rosenbaum 等, 1983)。倾向指数的估计量可以用 Probit 或 Logit 模型估计得到, 对于倾向指数为  $P(X_i)$  的处理者, 用  $P(X_i)$  附近的非处理者进行匹配。

在实际应用中, 倾向指数匹配法有分层匹配法、最小邻域匹配法、半径匹配法和核匹配法等几种方法。目前常用的是后 3 种方法, 其中核匹配法又分为高斯核匹配法和伊氏核匹配法两种 (Becker 等, 2002)。同时, 可以只对处于公共集内的样本进行匹配, 提高匹配的精确性。另外, 在估计倾向指数时, 平衡性条件可以帮助我们确定  $X_i$  的选择是否合适。所谓平衡性条件, 是指倾向指数相同的处理组和非处理组样本的控制变量的分布应该相同 (Becker 等, 2002)。

在本文中, 我们主要关注处理者的平均处理效应 (ATT)。但研究者可能还关注其他类型的处理效应。如非处理者的平均处理效应 (TUT), 指对于未上“211”大学的个体而言, 上“211”大学带来的平均潜在收益; 平均处理效应 (ATE), 指随机从接受过高等教育的样本中抽出一个个体, 上“211”大学与不上“211”大学的期望收入差距。

### 三、样本和数据

本文使用的数据来自国家统计局于 2005 年进行的城镇住户调查, 调查对象为城镇市区和县城关镇区内的住户, 反映的是 2004 年的情况, 研究涉及北京、浙江、黑龙江、山西、陕西和贵州。2004 年这 6 个省份的人均 GDP 分别为: 北京 37 058 元、浙江 23 942 元、黑龙江 13 897 元、山西 9 150 元、陕西 7 757 元、贵州 4 215 元 (国家统计局, 2005)。可见, 样本能够代表中国东部、中部和西部 3 个不同经济发展水平地区的情况。我们将分析样本限定在 1978 年后接受大学教育、年龄为 16~60 岁并且在 2004 年取得劳动收入的在职劳动者。受访者填写了自己最后一次高考后所上大学的名称, 根据教育部公布的“211”工程大学的最新名单, 可以判断受访者第一学历所上大学是否为“211”大学。经整理后, 用于分析的样本数为 753 人, 其中有 124 人的第一学历所上大学为“211”大学。

Imbens (2004) 指出, 匹配法中的控制变量应该是在处理之前就已经确定的变量, 即这些变量不会受到处理的影响。这与 Angrist 等 (2008) 提出的“好的控制变量”的概念类似。在本文中, 控制变量应该是那些影响个体能否上“211”大学的变量, 而不是受到个体高等教育经历影响的变量 (如工作地、职业等)。基于上述原则, 借鉴国外相关研究的经验 (Brand 等, 2003; Black 等, 2004; Long, 2008) 并参考中国国情, 我们选择了以下几类控制变量: (1) 性别、年龄和民族等个体人口学特征。(2) 以父亲的受教育年限作为劳动者个体能力的代理变量, 以减小遗漏能力变量引起的估计偏误<sup>①</sup>。(3) 以劳动者的兄弟姐妹数量和高中毕业时家

