

普通话能力对流动人口健康的影响： 来自 CGSS 的经验证据

刘国辉 单宝刚 张卫国

摘要：流动人口健康是一个重要的社会和公共卫生问题。利用中国综合社会调查(CGSS 2013)中的农民工样本数据,考察普通话能力对流动人口健康的影响及其群体间差异。研究发现,普通话能力与农民工健康呈正相关关系,即较好的普通话听辨和表达能力可以提高农民工身体健康的概率。同时,普通话能力的健康效应在农民工样本中存在着明显的异质性。在利用工具变量法克服了可能存在的内生性之后,上述结论依然稳健。发现和理解语言的健康效应,对于增强对语言经济社会效应的认识,对于语言扶贫、健康扶贫以及促进流动人口社会融入等都具有重要的意义。

关键词：普通话能力;健康效应;农民工;群体差异

DOI: 10.19836/j.cnki.37-1100/c.2020.03.014

一、问题的提出

健康是人类安身立命之本,影响着劳动者的劳动力市场表现,对于城市外来流动人口特别是农民工而言,尤为重要。但是现实生活中,农民工的健康状况并不乐观。在当前提倡“全民健康”社会背景下,农民工健康问题愈发突出,不仅涉及公共卫生,俨然成为一个亟待解决的社会问题。

关于农民工健康,大量文献围绕其影响因素而展开,取得了不少成果。综合来看,相关的讨论涉及了教育、社会资本、工作时长、工作类型、居住条件、工作环境、家庭交流、社会交往等诸多方面。如,黄乾利用上海、天津等五个城市的调查数据,发现教育及社会资本会对农民工健康产生积极影响^①。李珍珍、陈琳利用江苏和浙江农民工的调查数据,发现性别、教育、工作时长、工作类型、家庭负担等会对农民工健康产生影响^②。此外,也有研究表明,居住条件、工作环境、家庭交流、社会交往等同样能够影响农民工的健康状况^③。

但事实上,农民健康的影响因素是复杂且多元化的。近年来,国外针对移民的研究发现,流动人口的健康状况与其个人语言能力也密切相关。Mui等曾利用美国纽约市中韩两国老年移民的样本讨论了英语语言能力与个人健康之间的关系。实证结果显示,两组亚裔移民的健康水平因个体的语言能力不同而存在差异。不仅如此,两组移民的健康水平普遍低于美国人的平均水平,且语言能力越

收稿日期：2019-07-20

基金项目：国家语委重大科研项目“中国语言扶贫经验成效及相关理论问题研究”(ZDA135-9);教育部人文社科基金青年项目“语言经济学视域下当代东北亚语言政策及其战略变革研究”(19YJC740034)。

作者简介：刘国辉,延边大学外国语学院副教授,经济学博士(延吉 133002; guohuilu2014@126.com);单宝刚,民生银行地产金融事业部济南分部(济南 250002; shanbaogang@foxmail.com);张卫国(通讯作者),山东大学经济研究院教授,博士生导师(济南 250100; sdzuzw@126.com)。

① 黄乾:《教育与社会资本对城市农民工健康的影响研究》,《人口与经济》2010第2期。

② 李珍珍、陈琳:《农民工健康状况影响因素分析》,《南方人口》2010第4期。

③ 俞林伟:《居住条件、工作环境对新生代农民工健康的影响》,《浙江社会科学》2016第5期。

差,健康水平越低^①。Pottie 等则研究了加拿大移民的语言能力与其自评健康之间的关系。研究发现,在控制性别、年龄、医疗保健、工作满意度等协变量后,英语或法语(均为加拿大官方语言)水平与移民健康状况显著正相关,且在女性移民中关系更为明显^②。Ng 等对 7000 余名加拿大移民的健康情况进行了跟踪调查研究,分析结果显示,随着时间的流逝移民的自评健康患病率明显高于加拿大本土居民,而移民健康水平的降低与他们对英语或法语的掌握程度较低有着显著联系^③。这表明,较好地掌握目的国的语言能力有利于保持个人健康水平。Schachter 等对美国部分亚裔和拉丁裔移民的语言能力(包括母语和英语)与健康水平进行了对比研究,发现和只掌握母语的移民相比,熟练掌握母语与英语的双语移民其健康水平具有明显优势。他们认为,双语移民健康水平较高的原因可能在于他们熟练掌握了目的国语言而获得了更高的社会经济地位和家庭关系的支持^④。

总之,作为一个较为新颖的切入点,从语言能力的角度来研究思考流动人口健康正日益得到国外研究人员的关注。国内方面,相关研究往往是把语言和健康作为两个独立事物,对它们在农民工城市融入中所扮演的角色和作用进行分别讨论,关于语言能力与健康水平关系的研究还十分少见。

