

流动人口的创业决策与收入效应

——基于自选择与异质性视角

莫旋¹, 周建²

(1. 衡阳师范学院 经济与管理学院, 湖南 衡阳 421002; 2. 上海财经大学 经济学院, 上海 200433)

【摘要】基于中国流动人口动态监测数据,利用内生转换模型(ESRM)弥补现有研究方法的不足,将流动人口创业决策纳入实证分析框架,从自选择和异质性视角,研究流动人口创业的收入效应。研究发现:流动人口性别与教育等因素同时影响其创业选择和收入,控制样本选择性偏误,创业者的收入显著高于不创业者,但创业对流动人口的收入效应具有明显异质性,不创业者创业的收入效应大于创业者,发展型创业者的收入效应大于生存型创业者。研究结果为精准地评估流动人口创业收入效应提供了科学依据,对于鼓励流动人口进行创业、建设创业型经济具有一定的政策启示。

【关键词】创业 流动人口 收入效应 自选择 异质性

【DOI】10.15884/j.cnki.issn.1007-0672.2019.05.001

【收稿日期】2019-01-24

【中图分类号】F014.44

【文献标志码】A

【文章编号】1007-0672(2019)05-0001-11

【基金项目】国家自然科学基金项目“城乡分割、区域差异和国际经济冲击三重复杂现状下的中国宏观经济系统稳定性、抗冲击承受力及反冲击政策研究”(71673175);湖南省社会科学成果评审委员会项目“新时代背景下流动人口创业决策与收入效应——基于自选择与异质性视角”(XSP19YBC280);湖南省教育厅科学研究优秀青年项目“新时代背景下女性流动人口的就业状况与主观幸福感研究——基于家庭庇护、体制庇护与工作家庭冲突视角”(18B363);衡阳师范学院科研启动项目“中国流动人口的就业、收入与主观幸福感”(18D30)。

【作者简介】莫旋,男,湖南衡阳人,衡阳师范学院经济与管理学院副教授,经济学博士,研究方向:流动人口;周建,男,四川南充人,上海财经大学经济学院教授,博士生导师,研究方向:计量经济。

一、引言

改革开放以来,随着经济快速发展,在城乡间与地区间出现了大规模的人口流动。就业获得与收入提升是流动人口流入城市的主要目标,就业是民生之本,收入是流动人口在流入地立足的经济基础。已有文献较多地关注流动人口就业的影响因素(王子敏,2017^[1])及其收入的决定(莫旋,2018^[2])。我们将研究视野转向流动人口的创业决策:被人雇佣成为工资获得者,还是自我创业成为创业者。当前中国经济发展进入新常态,在经济增速放缓和就业压力增大的背景下,鼓励创业成为政策制定者和学者们普遍关注的焦点。流动人口创业不但可以解决自身的就业问题,还可创造更多的就业机会,有助于缓解社会就业压力,同时也是优化资源配置、提升经济效率、促进社会纵向流动和公平的重要手段。但是,流动人口在创业过程中经常遭受到一些制度性歧视和非制度性约束,从而影响其创业热情。流动人口的创业决策受哪些因素影响?如何评估流动人口创业的收入效应?分析这些问题,有利于流动人口更好地进行创业,提高资源配置效率,提升流动人口的收入水平,让更多流动人口能共享改革发展成果。因此,研究流动人口的创业决策以及评估创业的收入效应就具有重要的现实意义。

本文讨论的核心问题是:如何精准评估流动人口创业的收入效应。当前研究这一问题的主要方法有:基于OLS的多元线性回归法、倾向得分匹配法、处理效应模型以及双重差分法等。如果流动人口是否创业是外生给定的,那么多元线性模型的OLS回归就足以评估创业的收入效应。但流动人口是否选择创业并不是随机给定的,而是在效用最大化框架下,根据多种因素理性选择的结果,不可避免地存在样本自选择问题,OLS回归忽视了创业的自选择问题,将导致估计结果有偏误。因此,研究流动人口创业收入效应时,首先需要考虑流动人口的创业决策,解决样本自选择问题导致的选择性偏误。近期有一些文献开始关注流动人口的创业决策,但主要从户籍(宁光杰,2017^[3])、信贷约束(翁辰,2015^[4];郝朝艳,2012^[5])、社会网络(张博,2015^[6])、方言(魏下海,2016^[7])和宗教信仰(阮荣平,2014^[8])等角度展开,对流动人口创业本身存在的自选择问题关注不够。利用倾向得分匹配法分析流动人口创业的收入效应,在很大程度上可解决样本选择性偏误,但倾向得分匹配法只控制可观测变量的影响,对个人能力等不可观测变量则无能为力,因此,仍可能存在“隐性偏差”^[9]。处理效应模型通过估计流动人口创业决策模型,计算流动人口创业的概率值,并将其作为自变量纳入流动人口的收入决定方程,解决由不可观测变量带来的样本选择性偏误,从而弥补了倾向得分匹配法的不足。但该方法忽视了创业者和不创业者之间的结构性差异,未考虑处理效应的异质性。基于自然实验的双重差分法是评估政策效果最常用的方法,但该方法要求研究样本满足随机性与同质性假设,对流动人口创业收入效应的评估很显然不满足模型所需的前置条件。流动人口创业的收入效应受创业决策的影响,创业者与不创业者在可观测特征与不可观测特征方面并不是同质的,而是具有异质性。如果忽视创业者与不创业者之间的异质性,将会导致研究结论过于笼统,缺乏针对性。因此,研究流动人口创业收入效应需要考虑的第二个重要因素是异质性。在实证研究中,自选择与异质性分析通常是密不可分的。

