

互联网使用与用途如何影响农村居民工资性收入差距

吴彬彬 沈扬扬 卢云鹤 滕阳川*

内容提要 本文利用2014年和2016年中国家庭追踪调查(CFPS)数据,考察了在信息时代“数字鸿沟”的背景下,农村居民互联网的使用与用途选择对其整体工资收入不平等的影响。研究发现,首先,工具变量回归结果显示,农村居民使用互联网具有显著的工资溢价效应。其次,平均对数离差、泰尔指数的群组分解结果显示,将互联网用于工作用途正逐渐成为加深整体农村居民工资不平等程度的潜在推动力。再次,在控制人口学特征、家庭特征和地区特征的情形下,基于基尼系数、平均对数离差和泰尔指数的再中心化影响函数回归结果显示,将互联网用于工作用途能够推动农村网民工资不平等程度逐年增加。最后,进一步剔除了人口特征对于互联网用途选择的影响,基于再中心化影响函数回归的分解发现,2014-2016年农村网民整体工资不平等程度趋于增加,其中互联网用途差异的贡献度在5.37%~21.94%之间。

关键词 互联网 收入差距 群组分解

一 引言

中国正处于互联网发展的特殊和关键时期。一方面,全国互联网普及率迅速提高,根据第43次中国互联网络信息中心(China Internet Network Information Center,简称

* 吴彬彬,扬州大学商学院、北京师范大学中国收入分配研究院,电子邮箱:wbbbiabia@163.com;沈扬扬(通讯作者),北京师范大学经济与资源管理研究院,电子邮箱:18811782104@163.com;卢云鹤,山东财经大学财政税务学院,电子邮箱:yunhelu021@163.com;滕阳川,北京师范大学经济与工商管理学院,电子邮箱:tengyangchuan@foxmail.com。本研究得到国家社会科学基金一般项目“社会保障权利平等促进农民工收入分配平等的机制和效果研究”(项目编号:18BJY045)的资助。

CNNIC) 报告显示,截至 2018 年 12 月,中国网民规模达 8.29 亿,互联网普及率为 59.6%;另一方面,互联网在区域、城乡和不同人群之间的差距程度迅速扩大,其中重要原因在于落后地区和贫困人口在互联网接入上存在门槛效应。根据 CNNIC 数据,互联网普及率的城乡差距从 2013 年的 32.2% 上升到 2018 年的 36.2%。在增长和差距并存的时期,考察中国当期互联网使用对居民工资的影响,是一个重要的研究课题。早在 2006 年,中共中央办公厅和国务院办公厅印发的《2006-2020 年国家信息化发展战略》便强调“坚持政府主导、社会参与,缩小区域之间、城乡之间和不同社会群体之间信息技术应用水平的差距,创造机会均等、协调发展的社会环境”。在上述背景之下,本文进一步将目标群体锁定在农村居民,其背后的政策含义是考察互联网使用对农村居民工资性收入的影响,以观测农村互联网的发展效率,并分析在关键时期政策支持农村互联网建设的必要性。

新技术带来了工作效率的提升以及工资的提高。Krueger (1993) 利用 1984-1989 年当前人口调查 (Current Population Survey, 简称 CPS) 跨期微观数据发现,同等工作类型的劳动者,平时使用电脑的人员的工资水平要明显高于非使用者 10~15 个百分点; Jensen (2007) 针对印度南部捕鱼产业工作者的研究表明,使用手机对促进市场发展以及增进福利方面具有正向效果; Muto & Yamano (2009) 对 2005-2009 年乌干达的香蕉销售额与手机使用的关系进行研究也发现了类似结果。可以说,上述研究本质上指出了电脑、手机等新技术的工具性价值,以及善用这种工具价值所能够带来的效率提升和收益提升。当前阶段,手机和电脑的工具性价值随着普及度的提高而有所削弱;相对地,互联网的产生和发展形成了新的工具性价值。

信息时代,“数字鸿沟”、“数字贫困”问题开始引发全球关注。“数字鸿沟”概念源起美国国家电信和信息管理局 (NTIA) 连续发布的《在网络中落伍》报告 (张新红, 2008)。“数字鸿沟”意指互联网接入、技能和使用方面的差距;“数字贫困”则指向“数字鸿沟”劣势一端的状态。上世纪 80 年代,阿马蒂亚·森 (2002) 提出“可行能力”理论,指出所谓贫困是人们一系列基本功能 (如免受饥饿、接受教育、参与社区活动) 的丧失。对应可行贫困理论,“数字贫困”是指获取知识与信息能力和机会的丧失。由此可见,“数字贫困”是经典能力贫困理论的一部分。

随后,学者将研究视角延伸到互联网和信息技术对生产率和收入水平的影响。其中,Attewell & Battle (1999) 讨论了家用电脑对学习提高的工具性价值,研究表明家庭电脑有利于学生成绩的提高,同时形成没有电脑和使用电脑学生成绩之间的差距。Lio & Liu (2006) 发现互联网有助于现代农业发展,且富裕国家的互联网获

益效率要高于贫穷国家。互联网的增收机制包括提供实时信息 (Aker, 2011; Cole & Fernando, 2012)、通过互联网销售提高议价能力 (Nakasone et al., 2014)、购物降低成本以及通过互联网提高销售量等 (Briggeman & Whitacre, 2010), 但由于穷国人力资本水平相对较低, 从而影响了上述机制的高效传递 (Dewan & Kraemer, 2000)。