基于 2013 年中国综合社会调查的微观数据,研究了普通话能力对流动人口(农民工)健康的影响及其群体差异。研究发现,普通话能力与农民工健康水平存在着显著的正相关关系,同时二者关系存在着异质性。在克服了可能存在的内生性问题后,上述结论依然稳健。本文可能的贡献在于:第一,通过识别农民工的普通话能力与其健康水平之间的关系,增强政府和相关部门对语言能力经济社会效应的认识,为解决农民工健康问题提供较为新颖的切入点;第二,研究考察农民工群体的普通话能力和健康水平关系的异质性,可以更加精准地维护农民工的健康权益,有助于实现健康的人口城市化;第三,从提升农民工语言能力、提高健康水平等角度为推动流动人口的社会融入提供借鉴,有利于解决农民工流动过程中出现的农民工难题,对推动农民工社会融入、建设和谐经济社会也具有重要的现实意义。

二、模型、变量与数据

(一)模型构建与变量选择

以农民工健康作为被解释变量,农民工普通话能力作为核心解释变量,将影响健康的其他因素作为控制变量引入模型,由此建立计量模型对普通话能力与农民工健康的关系进行实证分析。

$$Health_i = \beta_0 + \beta_1 mandarin_i + \lambda X_i + \epsilon_i \quad (1)$$

其中, $Health_i$ 为农民工的身体健康情况, i 表示个体, $mandarin_i$ 代表个体的普通话能力, X_i 是影响农民工健康的一系列控制变量, ϵ_i 为误差项。

被解释变量:农民工健康水平。在 CGSS 2013 问卷中,体现健康水平的问题为“您觉得您目前的身体健康状况是怎样的”,因此借鉴以往研究的通用做法,选取自评健康作为衡量农民工健康的标准。

核心解释变量:农民工普通话能力。具体从听和说(“普通话听辨”[ML]和“普通话表达”[MS])两个方面进行衡量。普通话听辨和表达能够基本反映特定个体普通话的真实水平,良好的普通话能力既需要能听懂、理解对,也需要能说出口、表达正确。在 CGSS 2013 调查中,通过向被调查者询问

① Mui A., Kang S., Kang D., Domanski M., “English Language Proficiency and Health-Related Quality of Life among Chinese and Korean Immigrant Elders”, *Health & Social Work*, 2007, 32(2), pp.119-127.

② Pottie K., Ng E., Spitzer D., Mohammed A., Glazier R., “Language Proficiency, Gender and Self-reported Health”, *Canadian Journal of Public Health*, 2008, 99(6), pp.505-510.

③ Ng E., Pottie K., Spitzer D., “Official Language Proficiency and Self-Reported Health among Immigrants to Canada”, *Health Reports*, 2011, 22(4), pp.15-23.

④ Schachter A., Kimbro R., Gorman B., “Language Proficiency and Health Status: Are Bilingual Immigrants Healthier”, *Journal of Health and Social Behavior*, 2012, 53(1), pp.124-145.

“您觉得自己听普通话的能力是什么水平”和“您觉得自己说普通话的能力是什么水平”,让其根据自己的普通话听说水平进行评级,评分越高表明其普通话水平越高。

控制变量主要包括工作特征变量及个体特征变量等。工作特征变量主要包括工作时长、工作年限、劳动收入等。工作时长是指每周累计工作的时间,工作年限则指截至被访日累计参加工作的年数,工作时间越久、工作年限越长,更容易引发劳累,也加大了职业病及其他伤害发生的可能性,同时累积的工作压力也可能越大。劳动收入是指通过工作等途径获得的全年劳动收入情况,收入的高低能够影响是否有足够的经济基础获得医疗援助及保健等。为了控制个体特征可能对健康水平产生的影响,本文还引入了性别、年龄、婚姻情况、受教育年限等个体特征变量。其中,性别、年龄、婚姻情况是较为常见的个体特征变量,指标上的差异会影响健康水平的差异。受教育年限能够反映被访者的教育程度,而教育程度能够对农民工的健康情况产生影响^①。

需要指出,CGSS 2013 问卷中关于健康的问题选项包括“很不健康”“比较不健康”“一般”“比较健康”“很健康”五个层级,是离散变量且存在明显的由低到高排序。因此,在具体的计量回归上,我们按照经验采用了有序的 Probit 模型来估计普通话能力和健康关系。

