

教育收益率对中国收入差距的影响

——基于分位数回归模型的分析

张传国, 晋媛媛

(厦门大学经济学院, 福建 厦门 361005)

[摘要]采用中国家庭追踪调查数据(CFPS2014),运用分位数回归模型探讨教育收益率对中国收入差距的影响。研究表明:(1)随着收入分位点的提高,中国的教育收益率逐渐下降,教育收益率有助于缩小中国不同群体的收入差距;(2)配偶的受教育年限对低收入个体的影响更大;(3)城镇的教育收益率高于农村,教育收益率的城乡效应将加剧城乡收入差距;(4)东部地区的教育收益率高于中西部地区,教育收益率的区域效应将加剧区域收入差距。总的来讲,教育收益率虽然会缩小不同群体的收入差距,但会扩大城乡和区域收入差距。

[关键词]教育收益率;收入差距;受教育年限;人力资本积累;劳动力市场;教育回报率

[中图分类号] F832.51;F062.2 **[文献标志码]** A **[文章编号]** 2096-3114(2020)02-0091-11

一、引言

当前,中国收入差距问题主要表现为不同收入群体、城乡和区域之间的收入差距不断扩大。1985—2014年中国的基尼系数在同一时期从0.38上升到0.469,超过了0.4这一收入分配差距的“警戒线”^[1],其中城乡收入差距和区域收入差距是中国收入差距的重要组成部分,占到总体收入差距的70%以上。从城乡来看,中国目前的城乡收入之比是3:1,是世界上城乡收入比例较高的国家之一;从区域来看,2014年东部地区人均国内生产总值为67109元,中部地区人均国内生产总值为41523元,西部地区人均国内生产总值为37487元,区域差距十分明显。同时,在制造业高度集中的中国,2014年东部地区9个沿海省区市的制造业产值就占到了全国的75%。

随着中国的市场化改革和经济发展,收入的决定因素发生了深刻的变化,教育在收入中的作用越来越重要,成为影响收入差异的重要因素。中国政府一直将促进教育公平作为国家基本教育政策,努力使教育体现出最大的社会公平。然而,我们需要关注这样一个问题:即使贫困学生获得适当的教育,但是如果他们在教育上的投资无法得到足够的回报,那么收入差距情况也依然不会得到改善,甚至会恶化。因此,我们需要关注教育收益率在收入分配领域的作用。现有研究多数采用分位数回归模型来估计教育收益率对收入差距的影响,这种方法可以用来估计收入分布上不同群体的教育收益率。假设教育收益率随着收入分位点的提高而逐渐下降,那么相较于高收入群体,低收入群体通过相同的教育投资可以获得更多的收益,教育收益率会起到减小收入差距的作用;反之,教育收益率将扩大收入差距。

国外学者采用分位数回归模型得到的实证结果显示,教育收益率往往呈现出随条件收入分位点上升而增加的特征,教育收益率与收入不平等之间存在正向关系^[2-4]。国外学者普遍将这一现象归因于能力与

[收稿日期] 2019-11-04

[基金项目] 中央高校基本科研业务经费项目(20720140001)

[作者简介] 张传国(1972—),男,山东临沂人,厦门大学经济学院教授,博士生导师,主要研究方向为区域经济,邮箱:cgzhang@xmu.edu.cn;晋媛媛(1994—),女,安徽芜湖人,厦门大学经济学院硕士生,主要研究方向为教育经济。

教育存在互补关系,他们将收入分位点解释为个人的能力水平,处于收入分位点上半部分的群体往往能力也较强,可以从教育中获得更多的边际收益,教育收益率也相对较高^[5]。然而,这种解释隐含着能力与收入之间存在正向关系,这种正向关系会导致严重的内生性问题,降低解释的可信性。另一种可能的解释是过度教育,发达国家的教育体系可以提供工作中所需的知识和技能,在这种情况下教育在高能力群体中显得十分重要,因此在收入分位点较高的地方教育收益率也较高^[2]。这些解释与发达国家的教育体制密切相关,发达国家实行开放式教育,崇尚对学生实践和动手能力的培养,这种类型的教育通常可以提供工作中所需要的技能,能够帮助能力强的群体在未来工作中获得更高的收入。国内学者罗楚亮的研究表明,中国的教育收益率随着条件收入分位点的上升而下降,并且就业特征是教育收益分布特征不同的根本原因^[6]。赵显洲研究发现,中国东部地区的教育收益率普遍高于中西部地区,并且能力的异质性、生产率差异、劳动力市场发育程度差异这三个因素是教育收益率区域分布产生差异的根本原因^[7]。中国现代教育体制与发达国家存在较大差异,因此教育收益率分布特征以及背后的原因也可能与发达国家存在较大差异。

与以往研究相比,本文可能有以下四个方面的边际贡献:第一,本文研究中国教育收益率与收入差距的关系,并详细讨论中国教育收益率对收入差距产生影响的根本原因,有助于补充该领域研究的空白。第二,本文在分位数回归模型中加入能力这一代理变量,控制教育的内生性问题,同时考虑教育收益率的异质性和教育的内生性问题,有助于真正揭示教育收益率对收入差距的因果效应。第三,考虑到中国一般以家庭为单位进行投资和消费,本文的被解释变量不仅指户主工资性收入,还包括户主在家庭中的财产性收入和经营性收入,因而能够更加全面地衡量收入差距。第四,中国幅员辽阔,在教育投资和经济发展方面存在明显的区域差异,教育收益率对城乡和区域收入差距也会产生影响,因此对城乡和区域的教育收益率进行比较研究有助于指导城乡和区域的教育投资。

二、理论分析

人力资本作为生产要素之一,符合新古典经济学的要素边际效用递减理论。相较于高收入群体,低收入群体受教育水平相对较低,通过教育投资来促进收入增长的潜力更大。由于这种潜力存在一定的差异,因此低收入群体会选择通过增加教育投资来提高人力资本,进而获得更高的收入,教育收益率提高。教育对收入的边际贡献随着收入的上升而下降,如果一个要素对低收入群体收入的促进作用大于中高收入群体,那么这种要素就具有缩小收入差距的作用^[8]。

