

社会痛点信息约束对大学生创业预期的影响

——动态性创业的时间成本与有序 Logistic 回归

戴尔

(广州大学经济与统计学院, 广东广州, 510006)

[摘要] 基于“理性经济人”假设, 以大学生如何选择创新创业方向为导向, 提出以下基础假设: 通过缓解大学生创业者的社会痛点信息约束, 可以降低创业时间成本并提高整体创业意愿, 创新性地将创业所产生的动态时间成本引入创业意愿的成本分析框架, 弥补了现有文献局限于静态分析的不足。通过收集数据并结合 R 型聚类和有序 Logistic 模型来进行实证检验, 实证结果较好地验证了研究中提出的理论假设, 并据此得出: 宏观领域中的创业帮扶政策不能局限于“资金帮扶”, 还要做到“信息帮扶”; 在微观领域可以适当建立地方的社会痛点信息库, 让大学生创业者了解更多的社会需求, 降低他们收集信息的成本, 最终驱动创新创业的发展。

[关键词] R 型聚类; 创新创业; Logistic 模型; 痛点信息

[中图分类号] F49 **[文献标识码]** A **[文章编号]** 1674-893X(2023)02-0032-13

一、引言与文献综述

近几年国家提出创新驱动发展战略并大力支持大学生创新创业, 目的既是为科技发展和社会进步注入新鲜血液, 也是为社会创造更多就业机会^[1], 同时还意在加快科技成果的转化从而推进中国的经济转型^[2]。因此, 积极探寻创新创业预期的影响因素, 并借此提高社会整体的创新创业意愿从而增加创新创业的数量和质量, 对于缓解新冠疫情冲击下市场就业率低迷问题^[3]、美国在高端科技产业对中国的“卡脖子”问题^[4]都有着重要的现实意义, 也跟《政府工作报告(2020)》中所提及的“稳就业”和“稳预期”的中心思想不谋而合。

在关于如何提高社会创新创业意愿的研究中, 大多数学者的研究思路主要是围绕着创业者的内部特征和外部环境进行理论和实证的归因分析。在研究创业者的内部特征方面, 蔺楠、张

茜和周炜^[5]在计划生育时滞影响的背景下对独生子女创业意愿的影响机制展开实证研究, 发现子女的独生性会左右父母的教育方式, 使得独生子女的风险规避倾向高于一般小孩, 最终影响他们的创新创业预期。与此同时, 还有一些学者通过实证研究, 认为社会对性别的歧视会让创业者的创业心理评估产生负效应和心理成本, 女性可能会因为受到性别刻板印象的影响而主动逃离创业活动所带来的心理威胁^[6]。除此之外, 户籍所在地的地理特征和人文因素也有可能影响创业者的个人特性进而左右其创业决策^[7]。

在研究创业者的外部环境方面, 刘敏佳^[8], Anwar 和 Chan^[9]等学者在以往的理论基础上进行调查验证了外部环境的相关教育与政策在创业方面会对社会公众产生积极作用。曾婧婧和温永林^[10]对创新城市试点这一“准自然实验”进行深入探索和实证分析后, 发现创新城市的建

[收稿日期] 2022-03-22; **[修回日期]** 2022-11-25

[作者简介] 戴尔, 男, 广东茂名, 广州大学经济与统计学院学生, 主要研究方向: 金融与创新创业, 联系邮箱: 173824808@qq.com

设以及创业政策的持续优化能够提高社会的创新创业预期。而匡远凤^[11]基于熊彼特经典创新模型对返乡农民的创业活动进行分析后，发现土地、资金、人才、信息和技术等生产要素的流动也会牵动着创业成本的上下浮动来影响社会的创业热情，进而影响其创业行为决策。

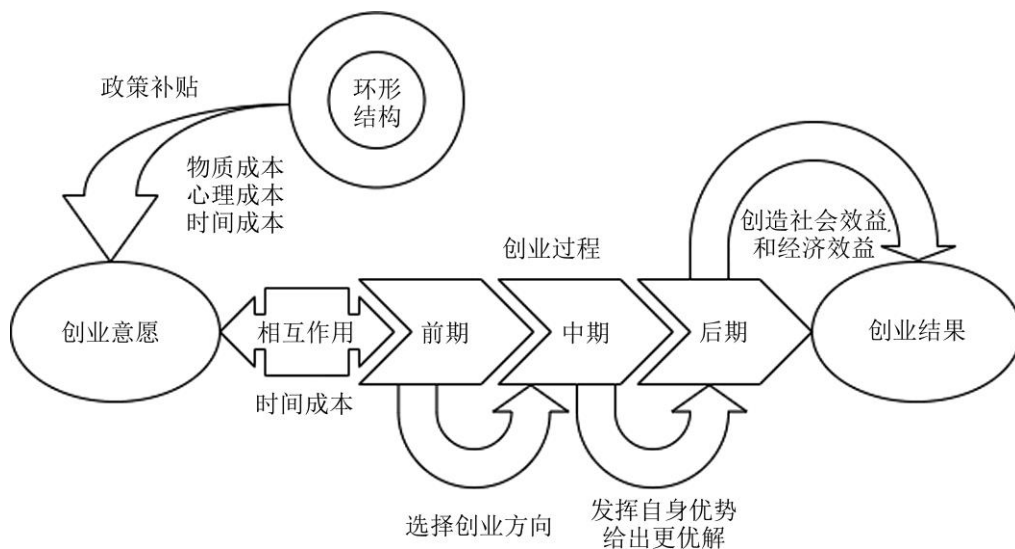
从当前的文献来看，提高创业意愿的根本机制就在于降低创业成本或者提高创业红利，其中，创业成本包括物质成本、心理成本以及时间成本，而创业红利包括创业平均收益和政策补贴。通过降低创业成本或提高创业红利可以帮助大学生群体感知创业价值，进而有效提高其创业意愿，从而完成驱动创业的目标^[12]。