在考虑自选择与异质性的基础上,如何精准评估流动人口创业的收入效应?目前相关研究尚未见于国内文献。本文将采用Maddala提出的内生转换模型(ESRM)^[10]来弥补现有研究方法的不足,内生转换模型考虑了可观测变量和不可观测变量导致的样本选择性偏误,解决了样本自选择问题,弥补了倾向得分匹配法的不足。同时,内生转换模型还分别拟合创业者和不创业者的收入决定方程,并进行反事实推断,分别计算出创业者和不创业者的处理效应,解决了处理效应模型中流动人口处理效应同质性的不合理假定,从异质性视角评估流动人口创业的收入效应。因此,内生转换模型是一种严谨处理流动人口创业收入效应的分析方法。本文基于中国流动人口动态监测数据,利用内生转换模型,从自选择与异质性视角,全面分析流动人口创业的收入效应,并在此基础上,对流动人口创业类型进行进一步细分,结合不创业者构成两个子样本,比较分析不同创业类型收入效应的差异。

二、内生转换模型与处理效应估计方法

(一)内生转换模型(ESRM)

本文主要讨论流动人口两种创业决策:创业或不创业(指工资获得者的雇员)。为分析流动人口创业决策对其收入水平的影响,我们首先构建流动人口收入决定方程:

$$y_i = \alpha S_i + \sum_{k=1}^K \beta_k x_{ik} + \varepsilon_i \quad (1)$$

其中, y_i 是流动人口*i*的收入, S_i 是流动人口*i*是否创业的虚拟变量, x_{ik} 为一组影响流动人口收入的其他解释变量, ε_i 表示随机误差项。

如果流动人口是否创业是随机给定的,那么基于方程式(1)的OLS回归能得到流动人口收入决定的无偏估计量, α 刻画了创业对流动人口收入的影响;但流动人口是否选择创业并不是随机给定的,

而是根据多种因素自我选择的结果。而且,有些因素可能同时影响流动人口的创业决策和收入决定,因而产生样本选择性偏误。

本文借鉴王存同(2017)^[11]、李雪松(2015)^[12]以及朋文欢(2017)^[13]等人的研究方法,构建流动人口的创业决策模型:

$$\begin{aligned} S_i^* &= z_i \gamma + u_i \\ S_i &= 1(S_i^* > 0) \end{aligned} \quad (2)$$

在方程式(2)中, S_i^* 是虚拟变量 S_i 的潜变量,如果流动人口选择创业,则 $S = 1$,否则 $S = 0$, z 是影响流动人口创业决策的一组变量, γ 为对应的系数, u 为随机误差项。需要强调的是,内生转换模型允许创业决策模型中的解释变量 z 与收入决定模型中的解释变量 x 重叠,但为了模型的识别,要求 z 中至少有一个变量不包含于 x ,且该变量直接影响流动人口的创业决策,而不直接影响其收入决定。

因此,流动人口在创业与不创业两种情境模式下,其收入决定方程分别为:

$$y_{1i} = \sum_{k=1}^K \beta_{1k} x_{ik} + \varepsilon_{1i} \quad \text{若 } S_i = 1 \quad (3)$$

$$y_{0i} = \sum_{k=1}^K \beta_{0k} x_{ik} + \varepsilon_{0i} \quad \text{若 } S_i = 0 \quad (4)$$

方程式(3)和(4)组成内生转换模型中的结果方程。其中, y_{1i} 和 y_{0i} 分别表示流动人口 i 创业和不创业时的收入水平,但在现实生活中,我们不可能同时观测到流动人口 i 创业和不创业两种情境下的收入,且随机误差项的条件期望不等于零,因此对方程式(3)和(4)分别采用OLS方法,得到的估计结果是有偏的。 x_{ik} 是影响流动人口收入的解释变量, β_{1i} 和 β_{0i} 分别是方程式(3)和(4)不同的系数向量, ε_{1i} 和 ε_{0i} 是各自方程的随机误差项。

若记 $\sigma_u^2 = \text{var}(u)$, $\sigma_{1u} = \text{cov}(\varepsilon_1, u)$, $\sigma_{0u} = \text{cov}(\varepsilon_0, u)$,将 σ_u^2 标准化为1, ε_1 和 ε_0 的条件期望可表示为:

$$E(\varepsilon_{1i}|S_i = 1) = E(\varepsilon_{1i}|z_i \gamma + u_i > 0) = \sigma_{1u} \frac{\phi(z_i \gamma)}{\Phi(z_i \gamma)} = \sigma_{1u} \lambda_{1i} \quad (5)$$

$$E(\varepsilon_{0i}|S_i = 0) = E(\varepsilon_{0i}|z_i \gamma + u_i < 0) = \sigma_{0u} \frac{-\phi(z_i \gamma)}{1 - \Phi(z_i \gamma)} = \sigma_{0u} \lambda_{0i} \quad (6)$$

因此,流动人口 i 在创业与不创业两种情境下收入的条件期望分别为:

$$E(y_{1i}|S_i = 1) = \beta_{1i} x_{1ik} + \sigma_{1u} \lambda_{1i} \quad (7)$$

$$E(y_{0i}|S_i = 0) = \beta_{0i} x_{0ik} + \sigma_{0u} \lambda_{0i} \quad (8)$$

方程式(5)-(6)中, $\phi(z_i \gamma)$ 和 $\Phi(z_i \gamma)$ 分别表示以 $z_i \gamma$ 为变量的标准正态分布密度函数和累计密度函数。 λ_{1i} 和 λ_{0i} 分别表示由不可观测变量带来的样本选择性偏误,根据方程式(5)和(6),如果 σ_{1u} 和 σ_{0u} 不为零,则有必要纠正由不可观测变量带来的样本选择性偏误。

