

北京农业经济学会 2017学术年会

2017.12.23



中國人民大學

RENMIN UNIVERSITY OF CHINA

互联网使用有助于农民工创业吗？

——来自流动人口动态监测调查的经验证据

汇报人：祝仲坤

中国人民大学农业与农村发展学院

➤ 提出研究问题

- ◆ 创业不仅是推动国家创新的重要动力（Aghion and Howitt, 1997, 还是推进经济长期增长的重要“引擎”（Banerjee and Newman, 1993; Samila and Sorenson, 2011），同时也是拉动就业的重要“抓手”（Glaeser et al., 2010）。
- ◆ 伴随着中国经济步入“新常态”，创新创业活动日渐成为社会各界关注的焦点。2014年9月，李克强总理在夏季达沃斯论坛首次提出“大众创业、万众创新”的号召。2015年3月，李克强总理在政府工作报告中明确指出：推动大众创业、万众创新，既可以扩大就业、增加居民收入，又有利于促进社会纵向流动和公平正义。
- ◆ 新常态的时代背景下，推动创新创业已经成为了重要的国家战略。



➤ 提出研究问题

- ◆ 2015年6月，国务院办公厅发布了《关于支持农民工等人员返乡创业的意见》明确支持农民工等人员的创业活动。
- ◆ 在推动创业的进程中，农民工是创业的重要支持对象。
- ◆ 实际上，农民工在“流动”过程中，不仅带动了信息与技术的流动，同时积累了一定的人力资本与物质资本（Parker, 1997; Hamilton and Barton, 2000），是一类潜在的创业者。
- ◆ 因此，积极发挥农民工的创业潜能，有可能掀起“草根创业”浪潮，对拉动农村就业、实现城乡融合发展、推动乡村振兴、新型城镇化等重大战略大有裨益。



➤ 提出研究问题

- 现有文献当中，已有不少研究关注影响农民工创业的因素，家庭禀赋（石智雷等，2010；陈文超等，2014）、人力资本（汪三贵等，2010；张广胜、柳延恒，2014；）社会资本（Ma，2002；朱红根、谢春艳，2012）、社会网络（庄晋财等，2014；Zhang and Zhao，2015）、信贷约束（Jia et al.，2013；翁辰、张兵，2015）语言技能（魏下海等，2016）、外出务工经历（周广肃等，2017）等众多因素备受关注。
- 不过令人遗憾的是，已有文献大多忽略了互联网使用对农民工创业行为的影响。
- 实际上，自20世纪90年代以来，伴随着信息通信技术的进步与扩散，互联网逐渐成为当今社会的“时代洪流”，将人们带入了前所未有的互联网时代。



