

非农就业对城郊农民收入的影响研究

刘 魏^{1,2}

(西南大学 1. 经济管理学院; 2. 农村经济与管理研究中心, 重庆市 400715)

摘 要:城镇化的外围式扩张导致城郊农民土地被征用, 非农就业将成为城郊农民重要的就业方式。基于中国综合社会调查(CGSS)2012年的统计数据, 运用倾向得分匹配模型, 分析了非农就业对城郊农民收入的影响。研究发现: 非农就业能够显著提高城郊农民收入水平, 非农就业的城郊农民收入比未非农就业高59%。基于不同受教育程度、不同社会地位和不同性别等异质型特征进一步研究发现, 低学历、女性人群由纯农就业转为非农就业, 其收入提升效应更为显著, 但现阶段低学历、女性人群非农就业比例却很低。因此, 鼓励城郊农民非农就业是很有必要的, 尤其要降低低学历、女性等人群的非农就业进入门槛, 使他们进入到城市服务行业, 以达到与城市高技能劳动者的行业互补和提高农业规模经济效益的目的。

关键词: 非农就业; 城郊农民; 收入效应; 学历; 女性人群; 中国综合社会调查

中图分类号: F323.8 **文献标识码:** A **文章编号:** 1673-9841(2016)05-0061-13

一、引 言

传统意义上, 农村非农经济发展被认为是低产出的, 将随着时间的变化逐渐消失, 农民收入获取的途径主要源于农业生产经营, 即通过土地、劳动力和资本的投入实现。但土地单位产出率的低下和农村信贷资本的限制, 导致剩余的劳动力集中在土地上, 农民不能通过资本投资获取新的生产技术。同时农产品价格的风险性和不确定性导致农民收入增长具有不稳定性。伴随着改革开放以后我国经济社会环境发生的变化, 农民非农就业越来越被重视, 其在农民收入增长、减少贫困和释放农村剩余劳动力等方面扮演着重要角色^[1-2]。农民收入增长不再单纯依靠农业生产, 而更多地依靠非农生产活动。据统计, 2015年底我国农民工总量达到2.77亿, 比上年增长1.3%, 其中外出农民工1.69亿, 增长0.4%^①。可以说, 农村经济中非农业对收入贡献的比例越来越大, 不仅提高了农民收入水平, 也为城市经济贡献了劳动力。

与此同时, 劳动力的流动促使城镇化快速发展, 城镇化的发展将掀起新一轮的非农就业浪潮^[3]。城镇化推进过程中, 城镇规模的外延式扩张会导致城郊失地农民的增加, 城郊农民生产行为将发生变化, 必将进一步影响其收入水平和收入结构^[4]。一方面, 城郊农民在城镇化发展过程中由于面临失去土地风险而选择非农就业, 另一方面, 城郊农民在传统农业生产过程中面临过密化^②导致的土地单位产出率低、农村资本匮乏、农业收入增长不稳定等供给侧问题, 会主动选择非农就业。

① 数据来源于国家统计局发布的《2015年国民经济和社会发展统计公报》。

② 黄宗智认为, 过密化是指经济在以单位工作日边际报酬递减为代价的条件下扩展, 即所谓的“没有发展的增长”。

收稿日期: 2016-02-10

作者简介: 刘魏, 西南大学经济管理学院, 博士研究生。

通讯作者: 张应良, 管理学博士, 西南大学经济管理学院, 教授, 博士生导师。

基金项目: 国家社会科学基金重大项目“三权分置、农地流转与农民承包权益保护”(15ZDA023), 项目负责人: 张应良。

因此, 非农就业成了城郊农民收入增长的关键。随着工业化、城镇化和农业现代化的发展, 城郊农民正面临收入结构的变化。非农就业不仅能够提高收入水平, 还能够为妇女和穷人创造更多的就业岗位^[5]。当前, 我国正在进行“供给侧改革”。劳动力作为供给侧一个重要的增长因素, 劳动力的优化配置将对供给侧改革产生深远影响。习近平同志在中央财经领导小组第 11 次会议和亚太经合组织 APEC 工商领导人峰会上两次强调, 要“加强供给侧结构性改革”。李克强同志在“十三五”规划纲要编制工作会议和 2015 年《政府工作报告》上指出, “要在供给侧和需求侧两端发力, 促进产业迈向中高端”。对于城镇化发展来说, 处理好城郊农民的非农就业问题无疑是供给侧改革的重要内容。

事实上, 许多学者认为提高收入的关键在于非农就业^[6]。但由于规避风险的偏好, 城郊农民中从事非农就业工作的通常是被征地或者受教育程度较高、具有非农工作经历的人群, 知识和思维结构、社会网络的变化使这部分人群更倾向于非农就业。Emran & Hou 指出穷人参与非农活动会面临进入壁垒^[7]。Cherdchuchai & Otsuka, Kijima, Matsumoto & Yamano, Lanjouw & Murgai 也认为从事非农活动需要娴熟的技能或者较高的受教育程度^[8-10]。那么, 非农就业对城郊农民收入的影响, 到底来自于非农就业的直接影响? 还是诸如受教育程度、性别、社会地位等异质性特征带来的间接影响? 这是本文需要甄别的一个问题。因此, 本文将运用倾向得分匹配(PSM)模型来检验非农就业对城郊农民收入的影响, 倾向得分匹配方法能够较好地剔除样本自选择所带来的内生性问题。

二、文献述评

近年来, 国内外学术界对农民的非农就业问题进行了广泛讨论和深入研究, 并取得了丰硕成果。不少学者探讨了发展中国家非农就业的特点和影响因素。Reardon et al. 认为非农就业在农村经济发展中扮演越来越重要的作用, 它为农业生产的上下游提供了产品和服务, 同时提供了较多的农村收入份额, 为粮食安全、缓解贫困、提高农业生产竞争力作出了较大贡献^[11]。Abdulai & Delgado 通过对加纳北部已婚男性和女性非农就业行为的研究, 得出教育、经验、基础设施、到首都的距离、人口密度能够显著提高劳动力非农就业的参与率, 同时非农就业工资具有性别差异^[12]。Barrett 发现非农就业与家庭福利存在正相关关系, 非农收入的多元化有利于刺激消费的快速增长^[13]。Canagarajah 发现自我雇佣方式的非农就业会导致收入不平等程度的加剧, 对于以女性为户主的家庭更甚, 而以工资为主要方式的非农就业可以减小收入不平等程度^[14]。此外, 他还发现地理位置、受教育程度、年龄和距离市场远近是影响非农就业的关键因素。Lanjouw 认为穷人由于受教育程度较低, 其参与非农就业活动的劳动生产率低下^[15]。任国强探讨了人力资本对非农就业的影响, 发现受教育程度越高, 非农就业参与度越高, 除高中外, 农民受教育程度越高, 非农报酬也越高^[16]。艾春荣、汪伟发现农户非农就业与否主要与由农户的农业生产率和非农劳动单位报酬决定, 与当期收入不相关^[17]。