<sup>①</sup> 由于样本中母亲受教育年限的缺失值太多和无法得知每一个个体的兄弟姐妹的教育水平, 所以没有使用母亲及兄弟姐妹的受教育年限。



庭在当地的消费水平(简称“家庭消费水平”)作为衡量家庭背景的代理变量<sup>①</sup>。(4)以劳动者就读的高中在当地的水平(简称“高中质量”)衡量高中教育质量。(5)中国高考制度的一个重要特征是,高校的招生指标是按省份确定。即在不同的省份,“211”大学的招生指标与高考考生的比例存在差异。个体和家庭特征相同的两个人如果在不同的省份参加高考,那么他们进入“211”大学的机会是不同的<sup>②</sup>。由于缺乏数据,我们无法计算各个省份每个年份进入“211”大学的入学机会。考虑到各个省份的大学入学机会与“211”大学入学机会存在很强的相关性,我们以前者作为后者的替代性指标。具体而言,我们使用李立峰、刘海峰(2007)计算的1978~2002年各省份的大学入学机会指数。其方法是,首先计算某省本、专科招生人数占全国本、专科招生人数的比例,然后计算该省高考报名人数占全国高考报名人数的比例,前者除以后者就得到该省的大学入学机会指数<sup>③</sup>。某省的大学入学机会指数大于1表明该省的大学入学机会高于全国平均水平,等于1表明与全国平均水平相当,小于1表明低于全国平均水平。大学入学机会指数的数值越大,说明该省高考考生进入大学的可能性越大,近似地意味着进入“211”大学的可能性也越大。本文使用的数据记录了劳动者最后一次参加高考的时间和省份,可以确定每一个劳动者的大学入学机会指数。

表1给出了相关变量的描述性统计。这里的年工资指2004年全年的工薪收入,即劳动收入,包括基本工资、津贴、奖金、提成和平均加班费等。上过和未上过“211”大学的样本的年工资均值分别为30816元和21316元,前者比后者高约45%。两个样本在年龄和性别方面差别不大,但在民族、兄弟姐妹数量、父亲受教育年限、家庭消费水平、高中学校质量及大学入学机会等方面存在明显差异。

## 四、高等教育质量收益的估计结果

### (一) 倾向指数的估计

我们使用Logit模型估计了倾向指数,因变量为是否上“211”大学的虚拟变量(取值为1表示上了“211”大学,反之取值为0),结果如表2所示。需要说明的是,所有控制变量都通过了平衡性条件的检验。可以看出,以父亲受教育年限衡量的个体能力对进入“211”大学有显著的正向影响。在家庭背景方面,兄弟姐妹数量对进入“211”大学的影响显著为负。高中

- ① 接受高等教育需要个人支付学费等直接成本,并承担因上学而放弃的收入,因此家庭经济状况会影响个体的高等教育选择。相关研究中常用父母的收入衡量家庭经济状况,但我们所用的数据没有这方面的信息,因此用家庭的消费水平替代。另外,在既定的家庭资源约束下,子女数量的增加会降低其质量(Björklund等,2011)。
- ② 乔锦忠(2007)计算了2006年中央部属高校在全国各个省份的计划录取率(部属高校在该省的计划招生数除以该省高考报名人数),其中,上海市最高(13.77%),河南省最低(1.89%)。
- ③ 由于缺乏各省每年高考报名人数的数据,李立峰、刘海峰(2007)用1999~2003年的数据对各省高中毕业生数与高考报名人数的相关性进行了分析,发现二者高度正相关,因此在计算中,实际使用的变量是高中毕业生数而非高考报名人数。

表 1 变量的描述性统计(均值)

变 量	处理组 (“211”大学)	非处理组 (非“211”大学)	变 量	处理组 (“211”大学)	非处理组 (非“211”大学)
年工资(元)	30816	21316	父亲受教育年限	9.62	8.02
年工资对数	10.10	9.70	家庭消费水平		
年龄	36.85	37.26	高 *	0.12	0.08
性别 *(男性 =1)	0.58	0.58	一般 *	0.67	0.68
民族 *(少数民族 =1)	0.10	0.07	低 *	0.21	0.24
最终学历			高中质量		
大专 *	0.21	0.42	好 *	0.46	0.41
大学本科 *	0.68	0.54	一般 *	0.27	0.21
研究生及以上 *	0.11	0.05	差 *	0.27	0.38
兄弟姐妹数量	2.93	3.48	大学入学机会指数	1.50	1.26
样本量	124	629	样本量	124	629

