

DOI: 10.16538/j.cnki.fem.20220516.106

三大创业数据库与创业研究

叶文平¹, 潘诗敏¹, 杨俊²

(1. 暨南大学管理学院, 广东广州 510632; 2. 浙江大学管理学院, 浙江杭州 310058)

摘要: 全球创业观察、创业动态跟踪调查、中国创业动态跟踪调查作为创业研究的三大权威数据库, 极大地推动了创业研究的情境化和独立发展。本研究采用文献计量方法, 对基于三大权威数据库并发表在国内外顶级期刊的219篇创业研究文献的主题和内容进行系统的梳理, 比较分析了三大数据库的设计思路、数据结构及其不足之处。本文从学科基础数据构建的视角探讨了创业研究发展的历史脉络, 揭示了基础数据构建对推动学科发展的重要价值, 也可以为中国本土创业研究构建大型基础数据提供参考和指引。

关键词: 创业研究; 文献计量分析; 研究主题; 数据库

中图分类号: F270 **文献标识码:** A **文章编号:** 1001-4950(2022)08-0003-21

一、引言

自20世纪80年代以来, 以 *Journal of Business Venturing* (1985)、*Entrepreneurship Theory and Practice* (1975/1988)、*Strategic Entrepreneurship Journal* (2007) 3本创业研究学术期刊的创办为标志, 创业研究在全球范围内有了长足的发展。然而, 创业研究作为一个独立研究领域所面临的“合法性”挑战并未随之消除。事实上, 当一个研究领域在发展繁荣的表象下并未寻找到支撑其存在的独特学术贡献时, 往往潜藏着更大的发展危机。正如Shane和Venkataraman (2000)所指出的, 一个独立的研究领域必须具有区别于其他领域的概念框架来解释和预测一系列经验现象, 否则就难以确认相关学者的独特贡献, 影响许多重要现象的理解, 甚至使该领域的关键问题浮于表面, 仅能作为一种研究背景或教学应用被排斥在学术研究之外。特别是, 现有的创业研究已突破特定边界, 实现跨越组织、行业、国家层面的拓展, 在全球经济形势下诱发了许多创业新问题与新趋势, 因此仅凭单向接受其他领域的理论, 无法解释丰富复杂的创业现象和问题, 为创业研究的拓展带来了挑战(周冬梅等, 2020; Arend, 2020)。拓展创业研究的情境化成为应对上述挑战的关键。Weick (1995)认为, 对于一个新兴的研究领域而言, 构建本领域研究框

收稿日期: 2022-01-13

基金项目: 国家自然科学基金重点项目(71732004); 国家自然科学基金青年项目(71902072); 国家自然科学基金面上项目(71872074); 广东省自然科学基金面上项目(2019A1515011921)

作者简介: 叶文平(1989—), 男, 暨南大学管理学院副教授;

潘诗敏(1998—), 女, 暨南大学管理学院硕士研究生;

杨俊(1980—), 男, 浙江大学管理学院教授(通讯作者, nkyangjun@163.com)。

架和解决本领域研究问题的基本前提是对本领域关注的研究问题的特征和属性要有清晰的描绘。因此,如何有效地探寻创业研究的情境化成为研究者关注的焦点。

创业学者从不同主题和理论视角聚焦创业研究情境化,绝大多数研究以定量研究为主导,依靠调查、档案数据和复杂的统计技术进行理论检验(Wei等,2019),专门的数据库为创业研究提供了必要基础,因此构建创业研究的高质量数据库成为深化创业研究情境化的必由之路。鉴于高质量数据库在创业研究中的重要性日趋凸显,创业研究学者积极参与了构建本领域的独特数据库的工作。迄今为止,创业领域已有三个大型数据库被研究者广泛采用:全球创业观察(Global Entrepreneurship Monitor, GEM)、创业动态跟踪研究(Panel Study of Entrepreneurial Dynamics, PSED)、中国创业动态跟踪研究(Chinese Panel Study of Entrepreneurial Dynamics, CPSED)。这三个大型创业数据库的实施推广,促进了创业研究的理论和方法探索,使得创业研究不只是聚焦在创业者特质上,而是逐渐转移到创业过程本身,与此同时,对情境化的重视也显著推进了创业理论和创业研究的快速发展。

尽管基于GEM、PSED和CPSED三大数据库的创业研究取得了一定的成果,但缺乏系统梳理降低了这3个代表性数据库在推动创业研究发展中的重要价值。鉴于此,本文分析了三大创业数据库的研究成果,识别了现有研究不足,进而针对研究现状和趋势为未来优化和构建更完善的创业数据库提供依据和指导。本研究可能的贡献在于:(1)揭示了创业数据库的构建对创业学科发展的深远影响。本研究基于三大创业数据库研究成果的系统梳理,分析其对创业研究发展的重要价值和推动作用,为进一步深化创业研究的情境化提供了可操作的途径。(2)探讨了创业学科在发展过程中对研究数据的创新要求。当研究发展达到一定程度后,更加依赖多维度、多层次的数据支持,需要寻求数据背后更复杂的关系和价值,因此数据库的收集和分析方法需要不断更新,本研究通过对三大数据库主题框架的考察,揭示了创业研究发展对高质量数据库构建的潜在需求。(3)总结了已有数据库设计中存在的不足,为后续构建大型基础数据库提供重要建议。本研究根据文献分析结果提出一个整合框架,为构建新的创业数据库提供有益的参考和借鉴。

正文内容安排如下:首先,详细介绍GEM、PSED、CPSED三个数据库的设计框架与研究方法。其次,选择在FT50期刊、中国国家自然科学基金委管理科学部认定的30种学术期刊以及*Small Business Economics*期刊上发表的219篇相关研究文献进行编码,系统梳理并总结文献的研究主题以及相关数据库的不足之处,构建本文的整合框架。最后,提出创业数据库的改进建议,为创业研究大型基础数据库的构建提供参考和指引,以期为进一步推动创业研究的情境化提供数据支撑。

二、期刊文献计量分析与数据库介绍

本文旨在梳理总结三大创业数据库的特点,并基于三大数据库的研究现状及不足之处,揭示数据库与创业研究发展的关联“黑箱”,推动数据库的完善工作,为未来创业相关研究提供参考和借鉴。本文采用文献计量学方法,选取FT50、中国国家自然科学基金委认定的30本期刊以及*Small Business Economics*为期刊来源,检索其中以“GEM”“PSED”“CPSED”为样本数据库的创业研究文献,并详细介绍三大数据库的设计框架。

(一)文献选择与编码

本文选取的期刊来源及原因如下:(1)FT50期刊,即英国《金融时报》评定的50本商学顶级期刊(Financial Times 50 Journals,简称FT50)。它以期刊影响因子作为入选的测量指标,具有稳定性。FT50期刊被用于评估商学院的研究能力,是《金融时报》对商学院排名的重要依据,在国

实际上具有很高的认可度和巨大的影响力。FT50期刊涵盖多个子领域,其中有9本期刊涉及创业领域。(2)中国国家自然科学基金委管理科学部认定的30种学术期刊。根据项目后评估工作的需要,管理科学部委托中国科学院文献情报中心以科学方法遴选符合管理科学部项目资助范围的我国管理科学重要期刊。经过4个阶段的专家调查和文献计量学统计分析,在充分尊重专家意见的基础上,本着总量控制、质量第一、增量从严的原则,最终遴选出重要期刊(A)22种和重要期刊(B)8种。(3) *Small Business Economics* (SBE)。首先,该刊是美国中小企业协会的会刊,对创业研究发展具有重要指导意义,其影响力仅次于FT50期刊中的ETP、JBV和SEJ。其次,SBE是发表采用GEM与PSED数据库文章最多的期刊,且被引用量可观,迄今已发表的基于GEM数据库的研究成果超过53%,显示GEM的权威性。最后,SBE的很多主编是推动GEM项目的重要参与者,如创始编辑Zoltan J. Acs自2001年就是GEM调研项目的重要团队成员,为推动项目发展与学术研究做出了重要贡献。因此本文将该刊纳入分析视野。

在确定检索期刊标准之后,本文对目标文献的筛选过程分成以下步骤:第一,分别以“Global Entrepreneurship Monitor / GEM / 全球创业观察”“Panel Study of Entrepreneurial Dynamics / PSED / 创业动态跟踪研究”“Chinese Panel Study of Entrepreneurial Dynamics / CPSED / 中国创业动态跟踪研究”为关键词全文检索EBSCO、JSTOR、Wiley、知网等数据库中的目标期刊,共下载论文221篇;第二,根据重要文献的作者所发表文献筛选补充遗漏的研究文献,共补充5篇论文;第三,针对现有文献中所引用的相关文献进行检索拾遗,共补充3篇论文;第四,阅读已选取文献全文进行筛查,人工判读后剔除非使用指定数据库研究的论文10篇。最后,讨论并确认219篇论文作为研究分析对象,其中,基于GEM数据库的140篇,基于PSED数据库的63篇,基于CPSED数据库的16篇。

本文采用文献计量学方法对219篇文献进行编码以便后续的统计分析,编码内容主要包括发表年份、发表期刊以及研究主题,所含信息有助于深入系统地研究三大数据库的使用情况,分析其对创业领域研究现状与发展趋势的贡献。依据编码内容差异,本文主要应用两种编码方式:第一,直接编码。期刊与发表年份信息直接从文章中获取进行编码;第二,提炼编码。研究主题无法直接获取,因此结合文章关键词、摘要等内容提取,梳理分析后将细分主题进行归类。为提高编码的准确性,本文作者多次阅读、检查目标文献,对比两次编码结果,针对分歧之处再次阅读并讨论原文,直至意见一致,最终形成本文研究的编码数据库。

(二) 创业三大数据库基本情况介绍

1. 全球创业观察

全球创业观察(GEM)项目始于1999年,是美国百森学院(Babson College in Wellesley)与英国伦敦商学院(London Business School)的联合研究项目。GEM侧重于三个主要目标:(1)揭示创业活动水平在不同国家(地区)间的差异;(2)揭示决定创业活动水平的因素;(3)确定可能提高创业活动水平的政策。项目为创业驱动经济增长提供必要基础知识,弥补了关于创业的国际数据空白。

