

DOI: 10.19361/j.er.2020.02.05

中国居民健康差距中的机会不平等

刘波 胡宗义 龚志民*

摘要: 缩小健康差距、提升居民健康水平是当前治国理政所关切的核心问题。在居民健康差距中,既有合理的差距,也有不合理的差距,即健康机会不平等。本文在改进已有测度方法的基础上构建测度模型,以中国家庭追踪调查数据(2010—2016年)为样本估计健康决定方程,通过夏普利值分解测度中国居民健康差距中的机会不平等。测度结果表明,健康机会不平等对健康差距的相对贡献度为18.316%~23.182%;在外生变量中,性别、3岁时所处的经济区域是形成健康机会不平等的主要因素。以实证结论为基础,本文从早期家庭背景、关注女性健康和提升健康素养等方面,提出抑制健康机会不平等、缩小健康差距的对策。

关键词: 健康不平等;机会不平等;夏普利值分解

一、引言

健康是国家发展之基,亦是民生幸福之本。党的十九大报告指出,“人民健康是民族昌盛和国家富强的重要标志”,居民健康是治国理政所关切的核心问题,也是舆论关注的焦点话题。近年来,医疗卫生领域的改革不断深化,新农合全覆盖、异地医保结算推行、格列宁等进口药品纳入医保范围、HPV疫苗推向市场等措施深受民众好评。2016年10月,中共中央、国务院颁布了《“健康中国2030”规划纲要》,提出“逐步缩小城乡、地区、人群间基本健康服务和健康水平的差异,实现全民健康覆盖,促进社会公平”。2019年7月15日发布的《国务院关于实施健康中国行动的意见》明确提出,到2030年“健康公平基本实现”。居民健康既有先天遗传因素的影响,也有后天生活、工作环境的作用,还有医疗保健的干预。居民健康原本是生物学事件,由于诸多社会因素掺杂到“生老病死”的过程中,从而把原本属于一个生物学事件的自然状态变成了一个复杂的社会事件,诸多社会经济因素左右了居民健康状态的公平性(李红文、毛新志,2015)。

在居民健康差距中,既有合理的差距,亦有不合理的差距。衰老导致的健康水平下滑是自然规律使然,不健康的生活习惯(如抽烟、酗酒)等因素造成的健康差异是个人自由选择的结果,二者所导致的健康差距符合正义法则,是合理的差距。反之,如果居民之间的健康差

*刘波(通讯作者),湖南大学金融与统计学院,邮政编码:410079,电子信箱:cq_liubo@163.com;胡宗义,湖南大学金融与统计学院,邮政编码:410079,电子信箱:zongyihu@163.com;龚志民,湘潭大学商学院,邮政编码:411105,电子信箱:gzm@xtu.edu.cn。

本文获得中国博士后科学基金第65批面上项目“收入分配中机会不平等测度方法的重构与应用研究”(项目编号:2019M652761)的资助。作者感谢匿名审稿人的建议,但文责自负。

距是由外生因素差异,如出生时的户口、地区、家庭背景等导致的,则是不合理的差距。以 Roemer 和 Silvestre(1993)、Roemer(1998)、Roemer 和 Trannoy(2016)的机会不平等理论为基础,诸多学者将上述不合理的差距定义为健康差距中的机会不平等(Rosa Dias, 2009; Fleurbaey and Schokkaert, 2009; Jusot et al., 2013; Donni et al., 2014; Fajardo-Gonzalez, 2016)。已有文献表明,我国居民的健康水平在城乡之间存在显著差异(鲁万波等, 2018),且存在“地位综合症”(Erreygers and Van Ourti, 2011; 王甫勤, 2012)与医疗服务利用不平等(解丕, 2009),健康机会不平等问题突出。

缩小健康差距是缩小不合理的差距——健康机会不平等,而不是缩小合理的差距。在医疗健康政策设计中,按照补偿原则(Compensation Principle),针对不合理的机会不平等,社会应予以补偿(马超等, 2018; Zhao et al., 2020);而对于合理的健康差距,按照回报原则(Reward Principle),社会应给予回报。因此,本文的主要边际贡献是以 Jusot 等(2013)、Juárez 和 Soloaga(2014)的方法为基础,遵循 Roemer 和 Silvestre(1993)、Roemer(1998)的理论框架,构建健康机会不平等的测度模型。在实证研究中,以本文构建的模型为基础,通过夏普利值分解方法,测度健康机会不平等对健康差距的相对贡献度,识别出导致健康机会不平等的主要因素,为推进“健康中国 2030”建言献策。

本文其余内容包括五个部分:首先,回顾国内外研究进展,提炼本文研究的理论基础,确定进一步推进健康机会不平等研究的工作方向。其次,给出机会不平等的测度方法与分解方法。第三,确定健康决定方程中的解释变量,给出变量的赋值方式,并估计健康决定方程;以健康决定方程为基础,结合夏普利值分解方法,测度健康不平等中的机会不平等。第四,调整健康决定方程的估计模型,验证测度结果的稳健性。最后,总结全文,提炼研究结论。

二、文献综述

与收入差距中的机会不平等类似,以健康差距中的机会不平等为主题的研究也包括三个方面:定义、测度与分解。本文将从国内外文献入手,分别从上述三个方面展开文献综述。

健康机会不平等是以机会不平等理论为基础,机会不平等理论最早是由 Rawls(1971)提出,他认为由个体选择差异所导致的结果不平等是可以接受的。随后, Dworkin(1981a, 1981b)、Arneson(1989)、Cohen(1989)对机会不平等理论做了进一步阐释, Roemer 和 Silvestre(1993)、Roemer(1998)则构建了“环境-努力”的二元分析框架,将机会不平等正式引入经济学,其后被广泛应用于收入与财富分配、教育资源分配与健康的研究中。然而,“健康、活力、智力和想象力都是自然的产物,尽管它们的拥有受到社会基本结构的影响,但它们并不是直接受其控制”(Rawls, 1971),因此, Rawls 并没有将健康列入受分配原则制约的清单之中,而是将其作为评价社会公正的核心变量。Daniels(1985)认为,虽然将健康作为评价社会公正的核心变量并不适当,但将有益于健康的医疗保健作为核心变量则是适当的,平等的健康权表现为平等地获得或享有医疗保健服务(Ruger, 2004),但公平并不意味着每个人都应该拥有相同的健康状况,或者不管需要与否,消耗相同数量的卫生服务资源(Whitehead et al., 1991)。Culyer 和 Wagstaff(1993)进一步将医疗保健公平界定为利用公平、按需分配公平、获得公平和健康公平,而健康公平是评价医疗保健分配公平的终极标准。

Braveman 和 Gruskin(2003)明确给出了健康公平与健康不平等的定义,“健康(及其关

键的社会决定因素)不存在与社会优势/劣势(财富、权力或声望)有系统联系的差异,健康不平等系统地使已经处于社会不利地位的人口(如贫困人口、女性、种族、少数族裔、宗教群体)在健康方面处于更不利的地位”。由此可见,在健康差距中,由财富、权力、身份导致的健康差距是健康机会不平等关注的焦点,也是形成健康机会不平等的现实基础。在量化研究中,已有文献仍然延续 Roemer 和 Silvestre(1993)、Roemer(1998)提出的机会不平等分析框架,健康差距同样可以分解为合理的不平等和不合理的不平等,其中家庭原生环境、性别等外生条件差异导致的健康差距被定义为不合理的机会不平等,而自我能够左右的生活方式差异导致的健康差距则被视为合理的不平等(Rosa Dias, 2009; Fleurbaey and Schokkaert, 2009; Jusot et al., 2013; Donni et al., 2014; Fajardo-Gonzalez, 2016),该定义与收入分配中机会不平等的定义如出一辙(Almås et al., 2011)。

在测度方法上,与收入分配中的机会不平等测度方法类似,亦可以分为非参数法和参数法(Ferreira and Gignoux, 2011),非参数法与参数法最大的差异在于构造反事实分布的方法不同。非参数方法包括事前法(Ex-ante)和事后法(Ex-post)(Donni et al., 2014),事前法和事后法均是以不平等指标分解理论为基础。在事前法中,首先将个体的属性值进行分组,再对各组进行平滑处理,以平滑处理后的数据为基础估计不平等程度并按组分解(Bourguignon et al., 2007; Pignataro, 2012)。显然,组间不平等是条件变量差异导致的不平等,即机会不平等。事后法的分组方法略有不同,且需以 Foster 和 Shneyerov(2000)提出的路径无关性为理论前提。首先将样本按照属性值分类(Type),然后依次将各类样本按照事先给定的分位点分成若干级别(Tranche),再将各组中的同一级样本别归为同一组,最后对各组样本进行平滑处理。以平滑处理后的数据为基础估计不平等程度并按组分解,组内不平等即为机会不平等。