当然, 上述发现主要集中于发展中国家, 或者基于互联网发展初期的发达国家经验性结论, 一部分原因在于互联网普及初期, 其较高的使用费用门槛、教育程度门槛、语言门槛、基础设施门槛, 导致贫困地区和贫困人群更容易在互联网接触方面被边缘化。伴随互联网普及率的提高, 一些新的研究表明互联网有助于缩小特定人群间的收入差距 (Banerjee et al., 2007; Postar, 2013)。那么, 在人类发展进程中应该如何看待互联网之于发展的作用? 更具体的问题是, 对中国的发展而言, 互联网应该扮演怎样的角色? 可以说, 互联网在拉动增收以及收入不平等方面所具有的作用也成为中国学者所讨论的焦点。

蒋琪等 (2018) 基于 2010 年和 2014 年两期中国家庭追踪调查面板数据, 发现互联网使用对不同居民带来 23.99% (年化为 5.52%) 的额外收入, 且对农村居民的增收效果更大。与之类似, 程名望和张家平 (2019) 利用 2015 年中国社会状况调查数据, 证实了现阶段互联网普及缩小城乡居民收入差距的直接原因在于, 微观层面上互联网使用对农村居民的收入效应要大于城镇居民。但同时, 刘晓倩和韩青 (2018) 也强调警惕互联网使用导致的农村居民内部收入差距过大的问题。2007 年和 2013 年, 在农村居民的所有收入细项中, 工资性收入对总体收入不均等的贡献比重都是最大的, 两年均超过了 40%; 且两年间工资性收入的不平等程度有所上升, 工资的分配比总收入的分配更不平等 (罗楚亮等, 2018)。

近年来农村互联网发展水平和覆盖率逐步提升, 但农村网民在性别、年龄、学历、职业结构和收入结构存在一定的差异性, 互联网使用对不同群体的作用具有异质性, 不同用途的互联网使用其影响效果也存在差异。比如, 从事服务业和制造业的人群使用电脑带来的工资回报率更加显著 (陈玉宇、吴玉立, 2008); 互联网对于非农就业的促进作用随劳动力教育水平的上升而增加 (马俊龙、宁光杰, 2017); 娱乐游戏会减少工资, 上网专业学习能够提高工资 (毛宇飞等, 2018)。基于此, 本文拟使用 2014 年和 2016 年中国家庭追踪调查数据, 考察其对农村居民的工资性收入差距变动的影响。本文的贡献主要在以下几个方面: 第一, 不仅从均值意义上度量了农村居民中各类互联网使用群体之间工资水平的差异, 还进一步考虑群体间及其内部的分布差异; 第二, 利用平均对数离差 (MDL)、泰尔 (Theil) 指数的群组分

解，从定量层面上研究互联网使用以及互联网用途差异对农村居民工资不平等的贡献度；第三，在控制一系列影响收入水平以及上网用途选择的人口特征的情况下，选用前沿的再中心化影响函数（Recentered Influence Function，简称 RIF）方法，探究互联网用途的差异对农村网民工资不平等（基尼系数、平均对数离差、泰尔指数）的影响。

二 数据来源与描述性统计

（一）数据来源及指标选取

本文使用中国家庭追踪调查（China Family Panel Studies，简称 CFPS）数据。CFPS 由北京大学中国社会科学调查中心（Institute of Social Science Survey，简称 ISSS）实施，样本覆盖全国 25 个省（市、自治区），目标样本规模为 16000 户，调查对象包含样本家户中的全部家庭成员。本文研究对象为农村 16 岁至 65 岁受雇样本，核心变量为受雇样本的年工资性收入（下文简称“工资”）及其互联网使用情况。

首先，工资性收入包括所有工资、奖金、现金福利、实物补贴在内，并扣除税和五险一金。计算 2016 年的工资时，采用 2014 年的不变价格，按照国家统计局公布的农村分省份的消费价格指数进行平减调整。

其次，将 CFPS 中互联网的使用变量做二分类处理：电脑上网赋值为 1、不上网赋值为 0。类似地，将互联网的用途变量也做二分类处理：工作用途赋值为 1、其他用途赋值为 0。具体而言，第一，本文将研究重点放于使用电脑上网群体，故将仅用手机上网的人群统一归类为“不上网”，在 2014 年数据中，本文将有家庭电脑且上网的群体定义为“电脑上网”。第二，CFPS 数据统计了不同互联网用途的使用频率。为了保证研究群体是真正意义上的互联网用户，本文仅将“几乎每天”、“一周 3~4 次”进行互联网工作的定义为“工作用途”。即将那些虽然将互联网用于工作用途，但频率较低的，视为“其他用途”。

最后，控制村内基础设施建设情况，仅考察村内居民电脑上网率大于零的样本。基于分析工资目的，只保留了工资大于零的样本。为了使结果更加精准，剔除了个人工资过低或过高的样本，即删除工资最低和最高的 0.1% 样本。最终形成两年跨期横截面样本，其中 2014 年、2016 年的样本量分别为 2500 人和 2923 人。

（二）互联网使用对农村居民工资的影响

作为研究起点，本文首先定量分析互联网使用对农村居民工资水平的影响。此处，

构建 Mincer (1974) 工资方程：

$$\ln w_i = net_i \beta^0 + X_i \beta^1 + u_i \quad (1)$$

式 (1) 中, $\ln w_i$ 表示农村居民年工资对数, net_i 是互联网使用变量 (电脑上网 = 1、不上网 = 0)。 X_i 为控制的人口学特征 (年龄、性别、婚否、民族)、人力资本特征 (受教育程度)、就业特征 (行业、职业、所有制)^①、家庭特征 (家庭规模、家庭有 65 岁以上老人) 和地区特征 (东、中、西、东北部地区)。 β^0 和 β^1 为对应的影响系数, u_i 为误差项。变量的描述性统计见附表 1。

