

信息与城市化

——基于跨国数据的实证研究

万广华 张 琰*

摘要 本文首次揭示近半个多世纪以来城市化进程在全球范围内加速的现象。我们构建一个两部门一般均衡模型,论证信息与城市化率的正相关关系。使用互联网使用率作为信息的代理变量,基于 109 个非经济合作组织国家或地区 1960—2013 年间的数据进行实证分析。结果显示:(1) 常规城市化决定因素不能解释城市化加速现象;(2) 作为信息技术代理变量的互联网使用率可以解释大部分的城市化加速;(3) 主要结论对于使用不同地区样本、不同控制变量以及不同代理变量都是稳健的。

关键词 城市化加速, 信息, 移民成本

DOI: 10.13821/j.cnki.ceq.2021.02.05

一、引言

在城市经济学或关于城市化的研究中,有一个未被发现或长期被忽视的现象,这就是城市化速度正在变得越来越快。图 1 展示了部分国家城市化率从约 10% 到约 50% 所花费的时间。那些城市化始于 16 世纪的先行者,比如葡萄牙、波兰、瑞士、瑞典、法国和英国,完成这一进程大约花费了 400 年。对于从 19 世纪中叶开始城市化的国家(如罗马尼亚、挪威、希腊和德国),时间缩短到了约 100 年。而对于较晚开始城市化的亚洲国家(如中国、印度尼西亚和泰国)以及非洲国家(如喀麦隆、科特迪瓦和利比里亚),所需时间仅仅为 60 年左右。需要强调的是,这一加速现象并非只存在于高速发展的经济体(如中国),在像多哥和安哥拉这样的经济最不发达国家,城市化加速也同样可见。另一个例子也可以用来说明城市化加速的现象,即伦敦经过了 130 年才从一个 100 万人口的小城市成长为 800 万人口的大城市,而曼谷和首尔完成同样的人口扩张分别仅用了 45 年和 25 年的时间(WHO, 2010)。

* 万广华, 复旦大学世界经济研究所; 张琰, 重庆工商大学经济学院。通信作者及地址: 张琰, 重庆市南岸区学府大道 19 号, 400067; 电话: 13817512257; E-mail: phd09010@grips.ac.jp。本文为国家自然科学基金项目(71833003、72073091)的成果之一, 作者感谢《经济学》(季刊)编辑及审稿人提出的富有建设性的意见与批评。文责自负。

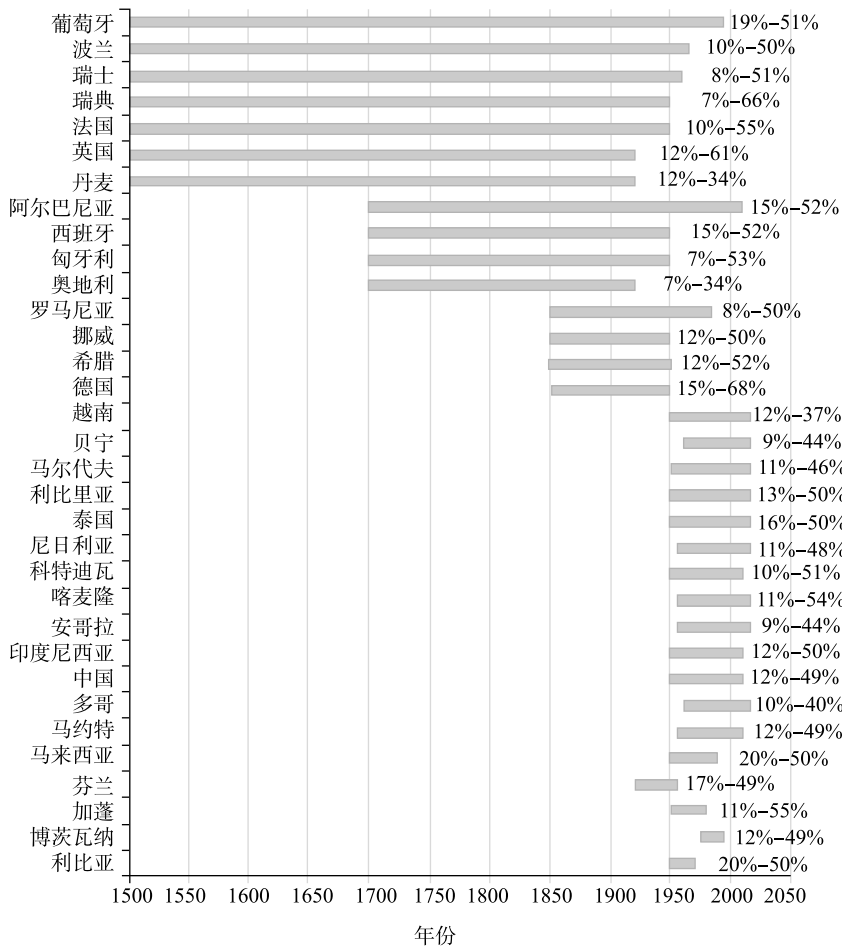


图1 部分国家城市化率从10%到50%所需年数

资料来源：作者根据两个数据来源整理。其一为Bairoch（1988）记载的2191个欧洲城市在公元800—1850年间的人口数，数据可以从<https://github.com/JakeRuss/bairoch-1988>获得，访问时间为2015年11月1日。根据Bairoch的数据加总得到各国城市人口，再除以OECD推测的各国人口数（来自OECD，2006，p. 241，表B-10）。其二为世界银行的WDI数据库，它提供了1950年及以后的城市化率。

