

# 家庭背景对个人教育回报率的影响<sup>\*</sup>

刘泽云 袁青青

**【摘要】**文章基于2018年中国家庭收入调查数据,从父母受教育水平、政治面貌、职业地位、社会关系和城乡背景等方面分析家庭背景对个人教育回报率的影响。研究发现:(1)父母受教育水平对个人的教育回报率存在显著的正向影响,意味着家庭教育背景的优势会通过子女的教育投资转化为子女收入的优势;(2)父母社会关系和父母从事非农就业对个人的教育回报率有显著负向影响,说明教育更有利于提高出身于家庭社会关系较弱和农民家庭的个人的收入;(3)父母的政治面貌及职业地位对个人的教育回报率的影响不显著,表明家庭政治背景和职业背景通过教育实现代际传递的作用已经弱化。基于以上结论,文章建议在优质教育资源的配置上向更加均衡的方向努力,为不同家庭出身的劳动者创造更为公平的竞争环境。

**【关键词】**家庭背景 人力资本 教育回报率

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

## 一、引言

教育回报率也称教育收益率,用于衡量个人从教育投资中获得的经济回报。无论是针对中国的教育回报率研究,还是针对其他国家的教育回报率研究,均证实个人的教育投资能够提高其未来的收入水平(刘泽云、刘佳璇,2020;Psacharopoulos等,2018)。作为最重要的人力资本投资形式,教育被寄予了改善收入分配和提高代际流动性的期望。如果家庭背景较差的个体能够从教育投资中获得更高的经济收益,那么教育可以在一定程度上有助于“寒门出贵子”,有利于缩小收入差距和增进代际收入流动。然而,如果家庭背景较好的个体获得更高的教育回报,教育则会阻碍“寒门出贵子”,加大收入差距导致社会阶层固化。

目前,国内外关于家庭背景对个人教育回报率影响的研究得出的结论并不相同。针对中国的相关研究认为,家庭背景对个人教育回报率具有正向影响,即家庭背景较好的个

<sup>\*</sup> 本文为国家社会科学基金重大项目“中国农村家庭数据库建设及其应用研究”(编号:18ZDA080)的阶段性研究成果。

人教育回报率更高(袁诚、张磊,2009;祁翔、周金燕,2015;吴延科、田茂再,2016)。然而,针对美国、德国、西班牙、意大利等发达国家的研究得出了不同的结论,主要有3种观点:(1)家庭背景越好,个人的教育回报率越高(Avram等,2017)。(2)家庭背景较差的个人教育回报率更高(Feigenbaum等,2019)。(3)父亲和母亲的影响存在差异。Deschênes(2007)针对美国的研究和Cornelissen等(2008)针对德国的研究均发现,父亲受教育水平对个人教育回报率的影响显著为正,母亲受教育水平对个人教育回报率的影响显著为负。可见,家庭背景和个人教育回报率的关系有其复杂性,需要进一步研究。另外,现有针对中国的相关研究主要使用的是2010年之前的数据,这些研究结论目前是否仍然成立,家庭背景对子女教育回报率的影响是否发生了变化,值得关注。

本文对相关文献进行梳理后,发现存在3个方面的不足。一是在研究内容方面,现有文献大多仅分析某一类家庭背景对子女教育回报率的影响。二是在研究方法方面,现有研究主要依据家庭背景变量分组估计明瑟收入方程,通过比较不同组别的教育回报率差异来判断家庭背景与教育回报率的关系(袁诚、张磊,2009;祁翔、周金燕,2015;吴延科、田茂再,2016)。这种方法假设明瑟收入方程中所有解释变量的回归系数均存在组间差异,是一个过于强的假设。此外,个人受教育年限的内生性是估计过程中必须考虑的问题,但在已有文献中只有少量研究基于工具变量法(IV)处理内生性问题(袁诚、张磊,2009)。三是在分析样本方面。考察家庭背景与个人教育回报率的关系,必须同时收集个体及其家庭背景信息,也就是需要两代人的信息。在国内已有相关研究中,要么以户主的子女作为分析对象,户主及其配偶作为父母(袁诚、张磊,2009);要么以户主及其配偶为分析对象,户主及其配偶的父母作为父母(祁翔、周金燕,2015;张世伟、吕世斌,2008;王孙禺、范静波,2011)。前者会导致同住样本选择偏差,即只使用住在同一家庭户的样本;后者则会遗漏被调查户中的户主子女样本。无论单独使用上述哪一类样本均会导致样本的代表性不足,造成估计结果的偏差。鉴于此,本文基于2018年中国家庭收入调查数据,从父母受教育水平、政治面貌、职业地位、社会关系和城乡背景等多个维度考察不同类型家庭背景对个人教育回报率的影响。