GEM数据逐年收集并随机抽选样本形成独特截面数据,收集主要包含四项活动:(1)成人人口调查(APS),APS探讨个人在创业过程中的作用,提供个体创业偏好、动机、能力、态度和活动的标准化数据。为提供宝贵纵向视角,APS的核心多年来始终保持一致,并于每年更改“特殊主题”部分,增加题项调查与国家(或地区)特别相关的问题,探讨尚未涵盖在核心APS中的创业方面的内容(见表1)。(2)国家专家调查(NES),NES确定了9个被认为对创业产生重大影响的因素(创业融资、政府政策、政府创业计划、创业教育、研发转让、商业和法律基础设施、内部市场动态和准入监管、有形基础设施以及文化和社会规范),即创业框架条件(EFCs)。(3)由

专家填写的自我管理问卷,其目的是建立标准化衡量指标以收集专家们对本国企业框架条件的看法。(4)基于现有跨国数据集汇编相关标准化措施,来自经合组织、世界银行、联合国、欧盟、劳工组织和全球竞争力报告等各种来源的标准化国家数据提供关于企业和一般国家框架条件的额外信息。

表1 GEM每年参与国家(地区)数量与特殊主题

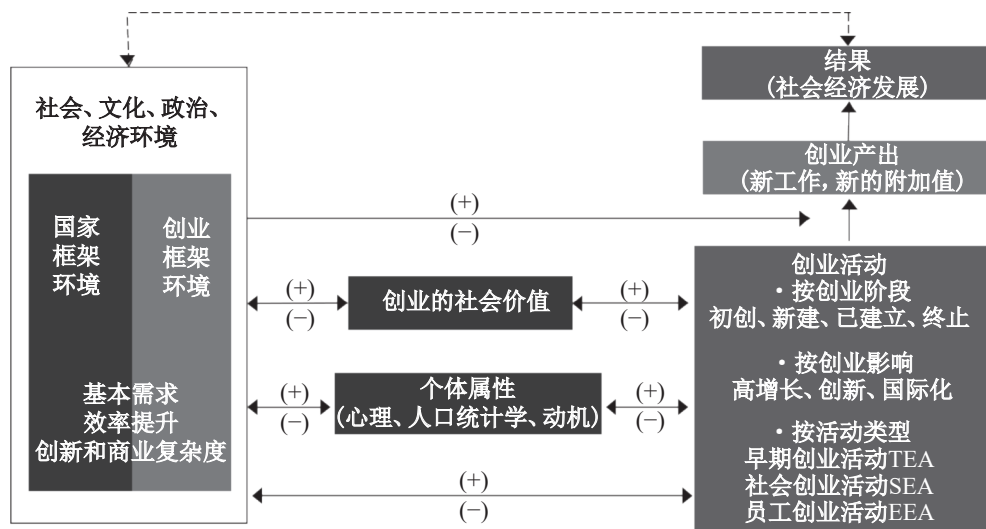
年份	参与国家(地区)数量	特殊主题
1999	10	-
2000	20	风险资本
2001	28	-
2002	37	女性创业,创业融资,家庭支持与创业
2003	31	高增长创业活动,政府政策、项目和创业
2004	33	-
2005	35	高期望创业,女性创业
2006	42	创业融资,女性创业
2007	42	女性创业,创新信心指数
2008	43	高增长创业,创业教育和培训,创新信心指数
2009	54	社会创业,经济衰退,女性创业,创新信心指数
2010	59	经济危机,女性创业
2011	59	员工创业活动(EEA),高成长创业
2012	69	移民创业,女性创业
2013	70	幸福感,青年创业
2014	73	青年创业
2015	60	创业、竞争力和发展
2016	65	创业、竞争力和发展
2017	54	创业和创新生态系统,社会创业
2018	49	家族创业,全球信息网络和共享经济
2019	50	青年创业
2020	46	新冠肺炎对创业的影响

资料来源:根据GEM报告整理。其中,地区主要是指GEM调查样本涵盖的中国香港、中国台湾地区。

GEM的理论分析框架如图1所示,从创业框架环境到国家层面的创业活动,最终到就业创造与经济增长,该模型描述了社会价值、个人属性与各种创业活动之间的假定关系。其中,社会、文化、政治和经济背景是指利用世界经济论坛的十二大支柱来分析经济发展阶段、调查竞争力和GEM国家创业状况的9个组成部分。创业的社会价值包括:社会如何把创业视为一种良好的职业选择,企业家是否具有较高的社会地位,以及媒体对创业的关注如何促进国家创业文化发展。个人属性包括人口因素(性别、年龄、地理位置)、心理因素(感知能力、感知机会、对失败的恐惧)和动机方面(生存型与机会型创业、基于改进的创业等)。创业活动根据创业生命周期阶段(初创、新建、已建立、终止)、活动影响(高增长、创新、国际化)和活动类型(早期创业活动TEA、社会创业活动SEA、员工创业活动EEA)确定。

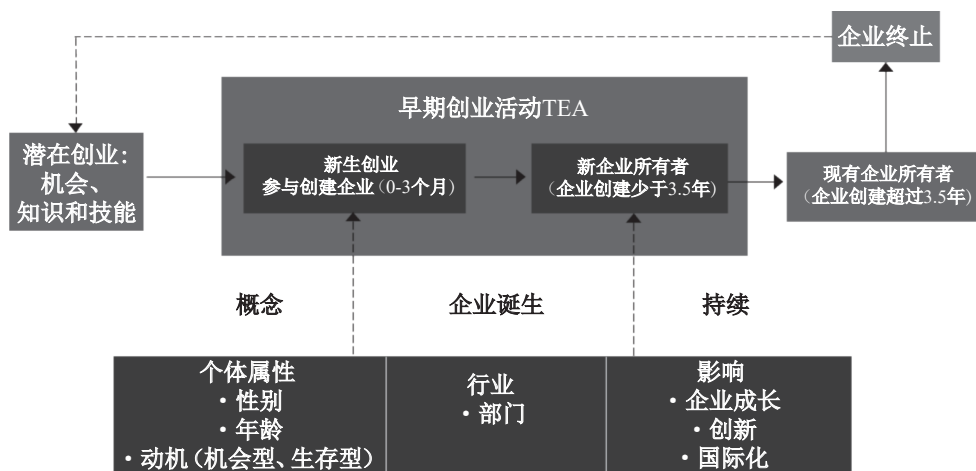
全球创业观察项目在三个方面具有独特性:第一,在全球范围内收集原始数据,每年参与调研的国家数量伴随项目影响力扩大而递增。GEM于1998年在6个参与国(加拿大、丹麦、芬兰、德国、英国和美国)进行预调研研究,并于每年发布一份全球创业报告呈现其研究成果,目前已累积来自全球115个国家(地区)的数据;第二,调查创业个体,充分了解与创业意愿、态度和活动相关的各种关键问题;第三,基于整个创业周期评估创业现象,始于创业机会,终于概念成熟或至其终止(见图2)。这些数据为更好地促进创业活动水平的跨国家比较、估计创业活动

在国家经济增长中的作用、确定造成国家创业活动水平差异的因素提供了重要支撑,对政策制定者也极具参考价值。



资料来源:根据现有文献与报告整理。

图1 GEM的理论框架



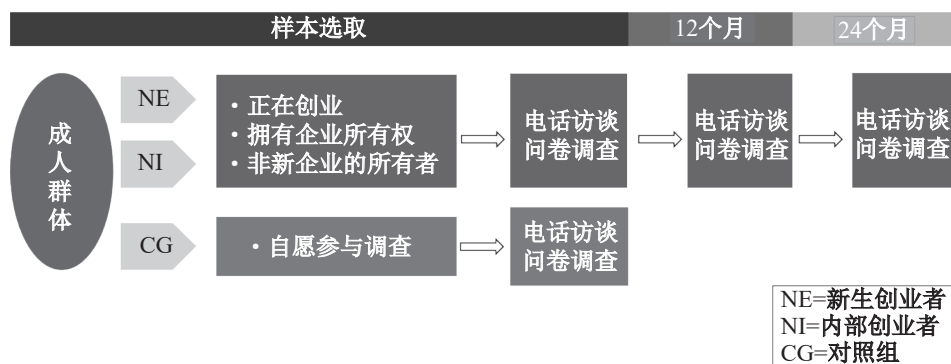
资料来源:根据现有文献与报告整理。

图2 GEM覆盖整个创业过程

2.创业动态跟踪研究

创业动态跟踪研究(PSED)项目的敲定过程经历了严谨的前期工作。以1993年为起点,威斯康星州住房和经济发展局(Wisconsin Housing and Economic Development Authority)对当地成年人进行初步调查,为创业过程提供了可靠的实证描述,建立了项目的基本设计——筛选具有代表性的成人样本,挖掘积极参与创业的个体获取创业详细信息,并跟踪采访对象以确定结果。在取得初步成功的基础上,密歇根大学调查研究中心于1996年对具有代表性的美国家庭进行了重复研究,证实了上述研究设计的可行性并促成了创业研究联盟的成立。直至1998年,由来自34个创业研究中心的120多名学者组成的PSED项目正式成立,旨在提高人们对创业方式的科学理解。

PSED以成人群体(18—64岁)为总体样本,分3个阶段收集数据(见图3)。第一阶段对家庭进行电话调查,以创建两个代表全国18岁及以上成年人人口的样本(新生创业者与对照组),如果受访者对“你是否正独自或与他人一起从事创业活动”以及“你是否正独自或与他人一起创建新企业或为你的雇主创建新企业”两个问题中的任意一个回答“是”,并且满足“受访者期望成为新企业所有者或部分所有者”“受访者已在12个月前就采取积极行动创办新企业”“新企业仍处于创建阶段”3个条件,则被认定为新生创业者。第二阶段中,对两个样本中的个人进行详细的电话采访并完成邮寄问卷。第三阶段是对新生创业者进行12个月和24个月的后继访谈(电话和邮件问卷)。

