

人力资本积累与收入差异

——来自 CFPS 的实证

秦 昌 才 孙 世 宇

(烟台大学 经济管理学院, 山东 烟台 264005)

摘 要:通过教育进行人力资本积累,是缩小收入差异,解决相对贫困,实现共同富裕的有效措施。基于中国国家追踪调查 2018 年的微观数据,使用工具变量分位数回归方法,从全样本、性别、城乡和区域的角度,系统分析条件收入分布下教育收益率的异质性。结果表明,我国大学教育收益率依条件收入分布呈 L 型形状,两端的低收入与高收入群体具有较高的收益率;这意味着低收入群体通过教育缩小收入差距,高收入群体通过教育扩大差距的“马太效应”同时存在。女性收益率大于男性,其中低收入群体的年度回报率,男性为 4.59%,女性为 10.62%。除了 10 分位点的低收入群体,城市收益率都大于农村。东部收益率随着收入水平的增加而递增,中西部则是递减。

关键词:教育收益率;工具变量分位数回归;异质性;CFPS

中图分类号:F323

文献标识码:A

文章编号:1673-0429(2024)04-0077-13

doi:10.19742/j.cnki.50-1164/C.240408

引 言

党的十八大以来,我国通过精准扶贫战略基本解决了绝对贫困问题,但居民间收入差异较大的相对贫困仍是一个任重道远的长期难题。根据 2023 年国民经济和社会发展统计公报的数据,全国居民前 20% 高收入组的人均可支配收入为 95 055 元,而后 20% 低收入组则仅为 9 215 元,两组差距超过 10 倍。城镇居民人均可支配收入 51 821 元,农村居民为 21 691 元,前者是后者的 2.4 倍。这表明,无论从全国还是城乡的角度来看,我国居民收入存在着显著的差异。根据传统人力资本理论,个体间收入差别的 60% 是源于教育方面的差异,而教育是人力资本积累的重要组成部分,教育通过提高个体的生产力、劳动技能和市场竞争力,从而提升个体的就业能力来直接影响收入水平,同时还通过改善健康状况、增强社会资本等间接渠道影响个体间的收入差异。宏观上,人口受教育水平的普遍提高减弱了财富向下极化的趋势,促进中产阶级的壮大和稳固^[1]。教育借助提高税收收入、促进人口流动、提升创业活力和促进高水平就业等途径促进共同富裕^[2]。正是基于此,通过教育进行人力资本投资,提高个体收入,成为

收稿日期:2024-06-12

作者简介:秦昌才(1980—),男,山东沂水人,经济学博士,烟台大学经济管理学院副教授,硕士研究生导师,主要研究方向:数量经济、发展经济。

基金项目:2021 年度山东省自然科学基金面上项目“新动能培育背景下山东省人力资本服务优化及促进经济增长的机制研究”(ZR2020MG036);山东省专业学位研究生教学案例库“农业硕士经济学实证方法教学案例库”(SDYAL20173)。

补偿弱势群体、解决贫富悬殊、缩小收入差异,从而实现共同富裕的有效措施。

教育人力资本积累主要是通过提高个体的能力来增加其收入,其内在影响机制非常复杂;而且教育收益率存在性别、城乡和区域间较大的异质性。教育对个体收入的影响渠道,宏观方面有经济结构的变迁影响劳动力的组成结构和劳动力的配置效率,从而影响劳动收入;微观方面有原生家庭因素,如父母的受教育水平等制度性文化资本影响个体的教育成就和人力资本积累,家庭经济资本和社会资本也会影响个人的收入获得能力等。部分研究认为,女性教育收益率会高于男性^[3],特别是城镇居民的性别差异更加明显^[4]。而我国公共教育资源无论从质量还是数量上都向城市较多倾斜,从而造成城乡教育质量存在较大不同,城市人口教育收益率大于农村^[5-6]。而经济发展水平的不同会造成区域间生产率的不同,从而造成教育收益率在我国东、中、西部之间具有明显的差异^[7]。

已有教育收益率研究主要面临测量误差、样本选择和遗漏变量等方面造成的内生性问题。由于诸多实证研究的数据基础来源于微观调查,在此过程中,教育程度往往容易出现向上报告的偏差,而收入则有向下汇报的可能,这会导致教育的收益率被低估^[8]。教育收益率研究中的收入,核算单位可以使用年、月或小时,核算形式可以是工资性收入或总收入等,不同的形式对于收益率测度也是有误差的;另外,教育的回报应该着眼于终生收入,而不是当期收入,否则,会产生生命周期偏误^[9]。我国户籍制度存在高学历的选择性迁移,使得农村样本是一个选择性样本,从而造成农村教育收益率会被低估^[10]。最后一个,也是最重要的问题是遗漏变量,特别是遗漏能力等变量造成的内生性偏差问题。

对于上述内生性问题,一般使用工具变量方法来克服,从而准确识别教育对个体收入的因果影响。但在教育收益率研究中寻找一个合适的工具变量并不容易,会产生较大争议。已有文献提供了下面多种选择:出生季度、吸烟习惯、性别筛选、距离大学的远近,父母受教育程度,配偶的受教育程度和家庭人口规模等^[11-17]。还有研究利用制度变量作为工具变量:义务教育法、最低劳动年龄法以及1999年的高校扩招等^[18-22]。根据布迪厄文化再生产理论,父母受教育程度属于制度化文化资本,是影响个人教育获得的重要因素。受教育水平相对较高的父母,本身会比较喜欢看书、阅读等文化活动,家庭藏书较多,会有书房等专门用于读书和学习的区域,家庭有较好的学习氛围和文化氛围,对后代的教育期望也较高,并且有对子女学习进行辅导的能力^[23-24],这都对个人教育获得有很大裨益。而家庭规模变量体现了兄弟姐妹个数等社会资本对家庭有限资源的稀释和分散强度,也会显著影响个人的教育成就。