(二)基于内生转换模型的处理效应估计

根据内生转换模型的估计结果,我们还可以估计出流动人口中创业者不创业时,以及不创业者创业时两种反事实情境下的收入。

$$E(y_{0i}|S_i = 1) = \beta_{0i} x_{1ik} + \sigma_{0u} \lambda_{1i} \quad (9)$$

$$E(y_{1i}|S_i = 0) = \beta_{1i} x_{0ik} + \sigma_{1u} \lambda_{0i} \quad (10)$$

因此,流动人口创业者的平均处理效应(ATT)为式(7)与式(9)之差:

$$ATT = E(y_{1i}|S_i = 1) - E(y_{0i}|S_i = 1) = (\beta_{1i} - \beta_{0i}) x_{1ik} + (\sigma_{1u} - \sigma_{0u}) \lambda_{1i} \quad (11)$$

流动人口不创业者的平均处理效应(ATU)为式(10)与式(8)之差:

$$ATU = E(y_{1i}|S_i = 0) - E(y_{0i}|S_i = 0) = (\beta_{1i} - \beta_{0i}) x_{0ik} + (\sigma_{1u} - \sigma_{0u}) \lambda_{0i} \quad (12)$$

流动人口创业对其收入的平均处理效应(ATE)则为创业者的平均处理效应(ATT)与不创业者的平均处理效应(ATU)的加权平均数。

三、数据、变量与描述性统计

(一)数据来源与处理

本文所使用的是中国流动人口动态监测数据(2014)。该数据的样本按照多阶段、分层、与规模成比例的PPS方法进行抽样。调查对象为在流入地居住一个月以上、非本区(市)户口的16~60岁流入人口。以浙江省嘉兴市、山东省青岛市、福建省厦门市、四川省成都市、河南省郑州市、广东省深圳市和中山市、北京市朝阳区八市(区)为调查地点,每个城市的样本量为2000人。我们对数据作如下处理:删除就业身份缺漏值和模糊值;剔除关键变量缺失和存在明显逻辑错误的样本;为减少测量误差的影响,对收入变量极端值进行截取,处理后有效样本数为13697个。

(二)变量选取

1. 创业

本文试图将流动人口创业决策过程纳入实证分析框架,根据调查问卷中“您现在就业身份属于哪一种”的作答,构建流动人口是否创业虚拟变量,选择“雇主”和“自营劳动者”,认定为创业者,取1,选择“雇员”,认定为不创业者,取0。同时,我们对流动人口创业类型按创业目的细分为两类:生存型创业和发展型创业。由于流动人口在劳动力市场处于相对弱势地位,一部分流动人口在流入地一时难以找到合适的工作岗位,为了生存的需要,不得不选择自雇形式的就业,即为了实现自身就业而进行的创业活动,称为生存型创业;另一种是为了寻找商业机会,在创造物质财富的过程中,更好地实现精神追求和自身价值而进行的创业活动,称为发展型创业。因此,我们将“自营劳动者”定义为生存型创业者,“雇主”定义为发展型创业者。

2. 收入

流动人口动态监测调查问卷中提供了流动人口的月收入,包括工资、奖金、津贴、加班费和单位包吃折算的金额,因为月收入未考虑到个体工作时间的差异,所以小时收入更为合理,由于调查问卷中没有小时收入,因此,我们用周工作小时数乘4来估计月工作小时数,流动人口的小时收入即为月收入除以月工作小时数。

3. 创业决策变量

流动人口的个体特征,如性别、受教育程度、工作经验、婚姻状况等因素,既会影响其收入决定,又会影响其创业决策。为了模型的识别,要求创业决策模型中至少有一个变量能直接影响流动人口的创业决策,而不直接影响其收入决定。根据周京奎(2014)^[14]和彭艳玲(2016)^[15]等人的研究结果可知,创业决策还受老家田地和住房、家庭经济状况等因素影响。结合调查可得的数据,我们选取影响流动人口创业决策的因素主要有:①性别。设定虚拟变量,男性取1,女性取0。由于创业活动中存在一定的风险,而女性多为风险规避者,因此,女性较男性而言,创业概率一般较低。②受教育程度。流动人口调查问卷中受教育程度分为七个层次,未上过学、小学、初中、高中、大学专科、大学本科、研究生依次赋值为1~7。教育是人力资本的主要形式,对流动人口创业决策有重要影响,一般而言,受教育程度越高者,越容易获得一份稳定且收入较高的工作,因此,创业的机会成本就越大,选择风险较高的创业活动概率就越低。③工作经验。假设个人上学年龄是6岁,以取得学历的最高教育时间为受教育年限,用年龄减出受教育年限,再减6,作为工作经验。一般认为,工作经验越丰富,创业所需的知识和能力积累越多,越有利于创业的开展。④婚姻状况。设定虚拟变量,在婚者取值为1,其他取值为0。

在婚者能更多获得来自家庭的经济支持和情感支持,其创业概率可能会更高。⑤老家田地和住房。老家田地越多和住房面积越大,一方面表明其抵御创业风险的能力越强,另一方面也反映其个人工作能力越强,因而创业概率越高。⑥家庭收入和家庭支出收入比。家庭收入用家庭在本地平均每月总收入来度量,家庭收入越高,越有能力从事前期投入较大的创业活动,且抵御创业风险的能力也越强,因此,创业的概率越高。家庭支出收入比是家庭每月总支出比家庭每月总收入,流动人口家庭支出收入比越大,越希望通过创业活动来改变自己当前生活的窘境,从而倾向于选择收入相对较高的创业活动。⑦城市。构造城市虚拟变量,以反映流动人口创业的区域性差异。