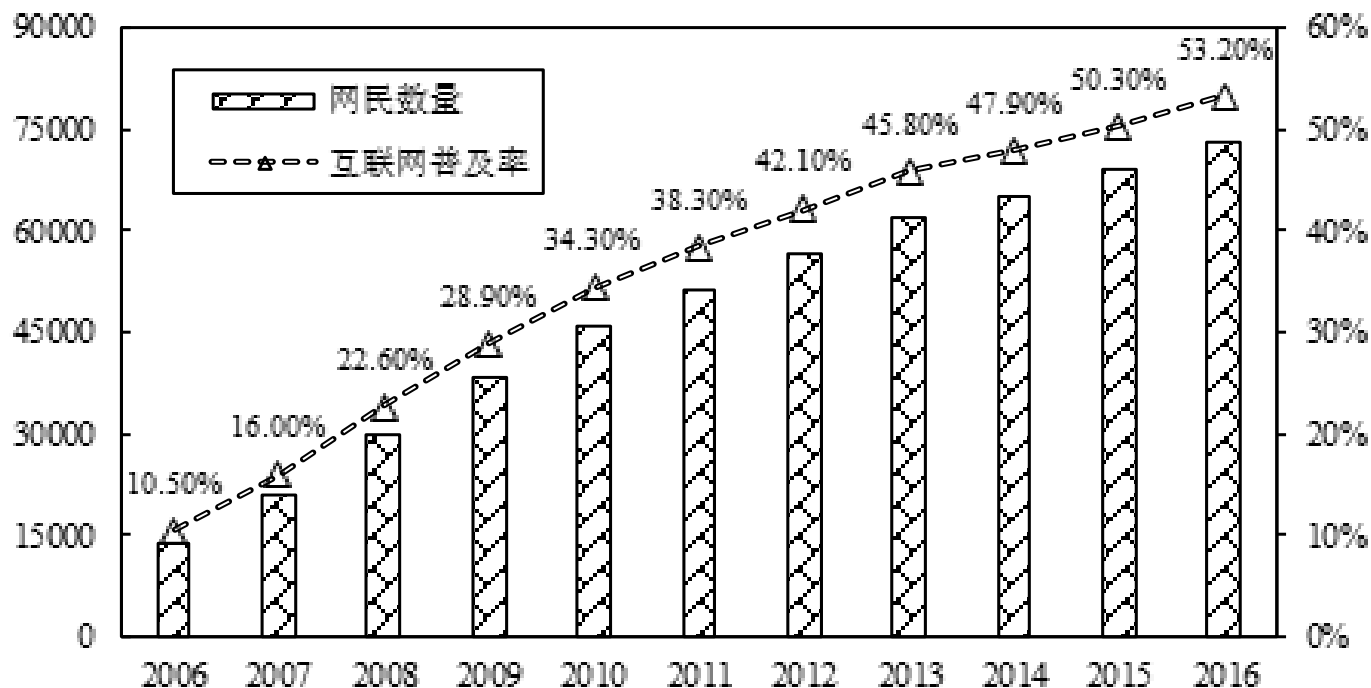


图 1 2006-2016 年中国网民数量及互联网普及情况

- ◆ CNNIC发布的2016年《中国互联网络发展状况统计报告》显示，中国互联网用户数量从2006年的13700万，迅速增加至2016年的73125万，互联网普及率从10.5%上升到53.20%，即便在农村地区，互联网的普及率更是从2006年的3.1%，快速攀升至2016年的33.1%。



➤ 提出研究问题

- ◆ 已有研究表明，互联网的普及对创新创业活动已经产生了深刻影响。
 - 首先，作为一种新技术，互联网不仅有助于减少市场摩擦、降低市场交易成本，还能够推动企业研发创新、催生新的商业模式，从而形成大量的商业机遇（李海舰等，2014；罗珉、李亮宇，2015）；
 - 其次，近几年，异军突起的互联网金融扩大了金融市场的服务对象，降低了金融市场的服务门槛，有助于缓解信贷约束，为创业者提供金融支持（王子敏、蔡羚奕，2016；刘柳、屈小娥，2017）；
 - 再者，伴随着移动互联网的兴起，移动社交媒体能够为人们提供更为即时便捷的交流互动平台，拉近了人与人之间的关系，由此形成的网络效应（史晋川、王维维，2017），使得人们创业的形式更具多样化，依托移动社交媒体的创业模式正如雨后春笋般涌现。
- ◆ 那么，互联网能够促进农民工创业吗？互联网对农民工创业影响是否具有代际差异？互联网影响农民工创业的作用机制又该作何解释呢？
- ◆ 本文将循着上述问题的思路，结合国家卫计委组织实施的2012年度流动人口动态监测调查数据，力求全面而系统地考察互联网使用对农民工创业行为的影响。



