

doi:10.3969/j.issn.1672-626x.2019.02.008

自评健康是否对工资收入存在溢价效应

——基于倾向得分匹配与分位数分解

俞 彤

(皖南医学院 人文与管理学院,安徽 芜湖 241002)

摘要:基于中国综合社会调查(CGSS)2012年和2015年构成混合截面数据,运用OLS与分位数回归估计工作单位性质、性别因不同的自评健康所产生的工资收入差异,并运用倾向得分匹配以降低健康选择的内生性问题。运用工资收入差异分解估计特征效应与系数效应。研究发现:自评健康的平均溢价效应约为18.07%。其中,非公共部门与女性就业群体的自评健康溢价效应较为明显。分位数回归表明随工资收入逐步递增,自评健康的溢价效应呈逐步递减趋势。工资收入差异分解表明,我国就业市场总体存在因自评健康所导致的工资收入歧视效应,且歧视效应是随工资收入成反向变动关系。各级政府、社会及企业应着力提升低收入就业人群的健康状况,其对缩小我国就业市场的工资收入差距,提升中低收入人群的工资收入水平有着重要的政策意义。

关键词:自评健康;分位数回归;劳动力的市场歧视

中图分类号:f036

文献标志码:A

文章编号:1672-626X(2019)02-0064-11

一、引言和文献综述

近10年来我国劳动者工资收入总体上呈现出较大幅度的提升。据国家统计局公布的数据,2017年我国城镇单位就业人员平均工资为74318元,较2007年的24721元实际增加了3倍^[1]。但另一方面,日益深化的市场改革也通过影响人力资本、工作经验、地区差异、行业垄断等各类要素价格与要素回报,进一步拉大了居民收入差距。虽然近10年内收入差距扩大的趋势得到了初步遏制^[2],但我国收入的基尼系数基本维持在0.45以上^[3],高于0.4的国际贫富差距警戒线。如何有效地缩小我国社会的贫富差距,着力提升低收入人群的工资收入对我国减贫政策的实施尤为关键。

马磊(2018)运用国务院发展研究中心的“中国民生调查”2014—2017年入户调查数据系统分析了我国就业者收入差距的影响因素,认为人力资本、性别歧视、户籍、教育水平、地区差异、就业稳定性及所有制等诸多非市场化因素对我国就业者的工资收入差距的形成均产生重要影响^[4]。候猛(2016)基于2012年“中国劳动力动态调查”数据,运用RIF回归分解模型分析了性别工资差异与性别歧视程度,研究认为提高工资水平,特征效应对性别收入的解释力逐步提升,系数效应的解释力逐渐下降。但研究也发现,在低收入就业群体中性别歧视对工资收入差距的影响明显高于高收入群体^[5]。户籍收入歧视对我国就业者工资收入差距的影响不容忽视。章莉与蔡文鑫(2017)基于2007年中国家庭住户收入调查数据(China Household Income Proj-

收稿日期:2018-12-07

基金项目:安徽省软科学项目(1502052069);安徽省高校人文社会科学重点项目(SK2014A417)

作者简介:俞彤(1972-),男,安徽芜湖人,皖南医学院副教授,研究方向为社会保障政策。

ects, CHIPS),运用RIF回归分解方法实证分析了我国就业市场的户籍歧视水平,结果显示我国的户籍歧视可解释城乡居民收入差距的1/3左右,且歧视效应随收入的递减而逐步提升。研究认为户籍制度因对居民进入高收入行业产生壁垒从而形成明显的劳动力市场分割是造成收入差距的重要因素^[6]。教育人力资本一直以来就是收入分配理论的研究热点。吴彬彬与李实(2018)认为,2002—2013年间,我国城乡与地区间的收入差距正逐步缩小。但城乡与地区之间,教育差异对收入差异影响正逐步扩大^[7]。齐亚强与梁童心(2016)基于2010年中国家庭追踪调查数据(China Family Panel Studies, CFPS)与相应的地区与行业数据,运用拟合交叉分类多层模型分析地区与行业的收入差距,研究显示地区与行业差异是造成我国居民收入差距的重要来源,两者可解释个人收入差距的1/5,且行业差异较地区差异对居民收入的影响更为重要^[8]。

目前国内外多数研究认同健康对收入具有较为明确的正向贡献率。自Schultz(1961)开创了人力资本理论,经Mushkin(1962)将健康拓展至人力资本中,健康与教育已成为人力资本理论的两大支柱。Grossman(1972)提出健康需求函数模型,认为就业者通过接受医疗服务对健康进行投资,从而能有效地增加健康资本存量^[9]。目前健康对就业与收入影响的研究在健康经济学与劳动经济学等领域渐成热点。Gong(2012)引入Arrow-Romer生产函数和Grossman健康需求函数形成扩展Ramsey模型,分析了健康人力资本对物质资本积累以及长期经济增长的影响。研究发现国民健康水平通过促进劳动生产率水平,进而影响物质资本积累,从而对经济增长产生正面影响^[10]。王鹏与刘国恩(2010)也认为健康人力资本是影响我国居民工资收入的重要因素^[11]。俞福丽与蒋乃华(2015)利用中国健康与营养调查数据(China Health and Nutrition Survey, CHNS)从微观层面发现,相比于高机械化程度地区,低机械化程度地区中健康对种植业收入影响较为显著^[12]。王秀芝与易婷(2017)基于中国家庭动态跟踪调查(CFPS),运用扩展的明瑟收入方程研究健康收入效应,认为健康人力资本可显著影响居民收入,且存在较为明显的地区差异与城乡差异^[13]。邓力源等(2018)通过Heckman-twosteps两阶段模型研究也佐证上述观点,即良好的健康状况对农村居民参与非农就业与获取非农就业收入具有正向效应,该影响对男性居民更为显著^[14]。

