

第二代农民工劳动收入影响因素

——基于 OLS 与 Multinomial Regression 方法

惠 源，胡宏伟

(武汉大学社会保障研究中心，湖北武汉 430072)

关键词：第二代农民工；劳动收入；影响因素

摘 要：伴随经济快速发展，中国工业化、城市化加速，进城农民群体出现并且规模迅速膨胀。与第一代农民工相比，第二代农民工在教育、城市化意愿等方面有不同特征，尤其是其在工作收入及收入影响因素方面有了新的变化趋势。基于明瑟扩展方程，通过对典型数据的分析，使用最小二乘法(OLS)，测量了年龄、教育、经验等因素对第二代农民工劳动收入的是否存在影响，以及影响的方向和程度，并使用多分变量逻辑斯蒂回归(Multinomial Logistic Regression) 方法进行了进一步验证。发现教育水平和性别差异对第二代农民工的劳动收入有显著影响，年龄、就业方式、工作单位性质等因素也有一定程度影响，但影响并不稳定。在研究结论的基础上提出了加强对第二代农民工的教育和就业扶持等对策性思考。

中图分类号：C913.2 **文献标志码：**A

基金项目：教育部“国际金融危机应对研究”应急课题资(2009JYJR004)、国家自然科学基金重点项目“中国补充养老保险制度研究”(70533040) 支持。

The Study on Influence Factors of Second Generation of Migrant Workers' Income —Based on OLS and Multinomial Regression

Methods

HUI Yuan, HU Hong-wei

(Social Security Studies Center, Wuhan University, Wuhan Hubei 430072,
China)

Key words: the second-generation of migrant workers; labor income; influence factors

Abstract: Accompanied by the rapid development of economy, Speed up of China's industrialization and urbanization, emergence of migrant groups and the rapid expansion of the scale. Compared to the first generation of migrant workers, the second generation of migrant workers have different characteristics in education, urbanization and other aspects, especially its impact on income and income factors that have new trends. Based on the Mincer equation extended by the analysis of typical data, the use of least squares (OLS) , measured the age, education, experience and other factors on the second generation of migrant workers of the existence of labor income impacts, aswell as the direction and extent of impact, and the use of multi-variable logistic regression (Multinomial Logistic Regression) method

for further validation. Found gender differences in educational level and the second generation of migrant workers on the labor income has a significant influence, age, employment, the nature of work and other factors have to some extent, but the impact is not stable. At the conclusion of the study based on the strengthening of the second generation of migrant workers to support the education and employment, such as thinking countermeasures.

一 研究背景与问题提出

自实行改革开放政策以来，我国的经济、社会、政治体制均发生了深刻的变化，亦取得了非凡的成就。但是，自 20 世纪 90 年代中期以来，城乡之间、区域之间、群体之间的收入差距却日益凸显。伴随工业化、城市化加速，进城农民工逐渐成为了日益庞大的社会群体。随着工业化、城市化和现代化加速，第二代(或新生代) 农民工群体出现并逐渐成为农民工群体主流。

但是，国内外对于中国第二代农民工群体的关注和研究却并不充分。王兴周(2008)认为新生代农民工的群体特性是积极的，他们具有较高的人力资本和个人素质；有理想、有目标；注重市场原则；更有平等意识；独立自主意识更强。张宝庆(2005)在呈现新生代农民工流动动机的基础上，凸显了新生代农民工的群体特征，也分析了由于他们的流动而引发的城乡社会的融合问题，以及这种隐患对建构和谐社会的挑战。许传新、许若兰(2007)以及许传新(2007)认为新生代农民工城市工作适应、人际适应都处于中等水平，而生活适应处于较低水平。回归分析结果表明，自评家庭阶层地位、媒介接触、城市经历、相对剥夺感、社区参与、组织支持对新生代农民工城市社会适应有显著性影响。邹英(2007)认为新生代农

民工产生身份认同困境的原因在于其个人身份建构与社会建构之间的矛盾,因而主张必须从内部条件、外部环境和政策建构等多个层面着手才能解决他们的身份认同的困境问题。许多国内学者关注了两代农民工在群体特征等方面的差异。吴红宇、谢国强(2006)认为,新生代农民工的利益诉求与第一代农民工不同,传统农民工进城是为了生存,新生代农民工则是为了谋求更大的发展前途。因此新生代农民工的角色也产生了变化:由传统的劳动力的供给者和家庭的生产者转变为人力资本的投资者和地区生活环境的消费者。与此类似的研究还有刘天金、蔡志强(2007),他们认为,第一代农民工从简单谋生的角度出发,涌入城市寻找生存机会,他们能够忍耐当时政策对他们行为和利益的节制还有不少学者对农民工的代际差异进行比较分析,突显了不同代际农民工在年龄、初次流动年代、婚姻状况和文化程度等方面的差异,为进行新老世代农民工的比较研究提供了分析的视角与思路(王东、秦伟,2002;成艾华、姚上海,2005;杜书云、张广宇,2008;等)。