非农就业的进入限制也是一个值得研究的问题。Knight 指出初中及以上学历的农民更容易获得贸易、制造业等非农工作, 其子女更容易受到良好的教育, 这会形成一个良性循环, 而那些受教育程度较低的农民往往会陷入恶性循环中^[18]。同样, 营养不良的农民由于工作效率低下, 失业的可能性更高, 从而陷入贫困的恶性循环之中^[19-20]。Jha 运用印度农村的实证研究证实了营养不良导致贫困的假设^[21]。Reardon 强调贫困家庭由于非农活动进入限制, 阻碍了他们投资非农活动, 这印证了贫困陷阱的存在, 也说明贫困家庭进入非农产业工作并非是一个自然过程^[22]。

还有学者研究了非农就业与收入之间的关系问题。Taylor J.E 认为非农就业所造成的农业劳动力流失会导致农业收入的下降^[23]。Haggblade et al. 认为发展中国家非农就业是农民收入的主要来源, 贡献了 35~50% 的份额, 同时吸纳了 33% 的农村劳动力^[24]。不少学者认为非农就业活动

在经济社会中的作用越来越大^[25],增加了农民的粮食产出和收入^[26],缩小了农业家庭和非农就业家庭的收入差距^[27]。刘洪银运用丹尼森方法构建回归模型,发现非农就业能够显著促进经济增长,同时非农就业对经济增长的作用符合边际递减规律^[28]。钱文荣等区分了非农就业对收入影响的直接和间接效应,直接效应显示非农就业人数的增加会导致农业收入下滑,间接效应显示非农就业的工资汇款可以补偿直接效应所带来的负面影响^[29]。宁光杰运用2008年农村住户调查数据分析了非农就业的影响因素及非农就业对农民收入的影响问题,研究发现教育、年龄、培训、健康以及非农工作经历会显著影响农民非农就业参与度。同时还发现外出非农就业的收入显著高于本地就业^[30]。冉璐等运用2008年中国综合社会调查数据,分析了非农就业经历对务农收入的影响,发现具有非农就业经历的农民务农收入显著高于无非农就业经历的农民^[31]。

通过对相关文献的梳理,本文可能的创新之处是:(1)已有研究内容主要集中在非农就业行为的决定因素上,较少涉及非农就业对收入影响问题。对于部分研究非农就业对农民收入影响的文献,则忽略了不可观测因素所导致的自我选择所产生的偏误,难以区分非农就业所产生的因果效应与不可观察因素所产生的效应,会导致回归估计偏误,从而给出一个错误的政策含义。因此本文利用倾向得分匹配模型(PSM),可以较好地地区分非农就业或是不可观测因素对农民收入的影响。(2)研究主体主要集中于城郊农民。现有文献主要研究农民的非农就业问题,其理论背景基于劳动力迁移理论和农业的弱质性问题,因此农民的非农就业主要考虑农业生产率和非农劳动单位报酬的高低。而城郊农民则是城镇化外围扩张所产生的一个特殊农民群体,其非农就业主要来自城镇化扩张所导致的失地问题和市民化过程积极融入城镇的自发就业过程。因此城郊农民与其他农村地区农民的非农就业动机是不一样的。(3)国外学者主要从特定国家研究非农就业问题,这与我国的城郊农民有本质不同,且研究数据不适用于国内问题研究。本文主要借助于中国综合社会调查数据,覆盖面广、代表性强。

三、理论模型

农户非农就业行为选择问题,本文借鉴了 Huffman 的理论模型^[32],Owusu et al.对该模型进行了发展^[33]。在给定物品消费量(Q)和闲暇消费量(H)的情况下,农户家庭效用函数(U)可以表示为:

$$U = U(Q, H) \quad (1)$$

同时,每个农户家庭还面临时间和资金预算约束。时间约束可以表示为:

$$T = L_1 + L_2 + H \quad (2)$$

T 表示总的的时间禀赋, L_1 和 L_2 分别表示分配给农业就业和非农就业的时间, H 表示闲暇时间。 L_1 和 L_2 满足:

$$L_1 \geq 0; L_2 \geq 0 \quad (3)$$

资金预算约束可以表示为:

$$PQ = p_1 y_1 = \omega_{L1} L_1 + \omega_2 L_2 + R \quad (4)$$

其中, P 代表市场上产品价格, ω_1 和 ω_2 分别表示农业就业、非农就业的劳动力单位时间报酬, y_1 和 p_1 分别表示农产品的销售量、销售价格, R 表示非劳动收入。

因此,农户获取最大利润的最优时间分配,可以表示为效用函数的一阶条件:

$$\partial U / \partial L_i = \omega_i \partial U / \partial Q - \partial U / \partial L = 0 \quad (5)$$

(5)式中 i 表示农业就业或非农就业,将(5)式化简,可得到农户在农业就业和非农就业的最优工资报酬:

$$\omega_i = (\partial U / \partial L) / (\partial U / \partial Q) \quad (6)$$

根据(5)式,农户分配到农业、非农就业和闲暇的最优时间,可以得农业就业和非农就业的劳动

力供给函数:

$$L_1 = L_1(\omega_1, \omega_2, p_1, p_2; Z) \quad (7)$$

$$L_2 = L_2(\omega_1, \omega_2, p_1, p_2, R; Z) \quad (8)$$

根据农户非农就业的实证研究,农户选择农业还是非农就业主要基于非农就业部门市场潜在工资(W_i^m)和保留工资^①(W_i^r)的比较,当 $W_i^m > W_i^r$,则农户选择非农就业,即 $L_i = 1$;当 $W_i^m \leq W_i^r$ 时,则农户选择农业就业,即 $L_i = 0$ 。然而现实中我们无法观察市场潜在工资和保留工资的差异,只能观察到农户就业行为的选择,因此 Huffman & Lange^[34]定义了一个包含不可观察变量(L_i^*)的指标函数:

$$\begin{aligned} L_i^* &= \beta Z_i' + \mu_i \\ L_i &= 1 \quad \text{if } L_i^* > 0 \\ L_i &= 0 \quad \text{if } L_i^* \leq 0 \end{aligned} \quad (9)$$

(9)式中, Z_i' 表示影响农户保留工资和非农就业工资的个体或家庭特征的因素矩阵, μ_i 表示随机干扰项。为了估计农户非农就业决策与收入的关系问题,我们进一步定义一个线性函数:

$$Y_i = \beta Z_i' + \gamma L_i + \epsilon_i \quad (10)$$

(10)式中, Y_i 表示农户收入, L_i 是一个二分变量,取值 1 表示非农就业,取值 0 表示农业就业, Z_i' 表示农户个体特征、家庭特征或地理、经济因素, ϵ_i 表示随机误差项。

