

农民工培训的收入效应[※]

张世伟¹ 武 娜²

[内容摘要] 依据进城务工人员调查数据,应用系统性样本选择方法对样本进行了筛选,并采用 Horvitz-Thompson 估计量分析了培训对农民工收入的影响。研究表明,农民工接受职前培训和在职培训将导致其收入分别提高 18.38% 和 9.45%,既接受职前培训又接受在职培训将导致其收入提高 20.24%。因此,政府和企业发展针对农民工的培训项目,将有助于农民工收入水平的提高和农村剩余劳动力向城镇有序流动。

[关键词] 培训;收入;农民工;倾向分

一、引 言

20 世纪 80 年代以来,随着中国经济体制改革的持续推进,大量农村剩余劳动力向城镇流动,成为城镇劳动力市场中的重要组成部分。然而,由于农民工人力资本水平较低,通常只能从事技术含量较低的工作,导致其收入水平明显偏低。为了提高农民工的人力资本水平,中国许多地方政府和企业陆续开展了各种形式的职前培训和在职培训。培训能否显著地提高农民工的收入水平?职前培训和在职培训对农民工收入的影响程度如何?研究这些问题不仅有助于我们对农民工工资决定机制的理解,而且有助于政府部门培训政策的评价和设计。

尽管经济理论认为,培训作为人力资本投资的重要形式,有助于个体收入水平的提高。然而,培训的实际作用效果通常需要通过经验研究途径加以确定。^[1]目前,关于培训收入效应的经验研究方法主要分为三类:(1)回归方法,该方法将培训作为一个解释变量直接加入收入方程,通过观察培训的回归系数判断培训对收入的贡献。然而,由于个体间相互作用和宏观运行趋势的干扰,回归方法估

※ 本文得到教育部人文社会科学重点研究基地重大项目(12JJD790042)的资助。

作者简介:张世伟(1964—),男,吉林大学数量经济研究中心(长春,130012),教授。研究方向:数量经济学。

武 娜(1988—),女,吉林大学数量经济研究中心(长春,130012),博士生。研究方向:数量经济学。

计结果通常存在外部和内部有效性问题。^{[2](151-161)} (2) 自然实验法, 该方法将培训作为一个自然实验, 通过比较实验组和对照组在培训前后收入的变化确定培训的收入效应。尽管自然实验法能够有效地解决估计结果的外部性和内部有效性问题, 但其对实验前后面板数据的要求使得该方法的应用受到限制。(3) 倾向分匹配法, 该方法根据倾向分对实验组和对照组中相似个体进行匹配, 通过比较匹配成功个体的收入差异度量培训的收入效应。由于倾向分匹配法具有“准自然实验”性质, 同时其对数据要求不高(截面数据), 因而许多经济学者用该方法研究培训的收入效应。^{[3](1053-1062)[4](605-654)[5](74-90)}然而, 倾向分匹配方法却常常会带来估计结果有效性的损失。^{[6](315-331)}

Crump 等人^{[7](187-199)}指出, 在估计平均处理效应时, 常常由于处理组和对照组之间解释变量的分布缺少重叠而使得估计结果不精确, 并且可能使得一些常用的估计量敏感于模型设定。他们提出了一种系统性样本选择方法, 即首先应用倾向分匹配方法选择最优子样本以解决解释变量分布的有限重叠问题, 然后应用一些标准方法对平均处理效应进行估计和推断。他们采用该方法对 Connors 等人^{[8](889-897)}的数据进行分析, 发现通过减少 18% 的样本量, 计算结果方差降低了 36%。

由于微观数据匮乏, 关于农民工培训的收入效应研究起步较晚。王德文等^{[9](1131-1148)}将农民工参加培训视为一种自愿参与的“准自然实验”, 采用平均处理效应模型对农民工培训的收入效应进行了估计。张世伟和王广慧^{[10](34-38)}、孟宪生等^{[11](52-55)}分别应用倾向分匹配方法研究了培训对农民工收入的影响。然而, 上述研究并未对样本进行系统化筛选, 研究结果势必存在一定程度的偏差。