国内外文献对健康有多种测量维度,如身高、BMI指数以及广泛运用的自评健康等指标。尽管国别之间有法律制度与文化传统的差异,但不论选择何种健康维度,就业市场上因健康所导致的收入差距普遍存在。BenHalima等(2014)研究表明,以法国为例,高自评健康的就业者其小时工资率高于低自评健康的就业者约14.2%,而英国则为13.2%^[15]。较早的研究中,Thomas和Strauss(1997)以巴西就业市场为例,研究表明身高及体重指数(BMI)对受教育程度低的就业者工资收入影响尤其显著^[16]。张玉华等(2015)的研究以身高为健康人力资本表征变量,认为在我国就业市场中身高对个人收入有显著正向贡献,而这种影响长期稳定并略有增加^[17]。Kim(2018)运用韩国国家健康筛查数据库(NHSD)和社区健康调查数据库(CHS)分析了2009—2014年体重指数(BMI)与收入之间的关系,认为BMI指数可较好表征地区总体健康水平,研究发现超重加剧了收入不平等,在女性中尤为明显^[18]。

综上,不同的健康水平导致不同就业群体的工资收入产生差距已有较为明确的研究佐证。但健康人力资本对就业及工资收入影响研究尚不清晰,目前基于健康视角探讨就业市场中工资收入的溢价效应研究较少,还存在不同的健康水平对不同收入群体中的歧视程度是否存在显著差异等问题。基于此,本文在对不同健康水平的就业人群进行工资收入差距分解时,考虑如何消除因健康选择所导致的内生性问题,系统地研究健康对不同就业人群工资收入的影响机制,这将有助于提升就业者健康人力资本存量,提高就业者的劳动生产率水平,缩小就业者之间工资收入差距,对实现我国经济发展从“人口红利”向“健康红利”转化具有重要的现实意义^[19]。

本文余下结构安排如下:第二部分为数据、变量与描述性统计;第三部分为实证研究与结果,包括OLS、分位数回归、倾向匹配得分等模型;第四部分为工资差异分解结果,包含Oaxaca分解、Neumark分解和分位数分解等模型;第五部分为结论与政策建议。

二、数据、变量与描述性统计

(一)数据来源

本文数据来源于中国人民大学社会学系与香港科技大学调查研究中心合作开展的中国综合社会调查数据(CGSS)。该调查始于2003年,目前最新调查为2015年。数据覆盖我国28个省、市与地区,具有较强的代表性。本文主要研究对象为健康对劳动者工资性收入的影响。鉴于此,本文遴选从事非农工作的就业人口为样本。其中,男性年龄为18~60岁,女性年龄为18~55岁。基于单年度样本容量限制,故合并CGSS2012、CGSS2015年数据,构建混合截面数据。经过整理,最终获得4732个有效样本。

(二)主要变量定义与描述

被解释变量为对数工资性收入。其在CGSS问卷中被定义为全年非农就业的职业收入,不计职业外收入。因CGSS中工资性收入存在较多的0,运用Stata15.0系统数据清理时,采取消除离群值影响,对主要变量进行1%与99%的Winsorize缩尾处理,比例设定1%。此外,工资性收入为0,可能样本人群主要因为从事农业,未能产生工资性收入。

核心解释变量为自评健康(SRH,Self-Rated Health)。该变量是反映人群客观健康状况和主观心理期望的综合指标,其与个体健康水平、死亡率存在高度相关性^[20]。相比传统健康指标,能够更有效地反映人群的健康状况。在CGSS问卷中,受访者的“自评健康”被分成5个档次,分别为“很不健康”、“比较不健康”、“一般”、“比较健康”、“很健康”。健康状况从低到高赋值为1~5。构建虚拟变量,将“比较健康、很健康”合并为高自评健康,设定为1;“很不健康、比较不健康、一般”合并为低自评健康,设定为0。

主要人口学变量:(1)户籍。始于20世纪50年代的户籍制度已成为我国社会资源分配、社会流动与代际传承的主要影响因素。户籍制度通过对职业选择影响社会流动,进而对工资收入的获得产生重要作用^[21]。CGSS问卷根据我国户籍制度的相关规定,将居民户籍划分为农业户口、非农业户口、蓝印户口、居民户口(以前是农业户口)、居民户口(以前是非农业户口)、军籍、无户口等7种类型。本文根据李代(2017)的划分将“蓝印户口”、“以前是非农户口的居民户口”及“军籍”统一编为非农户口^[22]。(2)婚姻状态。婚姻状态在CGSS问卷中存在7种状态。根据李代(2017)的划分,本文将初婚有配偶、再婚有配偶以及分居未离婚定义为已婚。未婚、同居、离婚以及丧偶定义为未婚。已婚设定为1,未婚设定为0。(3)性别。女性设定为1,男性设定为0。(4)年龄。(5)受教育年限。