(二)数据及描述性统计

数据选取自中国综合社会调查项目课题组公开发布的 2013 年微观调查(CGSS 2013)。在样本主体的选取上,我们参考既有文献的做法,将“农民工”的定义为:十八周岁以上、六十五周岁以下,具有农村户口、在城市从事非农工作的“广义农民工”群体。在对缺失值进行筛选和剔除后,本文累计有效的农民工样本数量为 1613 个。

各变量的描述性统计如表 1 所示。可以发现,对于被解释变量(被访农民工自评健康)的均值为 4.129,直观数据反映多数农民工认为自己处于相对健康的状态。从普通话听辨和表达水平来看,样本中的农民工认为自己的普通话听辨水平优于普通话表达水平,其均值分别为 3.864 和 3.235。整体而言,多数农民工认为自己的普通话能力一般,仍有较大的提升空间。从个体特征变量统计来看,样本中的农民工群体多数为男性,这也与农民工从事的工作性质密切相关。年龄平均值在 39.219 岁,大多数已婚,绝大多数为汉族农民工。受教育年限平均值为 9.18,相当于初中毕业水平,受教育程度较差。

表 1 变量定义及描述性统计

变量	变量说明	均值	标准差	最大值	最小值
被解释变量:					
自评健康	依次按“很不健康”“比较不健康”“一般”“比较健康”“很健康”赋值 1、2、3、4、5	4.129	0.898	5	1
核心解释变量:					
普通话听辨	依次按“完全听不懂”“比较差”“一般”“比较好”“好”赋值 1、2、3、4、5	3.864	0.892	5	1
普通话表达	依次按“完全不能说”“比较差”“一般”“比较好”“好”赋值 1、2、3、4、5	3.235	1.024	5	1
个体基本特征:					
性别	男=1;女=0	0.626	0.483	1	0
年龄	岁	39.219	10.812	65	18

① 黄乾:《教育与社会资本对城市农民工健康的影响研究》,《人口与经济》2010 第 2 期。

续表 1

变量	变量说明	均值	标准差	最大值	最小值
婚姻	未婚=1;其他=0	0.119	0.324	1	0
民族	汉族=1;其他=0	0.921	0.268	1	0
受教育年限	年	9.180	3.237	19	0
工作特征:					
工作时长	每周工作时长(小时)	56.430	17.041	112	2
工作年限	累计工作年数(年)	12.106	9.288	47	1
劳动收入	对“全年总收入”作对数处理	10.013	0.774	11.695	6.684

注:CGSS 2013 问卷关于受教育程度问题的选项为当前最高教育程度(从“未接受过教育”到“研究生及以上”),我们对其分别赋值相应的受教育年限。

三、实证结果与分析

(一)普通话能力对农民工健康的影响

1.基于有序 Probit 模型的基准回归结果。表 2 汇报了基于有序 Probit 模型普通话能力对农民工健康影响的初步结果。我们使用了嵌套模型形式在方程中依次加入个人基本特征和工作特征等控制变量,第 1-3 列为普通话听辨的结果,第 4-6 列为普通话表达的结果。可以看出,各组方程中年龄与农民工健康显著地负相关,说明年龄越大的农民工身体健康状况越差。年龄的增大会带来身体机能的下降,加重各项疾病的发生率,这符合人们普遍的生理规律。

在不加入工作特征、职业、行业和地区变量之前,性别、受教育年限对健康变量亦有影响(见第 2、5 列),但加入上述变量之后,性别、受教育年限,回归系数变得不显著,这说明性别、受教育年限对农民工健康的影响可能是通过工作特征(如劳动收入)等其他渠道产生的。具体的工作特征方面(见第 3、6 列),周工作时长和工作年限的回归系数符号为负,虽然统计上不显著,但说明工作时间越长,农民工健康有下降的趋势。全年劳动收入的回归系数显著为正(1%的统计水平上显著),说明收入越高,农民工接受医疗救助和医疗保障的可能性相对较大,同时也会加大对健康的重视程度,越有利于维系健康。

在本文最为关心的语言能力方面,可以看到在依次加入控制变量之后,普通话听辨和表达的回归系数相应地有所下降,但它们对农民工自评健康的影响在各个方程中都显著为正(均在 1%统计水平上显著)。因此,有证据表明农民工的普通话听辨和表达能力均能够对其身体健康情况产生积极作用,即普通话听辨和表达能力越好的农民工,其身体健康状况也表现得更好,这也与国外既有文献的研究成果相吻合。

表 2 普通话能力对农民工健康的影响(有序 Probit 模型的估计结果)

解释变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)
普通话听辨(ML)	0.267*** (8.62)	0.217*** (6.68)	0.199*** (5.75)				0.176*** (4.49)
普通话表达(MS)				0.175*** (6.09)	0.104*** (3.50)	0.092*** (3.03)	0.049 (1.25)