基于上述分析并借鉴已有研究,本文的边际贡献是采用父母受大学教育与否以及家庭规模作为个体受大学教育与否的工具变量;基于中国家庭追踪调查(CFPS)2018年的微观数据;使用工具变量分位数回归(IVQR)估计方法,分别从全样本、性别、城乡和区域的角度,研究我国条件收入分布下大学教育收益率的异质性,从而为我国人力资本教育收益研究、实施教育反贫困战略提供一定的实证依据。

一、理论分析与识别策略

(一) 理论模型

根据Card^[25]提出的观点,个人选择受教育水平是为了最大化其预期收益与成本之间的差值。教育不仅提升个人的就业与收入,还在精神层面提高生活满意度。假定个人选择受教育水平 ep 是为了使得以下目标效用函数最大化:

$$U(wap, ep) = \ln wap - c(ep) = \ln b(ep) - c(ep) \quad (1)$$

其中, $wap = b(ep)$ 为受教育水平对应的收益函数,包括物质方面的就业与收入,以及精神层面的生活满意度等方面;但为了便于量化并与已有研究进行比较,本文侧重指工资收入、总收入等物质收入方面。 $c(ep)$ 为受教育水平对应的成本函数,包括接受教育过程直接付出的货币成本、时间成本,以及精神层面无形的“痛苦”成本等^[22]。此外, $b(\cdot)$, $c(\cdot)$ 都是 ep 的单增凸函数。

一阶最优条件就是：

$$MB = \frac{b'(ep)}{b(ep)} = c'(ep) = MC \quad (2)$$

显然,在最优受教育水平 ep^* 上,教育的边际收益 MB 等于边际成本 MC 。个人最优受教育水平的异质性差异主要来自于成本和收益两个方向。

如果把工资收入的分位数分布作为能力的代理指标^[26],也即平均来看,较高能力者会有较高收入,而较高能力者获得同等教育水平的成本较低,因此,在收入分布的较高分位数上,个体的边际收益应该较小。

虽然 Card(1999) 是明瑟方程的一般化,但无法直接用于实证分析。已有研究表明,工资收入与受教育水平是对数线性的,同时对数线性函数比较好地描述了条件分布下工资收入与受教育水平的关系^[27-29]。

(二) 实证模型与优化问题

基于此,本文进行实证分析的基本模型形式是下列扩展的明瑟方程：

$$\ln wap_i = \alpha + \beta_e p_i + X' \gamma + u_i \quad (3)$$

其中, wap_i 是第 i 个人的工资性收入, ep_i 是其受教育水平, X 是控制向量,包括工作经验、工作经验的平方项、智力状况、健康状况、婚姻状况等。上述模型主要是针对平均处理效应 (ATE, average treatment effects) 的因果识别,但个体特征差异以及一些不可观测因素都会造成个体收入的异质性。此时, Koenker and Bassett^[30] 提出的分位数回归方法 (QR, quantile regression) 可以刻画内在机制较为复杂的分布效应^[31]。

因此,本文的问题就转化为以下线性规划问题：

$$\argmin_{\beta(\tau), \gamma(\tau)} E\{\rho_\tau[\ln wap - Q_{\ln wap}[\tau | ep, X]]\} \quad (4)$$

这里, τ 为分位点, $\rho_\tau(u_i) = [\tau - 1(u_i \leq 0)] u_i$ 为检验函数, $Q_{\ln wap}[\tau | ep, X] = \beta(\tau) ep_i + X' \gamma(\tau)$ 为 τ 分位点上的条件收入函数, $\beta(\tau), \gamma(\tau)$ 为对应分位点上的条件分布系数,可以刻画不同收入群体的组间异质性。QR 的前提假设是受教育水平变量 ep 与误差项 u_i 严格独立,由于上述工资收入决定方程中存在众多遗漏因素,也就是众所周知的内生性问题,因此上述假设过于严格。基于此,本文使用 Chernozhukov and Hansen^[32-34] 提出的工具变量分位数回归方法 (IVQR, Instrumental variable quantile regression), 在考虑内生性的情况下研究教育收益率的异质性。与传统工具变量方法的 2SLS 估计过程一样, IVQR 假定存在工具向量 Z 使得在 (X, Z) 条件下, 对于 $\forall ep \in EP$, 结构误差 u_{ep} 独立同分布, 这就是秩相似性 (rank similarity)。这意味着,

$$P[\ln wap \leq Q_{\ln wap}(\tau | ep, X) | (X, Z)] = \tau \quad (5)$$

在本文中就是寻找 $(\beta(\tau), \gamma(\tau))$ 以满足下列最优问题

$$0 = \argmin_{f \in F} E\{\rho_\tau[\ln wap - \beta(\tau) ep_i - X' \gamma(\tau) - f(X, Z)]\} \quad (6)$$

因此,本文的实证分析模型如下：

$$\ln wap_{-\tau_i} = \alpha_{-\tau}^{IVQR} + \beta_{-\tau}^{IVQR} ep_i + X' \gamma_{-\tau}^{IVQR} + u_i \quad (7)$$

其中, $\ln wap_{-\tau_i}$ 为第 i 个个体的工资性收入分位数水平, τ 依次取值 10、30、50、70 和 90, 依次对应 10 分位至 90 分位。分位数值越大, 相应的收入水平就越高。 $\alpha_{-\tau}^{IVQR}, \beta_{-\tau}^{IVQR}, \gamma_{-\tau}^{IVQR}$ 为对应分位数上条件收入分布系数。

二、数据来源与描述性统计

(一) 数据来源

本文实证分析的依据是北京大学中国社会科学调查中心的中国家庭追踪调查 (CFPS) 2018 年的微观截面数据, 这一数据集不仅具有广泛的代表性, 还能够深入反映我国不同地区、城乡以及收入阶层的

社会经济状况,是研究教育收益与收入差距问题的理想数据来源。本文主要利用其中的成人数据、家庭成员关系数据和家庭经济数据。以成人数据库中的个人 ID 为依据,将家庭数据与个人进行匹配对应,从而获得本文研究所需要的全部个人及家庭数据信息,这一数据匹配过程确保了数据的一致性和可靠性,有助于减少研究中的潜在误差。通过对这些数据的分析,本文能够系统地探讨教育在不同社会经济背景下的作用及其对收入分配的影响,为政策制定提供实证支持。