对式 (1) 进行估计可以验证互联网使用是否存在工资溢价效应, 考虑到互联网使用与工资可能呈反向因果关系, 本文使用两阶段最小二乘法 (2SLS) 对工资方程进行回归。在解决内生性问题过程中, 参考已有文献 (毛宇飞等, 2018; Agarwal et al., 2009), 本文选取个体所在县级互联网普及率作为工具变量进行内生性检验。工具变量 (IV) 选取的内在机理是: 一方面, 互联网普及率反映地区网络基础设施状况, 是决定个体上网的必要条件, 但与其他决定个体收入的变量没有直接关系, 符合 IV 的相关性假设; 另一方面, 互联网普及率不会直接对个体的工资产生影响, 符合 IV 的外生性假设。

式 (1) 的估计结果如表 1 所示, 其中 2014 年和 2016 年的第一阶段模型结果显示, 互联网普及率系数均显著为正, 表明随着周围居民互联网普及率的提高, 个人上网的概率也会增加。同时, 两年的第一阶段回归 F 值均大于 10, 即排除了存在弱工具变量的可能。从 2014 年和 2016 年第二阶段的结果来看, 互联网对工资的影响系数显著为正, 互联网普及率在一定程度上反映了该地区的经济状况, 可能会对个人的就业机会产生影响, 互联网使用可以产生工资溢价效应。

如表 1 所示, 控制一系列人口特征之后, 相较于不上网的农村居民, 使用互联网的农村居民工资收入更高。上述回归结果形成本文基本分析前提: 互联网使用能产生工资溢价效应, 使用与不使用互联网的农村居民群体之间将产生工资不平等现象, 甚至使用互联网的农村居民群体内部, 也会由于互联网用途的不同, 产生工资不平等现象。但不同群体之间的工资不平等程度有多大, 互联网使用在其中的贡献为多少, 尚待后文进一步分析。

① 根据职业声望 (李强, 2012), 将 7 种职业分为白领、蓝领和服务业三大类。其中白领包括国家机关党群组织、企事业单位负责人、专业技术人员、办事人员和有关人员, 蓝领包括生产、运输设备操作及有关人员。根据陈斌开和陈思宇 (2018) 的分类, 将行业中的服务业分为高端、低端、其他服务业。其中高端服务业包括金融业、技术研究行业、教育业和传媒业, 低端服务业包括交通运输业、零售业、餐饮住宿业和租赁服务业。

表 1 互联网使用对农村居民工资的影响

年份	2014 年		2016 年	
	第一阶段	第二阶段	第一阶段	第二阶段
互联网使用	—	0.3984 *	—	0.8801 ***
县级互联网普及率	1.0991 ***	—	0.6764 ***	—
控制人口特征	是	是	是	是
R ²	0.3039	0.1121	0.3351	0.0318
第一阶段 F	96.8801 ***	—	81.0827 ***	—
Wald χ^2	—	1144.4814 ***	—	5137.1399 ***
观测值	2500	2500	2923	2923

注：***、**、* 表示估计结果在 1%、5%、10% 的水平上显著；采用省级层面的聚类标准误。
资料来源：根据 2014 年和 2016 年 CFPS 数据计算得到。

(三) 互联网与农村居民工资差距

在互联网使用存在工资溢价效应的前提下，2014 年和 2016 年农村居民的上网比例及工资变动如表 2 所示。若以是否使用互联网将农村居民分为两类，使用互联网的农村居民工资的跨期值和增速均更高，该群体内的工资收入不平等程度也更高。比如，2014 - 2016 年期间上网的农村居民工资增长率达到 14.12%，而不上网的农村居民工资增长率只有 9.10%。

表 2 农村互联网使用与个人工资差距

年份	2014 年			2016 年		
	上网	不上网	整体	上网	不上网	整体
互联网使用						
年工资(万元)	3.1541	2.4505	2.6388	3.5993	2.6734	2.9557
基尼系数	0.3683	0.3563	0.3658	0.3850	0.3778	0.3882
平均对数离差	0.3085	0.2826	0.2960	0.3164	0.2983	0.3135
泰尔指数	0.2344	0.2191	0.2307	0.2517	0.2428	0.2562
样本量	669	1831	2500	891	2032	2923

资料来源：根据 2014 年和 2016 年 CFPS 数据计算得到。

表 3 进一步按照互联网的用途细分了农村网民群体，即划分为用于工作用途和其他用途的农村网民，并测算了与表 2 类似的指标。由表 3 可知，用于工作用途的农村

网民工资更高，且不平分化程度不断扩大。结合表 2 和表 3，说明在信息时代下的“数字鸿沟”极有可能造成贫富差距的进一步拉大，值得引起重视。为令本文基础论点更具有说服力，下文做进一步定量分析。

表 3 农村互联网用途差异与个人工资差距

年份	2014 年		2016 年	
	工作用途	其他用途	工作用途	其他用途
互联网用途				
年工资(万元)	3.6661	2.8506	4.0076	3.1675
基尼系数	0.3618	0.3630	0.3907	0.3662
平均对数离差	0.2930	0.3056	0.3334	0.2843
泰尔指数	0.2290	0.2251	0.2594	0.2255
样本量	249	420	458	433

资料来源：根据 2014 年和 2016 年 CFPS 数据计算得到。

三 互联网使用及其用途差异对农村居民工资差距的贡献

表 2 和表 3 从均值意义上度量了互联网使用及其用途差异对农村居民工资的影响。本文认为应进一步考察不同类型的农村居民群体之间及其内部的工资分布差异，这样才能更有效从定量层面上研究互联网使用与否，以及互联网用途差异对工资不平等的贡献度。