考虑到城市化与很多因素相关，尤其可能由工业化推动，而工业化进程在工业革命后明显加快，我们将城市化水平与非农业增加值在GDP中的占比做回归，同时在回归方程中加入10年期（decadal）虚拟变量。结果发现，这些虚拟变量的系数高度显著为正，且越是更近的年份，其系数估算值越大。当然，除了工业化或经济结构变化，还有其他变量影响城市化，比如所谓的拉动（pulling）和推动（pushing）因素。鉴于此，我们在表1的第（2）列报告了控制所有常规变量的城市化面板模型估算结果（模型设定和相关讨论见后文）。显然，这时的十年期虚拟变量仍然显著，且系数还是不断增大。这些

都表明，随着时间的推移，城市化越来越快，而现有城市经济学理论或模型并不能很好地解释这一加速现象。

表1 十年期虚拟变量系数

十年期虚拟变量	(1)	(2)
1960s	—	—
1970s	4.22*** (0.000)	3.48*** (0.000)
1980s	10.41*** (0.000)	7.40*** (0.000)
1990s	15.42*** (0.000)	11.15*** (0.000)
2000s	18.72*** (0.000)	14.26*** (0.000)
2010s	21.32*** (0.000)	15.77*** (0.000)
R^2	0.97	0.98

注：表1的第(1)列中控制变量仅包含工业化程度，第(2)列对应表5基准模型结果的第(2)列，所不同的是这里的虚拟变量以十年为一个单位，而表4以每一年为单位。控制变量包括传统的“推拉”(push-pull)因素以及其他因素(参见文章的第四部分)。括号中为 p 值。

必须指出，忽视城市化加速可能会带来严重的后果，最直接的就是政府及相关部门准备不足，在面对蜂拥而来的大规模城乡移民时变得措手不及，无法提供公共服务，从而引起大城市病，诸如城市贫民窟、交通拥挤和城市污染等。在这种情况下，政府往往变得更加反城市化，千方百计阻挠乡城移民。这不但影响了经济增长和结构变迁，而且使大量隐性失业的农民滞留在资源匮乏的农村，无法解决贫困问题和居高不下的城乡差距(万广华，2008)。换言之，忽视城市化加速很可能导致了众多发展中国家公平和效率双输的严重后果(万广华，2013；罗知等，2018)。

因此，研究城市化加速现象并揭示其成因，不但具有理论和学术上的创新意义，而且具有重要和紧迫的政策与现实意义。基于此，本文构建了一个两部门一般均衡理论模型，论证了信息作为迁徙决策的重要影响因素，与城市化率呈正相关的关系。互联网的使用不但高效且廉价地提供了迁徙决策所需要的种种信息，还有助于大幅度降低迁徙成本和风险。所以，我们使用互联网使用率作为信息的代理变量，并基于109个非经济合作组织国家或地区1960—2013年间的数据对理论模型进行了实证检验。基准回归和工具变量模

型的结果均显示:(1)现有文献中的常规城市化决定因素不能解释这一加速现象;(2)互联网的使用率可以解释大部分的城市化加速;(3)论文的主要发现和结论对于使用不同地区的数据样本、不同的控制变量组合以及不同的代理变量都是稳健的。

本文后续内容安排如下:第二部分简要回顾相关文献并讨论了信息通信技术和互联网的发展背景;第三部分构建了一个简单的理论模型,并推导信息和城市化之间的关系;第四部分设定实证模型;第五部分讨论数据和实证模型估算结果;第六部分聚焦稳健性检验;第七部分总结全文,并提供政策建议。

二、文献综述和信息通信技术的发展背景

现有经典文献认为,城市化的动力主要来自农村的“供给推动”(pushing)和来自城市的“需求拉动”(pulling)(Marshall, 1890; Todaro, 1969)。在这种双重动力的作用下,只要城乡收入差为正,农村移民将源源不断地涌向城市。然而,正如前文所述,这些经典的推拉因素却并不能解释本文所发现的城市化加速现象。