## 二、模型与方法

### (一) 模型设定

本文通过在明瑟收入方程中加入家庭背景及其与个人受教育年限的交互项来识别家庭背景对个人教育回报率的影响(Björklund, 1994):

$$\log Y = \beta_0 + \beta_1 S + \beta_2 F + \beta_3 S \times F + \beta_4 M + \varepsilon \quad (1)$$

式(1)中, $\log Y$ 为个人年工资性收入的自然对数; $S$ 为个人受教育年限; $F$ 为个人家

庭背景; $S \times F$ 为家庭背景与个人受教育年限的交互项; $M$ 为控制变量,包括性别、工作经验、工作经验的平方、户籍类型、住户类型和居住省份的虚拟变量; $\varepsilon$ 为随机误差项。本文没有控制个人的职业、行业 and 单位所有制性质等工作特征变量,原因在于这些变量是本文关注的解释变量(即受教育年限)的结果变量,是“不好的”控制变量(Angrist等,2008),控制了这些变量反而会导致教育回报率的估计出现偏误。

家庭背景与受教育年限的交互项的系数 $\beta_3$ 反映家庭背景对个人教育回报率的影响。在教育回报率为正的情况下,如果父母受教育年限与个人受教育年限的交互项系数为正,说明父母受教育年限对个人教育回报率有正向影响,反之有负向影响。

(二) 内生性问题

在教育回报率的估计中,内生性问题主要来自测量误差和遗漏变量偏误。测量误差是指由于收入变量或教育变量与其实际值之间存在偏差,从而导致教育回报率的 OLS 估计有偏且不一致。2018 年中国家庭收入调查提供两类个人收入数据:一类是个人自报的收入,与个人的实际收入可能存在较大偏差;另一类是从国家统计局过录的个人收入,是基于个人日常记账得到的收入数据,准确程度相对较高。本文使用后一类收入数据,收入变量存在测量误差的可能性较小。另外,本文使用受访者填写的接受正规教育的年数衡量个人的受教育年限,避免在教育变量上的测量误差。遗漏变量偏误是指个人能力和教育偏好等不可观测的因素被遗漏在误差项中,这些因素既与个人的受教育年限相关,又与个人的收入相关,会导致教育回报率的 OLS 估计有偏且不一致。本文使用兄弟姐妹数量作为个人受教育年限的工具变量来处理可能存在的遗漏变量偏误。兄弟姐妹数量越多,教育支出受家庭预算约束的可能性越大,故可能接受更少的教育,而兄弟姐妹数量与个人未来的工资收入并没有直接的联系,满足工具变量的使用条件。具体模型估计使用两阶段最小二乘法(2SLS)和控制函数法(CF)。

两阶段最小二乘法(2SLS)是解决内生性问题的标准方法。在式(1)中,个人受教育年限  $S$  是内生变量,家庭背景与个人受教育年限的交互项  $S \times F$  也是内生变量。对于这样的模型,可以分别选用工具变量及其与外生变量的交互项作为工具变量(Balli等,2013)。而且在有内生变量的模型中,如果加入内生变量与外生变量的交互项,会减轻模型的内生性问题(Bun等,2019)。本文以兄弟姐妹数量( $Z$ )作为个人受教育年限( $S$ )的工具变量,以兄弟姐妹数量与家庭背景的交互项作为个人受教育年限与家庭背景的交互项的工具变量,运用 2SLS 进行估计的步骤如下。

第一步:分别以内生变量  $S$  和  $S \times F$  为被解释变量,以工具变量  $Z$ 、 $Z \times F$  和控制变量  $M$  为解释变量进行 OLS 回归。

$$S = \alpha_0 + \alpha_1 F + \alpha_2 Z + \alpha_3 Z \times F + \alpha_4 M + \mu \tag{2}$$

$$S \times F = \delta_0 + \delta_1 F + \delta_2 Z + \delta_3 Z \times F + \delta_4 M + e \tag{3}$$

第二步:将第一步回归得到的拟合值 $\hat{S}$ 和 $\widehat{S \times F}$ 带入式(1)中进行 OLS 估计。

$$\log Y = \theta_0 + \theta_1 \hat{S} + \theta_2 F + \theta_3 \widehat{S \times F} + \theta_4 M + \varepsilon \quad (4)$$