在讨论互联网使用与否及其用途差异对工资收入差距的贡献时，本文选用平均对数离差（MLD）、泰尔（Theil）指数度量工资不平等程度。本文目的在于分析互联网使用及其用途差异对农村居民工资差距的贡献，首先关心的是互联网使用对农村居民工资差距的贡献（Shorrocks, 1980）：

$$I(y) = \sum_{r=0}^1 s_r I_r + I(\mu_1, \mu_0) \quad (2)$$

式（2）中， r 代表上网（ $r=1$ ）和不上网（ $r=0$ ）两类农村居民。 I_r 是收入不平等指数， s_r 分别代表两类群体人口份额（适用于平均对数离差）或收入份额（适用于泰尔指数）， μ_1 和 μ_0 分别为上网和不上网的农村居民的平均收入水平。两类群体内部的收入不平等指数的加权之和称为内部不平等（即“组内差距”）， $I(\mu_1, \mu_0)$ 为群体之间收入不平等（即“组间差距”）。

表 4 的整体工资不平等部分，给出了是否使用互联网的两类农村居民群体的组间和组内收入差距的分解结果，其中可以看到存在两个明显的趋势：一是互联网使用与否的组间差距不断提高；二是农村网民群体内部工资不平等程度跨期递增。上述结论说明上网与不上网的农村居民之间工资离散程度提高，而前文的表 2 显示使用互联网的农村居民工资均值更高且工资增速更快。此外，在使用互联网的农村居民中，位于收入分布高端一侧的群体的工资性收入提速更快，或者更多高工资的不使用互联网群体进入到了互联网使用群体当中。

本文进一步结合上网用途，将农村网民分为将互联网用于工作、其他用途两类群体，研究二者之间收入差距。参照 Gustafsson & Shi (2002) 的方法，得到上网群体内部的不平等分解式：

$$I_1(y) = \sum_{j=0}^1 s_{1j} I_{1j} + I(\mu_{11}, \mu_{10}) \quad (3)$$

式 (3) 中， j 表示上网用于工作用途 ($j=1$)、其他用途 ($j=0$)， s_{1j} 是上网人群中将网络用于不同用途的人口份额 (适用于平均对数离差) 或收入份额 (适用于泰尔指数)。式 (3) 的计算结果见表 4 的网民工资不平等部分。

由表 4 可知，2014 - 2016 年期间将互联网用于工作用途的农村居民内部工资差距不断拉大，且在 2016 年，其对农村网民群体内部整体不平等的贡献度超越了其他用途的群体。同时，可以得到与上文类似的推论：相比于将互联网用于其他用途的农村居民，将互联网用于工作用途的农村居民，其位于收入分布高端一侧群体的工资性收入提速更快，或者更多高工资农村居民进入到了互联网工作群体当中。上述结论说明，将互联网用于工作用途，正逐渐成为拉大整体农村居民之间工资不平等程度的潜在推动力，这一趋势值得重视。

表 4 农村居民工资不平等的分解 (互联网使用与否及其用途差异)

年份		平均对数离差		泰尔指数	
		2014 年	2016 年	2014 年	2016 年
整体工资不平等	是否上网之间	0.0065	0.0097	0.0067	0.0100
	(%)	2.19	3.10	2.91	3.92
	上网内部	0.0825	0.0965	0.0750	0.0934
	(%)	27.89	30.76	32.50	36.48
	不上网内部	0.2070	0.2074	0.1490	0.1527
	(%)	69.92	66.14	64.59	59.60

续表

年份		平均对数离差		泰尔指数	
		2014 年	2016 年	2014 年	2016 年
网民工资不平等	是否用于工作用途之间	0.0075	0.0069	0.0077	0.0068
	(%)	2.44	2.17	3.27	2.72
	工作用途内部	0.1091	0.1714	0.0991	0.1485
	(%)	35.35	54.17	42.26	58.97
	其他用途内部	0.1919	0.1382	0.1277	0.0964
	(%)	62.20	43.66	54.47	38.31

资料来源：根据 2014 年和 2016 年 CFPS 数据计算得到。

四 互联网用途对农村网民工资不平等的影响

前文得到的一个重要结论是，将互联网用于工作用途，正逐渐成为拉大整体农村居民之间工资不平等程度的潜在推动力。将互联网用于工作用途，也是信息技术时代的必然趋势。那么，在当前中国农村，将互联网用于工作用途究竟在工资不平的变动中具有多大的贡献？上文所使用的简单分解并无助于定量得到这一净影响。其原因在于，两个群组间很多拥有共性特征的人群，其对工资性收入所产生的交叠特征削弱了互联网效应（是否用于工作用途之间的不平等贡献程度呈现略微下降趋势）。故需要借助其他技术手段来验证本文研究的最终目标，本文选用前沿的 RIF 回归分析方法。

（一）不平等指数与 RIF 回归

人口学特征、家庭特征和地区特征能够影响个体的工资水平，从而影响群体间工资差距。为观测互联网用途对于农村网民工资不平等的净影响，需要对上述特征变量进行控制。本部分的讨论基于工资收入 (y)、不平等指数 (I) 的 RIF 函数。 $RIF(y; I)$ 是收入水平 (y) 的非单调变换 (Firpo et al., 2009, 2018):

$$RIF(y; I) = I + IF(y; I) \quad (4)$$

式 (4) 中， $IF(y; I)$ 为 Hampel (1974) 提出的影响函数 (Influence Function, 简称 IF)，衡量一个收入特定值的微小变化对目标统计量 (不平等指数) 的影响，具有期望为零的性质，这样可以得到 $RIF(y; I)$ 的期望等于不平等指数 (I): $E(RIF$

