

教育投资具有外部收益吗？

——来自中国的经验证据

刘泽云

[摘要] 使用2005年全国1%人口抽样调查数据,在微观明瑟方程中加入代表地区平均教育水平的变量,使用工具变量法估算我国地市层面和省级层面的教育外部收益率。主要结论是:第一,在我国城镇地区存在明显的教育外部性,城市劳动者平均受教育年限提高一年,将导致在该城市工作的劳动者工资提高35% - 38%,而城市劳动者中大学学历劳动者的比例提高一个百分点,将导致劳动者工资提高3.1% - 3.6%。第二,省级层面的教育外部收益率高于地市级层面。第三,高学历劳动者获得的教育外部收益率高于低学历劳动者。研究结论证明了政府从纠正外部性的角度投资教育的正当性,支持省级政府比基层政府承担更大教育责任的政策取向,同时表明地区平均教育受水平的提高更有利于高学历劳动者的工资增长。

[关键词] 教育;人力资本;外部收益率

一、引言

教育收益包括私人收益(private benefits)和外部收益(external benefits),前者指受教育者本人从教育中获得的收益,后者指外溢到其他社会成员的收益,即教育的外部性(externalities)。教育是最重要的人力资本投资形式, Lucas (1988)对人力资本的外部性给出了明确的解释。他将人力资本的作用分为内部效应(internal effect)和外部效应(external effect)两个方面。前者指个体的人力资本能够提高自身的生产率和收入;后者指平均人力资本的增加能提高总体的生产率,从而提高整体的收入水平,但没有人会在自己进行人力资本投资时

[收稿日期] 2016 - 03 - 01

[基金项目] 北京师范大学“中央高校基本科研业务费专项资金”

[作者简介] 刘泽云,北京师范大学经济与工商管理学院/首都教育经济研究院,电子邮箱地址:zeyun_liu@bnu.edu.cn。

考虑后者,所以称之为外部效应。Lucas(1988)认为这种外部性的来源是知识的分享以及劳动者之间正式和非正式的交流,而 Acemoglu(1996)认为源于研究开发的外部性或技能偏向型技术进步(skill-biased technical change)。

自上世纪六十年代人力资本理论创立以来,估算教育私人收益率的文献可谓汗牛充栋,关于我国的研究也积累了很多文献。但相对来说估算教育外部收益率的文献却为数不多,针对我国的高质量研究更是屈指可数。Lange 和 Topel(2006)援引诺贝尔经济学奖获得者贝克尔(G. Becker)的观点来描述这一状况:“一般而言,经济学家在估计各种投资的社会价值方面很少取得成功,不幸的是,教育也不例外。”

研究教育外部收益的经济学文献分为两类。第一类是分析某种特定的教育外部收益,如教育对政治稳定、减少犯罪或促进公民参与的影响(McMahon, 2010)。第二类是估算教育的总量外部收益(aggregate externalities),主要使用两种方法。一种方法是用宏观增长方程估算教育的社会收益率,用微观明瑟方程估算教育的私人收益率,前者减去后者就是教育的外部收益率(Breton, 2010)。另一种方法是基于个体数据,在微观明瑟方程中加入代表地区(州或城市)平均教育水平的变量,以这一变量的系数衡量教育的外部收益率。这种方法始于 Rauch(1993)的研究,近来得到越来越多的应用,如 Acemoglu 和 Angrist(2000)、Moretti(2004)以及 Ciccone 和 Peri(2006)对美国的研究, Kirby 和 Riley(2008)对英国的研究, Dalmazzo 和 Blasio(2007)对意大利的研究, Bauer 和 Vorell(2010)对德国的研究, Muravyev(2008)对俄罗斯的研究。

一些研究基于个体数据,使用估算教育总量外部收益的后一种方法估计了我国的教育外部收益率。例如 Liu(2007)使用“中国居民收入调查(CHIP)”数据,李小瑛等(2010)、徐大丰(2010)以及 Fan 和 Ma(2012)使用“中国健康营养调查(CHNS)”数据,都估算了我国教育的外部收益率或高等教育的外部收益率。但这几项研究存在的共同问题是均使用已公开的抽样调查数据,但涉及到的省或城市的样本数太少,缺乏足够的代表性。特别是李小瑛等(2010)以及 Fan 和 Ma(2012)估计的是省级层面的教育外部收益率,但其使用的 CHNS 数据只有八个样本省份,让我们不能不对其研究结论存有质疑。

本文使用 2005 年全国 1% 人口抽样调查数据,在微观明瑟方程中加入代表地区平均教育水平的变量,使用工具变量法估算我国城镇地区的教育外部收益率。与以往同类研究相比,本文的特色在于:第一,使用大样本的全国人口抽样调查数据,覆盖了所有的省份和地市,分析样本具有很好的代表性;第二,分别在地市和省两个层面进行估计,比较不同地域层面的教育外部收益率;第三,分别使用滞后变量和制度变量作为工具变量解决内生性问题,并得到了稳健的