劳动能力变量:(1)外语能力。目前有文献认为外语能力是重要的人力资本,具有提高劳动生产率以及向雇主表明个人能力的信号传递功能^[23]。外语能力与收入之间具有正相关性^[24]。CGSS问卷中表征外语能力的指标为外语的听、说能力。本文认为,良好的外语会话能力可综合反映样本的外语能力。CGSS问卷中外语会话能力分为“完全不能说”、“比较差”、“一般”、“比较好”以及“很好”5个层次。本文将“比较好”和“很好”设定熟练=1,其余层次设定为不熟练=0。(2)工作经验。代表性研究基本证实工作经验与职业收入存在较为明确的依存关系^[25]。即工作经验越长,则工资收入越高。本文使用CGSS问卷中“从您第一份非农工作到目前的工作,您一共工作了多少年”来定义样本人群的工作经验。

社会资本变量:(1)家庭背景。吴晓刚(2007)认为家庭背景对当前我国职业地位的获得与社会流动,进而对城乡间制度分割具有最重要的作用。本文借助父母的社会经济地位来表征样本人群的家庭背景,但CGSS2012、CGSS2015问卷只询问了受访者父亲与母亲的职业,所以本文采用Blau和Duncan(1967)的地位获得模型^[26],该模型是通过受访者的受教育年限、父亲的职业、父亲的受教育年限预测受访者当前的职业地位。将样本父母职业的ISCO-88代码转换为ISEI(International Socioeconomic Index,国际社会经济指数)。ISEI指标为取值16~90的连续型变量^[27]。(2)个人的社会等级。该变量为定序变量。在CGSS问卷中属于阶

层认同指标类。其中,社会等级变量分为“您目前的社会等级”、“10年前您的社会等级”、“10年后您的社会等级”以及“14岁时,你的家庭社会等级”4个分类。本文选取“您目前的社会等级”来表征个人的社会等级。社会等级从低到高赋值为1~10。(3)工作单位性质。针对于CGSS问卷,本文将党政机关(包括党委、政府、人大、政协、公检法、武装部、部队)、事业单位以及国有独资企业划归为公共部门。除公共部门以外的工作单位划归为非公共部门(包括集体企业、私营及个体企业、股份制企业以及三资企业等)。其中,公共部门设定为1;非公共部门设定为0。变量定义与描述性统计详见表1。

表1 全部变量定义与描述性统计

变量	定义	全样本	
		Mean	Std. Dev.
lnincome	全年职业收入对数	10.44823	1.38086
health	是否健康(1=高自评健康;0=低自评健康)	0.75	0.433059
yr2015	调查年份是否是2015年(2015年=1;2012年=0)	0.401793	0.490310
sex	是否为女性(1=女性;0=男性)	0.400398	0.490028
marri	是否已婚(1=已婚;0=未婚)	0.83506	0.371164
age	年龄	39.21673	9.702901
Hukou	是否为非农户口(1=非农户口;0=农业户口)	0.481076	0.499692
degree	1为社会最底层;10为社会最顶层	3.998207	1.767914
English	说英语能力是否熟练(1=熟练;0=不熟练)	0.019323	0.13767
healthcare	是否有社会医疗保险(1=有;0=无)	0.893825	0.308093
oldcare	是否有社会养老保险(1=有;0=无)	0.678685	0.467028
Work_year	工作经验	17.79422	11.10764
Father_eco	父亲社会经济地位指数	32.05916	17.71939
Mother_eco	母亲社会经济地位指数	26.92888	13.65563
Edu_year	受教育年限	11.25219	3.636248
middle	中部地区	0.265139	0.441451
west	西部地区	0.232669	0.422575
public	就业是否是公共部门(1=是;0=否)	0.198406	0.39884
样本总数		4732	

(三)工资差异的典型化事实

全样本统计中,高自评健康人群年均工资显著高于低自评健康人群11.12%,年工资收入存在约3672.5元的工资的健康溢价效应。公共部门内高自评健康人群与低自评健康人群年均工资之间无显著差异,而在非公共部门就业的高自评健康人群年均工资显著高于低自评健康人群14.83%。男性高自评健康人群年均工资显著高于低自评健康人群约11.2%,而女性高自评健康人群则显著高于低自评健康人群约11.13%。因此无论以工作单位性质或以性别为分组,自评健康对工资水平均存在明确的溢价效应(具体数据见表2)。

表2 基于不同分组下的工资差异典型化事实

类别	样本数	占比	年均工资(元)		T-value
			高自评健康	低自评健康	
全样本	4732	100	35590.7	31918.2	0.000
公共部门	964	20.3	37723.4	37884.1	0.466
非公共部门	3768	79.7	34997.9	30479.1	0.000
男性	1905	40.3	32072.0	28841.3	0.007
女性	2827	59.7	37915.7	34116.6	0.000