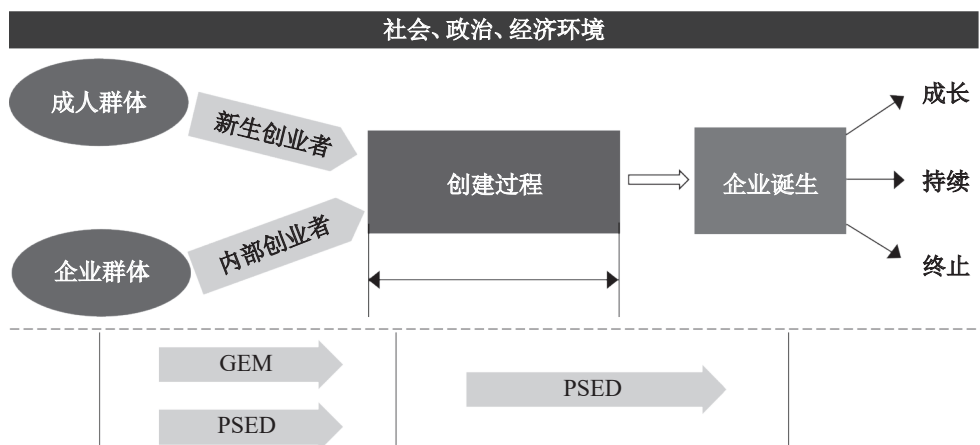


资料来源:根据现有文献与报告整理。

图3 PSED的设计过程

PSED项目分两期实行。PSED I于1998—2000年进行调查,通过3次后续访谈最终选择830名新生创业者提供有关其创业活动的信息,并将其与未参与公司创建的人员(对照组)进行比较。PSED II于2005—2006年进行调查,随后每年进行6次访谈,最终确认了1 214名新生创业者。实施PSED II有以下原因:首先,PSED I人群基于1998—2000年互联网热潮期间对公司创建高度关注时确定的,PSED II人群在2005—2006年这一典型时期确定。第二,鉴于PSED I以及GEM项目的大量经验,PSED II有了方法改进的基础。新研究方案提供了更可靠的措施并改进了关键议题。第三,探究不同种族群体在参与和经历等方面的差异。例如,非洲裔美国人进入新公司的机会更高,而在发展新公司方面的成功较少。这些差异得到了更多的关注。最后,影响结果的因素众多,在多变量分析中较大样本的初创企业将具有相当大的价值。因此PSED II将新生创业者样本增加了一倍多,从830人增至2 044人。

创业过程不断受到政治、社会和经济因素的影响。这在PSED理论框架中得到了充分体现(见图4)。模型左侧透露创业过程的第一阶段涉及个人参与者,其中一些人可能决定创业,这些人有两个潜在的来源:所有劳动者和现有企业的雇员。创业过程的第一个转折点是概念阶段,这两个来源的个人都可能选择创业。若创业目标是成立一家独立的创业公司,则参与其中的人被称为新生创业者(NE),若初创企业由现有企业发起,则参与其中的人被认为是内部创业者(NI),这两个群体统称为创业者。PSED的一个重要特点是在普通人口中确定一个不参与创业活动的成年人作为比较群体,从而将新生创业者样本与对照组的特征进行比较。PSED是创业领域中第一个大规模的动态跟踪数据,获得的信息包括新生创业者中活跃者的性质、创业过程中开展的活动以及成为新公司的初创企业特点的数据。模型下半部分显示,与GEM相比,PSED除了确认新生创业者,还重点关注创业行为与过程,便于研究者更好地了解企业创建过程。



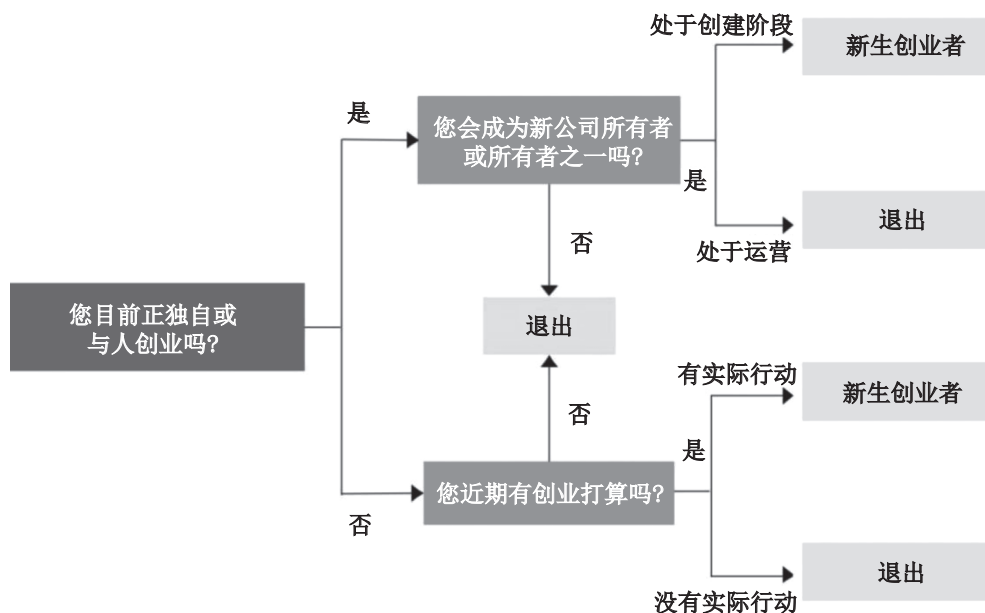
资料来源:根据现有文献与报告整理。

图4 PSED的理论框架

3.中国创业动态研究跟踪研究

中国创业动态跟踪研究(CPSED)于2009年由南开大学创业管理研究中心组织实施,是我国首个旨在揭示微观层面创业活动规律的调研项目。

CPSED的研究对象主要为18岁以上成人群体中的新生创业者,筛选过程见图5。CPSED结合分层抽样与随机跳号 household 电话抽样(RDD)方法,考虑各地区创业强度差异前提下,选择8个城市作为调查样本(分别是东部地区的北京、天津、杭州、广州,中部地区的武汉,东北地区的沈阳以及西部地区的成都、西安)。CPSED依据各被调查城市人口规模和18岁以上成人数量,确定每个被调查城市的调查样本权重,进而制定各被调查城市的样本量分配方案,随后在被调查城市以随机跳号 household 电话抽样方式同步接触被调查者,按照既定标准从中筛选新生创业者并进行电话访问调查。



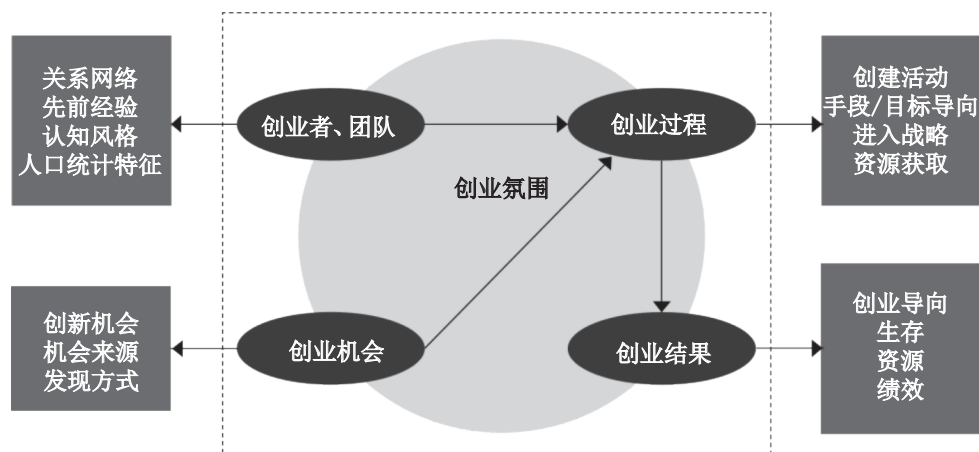
资料来源:根据现有文献与报告整理。

图5 CPSED新生创业者的筛选过程

该项目经历三轮数据收集过程,在2009—2011年内采用3次电话跟踪访问的手段追踪新生

创业者的创业活动进展状况。2009年首轮调查总共电话接触69 990户家庭,电话访问20 424人,从中识别出974名新生创业者,最终有601名新生创业者完成访谈。2010年对601个成功样本展开了第二轮电话跟踪访谈,最终有321位新生创业者完成。第三轮调查于2011年展开,针对第二轮的321个样本进行跟踪调查,并且在前两轮的基础上进行了变量拓展,增加了对创业退出者关于创业过程因素的询问,最终共有121位新生创业者完成电话访谈。

CPSED通过科学的研究设计对新生创业者进行动态跟踪,内容涵盖创业者、创业机会、创业环境、创业过程与创业结果等关键要素,设置了新生创业者的人口统计特征、先前经验、社会网络、机会发现方式、创业氛围等研究变量(见图6),力图从微观视角剖析我国特定情境下的创业活动路径与内在发展规律。



资料来源:根据现有文献与报告整理。

图6 CPSED的理论框架

三、基于三大创业数据库的相关研究主题分析

鉴于创业领域的日趋进步,需要大量研究数据支撑以规划研究过程、探索有趣和重要的问题。三大数据库以其独特优势,为创业研究的演进提供了重要指导。为了使创业研究发展与数据库建设相得益彰,本文简要介绍三大创业数据库在创业研究相关主题领域的应用现状,旨在为解决以下问题提供参考:第一,明确目前创业领域已经使用该数据库做了什么研究,还可以沿着什么方向前进;第二,展示创业研究者是如何通过重要数据库做研究的,现有的数据库存在哪些不足之处。

为识别三大数据库在推动创业研究发展中发挥的基础作用,以进一步说明三大数据库的重要贡献,本文通过人工编码提炼出三大数据库的研究主题,将其与顶级期刊发表的创业研究成果及其主题演变趋势比较^①。具体工作步骤及研究发现如下:第一,确认基于三大数据库的创业研究的主题分布。通过人工编码对三大数据库文献的研究主题进行划分,发现其研究主题主要涵盖创业者与团队、创业环境、创业机会、创业影响、创业决策等9个方向(参见表2)。第二,检索2000—2022年发表在“UTD24”期刊以及JBV、ETP、SEJ三个创业顶级期刊上的创业文献,确认这27个顶刊中的创业研究主题分布。选取这些期刊的原因在于,它们作为管理学与创业领域的顶级期刊具有较高的影响力和被引量,很大程度上引领了创业研究的发展,将其作为检索范围有较高的说服力。通过人工判读最终筛选630篇文献进行主题编码。第三,对比分析以上两者在创业研究主题分布及发展特点,以期揭示三大数据库对创业研究的重要推动力(如图7所