➤ 文献综述

- 伴随着第三次科技革命，信息通信技术实现了飞速发展，作为新生事物的“互联网”开始“野蛮生长”，不仅对经济社会发展产生了战略性、全局性的影响，同时也在潜移默化地影响着居民的行为观念与生活状态。
- 已有研究表明，互联网的推广与普及能够促进经济增长（Czernich et al., 2011；郭家堂、骆品亮，2016），带动国际贸易发展（Freund and Weinhold, 2004；施炳展，2016；李兵、李柔，2017），推动金融领域不断革新（谢平、邹传伟，2012；刘传明等，2017），同时还能扩大就业（Atasoy, 2013；周冬，2016），增强人们的投资兴趣（Liang and Guo, 2015）。
- 此外，互联网还正在深刻地影响着人们的生活，当前，探讨互联网对价值观、社会认同等的研究开始不断涌现（陈思宇等，2016；孟天广、季程远，2016）。



➤ 文献综述

- 关于互联网如何影响创业，现有文献已经展开了一定的探讨。从国外研究来看，研究的焦点集中在互联网基础设施以及互联网技术对创业带来的影响。
- Kim and Orazem（2012、2017）研究指出，宽带基础设施的普及对新企业的设立与进入具有积极作用。
- Audretsch et al.（2015）、Gillett et al.（2016）也证实了宽带对创业的促进作用。
- 关于互联网技术对创业的影响，Cumming and Johan（2010）指出互联网技术的进步对农村社区创业活动的正向影响；Kotnik and Stritar（2015）指出互联网及信息通信技术对新企业创立、中小企业发展存在积极影响。



➤ 文献综述

- ◆ 国内互联网发展日新月异，互联网与创业相关研究研究也开始涌现。
 - 王子敏、蔡羚奕（2016）利用抓取的人人贷平台交易数据探讨了互联网借贷对工薪阶层创业所发挥的作用；湛泳、徐乐（2017）CFPS2014，研究了“互联网+”背景下包容性金融对家庭创业决策的影响。
- ◆ 上述两份文献虽关乎互联网与创业，但实质上关注的是依托于互联网的金融手段对创业的影响。
 - 史晋川、王维维（2017）基于CFPS2010、2012，指出互联网的使用能够使个人创业概率提高4.1%-4.8%；周洋、华语音（2017）运用CFPS2014，研究发现互联网使用能够显著提升农村家庭的创业倾向。
- ◆ 上述文献与本文研究话题最为接近，但遗憾的是两者均未关注农民工群体。
 - 相比之下，王子敏（2017）、马俊龙、宁光杰（2017）关注了农民工群体的互联网使用问题，但均关注的是就业选择以及收入水平，并未考虑农民工的创业问题。



➤ 文献综述

- 综上所述，本文可能的边际贡献或者创新之处在于：
- 视角比较新颖：与过往研究略有不同，本文首次关注互联网对农民工创业决策的影响。
- 数据有一定代表性：本文使用的流动人口动态监测调查数据是由国家卫计委组织实施、符合严格抽样、覆盖全国的大样本调查数据，是研究农民工问题的权威性数据。
- 实证策略相对严谨：考虑到农民工使用互联网很可能是“自我选择”的结果，结合倾向得分匹配法构造反事实框架，以纠正选择性偏误。进一步，考虑到遗漏变量、互为因果等潜在的内生性问题，选取省级层面的“人均域名数”和“宽带用户比”为工具变量，利用Ivprobit模型进行更为稳健的估计。不仅如此，本文还运用中介效应模型探讨了互联网对农民工创业的影响机制，即互联网是否通过丰富社会网络提升农民工的创业倾向。



➤ 数据与变量

- 本文使用的数据是由国家卫计委发起并组织实施的**2012年流动人口动态监测调查数据**。
- 该调查基于分层、多阶段、与规模成比例的**PPS**抽样方法，样本覆盖了全国**31**个省、市、自治区以及新疆生产建设兵团共**32**个省级行政单位，调查对象为在流入地居住一个月以上，非本区（县、市）户口的**15-59**周岁流动人口。
- 考虑到本文关注的是互联网技能对农民工创业行为之间的关系，因此，本文仅考虑流动人口中处于就业状态的农业户籍部分，进一步，在对其他变量的缺失值、错误值处理后，最终获得包括**112296**个观测值的基准样本。