控制函数法(CF)通过控制第一阶段估计的误差项来解决遗漏变量偏误,本质上也是基于工具变量处理内生性问题的方法(Wooldridge, 2015)。运用 CF 估计的步骤如下。

第一步:将内生变量  $S$  对工具变量  $Z$  和控制变量  $M$  进行回归。

$$S = \pi_0 + \pi_1 F + \pi_2 Z + \pi_3 M + \omega \quad (5)$$

第二步:将第一步估计的残差 $\hat{\omega}$ 带入式(1)进行 OLS 估计。

$$\log Y = \gamma_0 + \gamma_1 S + \gamma_2 F + \gamma_3 S \times F + \gamma_4 M + \gamma_5 \hat{\omega} + \varepsilon \quad (6)$$

与 2SLS 估计不同,CF 通过考察式(6)中残差项的系数来检验模型是否存在内生性问题。另外,当模型中的内生变量以非线性形式出现时,CF 将不可分离的内生变量作为整体进行研究,增加了估计的有效性(Wooldridge, 2015)。

### 三、数据来源与变量设置

#### (一) 数据来源

本文使用的 2018 年中国家庭收入调查数据,由北京师范大学中国收入分配研究院中国居民收入分配课题依托国家统计局调查样本库,采取系统抽样的方法获得。该数据覆盖 15 个省份,调查涉及 2 万多个家庭 7 万余人。

如前所述,本文分析样本包括两类:一类以户主的同住子女为分析对象,另一类以户主及其配偶为分析对象。样本筛选还考虑以下几个方面:(1)由于 CHIP2018 问卷仅对 50 岁以下的户主及其配偶调查了其父母的背景信息,因此样本的年龄限定为 16~50 岁<sup>①</sup>。(2)分析对象为劳动力人口,即 2018 年末处于就业状态的劳动者,不包括离退休人员、在校学生、失业或待业人员、家务劳动者等不工作或上学的个人。(3)限定样本为 2018 年主要工作身份为雇员的个人,这也是教育回报率相关研究中通行的做法。(4)删去没有工资性收入的个体。(5)删除回归模型中相关变量缺失信息的个体。最终获得有效样本 12 465 人。

#### (二) 变量设置

##### 1. 被解释变量

本文被解释变量为个人年工资性收入的自然对数,使用从国家统计局过录的个人工资性收入数据。

<sup>①</sup> 在户主的同住子女样本中,经筛选后 51~60 岁的样本为 341 人,样本量很小,而且加入这部分样本并不影响本文的结论。为保证两类样本在年龄范围上的一致性,文中只使用 16~50 岁的样本。



2. 核心解释变量

本文核心解释变量包括个人受教育年限和家庭背景。其中,个人受教育年限指个人接受正规教育的年数(不包括跳级、复读和留级年数)。

家庭背景从5个方面衡量:(1)父母受教育水平。为了综合考虑父母双方的影响,用父母的最高受教育年限衡量。调查在询问户主及其配偶的父母信息时,只询问了受教育程度。因此,父母受教育年限根据受教育程度折算,即受教育程度为未上学、小学、初中、高中(职高或技校、中专)、大专、本科和研究生分别折算成受教育年限为0年、6年、9年、12年、15年、16年和19年。(2)父母政治面貌。用父母的党员身份衡量,如果父母任意一方为中共党员,父母党员=1,否则=0。(3)父母职业地位。调查问卷中将职业分为8种类型:单位(部门)负责人;专业技术人员;办事人员和有关人员;商业、服务业人员;农林渔牧和水利业生产人员;生产、运输设备操作人员及有关人员;军人;不便分类的其他从业人员。参考沈艳、张恺(2015)的研究,本文将单位(部门)负责人和专业技术人员定义为管理技术类职业,其他六类职业定义为非管理技术类职业。如果父母任意一方的职业为管理技术类,父母职业地位较高=1,否则=0。(4)父母社会关系。用个人在求职过程中是否依靠非正式网络途径来衡量(赵延东、罗家德,2005),如果个人通过家人联系、亲戚介绍或接班3种方式获得当前工作,社会关系良好=1,否则=0<sup>①</sup>。(5)父母城乡背景。本文用父母是否从事非农就业衡量,即如果父母任意一方从事非农就业工作,父母非农就业=1,否则=0。

相关研究还使用当前父母收入(袁诚、张磊,2009)或当前家庭收入(刘阳阳、王瑞,2017)衡量家庭背景。这样会面临反向因果的问题,即个人收入可能会影响父母的收入和整个家庭的收入。此外,在调查户主及其配偶的父母信息时,没有询问父母的收入。因此,本文没有考察收入因素的影响。