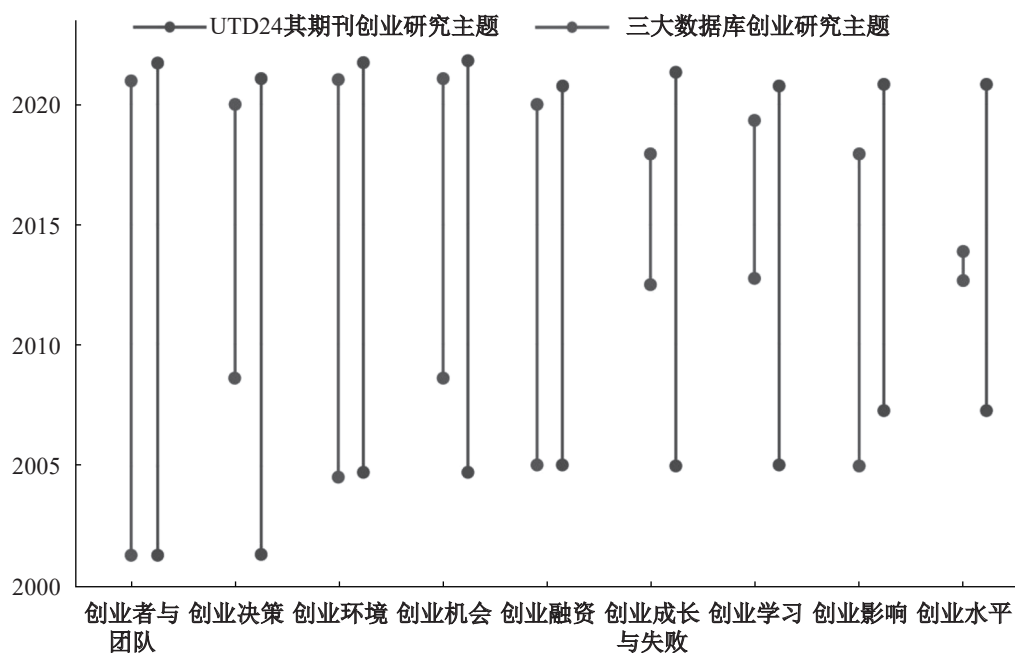
^①这一研究方法主要受到审稿人的启发,在此表示感谢。

示)。通过与顶刊论文发表的创业研究发展趋势比较,本文发现基于三大数据库的创业研究与已有的创业研究文献在趋势上高度契合。其中,两者在创业者与团队、创业环境、创业机会和创业融资这4个主题的时间分布上高度匹配,关于创业影响的研究在基于三大数据库的创业研究中具有先发优势。对比分析数据充分说明了三大数据库在推动创业研究发展中的重要作用。本文针对相关主题展开分析评述,以期对创业数据库的构建工作提供参考。

表 2 主要研究主题分布

序号	主题方向	文献占比(%)	序号	主题方向	文献占比(%)
1	创业者与团队	38.38	6	创业融资	2.95
2	创业环境	29.52	7	创业成长与失败	2.58
3	创业机会	9.23	8	创新水平	1.85
4	创业影响	6.27	9	创业学习	1.48
5	创业决策	7.75			

资料来源:根据编码文献整理。



资料来源:根据编码文献整理。

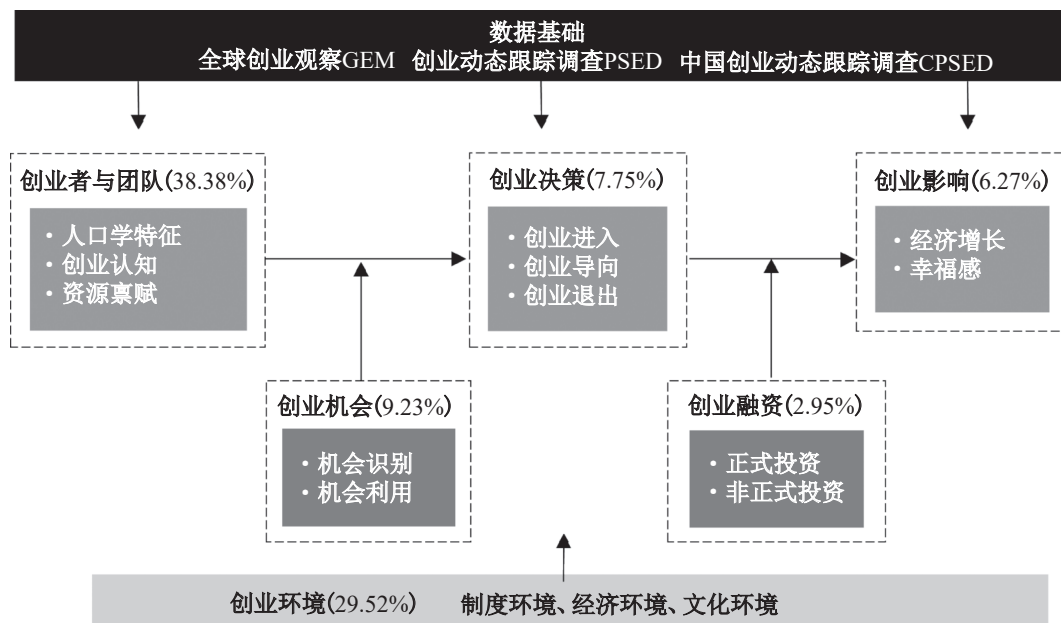
图 7 三大数据库与UTD24期刊创业研究主题分布对比

针对编码文献的统计发现,在2001—2021年基于三大数据库的创业研究中涉及创业者与团队的讨论最为活跃。随着20世纪90年代关于特质论的不断挖掘,过程视角的创业行为研究逐渐兴起并成为主流,相关研究聚焦于创业机会、创业行为、创业导向等问题,并开始关注创业环境的作用。图8展示了主要研究主题的框架。

(一)创业者与创业团队

创业者特质研究始于20世纪初期,这类研究结合心理学、社会学、管理学的方法来回答谁会成为创业者的关键问题。创业者及团队是创业活动的根本推动力,因此三大数据库为创业者特征提供的独立板块数据促进了研究者识别创业主体的独特特征。创业者是拥有创业意愿并采取行动的个人,创业团队由这些从事创业活动具有共同目标的个体组成,现有的针对创业者

与团队的研究主要涉及人口学特征、创业认知和资源禀赋三个方面。



资料来源：根据编码文献整理。

图8 主要研究主题框架

第一,针对人口学特征的研究多考虑与个人决定成为创业者相关的变量,如年龄、性别、教育程度等。尤其是,创业者性别研究已引起学者强烈兴趣,大量研究基于性别角色一致性理论提出,职场的性别不平等现象驱使女性离开薪酬性就业岗位去创办自己的企业,但由于存在明显的性别刻板印象,男性被认为更有能力,因此女性在创业生涯中面临更大的发展挑战(Minniti和Nardone,2007;Lim和Suh,2019)。也有研究从资源基础观对比了男性与女性在早期创业中面临的资源约束模式差异,解答女性早期创业的关键制约问题,以促进未来创业中的性别多元化(王汉瑛等,2021)。

第二,随着创业者特质受到更多的关注,认知视角的创业研究成为重要主题之一。创业者自我效能感(朱亚丽和郭长伟,2020;Vilanova和Vitanova,2020;张秀娥等,2021)、创业动机和感知(Chen等,2018)等因素对创业者的影响得到了更深入的研究。研究发现,认知因素能够影响创业者构建认知因果图,指导其资源配置决策。创业者个体的感知不确定性和认知倾向会影响创业判断,最终导致创业行为差异(Uygun和Kim,2016)。

第三,除了人口学特征和创业认知,包括人力资本、金融资本与社会资本在内的资源禀赋也会对创业活动造成深远影响。相关研究认为,在环境中易于转移的技能和知识存量、更容易获取资源的网络关系以及财富积累对创建新企业具有积极的作用(Semrau和Hopp,2016;Boudreaux和Nikolaev,2019)。

总体而言,由于三大数据库对创业者及团队特征探讨的支持,基于人口学特征、认知与资源禀赋的创业研究得到了更加有力的支撑,相关研究更加具体深入。然而,相关主题的研究也显示了现有数据库中样本数据存在的局限性。例如,由于人格特质具有较强的迁移性和适应性特征,这对研究观察创业者特质的长期数据提出了更高的要求(Lim和Suh,2019)。现有研究采用GEM数据集中的各种人力、财务和社会资本衡量指标,但这些变量也面临着测量精细度的挑战(Boudreaux和Nikolaev,2019)。之前的研究表明,不同类型的社会资本可能会以异质的方

式影响创业活动,在某些情况下强关系可能不如弱关系有价值,然而由于数据限制,无法区分强弱关系;只知道创业者是否与其他创业者有联系,却不能区分关系的性质。因此,未来的数据库需要进一步区分特定变量如何影响创业者及团队,以便完善现有的研究结论。

(二)创业机会

自shane和Venkataraman(2000)开创性地将创业定义为“如何、由谁以及在何种影响下发现、评估和利用创造未来产品和服务的机会”的研究,引发对创业行为与过程研究的持续深耕,研究由创业者特质论视角转向考察对机会的识别与作用机制。

三大创业数据库为机会识别与利用提供了独特的数据优势。GEM从2001年起调查受访者所感知的创业机会,创业研究开始识别机会的感知与行为模式,由此展开考察新机会和现有机会的识别与利用过程。研究主要体现在两方面:第一,微观层面,研究集中在个人先验知识与社会网络作用上,提出个体对机会的感知因其所处网络的不同以及人力资本而存在差异(Arenius和De Clercq,2005)。第二,宏观层面,研究强调外部资源对创业机会开发利用的支持是成就创业活动的重要因素,创业者在与外部环境的互动中充分利用外部资源并识别开发创业机会,当感知到的创业机会在经济上是可持续的,创业者将获取资源投入生产来采取创业行动。研究主要关注经济与文化支持、有利的商业环境条件、社会意识形态等,它们明显增强了个人利用创新型或模仿型机会的积极关系(Stuetzer等,2014;Young等,2018;Guerrero和Peña-Legazkue,2019;Bennett等,2022),为创业机会的重要性提供了有力依据。