4. 收入决定变量

根据孙猛(2018)^[16]和侯建明(2016)^[17]等人的研究结果可知,流动人口的收入受个体特征、职业特征、社会网络和地区经济等因素影响。结合调查可得的数据,我们选取影响流动人口收入决定的因素主要有:①性别。一般认为流动人口收入存在性别差异,男性收入要明显高于女性。②受教育程度。教育是人力资本的主要形式,对流动人口的收入决定有重要影响,一般而言,受教育程度越高者收入水平越高。③工作经验。流动人口收入随着工作经验的累积而增加,同时,引入工作经验的平方项,以考察流动人口收入与工作经验之间的非线性关系。④户籍身份。由于我国特有的二元经济结构,导致乡一城与城一城流动人口在劳动力市场诸多方面存在显著差异,进而会影响其收入的决定。设定户籍虚拟变量,乡一城流动人口取0,城一城流动人口取1,以反映流动人口收入的户籍差异。⑤社会资本。社会资本对流动人口收入有重要影响,社会资本越多者,收入水平越高。用流动人口在本地参加社会组织的数量作为其社会资本的代理变量,未参加任何社会组织者取0,参加1个社会组织者取1,参加2个及以上社会组织者取2,分别表示社会资本较少、中等和较多。⑥健康状况。身体健康状况是人力资本的一种重要形式,对流动人口的收入有重要影响。调查问卷中有流动人口的自评健康状况,健康状况差、一般、好、很好和非常好,依次赋值为0~4。⑦职业特征。流动人口职业特征包括就业单位的所有制和行业性质,就业单位所有制分为个体私营、国有集体和外资企业三种类型,就业行业分为社会服务、制造采供、建筑农牧、商业服务和高端服务五种类型。引入就业单位所有制和行业变量,以考虑流动人口收入的职业差异。⑧城市。构造城市虚拟变量,以反映城市经济发展水平对流动人口收入的影响。

(三)描述性统计

流动人口不创业者和创业者、生存型创业者和发展型创业者在收入水平、个体特征等方面呈现出明显的异质性,变量的统计描述能展现流动人口不同群体之间的差异。表1具体给出了主要变量的描述性统计。

从表1可知,不创业者的小时收入不但高于创业者,而且波动性更少;创业者的收入存在明显分化,生存型创业者的小时收入低于不创业者,但发展型创业者的小时收入明显高于不创业者,无论是生存型创业者还是发展型创业者,收入的波动性都明显大于不创业者。

创业者、尤其是发展型创业者中男性占比明显高于不创业者。不创业者的受教育程度和社会资本明显高于创业者,但波动性较大;生存性创业者的受教育程度和社会资本最低,且波动性最少。不创业者的工作经验明显少于创业者,但波动性较大;生存性创业者的工作经验最丰富。不创业者中在婚比例明显低于创业者,且波动性大。创业者的家庭收入和支出收入比明显高于不创业者,尤其是发展型创业者。

不创业者的小时收入低于发展型创业者,高于生存型创业者,上述判断是基于描述性统计而言,并没有控制其他相关因素,流动人口的收入决定还受其他许多因素的影响。不创业者与创业者、生存

表1 主要变量描述性统计表

变量	全样本 (N=13697)	不创业者 (N=9716)	创业者 (N=3981)	生存型创业者 (N=3051)	发展型创业者 (N=930)
小时收入	16.874(8.523)	17.033(7.969)	16.486(9.733)	15.628(9.119)	19.299(11.070)
性别	0.578(0.494)	0.562(0.496)	0.616(0.486)	0.611(0.488)	0.636(0.482)
教育	3.485(0.967)	3.574(1.009)	3.267(0.814)	3.227(0.775)	3.397(0.919)
工作经验	16.280(9.777)	14.799(9.782)	19.897(8.772)	20.010(8.884)	19.525(8.388)
婚姻状况	0.717(0.450)	0.640(0.480)	0.906(0.292)	0.904(0.294)	0.910(0.287)
老家田地	0.039(0.115)	0.039(0.094)	0.039(0.155)	0.037(0.094)	0.045(0.273)
老家住房	1.481(0.919)	1.498(0.912)	1.439(0.935)	1.422(0.916)	1.493(0.992)
家庭收入	0.598(0.487)	0.536(0.397)	0.750(0.632)	0.709(0.598)	0.885(0.717)
支出收入比	0.515(0.191)	0.497(0.191)	0.559(0.184)	0.557(0.184)	0.566(0.183)
户籍	0.131(0.338)	0.141(0.348)	0.107(0.309)	0.093(0.290)	0.154(0.361)
社会资本	0.369(0.648)	0.401(0.664)	0.289(0.600)	0.275(0.583)	0.334(0.650)
健康状况	2.771(0.967)	2.778(0.964)	2.753(0.973)	2.716(0.983)	2.874(0.927)

注:括号内为标准误。

型创业者与发展型创业者在个体特征等方面均存在明显差异,在研究流动人口创业决策和收入决定时,需对这些因素加以控制,因此,我们还需通过更严谨的实证分析以得到较为精准可靠的分析结论。