参数法以 Ferreira 和 Gignoux(2011)、Björklund 等(2012)的文献为代表,Davillas 和 Jones(2018)、Carrieri 和 Jones(2018)采用了参数法测度健康差距中的机会不平等。在参数法中,首先构建健康决定方程 $H=f(C)$,方程以条件变量 $C^{\text{①}}$ 为解释变量,健康指标 H 为被解释变量。其次,估计健康决定方程,并以方程为基础得到健康水平的预测值 $\tilde{H}$ 。最后,估计 $\tilde{H}$ 的不平等程度得到机会不平等的绝对值 $I(\tilde{H})$ 。相比于非参数法,参数法具有两方面的优势。一方面,参数法对于样本容量的要求较低。非参数法需以分组为基础,条件变量越多,分组越多,对样本容量的要求就越高。另一方面,参数法可以通过夏普利值分解方法,得到每个条件变量对机会不平等的贡献度。

在分解方面,量化研究中通常以自评健康水平(Self-Report Health, SRH)作为健康指标(Rosa Dias, 2009; Donni et al., 2014; Fajardo-Gonzalez, 2016; Gallardo et al., 2017),因而收入分配中的分解方法不能完全推广到健康差距的研究中。Jusot 等(2013)提出了一种基于方差的估计方法,该方法沿用了 Fields 分解方法(李莹、吕光明, 2016)。自评健康变量为有序

①在实证研究中,有文献将性别、种族作为导致机会不平等的变量(Ferreira and Gignoux, 2011; Bourguignon et al., 2007; Devooght, 2008),将性别、种族等个人的特征变量称为环境变量显然不合适。因此,本文沿用刘波等(2015)的做法,将环境变量和个人特征变量统一称为条件变量,并将超出自我控制的变量称为外生条件变量(Aaberge et al., 2010),剩余的变量称为其他条件变量。

变量,需通过 Ordered Logit/Probit 模型估计健康决定方程,基于健康决定方程得到的预测值并不是健康等级的预测值,而是每个健康等级的概率值。因此,需要给出一种适用于有序变量的测度方法与分解方式。

相比国外研究,国内研究主要集中于健康差距,而健康差距中的机会不平等还需进一步深入,尤其是在分解方面。在定义方面,史军和赵海燕(2010)基于罗尔斯的正义原则讨论了健康公平的内涵。李红文和毛新志(2015)从“地位综合症”的角度界定了健康公平,将健康公平描述为“消除各社会群体之间不公平的健康差别,消除那些由社会原因造成的健康不平等现象”,详细讨论了健康公平中的结果公平与手段公平,并从可避免性、自由选择性、责任主体三个方面给出了界定健康不公平的标准。

在测度方面,直接以健康差距中的机会不平等为主题的文献相对较少。赵广川(2017)以中国健康与营养调查(China Health and Nutrition Survey, CHNS)为样本测度了健康差距,并将有序变量转化为连续变量,估计健康差距中的机会不平等。鲁万波等(2018)从城乡的角度切入,测度中老年人健康差距中的机会不平等。除此之外,也有学者从社会地位和收入水平的角度,间接测度健康差距中的机会不平等。“地位综合症”或者“社会因果论”(Social Causation)是研究健康不平等的主要切入点之一,王甫勤(2012)、赵广川等(2016)从社会地位与健康生活方式、医疗服务利用之间的关系出发,分析了健康差距中的机会不平等。与收入相关的健康差距是近年来学者关注的热点,解垚(2009)以 CHNS 为样本数据的实证研究表明,我国存在亲富人的健康不平等、医疗服务利用不平等。黄潇(2012)对“与收入相关的健康不平等”进行测度和分解,研究结果表明,城乡居民中均存在亲富人的健康不平等,且累积效应在不断深化;陈东和张郁杨(2015)以中国健康与养老追踪调查(China Health and Retirement Longitudinal Study, CHARLS)中的中老年群体为样本的研究也得到了相类似的结论。

研究健康差距并不局限于测度健康差距本身,而是要挖掘出健康差距背后的现实基础,分解出各种因素对健康差距的贡献程度,从而为政策优化提供依据。已有文献为研究中国居民的健康机会不平等问题奠定了理论基础,但在关于中国居民健康差距的研究中,多采用分而治之的方式,评估单个因素对健康水平或者医疗保健服务利用的影响,而不是从全局出发,从总体上评估原生家庭背景、性别、城乡等因素在形成健康差距中的作用,揭示健康机会不平等的全貌,从而在众多因素中识别出导致健康机会不平等的主因。因此,有必要将影响居民健康的外生变量整合在一起,从总体上测度中国居民健康差距中的机会不平等,识别出导致健康机会不平等的主因。

三、测度模型

本文将采用参数法测度健康差距中的机会不平等,参数法共有四个步骤:(1)构建健康决定方程;(2)估计方程并得到相应的概率估计值;(3)基于预测值分解环境变量差异与努力变量差异对健康差距的相对贡献值;(4)将外生条件变量对健康差距的相对贡献值加总,从而得到健康机会不平等对健康差距的相对贡献值。

在测度健康差距中,通常以受访者的自评健康水平为对象(Rosa Dias, 2009; Donni et al., 2014; Fajardo-Gonzalez, 2016; Gallardo et al., 2017),自评健康水平为有序变量,因而健康决定

方程为离散选择模型,即 Ordered Logit/Probit 模型,为非线性模型。因此,基于健康决定方程得到的预测值,并非自评健康水平的预测值,而是每个自评健康等级的概率值。对于机会不平等的测度对象为有序变量,Juárez 和 Soloaga(2014)给出了一种测度方法,在其提出的方法中,根据受教育年限将样本个体的受教育程度划分成 5 个等级,并通过广义 Ordered Logit 模型预测概率 $\hat{p}_i = P(y_i > \tau | C_i)$,其中 τ 为受教育水平, $\tau = (1, 2, 3, 4, 5)$,然后将 $\hat{p}_i$ 作为测度机会不平等的基础。Juárez 和 Soloaga(2014)定义受教育水平的方式与自评健康水平具有相似性,因而可以将该方法用于测度健康差距中的机会不平等。

按照既定的步骤,首先是构建健康决定方程,影响健康的因素包括环境(C_i)、努力(E_i)和身份(D_i)三类变量,健康决定方程如式(1)所示:

$$H_i = f(C_i, E_i, D_i) = \alpha C_i + \beta E_i + \gamma D_i + \varepsilon_i \quad (1)$$

在健康机会不平等的语境下,努力程度通常包括生活习惯、健康素养,由于被解释变量为受访人的自评健康变量,受访人对自身健康状态的评价显然会影响个人的生活习惯。例如,如果受访人有“三高”(高血脂、高血糖、高血压),他可能会改变膳食结构、增加体育锻炼的频次、戒烟戒酒等,自评健康与当前的生活习惯存在互为因果的现象。因此,如果将抽烟、喝酒、运动、油炸食品等生活习惯变量作为努力程度变量纳入健康决定方程中,模型可能会存在内生性问题,从而影响系数估计值的无偏性与一致性。而且,居民的生活习惯存在显著的南北差异、地域之分,包含了条件变量的影响,如果将生活习惯变量纳入健康决定方程,可能会导致变量之间存在多重共线性。

在已有文献中,受教育水平常被作为努力程度的代理变量,个人早期所处的原生家庭环境、性别、出生年代是决定受教育水平的主要因素,因而需要将受教育水平中的原生家庭环境因素剔除。“地位综合症”是构建健康决定方程时必须考虑的问题,与受教育水平类似,个人当前的社会地位受原生家庭环境、身份变量等因素的影响,也包含个人努力。如果要将受教育水平、当前的社会地位作为解释变量,需要将两者包含的环境因素、身份变量因素剔除,并且可以将剥离原生环境变量与身份变量后的受教育水平、社会地位作为努力测度的代理变量。鉴于此,本文借鉴 Jusot 等(2013)、Juárez 和 Soloaga(2014)的方法,遵循 Roemer 和 Silvestre(1993)、Roemer(1998)的理论框架,通过两个辅助方程将受教育水平、当前社会地位中的环境因素剔除,辅助方程如(2)、(3)式:

$$edu_i = \delta_1 C'_i + \kappa_1 D_i + e_i \quad (2)$$

$$status_i = \delta_2 C'_i + \kappa_2 D_i + s_i \quad (3)$$

(2)、(3)式中: C' 为原生家庭环境, edu_i 与 $status_i$ 分别为受教育水平与个人的社会地位, e_i 与 s_i 为辅助方程的残差项。由于受教育水平与个人的社会地位是有序变量,因而需要采用 Ordered Probit/Logit 模型估计辅助方程。对于 Ordered Probit/Logit 模型不能得到一般意义上的残差项预测值,但通过 Gouriéroux 等(1987)给出的方法,可以得到式(2)、(3)的广义残差值估计值 $\hat{e}_i$ 、 $\hat{s}_i$,两者均为努力程度的代理变量。令 $\hat{E}_i = (\hat{e}_i, \hat{s}_i)$, (1)式可以改写为:

$$H_i = f(C_i, \hat{E}_i, D_i) = \alpha C_i + \beta \hat{E}_i + \gamma D_i + \varepsilon_i \quad (4)$$