研究迁徙的文献中,很多关注移民成本对移民决策的影响。Sjaastad (1962)首次在其研究中引入移民成本的概念,认为除了有形的货币成本以外,无形的机会成本和心理成本也是决定迁移与否的重要因素。其中,心理成本包括对新环境不熟悉所产生的种种不确定性、对亲朋好友的思念和在家乡积累的社会资本的流失等。一般来说,距离和移民成本正相关,因而和移民量负相关。远距离迁移不仅面临更高的交通费用,更由于信息流动相对较少而产生更高的心理成本。更为重要的是,进一步的研究发现,信息稀缺对移民决策的影响远远超过了交通成本(Greenwood, 1997)。换言之,“距离”的作用更多地体现在信息量方面,而不仅仅是以货币形式表现的交通费用。例如,Sahota (1968)估计了20世纪50年代巴西使用公路迁移330千米的交通成本,发现仅仅相当于一个单身工人1.5个月的工资,这笔费用并不会对生活产生重大影响。Vanderkamp (1971)对加拿大移民的研究以及Beals *et al.* (1967)对加纳移民的研究也得到了类似的结论。特别值得指出的是,Schwartz (1973)发现教育可以抵消距离对移民的负面作用,恰恰是因为高教育水平的人往往掌握更多的信息,这清晰地证明了信息对迁徙的决定性作用。

还有一些研究直接探讨了信息在移民决策中的作用。比如Banerjee (1984)论证了信息影响印度移民的重要性。他所提到的信息包括人际交往、报纸和以往的旅行经历。另外,移民网络可以传递信息,比如通过人脉关系获得就业、住房、教育资源等方面的信息。这些信息直接进入移民决策过程

(Massey *et al.*, 1993; Carrington *et al.*, 1996; Munshi, 2003)。Winters *et al.* (2001) 基于美国数据的研究发现, 目的地人脉网络密度, 对个人的移民决策有着重大影响。

需要强调的是, 以上所有研究都没有构建信息的量化指标, 也没有与城市化加速相关联。事实上, 这些文献没有考虑到信息会随着时间的变化而变化, 即假设了信息及其信息的作用是静态不变的, 所以不能用来考察我们所要研究的城市化加速现象。

另一方面, 信息和通信技术, 特别是在信息提供、端口普及以及信息传输等方面, 在近半个世纪发生了翻天覆地的变化。互联网的诞生革命性地改变了信息的传递方式与信息成本, 它使得人们无论何时何地都可以以极低的成本, 几乎同步地获得海量的信息, 在相当程度上消除了距离所带来的种种不确定性, 降低了人们移居前在就业、住房、生活环境、教育医疗、气候等方面面临的风险。脸书 (Facebook) 和照片墙 (Instagram) 这类社交网络的诞生几乎让“距离”消失了, 使得新移民在远离故土的时候, 可以继续与家人和朋友保持联系, 大大减轻了移民的情感负担和心理成本。

在本文第四部分的实证部分, 我们把互联网的使用作为信息的代理变量。需要说明的是, 我们将 1990 年作为互联网的起始年份来设计自然实验, 设 1990 年以前互联网使用率为 0。这是因为第一代应用 TCP/IP 协议的网络技术 (Advanced Research Projects Agency Network, ARPANET) 于 1990 年退役, 取而代之的是万维网 (World Wide Web, WWW) 的诞生, 而万维网正是互联网的前身。¹同时, 超文本文件、超文本标记语言 (HTML) 和统一资源定位符 (URL) 也都是在 1990 年出现的。另外, 在实证部分, 我们还将构造 1990 年以前的互联网使用率的工具变量, 试图解释互联网普及前信息对城市化加速的影响。

据我们所知, 目前尚没有文献将信息或者 ICT 技术的发展纳入城市化或者移民的理论与实证模型中, 虽然有不少研究论证了信息技术的其他影响。比如, Gaspar and Glaeser (1998) 考察了人与人的交流模式, 他们发现信息技术是增强而非取代了人们之间面对面的交流。另外 Czernich *et al.* (2011) 利用经合组织国家 1996—2007 年的面板数据估算了宽带技术对经济增长的影响, 结果显示宽带利用率每提高 10%, 可以带来 0.9%—1.5% 的人均 GDP 的增长。

¹ <https://online.jefferson.edu/communications/internet-history-timeline/>, 访问时间: 2018 年 5 月 3 日。

三、理论模型

我们的理论创新基于经典的两部门均衡模型 (Barrios *et al.*, 2006; Harris and Todaro, 1970; Lucas, 1988, 2004)。假设经济体由农村 (农业) 和城市 (工业) 两个部门组成, 所有人都在农村出生并只存活一期, 两部门的生产技术均由柯布-道格拉斯函数表示:

$$Y_r = L^{1-\gamma} N_r^\gamma, \quad (1)$$

$$Y_u = K^{1-\alpha} N_u^\alpha, \quad (2)$$

其中下标 r 和 u 分别表示农村和城市部门, L 、 K 、 N 表示土地 (相当于农业生产中的资本)、资本和劳动投入, γ 和 α ($0 < \gamma < 1$ 和 $0 < \alpha < 1$) 代表农村和城市劳动投入的产出弹性系数, 同时代表劳动在农村和城市部门所生产的增加值中的占比, 余下的是土地或资本的占比。所以, 式 (1)、(2) 中的系数之和为 1, 也意味着规模经济不变 (假设规模经济上升或下降, 并不会改变我们的理论模型结果)。假设劳动力总供给量 $N = N_r + N_u$, 我们将 N 标准化处理为 1, 这时的城镇化率可表示为 $U = \frac{N_u}{N} = N_u$ 。