而现阶段大部分关于创业意愿的研究主要是在成本机制分析框架下开展的，一般表现为通过对降低创业成本的归因来反溯创业意愿的影响因素。从学术界相关的研究内容来看，大部分文献主要是以创业者为核心向里向外延拓相关研究方向，由创业者的内部特质^[6]和外部的创业环境^[13]两方面组合成了一个如图 1 所示的环形研究结构。从知网的检索结果来看，近期的相关研究一方面是在这个环形结构下探寻新的影响因素并进行相关分析^[14]，另一方面是在引用新的经济理论并优化相关的数学模型和统计方法来对以往的研究结论做出进一步的数理解释^[11]。



图 1 创业意愿主流研究环形框架图

以创业者为核心的环形结构看似可以完美囊括成本机制下对创业意愿影响因素的所有研究，但实则却没有考虑到一系列创业流程和步骤对创业意愿的交互作用。创新创业作为一种实践活动，应该关注的是创业者与外界信息互动、创业步骤落实的过程^[15-16]，而不只是创业者本身。因此，在探索创业意愿影响因素的过程中，也不应该局限于创业者的内外特征，还要考虑到创业活动前期，比如相关痛点信息的收集、创业方向的选择等步骤所产生的时间成本对创业意愿的影响(见图 2)，而目前尚未有学者从这个角度来考究提高社会创业意愿的影响因素及其相关机制。



基于当前研究所存在的不足,本文在对比近些年来社会创业研究成果的基础上,结合实际创业过程的动态特征,拟定从创业前期的准备步骤——创新创业方向的选择——出发来研究在实际创业过程中该步骤对社会创新创业意愿的影响,并且通过对全国大学生展开抽样调查以及结合R型聚类和有序Logistic回归来为此研究提供数理证明,探索在动态视域中提高创新创业意愿的相关影响机制,最终在国家积极推动“双创”的背景下为政府驱动创新创业提供实证依据以及政策启示。

本研究存在的贡献有:第一,以创业步骤而不是创业者本身作为研究对象,从动态演变的角度考究了社会创新创业意愿的影响因素及其影响机制,是对当前创业意愿研究领域所存在的不足的有益补充。第二,设置一、二级指标来表示所研究的变量,并采用R型聚类和李克特五级量表对其进行修正和量化,有效消减指标之间的相关性过强的问题,最后再结合有序Logistic回归方法为本文结论提供实证依据。第三,本研究从创业步骤的角度探索了提高大学生创业意愿的影响因素及其相关机制,为政府或者高校提高大学生创新创业的数量和质量、完成驱动创业的目标提供了新的宏观政策制定角度以及微观调控行为的切入点。

二、理论分析和研究假设

首先,笔者认为创业前期的灵感发掘过程所产生的时间成本会对创业者的创业意愿产生影响,并且创业灵感的发掘也会影响创业方向的选择进而决定整个创业项目质量的高低。有学者认为创新创业的实质就是以可持续的、创新的方式探索和开发新的生产力来满足社会痛点并创造经济效益^[15],是遵循从感知创业机会、确定创业方向到组建创业团队、获取创业必需资源这个逻辑进程的^[17]。由此可见,大部分学者都认为发掘创业灵感、寻找社会痛点并确定创业方向是创业过程中最为基础的一部分。据《中国青年创业现状报告(2016)》统计的数据,47.2%的青年创业者认为好的创业方向是成功创业的关键因素。因

此,如何增加创业者发掘灵感的渠道,让广大创业者降低在寻找创业方向上的时间成本,使他们能便捷地找到适合自己的创业项目,对于提高社会创新创业的热情与信心来说是举足轻重的。有学者在研究智能手机的普及对农村家庭创业的影响机制时,提及智能手机可以通过缓解信息约束、降低创业成本来提高农村家庭创新创业的意愿^[14],而通过普及互联网也能凭借相似的信息约束缓解机制来激发社会的创业欲望^[18]。鉴于此,本文在现有的理论基础上结合创业前期灵感发掘的行为路径,认为通过缓解创业者在社会痛点信息方面的约束,可以降低创业者发掘创业灵感过程的时间成本,从而能够帮助创业者更广泛地了解社会需求,进而提高其信息资源禀赋,在理论上可以有效地增强创业者的创业意愿。

而在关于痛点信息约束方面的研究中,目前学界尚未对其形成系统的理论分析,大部分的相关研究仍是基于字面意义来理解“痛点信息”这个概念。本文化用了范炜昊和徐健^[19]对“痛点”的定义,认为“痛点信息约束”就是一种“被大多数人反复表述过的一个有待去满足但由于时间、空间的限制无法被创业者所感知的原始需求”。如何衡量并缓解这种认知约束,使得创业者和痛点信息紧密地结合到一起从而驱动大学生创业,是有待本文深入研究的问题。为此,本研究将围绕着痛点信息以及创业前期的方向选择问题,从动态演化的角度构建整合度、公示度、获取便捷度等三个维度来衡量痛点信息对大学生创业者的约束程度,并结合相关文献构建了如图3所示的理论机制图。

基于痛点信息对创业活动的互动模型以及上述的理论分析,本文提出了如下假设:

H₁: 提高痛点信息的整合度,在同一条件下使得大学生创业者有机会获得密度更高的痛点信息,有利于降低创新创业的时间成本,从而提高大学生的创新创业预期。

H₂: 提高痛点信息的公示度,在同一条件下可以提高大学生创业者获得痛点信息的概率,有利于降低创新创业灵感来源的不确定性,进而提

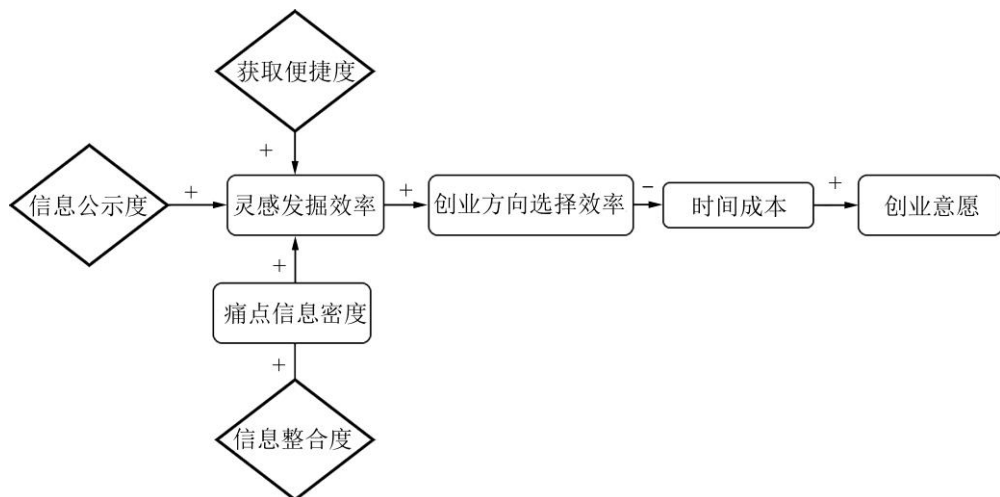


图3 理论机制图

高大学生的创新创业预期。

H₃: 提高痛点信息的获取便捷度, 在同一条件下可以缩短大学生创业者获取痛点信息的时间, 有利于降低创新创业的时间成本, 从而提高高大学生的创新创业预期。

三、实证分析

(一) 变量设置与数据

被解释变量: 创业预期。在国家大力推行创新创业的背景下, “创业”一词在日常语义中略带褒义, 而这可能会成为被调查者在填写问卷时产生自我欺骗行为的动机^[20], 如果直接询问被调查者的创业意愿, 调查结果可能会由于被调查者的自我欺骗行为而产生较大误差。为了避免出现这种情况, 本文选择了被调查者对社会大众的创业预期作为本研究的因变量。“预期”代表着人们对未来变量的估计, 从行为经济学的角度来看, 人们对某个变量的预期越高, 在未来则有更大的概率会朝着该预期的方向展开实践活动。也就是说, 创业预期在一定程度上能够衡量社会创业的热情和意愿, 并且能够有效规避被调查者的自我欺骗行为。

解释变量: ① 痛点信息整合度评估。为了更加全面地考究社会痛点信息整合度在被调查者心中的主观评估, 本研究并没有直接询问被调查者关于信息整合度的直接感受, 而是通过“你认为当下网络中关于社会痛点信息的分布是”“你对社会问题的总结与思考的频率”“你认为

在获取某一特定方面的社会痛点时所遇到的检索难度”“集中了解社会发展痛点是否会对创新创业产生促进作用”等四类问题构成痛点信息整合度评估的二级指标, 进而综合衡量被调查者对痛点信息整合度的主观感受。② 痛点信息公示度评估。同样, 本研究设计了三个二级指标来度量被调查者的信息公示度评估, 分别为: “您认为当下网络对社会问题的揭露程度如何”“您对披露社会痛点信息平台的了解情况”“您对当下社会痛点问题的了解程度”。③ 痛点信息获取便捷度评价。本研究设计了两个二级指标来度量被调查者的信息公示度评估, 分别为: “评估一下您在生活中获得相应痛点信息时所花费的时间长短”“您在日常生活中浏览社会痛点信息库的频率”。

笔者采用李克特五级量表对以上被解释变量和部分二级指标进行分级量化处理, 其中的关键变量由二级指标通过加权平均再结合分组的方式来量化。

控制变量: 本文为了简化分析模型并提高研究结论的广泛适用性, 决定从现有的研究中抽取少量代表性高的创业意愿影响因素作为控制变量(协变量)来进行实证分析, 从而进一步提高研究的严谨性和可靠性。通过查阅文献, 笔者发现在创业意愿内部特征影响因素中, 性别是最常被提到的研究对象之一, 同时也是对创业意愿最主要的影响因素之一, 甚至有研究表明在大学生群

体中性别差异所导致的两性之间的创新创业预期差异高达 2.437 倍^[21]。性别不仅能够影响个体的性格,还能通过社会刻板印象^[6]、自身的能力认知^[22]等渠道来左右个体的创新创业决策。而在对创业意愿外部环境的影响因素研究中,户籍是最常被提到的研究对象之一。个体的户籍所在地往往是个体出生以及成长的地方,所谓“一方水土养一方人”,户籍所在地的差异通常会通过金融环境的差异^[23]、政策环境的差异^[10]和创业氛围

的差异^[24]等来影响该区域在创业上的物质成本和时间成本,进而干扰不同个体的创业意愿,从而左右其创业决策。因此,本研究将性别和户籍性质这两个核心创业意愿的内外影响因素作为控制变量纳入实证模型当中,由此分析这两个变量对回归结果的影响,并对比原始模型的回归结果,来寻找并剔除它们对最终回归结果的潜在干扰。

以上变量如表 1 所示。

表 1 变量概述

因变量 Y	大学生创新创业形势预期(创新创业意愿)
自变量 X	你认为当下网络中关于社会痛点信息的分布是 你对社会问题的总结与思考的频率 你认为获取某一特定方面的社会痛点时所遇到的检索难度 集中了解社会发展痛点是否会对创新创业产生促进作用
	你认为当下网络对社会问题的揭露程度如何 您对披露社会痛点信息平台的了解情况 您对当下社会痛点问题的了解程度
	评估一下您在生活中获得相应痛点信息时所花费的时间长短 您在日常生活中浏览社会痛点信息库的频率
控制变量 Z	性别 Z ₁ 户籍性质 Z ₂

本次实证的样本数据来源于对全国高校大学生的抽样问卷调查。笔者共回收 143 份有效问卷,通过此途径不仅获取了被调查者的性别、户籍、自变量指标、因变量信息,还收集了他们对整合社会痛点并构建痛点信息库这个政策行为的态度,其中有 65.03% 的大学生认为构建痛点信息库有一定可能会激发创业灵感的发掘,仅有 23.77% 的人觉得可能性不大甚至没有可能。从初步的调查数据来看,大部分大学生认为社会痛点信息是激发创业灵感的影响因素之一。接下来,本文将围绕社会痛点信息的九个二级指标以及三个关键变量来构建有序 Logistic 回归模型,进而从数理的角度深入地验证这种激发灵感的有利因素是否能进一步影响创业者的创业意愿,最终为政府制定新的宏观创业扶持政策和微观调控行为方案提供理论和实证依据。

