

我国居民与收入相关的健康不平等实证研究

刘广彬*

东北财经大学富虹经济学院 辽宁大连 116025

【摘要】本文利用中国健康与营养调查(CHNS)2006年的调查数据,从定量的角度对我国居民与收入相关的健康不平等进行了分析。本文利用有序 Probit 模型获得了自评健康数据背后的实际健康得分,在此基础上计算健康集中指数衡量我国居民与收入相关的健康不平等程度。结果表明,我国居民的健康不平等问题较为严重,健康不平等问题在城乡之间和不同经济发展水平地区之间存在着较大的差别。

【关键词】中国健康与营养调查;健康不平等;健康集中指数

中图分类号:R197 文献标识码:A 文章编号:1674-2982(2008)03-0058-05

An empirical study on the income-related health inequality in China

LIU Guang-bin

Fuhong School of Economics, Northeastern University of Finance and Economics, Liaoning Dalian 116025, China

【Abstract】Using the data of self-assessed health status and income from the China Health and Nutrition Survey (2006), this paper analyzed the income-related health inequality in China. The latent health score was obtained from ordered probit model, and health concentration index was calculated. The study found that the inequality degree in China is much higher compared to other countries, and differentiates between urban and rural residents and among areas with different economic development levels.

【Key words】China health and nutrition survey, Health inequality, Health concentration index

1 导论

健康是一项珍贵的财富,也是人的基本权利,每一个社会成员都应该享有同等的健康权利。但是,无论在国家之间还是在国家内部,无论是发展中国家还是发达国家,也不论其总体健康水平是高还是低,都存在着相当程度的健康不平等问题。所谓健康不平等,是指按照社会经济地位、性别、种族等各种社会因素划分的各类人群之间健康指标的差异。大多数健康不平等不是生物学上不可避免的,而仅仅反映人群生活环境和行为之间的差异。我国也存在着较为严重的健康不平等^[1-4]。在影响居民健康及其不平等的众多因素中,居民收入是十分重要的要素,与收入相关的健康不平等问题是健康不平等研究领域内的热点之一。

但是,相对于国际上对健康不平等问题的研究,我国对居民的健康不平等进行的研究还不够充分,进行定量分析的研究寥寥无几。刘宝和胡善联利用健康自评数据测算了上海市四区(县)的健康集中指

数,考察了与收入相关的健康不平等问题。其数据来源是 2001 年在上海市四区(县)进行的家庭卫生服务调查,采用 Hopit 模型获得人群间可比性的伤残得分,在此基础上计算出健康集中指数来衡量上海市四区(县)的收入—健康不平等^[5-6]。胡琳琳利用第三次国家卫生服务总调查中的自评健康数据和收入数据,计算了我国及各样本县区的健康集中指数,并将计算结果与 Van Doorslaer 等的欧洲国家和美国的健康集中指数进行了比较,发现我国收入相关健康不平等程度相对较高,健康不平等有利于高收入者^[7-8]。在对我国居民健康不平等问题的考察中,在国外和国内得到广泛应用的中国健康和营养调查数据(CHNS)却并没有得到充分的利用。只有赵忠沿用了 Allison 和 Foster 提出的方法,利用 CHNS2000 年的数据对我国居民的健康不平等进行了研究^[9-10]。但是,Allison 和 Foster 框架的局限性决定了赵忠的分析不能够量化健康不平等的程度,不能够分析与社会经济因素相关的健康不平等,而只能分析纯粹的健康不平等。

* 作者简介:刘广彬,男(1981 年-),主要研究方向为劳动经济学、卫生经济学,E-mail:lecoliu@hotmail.com。

与上述提到的研究中国居民健康不平等问题的文献相比,本文利用中国健康与营养调查(CHNS)最新的 2006 年的调查数据对我国居民的与收入相关的健康不平等问题进行定量分析。我们利用 CHNS2006 的自评健康数据,采用 Probit 模型对其进行赋值,并在此基础上计算我国居民的健康集中指数,从数据利用和研究方法两方面进一步丰富对我国居民健康不平等现象的研究。本文的结构如下:第一部分导论,介绍本文的研究背景和任务;第二部分讨论健康自评数据及其赋值;第三部分介绍本文数据及相关变量,并以健康自评数据为基础研究了我国居民的健康分布;第四部分介绍健康集中指数及其计算结果;第五部分列示了主要结论和本文需要进一步研究的问题。