3. 控制变量

控制变量包括:(1)工作经验。指潜在的工作经验,计算公式为:工作经验=年龄-个人受教育年限-6。(2)性别。男性为1,女性为0。(3)户籍类型。非农户籍为1,农业户籍为0。参考赵西亮(2017)的研究,对于经历过户籍转换的个人,其户籍类型还原为初始的户籍类型。(4)住户类型。分为农村住户、城镇住户和流动人口三类。(5)居住省份。用一组代表个人当前居住地所在省份的虚拟变量表示。变量描述性统计如表1所示。

① 调查问卷中对应的问题是:“你是如何找到这份工作的?”。对应的选项为:政府分配、安排;政府职介;社区就业服务站;工作调动(含升职);商业职介(包括人才交流会);看到广告后申请;考试录取;家人联系;亲戚介绍;朋友或熟人介绍;雇主招工;接班;其他。

#### 四、父母受教育水平对个人教育回报率的影响

##### (一) 基准回归结果

从表 2 可以看出,用 OLS 方法估计的教育回报率为 10.96%。OLS、2SLS 和 CF 的估计结果均显示,个人受教育年限与父母受教育年限的交互项在 1%的水平上显著,父母受教育水平对个人教育回报率有显著正向影响,即家庭教育背景越好,个人的教育回报率越高。家庭教育背景对个人教育回报率的正向影响可能源于较高受教育水平的父母能够为子女做出更好的教育选择(如学校选择和专业选择),并更好地指导子女的学习(包括在辅导学校课程和作业方面做得更好、更合理地选择课外辅导班等),使子女从接受更高质量的教育中获得额外的经济回报。

在 2SLS 估计(模型 3)中,Kleibergen-Paaprk LM 检验和 Kleibergen-Paaprk Wald F 检验都表明不存在弱工具变量问题,但内生性检验(DWH 检验)显示,2SLS 和 OLS 估计不存在显著差异,即模型内生性问题得到较好控制。同

表 1 主要变量的描述性统计

变 量	均值	标准差	最小值	最大值	样本量
年工资性收入(元)	47522	37318	500	228259	12465
男性	0.58	0.49	0	1	12465
年龄	37.19	8.00	16	50	12465
受教育年限(年)	11.37	3.59	0	21	12465
工作经验(年)	19.81	9.73	0	44	12465
兄弟姐妹数量	1.69	1.40	0	10	12465
父母最高受教育年限(年)	7.75	3.72	0	19	12465
父母党员	0.18	0.39	0	1	12406
父母职业地位较高	0.11	0.31	0	1	11638
父母社会关系良好	0.14	0.34	0	1	12465
父母非农就业	0.70	0.46	0	1	12386

表 2 父母受教育水平对个人教育回报率的影响:基准回归结果(N=12465)

变 量	OLS		2SLS	CF
	模型 1	模型 2	模型 3	模型 4
受教育年限	0.1096*** (0.0030)	0.0933*** (0.0053)	0.1192*** (0.0300)	0.1237*** (0.0268)
父母最高受教育年限		-0.0188*** (0.0069)	-0.0314* (0.0177)	-0.0224*** (0.0075)
受教育年限×父母最高受教育年限		0.0020*** (0.0006)	0.0028* (0.0016)	0.0020*** (0.0006)
第一阶段残差项				-0.0309 (0.0268)
R <sup>2</sup>	0.387	0.388	0.382	0.388
内生性检验(DWH 检验)			2.216	
弱工具变量检验				
Kleibergen-Paaprk LM 检验			143.369***	
Kleibergen-Paaprk Wald F 检验			75.461***	

注:括号中数据为稳健标准误;所有回归方程均控制了性别、工作经验、工作经验的平方、户籍类型、住户类型和居住省份虚拟变量,表中略去这些变量的回归系数;\*、\*\*、\*\*\* 分别表示在 10%、5%、1%的水平上显著。