本文将收入看作是人力资本的函数,并把除受教育水平之外其他影响收入的因素放到残差项中,那么收入函数^[9-10]可以表示为:

$$\ln(Y_i) = f(S_i) \quad (1)$$

其中, S_i 代表个人的受教育年限, Y_i 代表个人的收入水平。由前文对收入差距的定义可知,对数收入差距的方差可以用来衡量收入差距,为了体现教育收益与收入差距之间的关系,本文对方程(1)两边同时取方差。此时,教育收益与收入差距的关系就可以表示为:

$$\text{Var} \ln(Y_i) = M^2 \text{Var}(S_i) \quad (2)$$

在方程(2)中, M 代表教育投资的市场收益,在完善的劳动力市场中 M 是常量。然而,中国劳动力市场存在一定程度的分割,教育投资的收益可能不是常量,它也许还会因职业、城乡、区域等的不同而不同。假如放宽对劳动力市场假设的限制,方程(1)可以改写为如下形式:

$$\ln(Y_i) = \theta [M_i f(S_i)] \quad (3)$$

其中, S_i 代表个人的受教育年限, Y_i 代表个人的收入水平, M_i 为教育收益率变量。假设教育水平和教

育收益率之间是相互独立的,且 $E(S_i) = \mu_s$ 、 $E(M_r) = \mu_m$,那么对方程(3)两边同时取方差可以得到下列等式:

$$\text{Var} \ln(Y_{ir}) = \mu_m^2 \text{Var}(S_{ir}) + \mu_s^2 \text{Var}(M_r) + \text{Var}(S_{ir}) \text{Var}(M_r) \quad (4)$$

在方程(4)中, μ_m 为教育收益率的期望值, μ_s 为受教育水平的期望值。从该公式中可以看出,收入差距不仅与受教育水平的差异 $\text{Var}(S_{ir})$ 和教育收益率的差异 $\text{Var}(M_r)$ 有关,而且和受教育水平的均值 μ_s 及教育收益率的均值 μ_m 有关。

从方程(4)中可以看到,如果教育投资的收益为零,即 $M = 0$,或者说教育无法改变收入水平,则 $\text{Var} \ln(Y_i) = 0$,在此种情况下,教育与收入差距之间不存在相关关系。如果教育收益率不为零,即 $M \neq 0$,且在 M 是一个变量的条件下,那么收入差距不仅和受教育水平及其差异有关,还和教育收益的大小及其差异有关。

综上,在给定受教育水平及其差异的情况下,收入差距由教育收益率及其差异所决定,劳动力市场中教育收益率越高,收入差异就越大;同时,如果教育收益率为一个变量,那么教育收益率的差异程度越大,收入差距也会越大。

Knight 和 Sabot 指出,教育对收入差距存在两种不同的作用:第一种是结构效应(Composition Effect),它假设教育收益率为常数,则教育结构变化使得接受更多教育的人拥有更高的生产和配置能力,可以在劳动力市场中获得更高的收入,进而导致接受较多教育与接受较少教育的人之间的收入差距扩大;第二种是压缩效应(Compression Effect),它是指在教育结构保持不变的情况下,教育收益率随着教育相对供给的增加而降低,即教育溢价随着教育供给相对于需求的增加而下降,这会导致收入差距的减小。压缩效应和结构效应的相对大小会影响教育对收入差距作用的方向^[11]。

当存在教育收益率差异时,不同居民的教育投资需求就会存在差异,这会影响居民对教育的投入,且教育投资差异会转化为人力资本差距,进而导致收入差距的出现。为了保持这种相对高收入水平,高收入群体会加大对教育的投资,人力资本差异扩大会进一步增大收入差距。如此一来,就形成了收入差距的“马太效应”。

由以上理论分析可知,深入分析教育收益率及其对中国收入差距的影响具有重要的理论价值。接下来本文将从四个方面实证分析教育收益率对中国收入差距的具体影响。

三、研究设计

(一) 样本选取与数据来源

本文数据来源于中国家庭动态跟踪调查(CFPS)2014年抽样调查数据。CFPS是由北京大学中国社会科学调查中心组织发起的,通过追踪个人、家庭和社区的信息来反映中国居民的经济、教育、家庭和健康的变化情况。目前CFPS数据更新到2014年,调查对象覆盖中国东部地区、中部地区和西部地区28个省区市,包含城镇居民和农村居民的相关信息,这便于本文研究城乡和区域收入差距。数据总共涵盖了16000户家庭和69848位居民的信息。

(二) 变量选取

1. 被解释变量:收入,包括工资性收入、财产性收入和经营性收入。明瑟方程一般以工资收入作为被解释变量来衡量教育带来的经济回报。然而,教育不仅可以提高劳动者的工资性收入,还会对劳动者的市场洞察力和投资能力产生重要影响^[12]。个体的市场洞察力和投资能力有助于提高其经营性收入和财产性收入。因此,为了更加全面地衡量教育对收入的影响,本文将居民的财产性收入和经营性收入加入被解释变量。

2. 解释变量:受教育年限。为了对受教育程度进行定量分析,本文将2014年CFPS数据中受教育程度

转化为受教育年限, 未上过学为0年、小学为6年、初中为9年、高中/中专为12年、大学专科为14年、大学本科为16年、硕士研究生为19年、博士研究生为22年。

3. 控制变量: 性别、健康状况、婚姻状况、家庭资产配置情况、城乡和区域。首先, 本文将常见的性别、健康状况、婚姻状况等作为控制变量加入方程中。其次, 为了更好地衡量教育对非工资性收入的影响, 本文将家庭资产配置的有关变量也纳入控制变量之中。最后, 为了消除劳动力市场因素对估计结果的影响, 本文在方程中加入与劳动力市场有关的虚拟变量, 如城乡和区域等, 以控制劳动力市场分割的影响。