虽然部分国内学者的研究涉及了第二代农民工的劳动收入及其影响因素,但是,大部分研究将重点放在了第二代农民工的收入分布,以及与第一代农民工收入的差异,较少有研究基于调查数据分分析影响第二代农民工劳动收入的影响因素。正是基于上述背景,在分析典型地区调研数据的基础上,本文将通过最小二乘法和明瑟扩展方程,分析影响第二代农民工劳动收入的典型因素。同时,本文还将通过多分变量逻辑斯蒂回归进一步验证分析结论。

二 理论基础、模型、数据与变量

1 理论基础与模型

现代人力资本理论认为,人力资本存在于劳动者本身,表现为知识、技能、

智力、体力(健康状况)价值的总和。一个国家的人力资本可以通过劳动者的数量、质量以及劳动时间来度量。同时，人力资本可以通过诸如教育、职业培训、医疗保健等多种投资渠道获得增长，而人力资本的增长是经济增长的主要源泉。而且，在所有的投资形式中，人力资本投资收益率最高，超过其他一切形态资本的投资收益率。研究市场经济中人力资本与收入差距之间关系最有代表性的分析工具是美国经济学家明瑟于 1972 年提出并首先使用的工资方程假定。本文主要基于明瑟的工资方程模型，来研究进城农民的人力资本因素在市场机制条件下如何与其收入产生关联作用。

明瑟方程认为，劳动者的工资性收入完全由劳动者人力资本和其他个体特征决定，个人收入可以看作一系列人力资本变量的函数。即， $Y_i=f(H)$ ，其中 Y =个人收入， H =人力资本变量。

为了更好地解释收入差距与人力资本之间的关系，可以同时从工资方程两端取方差，这样收入与人力资本之间的关系就可以表示为：

$$\text{Var}(Y_i)=M^2\text{Var}(H_i)$$

其中， M 为人力资本的市场回报。一般地，在一个完善的劳动力市场上， M 一般为常数。

但是，当劳动力市场不完善时，人力资本回报通常不能用一个常数来表达，它可能会随着劳动力就业的地区、行业、职业、及其工作经验等方面的不同而变化。在这种情况下，明瑟方程可以被进一步改写成如下形式：

$Y_i=\varphi[Mf(H_i)]$ ，其中， Y_i =个人收入， H_i =人力资本变量， M_r 为人力资本回报变量。

如果假定人力资本水平和人力资本回报之间是相互独立的，且 $E(H_i)=\mu_h$ 和

$E(M_r)=\mu_m$ 。同假设完善劳动力市场条件下的处理方法一样，对方程两端取方差，则收入与人力资本之间的关系式就可以表达为：

$$\text{Var}(Y_{ir})=\mu_m^2\text{Var}(H_{ir})+\mu_h^2\text{Var}(M_r)+\text{Var}(H_{ir})\cdot\text{Var}(M_r)$$

从上式可以看出，收入差距不仅依赖于人力资本水平的差异和人力资本回报的差异，即 $\text{Var}(H_{ir})$ 和 $\text{Var}(M_r)$ ，而且还依赖于人力资本的均值及人力资本回报的均值，即 μ_h 和 μ_m 。而且不难看出，如果人力资本回报为零，即 $M=0$ ，或者说人力资本无法得到市场回报时，则 $\text{Var}(Y_i)=0$ ，此时，收入差距和人力资本之间就不存在关系。

如果人力资本回报不为零，且不是一个常数的情况下，则收入不仅与人力资本及其差异有关，而且还同人力资本回报大小及其差异有关。概括来说，在给定人力资本水平及其差异的情况下，收入差距就由人力资本回报及其差异所决定，人力资本市场回报率越高，收入差异就将越大；同时，如果人力资本回报不是一个常数，则人力资本回报的差异程度越大，收入差距也会越大。由此可见，人力资本回报及其变化状况对收入水平有着重要意义。