续表 2

解释变量		(1)	(2)	(3)	(4)	(5)	(6)	(7)
个体基本特征	性别		0.1061* (1.85)	0.0909 (1.50)		0.1048* (1.80)	0.0615 (0.96)	0.1031 (1.55)
	年龄		-0.0160*** (-5.16)	-0.0149*** (-4.04)		-0.0162*** (-5.12)	-0.0142*** (-3.82)	-0.0145*** (-3.82)
	婚姻		-0.128 (-1.29)	-0.131 (-1.25)		-0.1318 (-1.33)	-0.127 (-1.26)	-0.128 (-1.22)
	民族		0.227** (2.23)	0.187* (1.82)		0.214** (2.11)	0.173* (1.69)	0.187* (1.82)
	受教育年限		0.0163* (1.73)	0.0133 (1.23)		0.0205** (2.13)	0.0124 (1.19)	0.0112 (1.02)
工作特征	周工作时长			-0.0017 (-1.03)			-0.0015 (-0.95)	-0.0017 (-1.04)
	工作年限			-0.0047 (-1.26)			-0.0044 (-1.18)	-0.0048 (-1.30)
	全年劳动收入			0.133*** (3.01)			0.139*** (3.15)	0.130*** (2.94)
其他	职业			控制			控制	控制
	行业			控制			控制	控制
	地区			控制			控制	控制
N		1613	1613	1613	1613	1613	1613	1613

注:括号内为 t 值,*、**、*** 分别表示在 10%、5%、1% 的水平上显著。

第 7 列汇报了一个比较有趣的现象。当普通话听辨和表达同时放入模型时,二者的回归系数较之前又有明显的下降,但是只有普通话听辨的回归系数统计上是显著的,而普通话表达的系数变得不再显著。这意味着相对于普通话表达而言,普通话听辨在维系农民工健康水平过程中扮演着更重要角色。可能的解释是:农民工的工作性质决定了他们大多数时间更需要听从管理指令,很少发表自己的看法,那么他们极有可能把这种语言习惯带到了求医问药的健康管理过程中,导致听辨对其健康的影响会更明显。无论怎样,以上实证结果表明:农民工的普通话能力对其自评健康水平具有显著的影响。

2.考虑内生性问题的工具变量回归。上文中我们采用了有序的 Probit 模型考察了普通话能力对农民工健康的影响,但回归中有可能存在着一定的内生性问题,其来源与方程中可能的遗漏变量、测量误差以及语言能力和健康之间反向因果关系等问题相关,进而导致回归结果存在偏误。为克服上述问题,最有效的方法是工具变量估计。不过,计量研究中合适的工具变量普遍难找,对于语言与健康关系研究更是如此。

表 3 普通话能力对农民工健康的影响(有序 probit IV 的估计结果)

解释变量	(1)	(2)	(3)
普通话听辨(<i>ML</i>)	0.351** (2.17)		0.251** (2.52)
普通话表达(<i>MS</i>)		0.250** (2.49)	0.081 (1.13)

续表 3

解释变量		(1)	(2)	(3)
个体基本特征	性别	0.1033 (1.56)	0.0988 (1.50)	0.1062 (1.60)
	年龄	-0.0147*** (-3.88)	-0.0137*** (-3.60)	-0.0134*** (-3.55)
	婚姻	-0.140 (-1.34)	-0.140 (-1.33)	-0.137 (-1.30)
	民族	-0.139 (-1.13)	-0.154 (-1.24)	-0.139 (-1.12)
	受教育年限	0.0100 (0.89)	0.0106 (0.94)	0.0084 (0.74)
工作特征	周工作时长	-0.0016 (-0.96)	-0.0014 (-0.85)	-0.0016 (-0.96)
	工作年限	-0.0045 (-1.25)	-0.0041 (-1.11)	-0.0047 (-1.27)
	全年劳动收入	0.130*** (2.95)	0.136*** (3.09)	0.128*** (2.90)
其他	职业	控制	控制	控制
	行业	控制	控制	控制
	地区	控制	控制	控制
corr(e.ML,e.Health)		0.006		0.043
corr(e.MS,e.Health)			0.001	0.005
N		1613	1613	1613

注：方程 1 中Ⅳ为同一地市级层面除受访者自身之外其他农民工普通话会听的平均水平和农民工个人的听英语的能力；方程 2 中Ⅳ为同一地市级层面除受访者自身之外其他农民工普通话会说的平均水平和农民工个人的说英语的能力；方程 3 中的Ⅳ包含了方程 2 和 4 的Ⅳ。括号内为 t 值，*、**、*** 分别表示在 10%、5%、1% 的水平上显著。