注: \* 表示该变量为虚拟变量。

质量的影响也是显著的,低质量高中的考生进入“211”大学的概率显著低于高质量高中的考生。另外,来自入学机会较高省份的考生更有可能进入“211”大学。

(二) 倾向指数匹配法的估计结果

表 2 Logit 模型倾向指数估计结果

解释变量	回归系数	标准误
年龄	-0.025	0.080
年龄的平方	0.000	0.001
男性	0.083	0.116
少数民族	0.264	0.197
父亲受教育年限	0.042***	0.015
兄弟姐妹数量	-0.104**	0.045
家庭消费水平(高为参照组)		
一般	-0.134	0.190
低	-0.015	0.222
高中质量(高为参照组)		
一般	0.045	0.144
低	-0.240*	0.132
大学入学机会指数	0.094*	0.053
常数项	-0.769	1.539
公共集	[0.03,0.54]	
样本量	753	

注: \* p<0.1, \*\* p<0.05,\*\*\* p<0.01。

在得到倾向指数之后,我们用最小邻域匹配法、半径匹配法和核匹配法对数据进行匹配得到各类处理效应。表 3 给出了利用公共集样本和全部样本得到的各种处理效应。其中,带宽是使用核匹配法或半径匹配法时利用交叉验证法所得到的最优带宽。我们还利用交叉验证法对这几种方法进行了比较,发现采用伊氏核匹配法能够得到最小平均平方误差,而最小邻域法和半径匹配法所得的平均平方误差较大,特别是最小邻域匹配法,这与 Black 等(2004)的结果相同<sup>①</sup>。因此,我们主要关注伊氏核匹配法的估计结果。另外,在使用核匹配法时发现带宽的选择对估计结果基本没有影响,只有在带宽很小时才会使处理效应有所变化。

① Frölich(2004)指出,由于最小邻域匹配法在构造反事实估计时选取样本点存在不稳定的因素,因此应谨慎看待其得到的结果。

从表 3 可以看出,使用公共集内样本得到的处理效应略大于使用全部样本得到的处理效应,由于前者的估计准确度更高,这里仅讨论公共集内的处理效应。伊氏核匹配法的估计结果表明,ATT 为 0.375,即对那些毕业于“211”大学的人而言,上“211”大学使其年工资增加了 37.5%,这是一个很可观的经济收益<sup>①</sup>。从数值上看,ATT 略大于 TUT 和 ATE。

表 3 倾向指数匹配法得到的处理效应

	公共集内样本(n=749)				全部样本(n=753)			
	核匹配法		半径	最小	核匹配法		半径	最小
	高斯核	伊氏核	匹配法	邻域法	高斯核	伊氏核	匹配法	邻域法
ATT	0.377	0.375	0.378	0.341	0.348	0.355	0.357	0.317
标准误	0.061	0.061	0.064	0.076	0.061	0.061	0.064	0.076
ATE	0.376	0.365	0.368	0.353	0.349	0.355	0.358	0.343
标准误	0.064	0.064	0.068	0.073	0.064	0.064	0.068	0.073
TUT	0.375	0.363	0.366	0.355	0.349	0.356	0.359	0.348
标准误	0.068	0.065	0.070	0.075	0.068	0.065	0.070	0.075
带宽	0.030	0.120	0.080	-	0.030	0.120	0.080	-

注:所有处理效应均在 1%的水平上显著;标准误估计使用 bootstrap 方法(500 次);带宽是使用交叉验证法得到的最优带宽。

在利用匹配法对美国的研究中,一些研究发现高校质量对工资的影响显著为正,有一些则不然。例如,Black 等(2004)发现在控制了劳动者的受教育年限后,高质量大学的收益率不显著,而只有在未控制劳动者受教育年限时高质量大学的收益率才是显著的。Brand 等(2003)发现,就读于精英大学只对劳动者早期的职业生涯产生影响,对中期和后期的职业生涯没有显著影响。Long(2008)用 5 个指标衡量高等学校质量,发现高等学校质量对年轻女性劳动者的小时工资没有影响,而用某些指标衡量的高等教育质量对年轻男性劳动者的小时工资有显著的正向影响。