$(y; I) = I$ ^①。假设 $RIF(y; I)$ 的期望是特征向量 X 的线性函数，则系数向量 β 可以通过 OLS 回归得到：

$$E(RIF(y; I) | X) = X'\beta \quad (5)$$

$$I = E(RIF(y; I)) = E_x[E(RIF(y; I) | X)] = E(X)'\beta \quad (6)$$

式 (6) 中，系数向量 β 表示在其他条件不变的情况下，对应特征均值的微小改变对于不平等指数的边际影响。Firpo et al. (2009, 2018) 利用影响函数得到了收入的各种分布统计量（如分位数、方差、基尼系数）的 RIF 估计值，Cowell & Flachaire (2002)、Gradín (2018) 进一步计算出了平均对数离差 (MDL)、泰尔 (Theil) 指数的 RIF 估计值。参考上述文献，本文基于基尼系数、平均对数离差以及泰尔指数三类不平等指数，测算出互联网工作用途对农村网民工资不平等的影响（表 5）。

表 5 互联网工作用途对农村网民工资不平等的影响

年份	2014 年			2016 年		
不平等指数	基尼系数	平均对数离差	泰尔指数	基尼系数	平均对数离差	泰尔指数
工作用途	0.0252	0.0226	0.0431	0.0350 *	0.0725 **	0.0469 *
控制人口特征	是	是	是	是	是	是

注：仅列出部分 RIF 回归的结果；***、**、* 表示估计结果在 1%、5%、10% 的水平上显著；采用省级层面的聚类标准误。

资料来源：根据 2014 年和 2016 年 CFPS 数据计算得到。

如表 5 所示，加入必要控制变量（人口学特征、家庭特征和地区特征），相较于将互联网用于其他用途的网民群体，工作用途的上网人群将使得农村网民整体工资不平等程度提高。以 2014 年基尼系数为例，平均而言，将互联网用于工作用途的农村居民使得农村网民整体工资不平等程度提高 0.0252，但未在 10% 的水平上通过显著性检验。2016 年，将互联网用于工作用途的农村居民使得农村网民整体工资不平等程度显著提高 0.0350（基尼系数）、0.0725（平均对数离差）、0.0469（泰尔指数），说明将互联网用于工作用途，农村网民工资不平等程度跨期提高。

（二）农村网民工资不平等变化的 Appleton 分解

本文将进一步探究互联网用途的比例变化对 2014 年和 2016 年两年不平等程度的具

① Firpo et al. (2018) 和 Gradín (2018) 详细介绍了 $RIF(y; I)$ 的估计方法，并证明了 RIF 估计的渐进正态性质。由于 $RIF(y; I)$ 的估计过程比较复杂，此处不对其估计方法进行具体讨论，有兴趣的读者可参阅 Firpo et al. (2018) 和 Gradín (2018)。

体贡献值。现有文献对不平等变化的分解方法主要有基于 RIF 回归的 Oaxaca-Blinder 分解 (Firpo et al., 2018; Gradín, 2018); 基于回归的 Shapley 值分解 (Wan, 2002; Wan & Zhou, 2005); 基于 RIF 回归的 Appleton 分解 (Appleton et al., 1999; 郭继强等, 2011)。本文考虑到互联网用途的选择也会受到人口学特征、就业特征、家庭特征和地区特征的影响, 这就类似于将“职业”作为收入方程的外生控制变量遇到的问题 (郭继强等, 2011)。基于此, 本文使用基于 RIF 回归的 Appleton 分解方法, 由此 2016 年和 2014 年农村网民工资不平等差异表示为:

$$\begin{aligned}
 I^{2016} - I^{2014} &= \sum_{j=0}^1 P_j^{2016} (\bar{X}_j^{2016'} \beta_j^{2016}) - \sum_{j=0}^1 P_j^{2014} (\bar{X}_j^{2014'} \beta_j^{2014}) \\
 &= \underbrace{\sum_{j=0}^1 P_j^* (\bar{X}_j^{2016'} - \bar{X}_j^{2014'}) \beta_j^*}_{\text{网络用途内的禀赋效应}} \\
 &\quad + \underbrace{\sum_{j=0}^1 P_j^* \bar{X}_j^{2016'} (\beta_j^{2016} - \beta_j^*) + \sum_{j=0}^1 P_j^* \bar{X}_j^{2014'} (\beta_j^* - \beta_j^{2014})}_{\text{网络用途内的系数效应}} \\
 &\quad + \underbrace{\sum_{j=0}^1 (P_j^{2016*} - P_j^*) (\bar{X}_j^{2016'} \beta_j^{2016}) + \sum_{j=0}^1 (P_j^* - P_j^{2014*}) (\bar{X}_j^{2014'} \beta_j^{2014})}_{\text{网络用途之间禀赋效应}} \\
 &\quad + \underbrace{\sum_{j=0}^1 (P_j^{2016} - P_j^{2016*}) (\bar{X}_j^{2016'} \beta_j^{2016}) + \sum_{j=0}^1 (P_j^{2014*} - P_j^{2014}) (\bar{X}_j^{2014'} \beta_j^{2014})}_{\text{网络用途之间系数效应}} \quad (7)
 \end{aligned}$$