一般而言，工具变量的选取需至少满足以下三个条件：第一，它要与所替代的随机解释变量（即模型中的内生变量）高度相关；第二，要与被解释变量不相关；第三，要与模型中其他解释变量以及随机误差项不相关。按照这一原则，我们首先选取了农民工听英语和说英语的个人能力水平分别作为其普通话听辨和表达的工具变量。具体原因如下：第一，英语能力和所替代的解释变量——普通话能力密切相关。它们都是语言能力，都需要后天习得而来，很大程度上体现了劳动者的认知能力，二者具有很强的相关性。这符合工具变量选取的第一个条件。第二，英语能力与被解释变量——农民工健康水平无关。虽然有研究表明英语能力在我国具有较高的经济回报^①，可以提高劳动者的工资收入，但一般作用于非农民工样本。由于农民工的工作性质、居住环境等因素，农民工在工作和生活中几乎使用不到英语，英语能力对农民工的收入和健康不会产生影响。我们利用本文样本数据对农民工英语听、说能力和健康关系进行了估计，发现英语听说能力的回归系数均不显著。这与我们的判断基本一致，即没有证据表明农民工英语听说能力与其健康存在关联。这与前面第二个条件相一致。第三，英语能力与模型中的其他解释变量以及随机误差项无关。因此，将农民工听英语和说英语的能力作为其普通话听辨和表达的工具变量满足工具变量选取的基本条件。

其次，借鉴了使用更高层级的变量作为低层级变量的工具变量的思路，与赵颖的研究相类似^②，我们把同一地市级层面除自身外其他农民工语言能力的平均水平作为工具变量。它也基本满足上述所列举的工具变量选取的三个条件。具体地，首先我们将普通话听话和表达“一般”以下定义为不会听和不会说，之上定义为会听和会说。然后，将同一地市级层面除受访者自身之外其他农民工普通话

① 刘国辉、张卫国：《中国城镇居民英语能力的经济回报率研究》，《云南师范大学学报(哲学社会科学版)》2015 年第 6 期。

② 赵颖：《语言能力对劳动者收入贡献的测度分析》，《经济动态》2016 第 1 期。

会听的平均水平作为普通话听辨的工具变量,将同一地市级层面除受访者自身之外其他农民工普通话会说的平均水平作为普通话表达的工具变量。

针对不同的因变量形式,内生自变量的工具变量回归有多种形式(如,因变量为连续变量的二阶段最小二乘法,因变量为 0 和 1 离散变量的 Logit/Probit 工具变量估计,等等),但是本文因变量为有序的离散变量,上述传统工具变量回归方法均无法正确地估计出其结果。为解决这一问题,我们采用了含有内生协变量的有序 Probit 估计方法(ordered probit with endogenous covariates,以下简称有序 Probit IV 估计。具体见 Stata 软件第 15 版的扩展回归模块[ERM])。该方法同时汇报了一阶段回归方程残差项和主方程残差项相关性的估计结果。根据 Stata 官方的帮助说明文件,当两个方程残差项相关性的估计结果的 p 值统计上显著(小于 0.1),即接受“回归方程存在内生变量”的原假设,同时说明所使用的工具变量及方程回归结果有效,由内生性问题造成的主方程估计偏误得到纠正。

表 3 汇报了采用有序 Probit IV 估计的工具变量回归结果。我们看到表格下方一阶段方程残差项与主方程残差项相关性的 p 值均小于 0.1,说明我们的工具变量回归结果有效。分别对照表 2 第 3、6、7 列的有序 Probit 模型回归结果,可以看出使用工具变量前后普通话听辨和表达的估计系数在符号方向和统计显著度上基本保持不变。唯一的差别在于,工具变量法得到的系数较之前的回归结果稍稍偏大。出现此种结果的原因可能在于,回归中一些向下的衰减偏误超过了那些因遗漏变量或者反向因果造成的向上偏误。事实上,这种情况并非个案,此现象出现于国内外劳动经济研究的众多文献中^①。这也说明我们之前使用有序 Probit 模型估计出来的结果可能偏低。

(二)普通话对农民工健康影响的群体差异

值得注意的是,农民工群体具有较强的异质性,那么语言对农民工健康的影响会随着农民工群体的异质性而变化吗?为更好地把握农民工语言能力和健康水平的关系,有必要对普通话对农民工健康影响的群体性差异做出进一步的考察。