2 健康自评数据及其赋值

本文所使用的健康数据为健康自评数据。健康自评数据具有众多的优点。首先,它允许个人自主决定关于健康多维因素的权重,避免了诸如健康调整生命年、质量调整生命年等综合指标需要主观决定权重的难题。其次,大量的文献证明,健康自评数据与死亡率等客观健康测量指标高度相关,从而能够有效地测量健康。比如, Mossey 和 Shapiro^[11]、Idler 和 Angel^[12]、Idler 和 Kasl^[13]、Marmot 等^[14] 文献证明健康自评数据能够有效预测死亡率、功能丧失、因病缺勤等客观健康测量指标。并且,自评健康数据比较容易获得,在众多的调查中均有数据,从而能够获得足够的样本数进行研究。

但是,健康自评数据的使用产生了如何测量健康不平等的问题。健康集中指数是测量健康不平等最常用的指标,健康集中指数以及相应的健康集中曲线要求健康数据是连续变量或者是二分类变量^[15]。因此,只有对序数形式的健康自评数据赋予基数形式的数值,才能够计算健康集中指数以测量健康不平等。当前,有四种不同的赋值方法^[16]。第一种方法是采用两分法将健康数据分为健康或者不健康。第二种方法是通过一定的计算方法构造指数进行赋值,根据调查获得的不同健康方面的得分获得一个总体健康指数。第三个方法是在某种假设条件下对健康自评数据进行赋值。Wagstaff 和 Van Doorslaer 假设在分类健康自评数据的背后是一个连续的、符合对数正态分布但是不能够被观测到的潜变量^[15]。将标准对数正态区间按照五个等级占样本

总数的比例进行划分,查正态分布表并进行指数换算,从而得到五个等级的实际得分,进而计算健康集中指数。胡琳琳所使用的正是这种方法^[7]。与此类似的第四种方法是将健康自评数据作为因变量,通过有序 Probit 模型对潜变量进行赋值,来计算集中指数^[17]。此外,在 Probit 模型的基础上加入情境变量(vignette)便会变成 Hopit 模型。本文使用有序 Probit 模型对潜在的健康数据进行赋值。

假设健康自评数据有 m 个取值,令 H_i 表示样本中第 i 个人的健康自评数据取值,则有序 Probit 模型可以表示为:

$$H_i = j \quad \text{当 } \mu_{j-1} \leq H_i^* \leq \mu_j, j = 1, L, m \quad (1)$$

其中, H^* 为潜变量。假设潜变量 H^* 是个人社会经济特征向量 x 的函数:

$$H_i^* = x_i \beta' + \varepsilon_i, \varepsilon_i: N(0, 1) \quad (2)$$

并且, $\mu_0 = -\infty, \mu_j \leq \mu_{j+1}, \mu_m = +\infty$ 。

因此,健康自评数据取 1 到 m 中任意一值的概率为:

$$P_{ij} = Pr(H_i = j) = \Phi(\mu_j - x_i \beta') - \Phi(\mu_{j-1} - x_i \beta') \quad (3)$$

其中, $\Phi(\cdot)$ 为标准正态分布的累积分布函数。

假设样本个体之间是相互独立的,则有序 Probit 模型的对数似然函数为

$$\ln L = \sum_i \sum_j H_{ij} \ln P_{ij} \quad (4)$$

其中, H_{ij} 是一个二分变量,如果 $H_i = j$, 则 $H_{ij} = 1$, 否则 $H_{ij} = 0$ 。通过最大似然估计方法我们就可以得到系数 β 和未知的分界点 μ_j 。估计值 $x_i \beta'$ 可以用来测量个体的健康水平,并通过下式的重新赋值将具体数据局限于 $(0, 1)$ 区间内^[18]:

$$H^2 = \frac{H^1 - H^{\min}}{H^{\max} - H^{\min}} \quad (5)$$

其中, H^1 表示由有序 Probit 模型计算出来的健康得分, H^{\min} 和 H^{\max} 分别为最小值和最大值, H^2 表示修正以后的健康得分。通过这种修正,从有序 Probit 模型获得的健康预测值将会重新赋值在 $(0, 1)$ 区间内。