三、实证研究及结果

(一)工资收入的OLS回归估计

Mincer收入模型自1974年提出已经成为研究人力资本的重要方法^[28]。经典Mincer收入模型是基于教育与工作经验等人力资本要素研究对工资收入的影响。本文在经典Mincer收入模型基础上引入自评健康变量构造扩展Mincer半对数收入模型。其中,被解释变量为对数工资收入,以是否为健康为核心解释变量,控制相关因素影响建立如下多元线性计量模型:

$$\ln income_i = \beta_0 + \alpha \cdot health_i + \sum_{i=1}^n \beta_i \cdot X_i \quad (1)$$

其中,个体特征变量 X_i 包括性别、年龄、婚姻状况、户口等人口学变量;外语能力、受教育年限等劳动能力变量;工作单位性质、个人的社会等级、家庭背景等社会资本变量;社会医疗保障、社会养老保障等社会福利变量;年份、地区等控制变量。表3为工资性收入的OLS回归结果,包括分工作单位性质、分性别测算的工资收入OLS回归结果。

表3 工资收入的OLS回归结果

变量	全样本	公共部门	非公共部门	男性	女性
health	0.0806*** (3.60)	0.0696* (1.75)	0.0824*** (3.18)	0.0661* (1.90)	0.0862*** (2.94)
sex	-0.261*** (-13.59)	-0.208*** (-5.92)	-0.271*** (-12.11)	-	-
public	-0.0629*** (2.63)	-	-	-0.0796** (-2.12)	-0.0524* (-1.68)
marri	0.142*** (4.87)	0.188*** (3.31)	0.131*** (3.95)	0.105*** (2.30)	0.173*** (4.53)
age	-0.0114*** (-6.33)	0.000282 (0.07)	-0.0139*** (-6.89)	-0.00989*** (-3.25)	-0.0124*** (-5.54)
Hukou	-0.00565 (-0.24)	-0.0887* (-1.74)	-0.00120 (-0.05)	-0.0242 (-0.66)	0.00425 (0.14)
degree	0.0400*** (6.82)	0.0327*** (3.19)	0.0413*** (6.02)	0.0347*** (3.86)	0.0439*** (5.67)
English	0.189*** (3.10)	-0.0339 (-0.41)	0.347*** (4.37)	0.125 (1.32)	0.252*** (3.23)
healthcare	0.0550 (1.54)	0.0309 (0.28)	0.0558 (1.48)	0.0646 (1.15)	0.0418 (0.90)
oldcare	0.0468* (1.93)	-0.00961 (-0.20)	0.0673** (2.48)	0.0587 (1.54)	0.0390 (1.24)
Work_year	0.0337*** (9.52)	0.0291*** (3.71)	0.0321*** (8.07)	0.0276*** (4.80)	0.0355*** (7.75)
Work_year2	-0.000610*** (-7.33)	-0.000485*** (-3.02)	-0.000594*** (-6.27)	-0.000444*** (-3.12)	-0.000651*** (-6.19)
Father_eco	0.000592 (0.95)	0.000498 (0.53)	0.000460 (0.58)	0.00105 (1.13)	0.000201 (0.24)
Mother_eco	0.00242*** (2.96)	0.00161 (1.44)	0.00303*** (2.74)	0.00311** (2.46)	0.00181* (1.67)
Edu_year	0.0710*** (19.81)	0.0935*** (12.50)	0.0640*** (15.44)	0.0745*** (12.68)	0.0692*** (15.04)
middle	YES	YES	YES	YES	YES
west	YES	YES	YES	YES	YES
yr2015	YES	YES	YES	YES	YES
N	4732	964	3768	1905	2827
R ²	0.3111	0.4056	0.2957	0.3356	0.2758

注:括号内数字为标准误。*、**、***表示在10%、5%和1%水平上显著,下同

本文使用稳健标准误控制异方差。在全样本的OLS回归中,控制地区与年份变量,高自评健康人群较低自评健康人群的工资收入平均溢价效应为8.06%。其中,公共部门的健康溢价效应为6.96%,显著低于非公共部门的8.24%。健康的溢价效应在不同性别群体中也呈现显著差异性,相比男性的6.61%,健康对女性就业群体中具有更大的溢价效应,达到了8.62%,且均在1%水平下显著。婚姻状况、年龄、社会阶层、外语能力、工作经验与工作经验平方、母亲的社会经济地位、受教育年限等变量对工资收入的影响符合经验预期。以上分析表明,我国就业市场工资收入存在显著的健康溢价效应。健康作为人力资本有效提升劳动生产

率,进而可以获得较高的工资收入。有良好健康状况的就业者在就业市场中更易获得有较高工资收入的工作岗位。另外,非公共部门和女性就业群体可能更多地面向以劳动力密集型为特征的二级劳动力市场,故而导致健康的溢价效应在非公共部门和女性群体内存在较大的差异度。