在实证研究中，明瑟工资方程通常被作为估计人力资本回报的基本方法，其一般表达形式为：

$$\text{Ln}Y=X'\beta+\mu$$

其中， $\text{Ln}Y$ 为收入的对数， X 为一系列代表人力资本的变量(如受教育程度、工作经验、性别)； β 为待估计参数，一般被解释为人力资本的投资回报； μ 为误差项，且假设 $E(\mu)=0$ ，也即 μ 满足期望值为零的假设。

本研究将扩展明瑟工资方程，并将扩展后的工资方程作为本研究的模型基础，结合调查数据中的问题选定相应变量纳入分析。

2 数据与变量

经验分析使用的数据是基于武汉市进城农民的调查,数据来源于中南财经政法大学梅建明副教授 2005 年组织的大规模进城农民状况的调查。武汉市位于中国中部地区,是人口流动的重要目的地,集中了大量进城农民,而且,武汉市重、轻工业和服务业都较为发达,进城农民行业、区域分布均匀,这都使武汉在进城农民研究方面具有较好的代表性。为了判定进城农民劳动收入与其教育和经验的函数关系模型,本研究根据“自身因素”、“教育与经验”、“就业方式”三个方面,选取了包括收入在内的 10 变量。

表 1 变量及缩写

应变量	劳动收入	进城农民月收入
	教育、经验	文化程度、外出打工时间
自变量	自身因素	性别、年龄、婚姻、来源地
		就业单位、从事行业、就业性质

三 经验分析结果

1 描述分析

由表 2 可知,第二代农民工的收入普遍偏低,主要集中在 1000 元以下,月收入少于 1000 元的农民工占本次调查样本总量的 66.6%,月收入能达到 2000 元及其以上者仅有 14 人、占样本总量的 4.4%。进城务工的第二代农民工中男性占大多数为 75.7%、女性仅占 24.3%,这与我国农村“男主外,女主内”的思想是相符合的,说明传统思想仍在影响年轻一代农村青年外出就业;第二代农民工年龄主要集中在 20—30 岁,这是外出务工的黄金年龄段;第二代农民工中已婚人数比例为 62.5%,这与我国农村普遍早婚、婚后外出务工这一传统相符合;超过半数农民工来自省内,表明地域因素仍对进城农民目标城市的选择发挥影响;

进城农民工文化程度普遍偏低，有 85.5%的农民工文化程度为初中及初中以下，文化程度普遍偏低应是造成农民工工作不稳定、收入低、难以与市民竞争的重要原因；外出打工时间大多不超过 5 年，并且随外出打工时间增长，人数急剧减少；“从业单位性质”“从事行业性质”和“就业方式”三项反映了第二代农民工的就业分布状况，三项均显示大约半数左右的进城农民没有稳定工作，多是就业于个体、私营等单位，多从事相对不稳定的餐饮、卫生等服务行业。

表 2 调查对象一般分布描述

变量类型	变量名	变量类型	变量取值 (均值)	频 数	分布 (%)
因 变 量	劳动收入 (元/月) N=317	连续变量	1070.940	—	—
	劳动收入分组 (元/月) N=317	分类变量	0—1000	211	66.6
			1000—2000	92	29.0
			2000以上	14	4.4
			2000以上	14	4.4
	性别 N=317	虚拟变量	男	240	75.7
			女	77	24.3
	年龄 (岁) N=317	连续变量	24.02	—	—
	年龄 (岁) N=317	序列变量	0—20	72	22.7
			20—30	245	77.3
自 变 量	婚姻状况 N=317	分类变量	已婚	198	62.5
			未婚	119	37.5
	来源地 N=317	分类变量	湖北省内	162	51.1
			湖北省外	155	48.9
	文化水平 N=316	序列变量	文盲	5	1.6
			小学	28	8.9
			初中	237	75.0
			高中及中专	25	7.9
			大专及以上	21	6.6
	外出打工时间 (年) N=317	连续变量	4.533	—	—
			0—5	218	68.8
	外出打工时间 (年) N=317	序列变量	5—10	89	28.1
			10—15	9	2.8
			15—20	1	0.3
	从事单位性质 N=316	分类变量	国有、集体、外资	78	24.7
			个体、私营、其他	238	75.3
	从事行业性质 N=317	分类变量	制造业、建筑业	134	42.3
			商业、餐饮、卫生	183	57.7
	就业方式 N=317	分类变量	雇佣型	171	53.9
			自办摊点、打零工等	146	46.1