变量定义见表1。

表1 变量定义

变量名称	变量代码	变量定义
收入	<i>Income</i>	年收入=工资性收入+(家庭经营收入和家庭财产性收入)/家庭人口
户主受教育年限(年)	<i>Householder's education</i>	在学校接受的教育年数
配偶受教育年限(年)	<i>Spousal education</i>	在学校接受的教育年数
工作经验	<i>Experience</i>	工作经验=年龄-受教育年限-6
家庭人口学特征变量		
是否未婚	<i>Married</i>	已婚取1, 未婚取0
性别	<i>Male</i>	男性取1, 女性取0
是否健康	<i>Health</i>	健康状况为极好、好和一般取1, 其他情况取0
家庭人口数(人)	<i>Family population</i>	户主所在家庭人口数
是否拥有自住房	<i>Owner-occupied house</i>	是取1, 否取0
是否拥有投资性住房	<i>Investment housing</i>	是取1, 否取0
家庭资产负债情况		
是否参与金融市场	<i>Participation in the financial market</i>	是取1, 否取0
是否有待偿房贷	<i>Unpaid mortgage</i>	是取1, 否取0
是否有非住房贷款	<i>Non-housing loans</i>	是取1, 否取0
是否借钱给别人	<i>Lends money to others</i>	是取1, 否取0
金融市场投资品种	<i>Financial market investment</i>	投资金融产品种类
是否城镇人口	<i>Urban</i>	是取1, 否取0
区域情况		
东部地区虚拟变量	<i>Eastern region</i>	东部地区: 河北省、北京市、天津市、山东省、江苏省、浙江省、上海市、广东省、海南省、福建省
中西部地区虚拟变量	<i>Midwestern region</i>	中西部地区: 山西省、河南省、安徽省、湖北省、江西省、湖南省、陕西省、四川省、云南省、贵州省、广西壮族自治区、甘肃省、青海省、宁夏回族自治区、西藏自治区、新疆维吾尔自治区、内蒙古自治区和重庆市

(三) 模型构建

现有研究一般采用Mincer方程来分析教育对收入的影响, 该模型是由Mincer在1974年提出的^[9]。

$$\ln(\text{wage}_i) = \beta_0 + \beta_1 \times \text{edu}_i + \beta_2 \times \text{exp}_i + \beta_3 \times \text{exp}_i^2 + \varepsilon_i \quad (5)$$

在该模型中, i 表示不同的个人 ($i = 1, 2, \dots, n$), $\ln(\text{wage}_i)$ 为小时工资的对数, β_0 为常数截距项, edu_i 为受教育年限, exp_i 为工作年限即工龄 (年龄-受教育年限-6), exp_i^2 为工作年限的平方, ε_i 为随机干扰项。 β_1 、 β_2 、 β_3 为各变量的回归系数, β_1 为平均的教育收益率, 表示在不考虑其他教育成本的情况下, 受教育年限每增加一年, 以小时工资为单位的人力资本相应增加的变化率, β_2 和 β_3 表示工作年限与收入之间的关系。该方程假定所有教育水平的收益率是相同的。

产业结构的升级和劳动力市场的发展促进了劳动者就业方式的多样化, 从而使得居民收入的来源趋于多元化, 不再仅仅依赖于工资性收入。国家统计局的数据显示, 2018年全国居民工资性收入占可支配收入的比重只有57.5%, 非工资性收入占比达到42.5%。因此为了更加全面地衡量教育对收入的影响, 本文将居民的财产性收入和经营性收入加入被解释变量中。被解释变量为户主收入, 计算方法如下:

$$Y_i = wage_i + (BI_i + PI_i) / N_i \tag{6}$$

其中, Y_i 表示不同家庭的户主收入 ($i=1, 2, \dots, n$), $wage_i$ 表示户主的工资性收入, BI_i 表示家庭经营性收入, PI_i 表示家庭财产性收入, N_i 表示家庭人口数。

本文将常见的性别、健康状况、婚姻状况和行业等变量作为控制变量加入方程中。在控制了其他不相关因素的影响后, 我们得到扩展的 Mincer 方程, 如式(7)所示:

$$\ln(Y_i) = \beta_0 + \beta_1 \times edu_i + \beta_2 \times exp_i + \beta_3 \times exp_i^2 + \sum_{m=1}^J \lambda_m X_m + \varepsilon_i \tag{7}$$

其中, $\ln(Y_i)$ 是不同家庭的户主收入取对数 ($i=1, 2, \dots, n$), X_m 是由其他控制变量组成的向量 ($m=1, 2, \dots, J$)。加入 $\sum_{m=1}^J \lambda_m X_m$ 可以控制除教育以外其他变量对收入的影响, 使得教育收益率的估计量更加接近真实值。

除了已经控制的可观测变量之外, 收入还会受到不可观测变量(如个人能力)的影响, 若在方程中忽略了能力变量就会导致教育收益率偏高, 同时考虑到工具变量法可能存在弱工具变量等问题, 故本文采用代理变量方法纠正遗漏变量带来的估计偏差。中国式婚姻关注的是门当户对的关系, 配偶的受教育程度可以在一定程度上反映对方的个人能力^[13]。因此, 本文选择配偶受教育年限作为个人能力的代理变量加入方程(7)中, 得到方程(8):

$$\ln(Y_i) = \beta_0 + \beta_1 \times edu_i + \beta_2 \times exp_i + \beta_3 \times exp_i^2 + \sum_{m=1}^J \lambda_m X_m + \gamma \times wedu_i + u_i \tag{8}$$