(二) 模型构建与实证思路

线性回归是常见的实证分析方法,它的算法本质可以理解为通过最小化 Y 预测值与真实值之间的误差来算出自变量和因变量之间的权重关系,进而得到最优的线性参数。而 Logistic 回归不同于线性回归,它的算法本质是根据历史数据求出因变量等于某个值时的概率,进而确定自变量和因变量之间的数量关系。因此,Logistic 回归的因变量常常是离散型变量,并且没有自变量与因变量之间必须为线性关系的强假设,也不要求变量完全满足正态分布。Logistic 回归方法的使用条件符合本研究的数据特点,并且对强假设的要求较低,拟合结果比较符合现实,故选用其作为本研究的主要统计工具。Logistic 回归模型一般以因变量是否呈现序列形式而被分为有序 Logistic 回归以及无序 Logistic 回归两种。考虑到本文的数据特点,本研究计划使用有序 Logistic

回归来对比展现实证数据的统计特征，最终为前文的理论假设提供数理证明。

本实证研究将因变量 Y 分成五个等级，分别是“很好”“比较好”“一般”“较差”“很差”，并计划先对二级指标进行聚类分析，主要目的是筛选并剔除相关性强的二级指标，进而提高本研究的科学性和准确性。接着再将聚类后的二级指标与因变量 Y 进行有序 Logistic 回归分析，探明在微观领域下影响大学生创业预期的行为因素。与此同时，本文还会将这些二级指标进行加权平均和分组分级来分别量化表示痛点信息整合度评估(X_1)、痛点信息公示度评估(X_2)、痛点信息获取便捷度评价(X_3)等三大关键变量，并将其对 Y 进行有序 Logistic 回归分析，进而在宏观层面上探明大学生创新创业预期的影响因素，为提高社会整体创新创业预期的政策启示提供实证依据。

本文第一部分实证模型计划把经过聚类分析的二级指标对 Y 进行有序 Logistic 回归，构建出模型一的回归结构，其表达式如：

$$P(y = j | X) = \frac{1}{1 + \exp(\alpha + \beta x_{ik})} \quad (1)$$

其中， i 表示第 i 个关键变量($i=1, 2, 3$)， k 表

示第 i 个关键变量下的第 k 个二级指标。而 j 表示大学生创新创业形势预期程度($j=1, 2, 3, 4, 5$)， j 越大，说明样本对社会创新创业形势的预期越高，其创业意愿就越明显。将上式转化为累计 Logistic 模型后可得：

$$\text{Logit}(P_j) = \ln \frac{P(y \leq j | x)}{P(y \geq j | x)} = \alpha + \beta X_{ik} \quad (2)$$

本文第二部分实证计划先对二级指标进行加权平均和分级量化，由此求出三大关键变量 X_i 的具体数值后，再让其对 Y 进行有序 Logistic 回归，进而构建出第三个回归模型，其表达式如：

$$\text{Logit}(P_j) = \ln \frac{P(y \leq j | X)}{P(y \geq j | X)} = \alpha_j + \sum_{i=1}^j \beta_j X_i \quad (3)$$

其中， P_j 表示 $y=j$ 时的概率， X 代表着三大关键变量，代入样本数据之后，可以求得三大关键变量对不同创业预期的回归系数 α 和 β ，进而可以逆推求出当 $y=j$ 时的概率值为：

$$P(y \leq j | X) = \frac{\exp(-\alpha_j - \sum \beta_j X_i)}{1 + \exp(-\alpha_j - \sum \beta_j X_i)} \quad (4)$$

上述实证过程的思路如图 4 所示。

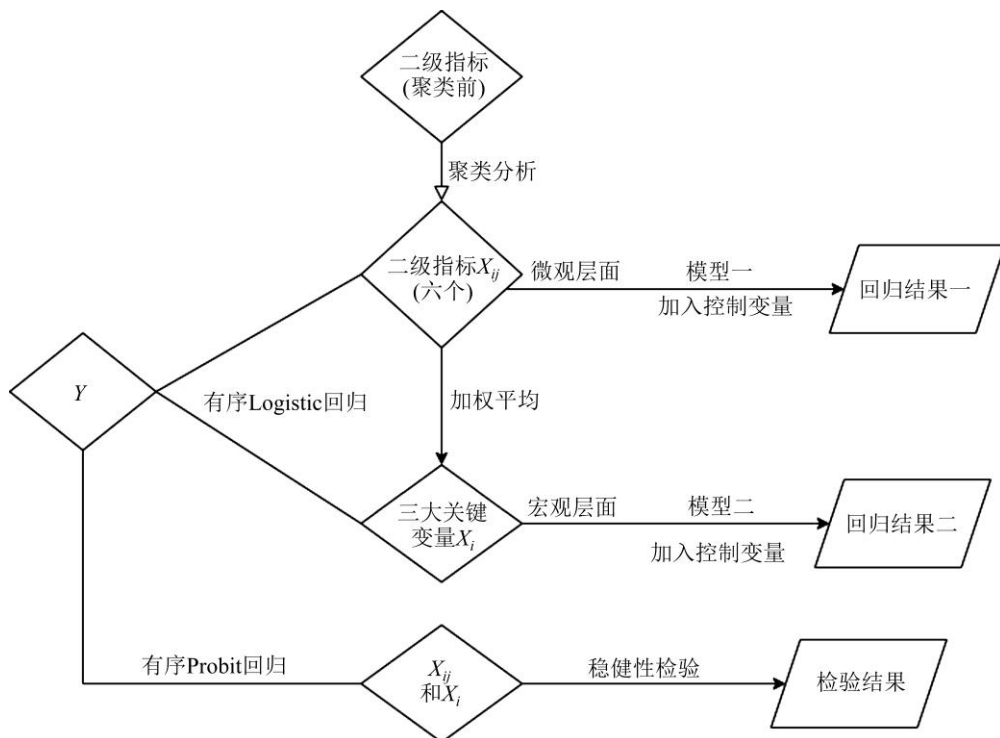


图 4 实证思路图

(三) 聚类分析

由于本文的实证分析在三个关键变量下设置了九个二级指标,为了防止二级指标之间存在较强的相关性从而对本文的回归分析造成较大的干扰,本研究需要对这九个二级指标进行聚类分析,进而削减部分同质性强的指标从而降低二级指标间的相关性,最终通过加权平均的方式来量化关键变量 X 。