3 数据与变量设定

本文所使用的健康数据和收入数据来源于 2006 年的 CHNS 横截面数据^①, 健康数据的调查形式为:与同龄人相比,你觉得自己的健康状况怎么样? 答案为非常好、好、一般和差, 分别用 1、2、3、4 表示。

① 2008 年 7 月 CHNS 提供了经过整理和更新的面板数据, 本文所使用的数据即从面板数据库中整理得来。

CHNS 是由美国北卡罗来纳大学和中国疾病预防控制中心联合开展的调查。该调查遵从多水平、随机群体样本的程序,收集了有关人口健康、医疗保健、保险、营养、计划生育、家庭收入、社会服务和社会人口学特征等方面的信息。本文所使用的样本是 CHNS2006 年的成人(18 岁及以上成人)调查数据,调查样本 11 202 个。其中,男性 5 485 人,占 48.96%,女性 5 717 人,占 51.04%;城镇人口 4 743 人,占 42.34%,农村人口 6 459 人,占 57.66%。统计结果表明,我国居民的健康水平在男女性别之间和在城乡之间都存在着明显的差异。在男性样本中,对自身健康评价为非常好、好、一般和差的比例分别为 15.17%、50.05%、29.17% 和 5.62%;女性样本的比例分别为 11.21%、47.87%、33.69% 和 7.22%。在城镇人口样本中,对自身健康评价为非常好、好、一般和差的比例分别为 14.06%、49.04%、31.06% 和 5.84%;农村人口样本的比例分别为 12.48%、48.86%、31.79% 和 6.87%。

表 1 CHNS(2006)健康自评数据分布情况表(%)

样本	非常好	好	一般	差
全国	13.15	48.94	31.48	6.44
个人收入组 1	9.66	44.21	36.47	9.66
个人收入组 2	12.02	49.91	31.28	6.79
个人收入组 3	13.15	50.55	30.46	5.84
个人收入组 4	17.87	50.89	27.77	3.46

注:资料来源于中国健康与营养调查(CHNS)。

表格 1 列示了健康自评数据在样本人群中的分布情况。可以看到,样本中被调查者对自己健康状态的评价为“好”的比例最大,“一般”、“非常好”和“差”的比例依次减少,表明健康自评数据基本符合对数正态分布。其中,个人收入变量指 CHNS 统计的个人年收入,以收入分布的四分位为参考标准将收入从而将样本划分为四组。各样本组之间的健康水平和健康不平等的定性分析,可以依照 Allison 和 Foster 提出的方法^[10]进行计算。由表格 1 我们可以看到,随着收入水平的提高,对自身健康评价为“非常好”和“好”的比例明显地呈现出上升趋势,收入组 3、4 中“非常好”和“好”的比例都超过了全国平均水平。相应的,对自身健康评价为“一般”和“差”的比例随着收入的提高呈现出下降趋势。比例数据本身已经表明,健康水平随着收入水平的提高而提高,不同收入组之间的健康水平差异较大。

4 健康集中指数及其计算结果

我们利用健康集中指数^[19]作为衡量与收入相关

的健康不平等的指标。因为它符合三个基本要求:(1)能够反映出健康不平等的社会经济特征;(2)能够反映全体人群的情况;(3)它对于人口社会经济分布的变化十分敏感^[20]。并且,Bommier 和 Stecklov 认为如果健康公平被定义为“如果一个社会中的每个人的健康获得,不是由社会经济状况或收入来决定,则在这个理想的公平社会中的健康分布就是健康公平^[21]”,相对于从社会福利函数所推导出来的不平等指数,健康集中指数更为合适。近年来,健康集中指数被广泛应用于儿童死亡率、儿童免疫、儿童营养不良、成人健康、健康补助以及医疗服务利用等各个方面^[16]。