综上所述, 基于 2007 年吉林省进城务工人员调查数据, 本文拟应用 Crump 等的系统性样本选择方法筛选样本, 并应用 Horvitz - Thompson 估计量分析培训对农民工收入的影响。本文的第二部分论述分析方法, 第三部分对数据进行统计描述, 第四部分对回归结果进行分析, 最后给出本文的研究结论。

二、研究方法

根据不同种类的培训, 本文首先从总体样本中抽取两个子样本, 实验组和对照组。其中, 实验组由接受该类培训的个体组成, 对照组由未接受培训的个体组成。个体接受该类培训的概率(倾向分)可以表示为一个 Logistic 模型:

$$\Pr(T_i = 1 | X_i = x) = e(x) = \frac{\exp(X_i' \gamma)}{1 + \exp(X_i' \gamma)} \quad (1)$$

其中, T_i 表示个体是否接受该类培训的二元指示变量(1: 接受培训, 0: 未接受培训), X_i 为影响个体是否接受培训的特征向量, γ 为待估计的系数向量。

为了确保每个接受培训的个体均存在一些具有类似特征的未接受培训的个体

与之匹配，即实验组和对照组之间个体特征存在充分重叠，需要对实验组和对照组样本进行筛选。依据 Crump 等的系统性样本选择方法，筛选的样本倾向分在 $[\alpha, 1-\alpha]$ 区间内，其中最优截断值 α 由倾向分的边际分布决定，即 α 是下式的解：

$$\frac{1}{\alpha(1-\alpha)} = 2E\left[\frac{1}{e(x)\{1-e(x)\}} \mid \frac{1}{e(x)\{1-e(x)\}} \leq \frac{1}{\alpha(1-\alpha)}\right] \quad (2)$$

根据倾向分的估计值 $\hat{e}(x)$ ，可以求得最优截断值的估计值 $\hat{\alpha}$ ($\hat{\alpha} \in [0, 1/2]$)，使其满足：

$$\frac{1}{\alpha(1-\alpha)} \leq 2 \frac{\sum_{i=1}^N [1_{\hat{e}(X_i)\{1-\hat{e}(X_i)\} \geq \alpha(1-\alpha)} / \hat{e}(X_i)\{1-\hat{e}(X_i)\}]}{\sum_{i=1}^N 1_{\hat{e}(X_i)\{1-\hat{e}(X_i)\} \geq \alpha(1-\alpha)}}$$

通过比较筛选后的实验组和对照组收入差异即可确定培训的收入效应 $\hat{\tau}$ (Horvitz - Thompson 估计量)：

$$\hat{\tau} = \frac{\sum_{i=1}^N \frac{T_i Y_i}{\hat{e}(X_i)}}{\sum_{i=1}^N \frac{T_i}{\hat{e}(X_i)}} - \frac{\sum_{i=1}^N \frac{(1-T_i) Y_i}{1-\hat{e}(X_i)}}{\sum_{i=1}^N \frac{1-T_i}{1-\hat{e}(X_i)}} \quad (3)$$

其中， Y_i 为农民工的月收入对数。由于无法得到 (3) 式中收入效应 $\hat{\tau}$ 标准差的具体表达形式，本文采用自举 (Bootstrap) 方法^[12]求解：从原始数据中重复地随机抽取样本规模为 n 的自举样本，计算该自举样本培训的收入效应 $\hat{\tau}$ ，将上述步骤重复进行 B 次，得到 B 个统计量 ($\hat{\tau}_1, \hat{\tau}_2, \dots, \hat{\tau}_B$)，计算 B 个统计量的标准差，即可得到原始数据 $\hat{\tau}$ 统计量的标准差。