需要说明的是,本模型估计教育的溢价效应为7.01%,低于健康的溢价效应的8.06%。近年来已有文献表明教育人力资本有呈边际报酬递减趋势^[29],而健康人力资本比教育人力资本呈现出更高的溢价能力。

(二)高自评健康与低自评健康人群的工资收入分位数回归估计

OLS 回归主要估计健康溢价效应的均值影响,而对整个工资收入分布区间的异质性较少涉及^[30]。分位数回归提供了条件分布的全面信息,在不同收入分位点上,健康的溢价效应有可能蕴含相应的政策意义。为进一步研究不同自评健康人群在不同分布区间的工资收入决定方程与回报差异,本文基于分位数回归将模型设定为:

$$Q_{\theta}(\ln income_i | X_i) = \beta_{\theta 0} + \alpha_{\theta} \cdot health + \sum_{i=1}^n \beta_{\theta} \cdot X_i \tag{2}$$

其中 $Q_{\theta}(\ln income_i | X_i)$ 表示在个体特征变量为 X_i 的条件下,高自评健康与低自评健康人群的工资收入对数处于不同分位数上相应的值; $\beta_{\theta 0}$ 为误差项; β_{θ} 为 θ 分位数的回归系数。本文选择0.25、0.50、0.75、0.90共4个代表性分位点,并通过(3)式估计回归系数。具体结果见表4。

$$\min \left\{ \sum_{i: \ln income_i \geq X_i \beta(\theta)} \theta |\ln income_i - X_i \beta(\theta)| + \sum_{i: \ln income_i < X_i \beta(\theta)} (1 - \theta) |\ln income_i - X_i \beta(\theta)| \right\} \tag{3}$$

表4 高自评健康与低自评健康人群的工资收入分位数回归结果

变量	分位数回归			
	.25	.50	.75	.90
health	0.0828***(2.65)	0.0655** (2.29)	0.0627* (1.96)	0.0279 (0.76)
sex	-0.280***(-11.69)	-0.261***(-12.00)	-0.250***(-8.59)	-0.250***(-7.50)
public	-0.0414(-1.24)	-0.0867**(-2.25)	-0.132***(-3.74)	-0.184***(-4.35)
marri	0.176***(4.51)	0.104***(3.50)	0.107***(2.61)	0.133***(2.70)
age	-0.0117***(-5.25)	-0.0107*** (-4.65)	-0.0111***(-3.58)	-0.0157***(-5.81)
Hukou	0.00699(0.25)	0.00493(0.22)	0.0253(0.74)	-0.00576(-0.15)
degree	0.0333***(4.45)	0.0384***(5.58)	0.0550***(7.68)	0.0691***(6.82)
English	0.189***(3.04)	0.221***(3.04)	0.176***(2.89)	0.0882(1.16)
healthcare	0.0553(0.91)	0.00292(0.06)	0.0577(1.18)	0.152**(2.36)
oldcare	0.0835***(2.68)	0.0682***(2.98)	0.0241(0.60)	-0.0505 (-1.12)
Work_year	0.0346***(8.25)	0.0316*** (9.35)	0.0292*** (5.24)	0.0302*** (5.34)
Work_year2	-0.000634***(-6.32)	-0.000554***(-6.00)	-0.000502***(-3.82)	-0.000417***(-3.12)
Father_eco	0.000414 (0.44)	0.000834(1.18)	0.00110 (1.34)	0.000883 (0.81)
Mother_eco	0.00218** (2.05)	0.00150* (1.74)	0.00202*(1.65)	0.00249* (1.95)
Edu_year	0.0843*** (18.08)	0.0749*** (14.55)	0.0701*** (14.85)	0.0609*** (12.50)
middle	YES	YES	YES	YES
west	YES	YES	YES	YES
yr2015	YES	YES	YES	YES
R ²	0.4056	0.2957	0.3356	0.2758

从表4可以看出,随着分位点上升,健康的溢价效应呈逐渐下降的趋势。在0.25分位点上,健康的溢价效应为8.28%,且具有1%显著性水平。至0.9分位点,健康的溢价效应已完全不显著。总体而言,健康对工

资收入具有明显的正向贡献率。但其影响主要作用于低收入就业人群,而高收入就业人群则影响较小。可能的原因在于,低收入就业人群主要处于工资收入水平较低的二级劳动力市场,以劳动力密集型行业为主,良好的健康状况意味着较高的劳动生产率,更易获得较高的工资收入。着力提升低收入就业群体的健康水平对提高收入水平,降低贫富差距具有明确的政策意义。而高收入就业人群多处于一级劳动力市场,健康可能并非是提升工资收入最为主要的驱动因素。

(三)倾向匹配得分估计

本文主要研究的是健康的溢价效应。若单纯使用高自评健康与低自评健康人群收入差异作为估计结果,则会因就业者个体初始禀赋存在差异,有可能对工资收入产生影响,从而导致估计结果并非单纯基于健康水平所带来的就业者工资收入的净效应。研究希望获得基于相同初始禀赋的个体在高自评健康与低自评健康条件下收入差异。倾向得分匹配(P propensity Score Matching, PSM)是基于反事实条件下的非参数估计,其原理是通过处理效应以消除样本自选择偏差与内生性问题。