聚类分析的种类有很多,根据本研究降低变量相关性的聚类目的,本次实证选择采用 R 型聚类方法来对多个二级指标进行聚类分析。 R 型聚类的基本原理是依据变量之间的相关性将其逐级分成多个大类,最后再对其进行筛选和删减进而剔除高相关性二级指标。由于各项二级指标都是由有序级别来进行衡量的,具有一致的量纲,因此本研究可以使用欧式距离来计算各项二级指标间的差异度,再凭借算出的差异度来对这些二级指标进行聚类化整。

本次聚类分析使用了最小距离法对由欧式

距离计算得出的差异度进行聚类划分,先把差异度最小的两个指标归为一个类,再把差异度最小的两个类归为一个较大的类,直到把全部的二级指标都归为一个类别为止。

将上述理论分析转化成代码输入 R 软件(4.03)之后,可以得到:在合并距离为 5 的水平线下, X_{21} 和 X_{23} 因为差异度较小而被归为一类, X_{13} 和 X_{14} 被归为一类, X_{11} 和 X_{31} 被归为一类。由此我们可知,第五个和第七个二级指标、第三个和第四个二级指标、第一个和第八个二级指标之间存在着较强的相关性,可以进行适当的修正,以确保整个实证检验的有效性和合理性。聚类结果如表 2 所示。

根据聚类结果以及实际需要,本文将去掉 X_{11} 、 X_{14} 和 X_{23} 这三个二级指标,剩下六个二级指标将通过加权平均和分组的方式来分别量化关键变量 X_1 、 X_2 、 X_3 ,为最终的有序 Logistic 回归提供样本特征,进而估计回归模型的参数并得到与模型相关的显著性水平。

表 2 R 型聚类结果

因变量 Y	大学生创新创业形势预期(创新创业意愿)	
自变量 X	痛点信息整合度评估	你对社会问题的总结与思考的频率 X_{12} 你认为获取某一特定方面的社会痛点时所遇到的检索难度 X_{13}
	痛点信息公示度评估	您认为当下网络对社会问题的揭露程度 X_{21} 您对披露社会痛点信息平台的了解情况 X_{22}
	痛点信息获取便捷度评价	评估一下您生活中获得相应痛点信息时所花费的时间长短 X_{31}
		您在日常生活中浏览社会痛点信息库的频率 X_{32}
控制变量 Z	性别 Z_1 户籍性质 Z_2	

(四) 回归分析

本文将构建四个 Logistic 回归模型,一个是二级指标对因变量的有序 Logistic 回归模型,目的是分析微观视角下各个二级指标与创新创业预期之间的关系。第二个模型是在第一个模型的基础上加入被调查者的性别和户籍性质作为控制变量,目的是将常见的创业预期影响因素纳入分析模型中,既能够分析这两个变量对回归结果的影响,同时也能够剔除它们对最终回归结论的潜在干扰。第三个模型是三大关键变量对因变量

的有序 Logistic 回归模型,创建该模型的目的是在较为宏观的角度上分析三大关键变量与创新创业预期之间的关系。模型四则是在模型三的基础上加入被调查者的性别和户籍性质作为控制变量,与模型三的回归结果产生对比,使本文的实证分析更加贴近实际情况。

1. 第一部分实证——二级指标与创新创业预期

这部分实证的主要内容是将聚类后的二级指标跟创新创业预期进行有序 Logistic 回归,分

析二级指标对创新创业预期的影响情况。另一个模型则是在前者的基础上加入创业者的性别和户籍性质作为控制变量，在分析这两个变量对回

归结果影响的基础上防止它们对最终的回归结论产生未知干扰。模型 1 的数据描述性统计如表 3 所示。

表 3 模型一描述性统计表格

自变量	样本数	取值范围	平均值	标准差	最小值	最大值
创业预期	143	1—5	3.210	0.740	1	5
X ₁₂	143	1—5	3.035	0.800	1	5
X ₁₃	143	1—5	3.643	0.791	2	5
X ₂₁	143	1—5	2.441	0.601	1	4
X ₂₂	143	1—5	2.224	0.696	1	4
X ₃₁	143	1—5	3.210	0.941	1	5
X ₃₂	143	1—5	3.364	1.417	1	5
性别	143	男性=1 女性=2	1.364	0.483	1	2
户籍	143	农村户口=1 城市户口=2	1.357	0.481	1	2

将收集上来的数据进行预处理并导入 SPSS 26.0 后，先对第一个模型进行有序 Logistic 回归、平行线检验和似然比检验，其中的平行线检验结果如表 4 所示。从中可得平行线检验的显著性水平为 0.940、 $\chi^2 = 49.112$ ，表明模型一满足有序 Logistic 回归的条件，可以通过六大二级指标来对创新创业预期 Y 进行有序 Logistic 回归。

表 4 模型一平行线检验

模型	-2 对数似然	卡方	自由度	显著性
原假设	262.912			
常规	213.800	49.112	66	0.940

从模型的似然比检验结果来看，数据检验的卡方为 29.951，显著性等于 0.120，说明在较小误差下可以认为模型自变量的回归系数不为零，样本数据能够有效地表现出自变量和因变量之间的关系，形成合理的统计规律，进而使模型一得以成立(见表 5)。

当模型通过了水平线检验、似然比检验之后，就可以进一步分析模型的参数估计值，其中模型一的回归结果如表 6 所示。

表 5 模型一似然比检验

模型	-2 对数似然	卡方	显著性
仅截距	292.863		
最终	262.912	29.951	0.120

表 6 模型一参数估计情况

变量	回归系数	OR 值	t 值	显著性级别
X ₁₂	-0.111	0.89	-0.45	
X ₁₃	0.045	1.04	0.20	
X ₂₁	0.198	1.21	0.63	
X ₂₂	0.576	1.77	2.02	**
X ₃₁	-0.030	0.97	-0.16	
X ₃₂	-0.053	0.94	-0.41	
性别	-1.112	0.32	-2.97	***
户籍性质	-0.076	0.92	-0.21	
Prob>chi2			0.04	
Observations			143	
F test			0.036 0	