根据经济理论和经济现实，教育作为人力资本的内核，通常会对个体接受培训产生重要影响；在不同的年龄段，个体接受培训的倾向会存在显著差异；由于男女在家庭中分工各有侧重，男性和女性接受培训的倾向会存在差异；许多企业在签订劳动合同时，会考虑到个体是否接受过培训；外出务工时间通常会对个体接受培训的倾向产生影响；为了获得技术等级证书，个体通常需要接受培训；此外，培训的举办组织可能会对个体接受培训有所影响。因此，本文将受教育程度 (虚拟变量，以小学及以下为参照组)、年龄、性别 (虚拟变量，以女性作为参照组)、是否签订劳动合同 (虚拟变量，以未签订劳动合同作为参照组)、外出务工时间 (虚拟变量，以一年以内作为参照组)、技术等级 (虚拟变量，以无技术等级为参照组) 和培训组织 (虚拟变量，以其他组织为参照组) 作为培训接受方程的解释变量。

三、数据描述

本文使用的数据来自 2007 年吉林省进城务工人员调查数据，调查内容涉及个体性别、年龄、来源地、受教育程度等人口统计学信息和收入、培训 (包括职

前培训和在职培训)、外出务工时间、劳动合同、技术等级等就业信息。本文关注职前培训和在职培训对农民工收入的影响,因此将农民工群体划分为四个子群体,即未接受培训群体、接受职前培训群体、接受在职培训群体和既接受职前培训又接受在职培训(简称职前与在职培训)群体,样本容量分别为2159、463、1128和1993。

表1给出了四个农民工群体收入和个体特征的统计描述。整体来看,62%左右的农民工接受过培训。农民工月平均收入在710元至947元之间,说明农民工收入水平较低。农民工平均年龄32岁左右,说明农民工主要为青壮年。农民工平均受教育年限不到9年,说明许多农民工未完成九年义务教育,受教育程度较低。农民工中男性占大多数,说明一部分农民工单独进城务工。农民工主要来自本省,符合劳动力流动理论预期。一半左右农民工外出务工时间一年以内,只有不到30%的农民工外出务工时间超过两年,说明许多农民工只是短期在城市务工。只有一半左右农民工签订了劳动合同,主要源于一方面其务工时间较短,另一方面说明企业未能认真执行劳动合同法。一半左右农民工不具有任何技术等级,而高级工仅占农民工总体的10%以下,说明农民工平均技能水平较低。

表1 不同群体农民工收入和个体特征的统计描述

个体特征	未培训	职前培训	在职培训	职前在职培训
月收入(元)	710.63 (7.00)	851.92 (19.23)	761.22 (9.50)	947.25 (9.09)
年龄(岁)	33.81 (0.20)	32.86 (0.39)	30.27 (0.26)	31.90 (0.19)
受教育年限(年)	8.65 (0.05)	8.87 (0.10)	8.94 (0.07)	9.10 (0.05)
小学及以下(%)	14.03 (0.01)	11.45 (0.01)	13.21 (0.01)	9.13 (0.01)
初中(%)	54.38 (0.01)	51.62 (0.02)	48.94 (0.01)	54.54 (0.01)
高中(%)	22.05 (0.01)	21.81 (0.02)	25.00 (0.01)	18.06 (0.01)
中专技校职高(%)	5.33 (0.00)	11.66 (0.01)	6.21 (0.01)	8.63 (0.01)
大专及以上(%)	4.21 (0.00)	3.46 (0.01)	6.65 (0.01)	9.63 (0.01)
男性(%)	60.95 (0.01)	66.31 (0.02)	61.61 (0.01)	71.85 (0.01)
来源于本省(%)	76.98 (0.01)	76.67 (0.02)	82.27 (0.01)	71.55 (0.01)
务工一年以内(%)	54.24 (0.01)	42.98 (0.02)	47.96 (0.01)	51.83 (0.01)
务工一年至两年(%)	21.26 (0.01)	32.18 (0.02)	25.80 (0.01)	19.77 (0.01)
务工两年及以上(%)	24.50 (0.01)	24.84 (0.02)	26.24 (0.01)	28.40 (0.01)
签订劳动合同(%)	33.95 (0.01)	56.59 (0.02)	67.55 (0.01)	76.72 (0.01)
无技术等级(%)	62.11 (0.01)	30.24 (0.02)	47.25 (0.01)	20.42 (0.01)
初级工(%)	20.15 (0.01)	22.68 (0.02)	28.46 (0.01)	25.09 (0.01)
中级工(%)	13.85 (0.01)	35.21 (0.02)	17.55 (0.01)	43.50 (0.01)
高级工(%)	3.89 (0.00)	11.88 (0.02)	6.74 (0.01)	10.99 (0.01)
样本量	2159	463	1128	1993

