

# 人力资本对农村迁移劳动力收入的影响研究

张世伟,赵亮,万相昱

(吉林大学 数量经济研究中心,吉林 长春 130012)

**摘要:**文章利用2005年吉林省进城务工人员抽样调查数据和2006年吉林省农村抽样调查数据,运用分位数回归方法和Heckman两步骤样本选择偏差方法,对人力资本对农民工收入的影响进行分析。研究结果表明,在不同的分位数回归结果中,教育回报率有所不同,在25%分位数回归中的教育回报率最大(1.89%),而中位数回归中的最小(1.40%)。在外出务工时间长的组,组内收入差距大于务工时间短的组。接受其他组织培训的组,组内收入差距比接受企业培训的组大。在进行了样本选择偏差修正之后,教育回报率上升到4.55%,所以在考察农村劳动力迁移收入问题时,不能忽视样本选择偏差的问题。

**关键词:**人力资本;农民工;收入;分位数回归;Heckman两步骤法

**中图分类号:**F241

**文献标志码:**A

**文章编号:**1008-5831(2009)01-0020-04

## 一、引言

农民工是中国经济社会转型时期的特殊概念,是当代中国城市的特殊群体。农民工也称为进城务工人员,是指兼具农民与工人身份的劳动者,户籍身份还是农民、有承包土地,但是主要从事非农产业、以工资为主要收入来源的人员,包括在第二、第三产业中的劳动者。

农民工的收入决定问题长期以来受到经济学家和社会学家的广泛关注。自人力资本理论创立以来,研究人力资本与收入关系的研究逐渐增多<sup>[1]</sup>。教育、培训、健康、迁移共同构成人力资本的四大要件。个人的人力资本积累是工资收入的重要决定因素<sup>[2]</sup>。Mincer运用收入函数对教育投资收益率进行估计,揭示了人力资本投资与收入之间的关系<sup>[3]</sup>。Barro认为具有较高人力资本的劳动力往往能够在城市劳动力市场上获得更好的就业机会和更高的收入<sup>[4]</sup>。Chiswick(1978)和Bojas(1987)对美国外来移民的研究表明,教育水平、工作经验和其他劳动技能等人力资本对其在经济上取得成功至关重要<sup>[5-6]</sup>。Strauss(1986)、Deolalikar(1988)及Haddad与Bouis(1991)的研究表明,营养与健康都对劳动生产率有显著的促进作用<sup>[7-9]</sup>。

国内一些学者也对人力资本对农民工收入的影响进行了初步的研究。周其仁(1997)、都阳(1999)对农村迁移劳动力进行深入的研究发现,教育、培训、专业技能等在劳动力迁移决策及收入的提高方面有决定性的作用<sup>[10-11]</sup>。高文书在对五个城市的调查数据基础之上,对进城农民工的就业状况进行了实证分析,分析得到进城农民工的特征、就业和社会保障状况及收入的影响因素<sup>[12]</sup>。卢志刚、宋顺锋利用天津的调查问卷数据,运用多元分析统计方法,对农民工收入的微观影响因素进行分析,发现户口状况、年龄、在本岗位的时间、经验、文化程度、

收稿日期:2008-06-20

基金项目:国家社会科学基金项目(05BYJ026);教育部人文社会科学重点研究基地重大项目(05JJD790079);教育部新世纪优秀人才支持计划(NCET-05-0318)

作者简介:张世伟(1964-),男,吉林长春人,吉林大学数量经济研究中心教授,经济学博士,博士生导师,主要从事数量经济研究。

欢迎访问重庆大学期刊网 <http://qks.cqu.edu.cn>

企业的所有权形式以及合同有效期长度等都有显著统计意义<sup>[13]</sup>。曾旭晖以成都市进城农民工为研究对象,从人力资本的回收机制上分析了非正式劳动力市场的发展状况。研究表明进城务工人员的教育收益率很低<sup>[14]</sup>。严于龙利用国家统计局农村司 2005 年农村外出务工劳动力调查数据进行了条件分析,并采用多元线性回归模型对农民工收入的影响因素进行分析<sup>[15]</sup>。栾敬东通过实际调查和分析表明,在经济发达地区打工的流动人口的性别、年龄、受教育年限、在本单位工作年限等基本社会特征均对其收入水平存在显著性影响<sup>[16]</sup>。苏群、周春芳利用 2004 年江苏省的调查数据,采用多元回归模型,在对农民工的人力资本定性描述的基础上,对农民工收入决定的人力资本因素进行了实证分析。结果表明文化程度、培训状况、打工时间、性别及婚姻状况等,对农民工收入的影响显著<sup>[17]</sup>。高嵩、焦宝松、王建运用经济学方法对决定农民工收入的户籍制度、教育水平和工作年限等方面进行分析,并尝试在此基础上提出一些切实可行的建议<sup>[18]</sup>。魏众根据 1993 年的中国营养调查数据,利用因子分析法提取健康指标,探讨了农村地区健康对非农就业及其工资决定的影响,研究发现对于劳动参与及非农就业机会,健康状况影响显著,但是对工资决定没有产生显著影响<sup>[19]</sup>。张车伟运用来自中国贫困农村的数据,系统地研究了营养、健康对劳动生产率或者说收入的影响,结果表明几乎所有的营养和健康方面都影响到农村的劳动生产率<sup>[20]</sup>。