注：* $p < 0.05$ ，** $p < 0.01$ ，*** $p < 0.001$ ；OR 值表示自变量对因变量的边际影响程度。

由表 6 中的模型回归结果可以看出，整个回归模型的 P 值为 0.04，小于 0.05，说明模型整体的统计效果显著，样本数据能够有效地表现出自

变量和因变量之间的关系，形成合理的统计规律。 X_{22} 、 Z_1 这两个变量在模型一中对 Y 的回归显著性较强且 OR 值大于零，结合有序 Logistic 回归原理，可以分别在 1% 和 0.1% 的误差内认为，大学生对痛点信息平台的了解程度和大学生性别对创业预期有着正向影响，说明增加大学生对披露社会痛点信息平台的了解可以增强潜在创业者的创业意愿，进而驱动大学生开展创新创业活动。这也容易理解，社会痛点信息库可以为创业者提供社会痛点信息，使其了解产业发展的不足之处，一方面能够减少创业者收集市场信息所花费的时间进而降低创业者的创业成本，最终提高社会创业意愿，另一方面还可以通过共享高密度的痛点信息来给予创业者不同的思考角度，从而帮助他们打开创业思路并且树立创业自信和激发创业热情。

另外，根据 OR 值推算出 X_{22} 、 Z_1 具体的创业

意愿转换效率，每提升一级大学生对社会痛点信息库的了解程度，至少使其整体的创新创业意愿提高 1.77 倍。与此同时，男性大学生的创业意愿比女性大学生高 32%，而其余的二级指标在所有水平下所表现出的 P 值都比较大，可知它们对因变量都没有统计学意义或者统计学意义较小，无法提供统计参考。当然也不排除可能是实证样本数量较少，现有的数据无法全面地展现总体的情况。

2. 第二部分实证——三大关键变量与创新创业预期

这部分实证首先把聚类后的二级指标经过加权平均并分级来求得三大关键变量 X ，进而将其跟因变量进行有序 Logistic 回归，分析痛点信息的整合度 X_1 、公示度 X_2 、获取便捷度 X_3 对创新创业的预期影响情况。模型二的数据描述性统计如表 7 所示。

表 7 模型二描述性统计情况

自变量	样本数	取值范围	平均值	标准差	最小值	最大值
创业预期	143	1—5	3.210	0.740	1	5
X_1	143	1—5	2.497	0.659	1	4
X_2	143	1—5	2.126	0.626	1	4
X_3	143	1—5	2.762	0.822	1	5
性别	143	男性=1 女性=2	1.364	0.483	1	2
户籍性质	143	农村户口=1 城市户口=2	1.357	0.740	1	2

将经过处理之后的三大关键变量导入 SPSS 26.0 后，对其进行有序 Logistic 回归，并通过了平行线检验和似然比检验，其中，似然比检验结果如表 8 所示。从中可得似然比检验的显著性水平为 0.011、 $\chi^2 = 29.958$ ，说明在极小的误差下可以认为关键变量 X 的回归系数不全为零，样本数据能够有效地表示自变量和因变量之间的关系，能形成合理的统计规律进而使得模型成立。

表 8 模型二似然比检验

模型	-2 对数似然	卡方	显著性
仅截距	175.507		
最终	152.549	22.958	0.011

当模型通过了显著性检验之后，就可以进一步分析模型的参数估计值，其中，模型二的回归结果如表 9 所示。

由表 9 中的回归结果可以看出，整个回归模型的 P 值为 0.02，远小于 0.05，说明模型的统计效果显著，样本数据能够有效地表现自变量和因变量之间的关系，形成合理的统计规律。 X_2 、 Z_1 这两个变量在模型一中对 Y 的回归显著性较强，且 OR 值大于零，结合有序 Logistic 回归原理可以 0.1% 的误差内认为 X_2 、 Z_1 对创业预期有着正向影响，说明提高社会痛点消息的公示度可以加强潜在创业者的创业意愿，进而驱动大学生

表9 模型二参数估计表

变量	回归系数	OR 值	t 值	显著性级别
X ₁	0.067	1.06	0.24	
X ₂	0.810	2.24	2.82	***
X ₃	-0.136	0.87	-0.66	
Z ₁	-1.197	0.30	-3.21	***
Z ₂	-0.186	0.82	-0.53	
Prob>chi2			0.02	
Observations			143	
F test			0.00203	

注：* $p < 0.05$ ，** $p < 0.01$ ，*** $p < 0.001$ ；OR 值表示自变量对因变量的边际影响程度。

开展创新创业活动。本文的假设 H₂ 得以验证。而本研究其他关键变量 X₁、X₃ 的 P 值较大，对因变量的统计关系不显著。也就是说，提高社会痛点信息的集中度和获取便捷度，可能对大学生的创业意愿没有积极作用，假设 H₁、H₃ 在模型二中没有得到证实。

另外，笔者根据 OR 值还推算出 X₂、Z₁ 具体的创业意愿转换效率，每提升一级社会痛点信息的公示程度，至少使大学生群体的创新创业意愿提高 2.24 倍，而 Z₁ 的 OR 值为 0.3，与第一部分实证结果相近，说明男性的创业意愿平均比女性高 30%。

值得说明的是，在两次回归中，作为控制变量的户籍性质仍然表现为统计上的不显著，与其他文献的结论恰好相反。这说明，无论是在微观层面还是在宏观层面，个体的户籍性质对大学生的创新创业预期都没有统计关系。但这不能说明外部环境对大学生的创业预期没有显著影响，因为近几年国家逐渐深化户籍改革的政策行为在一定程度上使得处于农村环境的个体农村户籍转变成了城市户籍，这在一定程度上削弱了外部环境与户籍性质的相关性，最终可能导致本研究关于户籍性质的实证结果与以往的研究结论有所不同。另外要补充的一点就是，在国家大力扶持“双创”活动的背景下，创业政策的扶持对象和扶持地区正在逐渐放宽，而这又使得不同地区的创业成本慢慢地趋于同一水平，进而减弱了户籍所在地对个体创业预期的差异性影响。这些因素可能会导致户籍性质在两次回归中对创业预