➤ 数据与变量

□ 被解释变量——农民工的创业行为。

- ✓ 问卷询问了被访者的就业省份，回答包括：雇员、雇主、自营劳动者和家庭帮工。考虑到家庭帮工往往属于不计薪酬的非正规就业形态，且比例极低（1.22%），故做删除处理（宁光杰、段乐乐，2017）。
- ✓ 根据创业理论可知，创业可以划分为广义创业和狭义创业。狭义创业主要指的是机会型创业，对应的选项为“雇主”（10.42%）；广义创业除了包括机会型创业外，还包括生存型创业，对应的选项为“自营劳动者”（31.53%）。
- ✓ 对于农民工而言，“自营劳动者”很可能不是主动的“创业”，而是生计所迫导致的无奈之选（周广肃等，2017）。这就意味着，“雇主”和“自营劳动者”样本很可能具有较强的异质性，回归过程很可能产生偏误；更为重要的是，就创业行为而言，机会型创业这种更具“创造性破坏”的活动才是企业家精神的核心表现，也是本文考察的重点所在。

□ 因此，本文将“雇主”定义为创业，赋值为1，比例为10.55%，“雇员”和“自营劳动者”则赋值为0，比例为89.45%。



➤ 数据与变量

- 核心解释变量——农民工的互联网使用情况。
- 问卷向被访者询问了“您平时是否有上网的习惯”，其中回答“是”的农民工有39479人，占比为35.59%，回答“否”的农民工有71443人，占比为64.41%。CNNIC数据显示，截止到2012年底，全国互联网普及率为42.1%，而农村地区互联网普及率为24.2%，农民工使用互联网比例恰好处于全国平均水平与农村地区平均水平之间，与实际情况基本吻合。
- 问卷还询问了被访者每周的互联网使用频率文中将被访者回答的小时数分成了10小时以下、10-20小时、20-30小时以及30小时以上四类。