估计结果。

估算教育的外部收益率有着重要的理论和现实意义。经济学家普遍认为,教育具有正的外部性是政府干预教育的重要理由之一,否则就会出现教育供给不足。^① Lange 和 Topel (2006) 的综述性文章也指出,几乎没有证据表明教育的外部性是负的。但笔者认为到目前为止,除了 Liu (2007) 的研究,其它基于微观数据估计我国的教育外部收益率的文献都缺乏说服力,而 Liu (2007) 的研究则存在样本代表性的问题。本文将利用大样本的全国人口抽样调查数据弥补相关研究的不足。毫无疑问,这将为教育政策的制定提供参考。同时,本文分别估计了省级层面和地市层面的教育外部收益率,可以为不同层级政府承担的教育责任的大小提供依据。此外,本文还比较了不同学历劳动者从地区教育水平提高中获得的外部收益率,这将有助于解释不同学历劳动者之间的收入差距。

二、模型和方法

(一) 理论模型

假定劳动者的边际产出(q)是个人人力资本(h)、地区人力资本水平(H)及其他生产性投入的函数:

$$q_{ij} = Ah_{ij}^{\beta_1} H_j^{\beta_2} \varphi(X_{ij}) \quad (1)$$

式(1)中, i 和 j 分别代表劳动者和地区, A 代表外生的技术因素, $0 < \beta_1 < 1$ 且, $0 \leq \beta_2 \leq 1$, φ 表示除个人人力资本和地区人力资本以外的其他因素。

假定劳动者的工资(W_{ij})按照其边际产出支付,即 $W_{ij} = q_{ij}$,则工资收入的对数为:

$$\ln W_{ij} = \ln A + \beta_1 \ln h_{ij} + \beta_2 \ln H_j + \ln \varphi(X_{ij}) \quad (2)$$

进一步,假定劳动者个人的人力资本是其受教育水平(S_{ij})的函数 $h_{ij} = e^{S_{ij}}$, $H_j = e^{\bar{S}_j}$ 是地区平均受教育水平(\bar{S}_j)的函数。^② 式(2)可改写为:

$$\ln W_{ij} = \beta_0 + \beta_1 S_{ij} + \beta_2 \bar{S}_j + \ln \varphi(X_{ij}) \quad (3)$$

在式(3)中, $\beta_0 = \ln A$, β_1 是教育的私人收益率, β_2 则衡量了教育的外部收益

^① 例如弗里德曼(1986)认为,政府干预教育的两个理由之一是教育具有相当多的“邻近影响”,即外部性。

^② Lucas(1988)在分析人力资本的外部性时,直接将地区人力资本水平定义为地区平均受教育年限。本文假设人力资本与受教育水平存在指数函数关系,这样式(3)中的系数可以方便地解释为教育收益率。

率。可见,劳动者的工资收入受到个人受教育水平的影响,也受到地区平均受教育水平的影响。

(二) 计量模型

基于式(3),本文在计量分析中使用以下模型:

$$\ln W_{ij} = \beta_0 + \beta_1 S_{ij} + \beta_2 \bar{S}_j + \gamma X_{ij} + \theta G_j + \mu_{ij} \quad (4)$$

其中,下标 j 表示地区(省或地市),下标 i 表示个人。 $\ln W_{ij}$ 为在地区 j 的个人 i 的工资的自然对数。 β_0 为截距项。 S_{ij} 表示个人的受教育水平,用个人的受教育年限衡量。 \bar{S}_j 表示地区 j 的劳动者的平均受教育水平,用两个指标衡量:其一,该地区劳动者的平均受教育年限;其二,该地区大学学历(包括大学专科、本科和研究生)劳动者占全体劳动者的比例。 X_{ij} 为个体层面的控制变量,包括性别、工作经验和工作经验的平方。 G_j 为地区层面的控制变量,包括在不同产业部门和不同性质单位中就业的劳动者的比例。^① 如果分析的是地市层面的教育外部收益率,则还加入表示地市所在省份的虚拟变量。 μ_{ij} 为随机误差项。

式(4)是明瑟工资方程的一种变形。 β_1 为教育的私人收益率,表示在控制其它变量的情况下,个人受教育年限提高一年导致的个人工资提高的百分比。 β_2 则为教育的外部收益率,表示在控制个人受教育年限以及其它变量的情况下,一个地区的劳动者的平均受教育年限提高一年(或大学学历劳动者比例提高一个百分点)导致的个人工资提高的百分比。

(三) 工具变量

使用 OLS 方法估计式(4)可能存在的问题是:影响个人工资的不可观测因素与个人所在地区的平均受教育水平相关,从而教育外部收益率(即 β_2) 的 OLS 估计结果不能反映地区平均受教育水平对工资收入的因果效应。解决这一内生性偏误问题的常用方法是工具变量法(IV),即找到与地区平均受教育水平相关但不直接影响个人工资收入的变量作为 IV。

经济学研究中经常使用滞后变量作为工具变量。本文根据公开出版的 2000 年人口普查数据(国务院人口普查办公室、国家统计局人口和社会科技统计司,2003),得到 2000 年各地区(省或地市)居民的平均受教育年限,并计算出各地区 6 岁以上人口中大学学历人口的比例,以此作为 2005 年各地区劳动