期表现为弱显著性，使得本实证得到了与以往研究结果相悖的结论。

(五) 稳健性检验

因为本文上一部分的内容已经对实证模型进行了适用性检验，判断出数据与模型具有高契合度以及模型本身的存在具有强统计性，因此笔者本部分主要是对模型的实证结果进行稳健性检验。传统的稳健性检验方法有很多，主要的类型有变换数据法、替换变量法以及相似回归法。由于本实证所用的 Logistic 回归方法与 Probit 回归方法在数理上有着相近之处，因此笔者决定采用相似回归法，使用有序 Probit 回归对实证数据进行模拟检验，进而判断原回归方法的有效性和指标解释能力的强弱。两次实证模型的稳健性回归结果分别见表 10 和表 11。

表10 Probit 回归参数估计表(一)

变量	回归系数	标准差	z 值	显著性级别
X ₁₂	-0.003	0.13	-0.02	
X ₁₃	-0.069	0.12	-0.58	
X ₂₁	0.115	0.18	0.65	
X ₂₂	0.278	0.16	1.72	**
X ₃₁	-0.001	0.10	-0.01	
X ₃₂	-0.029	0.07	-0.40	
性别	-0.613	0.20	-3.02	***
户籍性质	-0.024	0.20	-0.12	
Prob>chi2			0.05	
Observations			143	
F test			0.0360	

注：* $p < 0.05$ ，** $p < 0.01$ ，*** $p < 0.001$ ；OR 值表示自变量对因变量的边际影响程度。

表11 Probit 回归参数估计表(二)

变量	回归系数	标准差	z 值	显著性级别
X ₁	0.112	1.06	0.77	
X ₂	0.394	2.24	2.60	***
X ₃	-0.075	0.87	-0.66	
Z ₁	-0.650	0.30	-3.21	***
Z ₂	-0.073	0.82	-0.37	
Prob>chi2			0.04	
Observations			143	
F test			0.00388	

注：* $p < 0.05$ ，** $p < 0.01$ ，*** $p < 0.001$ 。

由表 10、11 可知,在变换了回归方法之后所得到的实证结果并没有发生太大的变化,首先是两次实证模型整体的 P 值分别为 0.05 和 0.04,没有超过一般规定的数值,说明两个模型本身的统计意义较强,能够反映数据变化的一般规律。其次,在有序 Probit 回归下,两次实证回归关键变量的显著性等级与原回归结果相同,说明大学生对痛点信息平台的了解程度、社会痛点信息的公示度和性别这三个变量依旧对大学生的创业预期有着显著影响。最后,根据上面的分析我们可知,变换了回归方法后对最终的实证结果并没有产生较大的影响,整体的显著性没有大的变化,表明本文的实证结果相对稳健。

四、结论与启示

本文的实证分析主要分为两大部分,一部分是以二级指标为自变量的有序 Logistic 回归,构建此实证的目的是利用二级指标的微观性来寻找与创业意愿联系紧密的微观行为因素,进而可以为政府的微观调控行为提供理论和实证依据。另一部分实证是三大关键变量对创业预期的有序 Logistic 回归,由于三大关键变量是由二级指标通过一定的权重量化而来,因此具有一定的宏观性,其回归结果对宏观创业扶持政策的制定具有重要的参考意义。

基于对实证结果的分析,本研究在微观方面的结论有:①在痛点信息约束存在的情况下,提高大学生在日常生活中关注社会痛点信息的频率能够有效增强其创业预期与热情;②大学生对社会痛点信息库的了解程度与其创业预期呈现较强的正相关性;③女大学生群体的创业预期普遍较低,存在较大的发掘空间。本研究在宏观方面的结论有:①高水平的社会痛点信息整合度对大学生创业预期的影响不显著,而在低水平下,痛点整合度评估对大学生的影响却呈现出较强的显著性,信息整合度对创业预期的影响可能呈现一个倒 U 型的效用曲线;②痛点信息公示度在所有水平下都表现出对创业预期的高显著性和正相关性;③与以往研究所得到的结论不同,个

体的户籍性质与其创业预期的相关关系在逐渐减弱,户籍性质对于相关群体的创业意愿在一定程度上可能已经失去了代表性。以上结论的政策启示和理论意义是明显的。

其一,在宏观的创业政策制定上,政府可以围绕着痛点信息的整合度和公示度来展开,进而对大学生创业者进行“信息帮扶”。比如,可以利用电视广播、报纸杂志、网络媒体等平台进行社会需求信息的公示,为广大的创业者以及潜在创业者提供创业方向。从实证结果来看,痛点信息公示度对创业预期的影响效果显著。基于此,笔者认为,政策制定可以从创业前期创业者对创业方向和市场需求的寻找入手,缓解创业者的痛点信息约束,让有能力解决问题的人了解更多的痛点问题,从而提高创新创业的概率和质量。

其二,在微观调控上,不同层级的政府可以构建地方发展的痛点信息库,并扩大痛点信息库在各大高校的宣传力度与浏览广度。在提高大学生对社会痛点信息库的了解度和使用频率的同时,也会在微观个体层面对创新创业预期产生积极作用。调研数据表明,有 65.03% 的大学生认为构建痛点信息库有一定可能会激发创业灵感的产生,仅有 23.77% 的人觉得可能性不大甚至没有可能。从初步的调查数据来看,大部分大学生认为社会痛点信息是激发创业灵感的影响因素之一,政府可以通过构建社会痛点信息库来缓解大学生乃至创业者的痛点信息约束问题。

其三,积极挖掘女性创业者的潜力,大力推广性别平等教育,普及正确的性别观念,降低不正确的性别观念对女性创新创业预期的影响。积极宣传女性创业者的成功经验,在让社会和家庭平等对待女性创业者的同时,也能提高女性创业者的创业信心和创业热情。本文实证结果表明,相对于比较冒险的创新创业,女性大学生可能会选择更加稳定的就业方向,这也体现在近几年来教师、护士、银行从业者等“岗位稳定”的职业群体女性化趋势明显的现象中。