当然,也可以直接以(1)式测度健康机会不平等,但由于没有剔除努力程度中的外生因素,从而可能会低估健康机会不平等。相比于 Ordered Probit 模型,Ordered Logit 模型可以得到解释变量的几率比,能够更为直观地解释变量的现实意义,因而本文采用 Ordered Logit 模

型估计健康决定方程。由 Ordered Logit 模型的性质可知,在给定健康水平(h)的条件下,相应的概率值 $P(H_i=h)$ ($h=1, \dots, M$) 为:

$$P(H_i=1)=\frac{\exp[\kappa_1-(\alpha C_i+\beta \hat{E}_i+\gamma D_i)]}{1+\exp[\kappa_1-(\alpha C_i+\beta \hat{E}_i+\gamma D_i)]} \quad (5)$$

$$P(H_i=h)=\frac{\exp[\kappa_h-(\alpha C_i+\beta \hat{E}_i+\gamma D_i)]}{1+\exp[\kappa_h-(\alpha C_i+\beta \hat{E}_i+\gamma D_i)]}-\frac{\exp[\kappa_{h-1}-(\alpha C_i+\beta \hat{E}_i+\gamma D_i)]}{1+\exp[\kappa_{h-1}-(\alpha C_i+\beta \hat{E}_i+\gamma D_i)]}, h=2, \dots, M-1 \quad (6)$$

$$P(H_i=M)=\frac{\exp[\kappa_{M-1}-(\alpha C_i+\beta \hat{E}_i+\gamma D_i)]}{1+\exp[\kappa_{M-1}-(\alpha C_i+\beta \hat{E}_i+\gamma D_i)]} \quad (7)$$

由(5)-(7)式可得 $P(H_i>h)$ 、 $P(H_i\leq h)$ ($h=1, \dots, M-1$) 的概率值:

$$\hat{p}_{h,i}=P(H_i>h)=\frac{1}{1+\exp[\kappa_h-(\alpha C_i+\beta \hat{E}_i+\gamma D_i)]} \quad (8)$$

$$1-\hat{p}_{h,i}=P(H_i\leq h)=\frac{\exp[\kappa_h-(\alpha C_i+\beta \hat{E}_i+\gamma D_i)]}{1+\exp[\kappa_h-(\alpha C_i+\beta \hat{E}_i+\gamma D_i)]} \quad (9)$$

根据 Jusot 等(2013)、Juárez 和 Soloaga(2014)给出的方法,在给定 h 的条件下,健康差距的绝对值为:

$$I(\hat{p}_{h,i})=I[P(H_i>h|C_i,\hat{E}_i,D_i)] \quad (10)$$

以此为基础,通过不平等指数分解方法,可以分解出每个解释变量对健康差距的相对贡献值。在已有文献中,Shorrocks(2013)提出的夏普利值(Shapley Value)分解方法是应用最为广泛的分解方法。令 $X_i=(C_i,\hat{E}_i,D_i)$,将 $I(\hat{p}_{h,i})$ 改写为 $I(\hat{p}_{h,i})=F(X_i)$,假定 X_i 包含 K 个变量,以第 k 个变量为例,将 S_{k1} 定义为第一轮效应:

$$S_{k1}=I(F(X))-I(F(X(k))) \quad (11)$$

(11)式中: $X(k)=[x_1, \dots, \bar{x}_k, \dots, x_K]$ 。相应地, S_{k2} 为第二轮效应:

$$S_{k2}=I[F(X(k))]-I[F(X(j,k))] \quad (12)$$

(12)式中: $X(j,k)=[x_1, \dots, \bar{x}_j, \dots, \bar{x}_k, \dots, x_K]$ 。以此类推, S_{kN} 为第 N 轮效应:

$$S_{kN}=I[F(X(1, \dots, k-1, k+1, \dots, K))]-I[F(X(1, \dots, K))] \quad (13)$$

(13)式中: $X(1, \dots, k-1, k+1, \dots, K)=[\bar{x}_1, \dots, \bar{x}_{k-1}, x_k, \bar{x}_{k+1}, \dots, \bar{x}_K]$, $X(1, \dots, K)=[\bar{x}_1, \dots, \bar{x}_K]$ 。最终可以得到因素 k 对机会不平等的相对贡献度 P_k :

$$\textcircled{1} \text{根据(8)、(9)式可以定义几率比: } OR=\frac{P(H_i>h)}{P(H_i\leq h)}=\frac{\hat{p}_{h,i}}{1-\hat{p}_{h,i}}=1/\exp[\kappa_h-(\alpha C_i+\beta \hat{E}_i+\gamma D_i)]。进一步恒等变$$

换可得: $\ln(OR)=\ln\left(\frac{\hat{p}_{h,i}}{1-\hat{p}_{h,i}}\right)=-\kappa_h+(\alpha C_i+\beta \hat{E}_i+\gamma D_i)$,按照几率比的解释方式,以变量集 C_i 为例,假定变量 C_{ij} 的系数 α_{ij} 为正值,在给定 h 的条件之下,变量 C_{ij} 增加一个单位,几率比将增加 $100\times\alpha_{ij}\%$ (若 α_{ij} 为负值,则降低 $100\times\alpha_{ij}\%$)。

$$S_k = \left(\sum_{j=1}^N S_{kj} \right) / N \quad (14)$$

$$P_k = S_k / \left(\sum_{k=1}^K S_k \right) \quad (15)$$

此外,与一般线性模型不同,在 Ordered Logit 模型中不涉及残差项问题(Devooght, 2008; Lefranc et al., 2009; Jusot et al., 2013),故不需要讨论残差项的归属问题。得到解释变量差异对健康差距的相对贡献值之后,再将外生条件变量对健康差距的相对贡献值加总就可以得到健康机会不平等的相对值。与此同时,根据变量的相对贡献值,可以识别出导致健康机会不平等以及健康差距的主要因素。

一般而言,随着 h 的增加,概率 $\hat{p}_{h,i}$ 的数值在总体上将随之下降。在 $\hat{p}_{h,i}$ 下降的同时, $\hat{p}_{h,i}$ 在个体之间的差异程度的变化却是不确定的。如果 $\hat{p}_{h,i}$ 在个体之间的差异程度随着 h 的增加而下降,则说明 $\hat{p}_{h,i}$ 的变化是趋同的,即在较高程度的健康水平上,居民健康差距更小。但如果 $\hat{p}_{h,i}$ 在个体之间的差异程度随着 h 的增加而增加,则说明 $\hat{p}_{h,i}$ 的变化是发散的,即在较高程度的健康水平上,居民健康差距更大。Juárez 和 Soloaga (2014) 以差异性指数(Dissimilarity Index, DI)为不平等指标,给出了一种分解 $D(\hat{p}_h)$ 随 h 变动的方法。该指数计算方法如(16)式所示:

$$D(\hat{p}_h) = \frac{1}{2N\bar{\hat{p}}_h} \sum_{i=1}^N |\hat{p}_{h,i} - \bar{\hat{p}}_h| = \underbrace{\frac{\sum_{i=1}^N |\hat{p}_{h,i} - \bar{\hat{p}}_h|}{2N}}_{\text{dispersion}} \cdot \underbrace{\frac{1}{\bar{\hat{p}}_h}}_{\text{level}} \quad (16)$$

(16) 式中: $\bar{\hat{p}}_h$ 为变量 $\hat{p}_{h,i}$ 的均值, $\bar{\hat{p}}_h = \sum_{i=1}^N \hat{p}_{h,i} / N$ 。差异性指数可进一步分解为分散(Dispersion)部分和水平(Level)部分,其中分散部分与 σ -收敛的形式非常接近^①。基于此,可将两个相邻健康水平的差异性指数之差 $\Delta_h D(\hat{p}_h) = D(\hat{p}_h) - D(\hat{p}_{h-1})$ 分解为分散效应和水平效应:

$$\begin{aligned} \Delta_h D(\hat{p}_h) = & \underbrace{\left[\frac{\sum_{i=1}^N |\hat{p}_{h,i} - \bar{\hat{p}}_h|}{2N\bar{\hat{p}}_h} - \frac{\sum_{i=1}^N |\hat{p}_{h-1,i} - \bar{\hat{p}}_{h-1}|}{2N\bar{\hat{p}}_h} \right]}_{\text{dispersion effect}} + \\ & \underbrace{\left[\frac{\sum_{i=1}^N |\hat{p}_{h-1,i} - \bar{\hat{p}}_{h-1}|}{2N\bar{\hat{p}}_h} - \frac{\sum_{i=1}^N |\hat{p}_{h-1,i} - \bar{\hat{p}}_{h-1}|}{2N\bar{\hat{p}}_{h-1}} \right]}_{\text{level effect}} \end{aligned} \quad (17)$$

(17) 式中分解方法与 Jenkins 和 Van Kerm (2006) 给出的分解方法具有相似之处,该分解方法同样适用于基尼系数。对于离散变量 y ,基尼系数的计算公式为:

$$G(\hat{p}_h) = \frac{\sum_{i=1}^N \sum_{j=1}^N |\hat{p}_{h,i} - \hat{p}_{h,j}|}{2\bar{\hat{p}}_h N(N-1)} \quad (18)$$