我们的研究结论与美国学者的研究结论之间存在一定差异,可能有以下 3 个原因:(1)用于衡量高等教育质量的指标不同,导致估计结果不同。由于我们缺乏高等学校的相关数据,无法用不同指标衡量中国高等学校的质量,难以检验结果的稳健性。(2)从理论上说,遗漏能力变量和家庭背景变量会高估教育质量的收益率。美国研究者使用的数据中有很多变量可作为二者的代理变量,如能力测试的成绩、高中时的成绩、父母的收入等。但我们的数据中缺乏上述信息,导致估计结果偏高。(3)一些特殊的因素使中国高等教育质量对劳动者收入的影响比美国大。如中国劳动力市场分割性较强,城市劳动力的流动性不高,毕业于名牌大学的劳动者不但在初次就业时更容易获得高收入的职位,而且能够在其终生都保持这种收入优势。

高等教育质量的经济收益可分为两个部分:一方面,在所受高等教育数量相同的情况下,较高的高等教育质量增加了个体的人力资本,提高其工资收入,可视为高等教育质量的直接收益;另一方面,较高的高等教育质量增加了个体接受更多高等教育的概率,最终提高

① 高斯核匹配法和半径匹配法的 ATT 分别为 0.377 和 0.378,与伊氏核匹配的 ATT 几乎没有差异。

其工资收入,可视为高等教育质量的间接收益<sup>①</sup>。因此,如果不控制个体的最终学历,得到的是高等教育质量的全部收益,而控制最终学历后得到的是高等教育质量的直接收益。显然,控制最终学历后上“211”大学的处理效应会小于不控制最终学历的处理效应。

我们在表2的控制变量的基础上加入个体的最终学历,得到倾向指数匹配法的各类处理效应。其中,基于公共集内样本的伊氏核匹配法的ATT为0.284(见表4),明显小于不控制最终学历时的ATT(0.375)。Black等(2004)针对美国的研究也发现,在匹配法中加入劳动者受教育年限作为控制变量后,高等教育质量收益的估计值变小了。但如前所述,匹配法中的控制变量应该是不受处理影响的变量,而劳动者的最终学历显然受到其第一学历是否为“211”大学的影响。因此,最终学历是一个“不好的控制变量”(Angrist等,2008),不应该在匹配法中用它作为控制变量。这里的讨论只是为了说明高等教育质量存在直接收益和间接收益,并非给出直接收益和间接收益的精确估计。

我们还根据式(1)利用OLS方法估计上“211”大学的经济收益,控制变量 $X_i$ 包括表2中的全部变量,结果如表4所示。从表4可以看出,两种方法的估计结果存在差异,这源于OLS估计没有考虑公共集假设,也没有考虑到高等教育质量对收入的影响可能是非线性的。

表4 使用不同方法得到的高等教育质量收益估计值		
	伊氏核匹配法(ATT)*	OLS
不加入最终学历作为控制变量	0.375	0.354
加入最终学历作为控制变量	0.284	0.260

注:使用公共集内样本。

五、结 语

本文根据中国城镇居民调查数据估计接受高质量高等教育对劳动者工资的影响,研究发现,个体能力、家庭背景、高中教育质量及大学入学机会都是影响个体能否进入高质量大学的因素。同时,匹配法的估计结果表明,在中国接受高质量高等教育具有较高的收益。从提高个人收入的角度看,接受高质量的高等教育是值得的。本研究表明,中国高质量高等教育存在显著的经济收益,意味着提高高等教育质量对于增加居民收入具有积极的作用。当然,本文的分析基于2004年的数据,随着时间的推移,高质量高等教育的收益可能会发生变化,这需要用近年份的数据来进行检验。此外,还应该分析高质量高等教育给个人带来的非货币性回报,以及对经济增长和社会发展的影响,这样才能更完整地理解高等教育质量的社会经济价值。

本研究存在的主要问题是:(1)限于数据原因,本文使用的代理变量有可能不能很好地纠正遗漏变量造成的估计偏误,从而导致高质量高等教育收益的估计结果偏高。(2)本文使用的匹配法比OLS估计方法更为合理,但匹配法基于可观测的个体特征比较处理组和非处理组的结果差异,无法纠正因不可观测异质性导致的选择偏差。