四、研究方法、变量与数据

(一)倾向得分匹配方法

根据是否选择非农就业,本文将样本分为激励组和控制组两类。激励组即非农就业的城郊农民,控制组即未非农就业的城郊农民。前文提到,城郊农民非农就业对收入的影响效果,可能来自于受教育程度、性别、社会地位等异质性特征,而非非农就业的直接影响,因此有必要控制样本的自选择偏误。文章采用 Rosenbaum & Rubin 提出的倾向得分匹配模型(PSM)^[35],其基本思想是找到与激励组相似的控制组样本,从而降低样本自选择偏误。这种选择方法即是倾向得分(propensity score),倾向得分既可以使激励组寻找一对一的控制组样本,也可以寻找一对多的控制组样本。倾向得分定义是在样本 X 给定的条件下,城郊农民选择非农就业的条件概率,即:

$$p(Z_i) = Pr\{L_i = 1 \mid Z_i\} E\{L_i \mid Z_i\} \quad (11)$$

其中, L_i 表示城郊农民选择非农就业与否,取值 0 表示未参与非农就业,取值 1 表示参与非农就业。 Z_i 表示影响就业行为的个体特征。倾向得分 $p(Z_i)$ 可以用 logit 或者 probit 模型进行估计^[36],在已知农户的倾向得分 $p(Z_i)$ 后,通常无法寻找到得分完全一样的样本,已有文献主要通过最近邻匹配、半径匹配和核匹配等方法来达到激励组与控制组的样本匹配。这三种方法的计算公式在此不一一列举,通过 Stata 13.1 软件可以很方便地计算出。样本匹配完成后,进一步估计城郊农民选择非农就业的平均处理效果(ATT),其公式为:

$$ATT = E\{E\{Y_i^1 \mid L_i = p(Z_i)\} - E\{Y_i^0 \mid L_i = 0, p(Z_i)\} \mid L_i = 1\} \quad (12)$$

其中 Y_i^0 和 Y_i^1 分别表示同一个城郊农户未非农就业和非农就业的收入水平。

(二)变量与数据

本文所用的数据主要来自于中国综合社会调查(Chinese General Social Survey, CGSS),该数据由中国人民大学中国调查与数据中心组织调查。数据样本始于 2003 年,调查范围覆盖中国大陆,调查内容涉及中国家庭及个人的个体特征、价值观、健康、收入等。本文选取(CGSS)2012 年的调查数据,数据中包含 s5a 和 a18 两个变量,s5a 的标签含义是受访者居住的地区类型,取值 1 代表

① 保留工资是指农户将时间完全分配给农业就业和闲暇的边际价值。

市/县城的中心地区,取值 2 代表市/县城的边缘地区,取值 3 代表市/县的城乡结合部,取值 4 代表市/县城区以外的镇,取值 5 代表农村,取值 6 代表其他。a18 的标签含义是“您目前的户口登记状况”,取值 1 代表农业户口,取值 2 代表非农业户口,取值 3 代表蓝印户口,取值 4 代表居民户口(以前是农业户口),取值 5 代表居民户口(以前是非农户口),取值 6 代表军籍,取值 7 代表没有户口,取值 8 代表其他。本文研究的城郊农民收入问题,因此同时选择 s5a 取值为 2 或 3 和 a18 取值为 1 的样本。通过这两个指标可以选取城郊农民样本,同时剔除了拒绝回答、不知道、不适用的样本,最终获得 849 个样本观测值。本文主要分析非农就业与城郊农民收入问题,结合明瑟收入方程常用的收入控制因素,本文选择了性别、年龄、年龄的平方、是否少数民族、是否宗教信仰、受教育程度、是否党员、健康状况、社会地位、是否结婚等个体特征变量,各变量的描述性统计见表 1 和表 2。

表 1 变量的选取和指标定义

变量名称	variable	变量描述	指标定义
收入	<i>lnincome</i>	您个人 2011 年全年的劳动收入是多少?	取全年劳动收入的对数值,对数产生的缺失值由 0 替换。
非农就业	<i>nonfarm</i>	是否非农就业?	是=1;否=0。
性别	<i>sex</i>	性别	男=1;女=2。
年龄	<i>age</i>	您的出生日期(年)	年龄=2012-出生日期
年龄平方	<i>agesq</i>	年龄的平方	$agesq = age^2$ 。
民族	<i>minority</i>	您的民族是	汉族=0;少数民族=1。
宗教信仰	<i>religion</i>	您的宗教信仰是	不信仰宗教=0;信仰宗教=1。
教育	<i>edu</i>	您目前的最高教育程度	没有受过任何教育=1,私塾=2,小学=3,……,研究生=13。
政治面貌	<i>party</i>	您目前的政治面貌是	非党员=0,党员=1。
健康状况	<i>health</i>	您的健康状况	很不健康=1,比较不健康=2,一般=3,比较健康=4,很健康=5。
社会地位	<i>class</i>	您认为自己目前处于哪个等级	取值 1-10,依次由最低层向最顶层过度。
婚姻	<i>marriage</i>	您的婚姻状况	未婚=0;同居、初婚再婚有配偶、分居未离婚、离婚、丧偶=1。

表 2 变量的描述性统计

变量	均值	标准差	最小值	最大值	观测值
<i>lnincome</i>	7.600	5.168	0.000	16.118	849
<i>nonfarm</i>	0.530	0.499	0.000	1.000	849
<i>sex</i>	1.463	0.499	1.000	2.000	849
<i>age</i>	44.311	15.947	17.000	90.000	849
<i>agesq</i>	2 217.458	1 545.058	289.000	8 100.000	849
<i>minority</i>	0.041	0.199	0.000	1.000	849
<i>religion</i>	0.154	0.361	0.000	1.000	849
<i>edu</i>	4.168	2.276	1.000	12.000	849
<i>party</i>	0.054	0.227	0.000	1.000	849
<i>health</i>	3.734	1.020	1.000	5.000	849
<i>class</i>	4.252	1.780	1.000	10.000	849
<i>marriage</i>	0.912	0.284	0.000	1.000	849

通过以上对变量和数据的选取,本文进一步观察非农就业所形成的收入分布差异。图 1 展示了城郊农民收入的核密度函数。从图中可以看到,在收入的低收入区间,未非农就业的人群概率密度高于非农就业人群,说明低收入区间未非农就业人群密度较高。在中等收入区间,非农就业人群概率密度高于未非农就业,说明非农就业人数较多。而在中高收入区间,未非农就业人群概率密度又超过非农就业人群。在高收入区间,非农就业人群超过未非农就业人群。这种现象说明在收入处于中等收入区间的人群主要是通过非农就业渠道获取收入,而低收入区间和高收入区间的人群