基于此,两个阶段的基尼系数之差 ($\Delta_h G(\hat{p}_h) = G(\hat{p}_h) - G(\hat{p}_{h-1})$) 可以分解为如下形式:

^① σ -收敛: $\sigma^2 = \frac{1}{N} \sum_{i=1}^N \left(\ln y_{it} - \frac{1}{N} \sum_{i=1}^N \ln y_{it} \right)^2$ 。

$$\Delta_h G(\hat{p}_h) = \underbrace{\left[\frac{\sum_{i=1}^N \sum_{j=1}^N |\hat{p}_{h,i} - \hat{p}_{h,j}|}{2 \bar{\hat{p}}_h N(N-1)} - \frac{\sum_{i=1}^N \sum_{j=1}^N |\hat{p}_{h-1,i} - \hat{p}_{h-1,j}|}{2 \bar{\hat{p}}_h N(N-1)} \right]}_{\text{dispersion effect}} + \underbrace{\left[\frac{\sum_{i=1}^N \sum_{j=1}^N |\hat{p}_{h-1,i} - \hat{p}_{h-1,j}|}{2 \bar{\hat{p}}_h N(N-1)} - \frac{\sum_{i=1}^N \sum_{j=1}^N |\hat{p}_{h-1,i} - \hat{p}_{h-1,j}|}{2 \bar{\hat{p}}_{h-1} N(N-1)} \right]}_{\text{level effect}} \quad (19)$$

由(19)式可知,如果离差效应不断扩大,则预示着在较高程度的健康水平上,是趋于更加不平等的;反之,则趋于更加平等。需要特别说明的是,自评健康水平(H)是一个包含 M 个数值的有序变量,因而将得到 $M-1$ 个概率预测值 $\hat{p}_{h,i}$ 以及 $M-1$ 个夏普利值分解结果,所以得到的机会不平等是一个区间值,而非一个固定值。Fajardo-Gonzalez(2016)给出了一种处理方式,即将“一般”(不含“一般”)以上的健康水平赋值为 1,其他赋值为 0,将有序变量转化为 0-1 变量,相应的机会不平等就转化为单个取值。

现实中,有诸多条件变量影响居民的健康水平,但是可观测的、可量化的条件变量却是有限的。在实证研究中,能够获取的条件变量也是有限的,不可能将所有条件变量囊括在内,但这并不影响测度结果的有效性。在条件变量有限的情况下,得到的机会不平等可视为基于该变量集得到的下限值;如果在该变量集上不断增加条件变量,机会不平等将不断逼近真实水平。因此,在现有调查数据的基础上,可以尽可能囊括更多的条件变量,即使不能包含所有的条件变量,但也可以得到机会不平等的下限值,通过下限值同样可以了解健康机会不平等的现状与成因。

四、实证研究

(一) 变量的选取

本文以中国家庭追踪调查数据(China Family Panel Studies, CFPS)2010年、2012年、2014年和2016年为样本构建平衡面板数据,为了充分反映社会经济环境对居民健康的影响,本文将研究对象设定为45岁及以上的受访者。在健康决定方程中,健康变量为自评健康水平,包括不健康、一般、比较健康、很健康、非常健康5个等级。环境变量包括受访人3岁时的户口与经济区域、父母是否为文盲、城乡、主食类型,身份变量包括性别、民族、年龄,努力变量为剔除外生条件变量影响后的教育与社会地位,辅助方程中的变量包括受访人3岁时的户口与经济区域、父母是否为文盲、性别、民族、年龄。主食类型包括大米、面粉和杂粮,主食类型可能兼具客观性与主观性,对于城市居民来说,地域对主食类型的限制不断减弱,但农村居民普遍存在“靠山吃山,靠水吃水”、“南人喜米,北人好面”的现象,因而将其作为受访人所处地理环境的代理变量。年龄增加导致的健康水平下降是合乎正义法则的,因而在测度健康机会不平等时,将受访人3岁时的户口与经济区域、父母的受教育水平是否为小学以下、性别、民族、主食类型共9个变量设定为考察健康机会不平等的内生条件变量,即将上述9个内生变量差异导致的健康差距定义为健康机会不平等。变量的定义详见表1。

与已有文献不同的是,本文以平衡面板数据估计机会不平等,可以在一定程度上提升估计结果的稳健性。经过数据整理,样本共包括7209个受访者连续4期的观测数据,共计28836个观测值,样本数据的描述性统计结果如表2所示。

表 1 变量的定义

变量名	符号	变量类型与定义
自评健康	<i>health</i>	1:不健康;2:一般;3:比较健康;4:很健康;5:非常健康
经济区域	<i>west₀₃</i> 、 <i>mid₀₃</i> 、 <i>east₀₃</i>	受访人3岁时所在的省份为西部时, <i>west₀₃</i> = 1,否则为0;若为中部省份, <i>mid₀₃</i> = 1,否则为0;若为东部省份, <i>east₀₃</i> = 1,否则为0
户口类型	<i>hukou₀₃</i>	受访人3岁时为非农户口, <i>hukou₀₃</i> = 1,否则为0
父母受教育水平	<i>fedu</i> 、 <i>medu</i>	父亲为文盲或半文盲, <i>fedu</i> = 1,否则为0;母亲为文盲或半文盲, <i>medu</i> = 1,否则为0
城乡	<i>urban</i>	根据国家统计局的城乡分类标准,受访地为城镇地区时, <i>urban</i> = 1,否则为0
主食	<i>food₁</i> - <i>food₃</i>	根据大米、面粉和杂粮依次赋值
性别	<i>gender</i>	受访人为男性, <i>gender</i> = 1,否则为0
民族	<i>minzu</i>	受访人为汉族, <i>minzu</i> = 1,否则为0
年龄	<i>age</i>	受访人在2010年、2012年、2014年、2016年的年龄
受教育水平	<i>edu</i>	根据受教育水平(1:文盲/半文盲;2:小学;3:初中;4:高中/中专/技校/职高;5:大专;6:大学本科;7:硕士)依次赋值
社会地位	<i>status</i>	根据受访人的自评结果(5个等级,1为最低,5为最高),依次赋值

表 2 变量的描述性统计

变量	样本量	均值	标准差	最小值	最大值
<i>health</i>	7 209	2.9147	1.3241	1	5
<i>west₀₃</i>	7 209	0.2448	0.4300	0	1
<i>mid₀₃</i>	7 209	0.2992	0.4579	0	1
<i>east₀₃</i>	7 209	0.4560	0.4981	0	1
<i>hukou₀₃</i>	7 209	0.1228	0.3282	0	1
<i>fedu</i>	7 209	0.3364	0.4725	0	1
<i>medu</i>	7 209	0.1160	0.3202	0	1
<i>urban</i>	7 209	0.4341	0.4957	0	1
<i>gender</i>	7 209	0.5099	0.4999	0	1
<i>minzu</i>	7 209	0.9424	0.2329	0	1
<i>age</i>	7 209	59.423	8.4788	45	94
<i>food₁</i>	7 209	0.5357	0.4987	0	1
<i>food₂</i>	7 209	0.3951	0.4889	0	1
<i>food₃</i>	7 209	0.0691	0.2537	0	1
<i>edu</i>	7 209	2.1820	1.1908	1	7
<i>status</i>	7 209	2.9344	1.0559	1	5

(二) 健康决定方程的估计

健康决定方程的估计分为两步,首先是估计辅助方程,然后是估计健康决定方程。在健康决定方程的估计中,本文采用面板 Ordered Logit 模型估计方程,估计结果如表 3 所示。似然比(LR)检验的结果表明,相比于混合效应模型,随机效应模型更优。在(1)-(2)列中列示了辅助方程的估计结果,在(3)-(4)列中列示了健康决定方程的估计结果,*edu_res_hat*、*status_res_hat* 为基于辅助方程得到的广义残差估计值。