^① 地区层面控制变量的选择参考了 Liu(2007)的研究。李小瑛等(2010)和徐大丰(2010)没有加入地区层面的控制变量,这显然不合适。Fan 和 Ma(2012)使用了地区人均 GDP 作为控制变量,但地区平均教育水平提高对总体生产率的影响将体现在地区人均 GDP 的变化中,如果控制了人均 GDP,就无法充分反映地区平均教育水平对个人工资的影响。

者平均受教育水平的工具变量^①。其基本思想是,一个地区滞后的教育变量与当前的教育变量是相关的,但不会直接影响当前的个人工资。即2000年的地区平均教育水平与2005年的劳动者平均受教育水平相关,但不会直接影响2005年的个人工资。

在经济学文献中,法律或政策的制定和实施也常常用来作为工具变量。义务教育对于一个地区的教育发展和居民受教育水平的提高最为重要,而我国1986年颁布的《义务教育法》和1992年颁布的《义务教育法实施细则》有一个突出的特点,即要求地方根据本地的经济和社会发展状况分阶段、有步骤地实现九年义务教育。这就意味着义务教育的普及并不是均衡推进的,而是因地方而异,这将在很大程度上导致地区间居民受教育水平存在差异。Liu(2007)指出,我国义务教育法的实施与地区的行政级别有很密切的关系(即行政级别越高,对义务教育法的贯彻和执行越好,也能越早普及义务教育),并以此构造工具变量。受到这一研究的启发,本文根据地市的级别构造了一个有序变量,以此作为地市劳动者平均受教育水平的工具变量。这一有序变量的取值为:北京、上海、天津三个直辖市为2;重庆、省会城市和计划单列市为1^②,其他地市为0。该变量与地市的平均受教育水平存在正相关关系,但与个人的工资收入应该没有直接关系。

另外,中国高考制度的一个重要特征是,高校的招生指标数根据省份来确定,但在不同的省份,高校招生指标数与高考考生数的比例并不相同,甚至差异很大。从而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(她)们进入大学接受高等教育的机会是不同的。因此,高等教育入学机会的差异将导致平均受教育水平的省际差异。本文使用了李立峰和刘海峰(2007)计算的1981-2002年各省平均的大学入学机会指数。在计算一个省某个年份的大学入学机会指数时,首先计算当年该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。^③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

① 使用2000年劳动力人口的平均受教育年限和大学学历人口比例更为合适,但根据公开的人口普查资料无法得到这些数据。

② 计划单列市包括深圳、厦门、宁波、青岛和大连。重庆虽为直辖市,但与其它三个直辖市相比教育发展水平较低(以2000年居民平均受教育水平为例,北京、天津和上海均超过了9年,而重庆仅为7.28年),因此将重庆与省会城市和计划单列市归为一类。

③ 由于缺乏各省每年高考报名人数数据,李立峰和刘海峰(2007)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。

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。显然,一个省的大学入学机会指数与该省居民的平均受教育水平相关,但不会直接影响个人的工资收入。

综上,本文在估计地市层面的教育外部收益率时,分别使用三个变量作为地市平均教育水平的工具变量,即2000年居民平均受教育年限、2000年6岁以上人口中大学学历人口比例、地市行政级别。而在估计省级层面的教育外部收益率时,分别使用三个变量作为省平均教育水平的工具变量,即2000年居民平均受教育年限、2000年6岁以上人口中大学学历人口比例、大学入学机会指数。

需要说明的是,在本文的所有回归分析中,均以地区的个体样本数作为权重进行加权。

三、样本和数据

本文使用的数据来源于2005年全国1%人口抽样调查中随机抽取的20%样本,原始数据涉及全国31个省、市、自治区的2585481名个体。对分析样本的选定基于以下几点考虑:第一,分析仅限于具有非农业户口且现居住地为城镇的劳动力,不考虑在农村工作的劳动力和进城农民工;第二,年龄限制在16-60岁且不在学校就读;第三,雇主或自营劳动者的收入受到非人力资本因素的影响较大,因此仅考虑并且就业身份为雇员的在职职工;第四,仅考虑月劳动收入大于0的样本。第五,西藏自治区情况特殊,而且经筛选后的样本量非常小(仅有270人),因此在分析中也不予考虑。最终用于分析的个体样本数为235120人。

但是,在估计地市层面的教育外部收益率时遇到两个问题。其一,湖北、海南、新疆三个省区的省直辖县级行政单位共用同一个地市级行政代码,但它们并不属于同一个地市,因此这些样本没有进入分析样本。其二,有些地市的个体样本数较少,可能会对估计结果造成影响。为此,在估计地市层面的教育外部收益率时将个体样本数少于100的地市删去。最终在地市层面的分析中使用的个体样本数为232052人,涉及四个直辖市和314个地市级行政区域。

在个体层面的变量中,月工资指被调查者在2005年10月份取得的劳动收入,财产性收入、转移性收入等非劳动收入不包括在内。个体的受教育程度分为三类,“初中及以下”包括未上过学、小学、普通初中和相当于初中程度的技工学校;“高中”包括普通高中、职业高中、中专和相当于高中程度的技工学校;“大学”包括专科、本科和研究生。个体的受教育年限根据其学历推算得出。