健康集中指数 C 的概念与健康集中曲线紧密相关,表示为集中曲线与对角线之间面积的两倍。假设连续变量 h_i 代表人群中每个人的健康得分,则健康集中曲线 $L_{(s)}$ 表示相对于人口累计百分比(按收入排序,从最低收入开始)的健康得分累计百分比。如果 $L_{(s)}$ 与对角线重合,则每个人都享有平等的健康。如果 $L_{(s)}$ 在对角线之上,则存在健康不平等,并且富裕人群健康状况好,低收入人群健康状况差;反之,如果 $L_{(s)}$ 在对角线之下,则表明低收入人群的健康状况较好。与对角线的距离越远,不平等性越高。健康集中指数 C,定义为 $L_{(s)}$ 与对角线之间面积的 2 倍。当 $L_{(s)}$ 与对角线重合时 C 值为 0,当在对角线之上(下)时 C 为负(正)。C 取值在 -1 和 1 之间,当 $C = -1$ 时,表示人群中的所有伤病都发生在收入最低者上;当 $C = 1$ 时,表示人群中的所有伤病都发生在高收入阶层。

由于我们使用的是个人数据,健康集中指数的计算公式表示如下^[22]:

$$C = \frac{2}{n \cdot h} \sum_{i=1}^n h_i R_i - 1, \quad (6)$$

其中, h_i 代表人群中每个人的健康得分, $h = \frac{1}{n} \sum_{i=1}^n h_i$ 表示样本的平均健康水平, R_i 表示第 i 个人收入的相对次序(relative rank)。 $i = 1$ 表示收入最低的人, $i = n$ 表示收入最高的人。

Kakwani 等在 1997 年指出,对于个人水平数据(individual-level data),获得健康集中指数最简捷的方式是对以下计量模型进行回归:

$$\sigma_R^2 [h_i/h] = \alpha + \beta R_i + u_i \quad (7)$$

其中, $2\sigma_R^2$ 为相对次序的方差,参数 β 就是健康集中指数的最小二乘估计值。由于公式(2)已经考虑了收入变量以外的社会经济变量,因此通过公式(7)所获得的健康集中指数是经过标准化以后的健

康集中指数,无须考虑人口结构以及其他干扰因素对健康集中指数计算的影响。

表格 2 列示了本文根据 CHNS 计算的全国样本以及分组样本的健康集中指数,同时根据 95% 置信区间计算出健康集中指数的上下限。在计算全国样本集中指数的同时,本文也同时计算了分城乡以及性别的健康集中指数,并且我们根据 2006 年的人均国内生产总值水平将 9 个样本省区划分为三个不同水平的经济发展地区,以更好的展示不同经济发展地区的健康不平等问题。其中,贵州省 2006 年人均 GDP 低于 1 万元被划为第一组(2006 人均 GDP 5 787 元),代表经济发展较为落后的地区;黑龙江(16 195)、河南省(13 313 元)、湖北省(13 296 元)、湖南省(11 950 元)、广西壮族自治区(10 296 元)人均 GDP 在 1 万元和 2 万元之间的省区被划分为第二组,为经济发展水平一般的地区;辽宁省(21 788 元)、山东省(23 794 元)和江苏省(28 814 元)人均 GDP 在 2 万以上被划为第三组,且经济发展水平较高的地区。

表 2 健康集中指数的计算结果

样本	健康集中指数	下限	上限	P 值
全国	-0.07468	-0.07816	-0.07120	<0.0001
城镇	-0.04769	-0.05404	-0.04135	<0.0001
农村	-0.09976	-0.10431	-0.09520	<0.0001
男性	-0.08078	-0.08614	-0.07542	<0.0001
女性	-0.07897	-0.08501	-0.07292	<0.0001
经济发展水平组 1	-0.07984	-0.09083	-0.06885	<0.0001
经济发展水平组 2	-0.05210	-0.05665	-0.04756	<0.0001
经济发展水平组 3	-0.08711	-0.09380	-0.08043	<0.0001