时,从模型4的CF估计结果看,第一阶段残差项的系数不显著,也说明模型在内生性问题得到较好控制。由于父母受教育水平能够较好地捕捉到个人能力和不可观测的家庭背景的信息,在回归中控制了父母受教育水平能在很大程度上克服遗漏变量偏误。同时,正如前文所述,在有内生变量的模型中,如果加入内生变量与外生变量的交互项,会减轻模型的内生性问题。因此,只需关注OLS估计结果。根据模型2的OLS估计结果,父母受教育年限每提高1年,个人的教育回报率将提高0.2个百分点。这意味着父母受教育水平为大学本科(受教育年限16年)的个体,其教育回报率比父母受教育水平为小学(受教育年限6年)的个体高2个百分点。另外,父母受教育年限的系数显著为负(即父母受教育年限越高,个人的收入越低),这一结论似乎与常理不符。但实际上,这一变量系数的含义是在个人受教育年限为0时,父母教育对个人收入的影响为负。在本文的样本中,个人受教育年限为0的样本只有42个(约占样本总数的0.3%),因此,虽然这一变量系数显著为负,但没有太大的实际意义<sup>①</sup>。

## (二) 稳健性检验

为了考察估计结果的稳健性,本文从以下3个方面进行稳健性检验。

### 1. 使用其他指标衡量父母受教育水平

在基准回归中,使用父亲和母亲中受教育水平较高一方的受教育年限衡量父母受教育水平。在稳健性检验中,分别使用父亲受教育年限、母亲受教育年限和父母平均受教育年限衡量父母受教育水平,使用与基准回归相同的模型进行了回归分析,结果如表3所示。其中,模型5、模型8、模型11使用父亲受教育年限,模型6、模型9、模型12使用母亲受教育年限,模型7、模型10、模型13使用父母平均受教育年限。可以看出,不论使用什么指标衡量父母受教育水平,2SLS和CF估计结果均表明模型内生性问题得到较好控制,只需要关注OLS估计结果。OLS估计结果均显示个人受教育年限与父母受教育水平交互项的系数显著为正,表明父母受教育水平对子女教育回报率存在显著正向影响。另外,使用父母平均受教育年限得到的交互项系数绝对值,稍大于单独使用父亲或母亲受教育年限得到的交互项系数绝对值,说明父母综合的教育水平对子女的影响大于父母某一方教育水平的影响,这是值得进一步研究的问题。

### 2. 使用自报的收入

为了尽可能减小收入变量的测量误差,基准回归中使用CHIP项目组从国家统计局过录的收入数据。但在整理数据的过程中发现,有相当数量的个体在国家统计局过录数

<sup>①</sup> 事实上,随着个人受教育年限的提高,父母教育对收入的影响会逐渐转变为正向影响。譬如,对于小学毕业、高中毕业和大学本科毕业的个体(即个人受教育年限分别为6年、12年、16年),父母受教育年限对个人收入的边际影响分别为-0.0068、0.0052、0.0132。

表 3 父母受教育水平对个人教育回报率的影响:用其他指标衡量父母受教育水平(N=12465)

变 量	OLS			2SLS			CF		
	模型 5	模型 6	模型 7	模型 8	模型 9	模型 10	模型 11	模型 12	模型 13
受教育年限	0.093*** (0.005)	0.099*** (0.004)	0.093*** (0.005)	0.117*** (0.031)	0.135*** (0.028)	0.125*** (0.028)	0.123*** (0.028)	0.136*** (0.027)	0.069*** (0.018)
父亲受教育年限	-0.021*** (0.007)			-0.037** (0.018)			-0.025*** (0.008)		
受教育年限×父亲受教育年限	0.002*** (0.001)			0.003** (0.002)			0.002*** (0.001)		
母亲受教育年限		-0.017** (0.007)			-0.0210 (0.017)			-0.019*** (0.010)	
受教育年限×母亲受教育年限		0.002*** (0.001)			0.0020 (0.002)			0.002*** (0.001)	
父母平均受教育年限			-0.024*** (0.008)			-0.034* (0.019)			-0.016* (0.010)
受教育年限×父母平均受教育年限			0.003*** (0.001)			0.003* (0.002)			0.003*** (0.001)
第一阶段残差项							-0.031 (0.028)	-0.037 (0.027)	0.0230 (0.017)
内生性检验(DWH 检验)				2.451	1.962	2.360			
弱工具变量检验									
Kleibergen-Paaprk LM 检验				134.1***	145.3***	154.9***			
Kleibergen-Paaprk Wald F 检验				70.4***	73.4***	81.3***			

注:括号中数据为稳健标准误。其他控制变量同表 2。\*、\*\*、\*\*\* 分别表示在 10%、5%、1%的水平上显著。

据中没有收入,却有自报的收入数据。由于不清楚是什么原因导致这一差异,为确保结果的稳健性,在稳健性检验中基于相同的模型和方法,用自报的收入数据进行分析,结果如表 4 所示。可以看出,使用自报收入数据后,样本量增加到 16 159 个,而使用国家统计局过录数据的样本量为 12 465 个,前者比后者多了近 30%的样本。2SLS 和 CF 估计结果表明模型内生性问题得到较好控制,而且在 OLS 估计结果中,交互项系数显著为正,依然表明父母受教育水平对子女的教育回报率存在正向影响。交互项系数的 OLS 估计值为 0.0021,与表 2 的 OLS 估计值(0.0020)非常接近。