注:()内为标准差。

由于不同群体间个体特征存在明显的差异，必须应用微观经济计量模型对个体特征异质性加以控制，才能得到培训对农民工收入影响的准确度量。

四、回归结果分析

根据本文的研究思路，首先针对农民工接受培训的三种情形，分别应用极大似然估计方法对（1）式进行回归，分析影响农民工接受职前培训、在职培训和职前与在职培训决策的因素（参见表2）。从回归结果中可以发现，年龄对接受三类培训决策的影响均显著为负，说明随着年龄的增长，农民工接受培训的意愿逐渐减弱，主要源于随着年龄的增长，人力资本投资回报逐渐降低。与女性相比，男性农民工更倾向于接受职前培训，主要源于男性在就业方面具有比较优势，培训有助于其就业。与外省农民工相比，来源于本省的农民工更倾向于接受在职培训，主要源于就业与迁移距离负相关，而在职培训有助于其长期就业。与“小学及以下”的学历水平相比，具有较高学历水平的农民工更愿意接受职前培训，主要源于教育作为人力资本的内核，有助于个体接受培训。具有中专（技校职高）及以上学历的农民工接受职前与在职培训概率明显较高，说明接受过专业化教育有助于其接受培训。农民工与用人单位签订劳动合同对接受在职培训和接受职前与在职培训均具有显著的正向影响，并且对接受职前与在职培训的影响更大。相对于其他职前培训的组织者，职前培训由“村里”或“县里”组织对农民工接受职前培训和接受职前与在职培训决策的影响均显著为正，且职前培训组织者的行政级别越高，对农民工接受培训决策的正向影响越大，可能主要源于组织者水平与培训效果正相关。与“一年以内”的外出务工时间相比，较长的外出务工时间对农民工接受在职培训和接受职前与在职培训存在显著的正向影响。相对于“无技术等级”，具有技术等级对农民工接受在职培训和接受职前与在职培训的决策均具有显著的正向影响，且随着技术等级的提高，农民工接受两类培训的意愿大致呈现逐渐上升的趋势。总的来看，回归结果基本符合经济现实和经济理论预期。

表2 Logistic模型的回归结果

解释变量	职前培训	在职培训	职前与在职培训
常数项	-1.8632***	-0.6829***	-1.7756***
年龄	-0.0166***	-0.0473***	-0.0314***
男性	0.3370***	-0.0396	0.0714
来源于本省	0.0088	0.5454***	0.0787
初中	0.2971*	-0.1319	0.1996*
高中	0.3636**	0.0998	0.2044
中专技校职高	1.1322***	-0.0564	0.8047***

大专及以上	0.3858	0.2993	1.2985***
签订劳动合同	-	1.3488***	1.6093***
村里组织培训	1.2894***	-	1.3001***
县里组织培训	1.4916***	-	1.5474***
务工一年至两年	-	0.4101***	0.0428
务工两年及以上	-	0.3641***	0.3553***
初级工	-	0.7114***	1.1377***
中级工	-	0.5146***	1.8791***
高级工	-	0.8433***	1.8487***
Log likelihood	-1140.2667	-1840.4283	-1975.2640
Prob > chi2	0.0000	0.0000	0.0000
Pseudo R2	0.0671	0.1294	0.3129
样本量	2622	3287	4152