$wedu_i$ 表示不同家庭户主配偶的受教育年限 ($i=1, 2, \dots, n$)。 u_i 表示剔除能力因素后方程的残差。有文献曾采用配偶受教育年限、父母受教育年限作为工具变量来解决内生性问题^[14-15], 但考虑到本文的解释变量包含部分家庭收入, 而家庭收入可能与配偶和父母的受教育年限有关, 工具变量要求解释变量与被解释变量无关, 配偶受教育年限不适合作为工具变量加入本文的方程中。

方程(8)所估计的系数 β_1 只是平均意义上的教育收益率, 它假设所有人的教育收益率都是一样的。然而, 考虑到不同收入群体的教育收益率可能存在差异, 本文选择分位数回归模型, 这一模型可以显示教育收益率在整个收入分布中的信息。基于分位数回归的估计教育收益率的模型如下:

$$\ln w_i = x_i \beta_\theta + u_{\theta i}, Quant_\theta(\ln w_i | x_i) = x_i \beta_\theta \tag{9}$$

其中, x_i 是外生变量向量, β_θ 是参数向量。 $Quant_\theta(\ln w | x)$ 给定 x 条件下 $\ln w$ 的第 θ 个分位点, $\theta \in (0, 1)$, 可由下式解出:

$$\min_{\beta \in R^k} \left\{ \sum_{i: y_i \geq x_{\beta\theta}} \theta | \ln w_i - x_i \beta_\theta | + \sum_{i: y_i < x_{\beta\theta}} (1 - \theta) | \ln w_i - x_i \beta_\theta | \right\} \tag{10}$$

标准式为:

$$\min_{\beta \in R^k} \sum_i \rho_\theta(\ln w_i - x_i \beta_\theta) \tag{11}$$

$\rho_\theta(\varepsilon)$ 是检验函数, 如果 $\varepsilon \geq 0$, 则 $\rho_\theta(\varepsilon) = \theta \varepsilon$; 如果 $\varepsilon < 0$, 则 $\rho_\theta(\varepsilon) = (\theta - 1) \varepsilon$ 。

分位数回归的目标函数使用的是绝对值偏差的加权求和, 即 LAD 估计式 (Least Absolute Deviations, LAD)。分位数回归通过分位数 θ 取 0~1 之间的任何值, 让解释变量估计不同分位数的被解释变量, 得到解释变量对被解释变量的不同特定分位数的边际效果 β_θ 。为此, 本文建立以下条件分位数方程:

$$Quant_\theta(y|x) = x \text{ 其中 } \theta = \int_{-\infty}^{x\beta_\theta} f(y|x) dy \tag{12}$$

其中, $Quant_\theta$ 表示不同样本数据中 x 取得固定值的条件下 y 的分位数。 $f(y|x)$ 代表条件收入分布, 即不

同样本数据中 x 取得固定值情况下 y 的条件概率分布。分位数回归模型侧重于条件收入分布的不同分位数, 并假设某个条件收入分布的分位数随解释变量变化而变化。

综上所述, 修正的分位数回归模型同时考虑了教育的内生性和教育收益率的异质性问题, 有助于精确测量不同收入分位点的教育收益率, 从而可以判断在收入分布的上端和低端教育收益率的差异。

为了更清楚地说明教育收益率差异与不同群体收入差距的关系, 本文将收入分布上不同群体的收入进行分解: 群体1的收入水平为 $\ln Y_1 = \hat{\beta}_1 \overline{edu}_1 + \hat{\gamma}_1 \overline{X}_1 + \varepsilon_1$; 群体2的收入水平为 $\ln Y_0 = \hat{\beta}_0 \overline{edu}_0 + \hat{\gamma}_0 \overline{X}_0 + \varepsilon_0$, 其中 Y 代表收入, edu 代表受教育水平, X 代表除教育之外影响收入的其他变量。两组人群的收入差距就可以分解为 $\ln Y_1 - \ln Y_0 = \hat{\beta}_1 (\overline{edu}_1 - \overline{edu}_0) + (\hat{\beta}_1 - \hat{\beta}_0) \overline{edu}_0 + \hat{\gamma}_1 (\overline{X}_1 - \overline{X}_0) + (\hat{\gamma}_1 + \hat{\gamma}_0) \overline{X}_0$ 。显然, 除了平均教育水平差异 $(\overline{edu}_1 - \overline{edu}_0)$ 外, 教育收益率差异 $(\hat{\beta}_1 - \hat{\beta}_0)$ 也是影响收入差距的重要因素。即使在教育已经实现完全公平的情况下, 只要不同群体之间教育收益率存在差异, 并且教育收益率差异还在不断增加, 那么收入差距依然会进一步扩大。因此, 在收入分配领域研究不同群体之间的教育收益率差异是十分重要的。

在教育收益率的城乡和区域效应的分析中, 目前可以采用两种不同的方法比较城乡和区域之间的教育收益率。第一种方法将样本分成不同的组, 然后分别计算教育收益率^[16]; 第二种方法将分组变量添加到回归模型中。本文将采用第一种方法, 它可以更加清晰地显示城乡和区域各自的收益率, 这将有助于达到本文研究的目的。

四、实证结果与分析

(一) 描述性统计

表2是收入的描述性统计结果。收入的均值接近9.10, 方差为1.25, 偏度小于0, 说明收入分布存在左偏现象; 峰度大于3, 说明收入分布存在厚尾现象。通过Shapiro-Wilk进行正态性检验发现, 被解释变量不服从正态分布, 所以本文采用分位数回归模型更合适。

表2 被解释变量的描述性统计结果

变量	样本量	均值	方差	偏度	峰度	Shapiro-Wilk 正态性检验	SW检验P值
$\ln Y$	12071	9.10	1.25	-1.01	5.87	0.95	0.000

表3显示了各个解释变量的描述性统计结果。参与金融市场和拥有投资性住房的比例分别为5%和15.4%, 所以在测量教育收益率对收入差距的影响中, 考虑家庭资产负债情况很有必要。另外, 东部地区居民的平均收入显著高于中西部地区。