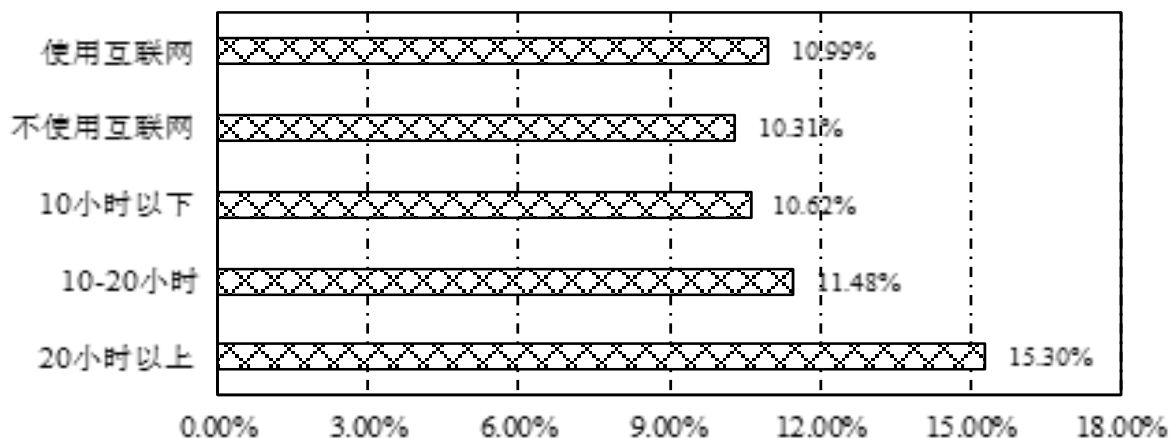


图 3 互联网使用与农民工的创业率



表1 描述性统计结果

变量名称	变量描述	观测值	平均值	标准差
被解释变量				
创业	创业=1, 非创业=0	110922	0.106	0.307
核心解释变量				
互联网使用	使用=1, 不使用=0	110922	0.356	0.479
控制变量				
年龄	受访者 2012 年的年龄	110922	33.855	9.131
年龄平方	年龄*年龄/100	110922	12.295	6.425
性别	男性=1, 女性=0	110922	0.600	0.490
教育程度	大学以下学历=0, 大学及以上学历=1	110922	0.045	0.207
未婚为参照				
已婚	--	110922	0.765	0.424
离婚或丧偶	--	110922	0.018	0.132
收入水平	收入水平的自然对数形式	110383	7.826	0.558
周工作时长	--	110922	60.606	16.860
社会保障	城镇职工医疗保险, 无=0, 有=1	110922	0.176	0.381
住房状况	有产权=1, 无产权=0	110922	0.105	0.307
居留时间	本次流入目前居住地时间	110922	4.389	4.667
迁移范围	跨省迁移=1, 省内迁移=0	110922	0.578	0.494
迁移模式	家庭化=1, 孤身外出=0	110922	0.802	0.399
吸烟情况	吸烟=1, 不吸烟=0	110922	0.320	0.466



➤ 实证结果分析

- 通过比较我们可以看出，互联网使用能够在1%水平上显著提高农民工的创业倾向，在加入了控制变量以及地区虚拟变量后，结论依然成立。
- 根据第（3）列结果来看，相比于不使用互联网的农民工，使用互联网的农民工创业倾向显著提升2.9%。
- 第（4）-（6）列检验了互联网使用频率对农民工创业倾向的影响，回归的步骤与第（1）-（3）列一致。结果显示，互联网使用频率也能显著提升农民工的创业倾向，结果在1%水平显著。这一结果进一步验证了互联网使用对农民工创业行为的正向影响。

表 2 互联网对农民工创业行为的影响（Probit 模型边际效应）

变量	(1)	(2)	(3)	(4)	(5)	(6)
互联网使用（使用=1）	0.007***	0.031***	0.029***			
	(3.51)	(15.07)	(13.86)			
互联网使用频率				0.001***	0.001***	0.001***
				(6.70)	(5.83)	(5.81)



➤ 实证结果分析

- 考虑到农民工使用互联网很可能是“自我选择”的结果，本文通过倾向得分匹配法构建反事实框架来纠正可能的选择性偏误。
- 文中首先进行倾向值匹配，通过一系列影响农民工创业的因素建立Probit模型，然后估计出农民工创业的倾向值，并在此基础上进行匹配。图4展示了匹配前后处理组与控制组倾向值的核密度分布情况。