式 (7) 中, j 分别代表网络的工作用途 ($j = 1$)、其他用途 ($j = 0$)。此处使用 Logit 模型, 在已有一系列人口特征的基础上, 进一步加入就业特征 (行业、职业、所有制) 控制变量, 估计 2014 - 2016 年互联网用途选择的平均边际效应 (工作用途 = 1、其他用途 = 0) (见附表 2)。 P_1^{2016} 代表 2016 年可观测样本中工作用途网民比例; P_1^* 代表 2014 和 2016 年混合样本中工作用途网民比例; P_1^{2016*} 代表根据 2014 和 2016 年混合样本 Logit 回归得到的互联网用途选择的系数, 反事实估计的 2016 年工作用途网民的比例; 其他 P 值的解释与之类似。 β_1^{2016} 代表 2016 年可观测样本中, 工作用途网民的 RIF 回归系数; β_1^* 代表 2014 和 2016 年混合样本的工作用途网民的 RIF 回归系数; 其他 β 的解释与之类似。

分解结果如表 6 所示。最终 2016 年与 2014 年农村网民工资不平等的差异 ($I^{2016} - I^{2014}$) 将被分为四部分: 网络用途内的禀赋效应 (RIF 回归方程中人口特征的均值差异)、网络用途内的系数效应 (RIF 回归方程中系数差异)、网络用途之间禀赋效应

(用途选择 Logit 回归方程中人口特征的均值差异)、网络用途之间系数效应(用途选择 Logit 回归方程中系数差异)。分解结果显示,在控制一系列影响收入水平以及上网用途选择的人口特征的情况下,互联网用途的选择差异(本文指网络用途之间系数效应)对 2014-2016 年农村网民工资不平等上升的贡献度在 5.37%~21.94% 之间。

表 6 农村网民工资不平等变化的分解

不平等指数	基尼系数		平均对数离差		泰尔指数	
	绝对值	%	绝对值	%	绝对值	%
分解贡献						
原始的 RIF 估计值差距	0.0167	100	0.0080	100	0.0173	100
网络用途内的禀赋效应	-0.0035	-20.93	-0.0103	-129.62	-0.0044	-25.70
网络用途内的系数效应	0.0183	109.13	0.0150	187.67	0.0187	108.35
网络用途之间禀赋效应	0.0011	6.42	0.0016	20.01	0.0017	9.96
网络用途之间系数效应	0.0009	5.37	0.0017	21.94	0.0013	7.39

资料来源:根据 2014 年和 2016 年 CFPS 数据计算得到。

五 结论与政策含义

本文利用 2014 年和 2016 年两期 CFPS 横截面数据,考察了信息时代“数字鸿沟”背景下,互联网使用与否及其用途差异对农村居民工资收入差距变动的影响。研究发现,互联网使用存在工资溢价效应,其对提高农村居民的工资性收入有正向影响;互联网使用,尤其是工作用途的互联网使用逐渐拉大了农村网民与非网民的工资性差距,以及农村网民内部个体的工资性差距;将互联网用于工作用途正逐渐成为拉大整体农村居民之间工资不平等程度的潜在推动力;基于基尼系数、平均对数离差和泰尔指数的 RIF 回归结果显示,在控制一系列影响工资水平的特征情况下,选择将互联网用于工作用途,显著扩大了农村网民的工资不平等;基于 RIF 回归的 Appleton 分解显示,2014-2016 年期间互联网网络用途的选择差异对农村网民工资不平等程度上升的贡献度在 5.37%~21.94% 之间。

总体来说,一方面,农村居民的网路使用及其用途选择之间的不平等对于总体工资不平等贡献比较低,主要是上网的人群份额(或收入份额)较低,但需要注意农村居民的网路使用之间不平等呈现扩大趋势。另一方面,随着上网的人群份额(或收入份额)升高(2014-2016 年已经揭示这个现象),在互联网发展初期,互联网用途选择之间系数效应为正,这可能会进一步扩大农村居民的工资不平等。

因此，在信息技术获取和使用方面，城乡之间的差异呈现扩大趋势，农村互联网普及率亟待提高。同时，要引导农村居民有效率地利用互联网，为提升其能力、提高经济水平创造条件。要缩小农村居民之间的“数字鸿沟”，缩小农村居民的工资性收入差距。具体看来：第一，应降低信息技术的经济和技术门槛，以发挥公共服务在互联网普及上的作用，如降低光纤入户费用，尤其对有需要的低收入农户予以优惠。第二，政府应引导性、协助性地帮助农户构建生产生活与互联网的利益链条，如搭建电商平台，为农户创造新型就业机会。第三，应发挥教育和培训的作用，为农户接入互联网提供智力和技术支持，如帮助劳动力打通更多就业渠道、利用互联网接受新技能、新知识，青少年学龄儿童也可以通过网络，接受更优质教育。

附录：

附表 1 人口特征的描述性统计

年份		2014 年			2016 年		
		工作用途	上网群体	整体	工作用途	上网群体	整体
人口学特征	年龄(岁)	30.05	30.52	37.26	29.36	30.94	37.85
	男性(%)	50.20	59.94	64.92	53.28	61.95	65.41
	已婚(%)	70.68	74.59	78.72	55.02	64.20	77.66
	汉族(%)	94.78	93.27	91.40	93.01	93.04	91.82
受教育程度 (%)	小学或以下	2.01	5.98	25.36	2.18	5.16	25.28
	初中	20.88	42.30	46.44	18.56	34.90	43.76
	高中或中专	29.32	26.16	17.44	23.80	24.35	17.72
	大专或以上	47.79	25.56	10.76	55.46	35.58	13.24
行业(%)	第二产业	35.34	50.52	51.08	30.35	41.98	46.94
	低端服务业	21.29	19.28	16.68	22.05	20.99	18.34
	高端服务业	25.70	13.30	7.64	26.86	17.28	8.72
	其他行业	17.67	16.89	24.60	20.74	19.75	26.00
所有制(%)	私营企业/个体工商户	56.22	61.29	62.40	57.86	62.85	64.39
	机关国有集体外资	36.55	28.10	20.68	36.03	29.07	20.05
	其他所有制	7.23	10.61	16.92	6.11	8.08	15.57
职业(%)	白领职业	59.04	32.74	19.36	62.45	41.19	21.76
	蓝领职业	17.67	43.95	59.64	12.45	34.01	53.16
	服务业及其他职业	23.29	23.32	21.00	25.11	24.80	25.08