依据 Harris and Todaro (1970)、Black and Henderson (1999) 的做法, 将农产品价格标准化为 1, 工业产品的相对价格为 P , 那么工资 (W) 即两部门各自的边际劳动产出可以表示为:

$$W_r = \gamma L^{1-\gamma} N_r^{\gamma-1} \quad (3)$$

$$W_u = P\alpha K^{1-\alpha} N_u^{\alpha-1}. \quad (4)$$

接着, 根据 Sjaastad (1962) 的建议和讨论, 我们引入正规的移民成本函数, 用 $C(I_c, Z)$ 代表, 其中 I_c 表示信息成本, Z 表示其他移民成本。进一步地, 借用 Amrhein (1985) 的成果, 信息成本函数可以定义为:

$$I_c = \begin{cases} 1 - \rho M & M < \delta, \rho > 0 \\ 0 & M \geq \delta \end{cases}, \quad (5)$$

其中, M 表示个人在移民前拥有的信息集, δ 是完全信息阈值, ρ 为一个正的系数。当 M 小于 δ 时, 即潜在移民对目的地不完全了解时, 信息成本为正, 但 I_c 是 M 的递减函数, 即已知信息量 M 越大, 信息成本越小。当已知信息量超过 δ 时, 移民已经充分掌握了与迁徙相关的信息, 这时移居所需要的信息的成本便为 0。基于式 (5), 我们得到:

$$\frac{\partial I_c}{\partial M} \leq 0, \quad \frac{\partial C}{\partial M} = C'(I_c, Z) \leq 0.$$

这时, 移民决策的均衡条件为 (参见 Harris and Todaro (1970)):

$$W_r = W_u - C(I_c, Z). \quad (6)$$

将式 (3) 和式 (4) 代入式 (6) 中可得:

$$\gamma L^{1-\gamma} N_r^{\gamma-1} = P\alpha K^{1-\alpha} N_u^{\alpha-1} - C(I_c, Z). \quad (7)$$

将 $N_r = (1 - N_u)$ 代入式 (7)，并在方程两边求关于 I_c 的导数，整理可得：

$$\frac{\partial N_u}{\partial I_c} = \frac{C'(I_c, Z)}{\gamma(\gamma-1)L^{1-\gamma}(1-N_u)^{\gamma-1}N_u^{\gamma-2} + (\alpha-1)(P\alpha K^{1-\alpha}N_u^{\alpha-2})}. \quad (8)$$

因为 $0 < \gamma < 1, 0 < \alpha < 1$ ，所以 $\frac{\partial N_u}{\partial I_c} < 0$ ，表示城市化率和移民成本负

相关。引入 $\frac{\partial I_c}{\partial M} \leq 0$ ，可以得到 $\frac{\partial N_u}{\partial M} \geq 0$ ，即城镇化率与信息变量之间呈正相关。

四、实证模型

为了实证检验上一部分的理论结果，我们参考 Davis and Henderson (2003)，设定了一个加入信息变量的跨国回归模型²：

$$urban_{it} = \beta Infor_{it} + \alpha' X_{it} + \theta' C_{it} + \mu_i + \lambda_t + \epsilon_{it}, \quad (9)$$

其中，下标 i 和 t 分别代表国家和年度， ϵ 为白噪声残差， $urban$ 是城市化率， $Infor$ 代表关键信息变量，用互联网普及率 ($inter$) 代理， λ 和 μ 分别为时间和国家固定效应。我们用 X 代表传统的“推拉”控制变量， C 代表除了传统“推拉”因素以外的控制变量（详细讨论见下文）。 β, α, θ 为待估算参数。

如前文所述，暂且假定 1990 年以前互联网对城市化的作用不显著，这就使得我们可以构造一个自然实验，其中 1990 年前的 $Infor = 0$ ，1990 年后各国或同一国家不同年份的 $Infor$ 不等。这样一来，式 (9) 可以看作一个 DID 方程， β 则代表信息对城市化加速的作用。在下文中，我们还将推出 1990 年之前的 $infor$ 的工具变量的观察值，以验证信息在 1990 年之前对城市化的促进作用。另外，我们还会使用短信数量 (SMS) 作为信息的代理变量，用以考察模型估算结果的稳健性。