笔者利用 2005 年吉林省进城务工人员抽样调查数据和 2006 年吉林省农村抽样调查数据,应用微观计量经济学方法对人力资本对农民工收入的影响进行分析。笔者的第二部分论述回归模型的设计,采用普通最小二乘回归、分位数回归和 Heckman 两步骤方法进行研究。第三部分对估计结果进行分析,最后给出研究结论。

## 二、模型

按照人力资本理论,农民工的自身人力资本要素会影响到收入,比如教育水平、培训、经验及健康状况。除此之外,性别、合同的长期有效性及所从事的职业等也都会影响到收入。因此建立如下多元线性回归模型:

$$Y = \alpha + X'\beta + \varepsilon \quad (1)$$

其中: $Y$ 表示月务工收入的对数; $X$ 表示所有自变量,包括:受教育年限;是否接受过企业培训,接受过为 1;是否接受过其他组织的培训;是否有技术等级,有技术等级为 1;一组务工时间的虚拟变量,以半年以内为省略变量;性别虚拟变量,男性为 1;年龄及年龄的平方;一组进城务工人员来源的虚拟变量,以本市为省略变量;一组行业虚拟变量,以其他行业为省略变量;合同虚拟变量,签订合同为 1。

普通最小二乘回归(OLS)建立了因变量 $Y$ 的条件均值与自变量 $X$ 之间的线性关系模型,它只能概括出数据集合的平均信息,估计出自变量对因变量的平均边际效果。为了弥补普通最小二乘法的缺陷,Koenker 和 Bassett 发展了有关变量的分位数估计的理论<sup>[21-22]</sup>。它根据因变量的条件分位数对自变量进行回归,能估计出不同分位数的估计结果。当自变量

对不同部分的因变量的分布产生不同的影响时,它能更加全面地刻画分布的特征。

第 $q$ 分位数回归估计值 $\hat{\beta}_q$ 是下面线性规划问题(2)最小化求解过程。

$$Q_N(\beta_q) = \sum_{i: y_i \geq X_i'\beta} q |y_i - X_i'\beta_q| + \sum_{i: y_i < X_i'\beta} (1-q) |y_i - X_i'\beta_q| \quad (2)$$

其中: $y_i$ 代表因变量的向量; $X_i$ 代表自变量的向量; $q$ 是所要估计的分位数值,0 到 1 之间任意值; $\beta_q$ 是一个系数向量。

为了研究人力资本对不同收入群体的影响,笔者采用分位数回归方法。公式(2)中的因变量为每月务工收入对数。自变量同公式(1)。

在研究人力资本对农民工收入的影响时,只有在农民外出打工时才有收入数据,仅仅利用这部分外出打工者的信息进行普通最小二乘回归时,估计结果是有偏的。Heckman(1974)提出的两步骤法(Two-Step)解决了样本选择偏差问题。

假设农民是否外出打工的方程如下:

$$p_i^* = z_i' + u_i \\ p_i = \begin{cases} 1 & \text{if } p_i^* > 0 \\ 0 & \text{if } p_i^* \leq 0 \end{cases} \quad (3)$$