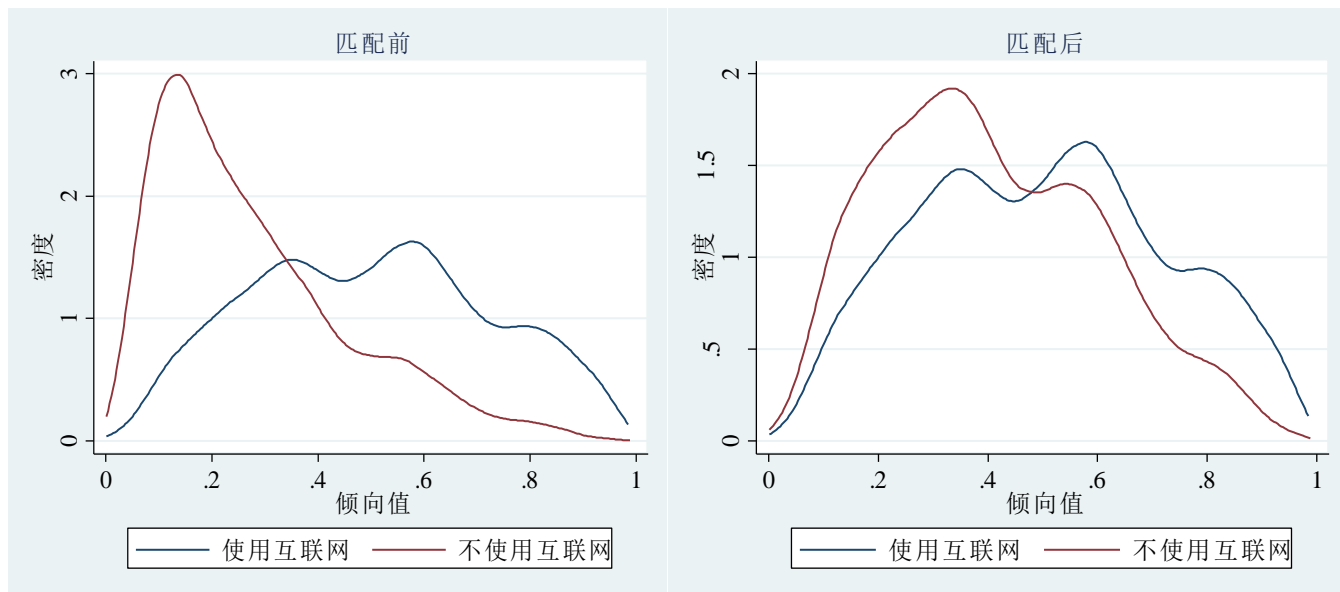


图 4 匹配前后的倾向值分布

➤ 实证结果分析

- ◆ 倾向得分匹配法能够纠正可观测变量引致的选择性偏误。然而，本文还可能面临遗漏变量、反向因果等潜在的内生性问题。
- ◆ 首先，一些难以衡量的遗漏变量（如农民工的个人经历、性格特征等）可能同时影响农民工互联网技能的掌握与使用、农民工的创业决策，从而产生遗漏变量问题。其次，某些农民工可能是因为创业行为才引致了互联网技能的需求，由此便产生了反向因果问题。
- ◆ 上述可能出现的内生性问题，将会导致估计系数不一致（Angrist and Pischke, 2010），这类问题常用的解决方法是工具变量法，但是互联网和创业决策都是离散型变量，基于连续变量的两阶段回归等工具变量方法不再有效（Angrist, 2001; Sajaia, 2008）。因此，本文利用工具变量Probit模型展开回归分析（周广肃、李力行，2016）。
- ◆ 为此，本文选用“各省份互联网域名数量的人均水平”（简称“人均域名数”）、“互联网宽带用户数量的人均水平”（简称“宽带用户比”）作为工具变量。



► 实证结果分析

- ◆ 本文首先利用传统的两阶段最小二乘法（2SLS）检验工具变量的合理性，然后利用Ivprobit模型展开进一步分析。
- ◆ Cragg-Donald Wald F统计量为45.13，远大于10%偏误下的临界值19.93，即拒绝弱工具变量的假设，可以排除弱工具变量问题（Stock and Yogo, 2002）。进一步，我们检验工具变量是否满足外生性，即工具变量是否满足过度识别检验条件。Sargan检验统计量为0.011，对应的p值仅为0.915，远大于0.1，说明我们无法拒绝工具变量符合外生性条件的原假设。再进一步，我们进行DWH检验，以判断模型是否存在内生变量。DWH检验的p值为0.000，结果在1%水平显著，说明互联网使用为内生变量。经过上述检验，我们可以确认工具变量法能够准确地估计互联网使用对农民工创业行为的影响。

表4 Ivprobit 模型分析

变量	工具变量： 人均域名数	工具变量： 宽带用户比	工具变量： 人均域名数+宽带用户比
	(1)	(2)	(3)
互联网使用（使用=1）	2.178***	1.840**	2.207***
	(4.10)	(2.26)	(4.31)