注:健康集中指数计算数据来源于中国健康与营养调查(CHNS),上下限取 95% 置信区间。

5 主要结论和需要进一步研究的问题

健康集中指数的符号和大小表示了健康变量和居民收入变量之间的相关关系的方向和强度。在健康集中指数为负的情况下,健康集中指数更为直观的意义是,将其与 75 相乘(取其绝对值),所得到的百分数表示需要从富的那一半人群中转移到穷的那一半人群中的健康再分配比例,以实现健康集中指数为零的健康平等^[23]。

由计算结果可以看到,我国居民的健康集中指数为 -0.07468。由于健康集中指数为负,所以可以断定我国当前的健康不平等不利于低收入人群,与高收入人群相比,低收入人群的健康状况更低。与

发达国家的健康集中指数相比较^[20],发现我国居民的健康不平等情况比较严重,健康集中指数小于欧洲大陆国家,但是大于英国和美国。而在我国,农村健康集中指数与城镇健康集中指数相比较,前者的绝对值是后者绝对值的两倍多,这表明相对于城镇而言,农村的健康不平等情况更为严重。与此形成对比的是,健康不平等水平在性别之间并没有太多的差异。这些结论与胡琳琳的计算结果和结论^[7]较为类似,但是与刘宝和胡善联的计算结果^[5]有细微的差别。

除此以外,从表 2 最后三行可以看出,经济发展水平与健康不平等之间没有确切的规律可循。三组经济发展水平地区分组中,介于中间的黑龙江、河南、湖北等省区的健康不平等程度最低,在样本省区中经济发展水平最高的辽宁、山东和江苏三省的健康不平等程度反而最为严重。因此,人均收入水平与健康不平等之间可能并没有紧密的联系。另一个引人注意的是收入不平等对健康不平等的影响,由公式(6)我们可以看到,本文计算的健康集中指数是一种“依赖于序列的指标”(rank-dependent indicator)^[24],即与收入相关的健康集中指数的大小只依赖于收入的排序,而不是收入差距的大小。因此本文计算的健康集中指数,不能够提供有用的信息使我们更好地理解收入不平等与健康不平等之间的关系。

与国外健康不平等研究的最新进展相比,本文仍然存在着诸多的不足和需要进一步研究的问题。近年来,传统的健康集中指数的局限性受到了越来越多的注意。Clarke 等人^[25]在 2002 年指出,以健康为衡量指标或者以发病率等疾病—健康指标来计算健康集中指数会获得不同的结果;Wagstaff^[26]在 2005 年指出,健康集中指数依赖于健康变量的均值,从而不能够比较不同健康均值的人群之间的比较;Erreygers^[27]则认为当健康变量为定性数据时健康集中指数的计算结果存在着不确定性。鉴于这些局限,Erreygers 对传统的健康集中指数进行了修正^[27]。但是由于技术原因,本文并没有利用这些新的进展。本文的第二个不足是,缺乏对健康不平等原因的进一步研究,也没有利用 Wagstaff 等提出的方法对健康集中指数进行分解^[28],因此还不清楚我国居民与收入相关的健康不平等的根源,以及健康不平等在城乡和地域之间存在较大差异的原因。第三个有待于改进的地方,是本文所利用的收入数据没有充分考虑收入水平、价格指数等方面的问题,和各地物价水平、生活成本的差异,使用未经处理的名义收入值有