表4将样本按收入等分为四组, 并给出了部分变量在组内的平均值。从表4中可以看出, 在从低到高的四个收入组中, 户主的平均受教育年限越来越长, 工作经验逐渐减少, 城镇化率逐步提高。另外, 在收入的四等分组中, 投资性住房拥有率从8.2%上升到25.1%, 提高了16.9个百分点, 金融市场参与率从0.05%上升到14.8%, 提高了14.75个百分点, 这两个因素的变化幅度较大。由于不同收入组的解释变量存在一定差异, 不同收入组的教育收益率也可能不同, 因此在研究其对收入差距影响的过程中, 有必要考虑收入的分布结构。

(二) 回归分析

1. 教育收益率在不同收入群体的分布情况

根据被解释变量非正态性和相关解释变量随收入变动而变化的情况, 本文采用修正的分位数回归模型检测各个分位点的教育收益率, 使用Stata13软件对不同收入群体的教育收益率进行估计, 结果如表5所示。

从收入的10%分位点到90%分位点, 教育收益率依次为7.2%、6.3%、4.5%、3.2%、2.8%, 说明随着收入水平的提高, 教育收益率逐渐降低。相较于高收入群体, 低收入群体通过相同的教育投资可以获得更多

的收益,这种教育收益率分布特征可能与中国的教育体系和经济体制有关。首先,中国的教育体制忽略了对学生能力的培养。由于中国人口众多,教育资源相对贫乏,教育的主要目的是筛选出优秀的人才享受优质的教育资源,而忽略了对学生能力的培养。虽然教育的筛选作用可以帮助毕业生初次进入劳动力市场时获得薪酬较高的工作,但由于教育忽略了对学生能力的培养,教育成果难以转化为劳动生产率,因此高收入群体也就无法通过教育进一步提高收入,从而导致高收入群体的教育收益率较低。其次,中国以公有制为主体的经济体制弱化了教育的报酬机制。张车伟与薛欣欣提出,中国国有部门的“共享式”收入决定模式提高了低技能员工的收入,压低了高技能员工的收入^[17]。这种收入模式造成了在收入分布的右半部分——高教育水平、高技能群体的集中区教育收益率较低,而在收入分布的左半部分——低教育水平、低技能群体的集中区教育收益率较高。

本文的研究结果表明,作为生产要素之一的人力资本符合新古典经济学的要素边际效用递减理论。相较于高收入群体,低收入

群体的受教育水平相对较低,通过增加教育投资来促进收入增长的潜力大。由于这种潜力存在一定差异,因此低收入群体会选择通过增加教育投资来提高人力资本,进而获得更高的收入,教育收益率更高。教育对收入的边际贡献随着收入的上升而下降。如果一个要素对于低收入群体收入的促进作用大于中高收入群体,那么这种要素就具有缩小收入差距的作用^[18]。中国的教育收益率分布特征与这一要求相一致,所以增加教育投资有助于缩小中国收入差距。

另一个值得我们关注的结果是配偶的受教育年限对低收入群体的影响更大,低收入群体匹配受教育程度较高的配偶可以一定程度上缩小收入差距。然而,在中国的婚姻市场上,受教育程度较高的个体往往会依据教育匹配原则来择偶,以提高或确保自己的经济社会地位,这使得低收入群体处于不利地位。研究表明,夫妻间教育的同质匹配程度增加会在一定程度上扩大家庭间收入差距^[19-20]。

2. 教育收益率在城镇和农村的分布情况

考虑到中国城乡二元经济特征,本文首先统计了城镇居民和农村居民受教育水平的分布情况,结果见表6。农村居民的受教育程度主要集中在基础教育阶段(小学、初中和高中),大多数农村居民只完成了初中教育,这意味着在完成九年义务教育之后他们不再继续接受教育,只有1.7%的农村居民接受了高等教育(大专、本科或研究生)。相比之下,城镇居民接受高等教育的比例(13.7%)要高得多。表7显示了不同教育水平下城镇居民和农村居民的年收入情况。基础教育阶段的城乡居民收入没有显著差异,而接受高等教育的城镇居民的年收入显著高于接受高等教育的农村居民,可见城乡教育收益差异主要集中在高等教育阶段。

表3 解释变量的描述性统计结果

变量	极小值	极大值	均值	标准差(%)
户主受教育年限(年)	0	19	7.037	4.646
配偶受教育年限(年)	0	22	6.243	4.198
工作经验	0	60	34.610	13.180
是否未婚	0	1	0.050	0.199
性别	0	1	0.523	0.448
是否健康	0	1	0.818	0.344
家庭人口数(人)	1	17	3.740	1.824
是否拥有自有房	0	1	0.864	0.312
是否拥有投资性住房	0	1	0.154	0.324
是否参与金融市场	0	1	0.050	0.197
是否有待偿房贷	0	1	0.074	0.016
是否有非住房贷款	0	1	0.052	0.201
是否借钱给别人	0	1	0.162	0.332
金融市场投资品种	0	5	0.063	0.303
是否城镇人口	0	1	0.472	0.441
东部地区虚拟变量	0	1	0.436	0.446
中西部地区虚拟变量	0	1	0.282	0.402

表4 收入四等分组中解释变量的组内均值

变量	最低25%组	中下25%组	中上25%组	最高25%组
户主受教育年限(年)	4.92	6.19	7.44	9.43
工作经验(年)	34.59	31.80	30.82	28.57
是否城镇人口	0.29	0.39	0.51	0.71
投资性住房拥有率(%)	8.2	11.2	17.0	25.1
金融市场参与率(%)	0.05	0.10	3.8	14.8

本文通过分位数回归模型研究教育对城镇和农村居民收入影响的差异,即教育收益率的城乡效应。表8显示了城镇居民和农村居民的教育收益率估计结果。在收入分布的各个分位点上,城镇居民的教育收益率都高于农村居民,教育收益率的城乡效应会加剧城乡收入差距。本文认为教育收益率对城乡收入差距的影响可能与教育投资和户籍制度有关。