➤ 实证结果分析

- ◆ 上文分析已经证实，使用互联网会显著提升农民工的创业倾向。不过，上述结果只是平均效应，并未考虑农民工群体的内部差异。接下来，我们重点分析互联网对新生代农民工、老一代农民工创业行为的异质性。按照惯例，出生于1980年以前为老一代农民工，出生于1980年及以后为新生代农民工。
- ◆ 老一代农民工的回归系数要明显高于新生代农民工的回归系数，这意味着，使用互联网对老一代农民工创业倾向的提升作用要更加明显。这其中可能的逻辑在于，一方面对于广大“80后”新生代农民工而言，互联网似乎早已司空见惯，而相比之下，互联网则很可能给老一代农民工带来的新鲜感与刺激性更强；另一方面，新生代农民工使用互联网注重娱乐性，而老一代农民工则更为注重知识性。

表5 新老农民工

变量	Probit模型		Ivprobit模型	
	新生代	老一代	新生代	老一代
	(1)	(2)	(3)	(4)
互联网使用（使用=1）	0.115***	0.225***	2.110***	2.604***
	(5.95)	(13.22)	(3.87)	(3.00)



► 实证结果分析

- 上文考察了互联网使用对农民工创业倾向的影响，并进一步分析了使用互联网对农民工创业影响的代际异质性，但并未考虑互联网使用对农民工创业“绩效”的影响。为此，本文参考周广肃、李力行（2016）的研究，利用Heckman选择模型估计互联网使用对农民工收入水平的影响。
- 本文同时运用传统的两步法和效率更高的MLE方法估计Heckman选择模型。我们观察到两步法中的逆米尔斯比率在1%水平上显著，MLE法中的 α 也在1%水平上显著，这意味着创业行为是内生虚拟变量，也就意味着Heckman选择模型要优于普通的OLS估计。
- 进一步来看，无论是两步法还是MLE法，第一阶段的回归结果都显示互联网使用会显著提升农民工的创业倾向，第二阶段的回归结果则显示，互联网使用能够显著提高农民工的收入水平。

表 6 互联网对农民工收入的影响

变量	Twostep		MLE	
	创业	收入	创业	收入
	(1)	(2)	(3)	(4)
互联网使用（使用=1）	0.150***	0.306***	0.149***	0.143***
	(12.03)	(5.80)	(11.98)	(11.54)



➤ 实证结果分析

- 上述分析表明，互联网使用对农民工的创业行为具有显著的正向效应。那么，为什么互联网能够促进农民工创业呢？或者说互联网对农民工创业行为的作用机制应该如何解释呢？
- 从逻辑上来讲，互联网作为信息沟通的重要载体，有助于增强人与人之间的联系，如当前大行其道的微信、微博等。这样一来，农民工能够借助互联网维持并强化个人的社会网络、进一步丰富社会关系网，进而获得掌握更多的创业信息、创业支持等，而且，已有文献已经证实了社会网络对创业行为的正向影响（Zhang and Zhao, 2015）。
- 为此，本文将社会网络作为互联网使用与农民工创业行为的中介变量，通过构建中介效应模型，探讨互联网是否通过丰富社会网络提升农民工的创业倾向。
- 问卷询问了“您业余时间在本地和谁来往最多”，被访者回答包括“同乡（户口在本地）、同乡（户口在老家）、其他本地人、其他外地人、很少与人来往”等五个选项，由于本文关注的是务工地社会网络，即农民工与本地人的交往是否融洽，故将来往最多的是“其他本地人”定义为1，其余选项则定义为0。



► 实证结果分析

- 上述分析已经表明，掌握互联网技能对农村居民主观幸福感有显著正向影响。那么，创业行为是否通过自主效应、收入效应和地位效应影响农民工的幸福感呢？
- 如表7所示，依照中介效应模型的检验顺序，第（1）列检验的是互联网使用对农民工创业行为的影响；第（2）列检验互联网使用对中介变量的影响；第（3）列检验的是互联网使用、中介变量同时纳入模型对农民工创业行为的影响。
- 温忠麟、叶宝娟（2014）提出，论证中介效应的存在性还需进行Sobel检验（Sobel, 1987）。本文利用Bootstrap自抽样法迭代500次得出Sobel检验值（Z统计量）为1.812，结果在统计层面具有显著性。
- 由此可得，社会网络是互联网使用影响农民工创业行为的渠道之一。