控制向量 X 包含了传统的“推拉”因素，包括：(1) 人均 GDP 的对数 ($lngdp$)；(2) 人均 GDP 增长率 ($gdpg$)；(3) 工业化水平 ($nagr$ ，用 GDP 中的非农增加值比重衡量)；(4) 教育 (sch ，以成年人口平均受教育年限衡量)，加入这个变量是因为平均受教育程度高的城市往往提供更高的工资，吸引各教育层次的移民 (Black and Henderson, 1999; Lucas, 1988; Rauch, 1993)；(5) 贸易 ($open$ ，用出口加进口占 GDP 的百分比衡量)，加入贸易变量是因为贸易通常要求市场营销和金融之类的配套服务，而这些服务在城市中更加集中 (Hofmann and Wan, 2013; 章元和万广华, 2013)；(6) 公路密度 ($density$ ，即每平方千米土地面积上的道路千米数)。公路密度越高，交通成本越低。交

² 豪斯曼检验结果表明固定效应模型优于随机效应模型。

通基础设施是研究卫星城市文献中通常考虑的因素 (Krugman, 1991)。

控制向量 C 包含余下的其他控制变量, 包括: (1) 城市化收敛程度 (*urban60*, 即 1960 年的城市化水平); (2) 城市首位度 (*primacy*, 最大城市人口占城市总人口的百分比), 城市首位度越高, 往往城市化率越高 (Davis and Henderson, 2003); (3) 人口密度 (*popdensity*, 即每平方千米土地上的人口数量) (Beals *et al.*, 1967); (4) 人口增长率 (*popgr*) (Jedwab *et al.*, 2017)。

五、数据和模型估算结果

(一) 数据

本文使用 109 个非经合组织国家或地区从 1960 年到 2013 年 54 年的相关数据。我们排除了经合组织国家, 是因为截至 20 世纪 60 年代, 经合组织国家已基本完成了城市化进程。三个单一城市国家或地区——中国香港、中国澳门和新加坡——也被排除在外。本文数据主要来源于: 世界银行世界发展指标数据库 (WDI), 世界电信/ICT 指标数据库 (ITU) 的 ICT 数据, Barro-Lee 受教育程度数据库 (BL)³ 的教育数据, 世界道路交通统计数据库 (WRS) 的道路密度数据。表 2 中列出了每个变量的定义和来源。

表 2 变量的定义和来源

变量	定义	来源
因变量:		
<i>urban</i>	城市化率 (城市人口占总人数的百分比)	WDI
关键自变量:		
<i>lninter</i>	每 100 人中互联网使用人数的对数	ITU
<i>lnsms</i>	短信数量的对数	ITU
传统控制变量: 推—拉因素		
<i>lngdp</i>	人均 GDP 的对数 (美元)	WDI
<i>gdpggr</i>	人均 GDP 增长率 (年度百分比)	WDI
<i>sch</i>	教育程度的对数, 匹配最近年度	BL
<i>nagr</i>	非农业占 GDP 百分比	WDI
<i>open</i>	国际贸易占 GDP 百分比	WDI
<i>density</i>	道路密度 (每平方千米土地面积的道路千米数)	WRS

³ <http://www.barrolee.com/>, 访问时间: 2015 年 9 月 1 日。由于这个数据库只提供五年期数据, 所以本文使用最接近年份的观察值。

(续表)

变量	定义	来源
其他控制变量：		
<i>popdensity</i>	人口密度（每平方千米土地面积人口数）	WDI
<i>popgr</i>	人口自然增长率（年度百分比）	WDI
<i>primacy</i>	最大城市中的人口数占城市人口的百分比	WDI
<i>urban60</i>	1960 年城市化率	WDI
<i>lnurbanpop</i>	城市人口增长率的对数	WDI
工具变量		
<i>tel_net</i>	1989 年固定电话用户数（每百人）	ITU

表 3 汇报了变量的描述性统计，有几点值得注意。首先，数据观察值的多少因变量而异，特别是道路密度和短信的样本体积远远小于其他指标。其次，城市化率的变动范围很大，最小值为 3.13%，最大值为 99.06%⁴，标准差为 22.5%。城市化率比较高的国家主要集中在拉丁美洲和加勒比海地区、北美洲和欧洲以及一些资源出口国（UN, 2014; Gollin *et al.*, 2016）。因此，后文将用分区域样本来估算回归模型，以检验基准模型的稳健性。最后，本文中的数据包括一些异常的观察值，例如战争期间，GDP 增长下降 65%，人口增长为负。在估算基准模型时，没有剔除这些观察值，因为我们认为这些宏观因素不会显著影响信息对城市化的效应。但我们在稳健性检验部分将考虑这些异常值的问题。

表 3 描述性统计

变量	样本体积	平均值	标准差	最小值	最大值
<i>urban</i>	3 865	45.4	22.5	3.1	99.1
<i>lngdp</i>	3 563	7.3	1.3	3.9	11.0
<i>gdpgr</i>	3 597	1.9	6.4	-65.0	102.8
<i>nagr</i>	3 229	78.9	14.8	25.7	99.7
<i>sch</i>	3 865	1.3	0.8	-3.1	2.5
<i>open</i>	3 522	72.8	39.5	0.3	280.4
<i>density</i>	2 390	0.6	5.5	0.0	256.0
<i>popdensity</i>	3 862	105.3	191.3	1.1	1 752.9