参考文献:

1. 国家统计局(2005):《中国统计年鉴(2005)》,中国统计出版社。  
2. 李立峰、刘海峰(2007):《中国高校招生考试中的区域公平研究》,华中师范大学出版社。

<sup>①</sup> Eide等(1998)的研究表明,本科的教育质量会影响个体接受研究生教育的概率。



3. 李宏彬等(2008):《中国城镇教育投资回报率:基于双胞胎的研究》,载于李宏彬、张俊森:《中国人力资本投资与回报》,北京大学出版社。
4. 任兆璋、范闽(2005):《中国教育质量与收益率的微观计量分析》,《华南理工大学学报(自然科学版)》,第9期。
5. 乔锦忠(2007):《优质高等教育入学机会分布的区域差异》,《北京师范大学学报(人文社科版)》,第1期。
6. 张海峰等(2010):《教育质量对地区劳动生产率的影响》,《经济研究》,第7期。
7. Angrist, J. & Pischke, J. (2008), *Mostly Harmless Econometrics*. Princeton University Press.
8. Becker, O.S. & Ichino, A. (2002), Estimation of Average Treatment Effects Based on Propensity Scores. *The Stata Journal*. 2(4), 358-377.
9. Björklund, A. & Salvanes, K.G. (2011), Education and Family Background: Mechanisms and Policies. in Hanushek, E.A., Machin, S. & Woessmann, L. (eds.) *Handbook of the Economics of Education* (Vol.3). Amsterdam: Elsevier. 201-248.
10. Black, D.A. & Smith, J. A. (2004), How Robust is the Evidence on the Effects of College Quality? Evidence from Matching. *Journal of Econometrics*. 121, 99-124.
11. Black, D. A. & Smith, J. A. (2006), Estimating the Returns to College Quality with Multiple Proxies for Quality. *Journal of Labor Economics*. 24(3), 701-728.
12. Brand, E. J. & Halaby, N.C. (2003), Regression and Matching Estimates of the Effects of Elite College Attendance on Career Outcomes. *2003 Annual Meeting of the American Sociological Association*. Atlanta.
13. Eide, E., Brewer, D. J., & Ehrenberg, R.G. (1998), Does it Pay to Attend an Elite Private College? Evidence on the Effects of Undergraduate College Quality on Graduate School Attendance. *Economics of Education Review*. 17(4), 371-376.
14. Frölich, M. (2004), Finite-sample Properties of Propensity-score Matching and Weighting Estimators. *The Review of Economics and Statistics*. 86(1), 77-90.
15. Hanushek, E.A. & L. Woessmann (2008), The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*. 46(3): 607-668.
16. Imbens, G. W. (2004), Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review. *The Review of Economics and Statistics*. 86(1): 4-29.
17. Liu, X., Thomas, S. & Zhang, L. (2010), College Quality, Earnings, and Job Satisfaction: Evidence from Recent College Graduates. *Journal of Labor Research*. 31(2), 183-201.
18. Long, M.C. (2008), College Quality and Early Adult Outcomes. *Economics of Education Review*. 27, 588-602.
19. Long, M.C. (2010), Changes in the Returns to Education and College Quality, *Economics of Education Review*. 29, 338-347.
20. Park, A., Shi, X., Hsieh, C., & An, X. (2010), Does School Quality Matter?: Evidence from a Natural Experiment in Rural China. <http://www.economics.ox.ac.uk/members/albert.park/papers/schoolquality.pdf>.
21. Psacharopoulos, G. & H.A. Patrinos (2004), Returns to Investment in Education: A Further Update. *Education Economics*. 12(2): 111-134.
22. Rosenbaum, P.R. & D.B. Rubin (1983), The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*. 70(1), 41-55.
23. Speakman, R. & F. Welch (2006), Using Wage to Infer School Quality. in Hanushek, E.A. & F. Welch (eds.) *Handbook of the Economics of Education* (Vol.1). Amsterdam: Elsevier. 813-864.

(责任编辑:朱 萍)