其中, $p_i^*$ 表示不可观测的决定个体 $i$ 是否外出打工的变量, $p_i$ 表示个体是否外出打工(1 表示外出打工,0 表示没有外出打工), $z_i$ 表示自变量向量,包括:个体性别、年龄及年龄平方、婚姻状况、受教育年限、技术等级(以无技术等级为省略变量)、健康状况(以不能工作为省略变量)、家庭收入、耕地、劳动力数、外出劳动力数、户主受教育程度、年龄及年龄平方。 $\alpha$ 表示个体属性向量系数,表示随机扰动项,服从 $N(0,1)$ 。个体 $i$ 的外出打工概率可以表示为:

$$\Pr(p_i = 1) = \Phi(z_i'\alpha) \quad (4)$$

其中, $\Phi(\cdot)$ 是服从标准正态分布的分布函数。

设农民工收入方程为

$$Y_i = X_i'\beta + \varepsilon_i \quad (5)$$

其中, $Y_i$ 表示个体 $i$ 务工月收入对数, $X_i$ 是影响个体收入的一系列因素,包括:个体受教育年限、技术等级、健康状况、性别、年龄、年龄平方及合同虚拟变量(以未签订合同为省略变量)。 $\varepsilon_i$ 是随机扰动项,服从 $N(0, \sigma_i)$ 。如果 $u_i$ 和 $\varepsilon_i$ 相关,则方程(3)和方程(5)构成了二元样本选择模型(bivariate sample selection model)。

$\lambda_i$ 为逆米尔斯比:

$$\lambda_i = \frac{\varphi(z_i'\alpha)}{\Phi(z_i'\alpha)} \quad (6)$$

## 三、结果分析

笔者采用两套数据样本,分别是 2005 年吉林省进城务工人员抽样调查数据和 2006 年吉林省农村抽样调查数据。第一套数据样本量足够大,而且在农民工外出务工地抽样调查,所以更能反映农民工收入的真实情况,可以用来分析人力资本对不同收入群体收入分布的影响。第二套数据因为是农村抽样调查数据,既包括外出劳动力信息,也包括非外出劳动力信息,在分析人力资本对农民工收入的影响

时,可以运用 Heckman(1974)提出的两步骤法(Two-Step)来解决自我选择偏差问题。

首先利用第一套数据,运用普通最小二乘回归方法,对方程(1)进行回归,分析人力资本对农民工收入的平均边际作用。为了研究人力资本对不同收入水平的影响,运用分位数回归方法,对方程(2)进行回归。结果如表1。表1第1列给出OLS估计结果。从总体样本回归结果来看, $F$ 值在1%水平上通过显著性检验,说明模型整体上线性显著。从方程的系数估计值来看,得到以下结论。男性的每月收入比女性高出16.20%。受教育年限每增加一年,收入增加1.84%。进城务工人员的年龄与收入正相关,年龄每增加1岁,月收入增加0.29%。年龄平方的系数是负的,说明随着年龄的增长,边际收入是递减的。这一结果符合理论上的预期。在其他条件相同的情况下,本省和外省的进城务工人员的月收入都高于本市的进城务工人员。务工时间在1年至2年和2年以上的农民工比务工时间在半年以内的月收入分别高出2.85%、6.32%和9.28%,随着务工时间的增加对月收入的影响也逐渐增加。企业培训和其他组织培训对进城务工人员月收入的增长有促进作用。有技术等级的进城务工人员的月收入比没有技术等级的务工人员的月收入要高出7.70%。住宿餐饮业、制造业、批发零售业和居民服务业进城务工人员的月收入相对于其他行业来说要低一些。建筑业的月收入与其他行业没有显著的差异。与用人单位签订合同的比没有签订合同的进城务工人员的月收入大约高出9.54%。在不同的分位数回归结果中,教育的回报率有所不同,10%、25%、50%、75%

和90%分位数回归中教育回报率分别为1.67%、1.89%、1.40%、1.65%和1.67%,其中在25%分位数回归中的教育回报率最大,中位数回归中的教育回报率最小。在高收入组中,男女收入差异最小。在务工时间长的组,组内收入差距大于务工时间短的组。接受其他组织培训的组,组内收入差距比接受企业培训的组大。

然后利用第二套数据,对方程(3)和(5)进行估计,结果如表2。表2的第1列运用普通最小二乘法估计的教育回报率是4.32%,而在进行了样本选择偏差修正之后,教育回报率上升到4.55%。同时,健康状况变量也在样本选择偏差修正前后有明显差异,边际贡献由1.68%上升到3.44%。在Heckman两步骤法样本选择方程中,大部分变量都通过了显著性检验。男性选择外出打工的概率要低于女性。年龄越大外出打工的概率越小,这可能是由于一般农村劳动力进城务工,进入的行业都属于劳动力密集型行业,需要较好的体力,所以年龄越大其流动的倾向性越小。已婚个体外出打工的可能性越高,由于他们担负着更大的家庭负担。家庭外出劳动力数越多,个体外出务工的几率越大。这可能是由于家庭外出劳动力在城市中获得了一定的经济利益,所以带动家庭其他成员也进城务工,形成一个社会关系网。户主年龄越大,个体越不容易外出打工。这是由于他们有比较重的赡养老人的责任,不容易外出务工。在收入方程中男性的收入比女性低。受教育年限和健康状况对外出打工的收入有正向的作用,但是技术等级对收入的影响不显著。签订合同的农民工比没有签订合同的收入要高25.90%。