四、实证分析结果

首先采用内生转换模型(ESRM)对全样本进行拟合,并根据模型的拟合结果进行反事实推断,计算出创业对流动人口收入的处理效应。然后,尝试将不同创业类型分别纳入实证分析框架,用生存型创业者和发展型创业者分别与不创业者组成子样本1和子样本2,并采用内生转换模型分别进行拟合,计算出两种不同创业类型中创业的处理效应。最后,基于全样本和子样本估计结果,分别与OLS模型和处理效应模型中创业收入效应进行比较,以精准地评价创业对流动人口的增收效果。

(一)全样本估计结果

表2显示了内生转换模型的全样本估计结果。就流动人口的创业决策模型而言,性别是影响流动人口创业的重要因素,男性多为风险偏好者,且抗风险能力一般强于女性,而创业是一种风险较高的活动,因此,男性创业的概率显著高于女性。流动人口受教育程度对其创业有显著的负作用,受教育程度越高,越容易找到合适的工作岗位,创业的机会成本越高,因此,创业的可能性越低。工作经验越丰富,越有利于开展创业活动。在婚者的家庭生活比较稳定,容易获得来自家庭的情感支持和经济支持,创业的概率会更高。老家田地和住房系数均为正,且较显著,这表明老家田地越多和住房面积越大,创业的可能性越大,这可能是由于老家田地和住房,不仅可体现其个人能力,还有助于抵御创业过程中的风险。流动人口家庭收入和支出收入比的系数均为正,且在1%显著性水平上显著;创业活动的前期投入较大,家庭收入越高者,越有能力开展创业活动,且抵御创业风险的能力越强,因而创业概率会越高;流动人口家庭支出收入比越大,因为生活所迫,越希望通过创业活动来提高自身收入,以改变当前生活的窘迫境况,因此创业的概率越大。

进一步比较流动人口中创业者和不创业者的收入决定方程,我们发现诸多变量对两类流动人口收入的影响存在较大差异,由于OLS模型和处理效应模型认为两类流动人口的收入决定机制是同质

表2 内生转换模型(ESRM)的全样本估计结果

变量	创业决策模型	变量	不创业者	创业者	OLS
性别	0.056 ^{**} (0.023)	性别	1.859 ^{***} (0.153)	2.059 ^{***} (0.311)	2.315 ^{***} (0.131)
教育	-0.234 ^{***} (0.015)	教育	2.773 ^{***} (0.098)	2.245 ^{***} (0.239)	2.351 ^{***} (0.089)
工作经验	0.012 ^{***} (0.001)	工作经验	0.064 ^{***} (0.023)	-0.050(0.068)	0.322 ^{***} (0.025)
婚姻状况	0.354 ^{***} (0.029)	经验平方	-0.366 ^{***} (0.056)	-0.090(0.154)	-0.733 ^{***} (0.061)
老家田地	0.163 ^{**} (0.075)	户籍	1.233 ^{***} (0.178)	0.977 ^{**} (0.493)	1.789 ^{***} (0.209)
老家住房	0.018 ^{**} (0.009)	社会资本	0.453 ^{***} (0.087)	1.674 ^{***} (0.247)	0.869 ^{***} (0.103)
家庭收入	0.351 ^{***} (0.018)	健康状况	0.209 ^{***} (0.059)	0.282 [*] (0.152)	0.359 ^{***} (0.067)
支出收入比	0.291 ^{***} (0.047)	城市	控制	控制	控制
城市	控制	所有制	控制	控制	控制
		行业	控制	控制	控制
		$\hat{\lambda}_0$ 或 $\hat{\lambda}_1$	0.503 ^{***} (0.002)	1.212 ^{***} (0.003)	
		σ_{0u} 或 σ_{1u}	-0.952 ^{***} (0.005)	-0.468 ^{***} (0.039)	
		创业			0.523 ^{***} (0.167)
常数项	-0.946 ^{***} (-0.086)	常数项	5.425 ^{***} (0.550)	15.850 ^{***} (1.468)	5.679 ^{***} (0.517)
样本量	13697	样本量	9716	3981	13697

注:***、**、*分别表示在1%、5%和10%显著性水平上显著,回归系数括号内为标准误。

的,这也从一个侧面反映了OLS模型和处理效应模型存在内在的缺陷。在表2的内生转换模型中,男性、社会资本和健康状况有助于提升不创业者和创业者的收入,但对创业者的贡献明显大于不创业者;受教育程度和城镇户籍有助于提升不创业者和创业者的收入,但对不创业者的贡献明显大于创业者,这与人们的直观感受一致。对不创业者而言,收入与工作经验呈倒“U”型关系,但对创业者而言,工作经验对收入的影响不显著。最后, $\hat{\lambda}_0$ 和 $\hat{\lambda}_1$ 的系数均为正, σ_{0u} 和 σ_{1u} 的系数均为负,都在1%的显著性水平上显著。 σ_{1u} 为负,这表明流动人口创业决策存在一个负向选择过程,也就是说,选择创业者若没有经过选择过程的话,会比实际上通过选择过程进行创业者的收入更高; σ_{0u} 为负,这意味着流动人口不创业决策也同样存在一个负向选择过程。这表明有必要纠正模型中由不可观测变量引起的样本选择性偏误。