首先,教育收益率促进了教育投资。根据理性人假设,由于农村居民的教育收益率显著低于城镇居民的教育收益率,因此农村居民的教育投资需求减少,会将更多的钱投资到其他方面。城镇与农村的教育投入差距形成了城乡人力资本水平差距,从而进一步扩大了城乡收入差距。而收入差距反过来又影响农村居民对教育的投入,农村经济陷入“低收入水平陷阱”。在收入差距和教育收益率的共同作用下,城乡之间很可能形成教育与收入之间的“马太效应”,不利于城乡之间的平衡发展。

其次,教育收益率对突破户籍限制具有激励作用。近年来,随着中国市场经济的发展和改革的深入,户籍制度已经不像原来那样将农民完全束缚在土地上,农村居民可以通过接受高等教育,突破户籍限制进入城市正规部门就业^[21]。教育收益率的城乡差异使得越来越多的来自农村的大学毕业生转变户籍状态,成为城市居民,进入城市公共福利体系。根据赵西亮的测算,2013年城镇居民中有36%的个体经历过户籍转换^[22]。能够突破户籍限制的往往是农村中接受高质量教育的群体,而留在农村的往往是接受过低等教育的妇女和老人,这会阻碍农业现代化的发展和农村居民收入水平的提高。

另外,接受高等教育的农村劳动力拥有未来成为高收入群体的潜力,一旦这部分群体选择离开农村就意味着农

表5 分位数模型回归结果

变量	q10	q25	q50	q75	q90
户主受教育年限	0.072*** (0.008)	0.063*** (0.004)	0.045*** (0.002)	0.032*** (0.002)	0.028*** (0.003)
配偶受教育年限	0.034*** (0.006)	0.031*** (0.004)	0.021*** (0.002)	0.016*** (0.002)	0.012*** (0.002)
工作经验	-0.021*** (0.005)	-0.014*** (0.004)	-0.011*** (0.002)	-0.008*** (0.002)	-0.006*** (0.002)
工作经验的平方	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
是否未婚	0.022 (0.124)	0.101 (0.076)	0.107** (0.044)	0.136*** (0.039)	0.111** (0.051)
是否男性	0.018 (0.059)	-0.056** (0.024)	0.014 (0.019)	0.047** (0.019)	0.073*** (0.021)
是否健康	0.227*** (0.046)	0.249*** (0.040)	0.200*** (0.027)	0.128*** (0.017)	0.103*** (0.025)
是否在职	0.168** (0.077)	0.044 (0.042)	-0.014 (0.020)	-0.023 (0.025)	-0.055** (0.025)
家庭人口数	-0.032*** (0.012)	-0.037*** (0.008)	-0.089*** (0.007)	-0.108*** (0.005)	-0.116*** (0.005)
是否拥有自住房	-0.226*** (0.012)	-0.117** (0.049)	-0.081*** (0.025)	-0.116*** (0.030)	-0.124*** (0.030)
是否拥有投资性房地产	0.311*** (0.090)	0.268*** (0.050)	0.272*** (0.024)	0.263*** (0.018)	0.317*** (0.030)
是否参与金融市场	0.723*** (0.157)	0.258*** (0.092)	0.321*** (0.067)	0.184** (0.082)	0.235* (0.121)
是否借钱给别人	0.454*** (0.058)	0.355*** (0.043)	0.278*** (0.025)	0.230*** (0.021)	0.280*** (0.038)
是否有待偿房贷	0.086 (0.120)	0.159*** (0.055)	0.168*** (0.036)	0.193*** (0.030)	0.195*** (0.044)
是否有非住房贷款	-0.260** (0.117)	-0.092 (0.078)	-0.005 (0.032)	0.062 (0.054)	0.075* (0.042)
金融市场投资品种	-0.036 (0.153)	0.145* (0.087)	-0.064 (0.042)	0.141** (0.070)	0.122 (0.094)
是否为城镇	0.495*** (0.067)	0.400*** (0.037)	0.243*** (0.019)	0.189*** (0.026)	0.143*** (0.031)
东部地区 (以中西部为参考)	0.513*** (0.054)	0.413*** (0.034)	0.327*** (0.023)	0.322*** (0.027)	0.285*** (0.025)
常数	6.623*** (0.138)	7.536*** (0.112)	8.755*** (0.058)	9.538*** (0.057)	10.074*** (0.071)

注:1.每一列表示不同分位点对应的分位数回归结果;2.***、**和*分别表示在1%、5%和10%水平下显著;3.括号中数值是标准误差。下同。

表6 城乡居民受教育水平分布情况

教育水平	农村		城镇	
	人数	百分比	人数	百分比
文盲/半文盲	1891	29.9%	898	15.6%
小学	1745	27.6%	1065	18.5%
初中	2023	32.0%	1867	32.4%
高中	547	8.7%	1139	19.8%
大专	82	1.3%	459	8.0%
本科	26	0.4%	295	5.1%
研究生及以上	1	0.0%	33	0.6%
总数	6315	100.0%	5756	100.0%

表7 居民年收入情况

教育水平	年收入(元)	
	平均值(农村)	平均值(城镇)
文盲/半文盲	18744	20212
小学	22296	24236
初中	24930	26368
高中	25530	27925
大专	34881	39147
本科	35966	47777
研究生及以上	30000	62200

村高收入群体的流失,就会导致农村居民收入水平的进一步下降,加剧城乡收入差距。

3.教育收益率的区域效应

中国幅员辽阔,地区间自然条件和经济发展水平迥异,东部地区与中西部地区的教育收益率可能存在差异。为了检验教育收益率的区域效应,本文首先统计了东部地区和中西部地区教育水平的分布情况,结果见表9。从表9中可以看出,东部地区接受高等教育的人数占14.04%,而

中西部地区接受高等教育的人数占11.52%,占比下降了2.52%。总体来看,东部地区与中西部地区的教育水平分布差异不大,说明中国促进区域教育均衡发展的政策取得了一定的成效。