1.男女农民工的差异。社会语言学认为,语言学习具有性别差异。这种差异不只在第二语言(外语)的学习上,也存在于对母语的学习、理解和使用上。一般来说,女性的语言能力往往要强于男性。我们对本文男女样本事先做了统计,也确实如此。女性农民工普通话听辨和表达能力的均值比男性均值分别高出 0.11 和 0.21。因此,我们首先对男女农民工普通话能力与健康关系进行了分析。这里采取了与前文相同的模型和方法,对男女农民工样本进行了分别回归,结果见表 4。第 1-2 列为普通话听辨的结果;第 3-4 列为普通话表达的结果;第 5-6 列为普通话听辨和表达同时放入方程的结果。其中,有序 Probit IV 估计结果的一阶段方程残差项与主方程残差项相关性的 P 值均小于 0.1,说明本表工具变量回归结果是有效的。

表 4 男女农民工普通话能力与健康的回归结果

解释变量	(1) 有序 Probit	(2) 有序 Probit IV	(3) 有序 Probit	(4) 有序 Probit IV	(5) 有序 Probit	(6) 有序 Probit IV
Panel A: 男性农民工						
普通话听辨	0.163*** (3.70)	0.338*** (4.13)			0.162*** (3.32)	0.290*** (3.08)
普通话表达			0.075 (1.63)	0.272* (1.93)	0.003 (0.06)	0.176** (2.48)
corr(e.ML, e.Health)		0.007				0.048

① Chiswick B., Miller P., *The Economics of Language: International Analyses*, London: Routledge, 2007.

续表 4

解释变量	(1) 有序 probit	(2) 有序 probit IV	(3) 有序 probit	(4) 有序 probit IV	(5) 有序 probit	(6) 有序 probit IV
$\text{corr}(e.MS, e.Health)$				0.024		0.001
N	1009	1009	1009	1009	1009	1009
Panel B: 女性农民工						
普通话听辨	0.260*** (4.47)	0.416* (1.76)			0.202*** (2.91)	0.354*** (2.71)
普通话表达			0.213*** (3.73)	0.238* (1.84)	0.106 (1.55)	0.052 (0.45)
$\text{corr}(e.ML, e.Health)$		0.004				0.002
$\text{corr}(e.MS, e.Health)$				0.007		0.008
N	604	604	604	604	604	604

注:本表各个方程均包含了表 3 模型中的所有变量,第 2、4、6 列方程所使用的工具变量也与表 3 中相同。限于篇幅,这里只给出了普通话能力的回归系数。括号内为 t 值,*、**、*** 分别表示在 10%、5%、1%的水平上显著。

从表 4 可以看出:第一,无论是普通话听辨还是普通话表达,总体上它们对男女农民工健康都产生一定的影响。简单地从回归系数上看,普通话听辨对女性农民工健康的影响大于男性农民工。然而,普通话表达则恰好相反,它对男性农民工健康的影响大于女性(见第 4 列有序 Probit IV 估计结果)。这种现象可能的解释是,绝大多数农民工在城市中经济社会地位低下,社会网络和交际渠道狭窄,对外交流的机会少。而相对于城市居民,农民工思想保守,男主外、女主内的传统观念让女性农民工与外界的交流更少。即使外界有交流,相对于男性,女性农民工更多的是陪同和倾听,说的机会也很少。在这种情形下,女性农民工听普通话的能力显得格外重要。同理,相对于女性,男性农民工说普通话的能力尤显重要。因此,女性农民工普通话听辨能力和男性农民工普通话表达能力成为“稀缺”资源,它们对健康的影响自然增大。

表 4 第 6 列有序 Probit IV 回归结果进一步证实了上述推断。当普通话听辨和表达同时放入方程时,女性农民工样本普通话听辨的回归系数大于男性农民工样本的系数(均在 1%的统计水平上显著),但她们普通话表达的回归系数统计上不显著,而男性农民工普通话表达的回归系数则是显著的(5%的统计水平),这说明拥有较强普通话表达能力的男性农民工和拥有较强普通话听辨能力的女性农民工更容易融入城市生活之中,更加适应身份和地域的转变,也更容易接受到良好的医疗保障和救助。

表 5 新、老农民工普通话能力与健康实证结果

解释变量	(1) 有序 Probit	(2) 有序 Probit IV	(3) 有序 Probit	(4) 有序 Probit IV	(5) 有序 Probit	(6) 有序 Probit IV
Panel A: 老一代农民工						
普通话听辨	0.152*** (3.68)	0.484** (2.01)			0.141*** (3.07)	0.341*** (2.89)
普通话表达			0.0895** (2.12)	0.289** (2.11)	0.0271 (0.58)	0.0859 (0.85)
$\text{corr}(e.ML, e.Health)$		0.000				0.054