则主要是通过农业就业渠道获取收入。这从侧面反映出农业就业既是致贫的主要因素,也是致富的因素,致贫的原因可能来自于传统落后的农业生产方式,而致富的原因可能来自于现代先进的农业生产方式,通过规模经济达到农业收益最大化。但中等收入群体占城郊农民的绝大多数,因此非农就业是提高城郊农民收入的有效途径。

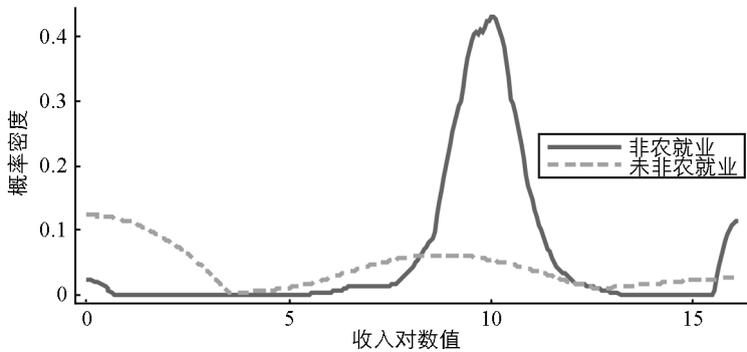


图 1 城郊农民收入的概率密度函数

数据来源:根据 CGSS 数据整理

从图 1 可以看出非农就业与未非农就业的城郊农民存在着显著的收入差异。表 3 进一步用 t 检验分析了非农就业造成的收入差异问题。如表 3 所示,非农就业的城郊农民收入对数值为 10.27,而未非农就业人群仅为 4.589,前者比后者高 124%,两者的收入对数均值存在显著差异。从影响收入的各因素看,性别、年龄、宗教信仰、受教育程度、健康状况等均存在显著差异,相较于未非农就业的城郊农民,非农就业人群男性居多、年龄较小、信教者较少、受教育程度更高、身体更健康。但这也隐含着内生性问题,因此需要运用 PSM 方法将各变量和非农就业对收入的影响分离出来。另外,是否少数民族、是否党员、社会地位、婚姻状况等方面,未非农就业人群和非农就业人群不存在显著差异。

表 3 基于非农就业分类的变量均值及均值差异

	未非农就业	非农就业	均值差异
<i>lnincome</i>	4.589	10.27	-19.116(0.000)
<i>sex</i>	1.617	1.327	8.824(0.000)
<i>age</i>	51.055	38.331	12.643(0.000)
<i>agesq</i>	2 903.361	1 609.291	13.401(0.000)
<i>minority</i>	0.05	0.033	1.228(0.220)
<i>religion</i>	0.208	0.107	4.116(0.000)
<i>edu</i>	3.551	4.716	-7.691(0.000)
<i>party</i>	0.045	0.062	-1.010(0.272)
<i>health</i>	3.431	4.002	-8.479(0.000)
<i>class</i>	4.256	4.249	0.055(0.956)
<i>marriage</i>	0.925	0.9	1.271(0.204)

注:均值差异栏中为 t 值,括号内为 p 值。

五、实证分析

(一)倾向得分匹配模型选择

从表 3 的分析结果可知,非农就业的城郊农民收入明显高于未非农就业的城郊农民。但进一步分析其个体特征发现,非农就业的城郊农民大多为男性、受教育程度较高、年纪更轻、身体更健康,这种特征隐藏着非农就业对城郊农民收入影响的内生性问题。我们无法区分非农就业对城郊农民收入的影响是来自于非农就业自身因素,还是来自于非农就业的城郊农民所具有的更能提高收入能力的个体特征因素。因此,前述分析为我们进一步进行 PSM 分析提供了基础。

PSM 分析的第一步是获得倾向得分(PS 值)。倾向得分主要通过 probit 或 logit 等概率回归模型进行估计而获取。为了达到匹配效果最佳,概率模型的选择很重要。Lian 等认为,倾向得分估计的概率模型设定主要通过调整 R^2 (pseudo- R^2)和 AUC 值进行判定^[37]。在概率模型中,被解释变量是 0—1 离散型变量,而通过概率模型估计获得的倾向得分(PS 值)则是连续变量,因此传统的判定方法失效^[38],而 AUC 值^①则能较好地反映模型设定效果。表 4 列举了 logit 和 probit 两种概率回归模型的估计结果。从回归结果来看,logit 和 probit 的调整均 R^2 为 0.255,logit 模型的 AUC 值高于 probit 模型的 AUC 值,Stürmer 等认为获取倾向得分值的概率模型,其值大于 0.8 时,匹配模型的效果会较好^[39]。而本文中 logit 模型的 AUC 值为 0.821,高于 0.8,同时也高于 probit 模型的 AUC 值。因此,本文选择 logit 模型作为获取倾向得分的回归模型。

表 4 概率模型估计

	<i>logit</i>	<i>probit</i>
<i>sex</i>	-1.558*** (-8.70)	-0.935*** (-8.97)
<i>age</i>	0.111** (2.56)	0.057** (2.36)
<i>agesq</i>	-0.002*** (-4.17)	-0.001*** (-4.13)
<i>minority</i>	-0.604 (-1.50)	-0.367 (-1.51)
<i>religion</i>	-0.207 (-0.85)	-0.134 (-0.92)
<i>edu</i>	0.051 (1.15)	0.027 (1.05)
<i>party</i>	0.335 (0.83)	0.217 (0.95)
<i>healty</i>	0.267*** (2.94)	0.163*** (3.03)
<i>class</i>	0.015 (0.32)	0.011 (0.39)
<i>marriage</i>	0.708** (1.99)	0.454** (2.20)
<i>constant</i>	-0.178 (-0.17)	0.049 (0.08)
<i>pseudo-R²</i>	0.255	0.255
AUC	0.821	0.820
N	849	849

注:1. 被解释变量为 0—1 变量 nonfarm,0 代表未非农就业,1 代表非农就业。

2. *, **, *** 分别表示在 10%、5%、1%水平上显著。

3. AUC 代表 ROC 曲线以下的面积。

(二)匹配效果分析

经过 logit 模型计算得到倾向匹配得分(PS 值后),选择合适的匹配方法,可以计算处理组(非农就业)与对照组(未非农就业)的平均处理效应(ATT)。匹配方法主要有最近邻匹配、半径匹配和核匹配。本文以最近邻匹配方法为主,半径匹配和核匹配作为稳健性检验。以最近邻匹配方式进行匹配后,可以通过核密度函数图和 ROC 曲线图检验匹配效果。本文显示了匹配后的 ROC 曲线图,如图 2 所示。从图中可以看出,ROC 曲线远离对角线,这反映出 AUC 值较大,AUC 值越大,