然而,在不同区域工作的个体即使拥有相同教育水平,收入也会存在较大差异。表10展示了不同教育水平下东部地区和中西部地区的年收入情况。表10显示,不同受教育阶段东部地区的年收入平均值都高于中西部地区。此外,本文还对区域居民收入差距进行了分析(独立t检验),结果表明区域居民收入水平存在显著差异($t=11.59, p<0.01$),这意味着东部地区整体的收入水平高于中西部地区。另外,基础教育阶段的东部地区与中西部地区的收入水平没有显著差异,而高等教育阶段东部地区的收入水平显著高于中西部地区,接近中西部地区的两倍。

因此,本文将教育收益率分析扩展到区域层面,以揭示东部地区和中西部地区教育收益率的差异,结果如表11所示。在收入分布的各个分位点上,东部地区的教育收益率都要高于中西部地区,教育收益率的区域效应不利于区域经济的协调发展。本文认为教育收益率对区域收入差距的影响可能与人力资本、产业结构和劳动力市场有关。

首先,教育收益率对人力资本积累具有激励作用。东部地区的教育收益率较高说明教育投资行为转化成人力资本的能力较强,相较于中西部地区,东部地区更愿意进行教育投资。因此,教育收益率促进了东部地区的人力资本积累,增强了劳动者对先进技术的吸收模仿能力,产出能力也获得了提升,从而促进了地区经济增长。参照彭国华的计算方法,本文测算得出2016年东部地区的人力资本存量为216983万人/年,显著高于中部地区(138353万人/年)和西部地区(121179万人/年)^[23]。

其次,教育收益率促进了产业升级。生产力水平较高的高端产业需要高学历人才,其相应的教育投资回报也较高。如果高端产业都聚集在东部沿海地区,形成了区域集聚效应,那么具有相同教育背景的个体在东部地区可以获得更高的收入和就业就会,因此会选择到东部地区就业,进而促进了相关产业的优化

表8 教育收益率的城乡效应

		q10	q25	q50	q75	q90
城镇	受教育水平	0.092*** (-0.007)	0.068*** (-0.004)	0.045*** (-0.003)	0.036*** (-0.003)	0.030*** (-0.004)
	工作经验	0.039*** (0.009)	0.037*** (0.004)	0.022*** (0.003)	0.019*** (0.003)	0.017*** (0.003)
农村	受教育水平	0.067*** (-0.013)	0.053*** (-0.006)	0.043*** (-0.003)	0.030*** (-0.003)	0.028*** (-0.003)
	工作经验	0.026*** (0.013)	0.023*** (0.005)	0.023*** (0.004)	0.034*** (0.004)	0.032*** (0.003)
其他控制变量		YES	YES	YES	YES	YES

表9 区域教育水平分布情况

教育水平	东部		中西部	
	人数	百分比	人数	百分比
文盲/半文盲	429	10.35%	700	12.86%
小学	799	19.28%	1255	23.05%
初中	1463	35.30%	1814	33.31%
高中	872	21.04%	1049	19.27%
大专	346	8.35%	402	7.38%
本科	222	5.36%	215	3.95%
研究生及以上	14	0.34%	10	0.18%
总数	4145	100.00%	5445	100.00%

表10 不同区域的教育阶段收入分布情况

教育水平	年收入(元)	
	平均值(东部)	平均值(中西部)
文盲/半文盲	15862	15397
小学	22790	19020
初中	25382	21158
高中	28257	21193
大专	46265	32153
本科	63778	37395
研究生及以上	118807	43307

表11 教育收益率的区域效应

		q10	q25	q50	q75	q90
东部地区		0.083*** (0.010)	0.064*** (0.008)	0.045*** (0.003)	0.035*** (0.003)	0.030*** (0.003)
中西部地区		0.071*** (0.014)	0.060*** (0.011)	0.045*** (0.007)	0.033*** (0.006)	0.027*** (0.005)
其他控制变量		YES	YES	YES	YES	YES

升级。

最后,教育收益率为劳动力市场整合创造了有利条件。教育收益率越高表明教育的信号作用越强。教育的信号作用使得市场价格能够更加合理地显示出各级劳动力市场的供需情况,这有助于优化劳动力配置,鼓励企业加大对人才的培养力度。根据田永坡对2000—2014年中国劳动力市场成熟度的测算,劳动力市场成熟度排名靠前的省区市集中在东部地区,排名靠后的省区市集中在中西部地区和东北地区^[24]。

五、结论性评述

本文基于分位数回归模型分析了教育收益率对中国收入差距的影响,所得结论如下:(1)随着收入分位点的提高,中国的教育收益率逐渐下降,教育收益率有助于缩小收入差距。(2)在收入水平不断提高的情况下,配偶受教育水平的边际收入效应会弱化,婚姻中的教育匹配可能会扩大中国不同家庭的收入差距。(3)城镇居民的教育收益率要高于农村居民,教育收益率的城乡效应会扩大城乡收入差距。(4)东部地区的教育收益率要高于中西部地区,教育收益率的区域效应也会扩大区域收入差距。

根据研究结果,本文对缩小中国收入差距提供几点建议:(1)坚持教育强国政策,重点扶持低收入群体。考虑到低收入群体的教育收益率更高,对该群体进行教育投资更有利于增加社会效益和促进收入公平,政府需要加强对低收入群体的教育扶持。(2)完善税收制度,防止阶层固化。税务部门可以考虑在个税制度中加入家庭信息(纳税人家庭特征、家庭负担、夫妻双方收入等),进一步完善税前扣除标准、优化税率结构,促进不同阶层群体之间的流动,缓解阶层固化的现象。(3)深化农村教育体制改革,促进职业教育发展。当前,在农业现代化和工业化快速发展的同时,农村教育体系与农村产业发展不协调的矛盾日益突出。当代农村教育体系需要进行一定程度的改革,以实现提高农村人力资本水平、加快推进农业现代化进程和促进城乡协调发展的目标。农村居民教育收益率较低的一部分原因是教育所学的知识与农村生产实际差距较大,不能学以致用,加强农村职业教育可以有效解决这一问题。(4)加大对中西部地区的教育投入,缩小区域发展差异。区域人才竞争越来越激烈,中西部地区学校的师资条件差,经济实力较弱,难以吸引东部地区的人才。人才逆差使得东部地区和中西部地区的差距越拉越大,需要出台中西部扶持计划等一系列政策,在中西部地区建设一些有水平、有特色的高等院校,通过产学研带动中西部经济的发展。同时,还需要改善当前的劳动力市场环境,消除劳动力流动的障碍,使得劳动者和工作可以实现更好的匹配,从而提高个体和社会的经济效益。