续表

年份		2014 年			2016 年		
群体分类		工作用途	上网群体	整体	工作用途	上网群体	整体
家庭特征	家庭规模(人)	1.94	1.94	1.94	1.81	1.87	1.87
	家庭有 65 岁以上老人(%)	15.66	13.45	14.04	12.01	13.24	14.68
地区(%)	东	52.61	48.73	39.12	37.55	39.06	36.74
	中	20.88	25.26	26.48	25.98	27.61	24.94
	西	18.88	16.29	22.44	26.86	22.90	28.12
	东北	7.63	9.72	11.96	9.61	10.44	10.20
样本量		249	669	2500	458	891	2923

资料来源：根据 2014 年和 2016 年 CFPS 数据计算得到。

附表 2 互联网工作用途选择的 Logit 回归

年份		2014 年	2016 年
人口学特征	年龄	0.0134	0.0134
	年龄平方	-0.0002	-0.0003
	男性	0.0099	-0.0403 *
	已婚	0.0256	-0.0213
	汉族	0.0028	0.0498
受教育程度	初中	0.0385	0.0768
	高中	0.1850 ***	0.1844 ***
	大专及以上学历	0.2989 ***	0.3270 ***
所有制	机关国有集体外资	-0.0296	-0.0556
	其他所有制	-0.0856	-0.0623
职业	蓝领	-0.2781 ***	-0.3133 ***
	服务业及其他职业	-0.1202 ***	-0.1310 ***
行业	低端服务业	0.0381	0.0355
	高端服务业	0.1281 **	0.1387 ***
	其他行业	0.0178	0.0575
家庭特征	家庭规模	-0.0045	-0.0380 **
	家庭有 65 岁以上老人	0.0722	-0.0355
人口学特征	年龄	0.0134	0.0134
	年龄平方	-0.0002	-0.0003
	男性	0.0099	-0.0403 *
	已婚	0.0256	-0.0213
	汉族	0.0028	0.0498

续表

年份		2014 年	2016 年
地区	中部	-0.0642	-0.0256
	西部	-0.0737	0.0344
	东北部	-0.0954 ***	-0.0925 ***
观测值		669	891

注：表中结果为 Logit 的平均边际效应；***、**、* 表示估计结果在 1%、5%、10% 的水平上显著；采用省级层面的聚类标准误。

资料来源：根据 2014 年和 2016 年 CFPS 数据计算得到。

参考文献：

- 阿马蒂亚·森 (2002), 《以自由看待发展》, 任贻、于真译, 北京: 中国人民大学出版社。
- 陈斌开、陈思宇 (2018), 《流动的社会资本——传统宗族文化是否影响移民就业?》, 《经济研究》第 3 期, 第 35 - 49 页。
- 陈玉宇、吴玉立 (2008), 《信息化对劳动力市场的影响: 个人电脑使用回报率的估计》, 《经济学 (季刊)》第 4 期, 第 1149 - 1166 页。
- 程名望、张家平 (2019), 《互联网普及与城乡收入差距: 理论与实证》, 《中国农村经济》第 2 期, 第 19 - 41 页。
- 郭继强、姜俐、陆丽丽 (2011), 《工资差异分解方法述评》, 《经济学 (季刊)》第 2 期, 第 363 - 414 页。
- 蒋琪、王标悦、张辉、岳爱 (2018), 《互联网使用对中国居民个人收入的影响——基于 CFPS 面板数据的经验研究》, 《劳动经济研究》第 5 期, 第 121 - 143 页。
- 李强 (2012), 《农民工与中国社会分层 (第二版)》, 北京: 社会科学文献出版社。
- 刘晓倩、韩青 (2018), 《农村居民互联网使用对收入的影响及其机理——基于中国家庭追踪调查 (CFPS) 数据》, 《农业技术经济》第 9 期, 第 123 - 134 页。
- 罗楚亮、史泰丽、李实 (2018), 《中国收入不平等的总体状况 (2007 - 2013 年)》, 载于李实、岳希明、史泰丽、佐藤宏主编《中国收入分配格局的最新变化——中国