具体的推算规则为:(1)如果未上过学,受教育年限为0年;(2)如果学业完成情况为毕业或肄业,受教育年限分别为:小学6年,初中9年,高中12年,大专14年,本科16年,研究生及以上19年;(3)如果学业完成情况为辍学或其他,则最高教育程度取半,即小学3年,初中7.5年,高中10.5年,大专13年,本科14年,研究生及以上17.5年。工作经验根据个人的年龄和受教育年限计算得出,计算公式为:工作经验 = 年龄 - 受教育年限 - 6。

地区层面的变量均根据本文使用的个体样本在地区层面(省或地市)上加总而得。产业划分根据《国民经济行业分类》(GB/T 4754 - 2002)的标准,第一产业指农、林、牧、渔业(不含农、林、牧、渔服务业);第二产业指采矿业(不含开采辅助活动),制造业(不含金属制品、机械和设备修理业),电力、热力、燃气及水生产和供应业,建筑业;第三产业即服务业,指除第一产业、第二产业以外的其他行业。在工作单位性质中,国有部门指机关团体事业单位和国有及国有控股企业;集体部门指集体企业;私营部门指个体工商户、私营企业、其它类型单位(包括外商投资和港、澳、台投资企业、民办非企业单位,等等)。

个体变量和地区变量的基本信息见表1。

表1 变量的描述性统计(均值)

	个体变量		地区变量	
	地市层面的分析	省层面的分析	地市层面的分析	省层面的分析
月工资	1272.23	1267.03	劳动者平均受教育年限	12.01 12.14
女性	0.4276	0.4275	大学学历劳动者比例	0.3772 0.3899
年龄	37.00	37.01	第一产业劳动者比例	0.0204 0.0207
初中及以下	0.2708	0.2717	第二产业劳动者比例	0.3122 0.3174
高中	0.3482	0.3480	第三产业劳动者比例	0.6674 0.6619
大学	0.3810	0.3803	国有部门劳动者比例	0.6814 0.6684
受教育年限	12.16	12.15	集体部门劳动者比例	0.0571 0.0554
工作经验	18.84	18.85	私营部门劳动者比例	0.2615 0.2762
样本数	232052	235120	样本数	318 30

另外,在估计地市层面的教育外部收益率时使用了三个工具变量:一是2000年居民的平均受教育年限,318个地市的均值为7.53年;二是2000年6岁以上人口中大学学历人口比例,均值为0.0332;三是地市行政级别,取值为2的是三个直辖市(北京、天津、上海),取值为1的为32个地市(包括重庆、5个

计划单列市和 26 个省会城市),取值为 0 的为 283 个地市。在估计省级层面的教育外部收益率也使用了三个工具变量:一是 2000 年居民的平均受教育年限,30 个省的均值为 7.67 年;二是 2000 年 6 岁以上人口中大学学历人口比例,均值为 0.0390;三是 1981-2002 年平均的大学入学机会指数,均值为 1.1135。

四、我国城镇地区教育外部收益率的估计结果

(一) 地市层面的估计结果

根据式(4)估计我国地市层面的教育外部收益率的结果见表 2。在 IV 估计中,分别使用了三个工具变量:IV1 表示 2000 年居民平均受教育年限;IV2 表示 2000 年 6 岁以上人口中大学学历人口比例;IV3 表示地市行政级别。可以看出,所有回归系数均至少在 1% 的水平上是显著的。在 IV 估计中,第一阶段回归结果表明,工具变量对内生变量(劳动者平均受教育年限或大学学历劳动者比例)的影响都是统计显著的。不可识别检验(underidentification test)和弱识别检验(weak identification test)的结果表明,模型是可识别的而且不存在弱工具变量问题。内生性检验(endogeneity test)结果表明可以拒绝劳动者平均受教育年限或大学学历劳动者比例是外生的假定,即这两个变量是内生变量。

表 2 教育外部收益率的估计结果:地市层面(N=232052)

	OLS (1)	IV			OLS (5)	IV		
		IV1 (2)	IV2 (3)	IV3 (4)		IV1 (6)	IV2 (7)	IV3 (8)
劳动者平均受教育年限	0.2534 (0.0063)	0.3499 (0.0095)	0.3746 (0.0123)	0.3522 (0.0117)				
大学学历劳动者比例					1.7179 (0.0452)	3.2508 (0.0881)	3.5638 (0.1175)	3.1547 (0.1050)
个人受教育年限	0.1246 (0.0004)	0.1242 (0.0004)	0.1241 (0.0004)	0.1242 (0.0004)	0.1247 (0.0004)	0.1240 (0.0004)	0.1239 (0.0004)	0.1241 (0.0004)
第一阶段回归中工具变量的估计系数								
2000 年居民平均受教育年限		0.3951 (0.0009)				0.0425 (0.0001)		