续表 5

解释变量	(1) 有序 Probit	(2) 有序 Probit IV	(3) 有序 Probit	(4) 有序 Probit IV	(5) 有序 Probit	(6) 有序 Probit IV
$\text{corr}(e, MS, e, Health)$				0.096		0.202
N	1075	1075	1075	1075	1075	1075
Panel B: 新生代民工						
普通话听辨	0.291*** (4.66)	0.541** (2.11)			0.278*** (4.62)	0.350** (2.02)
普通话表达			0.278*** (3.94)	0.364** (2.31)	0.054 (0.91)	0.182* (1.71)
$\text{corr}(e, ML, e, Health)$		0.000				0.000
$\text{corr}(e, MS, e, Health)$				0.001		0.000
N	538	538	538	538	538	538

注:关于变量说明,同表 4。括号内为 t 值,*、**、*** 分别表示在 10%、5%、1% 的水平上显著。

2. 新生代和老一代农民工的差异。年龄是影响健康的一个重要因素。我们需要对不同年龄层次的农民工普通话能力与其健康关系进行考察。但是受本文有效样本量的限制,如果将年龄按以往研究那样分为 18-29 岁、30-39 岁、40-49 岁、50 岁及以上等不同的年龄组,一些分组中样本量过小,将不具有统计上的意义。随着时代发展和环境变化,农民工群体发生代际更替和内部分化,出现了“新生代农民工”和“老一代农民工”两个群体,农民工的代际差异问题也日益成为研究的热点之一。因此,我们退而求其次,将农民工的年龄异质性分为新生代农民工和老一代农民工。根据当前学术界的共识,以 1980 年为限,此前出生的农民工被视为“老一代农民工”,此后出生的则为“新生代农民工”^①。代际差异是客观存在的,由于出生年代和社会背景的不同,不同的代际群体无论是价值观,还是态度反应、行为偏好,都存在一定的差异。研究新老农民工群体在语言能力和健康水平上的代际差异,可以更好把握农民工在社会融入中的特点和发展趋势。表 5 汇报了新老农民工子样本的回归结果。

我们看到,无论是有序 Probit 模型估计,还是有序 Probit IV 估计(其一阶段方程残差项与主方程残差项相关性的 p 值均小于 0.1,工具变量回归结果有效),结果都显示普通话能力(听辨和表达)对新老农民工健康产生显著性影响。同时,各个方程中新生代农民工普通话听辨和表达的回归系数都超过老一代农民工的系数。特别地,在将普通话听辨和普通话表达放入同一方程的有序 Probit IV 估计结果中(见表 5 第 6 列 Panel B),新生代农民普通话表达回归系数也是显著的。这与表 4 女性分样本的回归结果相类似。前面提到,新生代农民工出生于 1980 年之后,他们较老一代农民工接受了更好的普通话普及教育。在日常生活中新生代农民工会更容易或更熟练地运用普通话来扩大交际网络,可以利用自己的普通话优势接受更好的医疗,进而获得更好的身体健康水平。这进一步验证了普通话能力越高,农民工健康状况越好。

四、关于普通话与农民工健康关系的逻辑讨论

前面我们发现,有证据表明普通话能力更好的农民工会拥有更好的健康状态。那么,其背后有怎样的机理或传导路径? 这里有必要做进一步的讨论。

^① 王春光:《新生代农村流动人口的社会认同与城乡融合的关系》,《社会学研究》2001 第 3 期。

我们知道,语言是人们交流思想、表达情感,维系生产生活最基本的工具,它在流动人口的社会融入中发挥着不可忽视的作用。在国外较为成熟的移民社会融入研究及语言经济学研究中,语言常常被视为决定劳动者就业、工资收入及社会经济地位的重要因素之一。那么,语言对流动人口健康产生影响,很大程度上要归因于语言的交际作用和人力资本价值。

首先,语言作为交流的媒介,通过传递信息实现人与人之间的沟通,有利于人们更好参与到社会生活与人际交往之中。农民工在城市融入的过程中,首先面对的就是与城市人口的平等交流问题。但是农民工进入城市不同程度地都会面临着普通话理解和表达的挑战。语言障碍直接导致流动人口难以接受有效的预防保健和医疗诊治^①。由于不能有效地运用普通话交流,农民工一方面难以顺利地去寻医问药,另一方面农民工对看病抱有抵触情绪,即使主动就医,也很难使用普通话将自己的病情描述清楚,因而不能很好地得到救治。

其次,语言的交际功能可以使其对流动人口的社会关系网络产生积极的推动作用。农民工在城市里往往禁锢在乡音的小圈子里,群体外的社会交往范围有限,较差的普通话能力更是成为他们对外人际交往的一大阻碍。相反,较好的普通话能力可以使他们改善邻里关系,获得更大的交往空间。有研究表明,社会网络关系特别是来自家庭、朋友的支持将对个人健康产生有益影响^②。因此,语言可以通过影响农民工的社会网络关系进而对其健康产生影响。