表 3 基于面板 Ordered Logit 模型的健康决定方程

变量	(1)	(2)	(3)	(4)
	辅助方程		健康决定方程	
	<i>edu</i>	<i>status</i>	估计系数	几率比
<i>mid</i> ₀₃	0.2859 *** (0.0185)	-0.0723 *** (0.0174)	0.3117 *** (0.0512)	1.3658 *** (0.0699)
<i>east</i> ₀₃	0.2942 *** (0.0171)	-0.2037 *** (0.0161)	0.3642 *** (0.0480)	1.4394 *** (0.0691)
<i>hukou</i> ₀₃	0.9956 *** (0.0201)	-0.3088 *** (0.0198)	0.0204 (0.0605)	1.0207 (0.0618)
<i>fedu</i>	-0.3770 *** (0.0148)	-0.0475 *** (0.0144)	0.0570 (0.0421)	1.0587 (0.0445)
<i>medu</i>	-0.2727 *** (0.0216)	-0.0419 * (0.0216)	0.0821 (0.0630)	1.0856 (0.0684)
<i>urban</i>			0.0759 * (0.0397)	1.0789 * (0.0429)
<i>gender</i>	0.5831 *** (0.0133)	-0.0110 (0.0125)	0.5794 *** (0.0368)	1.7850 *** (0.0657)
<i>minzu</i>	0.0767 *** (0.0293)	-0.0365 (0.0273)	-0.0318 (0.0811)	0.9687 (0.0786)
<i>age</i>	-0.0336 *** (0.0008)	0.0102 *** (0.0008)	-0.0861 *** (0.0025)	0.9175 *** (0.0023)
<i>food</i> ₂			0.1123 *** (0.0386)	1.1189 *** (0.0432)
<i>food</i> ₃			-0.0209 (0.0633)	0.9793 (0.0620)
<i>edu_res_hat</i>			-0.0223 (0.0204)	0.9779 (0.0200)
<i>status_res_hat</i>			0.1868 *** (0.0138)	1.2054 *** (0.0166)
切点值[1]	-2.0964 *** (0.0563)	-0.8654 *** (0.0527)	-6.2273 *** (0.1714)	0.0020 *** (0.0003)
切点值[2]	-1.4464 *** (0.0557)	-0.2518 *** (0.0524)	-4.9919 *** (0.1681)	0.0068 *** (0.0011)
切点值[3]	-0.5486 *** (0.0553)	1.0334 *** (0.0526)	-3.5090 *** (0.1646)	0.0299 *** (0.0049)
切点值[4]	0.4825 *** (0.0568)	1.7563 *** (0.0533)	-2.1526 *** (0.1623)	0.1162 *** (0.0189)
切点值[5]	1.0036 *** (0.0602)			
切点值[6]	1.9650 *** (0.0947)			
σ_u^2			1.5265 *** (0.0558)	4.6023 *** (0.2567)
样本容量	28 836	28 836	28 836	28 836
观测对象			7 209	7 209

注：*、**、*** 分别表示通过了 10%、5%、1% 的显著性检验；括号中为标准差。下同。

由健康决定方程的估计结果可知,在 1%的水平上,3 岁时所处的经济区域、性别、年龄、主食类型、剥离环境因素后的社会地位对自评健康水平的影响是显著的;在 10%的水平上,城乡对自评健康水平存在显著影响。具体来看,在给定自评健康水平 h 的条件下,相比于西部地区的受访人,对于中部地区受访人的几率比 $P(H_i > h)/P(H_i \leq h)$ 将显著增加 36.58%,东部地区受访人的几率比将显著增加 43.94%;相比于乡村地区的受访人,城镇受访人的几率比将显著增加 7.89%;相比于女性受访人,男性受访人的几率比将显著增加 78.50%;如果受访人的年龄增加一个单位,几率比将显著下降 8.25%;相比于以大米为主食的受访者,以面粉为主食的受访人的几率比显著增加 11.89%;剥离外生条件影响后的个人社会地位增加一个单位,几率比将显著提升 20.54%。综上,地区、性别、年龄、剥离环境因素后的社会地位 ($status_res_hat$) 是影响自评健康的主要因素。

(三) 健康机会不平等的测度与分解

1. 效应分解

在估计健康决定方程之后,可以得到 $P(H_i = h)$ 的估计值,然后以此为基础反推 $P(H_i > h)$ 的估计值,表 4 中列示了基尼系数 $G(\hat{p}_h)$ 的估计结果,依次为 0.0731、0.1297、0.2109、0.2815,即 $G(\hat{p}_h)$ 的取值范围为 0.0731~0.2815, $\Delta_h G(\hat{p}_h)$ 的分解结果亦如表 4 所示。从总体上看,受访人处于不健康之外的 4 种健康状态(一般、比较健康、很健康、非常健康,即 $P(H_i > 1)$) 的概率较高,平均概率为 0.7911,相应的基尼系数仅为 0.0731。相比于 $P(H_i > 1)$,受访人健康状态处于非常健康 $P(H_i > 4)$ 的概率相对较低,平均概率为 0.1598,但基尼系数为 0.2815,约为前者的 4 倍。由此可见,位于不健康状态之上的概率相对较高,且个体之间的相对差距较小。与之形成鲜明对比的是,总体上能够实现最佳健康状态的概率相对较小,且个体之间的相对差距较大。

随着 h 的提升,基尼系数 $G(\hat{p}_h)$ 逐步增加,由基尼系数的公式可知, $G(\hat{p}_h)$ 随 h 的提升而增加,可能是因为 $\bar{\hat{p}}_h$ 不断下降,也可能是因为 $\hat{p}_h$ 在个体之间的差异扩大。进一步通过效应分解来确定 $\Delta_h G(\hat{p}_h)$ 变动的主要原因,分解结果如表 4 所示。效应分解的结果表明,从效应分解的结果来看,虽然总效应为正,但分散效应和水平效应却略有不同,分散效应不断缩小,水平效应不断扩大。随着自评健康水平 h 的提升, $\hat{p}_h$ 在个体之间的差异性不断缩小,平均值 $\bar{\hat{p}}_h$ 不断下降,且个体差异缩小的幅度远低于均值下降的幅度,基尼系数 $G(\hat{p}_h)$ 上升的动力主要来自于水平效应。由此可见,随着自评健康水平 h 的提升, $\hat{p}_h$ 在个体之间是不断趋同的,健康的相对差距并未随着自评健康水平的提升而扩大。

表 4 $\hat{p}_h$ 的基尼系数及效应分解

$\hat{p}_h$	$\bar{\hat{p}}_h$	基尼系数 $G(\hat{p}_h)$	总效应	分散效应	水平效应
$P(H_i > 1)$	0.7911	0.0731			
$P(H_i > 2)$	0.6049	0.1297	0.0593	0.0362	0.0231
$P(H_i > 3)$	0.3458	0.2109	0.0865	-0.0137	0.1002
$P(H_i > 4)$	0.1598	0.2815	0.0765	-0.1780	0.2545

2. 夏普利值分解

进一步结合夏普利值分解方法,分解出每个解释变量对健康差距的相对贡献度,分解结果如表 5 所示。3 岁时所处的地区差异对健康差距的相对贡献度为 4.072%~4.868%,3 岁时

的户口差异为 0.040%~0.045%, 父母受教育水平差异为-0.186%~0.011%, 城乡差异为 0.244%~0.393%, 性别差异为 13.266%~16.830%, 民族差异为-0.032%~-0.027%, 主食差异为 0.902%~1.075%, 年龄差异为 69.053%~75.185%。剥离环境因素后的受教育水平与社会地位作为努力程度的代理变量, 两者的差异对健康差距的相对贡献度为 6.499%~7.765%。其中, 年龄差异是影响健康差距的主要因素, 性别次之, 努力程度位居第三, 地区差异位列第四, 主食差异位列第五。

以分解结果为基础, 将外生条件变量导致的健康差距加总可得健康机会不平等对健康差距的相对贡献度为 18.316%~23.182%, 如果采用 Fajardo-Gonzalez (2016) 的处理方式, 将“3 比较健康、4 很健康、5 非常健康”赋值为 1, “1 不健康、2 一般”赋值为 0, 此时健康机会不平等对健康差距的相对贡献度为 20.715%。在外生变量中, 性别、3 岁时所处的经济区域是导致健康机会不平等的主要因素。

表 5 夏普利值分解结果

变量	$P(H_i > 1)$		$P(H_i > 2)$		$P(H_i > 3)$		$P(H_i > 4)$	
	绝对值	相对值	绝对值	相对值	绝对值	相对值	绝对值	相对值
mid_{03}	0.0011	1.779%	0.0025	1.901%	0.0046	2.021%	0.0061	2.066%
$east_{03}$	0.0014	2.293%	0.0035	2.661%	0.0065	2.839%	0.0082	2.802%
$hukou_{03}$	0.0000	0.045%	0.0001	0.045%	0.0001	0.043%	0.0001	0.040%
$fedu$	0.0000	-0.049%	0.0000	0.009%	0.0001	0.055%	0.0002	0.058%
$medu$	-0.0001	-0.137%	-0.0001	-0.077%	-0.0001	-0.044%	-0.0001	-0.050%
$urban$	0.0002	0.244%	0.0004	0.325%	0.0009	0.385%	0.0012	0.393%
$gender$	0.0083	13.266%	0.0196	14.907%	0.0376	16.398%	0.0494	16.830%
$minzu$	0.0000	-0.027%	0.0000	-0.028%	-0.0001	-0.030%	-0.0001	-0.032%
$food_2$	0.0005	0.868%	0.0012	0.935%	0.0023	1.008%	0.0030	1.033%
$food_3$	0.0000	0.034%	0.0000	0.037%	0.0001	0.041%	0.0001	0.042%
age	0.0469	75.185%	0.0948	72.254%	0.1598	69.705%	0.2026	69.053%
edu_res_hat	0.0001	0.098%	0.0001	0.078%	0.0002	0.073%	0.0002	0.081%
$status_res_hat$	0.0040	6.402%	0.0091	6.953%	0.0172	7.506%	0.0225	7.684%