2 第二代农民工收入决定因素：基于 OLS 回归方法

基于扩展的明瑟回归方程，本文使用最小二乘法(OLS)线性回归对第二代农

民工劳动收入决定因素进行分析。由表 3 可以看出，性别对第二代农民工收入影响很大，且影响是正向的，根据检验结果性别因素影响显著性通过检验；年龄对第二代农民工的收入也有正向影响，这是因为第二代农民工的年龄主要集中于 20—30 岁之间，随着年龄的增长他们的劳动经验和劳动能力都在不断增强，根据检验结果年龄因素影响显著($\text{sig}=0.002<0.1$)通过检验；婚姻对第二代农民工收入也有很大影响，影响为正，通过检验($\text{sig}=0.078<0.1$)，这可能是由于已婚的第二代农民工背后有了家庭的支持工作起来效率会更高，另外已婚的农民工年龄一般比未婚的偏大，不排除年龄因素在背后起了作用。文化程度对第二代农民工收入的影响最为明显，影响为正，通过检验，由此可见文化程度高低依然是影响第二代农民工收入多少的重要决定因素之一；就业单位性质对其收入影响较大，且影响为负，但未通过检验，另外来源地、外出打工时间从事行业以及就业单位性质均未通过检验。

表 3 第二代农民工收入决定因素回归表

		B	Std. Error	Beta	t	Sig.
	(Constant)	-597.586	309.449	—	1.931	0.054
自身因素	性 别	178.033	90.439	0.118	1.969	0.050
	年 龄	39.093	12.777	0.236	3.060	0.002
	婚 姻	167.106	94.399	0.125	1.770	0.078
	来 源 地	27.638	70.285	0.021	0.393	0.694
教育与经验	文化程度	180.186	50.332	0.195	3.580	0.000
	外出打工时间	-6.686	13.544	-0.032	-0.494	0.622
就业方式	就业单位性质	-110.723	84.979	-0.074	-1.303	0.194
	从事行业	84.790	82.150	0.065	1.032	0.303
	就业性质	-30.823	73.668	-0.024	-0.418	0.676

注：F=6.067 Sig.=0.000 Adjusted R Square=0.727

综上所述，只有性别、年龄、婚姻、文化程度这四个因素影响显著，通过了检验，显著影响第二代农民工劳动收入。

3 Multinomial Regression 分析结果

为了进一步验证 OLS 回归结果的有效性和稳定性，本文进一步使用

Multinomial Regression 结果加以验证。将第二代农民工月劳动收入分为三等，0~1000 元、1000~2000 元、2000 元以上三个水平，并将部分自变量由连续变量转成分类变量。回归结果见表 4 和表 5。表 4 和表 5 结果与上文中表 3 结果基本一致。由表 4 可知，模型有效、通过了检验，并且 Cox and Snell $R^2=0.473$ ，说明模型拟合的较好。由表 5 可以发现，性别、婚姻、文化程度、就业单位性质 4 个自变量的偏回归系数的显著水平小于 0.1，说明上述 4 个自变量对第二代农民工收入有显著影响。

表 4 模型的适应性检验

Model	Model Fitting Criteria (-2 Log Likelihood)	Likelihood Ratio Tests		
		Chi-Square	df	Sig.
Intercept Only	356.023			
Final	296.034	59.989	28.000	0.000

注：Cox and Snell $R^2=0.473$ ，Nagelkerke $R^2=0.221$

表 5 偏回归系数及其检验

Effect		Model Fitting Criteria		Likelihood Ratio Tests	
		-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
	Intercept	296.034	0.000	0.000	—
	性 别	303.311	7.277	2.000	0.026
	年 龄（分组）	297.986	1.952	2.000	0.37
自身因素	婚姻	306.912	10.878	2.000	0.004
	来源地	296.592	0.558	2.000	0.757
教育与经验	文化程度	315.011	18.977	8.000	0.015
	外出打工时间（分组）	299.854	3.820	6.000	0.701
	就业单位性质	302.736	6.702	2.000	0.035
就业方式	从事行业	300.477	4.443	2.000	0.108
	就业性质	298.410	2.376	2.000	0.305

四 研究结论与若干思考

1 研究结论

(1)教育是重要的人力资本，经检验显著影响第二代农民工的劳动收入。

(2)性别对于农民工而言往往意味着劳动能力差异，性别显著，表明性别人力资本也对第二代农民工收入产生影响。

(3)OLS 检验结果表明，年龄对第二代农民工劳动收入有影响。但是，Multinomial 检验没有通过，这表明年龄对第二代农民工收入的影响并不稳定。这可以理解，因为第二代农民工年龄大都集中在 20—30 岁，年龄较为集中，不够分散，年龄对其劳动收入的影响在一个“较短的区间内”显得不够稳定。