表1 2005年农民工收入分位数回归结果

自变量	OLS	QR				
		10分位	25分位	50分位	75分位	90分位
男性	0.161 9***	0.178 4***	0.189 7***	0.188 2***	0.175 0***	0.114 9***
教育年限	0.018 4***	0.016 7***	0.018 9***	0.014 0***	0.016 5***	0.016 7***
年龄	0.002 9***	0.003 9**	0.001 6*	0.001 9***	0.002 7***	0.004 2***
年龄平方	-0.000 3***	-0.000 3**	-0.000 1*	-0.000 2***	-0.000 3***	-0.000 4***
务工时间						
半年至1年	0.017 3	0.023 8	0.050 9***	0.038 7***	0.001 2	-0.037 9*
1年至2年	0.063 2***	0.080 7**	0.066 9***	0.063 7***	0.039 2***	0.025 0
2年以上	0.092 8***	0.114 1***	0.104 4***	0.091 2***	0.087 1***	0.064 8***
企业培训	0.033 3***	0.095 4***	0.073 6***	0.034 8***	-0.003 3	0.028 8*
非企业培训	0.114 3***	0.105 7***	0.098 1***	0.132 2***	0.117 0***	0.094 8***
有技术等级	0.077 0***	0.087 7***	0.070 0***	0.058 7***	0.066 8***	0.082 7***
本省	0.068 3***	0.078 8***	0.090 7***	0.088 2***	0.061 8***	0.062 1***
外省	0.097 0***	0.196 8***	0.119 8***	0.093 5***	0.049 1***	0.023 5
住宿餐饮业	-0.242 7***	-0.311 5***	-0.242 3***	-0.256 6***	-0.182 8***	-0.089 0***
建筑业	0.015 2	0.036 4	0.001 7	-0.041 2***	0.015 2	0.125 7***
制造业	-0.124 7***	-0.107 2**	-0.128 2***	-0.161 0***	-0.098 4***	-0.027 9
批发零售业	-0.119 0***	-0.098 3*	-0.125 0***	-0.158 1***	-0.124 3***	0.012 5
居民服务业	-0.159 5***	-0.149 5***	-0.152 6***	-0.208 0***	-0.091 8***	0.007 4
签订合同	0.095 4	0.083 9***	0.110 5***	0.118 6***	0.075 7***	0.090 7***
常数项	6.059 2***	5.491 7***	5.847 5***	6.172 5***	6.356 1***	6.437 4***
(Pseudo) $R^2$	0.261 7	0.126 8	0.164 0	0.202 3	0.155 6	0.099 3
样本量				5 743		

表2 2006年农民工收入回归结果

自变量	OLS	Heckman 两步骤法	
		选择方程	收入方程
男性	-0.619 8***	-0.242 4***	-0.540 2***
教育年限	0.043 2*	-0.040 5*	0.045 5**
年龄	-0.037 1**	-0.092 9***	-0.038 5**
年龄平方	0.000 5*	0.000 4	0.000 8***
技术等级	0.068 2	-0.665 5	0.129 7
健康状况	0.168 4**	-0.655 4***	0.343 8***
签订合同	0.244 1**		0.259 0**
婚姻状况		0.426 4***	
户收入		2.36E-05***	
劳动力数		-0.178 6***	
外出劳动力数		0.485 2***	
耕地面积		-0.028 1	
户主年龄		-0.168 2***	
户主年龄平方		0.001 8***	
户主教育年限		0.043 6**	
常数项	7.358 9***	6.026 1***	7.483 9***
lambda			-0.537 9***
R <sup>2</sup> (或 Wald chi2)	0.313 2		338.570 0
样本量	232	2 490	2 490

#### 四、结论

笔者利用2005年和2006年吉林省两套抽样调查数据,运用分位数回归方法和Heckman两步骤样本选择偏差方法,对农民工收入方程进行回归。研究结果表明,在不同的分位数回归结果中,在25%分位数回归中的教育回报率最大(1.89%),而中位数回归中的教育回报率最小(1.40%)。在务工时间长的组,组内收入差距大于务工时间短的组。接受其他组织培训的组,组内收入差距比接受企业培训的组大。在高收入组中,男女收入差异最小。在进行样本选择偏差修正之后,教育回报率上升到4.55%,所以在考察农村劳动力迁移收入问题时,不可以忽视样本选择偏差的问题。

#### 参考文献:

- [1] SCHULTZ T W. Investment in Human Capital [J]. The American Economic Review, 1961, 51: 1-17.
- [2] BECKER G. Investment in Human Capital a Theoretical Analysis [J]. Journal of Political Economy, 1962, 70 (5): 9-49.
- [3] MINCER J. Schooling, Experience and Earnings [M]. New York: Columbia University Press, 1974.