为反映创业对流动人口收入的影响,我们根据方程式(7)和(8),可计算出解决样本选择性偏误

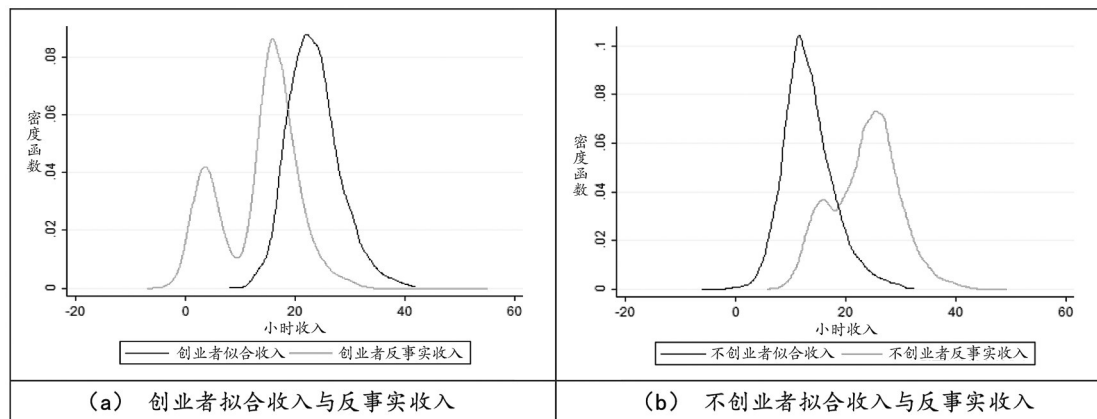


图1 两种情境下流动人口的收入概率密度

后,流动人口创业者和不创业者的收入,根据方程式(9)和(10)分别计算出创业者不创业、以及不创业者创业两种反事实情境下的收入水平。然后根据方程式(11)和(12)分别计算出ATT和ATU,数值分别为9.942和10.219,这表明创业能显著提高流动人口的收入,且不创业者进行创业,其增收的效果将更加明显。

为了更清晰地反映流动人口创业的收入效应,本文分别给出创业者和不创业者在两种情境下小时收入的概率密度分布,具体见图1(a)和图1(b)。图1(a)显示,创业者如果不创业,其收入的概率密度曲线会明显向左移动,且概率密度呈双峰分布,可能是由于创业者分为生存型和发展型两种类型,这表明剔除样本选择性偏误,创业者若不创业,其收入水平会明显下降,尤其是生存型创业者。图1(b)则显示,不创业者如果创业,其收入的概率密度曲线会大幅向右移动,且概率密度也呈双峰分布,

表3 内生转换模型(ESRM)的子样本估计结果

变量	创业决策模型		变量	子样本1			子样本2		
	子样本1	子样本2		不创业者	生存型创业者	OLS	不创业者	发展型创业者	OLS
性别	0.010*** (0.002)	0.028*** (0.006)	性别	1.973*** (0.152)	1.799*** (0.334)	2.254*** (0.131)	2.166*** (0.146)	3.463*** (0.712)	2.444*** (0.138)
教育	-0.280*** (0.015)	-0.226*** (0.020)	教育	2.903*** (0.097)	2.056*** (0.276)	2.364*** (0.089)	2.572*** (0.092)	2.653*** (0.525)	2.499*** (0.091)
工作经验	0.010*** (0.002)	0.005** (0.002)	工作经验	0.091*** (0.023)	-0.102 (0.073)	0.317*** (0.024)	0.227*** (0.024)	0.163 (0.166)	0.369*** (0.026)
婚姻状况	0.313*** (0.031)	0.271*** (0.045)	经验平方	-0.391*** (0.055)	0.025 (0.164)	-0.722*** (0.060)	-0.227*** (0.024)	-0.482 (0.379)	-0.828*** (0.065)
老家田地	0.162* (0.089)	0.263*** (0.098)	户籍	1.202*** (0.174)	0.266 (0.568)	1.746*** (0.210)	1.445*** (0.181)	1.965** (0.965)	1.963*** (0.215)
老家住房	0.016* (0.009)	0.021* (0.012)	社会资本	0.441*** (0.085)	1.504*** (0.275)	0.748*** (0.103)	0.478*** (0.092)	1.856*** (0.521)	0.693*** (0.106)
家庭收入	0.289*** (0.014)	0.305*** (0.016)	健康状况	0.186*** (0.058)	0.149 (0.164)	0.324*** (0.067)	0.215*** (0.063)	0.348 (0.353)	0.397*** (0.071)
支出收入比	0.205*** (0.047)	0.215*** (0.068)	城市	控制	控制	控制	控制	控制	控制
城市	控制	控制	所有制	控制	控制	控制	控制	控制	控制
			行业	控制	控制	控制	控制	控制	控制
			$\hat{\lambda}_0$ 或 $\hat{\lambda}_1$	0.433*** (0.002)	1.325*** (0.003)		0.216*** (0.001)	1.763*** (0.003)	
			σ_{0u} 或 σ_{1u}	-0.964*** (0.003)	-0.427*** (0.047)		-0.964*** (0.004)	-0.539*** (0.087)	
			创业			-0.179 (0.176)			2.359*** (0.253)
常数项	-0.781*** (0.087)	-1.378*** (0.121)	常数项	5.199*** (0.546)	15.375*** (1.605)	5.695*** (0.515)	6.110*** (0.532)	21.283*** (3.592)	5.075*** (0.533)
样本量	12767	10646	样本量	9716	3051	12767	9716	930	10646

注:***、**、*分别表示在1%、5%和10%显著性水平上显著,回归系数括号内为标准误。

这表明剔除样本选择性偏误,不创业者若创业,其收入水平会大幅上升,而且发展型创业者收入的增幅明显大于生存型创业者。

(二)子样本估计结果

全样本分析并未考虑不同类型创业者收入效应的差异,由于生存型创业者和发展型创业者的个体特征、创业目的等方面存在明显差异,如果将他们混为一谈,会模糊不同创业类型之间的差别。为进一步分析流动人口创业决策和收入决定的异质性,我们采用内生转换模型分别对子样本1和子样本2进行拟合,模型回归结果具体见表3。