五、稳健性检验

从上述测度机会不平等的流程可知, 机会不平等程度的绝对值依赖于概率的估计值 $\hat{p}_{h,i}$, 而概率估计值 $\hat{p}_{h,i}$ 又依赖于模型设定的形式, 因此有必要调整健康决定方程的估计模型。在面板 Ordered Logit 模型的实证研究中, 需要通过 $P(H_i = h)$ ($h = 1, 2, 3, 4, 5$) 的估计值得到 $\hat{p}_{h,i}$ 的估计值, 因而可以考虑直接以 $I(H_i > h)$ 为被解释变量, 将面板 Ordered Logit 模型转化为 4 个面板 Logit 模型, 逐一获取 $\hat{p}_{h,i}$ 的估计值。另一方面, 虽然似然比检验现实随机效应模型优于混合效应模型, 但在面板 Ordered Logit 模型或面板 Logit 模型中, 随机效应需要以较为严格的假定为前提, 有鉴于此, 在稳健性检验中, 本文采用广义 Ordered Logit 模型估计 $P(H_i > h)$ 的值。由于健康机会不平等对健康差距的贡献度是本文考察的重点, 故在稳健性检验中不再列示效应分解的结果。

(一) 基于面板 Logit 模型的稳健性检验

相比于面板 Ordered Logit 模型, 面板 Logit 模型的估计结果可以给出更多的信息, 估计

结果如表 6 所示,似然比检验表明,随机效应模型优于混合效应模型。例如,在两个极端的概率 $P(H_i>1)$ 与 $P(H_i>4)$ 中,以对自评健康较为突出的性别为例,在以 $I(H_i>1)$ 为被解释变量的估计结果中,相比于女性,男性的几率比将高出 136.52%;而在以 $I(H_i>4)$ 为被解释变量的估计结果中,相比于女性,男性的几率比则高出 66.13%。两相对比可知,对于能否逃离不健康状态 $I(H_i>1)$ 与实现最佳健康状态 $I(H_i>4)$ 而言,解释变量对两者的作用程度的大小以及作用方向并不一致。

表 6 基于面板 Logit 模型的健康决定方程

变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$I(H_i>1)$		$I(H_i>2)$		$I(H_i>3)$		$I(H_i>4)$	
	系数	几率比	系数	几率比	系数	几率比	系数	几率比
<i>mid</i> ₀₃	0.4640 *** (0.0807)	1.5904 *** (0.1283)	0.3003 *** (0.0617)	1.3503 *** (0.0833)	0.2753 *** (0.0447)	1.3170 *** (0.0589)	0.2012 *** (0.0614)	1.2228 *** (0.0751)
<i>east</i> ₀₃	0.5266 *** (0.0757)	1.6931 *** (0.1281)	0.3224 *** (0.0578)	1.3805 *** (0.0797)	0.3225 *** (0.0422)	1.3806 *** (0.0583)	0.2021 *** (0.0582)	1.2239 *** (0.0713)
<i>hukou</i> ₀₃	0.4086 *** (0.1043)	1.5047 *** (0.1569)	0.1520 ** (0.0744)	1.1641 ** (0.0866)	-0.1215 ** (0.0537)	0.8856 ** (0.0475)	-0.3252 *** (0.0774)	0.7224 *** (0.0559)
<i>fedu</i>	-0.1197 * (0.0682)	0.8872 * (0.0605)	0.0113 (0.0510)	1.0114 (0.0516)	0.0949 *** (0.0367)	1.0996 *** (0.0404)	0.1059 ** (0.0509)	1.1117 ** (0.0566)
<i>medu</i>	-0.1365 (0.1083)	0.8724 (0.0945)	-0.1093 (0.0780)	0.8964 (0.0699)	0.1166 ** (0.0547)	1.1236 ** (0.0615)	0.0670 (0.0749)	1.0693 (0.0801)
<i>urban</i>	0.2464 *** (0.0639)	1.2794 *** (0.0817)	0.0643 (0.0485)	1.0664 (0.0517)	0.0290 (0.0358)	1.0294 (0.0369)	0.0336 (0.0494)	1.0341 (0.0510)
<i>gender</i>	0.8609 *** (0.0599)	2.3652 *** (0.1416)	0.6676 *** (0.0447)	1.9496 *** (0.0871)	0.3749 *** (0.0321)	1.4548 *** (0.0467)	0.5076 *** (0.0448)	1.6613 *** (0.0744)
<i>minzu</i>	0.0653 (0.1260)	1.0675 (0.1345)	0.0788 (0.0972)	1.0820 (0.1052)	-0.1465 ** (0.0706)	0.8637 ** (0.0610)	0.0471 (0.1002)	1.0482 (0.1050)
<i>age</i>	-0.0775 *** (0.0038)	0.9254 *** (0.0035)	-0.0683 *** (0.0028)	0.9340 *** (0.0026)	-0.0669 *** (0.0023)	0.9353 *** (0.0022)	-0.0718 *** (0.0032)	0.9307 *** (0.0029)
<i>food</i> ₂	-0.0577 (0.0619)	0.9439 (0.0584)	0.1890 *** (0.0472)	1.2080 *** (0.0570)	0.1050 *** (0.0348)	1.1107 *** (0.0387)	0.1735 *** (0.0480)	1.1895 *** (0.0571)
<i>food</i> ₃	-0.2987 *** (0.0963)	0.7418 *** (0.0714)	0.0438 (0.0776)	1.0448 (0.0811)	0.0050 (0.0620)	1.0050 (0.0624)	0.0120 (0.0860)	1.0121 (0.0871)
<i>edu_res_hat</i>	0.2511 *** (0.0336)	1.2854 *** (0.0431)	0.1391 *** (0.0248)	1.1492 *** (0.0285)	-0.1192 *** (0.0181)	0.8876 ** (0.0161)	-0.1203 *** (0.0251)	0.8866 *** (0.0223)
<i>status_res_hat</i>	0.2109 *** (0.0216)	1.2348 *** (0.0267)	0.1884 *** (0.0175)	1.2073 *** (0.0211)	0.1668 *** (0.0153)	1.1816 *** (0.0181)	0.2534 *** (0.0208)	1.2884 *** (0.0268)
常数项	6.0432 *** (0.2651)		4.0644 *** (0.1941)		2.7299 *** (0.1471)		1.4770 *** (0.1965)	
样本容量	28 836	28 836	28 836	28 836	28 836	28 836	28 836	28 836
观测对象	7 209	7 209	7 209	7 209	7 209	7 209	7 209	7 209

值得关注的是,3岁时的户口类型、父母的受教育水平、主食类型对于 $I(H_i>h)$ 的影响并不一致。例如,如果受访人3岁时为城镇户口,几率比 $P(H_i>1)/[1-P(H_i>1)]$ 、 $P(H_i>2)/[1-P(H_i>2)]$ 会显著提升,但几率比 $P(H_i>3)/[1-P(H_i>3)]$ 、 $P(H_i>4)/[1-P(H_i>4)]$ 会

显著下降。由此可见,3 岁时的户口对于脱离不健康以及实现一般以上的健康水平是有利的,但对于实现比较健康以上的健康水平则不然。

以面板 Logit 模型的估计结果为基础,表 7 中给出了夏普利值分解的结果。3 岁时所处的地区差异对健康差距的相对贡献度为 1.444%~6.865%,3 岁时的户口差异为 0.595%~3.321%,父母受教育水平差异为 0.624%~2.413%,城乡差异为-0.163%~3.290%,性别差异为 10.709%~21.211%,民族差异为 0.118%~0.252%,主食差异为 1.049%~2.573%,年龄差异为 50.471%~70.759%,努力程度差异(经辅助方程调整后的受教育水平与社会地位)对健康差距的相对贡献度为 11.644%~16.751%。再将外生变量对健康差距的相对贡献度加总,健康机会不平等对健康差距的相对贡献度为 18.120%~34.525%,按照 Fajardo-Gonzalez(2016)界定方式,健康机会不平等对健康差距的相对贡献度为 29.961%。与上文的实证研究类似,性别与 3 岁时所处的经济区域仍然是导致健康机会不平等的主要因素。

表 7 基于面板 Logit 模型的夏普利值分解

变量	$P(H_i>1)$		$P(H_i>2)$		$P(H_i>3)$		$P(H_i>4)$	
	绝对值	相对值	绝对值	相对值	绝对值	相对值	绝对值	相对值
<i>mid</i> ₀₃	0.0010	2.076%	0.0024	2.237%	0.0035	2.013%	0.0038	1.473%
<i>east</i> ₀₃	0.0024	4.789%	0.0020	1.875%	0.0043	2.519%	-0.0001	-0.029%
<i>hukou</i> ₀₃	0.0017	3.321%	0.0010	0.912%	0.0010	0.595%	0.0073	2.805%
<i>fedu</i>	0.0007	1.364%	-0.0001	-0.067%	0.0008	0.464%	0.0014	0.552%
<i>medu</i>	0.0005	1.049%	0.0009	0.861%	0.0003	0.179%	0.0002	0.072%
<i>urban</i>	0.0016	3.290%	0.0005	0.502%	-0.0001	-0.085%	-0.0004	-0.163%
<i>gender</i>	0.0099	19.878%	0.0231	21.211%	0.0185	10.709%	0.0390	15.015%
<i>minzu</i>	0.0001	0.131%	0.0003	0.252%	0.0002	0.129%	0.0003	0.118%
<i>food</i> ₂	0.0001	0.137%	0.0024	2.208%	0.0019	1.077%	0.0067	2.584%
<i>food</i> ₃	0.0005	0.912%	0.0000	-0.030%	0.0000	-0.002%	0.0000	-0.011%
<i>age</i>	0.0252	50.471%	0.0633	58.036%	0.1222	70.759%	0.1580	60.833%
<i>edu_res_hat</i>	0.0035	7.027%	0.0048	4.395%	0.0059	3.416%	0.0067	2.590%
<i>status_res_hat</i>	0.0028	5.554%	0.0083	7.609%	0.0142	8.228%	0.0368	14.161%