可能导致较大的偏差。

参 考 文 献

- [1] 闫菊娥, 张静. 我国城乡居民健康公平性分析[J]. 卫生软科学, 1998, 12(6): 8-10.
- [2] 王丽敏, 张晓波, David Coady. 健康不平等及其成因——中国全国儿童健康调查实证研究[J]. 经济学季刊, 2(2): 435-452.
- [3] 卫生部统计信息中心. 中国卫生服务调查研究——第三次国家卫生服务调查分析报告[M]. 北京: 中国协和医科大学出版社, 2004.
- [4] 杨红燕. 我国城乡居民健康公平性研究[J]. 财经科学, 2007, (3): 69-75.
- [5] 刘宝, 胡善联. 收入相关健康不平等实证研究[J]. 卫生经济研究, 2003, (1): 14-16.
- [6] 刘宝. 人群自报健康水平与分布研究[M]. 上海: 复旦大学出版社, 2007.
- [7] 胡琳琳. 我国与收入相关的健康不平等实证研究[J]. 卫生经济研究, 2005, (12): 13-16.
- [8] Van Doorslaer E, Wagstaff A, Bleichrodt H. Income - related inequalities in health : some international comparisons [J]. Journal of Health Economics, 1997, (16): 93-112.
- [9] 赵忠. 使用自评健康数据度量我国健康的不平等[G]. 第五届中国经济学会年会论文, 2005.
- [10] Allison R, Foster J. Measuring Health Inequality Using Qualitative Data[J]. Journal of Health Economics, 2004, 23(3): 505-524.
- [11] Mossey J M, Shapiro E. Self-rated health: a predictor of mortality among the elderly[J]. Am J Public Health, 1982, 72(8): 800-808.
- [12] Idler E L, Angel R J. Self-Rated Health and Mortality in the NHANES-I Epidemiologic Follow-up Study[J]. Journal of American Public Health, 1990, 80(4): 446-452.
- [13] Idler E L, Kasl S V. Self-Ratings of Health: Do They Also Predict Change in Functional Ability? [J]. Journal of Gerontology, 1995, 50(6): 344-353.
- [14] Marmot M, Feeney A, Shipley M, et al. Sickness Absence as a Measure of Health Status and Functioning: From the UK Whitehall II Study [J]. Journal of Epidemiology and Community Health, 1995, 49(2): 124-130.
- [15] Wagstaff A, Van Doorslaer E. Measuring inequalities in health in the presence of multiple - category morbidity indicators[J]. Health Economics, 1994, 3(4): 281-291.
- [16] O'Donnell O, Doorslaer E V, Wagstaff A, Lindelow M. Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and Their Implementation [M]. The World Bank, 2007.
- [17] Groot W. Adaptation and scale of reference bias in self-assessments of quality of life[J]. Journal of Health Economics, 2000, 19(3): 403-420.
- [18] Van Doorslaer E, Jones A M. Inequalities in self-reported health: validation of a new approach to measurement [J]. Journal of Health Economics, 2003, 22(1): 61 -87.
- [19] Wagstaff A, Van Doorslaer E, Paci P. Equity in the finance and delivery of health care: some tentative cross-country comparisons [J]. Oxford Review of Economic Policy, 1989, 5(1): 89-112.
- [20] Van Doorslaer E, Wagstaff. Income-related inequalities in health: some international comparisons [J]. Journal of Health Economics, 1997, 16(1): 93-112.
- [21] Bommier A, Stecklov G. Defining health inequality: why Rawls succeeds where social welfare theory fails[J]. Journal of Health Economics, 2002, 21(3): 497-513.
- [22] Kakwani N, Wagstaff A, Van Doorslaer E. Socioeconomic inequalities in health: measurement, computation, and statistical inference [J]. Journal of Econometrics, 1997, 77(1): 87-103.
- [23] Koolman X, Doorslaer E V. On the Interpretation of a Concentration Index of Inequality [J]. Health Economics, 2004, 13(7): 649-656.
- [24] Erreygars G. Beyond the health Concentration Index: an Atkinson alternative for the measurement of the socioeconomic in-equality of health [R]. Helsinki: Advancing Health Equity, 2006.
- [25] Clarke P M, Gerdtham U G, Johannesson M, et al. On the measurement of relative and absolute income-related health inequality. Social Science & Medicine, 2002, 55(11): 1923-1928.
- [26] Wagstaff A. The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality [J]. Health Economics, 2005, 14(4): 429-432.
- [27] Erreygars G. Correcting the Concentration Index [R]. Antwerp, Belgium: University of Antwerp, 2006.
- [28] Wagstaff A, Paci P, Joshi H. Causes of Inequality in Health: Who Are You? Where Do You Live? Or Who Your Parents Were? [R]. World Bank, 2001.

[收稿日期:2008-10-21 修回日期:2008-11-12]

(编辑 田晓晓)