由于本文主要研究的是大学教育的经济回报,因此,首先将年龄限定在劳动法所规定的 16~60 岁区间、拥有劳动收入的样本,然后剔除了在校学生以及工资性收入和总收入均为零的样本,然后剔除了主要变量数据不适用或缺失的观测,最终得到 6 531 个有效样本,其中男性样本 4 010 个,女性 2 521 个,分别占 61.4%和 38.6%;城市样本 3 697 个,农村则是 2 834 个,占 56.6%和 43.4%。东部样本 2 978 个,中部样本 2 114 个,西部样本 1 439 个,分别占 45.6%、32.4%和 22.0%。

本文的主要解释变量是个人受大学教育与否的虚拟变量。在 CFPS 问卷中对应着成人数据库中的“个人最高学历”变量,分别有文盲/半文盲到博士 8 个等级。本文将大专及大专以上的四个分类处理为 1,而将高中及以下处理为 0。父母大学教育变量的处理方式也是这样。工作经验数据的处理则是用 2018 年调查时的年龄减去受教育时间以及学前时间(六年)。健康状况及智力水平,在 CFPS 问卷调查中有 1 至 7 个分类和赋值;工作时间为每周小时数。婚姻状况有未婚、同居、离婚、丧偶和在婚(有配偶)五个分类,本文将前四类处理为 0,而将在婚处理为 1。

(二)描述性统计

实证研究的软件是 Stata 14。所涉变量均值方面的描述性统计结果见表 1。

表 1 描述性统计(均值)

变量	性别		城乡		区域		
	男性	女性	城市	农村	东部	中部	西部
工资收入	32 407.41	22 778.65	31 040.55	25 625.18	32 659.91	26 445.11	23 775.2
总收入	32 446.5	22 885.45	31 123.76	25 666.95	32 743.62	26 501.91	23 814.56
大学教育	0.188	0.237	0.282	0.109	0.226	0.209	0.166
父亲大学教育	0.029	0.033	0.048	0.007	0.036	0.031	0.018
母亲大学教育	0.013	0.014	0.023	0.001	0.016	0.013	0.008
家庭规模	4.121	3.944	3.764	4.429	3.756	4.189	4.465
工作经验	21.079	19.553	20.583	20.368	20.403	20.69	20.376
工作经验平方/100	5.893	5.211	5.628	5.633	5.694	5.621	5.51
智力状况	5.708	5.691	5.792	5.584	5.732	5.747	5.571
健康状况	5.852	5.895	5.973	5.732	5.904	5.955	5.668
周工作时间	50.972	47.592	48.286	51.468	49.745	49.518	49.726
婚姻状况	0.774	0.756	0.768	0.765	0.752	0.79	0.764

表 1 中显示,无论工资性收入还是总收入,性别、城乡和区域之间均有显著差别。男性均大于女性,城市大于农村,东部最大、中部居中、西部最小。其中,男性工资性收入平均为 32 407.41 元,而女性为 22 778.65 元,只有男性同胞的 70.29%。城市居民的工资性收入和总收入分别高于农村 21.13%、21.26%;而东部大于中部大于西部,工资性收入方面,东部平均为 32 659.91 元,高于中部的 26 445.11 元,更高于西部的 23 775.2 元;而总收入方面,东部比中部高出 23.55%,比西部高 37.49%。大学教育

方面,男性均值为 0.188,低于女性的 0.237;城市为 0.282,高于农村的 0.109;东部与中部接近,分别为 0.226 和 0.209,都高于西部的 0.166。在父母受大学教育方面,男女两性别之间没有较大差异,但城市远高于农村,东部与中部接近,都高于西部。在家庭规模方面,城市为 3.764,小于农村的 4.429;东部小于中部及西部,符合农村家庭人口较多,中西部大于东部发达地区的实际。而在工作经验、健康状况、婚姻状况、智力水平和每周工作时间等方面,性别、城乡、区域之间差异不大。

总之,从均值的角度来看,个人收入与个人大学教育基本呈现正相关;个人与父母大学教育基本呈现正相关,而与家庭规模呈负相关。在接下来的分析中,相关的统计检验将表明,父母大学教育与家庭规模是个人大学教育的有效工具变量。

二、实证分析与结果讨论

这一部分本文将估计大学教育收益率。为了方便比较,首先基于方程(3)进行 OLS 估计;然后针对异质性进行普通 QR 估计,针对内生性问题进行 IV 估计;最后基于方程(7)进行同时考虑异质性和内生性的 IVQR 估计。

(一)基本回归

根据表 2 的回归结果,全样本大学教育收益率的 OLS 估计值是 0.545,这说明与未受大学教育的人相比,受过大学教育的工资性收入要提高 54.5%,这在 0.001 水平上是显著的,表明教育对收入的直接提升作用。由于普通最小二乘法是平均处理效应,其系数不随个人收入分布而变化。因此,第二至六列汇报了条件分位数回归(QR)的估计结果,进一步揭示了大学教育收益率在不同收入分布上的差异,特别是在收入分布的两端,这种差异更为显著,表明大学教育收益率在不同收入分布上差异明显。其中,10 分位上数值最大,是 0.664;70 分位上最小是 0.349,二者均在 0.001 水平上是显著的。其异质性的收益曲线一定程度上呈 L 型形状,95%置信区间上分位点的变化如图 1 所示。这说明教育投入对两端低收入和高收入人群的人力资本积累而产生的收入效果最为明显。其可能的原因是低收入群体由于教育资源的稀缺性,通过教育可以显著提升其收入水平,打破贫困的代际传递,从而缩小与其他群体的组间收入差距;教育程度提升可以有效减少低收入群体与其他群体的组间收入差异。而高收入群体则利用优质教育资源的获得进一步提升收入,扩大与中等收入群体的差距。