- 居民收入分配研究 V》，北京：中国财政经济出版社，第 31 - 79 页。
- 马俊龙、宁光杰 (2017)，《互联网与中国农村劳动力非农就业》，《财经科学》第 7 期，第 50 - 63 页。
- 毛宇飞、曾湘泉、胡文馨 (2018)，《互联网使用能否减小性别工资差距？——基于 CFPS 数据的经验分析》，《财经研究》第 7 期，第 33 - 45 页。
- 张新红 (2008)，《数字鸿沟测算方法比较》，《电子政务》第 11 期，第 16 - 23 页。
- Agarwal, Ritu, Animesh Animesh & Kislaya Prasad (2009). Research Note—Social Interactions and the “Digital Divide”: Explaining Regional Variations in Internet Use. *Information Systems Research*, 20 (2), 277 - 294.
- Aker, Jenny (2011). Dial “A” for Agriculture: A Review of Information and Communication Technologies for Agricultural Extension in Developing Countries. *Agricultural Economics*, 42 (6), 631 - 647.
- Appleton, Simon, John Hoddinott & Pramila Krishnan (1999). The Gender Wage Gap in Three African Countries. *Economic Development and Cultural Change*, 47 (2), 289 - 312.
- Attewell, Paul & Juan Battle (1999). Home Computers and School Performance. *The Information Society*, 15 (1), 1 - 10.
- Banerjee, Sarbani, Rama Parai & Amar Parai (2007). Computer Use and Wage Differentials: US and Foreign Born Male and Female Workers. *Applied Economics Letters*, 14 (6), 409 - 413.
- Briggeman, Brian & Brian Whitacre (2010). Farming and the Internet: Reasons for Non-use. *Agricultural and Resource Economics Review*, 39 (3), 571 - 584.
- Cole, Shawn & Asanga Fernando (2012). The Value of Advice: Evidence from Mobile Phone-based Agricultural Extension. *Harvard Business School Working Paper*, No. 13 - 047.
- Cowell, Frank & Emmanuel Flachaire (2002). Sensitivity of Inequality Measures to Extreme Values. *LSE STICERD Research Paper*, No. 60.
- Dewan, Sanjeev & Kenneth Kraemer (2000). Information Technology and Productivity: Evidence from Country-level Data. *Management Science*, 46 (4), 548 - 562.
- Firpo, Sergio, Nicole Fortin & Thomas Lemieux (2009). Unconditional Quantile Regressions. *Econometrica*, 77 (3), 953 - 973.
- Firpo, Sergio, Nicole Fortin & Thomas Lemieux (2018). Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics*, <https://doi.org/>

10. 3390/econometrics6020028.

- Gradín, Carlos (2018). Explaining Cross-state Earnings Inequality Differentials in India: An RIF Decomposition Approach. *WIDER Working Paper*, No. 2018/24.
- Gustafsson, Björn & Shi Li (2002). Income Inequality within and across Counties in Rural China 1988 and 1995. *Journal of Development Economics*, 69 (1), 179 – 204.
- Hampel, Frank (1974). The Influence Curve and Its Role in Robust Estimation. *Journal of the American Statistical Association*, 69 (346), 383 – 393.
- Jensen, Robert (2007). The Digital Divide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector. *The Quarterly Journal of Economics*, 122 (3), 879 – 924.
- Krueger, Alan (1993). How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984 – 1989. *The Quarterly Journal of Economics*, 108 (1), 33 – 60.
- Lio, Monchi & Meng-Chun Liu (2006). ICT and Agricultural Productivity: Evidence from Cross-country Data. *Agricultural Economics*, 34 (3), 221 – 228.
- Mincer, Jacob (1974). *Schooling, Experience, and Earnings*. New York: Columbia University Press.
- Muto, Megumi & Takashi Yamano (2009). The Impact of Mobile Phone Coverage Expansion on Market Participation: Panel Data Evidence from Uganda. *World Development*, 37 (12), 1887 – 1896.
- Nakasone, Eduardo, Maximo Torero & Bart Minten (2014). The Power of Information: The ICT Revolution in Agricultural Development. *Annual Review of Resource Economics*, 6 (1), 533 – 550.
- Postar, Dara (2013). The Effect of High-speed Internet Access on the Gender Wage Gap. Ph. D. diss., Georgetown University, pp. 35 – 53.
- Shorrocks, Anthony (1980). The Class of Additively Decomposable Inequality Measures. *Econometrica: Journal of the Econometric Society*, 48 (3), 613 – 625.
- Wan, Guanghua (2002). Regression-based Inequality Decomposition: Pitfalls and a Solution Procedure. *WIDER Discussion Papers*, No. 2002/101.
- Wan, Guanghua & Zhangyue Zhou (2005). Income Inequality in Rural China: Regression-based Decomposition Using Household Data. *Review of Development Economics*, 9 (1), 107 – 120.

How Does the Internet Usage Affect Wage Gap among Rural Residents?

Wu Binbin^{1,2}, Shen Yangyang³, Lu Yunhe⁴ & Teng Yangchuan⁵

(Business School, Yangzhou University¹;

China Institute for Income Distribution, Beijing Normal University²;

School of Economics and Resource Management, Beijing Normal University³;

School of Public Finance & Taxation, Shandong University of Finance and Economics⁴;

Business School, Beijing Normal University⁵)

Abstract: Using the China Family Panel Studies (CFPS) data from 2014 and 2016, this paper estimates the effect of internet usage on wage gap among rural residents in the context of the “Digital Divide”. The main results shows that, firstly, instrumental variable estimation shows that the internet usage of rural residents can create a significant wage premium for them. Secondly, through subgroup decomposition of the MDL and Theil indices, it is revealed that using the internet for work purposes is gradually becoming a potential driving force for the overall wage inequality of rural residents. Thirdly, RIF regression based on Gini, MDL, and Theil indices shows that using the internet for work purposes could enlarge the wage inequality year by year, after controlling demographic, family and regional characteristics. Finally, a further Appleton decomposition based on RIF regression finds that the overall wage inequality of rural residents increased from 2014 to 2016, and the contribution of the choice of internet usage is between 5.37 percent and 21.94 percent.

Keywords: internet, wage gap, subgroup decomposition

JEL Classification: J21, J31, O47

(责任编辑: 封永刚)