具体而言,创业者如何辨别和利用有吸引力的市场机会是衡量创业活动的重要前因。在确定的机会面前,有些人会采取行动成为创业者的一员,有些则不会。如果缺乏创业者的独特见解、观念与解释,那么机会不可能实现。如果没有组织行动就不会转化为财富。这表明有动机的个人与机会或组织模式构成了创业研究的中心,促进一系列产生共识与合法性的创业研究。同样值得注意的是,由于创业机会研究强调过程观视角,以及反映和揭示创业行为规律的概念,而机会的发现、识别、开发、利用过程是一个动态、复杂的有机整体,因此对三大数据库收集跨越更长时间维度的变量提出了新要求。研究者指出,具有更强解释力的研究结论依赖于更加长期有效的创业机会纵向跟踪数据(Stuetzer等,2014)。

(三)创业融资

创业融资是新创企业创建与发展绩效的重要资源。大量研究表明,融资对创业活动的开展具有重要作用。由于缺乏可用的跨国数据,部分因素一直受到忽略,因此为评估这一观点的有效性并对投资在创业过程中的作用提供新见解,三大数据库提供了关于创业融资作用和结构的重要数据。GEM等在2000年增加了融资这一话题,收集尽可能多的国家风险资本投资的总体数据,以探讨早期融资与创业活动的关系。独特的国际数据比较增加了对风险资本投资的走向以及投资在促进创业方面扮演的角色的洞察,极大地促进了创业融资研究在其后十年中的长足发展。

创业资本不仅影响新创企业进入市场的能力,也决定其进入市场后的绩效,正式投资与非正式投资是初创企业拓展新业务的两种重要途径。初始资本资源的充足性和规模能够促进创业者追求产品研发,推动规模扩张的实施(黄永春等,2020),提高新创企业的生存能力和增长能力(Ho和Wong,2007),但由于许多创业企业在启动阶段缺乏足够的商业经验与信用记录,从外部正式资源提供者如通过银行贷款筹集资金并不容易,尤其对于中小型企业来说极其困难,创业者如何筹集初始资金是初创企业生存的重要因素之一。一般而言,正式投资只流向高增长潜力的公司,极少数有前途的公司可以从风险资本公司获得正式投资的资金,而非正式投资则流向包括高增长潜力在内的各个领域的公司,大多数新公司最初的资金支持来源于家庭、

朋友、商业伙伴和其他个人关系的非正式投资(Ho和Wong,2007),因此非正式投资发挥了重要作用,进一步的研究也证实了非正式投资对机会驱动型创业以及具有高增长潜力的创业活动具有显著的积极影响(Maula等,2005;De Clercq等,2014)。

总体来看,三大数据库的投资数据使创业融资研究得到基本保证,但研究对象的选择大多基于研究项目可用性而非随机选择程序,这是制约覆盖更多国家与地区样本以及与时间效应有关的重要因素。例如,芬兰作为高税收国家,创业文化相对薄弱,这可能影响其非正规投资活动(Maula等,2005),忽略相关因素可能造成研究结论的国家偏见。因此,未来的数据收集需要考虑不同国家的发展水平等变量如何帮助释放一个国家的总体财政资源,而不是局限于现有的创业资源,以便更全面地预测创业融资活动(De Clercq等,2014)。

(四)创业决策

创业决策涉及创业进入、创业导向、创业退出等决策行为的研究,三大创业数据库项目包含的丰富测量指标为创业决策研究提供了良好的数据资源。相关研究可分为以下三类。第一类,创业进入。创业进入决策要求创业者决定是否、何时、以何种形式创建新企业,决策往往受到人力资本(Estrin等,2016)、社会认知与环境(Autio等,2013;Boudreaux等,2019)、社会信任多样性(宋渊洋和赵嘉欣,2020)等因素的影响。第二类,创业导向。在创业进入之后,创业者会采取不同的成长导向决策(Autio等,2013)。在早期研究中,张玉利等(2010)基于CPSSED数据发现,机会识别与成长导向具有相关性,且高成长导向的创业者往往进行更多的计划准备和资源整合工作,从而验证了成长导向在新企业生成中发挥着重要作用。然而并非所有的创业者都会选择高成长导向,基于家庭嵌入视角的创业研究揭露了工作与家庭的潜在冲突(田莉和张玉利,2018),工作—家庭冲突明显的制度环境加剧了女性创业者选择低成长导向创业的可能性(Thébaud,2015)。此后有学者基于资源配置视角,将创业导向划分为规模扩张导向和产品研发导向,并发现金融支持、政府项目、税收优惠等创业资助对创业导向的影响机制(黄永春等,2020)。第三类,创业退出。由于新创企业的失败率很高,创业退出和创业重新进入成为潜在的研究主题。虽然创业退出通常具有负面含义,但是企业终止并不仅仅意味着失败。Yusuf(2012)指出,创业退出是脱离创业的过程,脱离并不一定是一个消极的结果,所有脱离的情况也并非都是同质的。创业者由于得出商业机会不会成功的判断而选择的主动退出或战略脱离,往往是一种积极的模式。此外,退出的创业者决定如何开始后续创业也是极为关键的选择,探索重返社会过程的不同动态可能为未来的研究提供重要的机会。初步的研究显示,当破产法的友好程度、社会规范的支持程度以及风险投资可得性得到改善时,失败的创业者进行连续创业的概率会显著提升(郑馨等,2019)。更好地了解创业者重返社会的前因、背景和结果,将有助于失败创业者更有效地调整其重新进入时的创业决策(Simmons等,2019)。

简而言之,创业决策是对既有方案的评估和选择过程,决策来源于个体的有限理性和认知水平。现有的创业决策研究更多的是探讨创业者在当前环境中决定是否进入创业或者如何实现企业成长,而针对选择退出或从创业失败中东山再起的创业决策数据则较少。因此,今后的创业数据库需要考虑收集再创业活动的相关影响指标(Guerrero和Peña-Legazkue,2019)。另外,现有的创业决策数据还面临测量细致性不足的问题如GEM数据库目前对“失败恐惧”和“风险性导向”的测量题项雷同是造成创业决策研究主题局限性的重要因素,需要在未来数据库中加以完善(Arenius和Minniti,2005;Pindado和Sánchez,2017)。

(五)创业影响

创业活动能够帮助实现企业振兴、产业升级并带动就业增长。基于三大数据库的创业影响研究主要集中在创业活动对促进经济增长和提升国民幸福感的影响上。

1.对经济增长的影响。自1999年以来,基于GEM的创业研究已证明创业活动能够促进经济增长。虽然这种影响关系是一致的,但是联系的强度往往会因各国创业活动的性质而有差异(Du和O'Connor,2018)。2001年的数据显示,创业活动与各国的经济增长显著相关,但是弄清导致这种关系的因果机制需要长期跟踪数据的支持。跟踪创业活动与经济增长之间的关系是GEM项目的重要目标之一。基于GEM项目的进一步研究发现,创业活动之所以能够促进经济发展,是因为新企业创造了新的就业机会,提升了市场参与者的多样性并加剧了竞争,使其通过技术变革提高生产率(Acs等,2018)。Wennekers等(2005)发现,处于不同发展阶段的国家,创业活动具有不同的经济作用。与发达经济体相比,发展中国家的创业活动对经济增长的影响更小。Wong等(2005)认为,不同类型的创业会对一国的创新能力和经济增长率产生不同的影响,具有高增长潜力的初创企业和机会型创业能增强知识溢出和经济增长。

2.对国民幸福感的影响。由于人们日益关注创业实践的社会贡献,研究创业对幸福感的影响随之受到重视。2013年以后,新纳入GEM的项目包括主观幸福感、对当前工作的满意度以及工作与生活的平衡。相关研究表明,主观幸福感的普遍程度在世界各区域差别较大,参与创业活动可能与较高的主观幸福感有关(叶文平等,2018)。Naudé等(2014)指出,创业有助于提高国民的生活满意度和幸福感,但内在机制仍需要进一步探索。近期的研究表明,独立和工作控制是创业者较高主观幸福感的来源,会让机会型创业者受益匪浅,即更具活力和影响力的创业会带来更高的主观幸福感(Larsson和Thulin,2019)。

总体而言,创业影响的研究关注创业活动对经济发展和国民幸福的贡献,并与政策制定密切相关。同时,数据可用性仍然是一个较大的障碍。例如,Naudé等(2014)的样本仅限于GDP水平中等至较高的国家的34个幸福感指标。鉴于此,三大数据库未来需要扩展关于生活和工作满意度的数据,以便更深入地探讨个人层面上幸福感与创业活动之间的关系。

(六)创业环境

创业活动的开展依赖于相互交织的特定社会、地理情境。创业情境是一个跨越多分析层次的多元现象,为创业者提供创业机会并对其行为设定界限,对创业起促进或约束作用。深化创业情境化研究有助于理解创业的本质、丰富性和动态性(Welter,2011)。在创业研究中,情境涉及与创业活动相关的外部状况、条件、情况或环境,情境化研究需考察包含与市场、关系以及地理等因素相关的社会、空间和制度情境作用。在现有的研究中,创业环境作为创业情境在宏观层面的呈现,其不同维度的离散影响得到了比较充分的探讨。三大数据库对各地区宏观因素的调查为创业环境方面的研究提供了大量的实证成果。创业环境能够影响个体创业意愿与创业过程,支持创业的环境需要具有统一的制度特征和建设性的创业文化与经济条件。自2013年以来,围绕制度、文化、经济三类创业环境的研究数量呈现上升势头。

第一类,制度环境研究。主要基于制度理论解释制度环境与创业活动之间的关系。Urbano和Alvarez(2014)引入Scott的三个制度支柱,提出监管、认知和规范制度中的每一个都会影响新企业的合法性和动机,并且是理解新兴经济体和发达经济体的创业现象所必需考量的。随后,创业学者吸收了这些想法,认为不同的制度环境可能导致不同国家的经济增长率(Acs等,2018)、创业结构(Dilli和Westerhuis,2018)、创业发展(Bosma等,2018)、社会创业率(Stephan等,2015)以及正式与非正式创业率(Thai和Turkina,2014;Dau和Cuervo-Cazurra,2014;Mallon和Fainshmidt,2022)等方面的差异。