再次,从语言经济学视角来看,语言能力是人力资本的一种表现形式^③,而人力资本是提高个人经济社会地位的重要因素之一。国内外已有研究表明,具备更高语言能力的流动人口能够更好地实现社会融入,尤其体现在获得就业机会和更高的工作报酬等方面上^④。也就是说,语言能力有利于提升社会经济地位,而社会经济地位是健康水平的主要预测指标之一^⑤,因为农民工个体经济水平是平时是否能顺利接受医疗保险和保健服务、生病时能否及时得到救治的首要前提,这意味着语言能力还可以通过影响农民工的个人经济状况对其健康产生影响。

最后,语言能力较差的流动人口,在新的生活环境中可能会遭受更多的生存压力和社会歧视,而压力和歧视可能会导致个人健康水平下降^⑥。前面提到,语言有助于扩大流动人口的社会网络空间,而良好的人际关系对减少工作生活中的压力和歧视具有积极的调节作用,这对农民工健康也会产生积极的效果。

五、结论

利用 CGSS 2013 的微观数据,采用有序的 Probit 模型回归和含有内生协变量的有序 Probit 估计方法,研究考察了普通话能力对农民工健康的影响。我们发现,在控制了性别、年龄、民族、受教育年限、工作经验、劳动收入以及职业、行业 and 地区等个体及区域层面因素后,普通话能力(听辨和表达)与农民工健康水平之间存在着正相关关系。在分样本回归中,我们还发现在增强自身健康方面,听普通

① Lebrun L., "Effects of Length of Stay and Language Proficiency on Health Care Experiences among Immigrants in Canada and the United States", *Social Science and Medicine*, 2012, 74(7), pp.1062-1072.

② Kawachi I., Berkman L., "Social Ties and Mental Health", *Journal of Urban Health*, 2001, 78(3), pp.458-467.

③ Chiswick B., Miller P., "The Endogeneity between Language and Earnings: International Analyses", *Journal of Labor Economics*, 1995, 13(2), pp.246-288;张卫国:《作为人力资本、公共产品和制度的语言》,《经济研究》2008 第 2 期。

④ Chiswick B., Miller P., "The Economics of Language: International Analyses", London: Routledge, 2007.

⑤ Adler N. E., Newman K., "Socioeconomic Disparities in Health: Pathways and Policies", *Health Affairs*, 2002, 21(2), pp.60-76.

⑥ Williams D. R., Neighbors H. W., Jackson J. S., "Racial/Ethnic Discrimination and Health: Findings from Community Studies", *American Journal of Public Health*, 2003, 93(2), pp.200-208.

话的能力对于女性农民工贡献更大,而说普通话的能力对于男性农民工的贡献更大;同时,有证据表明相对于老一代农民工,新生代农民工能更好地利用普通话优势获得更佳的健康。在克服了可能存在的内生性问题后,结论依然稳健。

在当前人口流动性愈发增强的背景下,克服语言障碍对于加强农民工健康服务具有重要的现实意义。普通话作为国家的通用语言,与其他方言和少数民族语言相比,具有更强的交际作用和人力资本价值,可以通过帮助农民工改善医患沟通、扩大社会关系网络、提高经济社会地位以及减少压力与歧视等提高他们的健康水平。当前我国脱贫攻坚正进入决胜时刻,国家积极倡导包括语言扶贫、健康扶贫在内的全方位扶贫举措,从这个意义上讲,个体劳动者要增强语言能力意识,有关管理部门要树立语言服务健康的意识,国家则有必要继续强化普通话推广和相关培训工作,充分发挥语言的经济社会价值,服务健康,服务经济社会发展。

The Impact of Proficiency in Mandarin on Migrants' Health in China

Liu Guohui Shan Baogang Zhang Weiguo

(College of Foreign Languages, Yanbian University, Yanji 133002, P.R.China;

Center for Economic Research, Shandong University, Jinan 250100, P.R.China)

Abstract: Based on the micro-data of the China General Social Survey in 2013, this paper examines the impact of proficiency in Mandarin on health among migrant workers in China. We find that proficiency in Mandarin is positively associated with migrant workers' health. A higher level of Mandarin listening and speaking ability helps to increase the probability of being healthy for migrant workers. A further study shows that there is heterogeneity in the health effects of proficiency in Mandarin. After controlling for the possible endogeneity, the results are still robust.

Keywords: Proficiency in Mandarin; Health effect; Migrant worker; Heterogeneity

[责任编辑:郝云飞]