表 2 基本回归结果

	OLS(1)	QR					IV	
		(2)10 分位	(3)30 分位	(4)50 分位	(5)70 分位	(6)90 分位	(7)2SLS	(8)GMM
大学教育	0.545 *** (14.80)	0.664 *** (3.84)	0.476 *** (9.57)	0.382 *** (12.54)	0.349 *** (13.35)	0.427 *** (12.79)	1.345 *** (5.83)	1.211 *** (6.26)
工作经验	0.138 *** (13.29)	0.394 *** (15.53)	0.087 *** (11.96)	0.043 *** (9.62)	0.029 *** (7.65)	0.027 *** (5.66)	0.162 *** (15.36)	0.158 *** (12.96)
工作经验平方/100	-0.375 *** (-14.27)	-1.145 *** (-21.55)	-0.221 *** (-14.44)	-0.113 *** (-12.06)	-0.080 *** (-10.02)	-0.071 *** (-6.92)	-0.404 *** (-21.32)	-0.397 *** (-14.34)
智力状况	0.107 *** (4.25)	0.112 (1.67)	0.077 *** (4.00)	0.064 *** (5.40)	0.054 *** (5.36)	0.054 *** (4.17)	0.088 *** (4.00)	0.089 *** (3.52)
健康状况	-0.036 (-1.58)	-0.069 (-0.97)	-0.005 (-0.25)	-0.005 (-0.45)	-0.008 (-0.78)	-0.026 (-1.89)	-0.046 * (-1.99)	-0.039 (-1.69)

续表2

	OLS(1)	QR					IV	
		(2)10分位	(3)30分位	(4)50分位	(5)70分位	(6)90分位	(7)2SLS	(8)GMM
工作时间	0.381*** (7.02)	0.481*** (4.44)	0.257*** (8.25)	0.172*** (9.05)	0.112*** (6.86)	0.052* (2.49)	0.424*** (11.51)	0.417*** (7.43)
婚姻状况	0.009 (0.19)	-0.156 (-0.81)	0.073 (1.32)	0.062 (1.82)	0.083** (2.85)	0.074* (2.01)	-0.069 (-1.05)	-0.068 (-1.21)
常数项	7.029*** (26.48)	4.041*** (6.49)	7.541*** (42.08)	8.728*** (79.46)	9.430*** (100.07)	10.17*** (84.55)	6.598*** (28.17)	6.659*** (22.75)
样本数	6,531	6,531	6,531	6,531	6,531	6,531	6,531	6,531

注:(1)表中汇报的是系数值,下方括号是相应的t统计量(OLS和QR)和Z统计量(2SLS和GMM);(2)*、**、***分别代表显著性水平0.05,0.01和0.001,下同

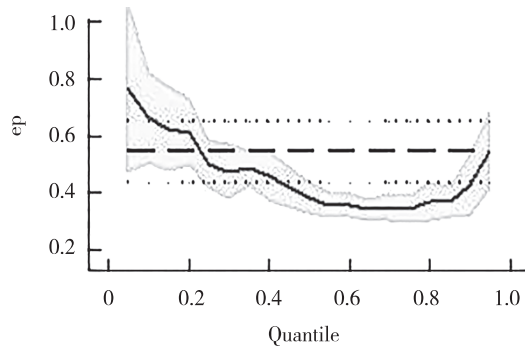


图1 95%置信区间上大学教育系数在各收入分位点上的变化

众多有关教育收益率的文献研究已经表明,教育对个体收入的因果关系识别面临的最大的挑战是遗漏变量产生的内生性问题。尽管本研究加入了工作经验、每周工作时间等影响收入的控制变量,但遗漏与个体教育获得与潜在收入都密切相关的能力变量,会使得OLS和QR的估计都存在较大的偏差。

接下来本文将使用工具变量(IV)方法来克服这一问题,由于考虑到本文使用的是微观截面数据,具有一定的异方差,因此采用两步最小二乘(2SLS)估计和考虑异方差的广义矩(GMM)估计。首先对个人大学教育进行内生性检验,2SLS的Wu-Hausman F统计量为13.202,而GMM的统计量为12.343,均在0.001水平上通过了显著性检验,即拒绝了个人大学教育外生的原假设。借鉴已有研究,本文使用父母大学教育和家庭规模作为个人大学教育的工具变量,以纠正估计偏差。2SLS估计的Sargan统计量为4.91,GMM估计的Hansen统计量为4.92,在0.05水平上均无法拒绝原假设,即所有工具变量都外生,与残差不相关;同时二者的F统计量值分别为134.64和132.15,在0.001水平上工具变量与内生变量显著相关。也即上述工具变量既满足外生性又满足相关性,是有效的。回归结果见表2的(7)(8)两列。加入工具变量之后,与OLS估计相比大学教育的IV系数显著增加,这与已有研究结论相似^[35]。2SLS估计和GMM估计的系数分别为1.345和1.211,均在0.001水平上通过了显著性检验。

(二)工具变量分位数回归(IVQR)

表中显示,与QR估计相比,在所有分位点上大学教育的IVQR估计值都显著增加,其中90分位点上大学教育收益率提高了56.9%,这与已有研究中工具变量的引入会提高回归系数的结论一致^[40]。即遗漏父母受教育程度和家庭规模等先天禀赋因素,可能低估大学教育在提高个人收入方面的正向作用。