第二类,经济环境研究。学者们重点考察宏观经济环境因素,如经济发展阶段、经济萧条、经济危机、经济自由度等,对创业活动的影响。Acs和Amorós(2008)基于其经济发展模式及所处的发展阶段,发现拉美国家在不同模式下遵循不同的竞争力路径,由此产生了企业动态差

异。Giotopoulos等(2017)考察2008年金融危机爆发前后欧洲国家高质量创业的变化,发现不利的经济条件会抑制创业活动的效果。Bruns等(2017)采用欧洲国家的GEM数据研究发现,由于全球金融危机和欧洲危机的冲击,2006—2014年的创业增长呈现消极结果。

第三类,文化环境研究。相关研究表明,作为特定群体或社会特有的潜在价值体系,文化塑造了某些人格特质的发展,并激励社会中的个人采取在其他社会中可能不那么普遍的行为,如创业。赵向阳和李海(2012)提出,集体主义、人文导向与权力距离较高时,中低GDP国家的创业活动比例大幅提升,但是在高GDP国家却抑制了创业活动。而高绩效导向、高未来导向和高不确定性规避文化则与高期望、高创新的创业活动有比较显著的正相关关系。Hechavarría和Brieger(2020)认为,一个社会的文化观念会影响社会创业的培养能力,当权力距离、人文导向和集体主义的文化观念程度较低,而未来导向和不确定性规避的文化观念程度较高时,女性创业者更有可能从事社会创业。

三大数据库为制度、经济、文化环境对创业活动的影响提供了扎实的数据基础。与此同时,学者们也提出,创业环境具有动态性,这对数据收集的长期可得性提出了更高的要求。鉴于不同城市创业环境数据的局限,未来需要持续跟踪各种创业环境的最新进展,以便揭示各地区创业环境与创业活动的关系,发掘创业环境因素对创业影响程度的差异(Bosma等,2018)。

整体来看,基于编码文献的研究主题分析发现(见表3):第一,GEM、PSED、CPSED三大数据库为创业研究贡献了必不可少的实证依据。创业数据的可得性和普适性为创业研究提供了活力源泉,有效推动了创业研究主题的演进。第二,创业研究的交叉融合充分说明其内在复杂性。本文认为,基于三大数据库的研究是在创业者与团队、创业机会以及更广泛情境下与时俱进的模式,因此探索各主题的发展并考虑主题边界能够更好地解释创业现象的复杂性与动态性。第三,三大创业数据库存在一定的局限性,仍需进一步完善。三大数据库包含了巨大的数据信息量,在逐年发展中已形成一个复杂的研究系统。学者们依托这些庞大的数据库能将更多的精力用于研究创业系统中不同要素的内在联系,收获更丰富多彩的创业研究成果。同时,有必要在新的情境下对已有结论和待解决问题采取更进一步的验证与考察,并就数据库建设如何更好地服务于创业研究发展提出改进建议。

四、三大创业数据库存在的不足和建议

由上述分析可见,三大数据库为推动创业研究做出了重要贡献,同时也存在不足之处。为进一步完善数据库建设以服务未来的创业研究,本文从五个方面进行总结(参见表4)。

(一)部分项目缺乏有效的长期跟踪,无法保证研究结论的动态性和连贯性。创业是动态化的过程,GEM缺乏纵向的跟踪数据,导致研究结论主要停留在相关性分析而缺乏因果关系推断,降低了结论的解释力(De Clercq等,2014;Stuetzler等,2014;Schillo等,2016)。由于数据集只涵盖较短的时间,缺乏持续的动态性的数据,无法估计组织和个人行为随时间推移产生的共同演变,仅能从较短时间或静态角度讨论参与创业过程任何阶段的可能性(Van Stel等,2005;Boudreaux和Nikolaev,2019;Mallon和Fainshmidt,2022)。部分变量如国家的社会资本和制度在短期内变化不大,需要在较长时间才能发现其内在规律。PSED数据仅在1998—2006年收集,调查过程存在窗口期,造成缺乏对创业者的长期持续观察,对数据的时间有效性存在潜在威胁(Newbert等,2013;Chen等,2018)。CPSED中对新生创业者的动态跟踪调查时间跨度较小,仅调查新生创业者过去一年的创业行为,不足以覆盖新企业生成过程(龙丹等,2013),影响结论的准确性。研究者考察创业实践的长期发展原因依赖于现有项目的长期跟踪,因此,未来的数据库需要匹配创业发展历程的差异和变化,把握创业长期趋势。

表3 主要研究主题分析

研究主题	理论基础	主要研究结论	数据来源	作者(年份)
创业者与团队	性别角色一致性理论	女性创业者相比男性参与创业的可能性较低	GEM 2002	Minniti和Nardone(2007)
	创业判断	创业认知影响个体判断从而导致创业行为差异	PSED 2005	Uygun和Kim(2016)
	社会网络理论	人力、社会和金融资本促进机会驱动型创业	GEM 2002—2012	Boudreaux和Nikolaev(2019)
	资源基础观	金融资本与心理资本联动主导的约束模式是女性早期创业的关键制约问题	GEM	王汉瑛等(2021)
创业机会	社会网络理论	居住在大聚集区的人更容易感知创业机会	GEM 2002	Arenius和De Clercq(2005)
	创业效果逻辑理论	创业者伦理意识水平有助于创业机会的识别	CPSED 2009—2010	李华晶等(2014)
	制度理论	创业者与外部环境的互动影响其识别创新型或模仿型创业机会	GEM	Young等(2018)
	制度经济学理论	国家领导人的民粹主义言论造成政权不确定性,导致创业者降低识别创业机会的动机	GEM	Bennett等(2022)
创业融资	-	能够获得任何类型外部融资的创业者更有可能创业	PSED	Gartner等(2012)
	制度理论	非正式投资的可获得性和早期创业活动水平之间存在积极而显著的关系	GEM 2003—2007	De Clercq等(2014)
	社会认知理论	创业资助促进新生创业者实施创业导向	GEM	黄永春等(2020)
创业决策	期望理论	创业失败恐惧影响创业者进入创业的决策	GEM 2002	Arenius和Minniti(2005)
	企业成长理论	高成长导向的创业者往往进行更多的计划准备工作与资源整合工作	CPSED	张玉利等(2010)
	-	主动或战略性的创业退出行为是积极的	PSED I	Yusuf(2012)
	制度理论	带薪休假、育儿补贴和兼职就业机会等制度安排缓解了工作-家庭冲突,显著提高女性在高成长导向创业中的比例	GEM	Thébaud(2015)
创业影响	增长理论	高增长潜力的初创企业和机会型创业增强了经济增长	GEM 2002	Wong等(2005)
	-	创业有助于提高国民生活满意度和幸福感	GEM 2000—2007	Naudé等(2014)
	文化二重性	创业对促进幸福感方面具有重要意义,且在不同国家文化影响下程度有所差异	GEM 2012—2016	叶文平等(2018)
创业环境	制度理论	创业环境支持性对创业活动有正向影响	CPSED	李华晶等(2012)
	经济发展理论	国家经济环境受冲击时,创业增长率呈现消极结果	GEM 2001—2014	Bruns等(2017)
	实践理论	国家文化观念影响创业实践	GEM 2009—2015	Hechavarría和Brieger(2020)

资料来源:根据编码文献整理,“-”指原文没有明确交代。

(二)数据库变量测量不够细致,难以全面体现创业概念的独特内涵。具体而言,GEM数据库的变量测量存在三方面的不足。第一,GEM变量测量水平忽略中观层面数据。根据Sternberg和Wennekers(2005)的分析标准,可以将GEM数据分为微观、中观、宏观三个层面的数据。然而,由于其变量收集范围的局限,大多数基于GEM的研究侧重于从微观(47.4%)和宏

表4 三大创业数据库的主要不足之处

数据库主要不足之处	可能造成的影响	主要研究
数据跟踪缺乏长期性	①当研究跨越长时间的变量时受限,不利于研究动态趋势型话题;②难以得到稳定的因果关系	Wong等,2005;李华晶等,2014;De Clercq等,2014;Mallon和Fainshmidt,2022
变量测量不够细致	①模型变量与实际测量之间存在差异,导致研究结论普适度不足;②研究水平覆盖面不够,易忽略跨层次研究	Arenius和Minniti,2005;Boudreaux和Nikolaev,2019;何良兴和张玉利,2020;张秀娥等,2021
指标测量缺乏客观性	数据的客观性有失偏颇,可能导致不同情境下的研究结果不相吻合	Minniti和Nardone,2007;Brieger和Gielnik,2020
调研主题分散	新兴话题缺乏连续性、系统性的追踪调查,使得理论研究落后于实践	Gartner等,2012;Álvarez等,2014
数据样本局限	新兴市场国家的创业活动快速发展,若样本仅局限于经济发达国家或地区将缺乏结论代表性	Uygur和Kim,2016;Bruns等,2017;Boudreaux等,2019

资料来源:根据编码文献整理。

观(45.3%)的角度分析创业活动,只有7.4%的研究文献侧重于中观层面的分析(Álvarez等,2014)。第二,GEM大部分数据的性质仅为二分变量,导致研究方法单一化。由于众多题项答案设置仅为二分变量(是或否),实证研究中在微观分析水平通常使用的统计方法为logit、probit和tobit模型(42%),其次是与宏观水平相关的多元线性回归分析(29%)、面板数据(13%)和其他技术(16%)(Álvarez等,2014)。然而,由于二分变量较低的解释力度,现有研究较少使用定性方法,无法解决实践中的复杂性与可延伸性问题(Darnihamedani等,2018)。第三,GEM对创业的特定概念通常以构念形式测量,比如对“失败恐惧”的测量,失败恐惧涉及感知失败经历风险及其后果,是创业者在失败后所感知的复杂情绪,然而目前对其测量主要通过“是否担心创业失败会阻碍创立企业”这一条目实现(Arenius和Minniti,2005;何良兴和张玉利,2020)。“创业自我效能感”的测量题项“是否认为自己具备创业所需的知识、技能或经验”(张秀娥等,2021)、“创业警觉性”的测量题项“是否认为未来6个月有创业的好机会”(Boudreaux等,2019),皆使用单一条目而非条目组合,无法全面反映概念的复杂性。GEM对“创业导向”的测量也缺少精确的界定。如创业导向的维度反映了企业创业行为的特征,Pindado和Sánchez(2017)的研究从风险性、主动性和创新性三个维度标识创业导向,对风险性导向的测量题项却与“失败恐惧”的测量项一致,影响了创业导向概念的精准性。因此,现有测量的标准细致程度不足将导致创业研究过程的粗略化,使得创业发展止步于目前阶段的概念,无法进一步对相关概念进行细分从而探索更多子领域的结构和关系。目前,创业特定概念的相关测量已有比较成熟的量表,未来数据库的建设可以从中汲取经验,改善测量指标。