① 医学中常用的受试者工作特征曲线(Receiver Operating Characteristic,ROC)以下的面积被定义为 AUC 值。

表明匹配效果越好。因此,ROC 曲线直观地表明倾向得分匹配效果良好。

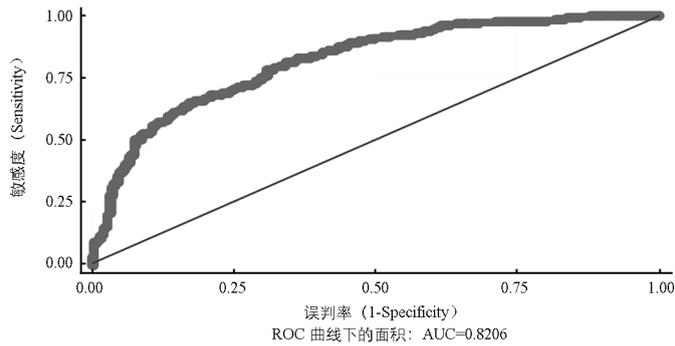


图 2 倾向得分匹配后的 ROC 曲线图

数据来源:根据 CGSS 数据整理

倾向得分匹配还必须满足平行假设,即城郊农民非农就业群体与未非农就业群体在倾向得分匹配后,其各个维度特征无显著差异。表 5 列举了倾向得分匹配后的平衡性检验,标准偏误低于 20%通常被认为是良好的匹配结果^[40]。从表中可以看出,性别、年龄的平方、受教育程度、是否少数民族、健康、社会地位等变量的标准偏误都低于 5%,而在匹配前,这些变量的标准偏误都较高。与此同时,通过 t 检验可以看出,性别、年龄、年龄的平方、宗教信仰、健康状况等变量在匹配前,控制组和处理组均存在显著差异,而在匹配后差异不再显著。是否少数民族、社会地位两个变量在匹配前和匹配后,控制组与处理组均不存在显著差异。受教育程度在匹配前,两组差异显著,匹配后,两组差异只在 10%的置信水平上显著。而政治面貌和婚姻状况两变量在匹配前,两组差异不显著,匹配后却变得显著。总体来看,经过倾向得分匹配后,对照组与处理组各个维度特征差异不显著,匹配效果较好。

表 5 倾向得分匹配后的平衡性检验

变量	样本	均值		标准偏误 (%)	误差消减 (%)	t 检验	
		处理组	控制组			t 值	p 值
sex	匹配前	1.327	1.617	-60.6		-8.82	0.000
	匹配后	1.327	1.313	2.8	95.4	0.43	0.669
age	匹配前	38.331	51.055	-86		-12.64	0.000
	匹配后	38.331	37.054	8.6	90.0	1.55	0.121
agesq	匹配前	1 609.3	2 903.4	-90.8		-13.40	0.000
	匹配后	1609.3	1537.8	5.0	94.5	1.03	0.301
minority	匹配前	0.033	0.050	-8.4		-1.23	0.220
	匹配后	0.033	0.029	2.2	73.5	0.38	0.701
religion	匹配前	0.107	0.208	-28.1		-4.12	0.000
	匹配后	0.107	0.084	6.2	78.1	1.13	0.257
edu	匹配前	4.716	3.551	52.9		7.69	0.000
	匹配后	4.716	5.001	-13.0	75.5	-1.68	0.093
party	匹配前	0.062	0.045	7.6		1.10	0.272
	匹配后	0.062	0.099	-16.1	-112.1	-2.00	0.045
health	匹配前	4.002	3.431	58.0		8.48	0.000
	匹配后	4.002	3.990	1.3	97.8	0.22	0.828
class	匹配前	4.249	4.256	-0.4		-0.06	0.956
	匹配后	4.249	4.125	6.9	-1 732.6	1.06	0.290
marriage	匹配前	0.900	0.925	-8.8		-1.27	0.204
	匹配后	0.900	0.819	28.6	-225.4	3.51	0.000

注:1. *、**、*** 分别表示在 10%、5%、1%水平上显著。2. 匹配前是指处理组与控制组(非农就业与未非农就业)没有经过倾向得分匹配所得到的估计结果,匹配后是指处理组与控制组经过倾向得分匹配后得到的估计结果。3. 处理组与控制组是指城郊农民是否非农就业所划分的组别。匹配前,处理组指非农就业的城郊农民样本,控制组指未非农就业的城郊农民样本;匹配后,处理组指非农就业的城郊农民样本,控制组指未非农就业、但其特征与非农就业群体特征类似的城郊农民样本。

表 6 进一步列举了匹配模型的总体检验。运用非农就业与匹配变量的回归,在匹配前,如表 6 所示,调整 R^2 为 0.255,这在概率模型中算比较良好的拟合效果,似然比检验显著,即匹配变量均为 0 的假设被拒绝,模型整体上显著。这种统计结果表明,在匹配前,控制组与对照组存在明显差异,城郊农民中非农就业群体与未非农就业群体的变量特征存在显著差异;匹配后,调整 R^2 为 0.015,拟合效果较差,似然比检验显著水平下降,模型在 1% 的置信水平上被拒绝。这表明,匹配后,控制组与对照组的差异不明显,通过倾向得分匹配,找到了与城郊农民非农就业群体特征相似的未非农就业群体样本,模型整体匹配效果较好。

表 6 倾向得分匹配模型的检验

样本	$pseudo-R^2$	似然比检验		标准偏误		B	R
		χ^2	p 值	均值	中位数		
匹配前	0.255	299.21	0.000	40.1	40.5	125.1*	0.40*
匹配后	0.015	18.77	0.043	9.1	6.6	28.9*	0.63

注: *、**、*** 分别表示在 10%、5%、1% 水平上显著。

(三) 结果分析

通过以上匹配模型效果的分析,本文的倾向得分匹配模型效果良好。进一步,通过最近邻匹配、半径匹配和核匹配方法,可以估计控制组与处理组的平均处理效应(ATT)。表 7 列举了基于三种匹配方式获得的 ATT 值。从表 7 可知,基于最近邻匹配方式,匹配前,处理组与控制组的收入对数值分别为 10.27 和 4.589,前者比后者高 123.8%,两者的差值(ATT)为 5.681,且在 1% 的水平上高度显著;匹配后,处理组与控制组的收入对数值分别为 10.27 和 6.459,前者比后者高 59%,ATT 差值为 3.811,在 1% 的置信水平上显著。这表明,经过匹配后,城郊农民的个体特征差异已经被消除,处理组与控制组的差异完全来自于非农就业行为影响。匹配后两组的收入对数值依然显著,说明非农就业确实能够促进城郊农民收入水平的提高。