续表

	IV			OLS (5)	IV			
	OLS (1)	IV1 (2)	IV2 (3)		IV3 (4)	OLS (5)	IV1 (6)	IV2 (7)
2000年大学学历人口比例		6.5797 (0.0231)				0.6916 (0.0034)		
地市行政级别		0.5006 (0.0016)				0.0559 (0.0002)		
工具变量估计中的相关检验								
不可识别检验	101761	60056	66419		61481	34610	43183	
弱识别检验	181207	81012	93038		83626	40670	53047	
内生性检验	186	131	100		414	292	231	

注:1. 模型(1) - (4)用劳动者平均受教育年限衡量地市平均受教育水平,模型(5) - (8)用大学学历劳动者比例衡量地市平均受教育水平。2. 所有模型还包括以下解释变量:个体层面:性别、工作经验、工作经验的平方;地市层面:第二产业就业的劳动者比例、第三产业就业的劳动者比例、国有部门就业的劳动者比例、私营部门就业的劳动者比例;省份虚拟变量。表中略去这些变量的回归系数。3. 括号内为标准误。4. 所有系数和统计量均至少在1%的水平上统计显著。5. 不可识别检验使用 Anderson canonical correlations test LM 统计量,弱识别检验使用 Cragg - Donald Wald F 统计量,内生性检验使用 C 统计量(也称为 difference - in - Sargan 统计量)。6. 在所有回归分析中均以地市的个体样本数作为权重进行加权。

表2显示,在不同的模型设定中,教育的私人收益率几乎不变,约为12.4%。当以劳动者平均受教育年限衡量一个地市的劳动者平均受教育水平时,三个工具变量的估计系数差异不大,在0.35 - 0.38之间。这意味着平均而言,一个城市的平均受教育年限增加一年,将导致在该城市工作的劳动者的工资提高约35% - 38%。OLS估计值为0.2534,小于IV估计值。

当以大学学历劳动者比例衡量一个地市的劳动者平均受教育水平时,三个工具变量的估计系数同样差异不大,在3.1 - 3.6之间。意味着平均而言,一个城市的大学学历劳动者比例提高一个百分点,将导致在该城市工作的劳动者的工资提高约3.1% - 3.6%。OLS估计值为1.7179,大约是IV估计值的二分之一。

在使用微观数据估计我国教育外部收益率的已有研究中,Liu(2007)使用的方法与本文最为接近。他使用CHIP1995年的数据,以劳动者平均受教育年

限衡量城市平均受教育水平,得到教育外部收益率的 OLS 估计值为 6%,IV 估计值为 11% - 13%。以大学学历劳动者比例衡量城市平均受教育水平,得到教育外部收益率的 OLS 估计值为 1.276,IV 估计值为 0.997 - 1.551。可见,从 1995 年到 2005 年,我国的教育外部收益率有了大幅提高。

考虑到直辖市的人口规模、经济规模和地域面积大于一般的地市,有可能对估计结果造成影响。本文还使用相同的模型,基于不包括四个直辖市的样本进行了估计。结果见表 3。可以看出,如果使用不含直辖市的样本,教育外部收益率的估计结果几乎没有变化,表明估计结果是稳健的。

表 3 使用不同地市样本的教育外部收益率

衡量地市平均教育 水平的变量	全体地市(N = 232052)				不含直辖市(N = 180154)			
	OLS	IV1	IV2	IV3	OLS	IV1	IV2	IV3
劳动者平均受教育年限	0.2534	0.3499	0.3746	0.3522	0.2627	0.3559	0.3755	0.3541
大学学历劳动者比例	1.7179	3.2508	3.5638	3.1547	1.7846	3.3152	3.5808	3.1771

注:所有回归系数均至少在 1% 的水平上显著。

(二) 省级层面的估计结果

根据式(4)估计我国省级层面的教育外部收益率的结果见表 4。

在 IV 估计中,分别使用了三个工具变量:IV1 表示 2000 年居民平均受教育年限;IV2 表示 2000 年 6 岁以上人口中大学学历人口比例;IV3 表示 1981 - 2002 年平均的大学入学机会指数。同样地,所有回归系数均在 1% 的水平是显著的。在 IV 估计中,第一阶段回归的结果表明,工具变量对内生变量的影响都是统计显著的。不可识别检验、弱识别检验和内生性检验的结果同样表明,模型是可识别的,不存在弱工具变量问题,可以拒绝劳动者平均受教育年限(或大学学历劳动者比例)是外生的假定。

表 4 显示,在不同的模型设定中,教育的私人收益率几乎不变,约为 12.7%。当以劳动者平均受教育年限衡量一个省的平均受教育水平时,三个工具变量的估计系数差异不大,在 0.43 - 0.46 之间。这意味着平均而言,一个省的平均受教育年限增加一年,将导致在该省工作的劳动者的工资提高约 43 - 46%。OLS 估计值为 0.4291,略小于 IV 的估计值。如果以大学学历劳动者比例衡量一个省的平均受教育水平,那么工具变量的估计系数在 5.2 - 6.1 之间。意味着平均而言,一个省的大学学历劳动者比例提高一个百分点,将导致在该省工作的劳动者的工资提高约 5.2% - 6.1%。OLS 估计值为 3.7168,低于 IV 估计值。