(三)部分指标测量缺乏客观性,以致真实评价存在偏差。GEM项目有两个主要的原始数据来源,即成人人口调查(APS)和国家专家调查(NES)。目前已有超过87%的文章使用APS数据,3%使用NES信息,10%同时使用两种信息来源(Álvarez等,2014)。APS的数据贡献显而易见,开发利用NES的信息也是未来研究的重要趋势。然而,GEM中的部分题项如创新性、创业技能的测量主要基于自我报告的形式,个人主观判断可能会产生认知偏见(Darnihamedani等,2018;Mickiewicz,2017),引起共同方法偏差,而NES数据是基于专家印象获取,缺失客观数据的测度(Brieger和Gielnik,2020),两者都存在主观性测量问题。在PSED中,关于创业原因的调查存在回顾性偏差和幸存者偏差,试图创业但已经退出、曾经创办了企业(但现在不存在)的创

业者均不在样本中。这些数据仅是创业者目前的观点,而不是在创业时的观点(Carter等,2003)。未来需要采用更多的客观衡量标准以提高研究结论的准确性。

(四)项目的固定调研主题单一,新增调研主题仍较为分散,连贯性不足。GEM每年拥有固定类型话题,现有研究中采用的因变量主要涉及创业活动(59%)、创业抱负指标(14%)(如增长抱负、创新、就业增长、出口)、女性创业(10%)、经济发展(5%)以及创业动机(5%)(Álvarez等,2014)。鉴于单一的固定研究主题,GEM每年会新增特殊主题,带来了正反两方面的影响。一方面,为研究者提供了有价值的一手数据,及时反映创业研究的关注焦点,突破理论与实践的结合,从而推动研究发展。例如,2020年新冠疫情爆发造成全球经济下滑,创业作为经济活动的重要组成部分也面临挑战,因此GEM将这一主题作为新的特殊主题纳入报告中展开研究。另一方面,主题众多导致研究问题缺少连贯性,尤其是GEM的主题调研占据调研问卷的绝大部分问题,诸多传统调研问题无法得到保留和延续。例如,2020年在新增疫情主题的同时却将2019年的青年创业主题删减,造成上一年存在的主题数据在下一年缺失的局面。因此,数据库对于实践主题的选择还需要结合研究趋势的考量,使得创业研究主题日益丰富的同时满足数据的可得性。

(五)数据库样本量不足,存在样本流失与收集分布不均问题,易形成研究样本数据两极分化。GEM数据库中可获得不同国家的新生创业者创业活动的可比信息,但是其横向数据的代表性仍不足以满足现有研究增长的需求,中高收入国家的样本量远超低收入国家,大多数国家来自西欧或东欧,少数国家来自南美,可能造成跨国异质性(Boudreaux等,2019)。此外,部分国家在调研后期退出项目,例如自2006年以来,新西兰退出数据收集(Frederick和Monsen,2011),可能妨碍形成更一般化的结论。在PSED数据库中同样存在数据缺失问题,限制了结果的统计有效性(Renko等,2012)。因此,对于国际数据库而言,保证调研国家的典型性和代表性是一大优势,国内数据库同样需要采取分层随机抽样选取调研样本的做法,以帮助获取更大的样本量和更广的覆盖面。

综上,虽然三大数据库项目为创业研究提供了重要依据,但是也存在缺乏有效长期跟踪、变量测量模糊、测量指标主观、调研主题分散以及样本量不足的问题。不断发展的创业研究对现有数据库需要满足更深入的机制研究提出的更高要求,在理论研究发展的同时也可带动创业数据库的迭代创新。因此,未来对现有数据库的改善或构建全新数据库时均应将这些因素纳入考虑范围。

五、结论与讨论

(一)研究结论

创业研究在成果不断丰富的同时也遭遇理论创新不足的瓶颈,因此亟待构建高质量的数据库,以便推进情境化研究,突破现有理论的局限。本文首先详细介绍了GEM、PSED、CPSED三个数据库的设计框架与研究方法,选择在英文FT50期刊、中国国家自然科学基金委管理科学部认定的30种学术期刊以及*Small Business Economics*期刊上发表的219篇相关研究文献进行编码,通过文献计量方法对源文献的研究主题进行梳理分析,得出以下结论。

首先,针对三个数据库的开发使用情况,GEM侧重于创业驱动因素的分析,PSED以创业过程与特征的探索为主,CPSED目前的研究仍较为分散,缺乏系统性。总体上看,现有理论研究对于分析框架与问题边界的讨论略显不足,跨层次多视角的研究是重要趋势。

其次,现有创业研究未能充分体现创业情境化的特点,尤其是针对中国情境的创业研究,意义重大但任务艰巨。创业研究的情境化对于理解创业何时、如何、为何发生以及谁参与其中

具有重要意义,可以极大地提升人们对创业规律的认识。现有研究仅针对创业环境,关注创业活动宏观层面的总体特征,将创业看作是一系列经济、制度、文化等要素的组合。然而,创业研究的情境化不同于创业环境研究,原因在于,创业情境化不仅要关注创业环境,还需要进一步考虑时间和历史情境、远端和近端的空间情境等其他情境特征(Welter, 2011)。目前创业研究大多关注离散的宏观环境变量,针对更综合的视角下创业情境化的研究不足,未来还需要更深入的探索。此外,改革开放以来,中国面临由计划经济向市场经济的深刻体制变革,中国情境意味着需要刻画中国这一新兴经济体在制度转型、现代化进程中深受独特文化影响的过程(蔡莉和单标安, 2013;董保宝, 2014),这可能诱发不同于西方国家的创业过程与作用机制,因此拓展对中国情境的解读具有深刻意义。情境化研究可以揭露更多中国情境的独特性。开展中国情境下的创业研究有赖于中国数据库的构建和完善,开发符合中国情境的创业数据库需要基于制度、经济以及文化等视角的深入剖析,以动态视角分析中国独特的创业问题。

最后,不断发展的创业研究对数据库构建提出了新的要求,本文基于现有研究主题框架对现有数据库的不足进行了分析和归纳,以便促进创业数据库的发展完善。

(二)研究贡献

本文的研究贡献和启示主要包括以下几方面:

第一,本文揭示了三大创业数据库对推动创业研究进步的重要价值。创业领域一直以基于理论的定量研究为主导,依靠调查、档案数据和复杂的统计技术进行数据分析,而高质量的创业数据对于提升创业研究的严谨性和稳健性、深化研究意义具有重要的支撑作用。本研究基于现有的创业研究成果,揭示了数据库在创业研究中发挥的多层次作用及其驱动机制。未来可以立足于特定创业情境,借助数据库平台开展更多理论驱动的融合性研究,进一步发挥数据库在创业研究的潜在作用,深化既有研究问题和拓展新议题。事实上,随着基于创业数据库研究成果的不断增加,最近已有学者呼吁对创业研究进行复制研究,以揭示创业的最新图景(Crawford等, 2022)。

第二,本文提出了创业研究发展对研究数据的创新要求。创业研究在发展中逐渐形成特有问题,研究重点从创业者特质转向对创业行为与过程的关注,对创业导向、创业机会等理论的持续探索和总结,以及对社会创业、女性创业等现象的关注。随着创业作为学术领域的制度化和学术合法性的提高,一批成功且日益壮大的创业学者聚焦于更前沿的研究问题,提供强有力的知识,从而发展出更诱人的创业领域。因此创业研究的发展对构建一个成熟数据库提出紧迫的要求,不同视角的理论以及多元主题的探讨均需要更加完善且集中持续的数据支持,本文证实了系统搭建数据平台的重要作用,为创业学者打开了新的视角。

第三,本文为创业数据库的深化构建提供了重要的理论支撑。创业研究必须建立复杂的数据库(Pindado和Sánchez, 2017),以便收集完整的数据,更有力地描述和预测创业行为与过程。本文首先通过呈现不同数据库的优点与不足,为创业研究者更好地运用这些数据库提供理论指引。此前大多学者仅针对单个数据库进行分析,缺乏多数据库的有效整合,本研究将三个数据库加以对比分析,有助于研究者多元地揭示创业活动规律。数据收集难是长期制约创业研究的问题,本文的整合框架有助于探索更有效的调查设计方法体系,促进创业研究的知识共享,也为后续设计创业研究的大型数据库提供指引。

第四,本文强调了创业情境化研究的重要作用及其与数据库的紧密关联。创业比其他领域更需要情境化,因为创业不仅范围广泛形式多样,而且过程经常复杂多变,情境化的研究能帮助创业学者探讨更深层的创业机理和更多元的创业模式。因此本文强调创业情境化研究的重要性,有助于研究者借助数据库考虑情境问题,突破已有研究局面,深化创业情境化中的问题理解与现象解释。

(三)研究展望

翔实完备的数据是开展实证研究的基础,在很大程度上决定了研究结论的科学性和普适性。因此,建设科学、可靠的创业数据库具有重要的价值,对学术研究以及政策实践都能够发挥重要作用。结合现有的研究成果及其不足,本文认为未来大型数据库的构建和使用需要关注以下六个方面问题。

第一,开展长期动态跟踪调查。相关关系和因果关系的研究可以从跨越更长时间的纵向数据获益,因而需要将动态因素纳入假设的关系中。长期有效的跟踪调查是确保真实反映创业变量关系的重要条件,是解释创业现象复杂性的关键,可为科学地评估创业发展、了解创业者与创业过程以及制定跨越性的创业政策提供可靠依据。高水准的实证研究需要高质量的数据支持。长期跟踪调查可能面临成本昂贵、设计复杂和操作困难的难题,但与截面调查和趋势调查相比,它能够采集更多有价值的信息,为科学研究带来更大的回报。因此,未来的研究应努力收集和分析微观层面的纵向数据和宏观层面的面板数据,以提炼稳定性结论。