(4)就业方式对第二代农民工劳动收入有影响，但是其影响并不稳定。Multinomial 检验通过了，而 OLS 检验没有通过，这可能是因为分析方法不同所致，OLS 对连续自变量较为敏感，而 Multinomial 对分类或定序变量较为敏感。

(5)结合调查中的发现，可以基本肯定，工作单位的性质对农民工的劳动收入产生影响。

2 若干思考

结合上文分析结果及研究结论，本文主要提出以下几点建议：

(1)加强教育培训，提高第二代农民工的综合能力与素质

我们的分析结果表明，第二代农民工的文化程度对其劳动收入影响非常显著。因此必须加强对第二代农民工的教育和培训，提高其文化素质和职业技能。就目前条件来看，城市对农民工的基础教育投入和职业培训投入远远难以满足其发展的需要。针对这种情况，农业、劳动保障、教育、科技、建设、财政等部门从 2003 年 4 月份开始调研的基础上已共同制定了农民工的培训规划，国务院办公厅已于 2003 年 9 月 18 日批准并转发了 6 部门共同制定的《2003—2010 年全国农民工培训规划》，确定目标为：2003—2005 年和 2006—2010 年，分别

对拟向非农业和城镇转移的 100 万和 500 万农村劳动力开展转移就业前的引导性培训，并对其中的 500 万人和 3000 万人开展职业技能培训，对已进入非农产业就业的 5000 万和 2 亿多农民工开展岗前培训。针对政策需要，各级地方政府应当建立起农民工教育与培训的专项基金，并为其提供相应的设施设备和师资力量。

(2)各级政府要出台政策对农民工创业予以扶持

各级政府应制定相应政策，鼓励农民工创业，尤其应对农民工创办投资小、效益好、服务性强的小企业(摊点)予以政策和资金上的扶持，以创业促进就业。同时，鼓励有条件的进城农民工做强、做大企业。建立由地方政府主导的信用担保体系，完善农民工小额贷款的金融服务体系建设。当然，农民工贷款、融资体系建立仍需更多科学论证和合理建议，这将是一个长期过程。

(3)消除劳动力市场的性别歧视

前文分析结果表明，性别因素显著影响第二代农民工的劳动收入。因此，政府应当完善相关法律法规，惩罚用人单位在招聘、待遇方面的性别歧视，使男女农民工在劳动力市场上拥有平等的就业机会，其从事同等工作的工资待遇水平不会因性别因素而出现差异，努力构建一个公正、平等、和谐的劳动力市场。

参考文献：

[1]王兴周.新生代农民工的群体特性探析——以珠江三角洲为例[J].广西民族大学学报，2008(4).

[2]张宝庆.新生代农民工流动研究——来自山东潍坊的调查与思考[JD].中国农业大学农村与区域发展专业，2005(5).

- [3]吴红宇, 谢国强.新生代农民工的特征、利益诉求及角色变迁——基于东莞塘厦镇的调查分析[J].南方人口, 2006(2).
- [4] 许传新.落地未生根——新生代农民工城市社会适应研究[J].南方人口, 2007(4).
- [5]许传新, 许若兰.新生代农民工与城市居民社会距离实证研究[J].人口与经济, 2007(5).
- [6]邹英.新生代农民工自我身份认同困境的社会学分析——以长春市为例[D].吉林大学社会学专业, 2007(4).
- [7]杜书云, 张广宇.农民工代际差异问题调查与思考[J].农村经济, 2008(2).
- [8]成艾华, 姚上海.农民工的代际差异分析[J].统计与决策, 200(20).
- [9]王东, 秦伟.农民工代际差异研究——成都市在城农民工分层比较[J].人口研究, 2002(5).
- [10]傅丕毅, 郭奔胜, 李亚彪, 钟玉明.流动人口:政治边缘人[J].浙江人大, 2005(5).
- [11]舒尔茨.人力资本投资[M].北京:商务印书馆.1990年版.第48页.
- [12]张车伟.人力资本回报率变化与收入差距:“马太效应”及其政策涵义[J].经济研究, 2006(12).
- [13]刘建荣.新生代农民工的制度保障诉求.学习论坛.2008(7).