表 3 工具变量分位数回归结果

	(1) 10 分位	(2) 30 分位	(3) 50 分位	(4) 70 分位	(5) 90 分位
大学教育	1.031 *	0.978 **	0.855 **	0.841 **	0.996 *
	(2.03)	(3.19)	(2.89)	(2.69)	(2.44)
工作经验	0.397 ***	0.097 ***	0.059 ***	0.043 **	0.041 *
	(17.12)	(6.98)	(4.42)	(3.05)	(2.21)
工作经验平方/100	-1.144 ***	-0.223 ***	-0.133 ***	-0.101 ***	-0.092 **
	(-27.41)	(-8.84)	(-5.47)	(-3.94)	(-2.76)
智力状况	0.083	0.051	0.057 *	0.051	0.040
	(1.70)	(1.73)	(2.00)	(1.69)	(1.03)
健康状况	-0.060	-0.015	-0.010	-0.022	-0.021
	(-1.17)	(-0.48)	(-0.36)	(-0.72)	(-0.53)
工作时间	0.472 ***	0.320 ***	0.206 ***	0.145 **	0.070
	(5.81)	(6.54)	(4.36)	(2.90)	(1.07)
婚姻状况	-0.091	0.000 159	0.014	0.056	0.040
	(-0.63)	(0.00)	(0.17)	(0.64)	(0.34)
常数项	3.990 ***	7.202 ***	8.381 ***	9.206 ***	10.01 ***
	(7.73)	(23.12)	(27.89)	(28.99)	(24.09)
样本数	6,531	6,531	6,531	6,531	6,531

与上述 QR 估计类似,依条件收入分布的大学教育收益率仍然呈 L 型的异质性特征。这不同于已有研究中条件收入分布的教育收益率成“W 型”变动趋势^[16]、不断下降的趋势^[36],抑或不断上升的趋势^[37]。其中,在 10 分位点上系数最大,为 1.031(0.05 水平上显著);在 70 分位点上系数最小,为 0.841(0.01 水平上显著)。90 分位点上系数仅次于 10 分位数。这说明两端的低收入与高收入群体具有较高的大学教育收益率,其通过教育进行人力资本积累带来的收入提高最为明显。此信号指示作用则会引导低收入群体通过大学教育进行人力资本的梯度升级,缩小与中收入群体的收入差距;而高收入群体则通过大学教育扩展进一步提高收入,扩大了与其他群体的差距,验证了“马太效应”在一定程度上的存在。导致这种情况的解释可能为,改革开放以来我国市场经济体制不断完善,就业部门的所有制分布更加多元,户籍等影响就业的制度变迁加速了劳动力资源的有效配置;使得受教育水平较高、生产率较高的人力资本配置在高精尖的产业和产业,从而获得较高的经济回报;并伴随产业的进一步升级,通过教育扩展进一步提升生产力水平,获得更高的收入,一定程度上出现了收入差距扩大的“马太效应”。而中国转轨经济结构中,中低端产业在一段时间内将会占有相当的比重,其对生产率较低的劳动力就业是相当重要的。伴随产业的升级,此方面的人力资本通过教育扩展将有利于提高生产率。考虑到我国劳动力市场完全竞争程度日益提高,并按照生产要素的贡献实现人力资本的分配性功能,因此,会有效增加这方面劳动力的收入总量,减少收入差别。

控制变量方面,工作经验变量前的系数显著为正,而其平方项的系数则显著为负,说明工作经验与收入的相关关系呈非线性的倒 U 型形状,这与众多理论研究一致;工作时间与收入是显著的正相关,而智力水平、健康状况与婚姻状况与工资性收入没有显著的关系。

三、异质性分析与稳健性检验

(一) 异质性分析

已有研究表明,教育收益率在性别、城乡和区域等方面都有显著不同。本文使用工具变量分位数估计方法得到了表 4 的结果,并每隔 10 分位得到了 10 至 90 分位完整的依条件收入分布的大学教育收益率,见图 2。

表 4 性别、城乡和区域异质性

分位数	性别		城乡		区域		
	男	女	城市	乡村	东部	中部	西部
10 分位	1.356 [*] (1.910)	1.985 ^{***} (2.770)	1.745 ^{***} (2.710)	2.243 ^{**} (2.220)	1.099 (0.420)	2.059 ^{***} (4.300)	2.639 (1.430)
30 分位	0.935 ^{**} (1.960)	1.522 ^{**} (2.420)	1.112 ^{***} (3.050)	1.021 ^{**} (2.400)	1.156 ^{**} (2.080)	1.052 ^{**} (2.100)	1.365 ^{**} (2.110)
50 分位	0.540 (1.420)	1.316 ^{**} (2.220)	0.691 ^{**} (1.980)	0.562 (1.580)	1.090 ^{***} (2.860)	0.817 (1.610)	0.894 (0.560)
70 分位	0.605 [*] (1.760)	1.141 ^{**} (2.500)	0.905 ^{***} (3.250)	0.690 (1.020)	1.039 ^{***} (3.290)	0.850 ^{**} (2.190)	0.223 (0.200)
90 分位	0.367 (1.120)	1.499 (1.490)	1.122 ^{***} (3.080)	0.611 ^{**} (2.110)	1.506 ^{**} (1.970)	0.606 ^{***} (2.840)	0.198 (0.130)
控制变量	是	是	是	是	是	是	是
样本数	4 010	2 521	3 697	2 834	2 978	2 114	1 439

注:(1)表中汇报的是系数值,下方是相应的 Z 统计量;(2) *、**和 ***分别代表显著性水平 0.1、0.05 和 0.01,下同

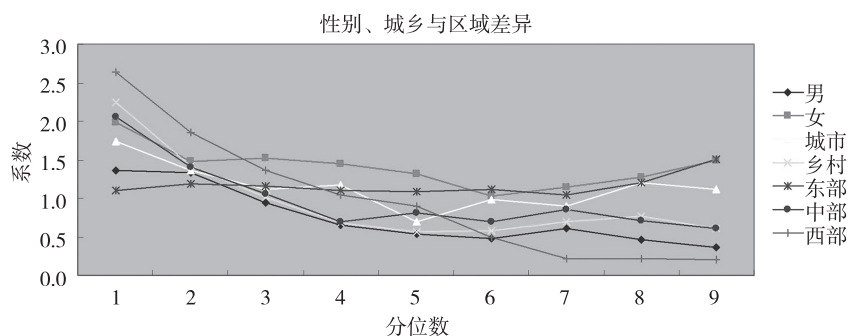


图 2 10 至 90 分位上大学教育收益率的变化

回归结果显示,大学教育收益率存在显著的男女性别差异。在所有分位点上,男性的大学教育系数都小于女性,这说明通过教育途径进行人力资本投资可以更好地增加女性的收入水平。其中 90 分位点上二者的系数差别最大,为 1.132,但未通过显著性检验。10 分位的系数都是各自最大的,为 1.356 和 1.985,分别在 0.1 和 0.01 水平上是显著的。这说明不分男女性别,低收入群体的大学教育收益率都是各自最大的,且女性比男性大 62.9%。这意味着无论男性还是女性,通过大学教育的人力资本积累都可