- [4] BARRO R, SALA-I-MARTIN X. Convergence [J]. Journal of Political Economy, 1992, 100 (2): 223-251.
- [5] CHISWICK B. The Effect of Americanization on the Earnings of Foreign-born Men [J]. Journal of Political Economy, 1978, 86 (8): 897-921.
- [6] BOJAS G. Self-selection and the Earning of Immigrants [J]. The American Economic Review, 1987, 77 (4): 531-553.
- [7] STRAUSS J. Does Better Nutrition Raise Farm Productivity? [J]. The Journal of Political Economy, 1986, 94 (2): 297-320.
- [8] DEOLALIKAR A B. Nutrition and Labor Productivity in Agriculture; Estimates for Rural South India [J]. The Review of Economics and Statistics, 1988, 70 (3): 406-413.
- [9] LAWRENCE J, HADDAD, HOWARTH E BOUIS. The Impact of Nutritional Status on Agricultural Productivity: Wage Evidence from the Philippines [J]. Oxford Bulletin of Economics and Statistics 1991, 53 (1): 45-68.
- [10] 周其仁. 机会与能力——中国农村劳动力的就业和流动[J]. 管理世界, 1997(5): 81-100.
- [11] 都阳. 贫困地区农户参与非农工作的决定因素研究[J]. 农业技术经济, 1999(4): 33-37.
- [12] 高文书. 进城农民工就业状况及收入影响因素分析——以北京、石家庄、沈阳、无锡和东莞为例[J]. 中国农村经济, 2006(1): 28-34.
- [13] 卢志刚, 宋顺锋. 农民工收入微观影响因素统计分析[J]. 现代财经, 2006(10): 77-81.
- [14] 曾旭晖. 非正式劳动力市场人力资本研究——以成都市进城农民工为个案[J]. 中国农村经济, 2004(3): 34-38.
- [15] 严于龙. 对农民工收入影响因素初步分析[J]. 中国统计, 2006(10): 19-20.
- [16] 栾敬东. 流动人口的社会特征及其收入影响因素分析[J]. 中国人口科学, 2003(2): 70-75.
- [17] 苏群, 周春芳. 农民工人力资本对外出打工收入影响研究——江苏省的实证分析[J]. 农村经济, 2005(7): 115-118.
- [18] 高嵩, 焦宝松, 王建. 农民工收入决定的实证分析——以北京市为例[J]. 经济问题, 2006(10): 62-64.
- [19] 魏众. 健康对非农就业及其工资决定的影响[J]. 经济研究, 2004(2): 64-74.
- [20] 张车伟. 营养、健康与效率[J]. 经济研究, 2003(1): 3-9.
- [21] KOENKER R, BASSETT Jr G. The Asymptotic Distribution of the Least Absolute Error Estimator [J]. Journal of the American Statistical Association, 1998, 73: 618-662.
- [22] KOENKER R, BASSETT Jr G. Regression quantiles [J]. Econometrica, 1978, 46: 33-50.

## The Effect of Human Capital on the Income of Peasant Workers

ZHANG Shi-wei, ZHAO Liang, WANG Xiang-yu

(Center for Quantitative Economics, Jilin University, Changchun 130012, China)

**Abstract:** By using the survey data of peasant workers in Jilin Province and sample survey data in rural areas of Jilin Province, the paper applies quantile regression and Heckman Two-Step sample selection bias methods in the study of the effect of human capital on the income peasant workers. The study shows that in different points of the quantile regression, return to education is different. In 25% quantile regression return to education is the largest (1.89 percent), while the median regression return to education is the smallest (1.40 percent). The income gap of long-time migrant workers group is greater than a short period of time working group. The income gap of the receiving Other organizations training group is greater than the group of receiving business training. As a result of amendments to sample selection bias, return to education rises to 4.55 percent, so the study in the income of migration of rural labor does not lose sight of the sample selection bias.

**Key words:** human capital; peasant workers; income; quantile regression; Heckman Two-Step

(责任编辑 傅旭东)