⁴ 这些国家在 20 世纪 60 年代的城市化率不是很高，其城市化过程大多发生在我们的样本区间（1960—2013）之内，而本文正是要研究信息在城市化过程中的作用，故未从样本中剔除这些国家。

(续表)

变量	样本体积	平均值	标准差	最小值	最大值
<i>primacy</i>	3 384	33.2	14.6	2.6	100.0
<i>popgr</i>	3 855	2.1	1.6	-7.6	17.5
<i>lninter</i>	3 326	7.2	6.3	0.0	16.0
<i>lnsms</i>	524	21.0	2.7	8.0	27.4

(二) 基准模型结果

在进行模型估计之前,我们先进行面板单位根检验。表4中的结果表明,ADF和ADF-PP的检验结果均拒绝了变量存在单位根为零假设。表4的最后一行中,基准模型的残差(对应于表5中第(3)列的模型)是平稳的,因而排除了伪回归的可能性。

表4 面板单位根检验结果

	Inverse χ^2	<i>p</i> -值
	(1)	(2)
<i>urban</i>	2 087.25	0.000
<i>lninter</i>	2 779.57	0.000
<i>lnsms</i>	807.94	0.000
<i>lngdp</i> [#]	455.56	0.000
<i>gdpggr</i>	2 009.13	0.000
<i>nagr</i>	384.73	0.000
<i>sch</i> [#]	597.21	0.000
<i>open</i>	378.03	0.000
<i>density</i>	327.59	0.000
<i>primacy</i>	747.95	0.000
<i>popdensity</i>	660.12	0.000
<i>popgr</i> [#]	682.22	0.000
Regression Residual [#]	239.03	0.000

注: #表示变量是漂移平稳(drift stationary),数值Inverse χ^2 对于ADF和ADF-PP检验,其统计值相同。

基准模型如表5所示。表5的上半部分(A组)报告了主要变量的系数估算结果,下半部分(B组)报告了年度虚拟变量的系数估算结果。为了节

省空间，表 5 中只列出了 1990 年起的结果。第 (1) 列只包含“推拉”控制变量，第 (2) 列包括了所有的控制变量，第 (3) 列是加入了信息代理变量后的回归结果。⁵

表 5 基准模型结果

变量	(1)	(2)	(3)
A 组			
<i>lngdp</i>	3.289*** (0.62)	4.115*** (0.61)	2.706*** (0.72)
<i>gdpgr</i>	-0.109*** (0.02)	-0.117*** (0.02)	-0.111*** (0.02)
<i>nagr</i>	-0.019 (0.02)	-0.002 (0.02)	0.022 (0.03)
<i>sch</i>	3.909*** (0.77)	2.588*** (0.72)	2.674*** (0.85)
<i>open</i>	0.006 (0.01)	0.016** (0.01)	0.024*** (0.01)
<i>density</i>	-0.008 (0.01)	0.003 (0.01)	0.016* (0.01)
<i>urban60</i>		0.440*** (0.04)	0.371*** (0.06)
<i>primacy</i>		-0.02 (0.04)	-0.011 (0.04)
<i>popdensity</i>		-0.071*** (0.01)	-0.072*** (0.01)
<i>popgr</i>		-0.157 (0.15)	-0.740*** (0.23)
<i>lninter</i>			0.527*** (0.13)
<i>constant</i>	-0.942 (5.29)	41.301*** (5.76)	73.918*** (7.22)
<i>N</i>	1 923	1 710	1 386
<i>adj. R²</i>	0.972	0.977	0.977

⁵ 因为表 5 中的大多数系数显著，所以多重共线性不是问题。

(续表)

变量	(1)	(2)	(3)
B组			
1990	11.700***	15.824***	0.000
1991	12.143***	16.494***	-2.946
1992	12.552***	16.888***	-3.220*
1993	13.030***	17.468***	-2.618
1994	13.460***	17.942***	-2.353
1995	13.574***	18.390***	-2.461
1996	13.917***	18.829***	-2.382*
1997	14.564***	19.527***	-2.512*
1998	14.883***	19.588***	-2.695**
1999	15.893***	20.466***	-2.241*
2000	14.593***	20.096***	-2.853**
2001	14.967***	20.664***	-2.698**
2002	15.290***	20.740***	-2.851***
2003	15.947***	21.666***	-2.285**
2004	17.062***	22.763***	-0.893
2005	16.505***	22.005***	-1.807*
2006	16.433***	22.347***	-1.594
2007	16.994***	23.062***	-0.691
2008	17.191***	23.000***	-1.139
2009	16.921***	22.775***	-1.342
2010	17.010***	23.201***	-1.143
2011	17.966***	24.004***	-0.450
2012	18.261***	24.446***	-0.020
2013	18.020***	24.415***	0.000