以显著缩小低收入群体与其他群体之间的收入差距。根据本文的截面数据,受过大学教育的人受教育时间平均是 14.74 年,而未受大学教育的人受教育时间平均是 7.95 年。因此,10 分位点上低收入群体的大学教育年度回报率,男性为 4.59%、女性为 10.62%,二者相差 6.03%。结合之前描述性统计的结果,男性的平均收入比女性要大,但大学教育收益率却比女性要小。这与已有研究的结论一致^[38]。这可能与女性在劳动力市场中面临的同工不同酬问题有关。女性通过接受更高水平的教育,能够显著提升自身的市场竞争力,从而获得更高的收入回报。此外,教育对女性收入的提升效果更为显著,可能是由于女性在职业发展中更依赖于教育背景来突破职业瓶颈,增加自己在就业市场的筹码,获得较高且公平的回报。女性在教育上投入更多,以获得较高且公平的收入回报,这也解释了为什么教育对女性收入的提升效果更为显著^[39]。

回归结果显示,大学教育收益率存在显著的城乡差异。其中,与性别相似,10 分位点低收入群体的城乡教育收益率都是最大的;但城市小于农村,分别是 0.01 水平上显著的 174.5%和 0.05 水平上显著的 224.3%,年度教育收益率分别为 8.55%和 12.63%。研究显示,城乡之间的教育回报率差异显著,且随着时间的推移,这种差异的演变表现出一定的趋势,城市教育回报率在高分位点上占优,而农村的教育回报率在低分位点上更为突出^[40]。这说明通过大学教育的人力资本积累可以较好地增加低收入群体的收入,减少收入差距,尤其表现在农村。但在其他分位点上,城市的教育回报率都大于农村,特别是 90 分位上二者的差距最大,为 51.1%。一个合理的解释可能是,我国城乡二元经济结构以及户籍体制的限制,造成农村劳动力难以获取城市正规部门就业的机会,难以获得城市户籍拥有者所能享受到的公共福利。即使高等教育是农村人口突破户籍限制的一条重要途径,但在许多城市以住房等固定资产投资为户籍进入的条件之下,原生家庭禀赋较少、经济资源和社会资源均面临稀缺硬约束的农村劳动力,无法获得与城市居民等量的就业机会与潜在收益。另外,包括城乡教育资源的不均衡配置等在内的外在环境因素,不仅通过影响个体努力等内在因素,影响个体的大学教育获得、就业机会获得和收入不平等,还会直接影响收入^[41]。根据回归结果,农村人口依条件收入分布的大学教育收益率不断下降,这有区别于已有研究中的“两头低、中间高”的“倒 U 型”结论^[42]。

回归结果显示,大学教育收益率存在显著的区域差异。东部,大学教育收益率依条件收入分布不断上升,其中 90 分位的系数最大为 1.506,在 0.05 水平上显著;而中西部则是不断下降,中部 10 分位的系数最大为 2.059,在 0.01 水平上显著;西部也是 10 分位的系数最大,但不显著。考虑到东部省份经济发展程度较高,中西部省份相对落后,上述结论可以解读为大学教育收益率在发达的东部依收入水平的提高而变大,而在欠发达的中西部地区则是呈现递减的情况。这与 Patrinos et al^[43]对发展程度不同国家的教育回报率分析相一致。导致这种情况的解释可能为,劳动力市场供需的不均衡造成了教育收益率的区域差异。东部地区对于劳动力的需求主要是能力较高的创新型高学历人才,根据资本论,其较高的生产率决定了较高的收益率。收入是能力的重要信号与表征,从而体现为东部收入分布较高分位上的回报系数较大。而东部对中低端劳动力的市场需求趋于饱和,均衡数量比较稳定,促使该部分劳动力向中西部迁移。而在全国产业布局中,中西部的比较优势在于中低端的产业类型,吸纳的就业人群也是生产率较低、收入较低的劳动力。依从教育的扩展,其生产率不断提升,从而带来收入的增幅会更明显,因此表现为中西部低收入群体的教育回报率会更高。另外,三个地区 30 分位的系数均显著,分别为 1.156、1.052 和 1.365,对应的年度教育收益率分别为 2.16%、0.75%和 4.59%。

(二) 稳健性检验

基于稳健性的考量,本文最后将个人工资性收入替代为个人年总收入再进行 IVQR 估计,如表 5 所示。结果表明,上述异质性结论是稳健的。

表 5 稳健性检验

分位数	性别		城乡		区域		
	男	女	城市	乡村	东部	中部	西部
10 分位	1.299 ** (2.040)	2.378 ** (2.290)	1.755 *** (2.770)	2.204 ** (2.100)	0.746 (0.520)	2.059 *** (4.330)	2.438 (1.450)
30 分位	0.935 ** (1.950)	1.465 ** (2.480)	1.115 *** (3.070)	1.021 ** (2.400)	1.159 ** (2.190)	1.052 ** (2.110)	1.365 *** (2.120)
50 分位	0.540 (1.420)	1.302 ** (2.320)	0.724 ** (1.950)	0.562 (1.580)	1.086 *** (2.790)	0.817 (1.610)	0.894 (0.560)
70 分位	0.593 * (1.750)	1.128 ** (2.490)	0.891 *** (3.150)	0.690 (1.020)	1.035 *** (3.220)	0.850 ** (2.190)	0.223 (0.200)
90 分位	0.367 (1.120)	1.495 (1.470)	1.108 *** (3.050)	0.611 ** (2.110)	1.488 ** (1.940)	0.606 *** (2.850)	0.192 (0.130)
控制变量	是	是	是	是	是	是	是
样本数	4 010	2 521	3 697	2 834	2 978	2 114	1 439