将表2和表4进行比较可以发现,省级层面的教育外部收益率高于地市层面。也就是说,教育的外部性在较大的地理区域内会更强。对于这一结果,笔者认为有两个解释。其一,人力资本的外部性来源于劳动者之间进行的知识和经验的交流(Lucas,1988),因此,地理范围越大,劳动者之间的异质性就越强,从交流互动中获得收益就越大,外部性就越强。另外,在一个更大的地理范围内,市场竞争更激烈,促使个体有更大的激励去学习,劳动者之间的交流会更频繁,从而产生更大的外部性。其二,人力资本理论的创始人之一舒尔茨认为,教育的经济价值主要在于能够提高个人“处理不均衡状态的能力(the ability to deal with disequilibria)”或配置能力(allocative ability),即能够根据外部条件的变化(如技术变革、制度转型、产业升级、供需变化等)重新配置资源,从而实现自身利益最大化的能力。并指出:“有充分的证据表明教育能够提高个人成功地处理经济不均衡的能力,这是教育在现代经济中的主要收益之一”(Schultz,1975)。而配置能力的实现很大程度上取决于市场半径(赖德胜,2011),也就是说,市场半径越大,经济机会就越多,生产要素的流动就越充分,给配置能力的发挥留下了更大的空间。地区平均受教育水平的提高将增强地区内劳动者整体的配置能力,而这种配置能力在更大的地理区域内能够得到更充分的实现,因此教育的外部性在更大的地理区域内会更强。

表4 教育外部收益率的估计结果:省级层面(N=235120)

	OLS (1)	IV			OLS (5)	IV		
		IV1 (2)	IV2 (3)	IV3 (4)		IV1 (6)	IV2 (7)	IV3 (8)
劳动者平均受教育年限	0.4291 (0.0053)	0.4377 (0.0066)	0.4451 (0.0062)	0.4565 (0.0064)				
大学学历劳动者比例					3.7168 (0.0512)	6.1366 (0.0930)	5.2037 (0.0727)	5.2709 (0.0736)
个人受教育年限	0.1271 (0.0005)	0.1270 (0.0005)	0.1270 (0.0005)	0.1268 (0.0005)	0.1280 (0.0005)	0.1260 (0.0005)	0.1268 (0.0005)	0.1267 (0.0005)
第一阶段回归中工具变量的估计系数								
2000年居民平均受教育年限		0.2578 (0.0004)				0.0184 (0.0001)		

续表

	OLS (1)	IV			OLS (5)	IV		
		IV1 (2)	IV2 (3)	IV3 (4)		IV1 (6)	IV2 (7)	IV3 (8)
2000年大学学历人口比例		5.4667 (0.0067)				0.4676 (0.0010)		
大学入学机会指数		0.2236 (0.0003)				0.0194 (0.0000)		
工具变量估计中的相关检验								
不可识别检验		154195	174399	165934		72019	117142	114275
弱识别检验		447980	675266	563878		103815	233444	222329
内生性检验		24.98	25.93	63.44		985	836	870

注:所有模型还包括以下解释变量:个体层面:性别、工作经验、工作经验的平方;省级层面:第二产业就业的劳动者比例、第三产业就业的劳动者比例、国有部门就业的劳动者比例、私营部门就业的劳动者比例。表中略去这些变量的回归系数。在所有回归分析中均以省份的个体样本数作为权重进行加权。其他注释同表2。

(三) 不同受教育程度劳动者的教育外部收益率

相关研究普遍关注的一个问题是不同受教育程度劳动者获得的教育外部收益率。Liu(2007)指出,一个地区劳动者的平均教育水平提高会有两种效应:其一,地区总体生产率提高,从而各类受教育程度劳动者的工资都会提高。其二,平均受教育水平提高意味着高教育水平劳动者的比例增加而低教育水平劳动者的比例下降,由于两类劳动者具有不完全替代性,那么高教育水平劳动者的供给增加会压低其工资水平,而低教育水平劳动者的供给减少会提高其工资水平。对于低教育水平劳动者,两种效应都会导致其工资上升。对于高教育水平者,前一种效应导致其工资上升而后一种效应导致其工资下降,因此教育的外部收益率可能为正也可能为负。他利用CHIP1995年的调查数据,发现高中及以上学历劳动者和高中以下学历劳动者的教育外部收益率都为正,但前者明显低于后者。李小瑛等(2010)和徐大丰(2010)利用CHNS1989-2006年的混合截面数据,也发现教育的外部收益率对于学历较低者较高,对学历较高者较低。

根据式(4),本文分样本估计了高中及以下学历者和大学学历者两类人群的教育外部收益率,结果见表5。可以看出,无论使用哪一个变量衡量地区平

均教育水平,无论是地市层面还是省级层面的分析,无论使用 OLS 估计还是 IV 估计,大学学历劳动者的教育外部收益率均高于高中及以下学历的劳动者,即教育的外部收益率对于高学历者而言较高。这一发现与之前的研究结论不同,笔者认为主要原因是 Liu(2007)使用的是 1995 年的数据,李小瑛等(2010)和徐大丰(2010)使用的是 1989-2006 年的混合截面数据,不能反映进入二十一世纪以后我国的教育外部收益率。进入二十一世纪以后,一方面,知识经济对劳动者的技能要求提高,高技能劳动者更能从总体生产率的提高中获益。另一方面,我国在经济发展中存在技能偏向型技术进步,对技能型劳动的相对需求增加(宋东林等,2010),抵消了高技能劳动者供给增加的影响。从而在劳动者平均受教育水平提高的情况下,高学历劳动者的工资增加得更快。事实上,在徐大丰(2010)利用 2006 年横截面数据的分析中,也发现大学学历劳动者的教育外部收益率高于高中及以下学历劳动者,而且后者是统计不显著的。