注：***、**和*分别表示在1%、5%和10%水平下显著。

其次，根据表2的回归结果计算相应子群体中每个个体的倾向分，并依据(2)式计算最优截断值 $\hat{\alpha}$ ，进而选择倾向分在区间 $[\hat{\alpha}, 1 - \hat{\alpha}]$ 之内的样本。表3给出了通过计算得到的 $\hat{\alpha}$ 值和进行系统性样本选择之后的样本量，可以发现每一组样本量均有所减少，尤其是接受职前与在职培训群体和未接受培训群体经过样本选择后，样本量分别减少540个和385个。通过系统性样本选择，实验组和对照组样本个体属性存在比较充分的重叠，从而能够提高分析的精度。

表3 样本选择结果

样 本	职前培训		在职培训		职前与在职培训	
	实验组	对照组	实验组	对照组	实验组	对照组
全样本	463	2159	1128	2159	1993	2159
选择样本	463	2154	1113	1960	1608	1619
$[\hat{\alpha}, 1 - \hat{\alpha}]$	[0.0648, 0.9352]		[0.1022, 0.8978]		[0.0889, 0.9111]	

最后，基于系统性选择后的样本，根据(3)式分别计算农民工职前培训、在职培训、职前与在职培训的收入效应(参见表4)。从计算结果中可以发现，三种培训收入效应的计算结果均显著为正，说明培训可以显著地提高农民工的收入水平。其中，职前培训、在职培训和职前与在职培训将导致农民工收入分别提高18.38%、9.45%和20.24%，说明职前培训对农民工的增收效果优于在职培训，农民工在接受职前培训的基础上继续接受在职培训仍能显著地提高其收入水平，但边际回报递减。由于农民工培训周期通常较短，培训收益率明显高于教育收益率，因此培训是农民工增收的有效途径。

表 4 农民工培训收入效应的估计结果

培训的收入效应	系统选择样本		全样本	
	估计值	标准差	估计值	标准差
职前培训	0.1838	0.0279	0.1840	0.0288
在职培训	0.0945	0.0224	0.0936	0.0268
职前与在职培训	0.2024	0.0207	0.1618	0.0261

注：标准差采用 Bootstrap 方法重复抽样 500 次得到。

出于比较的目的，表 4 还给出了应用全样本的估计结果。可以发现，基于全样本的职前培训和在职培训收入效应的估计结果与基于系统化样本选择的估计结果比较接近，但通过系统化样本选择，标准差明显降低，进而提高了计算结果的精度。此外，需要引起注意的是，在全样本的分析中，职前与在职培训的估计结果明显偏低，甚至低于职前培训的估计结果，即在在职培训的基础上接受在职培训将导致收入降低，这显然是无法解释的，主要源于实验组和对照组中存在许多个体特征差异较大的个体，对估计结果产生较大干扰。通过系统化样本选择，删除个体特征差异较大的个体，可以比较准确地估计职前与在职培训的收入效应。

五、结 论

依据 2007 年吉林省进城务工人员调查数据，本文应用系统性样本选择方法对样本进行了筛选，并应用 Horvitz - Thompson 估计量计算了培训对农民工收入的影响。研究表明，农民工接受职前培训和在职培训将导致其收入分别提高 18.38% 和 9.45%，而既接受职前培训又接受在职培训将导致其收入提高 20.24%。因此，培训是提高农民工收入的有效途径。

农民工收入偏低的一个主要原因是其技能水平较低。从统计数据中可以发现，57% 左右的农民工在进城务工以前未经过培训，38% 左右的农民工未接受过任何培训。同时，从回归结果中可以发现，培训不仅有助于农民工提高收入水平，而且有助于农民工获得技术等级证书，有助于农民工稳定与企业的劳动关系。因此，政府部门加大培训投入，扩大培训覆盖面，将有助于农民工技能水平的提高，进而有助于农民工收入水平的提高，有助于农村剩余劳动力向城镇有序流动。