从表3可知,流动人口创业决策模型的拟合结果与全样本估计结果非常接近,所有解释变量无论是系数方向,还是显著性水平都表现出高度一致性,但是,男性、老家田地和住房越多、家庭收入和支出收入比越高,选择发展型创业的概率越大。就流动人口收入决定方程而言,不创业者收入决定方程的拟合结果与全样本估计结果十分接近,各解释变量的系数方向和显著性水平都表现出高度一致性;但对创业者则存在差异,健康状况变量对生存型创业者和发展型创业者均不显著,且性别、受教育程度、城镇户籍和社会资本对发展型创业者的作用明显大于生存型创业者。表3中 $\hat{\lambda}_0$ 和 $\hat{\lambda}_1$ 的系数均为正, σ_{0u} 和 σ_{1u} 的系数均为负,都在1%的显著性水平上显著,这同样表明有必要纠正模型中由不可观测变量引起的样本选择性偏误。

(三)不同研究方法的比较分析

为了进行比较分析,表4给出了OLS模型、处理效应模型和内生转换模型(ESRM)中创业收入效应的回归结果。从OLS模型回归结果可知,创业在总体上使流动人口的小时收入提高了0.523元,但不同类型创业者的收入效应存在明显差异,创业使发展型创业者的小时收入提高了2.359元,对生存型创业者而言,创业的收入效应不显著。由于OLS模型忽视了流动人口创业决策的内生性问题,导致研究结论存在严重偏误。从处理效应模型回归结果可知,处理效应模型在一定程度上解决了内生性问题导致的样本选择性偏误,创业显著提高了流动人口的收入,就全样本而言,创业使流动人口的小时收入提高了7.016元,就分样本而言,创业对发展型创业者收入的提升作用远大于生存型创业者,但处理效应模型没有考虑到创业者与不创业者处理效应的异质性。

内生转换模型既可解决内生性问题导致的样本选择性偏误,又考虑到流动人口创业处理效应的异质性,是一种能严谨处理流动人口创业收入效应的研究方法。根据方程式(7)~(10),在解决样本选择性偏误后,我们可计算出全样本与分样本流动人口创业者和不创业者的收入、以及他们在反事实

表4 不同研究方法的比较

	全样本			分样本1			分样本2		
	OLS	处理效应模型	ESRM	OLS	处理效应模型	ESRM	OLS	处理效应模型	ESRM
回归系数 (或 ATE)	0.523*** (0.167)	7.016*** (0.464)	10.138	-0.179 (0.176)	5.066*** (0.529)	9.790	2.359*** (0.253)	10.864*** (0.756)	13.562
ATT			9.942* (5.113)			9.498* (5.651)			12.927** (5.126)
ATU			10.219** (4.274)			9.882* (5.812)			13.623*** (4.578)
样本量	13697	13697	13697	12767	12767	12767	10646	10646	10646

注:***、**、*分别表示在1%、5%和10%显著性水平上显著,回归系数括号内为标准误。ATE为ATT和ATU的加权平均值,故未给出显著性水平。

情境下的收入,然后根据方程式(11)和(12)分别计算出ATT和ATU。从表4可知,无论是全样本,还是分样本,流动人口创业的收入效应具有明显的异质性,ATT和ATU的值都显著为正,但ATU的数值都明显大于ATT,这表明对所有类型创业者而言,创业都能显著提高流动人口的收入,但不创业者创业的增收效果明显大于创业者。对比全样本和分样本回归结果发现,不同类型创业者的收入效应存在明显差异,发展型创业者创业收入效应明显大于生存型创业者。对比OLS模型和内生转换模型的回归结果可以发现,由于忽视了创业决策的内生性问题,OLS模型严重低估了流动人口创业的收入效应。对比处理效应模型和内生转换模型的回归结果可以发现,内生转换模型不仅考虑了样本自选择问题,而且考虑了处理效应的异质性,因而得到的分析结论较为精准可靠。

五、研究结论与政策启示

基于中国流动人口动态监测数据,将流动人口创业决策纳入实证分析框架中,并采用内生转换模型(ESRM)研究流动人口创业对其收入的影响。主要研究结论如下:1.基于不同研究方法的比较发现,内生转换模型既解决了内生性问题导致的样本选择性偏误,又考虑到流动人口创业处理效应的异质性,具有其他研究方法无可比拟的优势,是研究流动人口创业收入效应较为严谨的分析方法。2.流动人口创业决策模型显示,受教育程度对流动人口创业有显著负向作用,老家田地和住房越多、家庭收入和支出收入比越高,越倾向于创业,工作经验丰富、在婚和男性流动人口也有较高的创业倾向。3.对流动人口所有类型的创业者而言,创业都能提升其收入水平,但创业对流动人口的收入效应具有明显的异质性,不创业者创业时收入的提升幅度明显大于创业者创业时收入提升幅度。4.创业显著提升了流动人口的收入水平,对任一随机流动人口创业的平均处理效应为10.138,但不同类型创业者的收入效应存在明显差异,发展型创业者创业的平均处理效应明显大于生存型创业者。

上述研究结论有助于人们更好地理解流动人口创业决策过程和评估创业收入效应,对鼓励流动人口进行创业和建设创业型经济具有重要意义。一个基本的政策启示是要大力促进流动人口创业。

首先,鉴于受教育程度对流动人口创业存在显著的负向作用,需要从两方面提升教育对流动人口创业的正面影响。一方面要构建与“大众创业”相适应的创业教育新格局,培养国民的创业精神,创业教育不只是解决当前社会就业问题的权宜之计,更是推动经济社会持续发展的长远大计;另一方面要积极鼓励和支持高学历者创业,当前政府出台一系列支持高学历层次人员创业的政策,从侧面凸显了本研究的应用价值。