第二,注重设计方法的合理性与变量测量的精细化。目前数据库中大多数问题测量基于自我报告形式,未来需要结合客户的看法或绩效指标,从更客观的视角收集数据。此外,问卷设计需要阐明相关结果变量之间的复杂关系,满足更精细的数据要求依托于复杂数据库,以便收集同一创业者长期的内外部变化。未来的研究可增加考察创业的潜在个人层面决定因素的作用。目前,父母身份等因素尚未纳入分析中。

第三,优化调研主题,增加调查样本。创业数据库的研究主题应聚焦前沿科学问题,促进跨学科交流,在多样性与前沿性之间寻找平衡,以期为实践活动提供更多的理论支持。以GEM为例,问卷主要由核心主题与特殊主题两部分组成,特殊主题根据研究需要进行变动。因此在未来的工作中,可以把有研究意义和价值的主题增补进下一次的调研中,完成数据库的创新变更,这样能够更好地匹配现实需要。此外,进一步的研究涉及大样本的国家因素分析和多变量数据分析,需要进一步完善测量数据、增加样本数量,提高结论信度。GEM是目前为止关于社会创业的唯一大型数据库(Estrin等,2016),因此,关于特定主题的研究严重依赖数据库的使用,未来相关主题的探索需要更集中,以提高数据利用率。在数据库建设之前对社会进行持续不断的观察,既是探索创业中的现实问题的需要,也可以与其他跨国数据库进行关联,提升探索新议题的可能性。

第四,开展多学院联合发布数据库,扩大数据收集范围。目前三大数据库多由一到两个学校组织开展实施,未来针对中国情境的创业数据库的构建与落地需要联合多个学校主导,结合各省市组织实施开展,每年一次不间断地进行全国抽样调查,收集全面、系统、翔实的调查数据,充分记录创业活动及创业者的成长过程,形成强大、完备的时间维度数据。数据库是战略研究中的重要资源,未来需要进一步组建和巩固数据库核心团队,开发具有及时性的创业政策报告,这将发挥更大的作用。此外,通过出版平台可以有力地促进研究者与创业者的交流,打破理论与实践的壁垒,更好地发挥创业研究的功能。

第五,逐步实现数据库的共享使用,为创业研究学者打造开放获取平台。创业研究的迅猛发展在一定程度上得益于数据库的高质量建设与利用,透过数据能够更好地揭示创业活动的方方面面及其蕴含的机制,因此充分利用好数据库具有时代意义。目前,国内外部分数据库均实现开放共享,研究者可以实行注册登记申请获取;未来,创业数据的免费或付费共享也将成为趋势。

第六,将国家战略需求与科学探索目标相结合,探索开发适应中国情境的创业研究数据库,深化服务中国本土创业实践。在中国企业数字化转型的大趋势下,科技革命与产业革命推动数字经济快速发展,数字化、智能化、可持续化作为提升竞争力的基点已成为企业重要的发

展趋势。因此,未来中国创业数据库的发展既要借鉴国外重要研究成果,也要结合中国转型经济的具体情境,紧密围绕中国重要创业发展路线,探讨中国情境下更多的创业问题,为中国“双创”的纵深发展如商业模式创新(杨俊等,2020),以及企业数字化转型提供深刻的应用价值与社会贡献。

主要参考文献

- [1]董保宝. 创业研究在中国: 回顾与展望[J]. 外国经济与管理, 2014, 36(1): 73-80.
- [2]宋渊洋, 赵嘉欣. 地区社会信任多样性对个体创业进入决策的影响研究——来自全球33个国家的经验证据[J]. 管理评论, 2020, 32(9): 118-131.
- [3]田莉, 张玉利. 创业者的工作家庭冲突——基于角色转型的视角[J]. 管理科学学报, 2018, 21(5): 90-110.
- [4]杨俊, 张玉利, 韩炜, 等. 高管团队能通过商业模式创新塑造新企业竞争优势吗?——基于CPSED II数据库的实证研究[J]. 管理世界, 2020, 36(7): 55-76, 88.
- [5]叶文平, 杨学儒, 朱沆. 创业活动影响幸福感吗——基于国家文化与制度环境的比较研究[J]. 南开管理评论, 2018, 21(4): 4-14.
- [6]张秀娥, 王超, 李帅. 制度环境、创业自我效能感与创业意愿[J]. 科研管理, 2021, 42(10): 1-17.
- [7]张玉利, 闫丽平, 胡望斌. 新企业生成中创业者成长期望研究——基于CPSED首轮调查数据分析[J]. 管理学报, 2010, 7(10): 1448-1454, 1471.
- [8]周冬梅, 陈雪琳, 杨俊, 等. 创业研究回顾与展望[J]. 管理世界, 2020, 36(1): 206-225.
- [9]Acs Z J, Estrin S, Mickiewicz T, et al. Entrepreneurship, institutional economics, and economic growth: An ecosystem perspective[J]. *Small Business Economics*, 2018, 51(2): 501-514.
- [10]Álvarez C, Urbano D, Amorós J E. GEM research: Achievements and challenges[J]. *Small Business Economics*, 2014, 42(3): 445-465.
- [11]Bosma N, Content J, Sanders M, et al. Institutions, entrepreneurship, and economic growth in Europe[J]. *Small Business Economics*, 2018, 51(2): 483-499.
- [12]Boudreaux C J, Nikolaev B N, Klein P. Socio-cognitive traits and entrepreneurship: The moderating role of economic institutions[J]. *Journal of Business Venturing*, 2019, 34(1): 178-196.
- [13]Chen H S, Mitchell R K, Brigham K H, et al. Perceived psychological distance, construal processes, and abstractness of entrepreneurial action[J]. *Journal of Business Venturing*, 2018, 33(3): 296-314.
- [14]Crawford G C, Skorodzyevskiy V, Frid C J, et al. Advancing entrepreneurship theory through replication: A case study on contemporary methodological challenges, future best practices, and an entreaty for communality[J]. *Entrepreneurship Theory and Practice*, 2022.
- [15]Dau L A, Cuervo-Cazurra A. To formalize or not to formalize: Entrepreneurship and pro-market institutions[J]. *Journal of Business Venturing*, 2014, 29(5): 668-686.
- [16]Du K, O'Connor A. Entrepreneurship and advancing national level economic efficiency[J]. *Small Business Economics*, 2018, 50(1): 91-111.
- [17]Hechavarría D M, Brieger S A. Practice rather than preach: Cultural practices and female social entrepreneurship[J]. *Small Business Economics*, 2022, 58(2): 1131-1151.
- [18]Lim Y, Suh C S. Where is my partner? The role of gender in the formation of entrepreneurial businesses[J]. *Small Business Economics*, 2019, 52(1): 131-151.
- [19]Naudé W, Amorós J E, Cristi O. “Surfeiting, the appetite may sicken”: Entrepreneurship and happiness[J]. *Small Business Economics*, 2014, 42(3): 523-540.
- [20]Pindado E, Sánchez M. Researching the entrepreneurial behaviour of new and existing ventures in European agriculture[J]. *Small Business Economics*, 2017, 49(2): 421-444.
- [21]Thébaud S. Business as plan B: Institutional foundations of gender inequality in entrepreneurship across 24 industrialized countries[J]. *Administrative Science Quarterly*, 2015, 60(4): 671-711.

[22]Urbano D, Alvarez C. Institutional dimensions and entrepreneurial activity: An international study[J]. *Small Business Economics*, 2014, 42(4): 703-716.

[23]Yusuf J E. A tale of two exits: Nascent entrepreneur learning activities and disengagement from start-up[J]. *Small Business Economics*, 2012, 39(3): 783-799.

The Three Major Databases of Entrepreneurship and Entrepreneurship Research

Ye Wenping¹, Pan Shimin¹, Yang Jun²

(1. *School of Management, Jinan University, Guangzhou 510632, China*;

2. *School of Management, Zhejiang University, Hangzhou 310058, China*)

Summary: As the three authoritative databases of entrepreneurship research, Global Entrepreneurship Monitor (GEM), Panel Study of Entrepreneurial Dynamics (PSED), and Chinese Panel Study of Entrepreneurial Dynamics (CPSED) have greatly contributed to the contextualization and independent development of entrepreneurship research. However, the existing research findings are far from enough, lacking the integrated framework that can fully reflect the contribution of three entrepreneurship databases. In view of this, this paper systematically combs the entrepreneurship research themes and three major database deficiencies, and then constructs a comprehensive analytical framework to promote the flourishing of entrepreneurship research. Firstly, this paper provides an overall introduction to the three major entrepreneurship databases, and compares and analyzes their design ideas and data structures, thus helping scholars further identify what can be used in the three major databases. Secondly, by adopting the method of bibliometric analysis, this paper reviews 219 entrepreneurship literatures which use the three databases and are published in top journals, and analyzes their publication years and journals to understand the basic distribution. Thirdly, this paper summarizes the development of research themes and constructs an integrated framework to understand the unique advantages of the three databases that provide important guidance for the evolution of entrepreneurship research. Among them, this paper focuses on six hot research themes, including entrepreneurs and teams, entrepreneurial opportunities, entrepreneurial financing, entrepreneurial decision-making, entrepreneurial impact, and entrepreneurial environment. Finally, based on the above analysis, this paper summarizes the five main shortcomings of the three major databases, namely, lack of effective long-term tracking, vague variable measurement, subjective variable measurement, scattered research themes, and insufficient sample size, and proposes suggestions for the construction of large-scale databases in the future. This paper innovatively reveals the great value of the three major entrepreneurship databases in advancing entrepreneurship research, proposes the innovative requirements for research data for the development of entrepreneurship research, provides important theoretical support for the construction of entrepreneurship databases, and emphasizes the significant role of entrepreneurship contextualization and its close connection with databases. The study has certain theoretical significance for further promoting entrepreneurship research and achieving a better construction of entrepreneurship databases.

Key words: entrepreneurship research; bibliometric analysis; research themes; database

(责任编辑: 宋澄宇)