农民工收入偏低的另一个主要原因是其知识水平较低，尚存在许多农民工未完成 9 年义务教育。同时，从回归结果中可以发现，教育作为人力资本的内核，有助于农民工接受培训。因此，政府加大农村基础教育的投入，将有助于农民工接受各种职前培训和在职培训，有助于农民工收入水平更大程度的提高，进而有助于农村剩余劳动力向城镇有序转移。

最后，如何消除异常数据的干扰，提高估计结果的准确度，一直是计量经济

学方法及其应用的一个难题。通过比较系统性选择样本和全样本的估计结果可以发现,经过系统性样本选择,不仅降低了农民工培训收入效应的标准差,而且提高了估计培训收入效应的准确度。在许多经济问题研究过程中,尤其是在群体间绩效比较和政策评价过程中,对样本进行系统化筛选,将会有效地降低估计结果的标准差,提高估计结果的准确度。☆

主要参考文献:

- [1] 卡赫克,齐尔贝尔博格. 劳动经济学 [M]. 上海:上海财经大学出版社,2007.
- [2] Meyer B. Natural and Quasi - Experiments in Economics [J]. Journal of Business and Economic Statistics, 1995, 13 (2).
- [3] Dehejia R, Wahba S. Causal Effects in Non - Experimental Studies: Re - Evaluating the Evaluation of Training Programs [J]. Journal of American Statistical Association, 1999, 94 (448).
- [4] Heckman J, Ichimura H, Todd P. Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program [J]. Review of Economic Studies, 1997, 64 (4).
- [5] Lechner M. Earnings and Employment Effects of Continuous Off - the - Job Training in East Germany after Unification [J]. Journal of Business and Economic Statistics, 1999, 17 (1).
- [6] Hahn J. On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects [J]. Econometrica, 1998, 66 (2).
- [7] Crump R, Hotz V, Imbens G. Dealing with Limited Overlap in Estimation of Average Treatment Effects [J]. Biometrika, 2009, 96 (1).
- [8] Connors A, Speroff T, Dawson N. The Effectiveness of Right Heart Catheterization in The Initial Care of Critically Ill Patients [J]. Journal of American Medical Association, 1996, 276 (11).
- [9] 王德文,蔡 昉,张国庆. 农村迁移劳动力就业与工资决定:教育与培训的重要性 [J]. 经济学家, 2008 (4).
- [10] 张世伟,王广慧. 培训对农民工收入的影响 [J]. 人口与经济, 2010 (1).
- [11] 孟宪生,关凤利,唐哲一. 农民工参与就业培训的决定因素及对收入影响的实证分析 [J]. 东北师大学报 (哲学社会科学版), 2011 (4).
- [12] Efron B, Tibshirani R. An introduction to the Bootstrap [M]. Chapman & Hall, 1993.

Income Effect of Training on Migrant Workers

Zhang Shiwei¹ Wu Na²

Abstract: With the survey data of migrant workers, the paper selects sample using a systematic sample selection method, and analyzes the income effect of training on migrant workers using Horvitz - Thompson estimator. The results show pre - job training and on - the - job training improve the income of migrant workers by 18. 38% and 9. 45% respectively, migrant workers who receive both pre - job and on - the - job training will have an income increase by 20. 24%. As a result, the government and enterprises should develop training programs, which can help increasing the income level of migrant workers, and prompting the well - organized rural - urban migration of rural surplus labor force.

Key words: Training; Income; Migrant Workers; Propensity Score

[收稿日期: 2013.10.5 责任编辑: 陈健生]

[中图分类号] F064.1 [文献标识码] A [文章编号] 1000 - 8306 (2013) 12 - 0065 - 08