注：第(1)列只包含“推拉”控制变量；第(2)列包括所有控制变量；第(3)列是自然实验，1990年以前的互联网数据设为0。括号内为稳健标准差；*、**和***分别表示10%、5%和1%显著性水平。

表5的第(1)、(2)列确认了传统的城市化模型无法解释城市化的加速，因为对应于这两个模型的所有年份虚拟变量的系数都正显著，且随时间单调增加，从11.7增长到18.02。与此恰恰相反，第(3)列的结果显示，互联网的作用显著为正，表明信息对城市化的促进作用，这印证了理论模型的结果。同时，对应于第(3)列的年度虚拟变量的系数大部分都不显著，且估算值比

较稳定。这说明，信息变量可以解释城市化加速这一重要现象。

使用表 5 下半部分报告的年度虚拟变量的系数估算值作图（见图 2），可以更为直观地获得本文的关键发现与结论。点线和虚线对应表 5 中的第（1）和第（2）列，曲线大体上呈单调递增趋势。实线对应包含了信息变量的模型（表 5 的第（3）列）。可以看出 1990 年以前实线与点线、虚线基本平行，但加入 1990 年后的互联网变量后，实线变得较为平稳并接近于零。

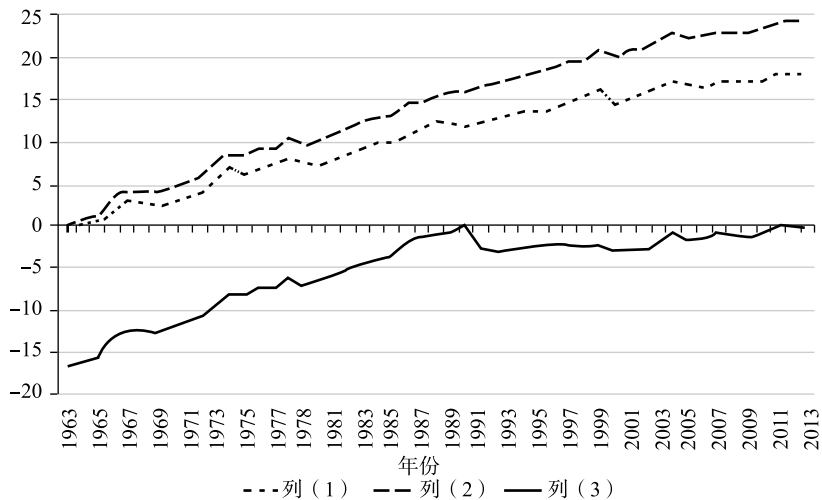


图 2 对应表 5 中的年度虚拟变量估算值

就控制变量而言，人均 GDP、经济增长和教育的结果与预期一致，工业化水平不显著，道路密度仅在列（3）微弱显著，而贸易指标在列（2）、（3）显著，符合预期。

（三）IV 固定效应模型结果

城市化和信息之间可能存在反向因果关系，即城市化水平越高的国家，其互联网普及程度越高，因为城市通常比农村更能提供较多较好的服务，包括基础设施如互联网。内生性也可能由第三方因素引起，即可能存在的遗漏变量同时促进城市化和互联网普及率。这样的因素可能与技术进步或经济增长相关，比如随着经济的发展，越来越多人迁往城市，同时经济水平的提高也增加了人们对互联网的需求（Czernich *et al.*，2011）。为了解决这些潜在的内生性问题，本文借鉴 Czernich *et al.*（2011）的研究成果，依据技术扩散理论，构造互联网普及率的工具变量（IV）。

本文工具变量（IV）的构造基于这样一个事实，即在 20 世纪 90 年代初互联网的使用依赖于固定电话网络，人们使用配线架并通过所谓的拨号接入互联网服务提供商（ISP）。换句话说，使用互联网需要访问与互联网无本质关系的、原先就已经独立运行的语音电话网络。基于 Griliches（1957）有关

技术创新过程的开创性研究, Czernich *et al.* (2011) 将一个国家互联网普及率的最大值 (ω_i), 设为互联网存在之前的固定电话普及率 (tel_net_{i0}) 的线性函数:

$$\omega_i = \omega_0 + \varphi_1 tel_net_{i0}, \quad (10)$$

其中, φ_1 和 ω_0 为参数, tel_net_{i0} 为 1989 年每百人中的固定电话用户数。观察到的互联网普及率 $inter$ 可以表示为 ω_i 和时间 t 的逻辑方程 (见 Comin *et al.*, 2006; Geroski, 2000):

$$inter_{it} = \frac{\omega_i}{1 + \exp[-\eta(t - \tau)]} + \pi_{it}, \quad (11)$$