通过半径匹配和核匹配两种匹配方式的分析,可以对结果的稳健性进行验证。基于半径匹配方法,匹配前,处理组与控制组的收入对数值分别为 10.27、4.589,前者比后者高 123.8%,在 1% 的水平上显著;匹配后,两组收入对数值分别为 10.141、7.29,前者比后者高 39.11%,在 1% 的水平上显著。半径匹配方法表明非农就业能够显著提高城郊农民收入水平。同理,基于核匹配方法,匹配前,处理组与控制组的收入对数值分别为 10.27、4.589,在 1% 的水平上显著;匹配后,两组收入对数值分别为 10.27、6.983,在 1% 的水平上显著。核匹配方法也表明非农就业显著提高了城郊农民收入水平。两种匹配方法均验证了结果的稳健性。

表 7 三种匹配方式的 ATT 比较

匹配方式	样本	处理组	控制组	ATT	标准误	t 值
最近邻匹配	匹配前	10.270	4.589	5.681	0.297	19.120***
	匹配后	10.270	6.459	3.811	0.633	6.020***
半径匹配	匹配前	10.270	4.589	5.681	0.297	19.120***
	匹配后	10.141	7.290	2.850	0.669	4.260***
核匹配	匹配前	10.270	4.589	5.681	0.297	19.120***
	匹配后	10.270	6.983	3.287	0.560	5.870***

注: *、**、*** 分别表示在 10%、5%、1% 水平上显著。

(四) 异质性分析

通过前述分析发现,城郊农民的受教育程度、社会地位和性别是重要的异质性变量。不同受教育程度、不同社会地位、不同性别的人群在非农就业行为选择时,会做出不同的抉择。因此,本文有必要基于受教育程度、社会地位、性别等变量,区分不同受教育程度、不同社会地位、不同性别的人群非农就业行为选择问题。表 1 所列的变量定义中,受教育程度的指标描述是从 1—13 分别表示未受过教育,⋯,研究生。本文为了方便起见,将受教育程度为 1、2、3 的样本归纳为小学及以下,将受教育程度为 4、5、6、7、8 的样本归纳为中学,将受教育程度为 9、10、11 的样本归纳为大学及以上。

同理,社会地位为 1、2、3 的样本归纳为最低层,社会地位为 4、5、6、7 的样本归纳为中产阶级,社会地位为 8、9、10 的样本归纳为最顶层,性别变量 1 代表男性、2 代表女性。

表 8 列举了不同受教育程度的城郊农民的 *ATT* 估计结果。从表中可以看出,基于最近邻匹配,匹配前,非农就业均促进了不同教育背景城郊农民的收入水平,小学及以下、中学在 1% 水平上高度显著,大学及以上在 5% 水平上显著。值得注意的是,非农就业对城郊农民收入的促进作用,在受教育程度较低人群更为明显;匹配后,小学及以下、中学学历的城郊农民,其非农就业行为显著促进了收入水平的提高,小学及以下在 1% 水平上显著,中学在 10% 水平上显著,而大学及以上群体,非农就业对收入的促进作用不显著。从 *ATT* 差值上可以看出,教育程度越低,非农就业对收入的促进作用更明显,小学及以下、中学、大学及以上的 *ATT* 差值分别为 3.632、1.952、1.654。由此可见,非农就业对受教育程度较低的城郊农民的收入提升作用是非常显著的。同理,半径匹配、核匹配方法也都得出了相似的结果,验证了结果的稳健性。但如图 3 所示,小学及以下、中学、大学及以上的城郊农民非农就业比例分别为 33.23%、63.24%、75.93%,可见,受教育程度较低的人群,其非农就业比例也较低。因此,要鼓励受教育程度较低的城郊农民积极选择非农就业,通过破除低技能劳动者非农就业进入障碍,免除低受教育人群进入非农产业的后顾之忧。

表 8 不同受教育程度的估计 *ATT* 结果

匹配方式	样本	小学及以下		中学		大学及以上	
		<i>ATT</i>	<i>t-stat</i>	<i>ATT</i>	<i>t-stat</i>	<i>ATT</i>	<i>t-stat</i>
最近邻匹配	匹配前	5.883	10.650***	4.959	12.900***	3.840	3.080**
	匹配后	3.632	3.970***	1.952	1.970*	1.654	0.660
半径匹配	匹配前	5.883	10.650***	4.959	12.900***	3.840	3.080**
	匹配后	4.618	3.420***	3.636	3.540***	-0.075	-0.020
核匹配	匹配前	5.883	10.650***	4.959	12.900***	3.840	3.080**
	匹配后	3.431	4.400***	2.402	2.650**	2.144	0.730

注: *、**、*** 分别表示在 10%、5%、1% 水平上显著。

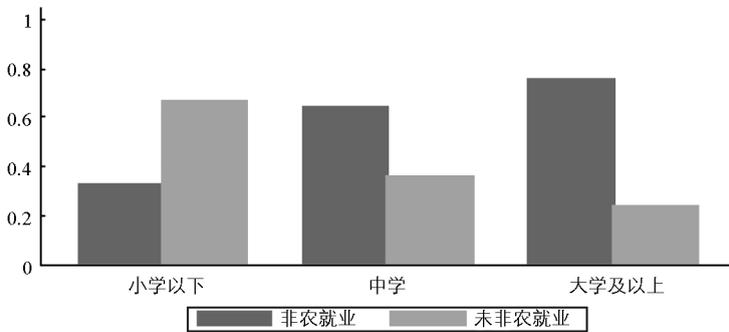


图 3 不同受教育程度的城郊农民非农就业比例

数据来源:根据 CGSS 数据整理

表 9 列举了不同社会地位的城郊农民 *ATT* 差值估计结果。基于最近邻匹配,从表中可以看出,匹配前,非农就业均促进了不同社会地位的城郊农民收入水平,最低层、中产阶级、最顶层的 *ATT* 差值分别为 5.47、5.549、7.268,依次呈递增态势,且均在 1% 的置信水平上高度显著;匹配后,最底层的差值为 2.733,在 5% 水平上显著,中产阶级的 *ATT* 差值为 2.851,在 1% 水平上显著,最顶层的 *ATT* 差值为 5.739,在 5% 水平上显著。从三者 *ATT* 差值大小看,最低层到最顶层依次呈递增态势,这表明非农就业均促进了不同社会地位人群的收入水平,对于社会地位较高的人群,其促进作用更为明显。同理,半径匹配、核匹配方法也都得出了相似的结果,验证了结果的稳健性。表 10 报告了不同社会地位的城郊农民非农就业情况。从表中可以发现,不同社会地位的城郊农民非农就业比例大体相当,均超过 50%,但社会地位较低的城郊农民基数较大,还有相当数量的城郊农民未选择非农就业,社会地位在一定程度上限制了中下层人群进入非农产业。因此,鼓励社会地位较低的城郊农民选择非农就业,破除就业歧视和进入障碍,保障非农就业的均等化。