### 3. 非线性影响

在基准回归中,父母受教育水平对个人教育回报率的影响是线性的,也就是说无论父母是何种受教育水平,父母受教育水平对个人教育回报率的影响均不变。为了检验父母教育对个人教育回报率是否存在非线性影响,在基准模型中加入个人受教育年限与父母受教育年限平方的交互项进行了回归分析。结果显示,个人受教育年限与父母受教育年限平方的交互项不显著,说明父母教育对个人教育回报率不存在非线性影响。



表4 父母受教育水平对个人教育回报率的影响:使用自报的收入数据(N=16159)

变 量	OLS	2SLS	CF
受教育年限	0.0600*** (0.0039)	0.0827*** (0.0274)	0.0745*** (0.0255)
父母最高受教育年限	-0.0216*** (0.0049)	-0.0020 (0.0131)	-0.0235*** (0.0059)
受教育年限×父母最高受教育年限	0.0021*** (0.0004)	0.0002 (0.0012)	0.0021*** (0.0004)
第一阶段残差项			-0.0147 (0.0255)
内生性检验(DWH 检验)		2.853	
弱工具变量检验			
Kleibergen-Paaprk LM 检验		124.078***	
Kleibergen-Paaprk Wald F 检验		62.426***	

注:括号中数据为稳健标准误。其他控制变量同表2。\*、\*\*、\*\*\* 分别表示在10%、5%、1%的水平上显著。

### 五、其他家庭背景对个人教育回报率的影响

下面使用父母的政治面貌、职业地位、社会关系、城乡背景等衡量家庭背景,分析其对个人教育回报率的影响。与前文一样,重点考察这些家庭背景变量与个人受教育年限的交互项(见表5)。由于父母受教育水平对其政治面貌、职业地位、社会关系、是否从事非农工作等具有因果性影响,因此在考察后者对个人工资收入的影响时,需要控制父母受教育水平的相关变量。同时,前面的分析表明,在控制父母受教育水平及其与个人受教育年限的交互项后,回归模型中内生性问题得到较好控制,因此文中只呈现 OLS 的估计结果。表5呈现的结果如下。

第一,个人受教育年限与父母党员身份的交互项系数不显著,表明以父母党员身份衡量的家庭政治面貌对个人的教育回报率没有影响。这一发现与已有研究结论不同,可

表5 其他家庭背景对个人教育回报率的影响(OLS)

变 量	父母政治面貌	父母职业地位	父母社会关系	父母城乡背景
受教育年限	0.0937*** (0.0053)	0.0963*** (0.0055)	0.0953*** (0.0053)	0.1054*** (0.0057)
父母党员	-0.0056 (0.0824)			
受教育年限×父母党员	0.0010 (0.0062)			
父母职业地位较高		-0.0943 (0.1154)		
受教育年限×父母职业地位较高		0.0071 (0.0082)		
父母社会关系良好			0.1532** (0.0751)	
受教育年限×父母社会关系良好			-0.0200*** (0.0067)	
父母非农就业				0.1939*** (0.0556)
受教育年限×父母非农就业				-0.0245*** (0.0047)
父母最高受教育年限	-0.0186*** (0.0072)	-0.0151** (0.0074)	-0.0175** (0.0069)	-0.0234*** (0.0070)
受教育年限×父母最高受教育年限	0.0019*** (0.0006)	0.0017*** (0.0006)	0.0019*** (0.0006)	0.0026*** (0.0006)
样本量	12406	11638	12465	12386
R <sup>2</sup>	0.389	0.386	0.389	0.392