表 7 影响机制分析

变量	被解释变量：创业	被解释变量：社会网络	被解释变量：创业
	(1)	(2)	(3)
互联网使用（使用=1）	0.179***	0.153***	0.173***
	(13.85)	(16.04)	(13.38)
社会网络			0.110***
			(9.08)



► 研究结论

- 本文结合**2012年度流动人口动态监测调查数据**，系统考察了使用互联网对农民工创业倾向的影响。研究表明：
 - ◆ 第一，使用互联网能够显著提升农民工的创业倾向，这一结果在加入控制变量、地区虚拟变量后依然成立。
 - ◆ 第二，考虑到掌握互联网技能很可能是农民工“自我选择”的结果，结合倾向得分匹配法构造反事实框架纠正选择性偏误，所得结果仍保持一致。进一步，考虑到模型中遗漏变量、互为因果等潜在的内生性问题，本文选取“人均域名数”和“宽带用户比”为工具变量，利用**lvprobit**模型重新估计，最终所得结果依然稳健。
 - ◆ 第三，异质性分析结果显示，使用互联网对老一代农民工创业的影响明显高于对新生代农民工的影响。
 - ◆ 第四，利用**Heckman**选择模型的扩展性分析显示，使用互联网还能够显著提高农民工的收入水平。
 - ◆ 第五，基于中介效应模型的影响机制分析显示，使用互联网能够通过丰富社会网络提升农民工的创业倾向。



➤ 政策启示

- ◆ 本文的研究结论具有重要的政策启示。
- ◆ 总体上来说，要将互联网作为推动农民工创业的重要“抓手”，紧密结合“互联网+”的重大时代背景，着力提升互联网的普及率，为农民工提供用得上、用得起、用得好的信息服务，力争使互联网成为农民工创业的重要工具，让农民工能够共享互联网时代的发展成果，拥有更多的获得感与满足感。具体而言：
 - ◆ 一方面，要加快补齐互联网基础设施“短板”、努力填平地区间、城乡间的“数字鸿沟”，尤其要重视发展移动互联网，大力投入移动基站建设，加快新通信技术在农民工群体中的推广与普及。
 - ◆ 另一方面，要结合“互联网+”的时代背景，加强互联网知识、互联网观念的宣传引导，可通过组织多种形式的培训，让农民工、尤其是老一代农民工了解互联网、掌握互联网技能，以推进互联网带来的“创业效应”。



➤ 研究不足之处

- ◆ 最后需要指出的是，尽管本文采用了国家卫计委组织的全国性大样本调研数据，并运用多种方法进行实证分析，但依然存在一些缺陷。
- 其一，由于流动人口动态监测数据为截面数据，本文仅能观测到农民工在某一时刻的静态信息，无法捕捉农民工使用互联网以及创业的动态过程，而这很可能导致估计结果出现偏差。
- 其二，本文通过倾向得分匹配法纠正了可能的选择性偏误，并通过工具变量法得到了相对一致的研究结论，但本文所选工具变量仅是省级层面变量，不够完美。因此对于本文所得结论仍需保持谨慎态度。
- 其三，囿于数据，本文仅从社会网络层面识别互联网对农民工创业行为的作用机制。
- 总而言之，从严格意义来讲，本文并未得出互联网使用与农民工创业严谨的因果推断意义，但对于相关研究仍然具有比较重要的参考价值。



谢谢！ 恳请各位专家
评委批评指正！

电子邮箱：zzkandqq@ruc.edu.cn

联系电话：15910365152