四、研究结论与相关建议

(一) 研究结论

基于中国家庭追踪调查的微观数据,本文利用考虑异质性与内生性的工具变量分位数回归模型系统研究了我国人力资本教育收益率在性别、城乡和区域之间的异质性特征。

第一,我国教育收益率依条件收入分布呈现 L 型形状,即低收入与高收入群体的收益率较高。这一发现意味着,低收入群体通过接受大学教育可以显著提升收入水平,缩小与其他群体的收入差距,从而打破贫困的代际传递;而高收入群体则利用优质教育资源进一步扩大其经济优势,这种“马太效应”在我国收入分配中尤为明显。这表明,教育在收入分布的不同层次中起到的作用存在显著差异,尤其在资源分配不均的背景下,这种差异更加突出。

第二,女性的教育回报率普遍高于男性,特别是在低收入群体中,女性的年度教育收益率达到 10.62%,显著高于男性的 4.59%。这一现象可能与女性在劳动力市场中面临的性别歧视及机会不平等有关,因此,通过高等教育,女性能够提升自身市场竞争力,获得更高的收入。由此可见,教育对缩小性别收入差距具有关键作用,尤其是在社会中下层,教育可以作为性别平等的重要推动力。

第三,城乡之间的大学教育回报率存在显著差异。总体来看,城市的大学教育回报率高于农村,这可能是由于城市教育资源更为丰富且质量更高,然而在低收入群体中,农村的教育回报率反而更高,表明农村地区的低收入群体通过教育获得显著的收入提升潜力,教育在这些群体的脱贫中起到重要作用。这进一步说明了教育资源的分布应更加注重公平,特别是要注重对农村和低收入群体的支持。

第四,区域经济发展水平显著影响教育回报率。东部地区由于经济发达、产业结构较为完善,教育收益率随着收入水平的提高而增加;而中西部地区的教育回报率则呈现递减趋势,反映出区域经济发展对教育投资回报的重要性。因此,提升中西部地区的经济水平和教育质量,成为提高这些地区教育回报率的关键因素。

(二) 相关建议

随着市场经济配置资源的机制更加完善,教育在收入分配中的作用将更加明显。根据本文的研究结论,大学教育扩展将提高人力资本积累带来的收入水平,有效解决收入悬殊问题。基于教育在我国乡村振兴中的基础性、先决性和关键性作用,加快建立健全教育缓解相对贫困的政策体系^[44]。

第一,应该深入实施教育反贫困战略,加大对相对贫困地区的高等教育政策倾斜,特别是在中西部和农村地区,增加录取名额,扩大招生规模,并提供更多的财政支持。通过这样的政策调整,可以有效减少区域间的教育不公平问题,进一步促进社会的整体公平发展。

第二,应提升中小学教育质量,特别是在中西部地区,通过加强师资力量的培训和改善教育设施,为学生提供更好的基础教育,从而为未来的高等教育奠定坚实基础。这不仅有助于提高这些地区的教育回报率,还能缩小城乡和区域间的教育差距。

第三,进一步加大对贫困家庭学生的资助,确保他们能够顺利完成高等教育,更好地融入劳动力市场。通过提供奖学金、助学贷款和其他形式的财政支持,可以有效减少因经济压力导致的辍学现象,增强教育的公平性。

第四,鉴于女性的教育收益率较高,政府应制定专门的政策,鼓励并支持女性接受高等教育,同时提供职业发展支持,以缩小性别收入差距。这对于实现更广泛的社会公平具有重要意义。

第五,优化区域教育资源配置,确保其与当地经济发展水平相匹配。特别是在东部地区,推动教育与高端产业结合,培养创新型人才;而在中西部地区,则根据当地的经济需求,调整教育资源配置,培养应用型人才,从而提高这些地区的教育回报率,促进区域经济的均衡发展。

[参 考 文 献]

- [1] 王晶,董日宇,高艳云. 中国家庭财富分布的极化特征:基于相对分布法的测度与分析[J]. 人口与经济, 2024(04): 70-82.
- [2] 张茂聪,黎敏辉,范晓婷. 教育何以促进共同富裕:基于 2003-2020 年省级面板数据的实证检验[J]. 教育研究, 2024(05): 132-149.
- [3] 杨宗凯. 高等教育数字化转型的路径探析[J]. 中国高教研究, 2023(03): 1-4.
- [4] 刘红奎. 反思与发展:高等教育个人收益理论研究[J]. 教育与经济, 2022(05): 89-96.
- [5] 管振,孙志军. 教育收益中的人力资本与信号效应估计:基于学制改革的证据[J]. 劳动经济研究, 2020(04): 3-20.
- [6] 王德文. 中国农村义务教育:现状、问题和出路[J]. 中国农村经济, 2003(11): 4-11.
- [7] 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.
- [8] 王轶,熊文,赖德胜. 乡村振兴战略下返乡农民教育收益与收入不平等:基于比较分析的视角[J]. 教育研究, 2019, 40(09): 120-138.
- [9] 刘生龙,胡鞍钢. 大学教育回报:基于大学扩招的自然实验[J]. 劳动经济研究, 2018(04): 48-70.
- [10] 于洪霞. 生命周期偏误、终身收入与中国教育收益率的估计[J]. 管理世界, 2014(12): 51-61.
- [11] 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.
- [12] Angrist J D, Krueger A B. Does compulsory school attendance affect schooling and earnings? The Quarterly Journal of Economics, 1991 (04): 979-1014.
- [13] Evans W N, Montgomery E. Education and health: Where there's smoke there's an instrument. NBER Working Paper 4949, National Bureau of Economic Research, 1994.
- [14] Butcher K F, Case A. The effect of sibling sex composition on women's education and earnings. The Quarterly Journal of Economics, 1994 (03): 531-563.
- [15] Card D. Earnings, schooling, and ability revisited. Research in Labor Economics, 1995(14): 23-48.