注:括号中为稳健标准误。其他控制变量同表2。\*、\*\*、\*\*\* 分别表示在10%、5%、1%的水平上显著。

能是已有文献反映的是较早年份的情况,而随着市场经济体制的逐步完善,家庭政治面貌在通过教育实现代际传递中的作用已经弱化。

第二,个人受教育年限与父母职业地位的交互项系数不显著,表明以父母是否从事管理技术类职业衡量的家庭职业地位并不显著影响个人的教育回报率。本文还根据父母是否在体制内单位工作来衡量其职业地位(如果父母任意一方在党政机关团体、事业单位或国有企业工作,父母职业地位较高=1,否则=0),发现研究结论没有变化。这一发现也与已有研究不同,其原因除了随着时间的推移,父母职业地位的影响已发生变化外,父母的受教育水平也会影响父母的职业地位,因此后者的影响已被前者很好地解释。在关于父母职业地位的回归中去掉父母最高受教育年限及其与个人受教育年限的交互项后,发现个人受教育年限与父母职业地位的交互项系数从 0.0071 大幅提高到 0.0135,且在 10%的水平上显著,一定程度上证实了上述观点。因此,如果在考察父母职业地位的影响时不控制父母受教育水平,可能会得到错误的估计结果。此外,父母职业地位较高的家庭能够为子女的教育和就业提供更多的支持,但这些父母有可能将更多的时间投入工作,导致对子女的时间投入减少。也就是说,父母职业地位较高对个人教育回报率的影响可能为正也可能为负,从而整体的影响方向是不确定的。

第三,个人受教育年限与父母社会关系良好的交互项系数为 -0.02,且统计显著,即父母社会关系良好的个人比父母社会关系较差的个人的教育回报率低 2 个百分点。表明父母具有良好的社会关系对于个人教育回报率有显著负向影响。在界定父母社会关系时,本文未将通过“熟人或朋友介绍”获得当前工作视为父母社会关系良好,因为这更多地反映了个人的社会关系而不是父母的社会关系。但进一步分析发现,即使将通过“熟人或朋友介绍”获得当前工作个人也定义为父母社会关系良好,研究结论也没有发生改变。其原因可能源于两个方面。一是如果父母缺乏社会关系,会将提高子女社会经济地位的希望寄托于教育,并给予子女更多的教育投入。同时,父母社会资本少的个人也会更加意识到教育对改变自身命运的重要性,从而更加努力学习。因此,与家庭社会资本丰厚的个人相比,家庭社会资本不足的个人会更有效地利用受教育的机会,在受教育期间积累更高质量的人力资本,进而从教育中获得更高的经济回报。二是父母社会关系在个人求职过程中发挥重要作用,在劳动力市场化程度较低的地区,个人依靠父母社会关系获得工作的比例较高,教育回报率较低,父母社会关系与个人教育回报率呈负相关。此外,从表 5 可以看出,父母社会关系良好对个人收入的影响为正,体现出“寒门难出贵子”。然而,父母社会关系对个人教育回报率有负向影响,父母社会关系较弱的个人虽然没有太多的社会关系可以依赖,但可以从教育中获得更高的经济收益,教育起到了缓解“寒门难出贵子”的作用。

第四,个人受教育年限与父母非农就业的交互项系数为 -0.0245,且统计显著,即父

母中至少有一方从事非农就业的个人的教育收益率比父母均务农的个人低 2.45 个百分点。这表明,农民家庭出身的个人从教育中获得了更高的经济回报。可能的原因是农民家庭在教育资源获取、工作获得等方面的条件较差,农民家庭出身的个人会更加珍惜教育机会,会更加努力地学习以积累人力资本,故得到较高的教育回报。

表 5 显示,在分别加入父母政治面貌、职业地位、社会关系或城乡背景后,个人受教育年限与父母受教育年限交互项的系数仍显著为正,这也说明父母受教育水平对个人教育回报率的影响是稳健的。

## 六、结论与启示

本文基于 2018 年中国家庭收入调查数据,分析了不同类型的家庭背景对个人教育回报率的影响。现有关于中国的研究都发现家庭背景较好的个人的教育回报率更高,但本文发现家庭背景对个人教育回报率的影响较为复杂,使用不同指标衡量的家庭背景对个人教育回报率的影响有所不同。本文的主要研究结论是:(1)父母受教育水平对个人的教育回报率存在显著的正向影响,父母受教育年限提高 1 年,个人的教育回报率将提高约 0.2 个百分点。(2)父母社会关系及父母从事非农就业对个人的教育回报率存在显著负向影响,父母社会关系良好的个人比父母社会关系较差的个人教育回报率低 2 个百分点,父母中至少有一方从事非农就业的个人的教育收益率比父母均务农的个人低 2.45 个百分点;(3)父母的政治面貌及职业地位对个人教育回报率的影响不显著。