表 9 不同社会地位的估计结果

匹配方式	样本	最低层		中产阶级		最顶层	
		ATT	t-stat	ATT	t-stat	ATT	t-stat
最近邻匹配	匹配前	5.470	10.770***	5.549	14.150***	7.268	7.270***
	匹配后	2.733	2.980**	2.851	3.520***	5.739	3.030**
半径匹配	匹配前	5.470	10.770***	5.549	14.150***	7.268	7.270***
	匹配后	3.266	2.140*	2.439	2.390*	2.558	0.440
核匹配	匹配前	5.470	10.770***	5.549	14.150***	7.268	7.270***
	匹配后	2.987	3.510***	3.039	4.050***	6.077	3.140**

注：*、**、*** 分别表示在 10%、5%、1%水平上显著。

表 11 报告了不同性别城郊农民的 ATT 差值结果。基于最近邻匹配,匹配前,男性和女性的 ATT 差值分别为 5.254、5.730,在 1%水平上高度显著;匹配后,男性的 ATT 差值为 2.03,在 10%水平上显著,女性的 ATT 差值为 4.898,在 1%水平上高度显著,女性的 ATT 差值大大高

于男性,这表明非农就业的女性收入比未非农就业的女性收入增加幅度高于男性,非农就业对女性的收入增加效应更为明显。半径匹配、核匹配方法也得到了相似的结果,验证了结果的稳健性。然而如表 12 所示,男性群体中非农就业比例高达 66.45%,而女性的非农就业比例仅为 37.40%。女性非农就业收入增加效应更为明显,非农就业比例却很低,导致女性收入水平明显低于男性。因此,政府应当把女性非农就业作为就业政策的优先考虑方向。

表 11 不同性别的估计结果

匹配方式	样本	男		女	
		ATT	t-stat	ATT	t-stat
最近邻匹配	匹配前	5.254	14.490***	5.703	11.020***
	匹配后	1.707	2.030*	4.898	7.320***
半径匹配	匹配前	5.254	14.490***	5.703	11.020***
	匹配后	1.384	1.200	5.882	6.390***
核匹配	匹配前	5.254	14.490***	5.703	11.020***
	匹配后	1.695	2.110*	4.400	7.580***

注：*、**、*** 分别表示在 10%、5%、1%水平上显著。

表 12 不同性别的城郊农民非农就业比例

性别	非农就业比例	观测值
男性	66.45%	456
女性	37.40%	393

六、结论和政策含义

本文基于中国综合社会调查(CGSS)2012年的统计数据,运用倾向得分匹配模型,分析了非农就业对城郊农民收入的影响问题。同时基于受教育程度和社会地位是重要的异质性变量,本文重点分析了不同受教育程度和不同社会地位的城郊农民非农就业对收入的影响。结果发现:(1)基于最近邻匹配方法,匹配前非农就业的城郊农民比未非农就业的城郊农民收入对数值高 123.8%,在 1%的水平上显著;匹配后两组收入对数值分别为 10.27、6.459,前者比后者高 59%,在 1%的水平上显著,表明非农就业能够显著提高城郊农民收入水平。(2)受教育程度越低,非农就业对收入的促进作用更明显,小学及以下、中学、大学及以上的 ATT 差值分别为 3.632、1.952、1.654。随着教育背景的提升,ATT 差值呈递减趋势,表明非农就业对高学历城郊农民的收入增加效应不显著,而

对受教育程度较低的城郊农民的收入提升作用是非常显著的。(3)非农就业促进了不同社会地位人群的收入水平,对于社会地位较高的人群,其促进作用更为明显,最底层、中产阶级、最顶层的城郊农民的 ATT 差值分别为 2.733、2.851、5.739。(4)非农就业对女性收入增加效应更为显著,但女性非农就业比例却很低。

基于以上结论,本文的政策含义是:(1)政府应当鼓励城郊农民非农就业。非农就业能够显著增加家庭收入、增加消费和减少贫困,通过制度相应的政策措施,把劳动者分流到非农行业,既可以解决城镇化扩张造成的城郊农民土地流失问题,也可以使城郊农民有效参与市民化过程。(2)政策制订者要着力关注受教育程度较低、社会地位较低、性别为女性的城郊农民非农就业问题。受教育程度较低、社会地位较低、性别为女性的城郊农民非农就业能够显著增加收入,但他们面临非农就业进入障碍,导致这些人群的非农就业参与度较低。政府要破除就业歧视和进入障碍,免除他们非农就业的后顾之忧,鼓励他们从事非农活动。(3)上述第二点中提到“城郊农民面临非农就业进入障碍”“要降低低学历、女性等人群非农就业进入障碍”,这涉及到政府、企业及各种非政府组织、社会“必须做什么、能够做什么、不能做什么、做的效果”等问题。具体来说,政府应当制订针对低学历、女性等人群的针对性就业政策、营造宽松的就业环境;企业应当实行同工同酬的工资待遇、减少对低学历、女性等人群的就业歧视;非政府组织及社会应当积极宣传公平的就业政策,为他们提供各种就业帮扶和爱心救助。(4)当然,研究结论已经证明“低学历、女性人群由纯农就业转为非农就业,其收入提升效应更为显著”,但由于专业技术、性别等方面的差异,政府应当鼓励低学历、女性等人群进入到就业门槛较低的服务行业。相较于非农产业,农业是一个比较经济效益不好而比较社会效益好的产业,农业比较经济效益低是一个普遍现象。因此,鼓励城郊农民中的低技能劳动者转移到城市第三产业中来,既能实现低技能劳动者经济效益的提升,也能实现城郊农业规模经济效益。梁文泉和陆铭认为低技能劳动者能够与城市高技能劳动实现互补、实现人力资本外溢^[41],而这些低技能人群进入到服务行业,正好能够促进城市高低技能劳动者的共同发展,促进城镇化快速发展。

参考文献:

- [1] KUNG J K S, LEE Y. So What If There Is Income Inequality? The Distributive Consequence of Nonfarm Employment in Rural China[J]. *Economic Development and Cultural Change*, 2001, 50(1): 19-46.
- [2] HAGGBLADE S, HAZELLI P, REARDON T. The Rural Non-farm Economy: Prospects for Growth and Poverty Reduction[J]. *World Development*, 2010, 38(10):1429-1441.
- [3] 丁守海. 中国城镇发展中的就业问题[J]. *中国社会科学*, 2014(1):30-47.
- [4] 李永友, 徐楠. 个体特征、制度性因素与失地农民市民化——基于浙江省富阳等地调查数据的实证考察[J]. *管理世界*, 2011(1):62-70.
- [5] IMAI K S, GAIHA R, THAPA G. Does non-farm sector employment reduce rural poverty and vulnerability? Evidence from Vietnam and India[J]. *Journal of Asian Economics*, 2015, 36:47-61.
- [6] 钟甫宁, 何军. 增加农民收入的关键:扩大非农就业机会[J]. *农业经济问题*, 2007(1):62-70.
- [7] SHAHE EMRAN M, HOU Z. Access to Markets and Rural Poverty: Evidence from Household Consumption in China[J]. *Review of Economics & Statistics*, 2011, 95(2):682-697.
- [8] CHERDCHUCHAI S, OTSUKA K. Rural income dynamics and poverty reduction in Thai villages from 1987 to 2004[J]. *Agricultural Economics*, 2006, 35(Supplement s3):409-423.
- [9] KIJIMA Y, MATSUMOTO T, YAMANO T. Nonfarm employment, agricultural shocks, and poverty dynamics: evidence from rural Uganda[J]. *Agricultural Economics*, 2006, 35(s3):459-467.
- [10] LANJOUW P, MURGAI R. Poverty decline, agricultural wages, and nonfarm employment in rural India: 1983-2004[J]. *Agricultural Economics*, 2009, 40(2):243-263.
- [11] REARDON T, STAMOULIS K, BALISACAN A, et al. Rural non-farm income in developing countries[J]. *The state of food and agriculture*, 1998, 1998: 283-356.
- [12] ABDULAI A, DELGADO C L. Determinants of nonfarm earnings of farm-based husbands and wives in Northern Ghana[J]. *American Journal of Agricultural Economics*, 1999, 81(1): 117-130.
- [13] BARRETT C B, REARDON T, WEBB P. Nonfarm income diversification and household livelihood strategies in rural Africa:

- concepts, dynamics, and policy implications[J]. *Food Policy*, 2001, 26(4): 315-331.
- [14] CANAGARAJAH S, NEWMAN C, BHATTAMISHRA R. Non-farm income, gender, and inequality: evidence from rural Ghana and Uganda[J]. *Food Policy*, 2001, 26(4):405-420.
- [15] LANJOUW P. Nonfarm Employment and Poverty in Rural El Salvador[J]. *World Development*, 2001, 29(3):529-547.
- [16] 任国强. 人力资本对农民非农就业与非农收入的影响研究——基于天津的考察[J]. *南开经济研究*, 2004, 03:3-10.
- [17] 艾春荣, 汪伟. 非农就业与持久收入假说:理论和实证[J]. *管理世界*, 2010(1):8-22.
- [18] JOHN Knight, Li Shi, Deng Quheng. Education and the poverty trap in rural China; setting the trap.[J]. *Oxford Development Studies*, 2009, 37(4):311-332.
- [19] BLISS C, STERN N. Productivity, wages and nutrition ☆ : Part I: the theory[J]. *Journal of Development Economics*, 1978, 5(4):331-362.
- [20] DASGUPTA P, RAY D. Inequality as a determinant of malnutrition and unemployment: Theory[J]. *The Economic Journal*, 1986, 96(384): 1011-1034.
- [21] JHA R, GAIHA R, SHARMA A. Calorie and Micronutrient Deprivation and Poverty Nutrition Traps in Rural India[J]. *World Development*, 2009, 37(5):982-991.
- [22] REARDON T, TAYLOR J E, STAMOULIS K, et al. Effects of Non-Farm Employment on Rural Income Inequality in Developing Countries: An Investment Perspective[J]. *Journal of Agricultural Economics*, 2000, 51(2):266-288.
- [23] TAYLOR J E, MARTIN P L. Human capital: Migration and rural population change[M]. Elsevier, 2001.
- [24] HAGGBLADE S, HAZELL P, REARDON T. The rural non-farm economy: Prospects for growth and poverty reduction[J]. *World Development*, 2010, 38(10): 1429-1441.
- [25] LANJOUW J O, LANJOUW P. The Rural Non-Farm Sector: Issues And Evidence From Developing Countries[J]. *Agricultural Economics*, 2001, 26(1):1-23.
- [26] BABATUNDE R O, QAIM M. Impact of off-farm income on food security and nutrition in Nigeria[J]. *Food Policy*, 2010, 35(4): 303-311.
- [27] HOLDEN S, SHIFERAW B, PENDER J. Non-farm income, household welfare, and sustainable land management in a less-favoured area in the Ethiopian highlands[J]. *Food Policy*, 2004, 29(4):369-392.
- [28] 刘洪银. 我国农村劳动力非农就业的经济增长效应[J]. *人口与经济*, 2011(2):23-27.
- [29] 钱文荣, 郑黎义. 劳动力外出务工对农户农业生产的影响——研究现状与展望[J]. *中国农村观察*, 2011(1):31-38.
- [30] 宁光杰. 自选择与农村剩余劳动力非农就业的地区收入差异——兼论刘易斯转折点是否到来[J]. *经济研究*, 2012(S2):42-55.
- [31] 冉璐, 谢家智, 张明. 非农工作经历与农民务农收入:基于分位数回归与分解的实证研究[J]. *农业技术经济*, 2013(6):32-39.
- [32] HUFFMAN W E. Agricultural household models: survey and critique[J]. *Multiple job-holding among farm families*, 1991: 79-111.
- [33] OWUSU V, ABDULAI A, ABDUL-RAHMAN S. Non-farm work and food security among farm households in Northern Ghana [J]. *Food Policy*, 2011, 36(2):108-118.
- [34] HUFFMAN W E, LANGE M D. Off-farm work decisions of husbands and wives: joint decision making[J]. *The Review of Economics and Statistics*, 1989: 471-480.
- [35] ROSENBAUM P R, RUBIN D B. The central role of the propensity score in observational studies for causal effects[J]. *Biometrika*, 1983, 70(1):41-55.
- [36] FALTERMEIER L, ABDULAI A. The impact of water conservation and intensification technologies: empirical evidence for rice farmers in Ghana[J]. *Agricultural Economics*, 2009, 40(3):365-379.
- [37] LIAN Y, SU Z, GU Y. Evaluating the effects of equity incentives using PSM: Evidence from China[J]. *Frontiers of Business Research in China*, 2011, 5(2):266-290.
- [38] HOSMER JR D W, LEMESHOW S, STURDIVANT R X. *Applied logistic regression*[M]. New York:John Wiley & Sons, 2013.
- [39] STÜRMER T, JOSHI M, GLYNN R J, et al. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods[J]. *Journal of clinical epidemiology*, 2006, 59(5): 437. e1-437. e24.
- [40] ROSENBAUM P R, RUBIN D B. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score[J]. *The American Statistician*, 1985, 39(1): 33-38.
- [41] 梁文泉, 陆铭. 城市人力资本的分化:探索不同技能劳动者的互补和空间集聚[J]. *经济社会体制比较*, 2015(3):185-197.