表 5 不同受教育程度劳动者的教育外部收益率

	以劳动者平均受教育年限 衡量地区平均教育水平				以大学学历劳动者比例 衡量地区平均教育水平			
	OLS	IV1	IV2	IV3	OLS	IV1	IV2	IV3
地市层面								
全体劳动者	0.2534	0.3499	0.3746	0.3522	1.7179	3.2508	3.5638	3.1547
高中及以下学历劳动者	0.2314	0.3026	0.3183	0.3064	1.6203	2.7118	2.8421	2.6053
大学学历劳动者	0.2870	0.4916	0.5970	0.5172	1.7086	5.2733	6.9836	5.4842
省级层面								
全体劳动者	0.4291	0.4377	0.4451	0.4565	3.7168	6.1366	5.2037	5.2709
高中及以下学历劳动者	0.4178	0.4059	0.4076	0.4191	3.5485	5.8771	4.7808	4.8136
大学学历劳动者	0.4719	0.4715	0.4895	0.5083	4.1069	6.5547	5.7719	5.9642

注:所有回归系数均至少在 1% 的水平上显著。

(四) 基于 Moretti(2004) 方法的估计结果

Moretti(2004)在根据美国人口普查数据估计城市层面的教育外部收益率时,认为人口普查数据的个体样本量很大,应该使用两阶段的估计策略。笔者认为这一策略在理论上不甚严谨,也缺乏直观性,但为了检验前述估计结果的稳健性,本文使用这一策略估计了地市层面的教育外部收益率。^① 具体过

^① 如果使用这种策略估计省级层面的教育外部收益率,则在第二阶段只有 30 个省份样本,估计缺乏精确度,故而只用这种策略进行地市层面的估计。

程为:

第一阶段使用个体数据,估计经过回归调整的地市平均工资(regression-adjusted mean wage):

$$\ln W_{ij} = \varphi_j + \alpha S_{ij} + \beta X_{ij} + \varepsilon_{ij} \quad (5)$$

在式(5)中, $\ln W_{ij}$ 、 S_{ij} 、 X_{ij} 的含义同式(4), ε_{ij} 为随机误差项。 φ_j 是表示个体所在城市的虚拟变量,可解释为经过回归调整的地市平均工资。

第二阶段使用地市层面的数据,以第一阶段的估计值 $\hat{\varphi}_j$ 为因变量,对地市层面的变量进行回归:

$$\hat{\varphi}_j = \pi + \gamma \bar{S}_j + \theta G_j + \rho P_j + v_j \quad (6)$$

在式(6)中, π 为截距项, v_j 为随机误差项, \bar{S}_j 、 G_j 的含义同式(4), P_j 为表示地市所在省份的虚拟变量。 γ 即为教育的外部收益率。在两个阶段的回归估计中,均以地市的个体样本数作为权重进行加权。

表6 地市层面教育外部收益率的估计结果(基于 Moretti (2004)的方法)

	以劳动者平均受教育年限衡量 地市平均教育水平(N=318)				以大学学历劳动者比例衡量 地市平均教育水平(N=318)			
	OLS	IV1	IV2	IV3	OLS	IV1	IV2	IV3
全体劳动者	0.2275	0.2879	0.2983	0.2669	1.5751	2.6870	2.5598	2.2175
高中及以下学历劳动者	0.2033	0.2509	0.2598	0.2301	1.4081	2.3005	2.2095	1.8890
大学学历劳动者	0.3069	0.3823	0.4015	0.3576	2.0144	3.7047	3.5112	3.0374

注:所有回归系数均至少在1%的水平上显著。

基于式(5)和式(6)得到的地市层面的教育外部收益率的估计结果见表6。与表(5)的结果相比,两阶段估计策略的教育外部收益率要小一些,但仍然是统计显著的,而且同样发现大学学历劳动者的教育外部收益率更高。

五、结语

本文使用2005年全国1%人口抽样调查数据,在微观明瑟方程中加入代表地区平均教育水平的变量,使用工具变量法估算我国地市层面和省级层面的教育外部收益率。主要结论是:第一,在我国城镇地区存在明显的教育外部性,工具变量法的估计结果显示,城市劳动者平均受教育年限提高一年,将导致在该城市工作的劳动者工资提高35% - 38%,而城市劳动者中大学学历劳动者

的比例提高一个百分点将导致劳动者工资提高 3.1% - 3.6%。第二,省级层面的教育外部收益率高于地市级层面。第三,高学历劳动者获得的教育外部收益率高于低学历劳动者。