基于上述研究结论,在中国现阶段,对于教育在改善收入分配和促进代际流动性的作用,需要从以下两个角度去认识。

第一,对于父母社会关系较弱或务农的个人,教育回报率更高,意味着从人力资本投资的角度看,目前的教育体系和劳动力市场环境在一定程度上有利于这些处于家庭背景不利地位的群体,从而教育在缩小收入差距和提高代际流动性方面发挥了一定的作用。父母社会关系较弱或务农的个人的教育回报率更高,可能是因为这些人更加珍惜教育机会,会更加努力地学习以积累人力资本,进而从教育中获得更高的边际回报。然而,教育回报率更高并不必然意味着收入更高,也不意味着通过教育就一定能实现阶层的跨越。因此,应该在教育体制改革和教育资源配置中对家庭背景不利的群体继续给予更多关注;同时应积极推进劳动力市场的市场化改革,为不同家庭出身的劳动者创造更公平的竞争环境,让人才的招聘和任用取决于个人的能力而不是家庭出身或社会关系。

第二,父母受教育水平较高的子女能够从其教育投资中获得更高的收益,表明教育投资在一定程度上起到了阶层固化的作用,不利于缩小收入差距和提高代际流动性。家庭教育背景对个人教育回报率的正向影响可能源于高教育水平的父母能够为子女做出更好的教育选择,并更好地指导子女的学习,从而子女能够接受更高质量的教育。因此,如果想要让教育在促进代际流动中发挥积极作用,应该在优质教育资源(特别是学前教

育资源和义务教育资源)的配置上向更加均衡的方向努力。

考虑到本文存在的局限,未来研究应关注以下两个方面:一是对家庭背景影响个人教育回报率的机制展开深入的理论探讨和经验分析;二是寻找合适的微观数据资源,分析家庭经济因素对个人教育回报率的影响。

#### 参考文献:

1. 刘泽云、刘佳璇(2020):《中国教育收益率的元分析》,《北京师范大学学报(社会科学版)》,第7期。
2. 祁翔、周金燕(2015):《教育回报率的家庭背景差异》,《教育科学》,第3期。
3. 沈艳、张恺(2015):《家庭背景对我国高等教育入学机会的影响——基于2013届高校毕业生调查的实证分析》,《教育学术月刊》,第5期。
4. 王孙禺、范静波(2011):《文凭信号、职业因素与家庭背景对教育收益的影响研究》,《华东师范大学学报(哲学社会科学版)》,第6期。
5. 吴延科、田茂再(2016):《社会经济地位与教育回报率的无条件分位回归分析——基于CGSS2010年调查数据》,《数理统计与管理》,第4期。
6. 袁诚、张磊(2009):《对低收入家庭子女大学收益的观察》,《经济研究》,第5期。
7. 赵西亮(2017):《教育、户籍转换与城乡教育收益率差异》,《经济研究》,第12期。
8. 赵延东、罗家德(2005):《如何测量社会资本:一个经验研究综述》,《国外社会科学》,第2期。
9. 张世伟、吕世斌(2008):《家庭教育背景对个人教育回报和收入的影响》,《人口学刊》,第3期。
10. Angrist J.D., Pischke J.S. (2008), *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.
11. Avram S., Cantó-Sánchez O. (2017), Labour Outcomes and Family Background: Evidence from the EU during the Recession. ISER Working Paper Series No 2017-15.
12. Balli H.O., Sørensen B.E. (2013), Interaction Effects in Econometrics. *Empirical Economics*. 45(1):583-603.
13. Björklund A. (1994), The Impact of Family Background on the Returns to and Length of Schooling in Sweden. *Human Capital Creation in an Economic Perspective*. Physica-Verlag. 95-116.
14. Bun M.J.G., Harrison T.D. (2019), OLS and IV Estimation of Regression Models Including Endogenous Interaction Terms. *Econometric Reviews*. 38(7):814-827.
15. Cornelissen T., Jirjahn U., Tsertsvadze G. (2008), Parental Background and Earnings: German Evidence on Direct and Indirect Relationships. *Jahrbuecher fuer Nationaloekonomie und Statistik*. 228(5-6):554-572.
16. Deschênes O. (2007), Estimating the Effects of Family Background on the Return to Schooling. *Journal of Business & Economic Statistics*. 25(3):265-277.
17. Feigenbaum J.J., Tan H.R. (2019), The Return to Education in the Mid-20th Century: Evidence from Twins. NBER Working Paper No 26407.
18. Psacharopoulos G., Patrinos H.A. (2018), Returns to Investment in Education: A Decennial Review of the Global Literature. *Education Economics*. 26(5):445-458.
19. Wooldridge J.M. (2015), Control Function Methods in Applied Econometrics. *Journal of Human Resources*. 50(2):420-445.

(责任编辑:李玉柱)