最后值得一提的是,在以往的一些对创业意愿的研究中,个体的户籍性质常常会是影响创新

创业意愿的主要因素。但本文的实证结果表明, 个体的户籍性质与其创业意愿的相关性已经变得不再显著。在对比当下政府所实施的国家政策之后, 我们不难发现这个异象背后的合理性: 第一, 近几年国家逐渐落实深化户籍改革政策, 有序地将农村人口转化为城市人口, 使得部分基础设施较为落后的农村地区人口的户籍转化为城市户籍, 这在一定程度上直接减弱了户籍性质与个体成长环境、融资环境和创业环境的相关性, 进而降低了户籍性质对创业意愿的解释力。第二, 在全面提倡创新创业的背景下, 国家创业政策的扶持对象和扶持地区正在逐渐放宽, 而这使得不同地区的创业成本慢慢地趋于同一水平, 进而减弱了户籍所在地的差异对个体创业意愿的差异性影响, 从而直接导致了户籍差异与个体创业意愿相关性的弱化。因此, 笔者认为, 在往后的相关研究中, 不能再以户籍差异来直接表示个体外部环境的差异, 而应该通过构建个体所在区域的经济指标、政治指标、文化指标等来综合考虑个体的环境变量, 进而再借助新的环境变量与研究变量来构建逻辑链条, 最后再依据此逻辑链条来设计实证研究, 从而取得更加符合实际情况的研究结论和新的学术突破。

参考文献:

- [1] 叶文平, 李新春, 陈强远. 流动人口对城市创业活跃度的影响: 机制与证据[J]. 经济研究, 2018, 53(6): 157-170.
- [2] 崔宏桥, 吴焕文. 创业环境如何影响科技人员创业活跃度——基于中国 27 个省市的 fsQCA 分析[J]. 科技进步与对策, 2021, 38(13): 126-134.
- [3] 宋思成, 张继伟. 疫情下就业政策对高校毕业生就业预期的影响[J]. 合作经济与科技, 2021(11): 93-95.
- [4] 肖广岭. 以颠覆性技术和“卡脖子”技术驱动创新发展[J]. 人民论坛·学术前沿, 2019(13): 55-61.
- [5] 蔺楠, 张茜, 周炜. 独生子女与创业意愿——基于“差别养育”视角的研究[J]. 经济管理, 2020, 42(6): 116-134.
- [6] 张慧, 周小虎, 宋迪. 从 0 到 1: 刻板印象威胁与女性创业者机会评估[J]. 经济管理, 2020, 42(4): 72-90.
- [7] 蔡静俏. 青年学生自主创业意愿影响因素研究——以广东湛江为例[J]. 现代商贸工业, 2021, 42(24): 62-64.
- [8] 刘敏佳. 经管类专业大学生创新创业意愿影响因素及对策研究[J]. 豫章师范学院学报, 2021, 36(3): 86-90.
- [9] ANWAR R P, CHAN C. Contrasting return migrant entrepreneurship experiences in Javanese Villages[J]. International Migration, 2016, 54(4): 150-163.
- [10] 曾婧婧, 温永林. 政府创业政策对城市创业的影响及其作用机制——基于国家创业型城市的准自然实验[J]. 经济管理, 2021, 43(4): 55-70.
- [11] 匡远凤. 人力资本、乡村要素流动与农民工返乡创业意愿——基于熊彼特创新视角的研究[J]. 经济管理, 2018, 40(1): 38-55.
- [12] 吴凌菲. 基于感知创业价值的大学生创业意愿形成模型[J]. 武汉理工大学学报, 2010, 32(1): 200-204.
- [13] 张娜依. 我国大学生创业影响因素分析[J]. 科技视界, 2014(8): 27-28.
- [14] 王阳, 贾晋. 智能手机与农户创业决策——基于中国农村家庭数据的实证[J]. 软科学, 2021, 35(10): 138-144.
- [15] 徐虹, 张妍, 翟燕霞. 社会创业研究回顾与展望[J]. 经济管理, 2020, 42(11): 193-208.
- [16] 王晶晶, 王颖. 国外社会创业研究文献回顾与展望[J]. 管理学报, 2015, 12(1): 148-155.
- [17] 张玉利, 杨俊. 企业家创业行为的实证研究[J]. 经济管理, 2003(20): 19-26.
- [18] 周洋, 华语音. 互联网与农村家庭创业——基于 CFPS 数据的实证分析[J]. 农业技术经济, 2017(5): 111-119.
- [19] 范炜昊, 徐健. 基于网络用户评论情感计算的用户痛点分析——以手机评论为例[J]. 情报理论与实践, 2018, 41(1): 94-99.
- [20] 钟罗金, 汝涛涛. 自我欺骗的认知加工: 模型及其构建[J]. 华南师范大学学报(社会科学版), 2021(3): 81-89, 206.
- [21] 何兴旺, 孙礼辉, 汪圣佑. 基于 Logistic 回归模型的大学生创业能力影响因素分析[J]. 红河学院学报, 2021, 19(2): 131-133.
- [22] 王乙杰, 杨大利. 性别角色观念对创业行为性别差异的影响[J]. 人口与经济, 2021(4): 126-142.
- [23] 徐永兵, 张橙. 金融支持对大学生创业意愿的影响研

- 究——以江苏省为例[J]. 中国商论, 2021(14): 173-175.
- [24] 田楠. 广州地区大学生创业意愿影响因素分析——基于 964 份调查问卷[J]. 韶关学院学报, 2020, 41(8): 67-73.

The influence of social pain point information constraints on college students' entrepreneurial expectations: The time cost and orderly Logistic regression of dynamic entrepreneurship

DAI Er

(School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China)

Abstract: Based on the hypothesis of “rational economic man” and guided by the question of how college students choose the direction of innovation and entrepreneurship, this paper proposes the following basic assumptions: By alleviating the information constraints of social pain points, the time cost of starting a business can be reduced and the overall entrepreneurial intention can be improved, and the dynamic time cost generated by starting a business is innovatively introduced into the cost analysis framework of entrepreneurial intention, which makes up for the deficiency of the existing literature limited to static analysis. Then, empirical test is carried out by collecting data and combining R-type clustering and ordered Logistic model. The empirical results verify the theoretical hypothesis proposed in this study. Therefore, this paper proposes that entrepreneurship support policies in the macro field should not be limited to “capital support”, it should include “information support”. In the micro field, the local information database of social pain points can be established appropriately, so that college student entrepreneurs can understand more social needs, their cost of information collection would be reduced, thus ultimately the development of innovation and entrepreneurship would be driven forward.

Key Words: R-type clustering; innovation and entrepreneurship; Logistic model; pain points information

[编辑: 苏慧]