(二) 基于广义 Ordered Logit 模型的稳健性检验

由于面板 Ordered Logit 模型与面板 Logit 模型中,随机效应的假定条件较为严格,因而在本部分的稳健性检验中,直接采用混合效应模型。为了获取概率 $P(H_i>h)$ 的估计值,可直接采用广义 Logit 模型估计健康决定方程。相比于 Ordered Logit 模型,广义 Ordered Logit 模型同样可以捕捉到系数估计值在不同门限值 h 下的异质性(Williams,2016),估计结果与面板 Logit 模型的估计结果具有一致性,故不再赘述。以估计结果为基础,进一步通过夏普利值分解方法分解变量对健康差距的相对贡献度,如表 8 所示。

由分解结果可知,3 岁时所处的地区、户口、父母受教育水平、城乡、性别、民族、主食、年龄、努力程度差异对健康差距的相对贡献度分别为 1.438%~6.682%、0.759%~3.612%、0.376%~4.649%、-0.241%~6.116%、11.787%~24.639%、0.066%~0.224%、1.549%~3.184%、27.054%~64.627%、16.002%~25.803%,健康机会不平等对健康差距的相对贡献度为 19.371%~47.144%。如果采用 Fajardo-Gonzalez(2016)的处理方式,健康机会不平等对健康差距的相对贡献度为 35.691%。由两组分解结果可知,性别与 3 岁时所处的经济区域仍

然是导致健康机会不平等的主要因素。

表 8 基于广义 Ordered Logit 模型的夏普利值分解

变量	$P(H_i>1)$		$P(H_i>2)$		$P(H_i>3)$		$P(H_i>4)$	
	绝对值	相对值	绝对值	相对值	绝对值	相对值	绝对值	相对值
<i>mid</i> ₀₃	0.0012	2.457%	0.0019	2.413%	0.0034	2.500%	0.0034	1.696%
<i>east</i> ₀₃	0.0020	4.225%	0.0012	1.537%	0.0030	2.184%	-0.0005	-0.258%
<i>hukou</i> ₀₃	0.0016	3.304%	0.0006	0.759%	0.0012	0.904%	0.0073	3.612%
<i>fedu</i>	0.0009	1.986%	0.0001	0.097%	0.0005	0.392%	0.0008	0.379%
<i>medu</i>	0.0013	2.662%	0.0015	1.957%	0.0000	0.011%	0.0000	-0.003%
<i>urban</i>	0.0029	6.116%	0.0012	1.506%	-0.0002	-0.120%	-0.0005	-0.241%
<i>gender</i>	0.0110	23.143%	0.0192	24.639%	0.0161	11.787%	0.0316	15.714%
<i>minzu</i>	0.0000	0.066%	0.0002	0.224%	0.0002	0.165%	0.0004	0.183%
<i>food</i> ₂	0.0002	0.418%	0.0019	2.465%	0.0020	1.488%	0.0055	2.746%
<i>food</i> ₃	0.0013	2.766%	0.0001	0.094%	0.0001	0.062%	0.0000	-0.006%
<i>age</i>	0.0129	27.054%	0.0301	38.532%	0.0883	64.627%	0.1068	53.033%
<i>edu_res_hat</i>	0.0053	11.150%	0.0065	8.333%	0.0036	2.645%	0.0038	1.886%
<i>status_res_hat</i>	0.0070	14.653%	0.0136	17.444%	0.0182	13.357%	0.0428	21.260%

在稳健性检验中,基于面板 Logit 模型与广义 Ordered Logit 模型得到的健康机会不平等对健康差距的相对贡献值依次为 18.120%~34.525%、19.371%~47.144%,在上文基于面板 Ordered Logit 模型的实证研究中,该取值范围为 18.316%~23.182%,三者取值范围上存在交集。与此同时,在上述测度中,性别与 3 岁时所处的经济区域均是导致健康机会不平等的主要因素。由此可见,即使调整健康决定方程的估计模型,健康机会不平等对健康差距的相对贡献度仍存在交集,且在识别健康机会不平等的主要成因上完全一致。综上,测度结果具有较好的稳健性。

六、结论与政策建议

居民健康并不是一个单纯的生物学事件,其中掺杂了许多社会经济因素。在导致健康不平等的因素中,既有遵从自然法则、合乎正义的因素,也存在有违公平的因素,存在机会不平等的现象。本文的主要工作在于测度出健康不平等中的机会不平等,分解出各个条件变量对机会不平等的相对贡献度。首先,根据自评健康水平变量的性质构建健康决定方程,再以 Jusot 等(2013)、Juárez 和 Soloaga(2014)的测度方法为基础,给出健康机会不平等的测度模型;其次,以 CFPS(2010—2016 年)的数据为样本估计面板 Ordered Logit 随机效应模型,再以估计结果为基础,通过夏普利值分解量化健康机会不平等对健康差距的相对贡献度,识别出导致健康机会不平等的主要因素。

基于上述分析,得到以下研究结论:(1)在不健康、一般、比较健康、很健康、非常健康 5 个健康水平中,处于不健康状态的概率相对较小,平均概率为 20.89%,且外生条件变量作用相对较小,基于 $P(H_i>1)$ 得到的基尼系数仅为 0.0731;处于非常健康状态的概率亦比较低,平均概率为 15.98%,但外生条件变量导致的差异较大,基于 $P(H_i>4)$ 得到的基尼系数为 0.2815,约为前者的 4 倍。(2)年龄、性别、努力程度、3 岁时所在的经济区域是导致健康差距的主要因素,以分解结果为基础,将外生条件变量导致的健康差距加总,可得健康机会不平

等对健康差距的相对贡献度为 18.316%~23.182%。如果采用 Fajardo-Gonzalez (2016) 的测度方式,健康机会不平等对健康差距的相对贡献度为 20.715%。在外生变量中,性别、3 岁时所处的经济区域是导致健康机会不平等的主要因素。(3) 考虑到测度结果依赖于模型的设定,因而采用面板 Logit 模型、广义 Ordered Logit 模型验证测度结果的稳健性。稳健性检验结果表明,三类模型的测度结果在取值范围与主要因素识别上具有一致性,实证研究结论具有较高的稳健性。

基于上述结论,本文认为可从以下方面抑制健康不平等中的机会不平等:

首先,通过政府与社会的力量,弱化早期家庭环境对婴幼儿健康的不利影响。毋庸置疑,儿童时期健康状况会对个人健康产生长期影响。一方面应提升优生优育政策力度,实现新生儿疾病筛查、基因检测、遗传性疾病筛查等项目全覆盖;另一方面应从营养健康上给予经济欠发达地区、农村地区儿童更多的关爱,尤其是留守儿童。

其次,给予女性健康更多关注,提升女性保健水平。切实提升“两癌”筛查的覆盖面,尤其是对于经济欠发达地区和农村地区,力争实现全覆盖。医疗保健政策适度向女性倾斜,如将 HPV 疫苗纳入免费接种计划或者医保范围,降低接种费用。

第三,加强健康素养教育与健康监控,居民健康与良好的生活习惯密不可分,而良好的生活习惯又与健康知识紧密相关,健康知识普及有利于降低外生条件变量对健康素养的不利影响。因而有必要构建多层次、多主体、有针对、广覆盖的国民健康教育体系,积极建设和规范各类广播电视等健康栏目,利用新媒体拓展健康教育。加强体医融合和非医疗健康干预,开展国民体质测试,完善体质健康监测体系,制定实施青少年、妇女、老年人、职业群体及残疾人等特殊群体的健康干预计划。

第四,弱化医疗资源分配与社会地位的关联性,促进医疗服务在各个收入群体、地区之间的均等化。相比于经济发达地区、高收入群体,边远地区与低收入群体看病难、看病贵的问题更为突出,小病拖成大病、因病致贫等现象屡见不鲜。医疗服务均等化,尤其是高质量的医疗服务均等化对于边远地区、低收入群体的意义更为重大。

参考文献:

1. 陈东、张郁杨,2015:《与收入相关的健康不平等的动态变化与分解——以我国中老年群体为例》,《金融研究》第 12 期。
2. 黄潇,2012:《与收入相关的健康不平等扩大了吗》,《统计研究》第 6 期。
3. 刘波、王修华、彭建刚,2015:《我国居民收入差距中的机会不平等——基于 CGSS 数据的实证研究》,《上海经济研究》第 8 期。
4. 李红文、毛新志,2015:《论健康公平》,《伦理学研究》第 2 期。
5. 李莹、吕光明,2016:《机会不平等在多大程度上引致了我国城镇收入不平等》,《统计研究》第 8 期。
6. 鲁万波、于翠婷、高宇璇,2018:《中老年人健康机会不平等的城乡分解》,《财经科学》第 3 期。
7. 马超、曲兆鹏、宋泽,2018:《城乡医保统筹背景下流动人口医疗保健的机会不平等——事前补偿原则与事后补偿原则的悖论》,《中国工业经济》第 2 期。
8. 史军、赵海燕,2010:《公平与健康:罗尔斯正义原则的健康伦理意蕴》,《自然辩证法研究》第 9 期。
9. 王甫勤,2012:《社会经济地位、生活方式与健康不平等》,《社会》第 2 期。
10. 解垚,2009:《与收入相关的健康及医疗服务利用不平等研究》,《经济研究》第 2 期。
11. 赵广川、顾海、郭俊峰,2016:《社会经济地位变迁与医疗服务利用不平等:2000—2011》,《公共管理学报》第 2 期。

- 12.赵广川,2017:《国民健康不平等及其内在影响机制、演变过程》,《世界经济文汇》第5期。
- 13.Aaberge, R., M. Mogstad, and V. Peragine. 2010. “Measuring Long-term Inequality of Opportunity.” *Journal of Public Economics* 95(3): 193–204.
- 14.Almås, I., A. W. Cappelen, J. T. Lind, E. Ø. Sørensen, and B. Tungodden. 2011. “Measuring Unfair (In)equality.” *Journal of Public Economics* 95(7–8): 488–499.
- 15.Arneson, R. J. 1989. “Equality and Equal Opportunity for Welfare.” *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 56(1): 77–93.
- 16.Björklund, A., M. Jäntti, and J. E. Roemer. 2012. “Equality of Opportunity and the Distribution of Long-run Income in Sweden.” *Social Choice and Welfare* 39(2–3): 675–696.
- 17.Bourguignon, F., F. H. Ferreira, and M. Menéndez. 2007. “Inequality of Opportunity in Brazil.” *Review of Income and Wealth* 53(4): 585–618.
- 18.Braveman, P., and S. Gruskin. 2003. “Defining Equity in Health.” *Journal of Epidemiology & Community Health* 57(4): 254–258.
- 19.Carrieri, V., and A. M. Jones. 2018. “Inequality of Opportunity in Health: A Decomposition-based Approach.” *Health Economics* 27(12): 1981–1995.
- 20.Cohen, G. A. 1989. “On the Currency of Egalitarian Justice.” *Ethics* 99(4): 906–944.
- 21.Culyer, A. J., and A. Wagstaff. 1993. “Equity and Equality in Health and Health Care.” *Journal of Health Economics* 12(4): 431–457.
- 22.Daniels, N. 1985. *Just Health Care*. Cambridge, MA: Cambridge University Press.
- 23.Davillas, A., and A. M. Jones. 2018. “Ex-ante Inequality of Opportunity in Health, Decomposition and Distributional Analysis of Biomarkers.” ISER Working Paper Series, No. 2018–13.
- 24.Devooght, K. 2008. “To Each the Same and to Each His Own: A Proposal to Measure Responsibility-Sensitive Income Inequality.” *Economica* 75(298): 280–295.
- 25.Donni, P. L., V. Peragine, and G. Pignataro. 2014. “Ex-ante and Ex-post Measurement of Equality of Opportunity in Health: A Normative Decomposition.” *Health Economics* 23(2): 182–198.
- 26.Dworkin, R. 1981a. “What Is Equality? Part 1: Equality of Welfare.” *Philosophy & Public Affairs* 10(3): 185–246.
- 27.Dworkin, R. 1981b. “What Is Equality? Part 2: Equality of Resources.” *Philosophy & Public Affairs* 10(4): 283–345.
- 28.Erreygers, G., and T. Van Ourti. 2011. “Measuring Socioeconomic Inequality in Health, Health Care and Health Financing by Means of Rank-dependent Indices: A Recipe for Good Practice.” *Journal of Health Economics* 30(4): 685–694.
- 29.Fajardo-Gonzalez, J. 2016. “Inequality of Opportunity in Adult Health in Colombia.” *The Journal of Economic Inequality* 14(4): 395–416.
- 30.Ferreira, F. H. G., and J. Gignoux. 2011. “The Measurement of Inequality of Opportunity: Theory and Application to Latin America.” *Review of Income and Wealth* 57(4): 622–657.
- 31.Fleurbaey, M., and E. Schokkaert. 2009. “Unfair Inequalities in Health and Health Care.” *Journal of Health Economics* 28(1): 73–90.
- 32.Foster, J. E., and A. A. Shneyerov. 2000. “Path Independent Inequality Measures.” *Journal of Economic Theory* 91(2): 199–222.
- 33.Gallardo, K., L. Varas, and M. Gallardo. 2017. “Inequality of Opportunity in Health: Evidence from Chile.” *Revista de Saude Publica* 51(110): 1 – 10.
- 34.Gourieroux, C., A. Monfort, E. Renault, and A. Trognon. 1987. “Generalised Residuals.” *Journal of Econometrics* 34(1–2): 5–32.
- 35.Jusot, F., S. Tubeuf, and A. Trannoy. 2013. “Circumstances and Efforts: How Important Is Their Correlation for the Measurement of Inequality of Opportunity in Health?” *Health Economics* 22(12): 1470–1495.

36. Juárez, F. W. C., and I. Soloaga. 2014. "Iop: Estimating Ex-ante Inequality of Opportunity." *The Stata Journal* 14(4): 830–846.
37. Jenkins, S. P., and P. Van Kerm. 2006. "Trends in Income Inequality, Pro-poor Income Growth, and Income Mobility." *Oxford Economic Papers* 58(3): 531–548.
38. Lefranc, A., N. Pistoiesi, and A. Trannoy. 2009. "Equality of Opportunity and Luck: Definitions and Testable Conditions, with an Application to Income in France." *Journal of Public Economics* 93(11–12): 1189–1207.
39. Pignataro, G. 2012. "Equality of Opportunity: Policy and Measurement Paradigms." *Journal of Economic Surveys* 26(5): 800–834.
40. Rawls, J. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
41. Roemer, J. E., and J. Silvestre. 1993. "The Proportional Solution for Economies with Both Private and Public Ownership." *Journal of Economic Theory* 59(2): 426–444.
42. Roemer, J. E. 1998. *Equality of Opportunity*. Cambridge, MA: Harvard University Press.
43. Roemer, J. E., and A. Trannoy. 2016. "Equality of Opportunity: Theory and Measurement." *Journal of Economic Literature* 54(4): 1288–1332.
44. Rosa Dias, P. 2009. "Inequality of Opportunity in Health: Evidence from a UK Cohort Study." *Health Economics* 18(9): 1057–1074.
45. Ruger, J. P. 2004. "Health and Social Justice." *The Lancet* 364(9439): 1075–1080.
46. Shorrocks, A. F. 2013. "Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value." *Journal of Economic Inequality* 11(1): 99–126.
47. Wagstaff, A., E. Van Doorslaer, and P. Paci. 1991. "On the Measurement of Horizontal Inequity in the Delivery of Health Care." *Journal of Health Economics* 10(2): 169–205.
48. Williams, R. 2016. "Understanding and Interpreting Generalized Ordered Logit Models." *The Journal of Mathematical Sociology* 40(1): 7–20.
49. Zhao, G., X. Cao, and C. Ma. 2020. "Accounting for Horizontal Inequity in the Delivery of Health Care in China: A Framework for Measurement and Decomposition." *International Review of Economics and Finance* 66: 13–24.

Inequality of Opportunity in the Health Gap of Chinese Residents

Liu Bo¹, Hu Zongyi¹ and Gong Zhimin²

(1: College of Finance and Statistics, Hunan University;

2: Business School, Xiangtan University)

Abstract: Reducing the health gap and improving people's health are the core issues to the government. In the health gap of residents, the reasonable gap and the unreasonable one, i.e. the inequality of opportunity in health, coexist. Based on the CFPS dataset (2010–2016) and the Shapley value decomposition method, this paper modifies the existing methods and constructs a model to measure the inequality of opportunity in the health gap of Chinese residents. The results show that, inequality of opportunity contributes 18.316%~23.182% of the health gap; gender and the economic area in 3 are the main influencing factors among the exogenous variables. Based on the conclusions above, this article puts forward countermeasures to curb inequality of opportunities and reduce health gap from three aspects: original family background, gender and health literacy.

Keywords: Inequality of Health, Inequality of Opportunity, Shapley Value Decomposition

JEL Classification: I10, I14

(责任编辑:陈永清)