其次,要关注和扶持女性流动人口进行创业。女性流动人口在劳动力市场上兼顾流动人口和女性双重劣势,创业作为相对灵活的就业方式,不仅可以解决女性自身的就业问题,提升其收入水平,还可以对其子女未来的创业活动提供积极的影响,因此,关注和扶持女性流动人口创业对促进流动人口整体的就业具有积极意义。

再次,要加大力度扶持流动人口创业,尤其是发展型创业。发展型创业对营商环境的要求较高,需进一步弘扬企业家精神,鼓励更多家庭经济条件好、自身素质高的流动人口投身创业。同时,加快推动信用体系建设,降低信息不对称,缓解创业过程中的金融约束,增加流动人口创业的金融供给,加快完善流动人口社会保障和创业保险市场,减少和分散流动人口创业中的潜在风险。

最后,政府要积极帮助流动人口创业者解决在创业过程中遇到的困难和问题,要破除一切束缚创业的体制机制障碍,让每个有创业意愿的人都有自主创业空间,激发人们的创业热情,为流动人口创业营造良好的社会氛围,从而掀起“草根创业”的新浪潮,形成“大众创业”的新态势,进而促进社会纵向流动和公平,以实现社会平衡充分发展。✻

参考文献:

- [1] 王子敏. 互联网、技能偏向与农村流动人口就业[J]. 人口与经济, 2017(2): 107-115.
- [2] 莫旋, 阳玉香, 唐成千. 分层异质视角下流动人口收入决定研究——基于分层线性模型的实证分析[J]. 财经理论与实践, 2018(2): 123-129.
- [3] 宁光杰, 段乐乐. 流动人口的创业选择与收入——户籍的作用及改革启示[J]. 经济学(季刊), 2017(2): 771-792.
- [4] 翁辰, 张兵. 信贷约束对中国农村家庭创业选择的影响——基于CHFS调查数据[J]. 经济科学, 2015(6): 92-102.
- [5] 郝朝艳, 平新乔, 张海洋, 梁爽. 农户的创业选择及其影响因素——来自“农村金融调查”的证据[J]. 中国农业经济, 2012(4): 57-65+95.
- [6] 张博, 胡金焱, 范辰辰. 社会网络、信息获取与家庭创业收入——基于中国城乡差异视角的实证研究[J]. 经济评论, 2015(2): 52-67.
- [7] 魏下海, 陈思宇, 黎嘉辉. 方言技能与流动人口的创业选择[J]. 中国人口科学, 2016(6): 36-46+126-127.
- [8] 阮荣平, 郑风田, 刘力. 信仰的力量: 宗教有利于创业吗?[J]. 经济研究, 2014(3): 171-184.
- [9] 阳玉香. 自选择、政府培训与流动人口收入提高[J]. 教育与经济, 2017(4): 42-48.
- [10] Maddala G S. Limited-dependent and Qualitative variables in Econometrics[M]. Cambridge: Cambridge University Press, 1983: 80-81.
- [11] 王存同. 进阶回归分析[M]. 北京: 高等教育出版社, 2017: 253-258.
- [12] 李雪松, 黄彦彦. 房价上涨、多套房决策与中国城镇居民储蓄率[J]. 经济研究, 2015(9): 100-113.
- [13] 朋文欢, 黄祖辉. 农民专业合作社有助于提高农户收入吗?——基于内生转换模型和合作社服务功能的考察[J]. 西北农林科技大学学报(社会科学版), 2017(4): 57-66.
- [14] 孙猛, 许世存. 流动人口就业收入的禀赋效应与地区效应分析[J]. 人口学刊, 2018(6): 52-61.
- [15] 侯建明, 李晓刚, 叶淑萍. 吉林省流动人口收入状况及其影响因素分析[J]. 人口学刊, 2016(6): 54-61.
- [16] 周京奎, 黄征学. 住房制度改革、流动性约束与“下海”创业选择——理论与中国的经验研究[J]. 经济研究, 2014(3): 158-170.
- [17] 彭艳玲, 孔荣, Calum G. Turvey. 农村土地经营权抵押、流动性约束与农户差异性创业选择研究——基于陕、甘、豫、鲁1465份入户调查数据[J]. 农业技术经济, 2016(5): 50-59.

**Floating Population's Entrepreneurship and Income Effect:
From the Perspective of Self-selection and Heterogeneity**

MO Xuan¹, ZHOU Jian²,

(1. Hengyang Normal University, Hengyang Hunan 421008, China;

2. Shanghai University of Finance and Economics, Shanghai 200433, China)

[Abstract] Based on the dynamic monitoring data of China's floating population and using the endogenous switching regression model (ESRM) to make up the shortages of the existing research methods, and attempts to bring entrepreneurship into the analysis framework. From the perspective of self-selection and heterogeneity, Study the income effect of floating population's entrepreneurship. The study found that: Factors such as the gender and education of floating population affect both their entrepreneurship and income. control the sample selection bias, Entrepreneurship has raised the income of floating population. However, the income effect of entrepreneurship on floating population is significantly heterogeneous, the income effect of non-entrepreneur is greater than entrepreneur, developmental entrepreneur is greater than survival entrepreneur. This study provides a scientific basis for accurately assessing the entrepreneurial income effect of the floating population and has certain policy implications for encouraging floating population to start their own businesses and establish an entrepreneurial economy.

[Key words] Entrepreneurship; Floating Population; Income Effect; Self-selection; Heterogeneity