当然,本文在未来可以进行更深入的研究,如:研究人口流动对教育收益率的影响,通过比较流动人口和非流动人口的教育收益率来进行深入分析;研究教育收益率对收入差距可能存在的其他影响机制;考虑加入2016年和2018年的CFPS数据,对比分析不同年份的教育收益率分布特征是否存在差异,并探讨存在差异的根本原因。

参考文献:

- [1] Shi L. Recent changes in income inequality in China[R]. Working Paper, 2016.
- [2] Harmon C, Oosterbeek H, Walker I. The returns to education: Microeconomics [J]. Journal of Economic Surveys, 2003, 17(2): 115-156.
- [3] Martins P S, Pereira P T. Does education reduce wage inequality? Quantile regression evidence from 16 countries [J]. Labour Economics, 2004, 11(3): 355-371.
- [4] Arabshuibani G R, Staneva A. Returns to education in Russia: Where there is risky sexual behaviour there is also an instrument [R]. Discussion Papers, 2012.
- [5] Chernozhukov V, Hansen C. Quantile models with endogeneity [J]. Annual Review of Economics, 2013, 5(1): 57-81.
- [6] 罗楚亮. 城镇居民教育收益率及其分布特征 [J]. 经济研究, 2007(6): 119-130.

- [7] 赵显洲. 教育收益率的地区差异研究[J]. 技术经济与管理研究, 2015(1):88-91.
- [8] 谭远发, 黄健, 曾尹嫵. 中国教育收益率分位回归的研究前沿与新进展[J]. 教育学术月刊, 2016(9):3-9.
- [9] Chiswick, Barry R. Earnings inequality and economic development[J]. The Quarterly Journal of Economics, 1971, 1(85):21-39.
- [10] Mincer J. Schooling, experience, and earnings[M]. New York: Columbia University Press, 1974.
- [11] Knight J B, Sabot R H. Educational Expansion and the Kuznets Effect[J]. American Economic Review, 1983, 73(5):1132-1136.
- [12] Iversen J, Malchow-Møller N, Sørensen A. Returns to schooling in self-employment [J]. Economics Letters, 2010, 109 (3) : 179-182.
- [13] 黄静, 王周伟, 杜永康. 收入差距分化、城市化发展与教育收益率——基于分位数回归的实证分析[J]. 教育与经济, 2015 (6):18-24.
- [14] Wang L. How does education affect the earnings distribution in urban China? [J]. Oxford Bulletin of Economics & Statistics, 2012, 75 (3):435-454.
- [15] McLaughlin J S. Does communist party membership pay? Estimating the economic returns to party membership in the labor market in China[J]. Journal of Comparative Economics, 2017, 45(4): 963-983.
- [16] Zhang J, Zhao Y, Park A, et al. Economic returns to schooling in urban China, 1988 to 2001 [J]. Journal of Comparative Economics, 2005, 33(4):730-752.
- [17] 张车伟, 薛欣欣. 国有部门与非国有部门工资差异及人力资本贡献[J]. 经济研究, 2008(4):17-27.
- [18] 谭远发, 黄健, 曾尹嫵. 中国教育收益率分位回归的研究前沿与新进展[J]. 教育学术月刊, 2016(9):3-9.
- [19] 潘丽群, 李静, 踪家峰. 教育同质性婚配与家庭收入不平等[J]. 中国工业经济, 2015(8):35-49.
- [20] 李代. 教育的同型婚姻与中国社会的家庭工资收入不平等:1996—2012[J]. 社会, 2017(3):103-130.
- [21] 邢春冰, 贾淑艳, 李实. 教育回报率的地区差异及其对劳动力流动的影响[J]. 经济研究, 2013(11):114-126.
- [22] 赵西亮. 教育、户籍转换与城乡教育收益率差异[J]. 经济研究, 2017(12):164-178.
- [23] 彭国华. 国地区收入差距、全要素生产率及其收敛分析[J]. 经济研究, 2005(9):19-29.
- [24] 田永坡. 中国劳动力市场的成熟度测度:2000—2014[J]. 改革, 2016(10):96-105.

[责任编辑:王丽爱]

The Impact of Returns to Education on Income Gap in China: An Analysis Based on the Quantile Regression Model

ZHANG Chuanguo, JIN Yuanyuan

(The School of Economics, Xiamen University, Xiamen 361005, China)

Abstract: Taking the data from China Family Panel Studies (CFPS 2014) and Quantile Regression Model, this paper explores the impact of education return on income gap in China. The results show that: (1) with the increase of income quantile, the returns to education in China has gradually declined, which helps to narrow the income gap among different groups in China; (2) The spouse's years of education have a greater impact on low-income individuals; (3) The return to education in urban areas is higher than that in rural areas. The urban-rural effect of the return to education will aggravate the income gap between urban and rural areas; (4) The return to education in the eastern region is higher than that in the central and western region, and the regional effect of the returns to education will aggravate regional income gap. In general, although the returns to education will narrow the income gap among different groups, it will widen the income gap between urban and rural areas and between regions.

Key Word: returns to education; income gap; years of education; human capital accumulation; labor market; the true rate of return to education