- [16] Harmon C, Walker I. The returns to the quantity and quality of education: Evidence for men in England and Wales. *Economica*, 2000(265):19-35.
- [17] Trostel P, Walker I, Woolley P. Estimates of the economic return to schooling for 28 countries. *Labour Economics*, 2002(01):1-16.
- [18] 郑猛. 教育扩张下流动人口教育收益率与收入差距[J]. *教育与经济*, 2017(05):48-60.
- [19] Brunello G, Fabbri D, Fort M. The causal effect of education on body mass evidence from Europe. *The Journal of Labour Economics*, 2013(01):195-223.
- [20] Brunello G, Fort M, Weber G. Changes in compulsory schooling, education and the distribution of wages in Europe. *Economic Journal*, 2009(536):516-539.
- [21] Fang H, Eggleston K N, Rizzo J A, et al. The returns to education in China: Evidence from the 1986 compulsory education law. NBER Working Paper 18189, National Bureau of Economic Research, 2012.
- [22] 刘生龙,周绍杰,胡鞍钢. 义务教育法与中国城镇教育回报率:基于断点回归设计[J]. *经济研究*, 2016(02):154-167.
- [23] 赵延东,洪岩璧. 社会资本与教育获得:网络资源与社会闭合的视角[J]. *社会学研究*, 2012(05):47-69+243-244.
- [24] 仇立平,肖日葵. 文化资本与社会地位获得:基于上海市的实证研究[J]. *中国社会科学*, 2011(06):121-135+223.
- [25] Card D. The causal effect of education on earnings. In: Ashenfelter, O., Card, D. (Eds.), *Handbook of Labor Economics*, *Handbook of Labor Economics* 3. Elsevier, 1999:1801-1863.
- [26] Arias O, Hallock K F, Sosa-Escudero W. Individual heterogeneity in the returns to schooling: instrumental variables quantile regression using twins data. *Empirical Economics*, 2001(26):7-40.
- [27] Card D, Krueger A B. Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of Political Economics*, 1992(01):1-40.
- [28] Heckman J, Polachek S. Empirical evidence on the functional form of the earnings-schooling relationship. *Journal of the American Statistical Association*, 1974(346):350-354.
- [29] Hungerford T, Solon G. Sheepskin effects in the returns to education. *Review of Economics Statistics*, 1987(01):175-177.
- [30] Koenker R, Bassett G. Regression quantiles. *Econometrica*, 1978(01):33-50.
- [31] Angrist J D, Pischke J S. *Mostly Harmless Econometrics: an Empiricist's Companion*. Princeton College Press, Princeton (USA), 2008.
- [32] Chernozhukov V, Hansen C. Instrumental quantile regression inference for structural and treatment effect models. *Journal of Economics*, 2006(02):491-525.
- [33] Chernozhukov V, Hansen C. Instrumental variable quantile regression: a robust inference approach. *Journal of Economics*, 2008(01):379-398.
- [34] Chernozhukov V, Hansen C. Quantile models with endogeneity. *Annu. Review of Economics*, 2013(01):57-81.
- [35] Li H, Luo Y. Reporting Errors, Ability Heterogeneity and Returns to Schooling in China, *Pacific Economic Review*, 2010(03):191-207.
- [36] 黄斌,高蒙蒙,查晨婷. 中国农村地区教育收益与收入差异[J]. *中国农村经济*, 2014(11):28-38.
- [37] Le W. How Does Education Affect the Earnings Distribution in Urban China? *Oxford Bulletin of Economics and Statistics*, 2013(03):1-48.
- [38] 张车伟. 人力资本回报率变化与收入差距:“马太效应”及其政策含义[J]. *经济研究*, 2006(12):59-70.
- [39] Liu, Y, Ning X. Literature review on education expansion and income gap in China. *International Journal of Science and Research Archive*, 2022(01):279-283.
- [40] Fu Q, Ren Q. Educational Inequality under China's Rural-Urban Divide: The Hukou System and Return to Education. *Environment and Planning A*, 2010(3):592-610.
- [41] 张永丽,李青原,郭世慧. 贫困地区农村教育收益率的性别差异[J]. *中国农村经济*, 2018(09):110-130.

- [42] 史新杰,卫龙宝,方师乐,高叙文. 中国收入分配中的机会不平等[J]. 管理世界,2018(03): 27-37.
- [43] Patrinos H A, Ridao-Cano C, Sakellariou C N. Estimating the Returns to Education: Accounting for Heterogeneity in Ability, World Bank Policy Research Working Paper No. 4040, 2006.
- [44] 王建. 教育缓解相对贫困的战略与政策思考[J]. 教育研究,2020(11): 11-21.

Human Capital Accumulation and Income Difference: Evidence from CFPS

Qin Changcai Sun Shiyu

(School of Economics and Management, Yantai University, Yantai Shandong 264005, China)

Abstract: Human capital accumulation through college education has become an important measure to narrow the income gap. Based on the micro-data of China Family Panel Studies, using the instrumental variable quantile regression method, the heterogeneity of returns to college education in China under the conditional income distribution was systematically analyzed from the perspective of full sample, gender, urban and rural, and regions. The results show that returns to college education in China is L-shaped according to the conditional income distribution, which means the low-income and high-income groups have higher yields. This means that low-income groups can narrow the income gap with other groups through college education. The “Matthew effect” of high-income groups to expand the income gap through college education exists at the same time. For all income groups, returns to college education for women are higher than men; the annual return rate for low-income groups is 4.59% for men and 10.62% for women. In addition to the low-income group of 10 points, the returns of urban are greater than of rural areas. The returns in the east increases with income levels, while the central and western regions are declining.

Keywords: returns to education; instrumental variables quantile regression; heterogeneity; CFPS

[责任编辑:左福生 邹柳馨]