在特定时期内,假定内就业者只选择高自评健康与低自评健康中某一特定情况。设 $health_i$ 表示第 i 个就业者是否为高自评健康者。当就业者为高自评健康时, $health_i = 1$, 设为处理组;反之, $health_i = 0$, 设为控制组。在控制组中匹配与处理组初始禀赋相同的个体,以高自评健康者与匹配者工资收入差别作为健康的溢价效应,即平均处理效应 ATT(Average Treatment Effect on The Treated)。

本文使用 logit 概率回归模型估计倾向得分。根据倾向得分结果,对处理组和控制组的个体进行包括近邻匹配、半径匹配、核匹配以及局部线性匹配的 4 种匹配。匹配前,本文只报告基于核密度匹配的检验结果。匹配后,运用这 4 种匹配方法,以相互印证匹配结果的稳健性(见表 5)。

表 5 变量的匹配平衡检验

匹配变量	处理组均值	控制组均值	标准化偏差(%)	标准化偏差减少(%)
sex	0.3980	0.4051	-1.5	61.4
public	0.2050	0.2065	-0.4	77.2
marri	0.8236	0.8278	-1.2	92.2
age	38.252	38.593	-3.6	91.8
Hukou	0.4945	0.4813	2.6	63.2
degree	4.0387	4.0424	-0.2	98.8
English	0.0198	0.0135	5.1	50.3
healthcare	0.8951	0.8884	2.2	-697.4
oldcare	0.6797	0.6848	-1.1	24.9
Work_year	17.118	17.419	-2.7	91.5
Work_year2	414.3	422.62	-1.8	93.9
Father_eco	31.736	31.917	-1.0	73.4
Mother_eco	26.488	26.712	-1.6	72.1
Edu_year	11.362	11.387	-0.7	96.2

从表 5 可知,匹配变量的标准化偏差绝对值基本在 5% 以内,可认定匹配变量合理。由于假设倾向得分为真实值,然后推导标准误,且标准误应满足同方差假设。考虑运用自助法估计匹配后的平均处理效应 ATT。

由表 6 可见,匹配前处理组平均处理效应(ATT)高自评健康人群的工资收入较低自评健康人群工资收入高出 17.14%。表 7 表示,匹配后处理组平均处理效应(ATT)基于自助法计算标准误条件下核匹配为 6.6%,最近邻匹配为 6.53%,样条匹配为 6.04%,卡尺内匹配为 6.36%,且具有 1% 或 5% 显著性水平。对照表 4 分位数回归结果,也基本吻合了 0.5 分位点与 0.75 分位点下健康对工资收入的回归系数。这表明当解决了

样本“自选择”后,我国就业市场中健康的溢价效应依然存在。

表6 倾向得分匹配前处理组的平均处理效应

处理效应	ATT	标准误	T值
匹配前	0.1714	.0256	6.69

表7 倾向得分匹配后处理组的平均处理效应

匹配后	ATT	自助标准误	Z值	P值
核匹配	0.0660	0.0240	2.75	0.006
最近邻匹配(K=4)	0.0653	0.0331	1.97	0.049
样条匹配	0.0604	0.0256	2.36	0.018
卡尺内匹配(Radius=0.01)	0.0636	0.0262	2.43	0.015

四、工资差异分解结果

工资差异分解是将两个不同群体的工资收入差距分解为不同因素影响之和。自Oaxaca(1973)^[31]提出Oaxaca分解,最初是研究工资性别差距的主要手段,后拓展为研究两个不同就业群体之间平均工资差距。其后为了修复Oaxaca分解在部门选择与指数基准上存在的缺陷,先后提出了Brown分解(1980)^[32]、Cotton分解(1988)^[33]与Neumark分解(1988)^[34]等改进方法。各类分解方法主要思想均是将两个不同就业人群之工资差异分解为各种因素的影响之和。其主要有两个研究目的:一是比较不同群体间收入差异及成因并检验特定因素对总体收入差异的贡献^[35];二是可定量刻画就业市场中的歧视程度^[36]。据此工资差异分解成特征效应与禀赋效应。其中,特征效应是源于就业者自身禀赋差异所导致的收入差距。禀赋效应则是就业者虽有相同的自身禀赋,但由于非市场因素而导致不同的收入差距,反映了市场对某一特定人群的歧视效应。

表6结果显示,匹配前平均处理效应为0.1714,表明高自评健康人群的工资收入比低自评健康人群高18.7%($e^{0.1714} - 1$)。表8为匹配前平均处理效应分解结果。其中,Oaxaca分解中特征效应占比约为37.07%,禀赋效应占比约为62.93%。Neumark分解中特征效应占比约为37.39%,禀赋效应占比约为62.61%,两类分解结果基本接近。其中,禀赋效应表征我国的就业市场存在着明确的健康收入歧视效应。

表8 Oaxaca分解与Neumark分解结果

分解方法	特征差异		系数差异		总差异 差异值对数
	差异值对数	占比	差异值对数	占比	
Oaxaca分解	0.0635	37.069%	0.1078	62.931%	0.1714
Neumark分解	0.0641	37.388%	0.1073	62.612%	0.1714