本文发现教育存在正的外部收益,表明了政府投资教育的合理性。特别是高等教育劳动者比例的增长有助于提高劳动者的工资收入,表明了政府投资高等教育的正当性。当然,本文考虑的仅仅是教育的外部性中的货币性收益,而事实上教育的外部性中还包括非货币性的收益,如推进民主、减少犯罪、提高居民整体健康水平,等等。考虑到这些非货币性的外部收益,政府应该在教育投资中发挥更大的作用。另外,省级层面的教育外部收益率大于地市级层面,说明从纠正教育外部性的角度看,省级政府应该比地市政府承担更大的教育责任。进入二十一世纪以后,我国政府一直致力于扭转基础教育过于依赖基层政府的状况,强调省级政府对基础教育发展特别是义务教育发展的责任,本文的研究结论支持了这一政策取向。本文还发现教育的外部收益率对于高学历劳动者更高,意味着在知识经济和技能偏向型技术进步的背景下,高技能劳动者更有能力享受到技术进步和生产率提高带来的外溢性,从而教育扩张将拉大高技能劳动者和低技能劳动者的工资差距。已有的一些研究发现,自二十世纪八十年代末期以来,我国城镇地区的教育收益率持续上升(邓峰和丁小浩,2013),而且高等教育的收益率也在持续上升(刘泽云,2015)。本文的发现可以在一定程度上解释上述研究的结论。

遗憾的是,笔者没有获得 2010 年人口普查数据,否则会使得研究更具有现实意义,而且将 2005 年与 2010 年的数据结合起来构成混合截面数据,将得到更为准确的估计。希望在将来有机会获取相关数据以改进现有研究。

[参考文献]

- 邓峰、丁小浩,2013:《中国教育收益率的长期变动趋势分析》,《统计研究》第7期。
- 弗里德曼,1986:《资本主义与自由》,张瑞玉译,北京:商务印书馆。
- 国务院人口普查办公室、国家统计局人口和社会科技统计司,2003:《2000 人口普查分县资料》,北京:中国统计出版社。
- 赖德胜,2011:《教育、劳动力市场与创新型人才的涌现》,《教育研究》第9期。
- 李立峰、刘海峰,2007:《中国高校招生考试中的区域公平研究》,武汉:华中师范大学出版社。
- 李小瑛、陈广汉、张应武,2010:《中国城镇地区高等教育外部回报率估算》,《世界经济文汇》第1期。

- 刘泽云,2015:《上大学是有价值的投资吗?——中国高等教育回报率的长期变动》,《北京大学教育评论》第4期。
- 宋东林、王林辉、董直庆,2010:《技能偏向型技术进步存在吗?——来自中国的经验证据》,《经济研究》第5期。
- 徐大丰,2010:《社会教育程度提高的工资效应》,《上海经济研究》第5期。
- Acemoglu, D., 1996, "A Microfoundation for Social Increasing Returns in Human Capital Accumulation", *Quarterly Journal of Economics*, 113 (3), 779-804.
- Bauer, T. K. and M. Vorell, 2010, "External Effects of Education: Human Capital Spillovers in Regions and Firms", *Ruhr Economic Papers*, No. 195.
- Breton, T. R., 2010, "Schooling and National Income: How Large are the Externalities?", *Education Economics*, 18(1), 67-92.
- Ciccone, A. and G. Peri, 2006, "Identifying Human Capital Externalities: Theory with Applications", *The Review of Economic Studies*, 73(2), 381-412.
- Dalmazzo, A. and G. Blasio, 2007, "Production and Consumption Externalities of Human Capital: An Empirical Study for Italy", *Population Economics*, 20, 359-382.
- Fan, W. and Y. Ma, 2012, "Estimating the External Returns to Education: Evidence from China". Working Paper, http://www.ucd.ie/t4cms/WP12_20.pdf.
- Kirby, S. and R. Riley, 2008, "The External Returns to Education: UK Evidence using Repeated Cross - sections", *Labor Economics*, 15, 619-630.
- Lange, F. and R. Topel, 2006, "The Social Value of Education and Human Capital", Hanush-ek, E. A. and Welch, F. (eds.) *Handbook of the Economics of Education* (Volume 1). Amsterdam: Elsevier. 460-506.
- Liu, Z., 2007, "The External Returns to Education: Evidence from Chinese Cities", *Journal of Urban Economics*, 61, 542-564.
- Lucas, R. E., 1988, "On the Mechanics of Economic Development", *Journal of Monetary Economics*, 22, 3-42.
- McMahon, W. W., 2010, "The External Benefits of Education", Brewer, D. J and McEwan, P. J. (eds.) *Economics of Education*, Amsterdam: Elsevier. 68-79
- Moretti, E., 2004, "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross - sectional Data", *Journal of Econometrics*, 121, 175-212.
- Muravyev, A., 2008, "Human Capital Externalities: Evidence from the Transition Economy of Russia", *Economics of Transition*, 16(3), 415-443.
- Rauch, J. E., 1993. "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities", *Journal of Urban Economics*, 34(1), 380-400.
- Schultz, T. W., 1975, "The Value of the Ability to Deal with Disequilibria". *Journal of Economic Literature*, 13(3), 827-846.