从表9可见,处于工资收入分布的低端分位数上,高自评健康人群与低自评健康人群工资收入差异较大。随分位点提高,两类人群的工资收入差异逐渐缩小。在0.25、0.50、0.75、0.90分位点上,对数工资收入差异总效应分别为0.2306、0.1619、0.1041、0.0764。高自评健康人群比低自评健康人群工资收入分别高出25.94%、17.57%、10.97%、7.75%。由此可见,低收入就业人群中健康作为必要的人力资本对工资收入影响尤其重要,良好的健康状况是拉开收入差距的重要因素。而在高收入就业人群中,健康状况可能并非是影响工资收入差异的关键因素。可能的原因是:在工资收入分布的低端分位数下,就业人群多为体力劳动者或非稳定受雇者,缺乏稳定而必要的社会保障,健康管理意识相对不足。因此,健康所导致的市场歧视度相对严重。而工资收入处于较高分位数时,就业人群多以从事相对稳定工作为主,社会保障覆盖相对完善,健康

管理相对到位。因此健康水平对高收入就业人群的市场歧视度相对较弱。

此外,随着工资收入的提高,特征效应占比逐渐变大,对工资收入差异的解释力逐渐增强。禀赋效应占比逐渐变小,表明市场歧视度逐渐变弱。在0.25分位点上,特征效应占比为54.16%,在0.50分位点上,特征效应占比为65.91%;在0.75分位点上,特征效应占比为89.82%;在0.90分位点上,特征效应占比为94.90%。这表明随着工资收入的提高,高自评健康人群比低自评健康人群拥有更高的人力资本优势,人力资本回报率差异更加明显,但在较低端工资分位数上禀赋效应占比较高,意味着在低收入就业人群中高自评健康与低自评健康的两类群体中存在较为严重的门槛障碍、就业歧视、劳动力二元化市场分割等一系列非市场因素。

表9 高自评健康人群与低自评健康人群工资收入差异分位数分解结果

收入效应	分位点			
	.25	.50	.75	.90
总效应				
差异值对数	0.2306(0.0162)	0.1619(0.0142)	0.1041(0.0146)	0.0764(0.0171)
特征效应				
差异值对数	0.1249(0.0288)	0.1067(0.0247)	0.0935(0.0286)	0.0725(0.0307)
占比(%)	54.16	65.91	89.82	94.90
系数效应				
差异值对数	0.1057(0.0309)	0.0552(0.0245)	0.0107(0.0262)	0.0039(0.0323)
占比(%)	45.84	34.01	10.28	5.10

五、结论与政策建议

本文通过运用CGSS2012、CGSS2015混合截面数据,运用OLS与分位数回归模型估计公共部门与非公共部门、不同性别中自评健康所产生的工资收入差异,通过倾向匹配得分消除样本选择性偏差,进一步估计健康的溢价效应。其次,通过Oaxaca与Neumark分解计算工资差异中的特征效应与禀赋效应。结果表明,全样本中健康的溢价效应均值为8.06%。其中,非公共部门健康的溢价效应明显高于公共部门,女性群体健康溢价效应明显高于男性群体。由此可知,相对于公共部门与男性就业群体,非公共部门与女性就业群体中健康的溢价效应更为明显。分位数回归结果表明,随着工资收入的提高,健康的溢价效应逐渐呈下降趋势。相对于高收入就业群体,自评健康对中低收入就业群体影响更为明确。即使运用倾向得分匹配结果也表明,匹配前劳动力市场健康工资收入的平均处理效应约17.14%。当消除就业者初始禀赋差异,就业市场健康溢价的平均处理效应也可达6%以上。Oaxaca分解与Neumark分解结果表明,高自评健康人群与低自评健康人群之间工资收入差异约37%可由特征效应解释,工资收入差异的约63%可由禀赋效应所解释,反映我国就业市场存在着明显的市场分割与健康歧视现象。

健康是影响劳动者工资收入差异的主要因素,特别是在一级劳动力市场上其作用更甚于教育的影响。各级政府应采取多种政策与措施,如加强健康教育、加大重大疾病的保障力度,加强职业病防治力度等。着力改善中低收入就业人群的健康状况。该举措对缩小我国就业市场的工资收入差距,提升中低收入人群工资收入水平有着一定的政策参考。对于处于二级劳动力市场的就业者,可通过积极改善自身健康状况以提升工资收入水平。企业应充分认识健康人力资本对企业发展的作用,让就业者明确良好的健康状况对提升工资水平的意义,引导并鼓励就业者有效地改善自身健康状况,为提升劳动生产率提供坚实的保障。

参考文献:

- [1] 国家统计局年度数据[EB/OL].<http://data.stats.gov.cn/easyquery.htm?cn=C01&zb=A0A07&sj=2017>.
- [2] 李实.中国收入分配制度改革四十年[J].China Economist, 2018, 13(4):2-33.
- [3] 李实,高霞.居民收入差距的测量及其合理判断[J].统计与决策,2015,(10):11-16.
- [4] 马磊.我国劳动收入差异影响因素研究——基于“中国民生调查”2014—2017年数据的分析[J].重庆理工大学学报(社会科学),2018,(10):1-13.
- [5] 侯猛.性别工资差异与工资歧视——基于RIF回归的分解方法[J].南方人口,2016,(1):18-25.
- [6] 章莉,蔡文鑫.中国劳动力市场收入户籍歧视的无条件分位数分解[J].复旦学报(自然科学版),2017,(1):12-18.
- [7] 吴彬彬,李实.中国地区之间收入差距变化:2002—2013年[J].经济与管理研究,2018,(10):31-44.
- [8] 齐亚强,梁童心.地区差异还是行业差异?——双重劳动力市场分割与收入不平等[J].社会学研究,2016,(1):168-190.
- [9] Grossman, M. On the Concept of Health Capital and the Demand for Health [J]. The Journal of Political Economy, 1972, 80(2): 223-255.
- [10] Gong, L.T., & H.Y. Li., & D.H. Wang. Health Investment, Physical Capital Accumulation, and Economic Growth[J].China Economic Review, 2012, 23(4): 1104-1119.
- [11] 王鹏,刘国恩.健康人力资本与性别工资差异[J].南方经济,2010,(9):73-84.
- [12] 俞福丽,蒋乃华.健康对农民种植业收入的影响研究——基于中国健康与营养调查数据的实证研究[J].农业经济问题,2015,(4):66-71.
- [13] 王秀芝,易婷.健康人力资本的收入效应[J].首都经济贸易大学学报,2017,(4):20-26.
- [14] 邓力源,唐代盛,余驰晨.我国农村居民健康人力资本对其非农就业收入影响的实证研究[J].人口学刊,2018,(1):102-112.
- [15] Ben Halima, M.A., & E. Rococo. Wage Differences According to Health Status in France [J]. Social Science & Medicine, 2014, 120(11): 260-268.
- [16] Thomas, D., & J. Strauss. Health and Wages: Evidence on Men and Women in Urban Brazil[J]. Journal of Econometrics, 1997, 77(1): 159-185.
- [17] 张玉华,赵媛媛.健康对个人收入和城乡收入差距的影响[J].财经问题研究,2015,(8):11-16.
- [18] Kim, I., & J. Bahk. Prevalence of Overweight and Income Gaps in 245 Districts of Korea: Comparison Using the National Health Screening Database and the Community Health Survey, 2009-2014[J].Journal of Korean Medical Science,2018,33(1): 553-560.
- [19] 高凯,汪泓,刘婷婷.劳动人口健康水平影响因素及健康状况演变趋势[J].社会科学研究,2018,(1):38-47.
- [20] Idler, E.L., & Y. Benyamini. Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies[J]. Journal of Health and Social Behavior,1997,38(1):21-37.
- [21] 吴晓刚.中国的户籍制度与代际职业流动[J].社会学研究,2007,(6):38-65.
- [22] 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.
- [23] 刘泉.外语能力与收入——来自中国城市劳动力市场的证据[J].南开经济研究,2014,(3):137-153.
- [24] Gao,W.,& R.Smyth. Economic Returns to Speaking 'Standard Mandarin' among Migrants in China's Urban Labor Market[J]. Economics of Education Review, 2011, 30(2):342-352.
- [25] 董良.教育、工作经验与家庭背景对居民收入的影响——对明瑟方程和“布劳—邓肯”模型的综合[J].中国社会科学院研究生院学报,2016,(4):103-109.
- [26] Blau, P.M., & O.D.Duncan. The American Occupational Structure [M]. Hoboken, NJ, US: John Wiley & Sons Inc, 1967.
- [27] 阳义南,连玉君.中国社会代际流动性的动态解析——CGSS与CLDS混合横截面数据的经验证据[J].管理世界,2015,(4):79-91.
- [28] 杨国涛,段君,刘子詠.明瑟收入方程的若干改进和思考[J].统计研究,2014,(7):81-84.
- [29] 娄艳霞.教育对中国城镇居民的收入影响分析[J].行政事业资产与财务,2018,(7):91-92.
- [30] 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):15-25.
- [31] Oaxaca, R. Male-Female Wage Differentials in Urban Labor Markets [J]. International Economic Review, 1973, 4(3):693-709.
- [32] Brown, R., &M.Moon &S.Barbara. Incorporating Occupational Attainment in Studies of Male-Female Earnings Differentials[J]. The Journal of Human Resources, 1980, 15(1): 3-28.
- [33] Cotton, J. On the Decomposition of Wage Differentials[J]. The Review of Economics and Statistics, 1988, 70(2): 236-243.

- [34] Neumark, D. Employers Discriminatory Behavior and the Estimation of Wage Discrimination[J]. The Journal of Human Resources, 1988,23(3):279-295
- [35] 阳建辉,杨继生.交互效应,动态面板模型的Blinder-Oaxaca分解与仿真[J].系统工程理论与实践,2017,(4):855-864.
- [36] 葛玉好,赵媛媛.工资差距分解方法之述评[J].世界经济文汇,2011,(3):110-120.

(责任编辑:卢 君)