其中, 参数 η 和 τ 分别是式 (11) 技术扩散曲线的速度和拐点⁶, π_{it} 为误差项。

将式 (11) 代入式 (10) 可以得到一个完全由外生变量决定的非线性模型, 用于估算互联网普及率, 并作为信息的工具变量:

$$inter_{it} = \frac{\omega_0 + \varphi_1 tel_net_{i0}}{1 + \exp[-\eta(t - \tau)]} + \pi_{it}. \quad (12)$$

表 6 汇报了 IV 模型式 (12) 的估计结果, 分地区回归拟合度较好, R^2 高达 0.81 和 0.91。

表 6 IV 第一阶段估计结果

被解释变量: 互联网普及率	亚洲和非洲	欧洲和美洲
固定电话普及系数 (φ_1)	4.20*** (0.000)	2.81*** (0.000)
扩散速度 (η)	0.25*** (0.000)	0.30*** (0.000)
拐点 (τ)	2 009.34*** (0.000)	2 007.84*** (0.000)
常数项	18.37*** (0.000)	28.44*** (0.000)
R^2	0.81	0.91
N	1 342	760

注: 括号内为 p 值, STATA 在此处没有汇报标准差值, *** 表示 1% 显著性水平。

表 7 报告了 IV 固定效应回归的结果。根据联合国 2014 年的数据 (UN, 2014), 亚洲和非洲的城市化速度明显快于其他地区, 所以使用亚洲、非洲和其他地区 (欧洲和美洲) 三个子样本来估算 IV 模型。回归结果与表 5 中的基准结果基本一致, 特别是 B 组里的大部分系数变得不太显著且更平稳。这证实了互联网普及率能很好地解释城市化的加速。

⁶ 在拐点 τ 处, 扩散曲线达到最大增长率。

表 7 IV 的回归结果

因变量	所有国家	亚洲	非洲	美洲和欧洲
	(1)	(2)	(3)	(4)
A 组				
<i>lninter</i>	2.047*** (0.55)	10.097*** (2.13)	6.305*** (0.86)	1.214 (1.69)
<i>lngdp</i>	3.766*** (0.60)	3.403*** (1.12)	4.283*** (1.15)	9.195*** (1.03)
<i>gdpgr</i>	-0.117*** (0.02)	-0.161*** (0.05)	-0.078*** (0.03)	-0.062** (0.03)
<i>nagr</i>	0.001 (0.02)	0.151*** (0.05)	-0.036 (0.04)	0.254*** (0.04)
<i>sch</i>	3.180*** (0.77)	-1.198 (1.66)	1.348 (1.06)	5.865*** (1.14)
<i>open</i>	0.014* (0.01)	-0.019 (0.01)	0.003 (0.01)	-0.042*** (0.01)
<i>density</i>	0.008 (0.01)	0.054 (0.08)	-0.194*** (0.07)	0.372 (0.30)
<i>primacy</i>	-0.016 (0.04)	-0.083 (0.09)	0.033 (0.06)	-0.482*** (0.07)
<i>popdensity</i>	-0.064*** (0.01)	-0.079*** (0.01)	-0.076*** (0.02)	-0.032 (0.02)
<i>popgr</i>	-0.104 (0.14)	2.316*** (0.78)	0.289** (0.13)	0.001 (0.76)
<i>urban60</i>	0.440*** (0.04)	-1.276*** (0.39)	1.166*** (0.11)	0.584*** (0.13)
constant	49.112*** (7.81)	52.698*** (14.64)	-18.125** (8.64)	-52.276*** (9.99)
B 组				
1990	1.314	9.140**	12.562***	8.491***
1991	1.483	7.334*	11.833***	8.317***
1992	1.351	5.889	10.812***	8.054**
1993	1.392	4.05	9.894***	7.995**

(续表)

因变量	所有国家	亚洲	非洲	美洲和欧洲
	(1)	(2)	(3)	(4)
1994	1.335	2.419	8.944***	8.458**
1995	1.149	0.781	7.926**	7.830*
1996	1.078	-1.1	6.941**	7.701
1997	1.204	-2.057	4.935	7.603
1998	0.724	-5.235	3.762	6.848
1999	0.973	-5.381	3.858	7.221
2000	0.206	-7.109	2.506	6.62
2001	0.244	-8.224	-0.064	6.894
2002	-0.081	-10.819	0.359	6.547
2003	0.316	-10.523	-0.56	7.219
2004	1.044	-11.46	0.85	7.307
2005	-0.25	-12.117	-1.863	6.316
2006	-0.209	-13.876	-2.569	6.298
2007	0.078	-14.513	-0.614	6.112
2008	-0.248	-15.157	-3.69	6.223
2009	-0.749	-16.596	-5.173	5.832
2010	-0.616	-17.57	-5.79	5.967
2011	-0.059	-16.121	-5.329	6.232
2012	0.211	-16.422	-6.072	6.12
2013	0	-15.933	-6.579	6.318
N	1 710	365	748	424
adj. R^2	0.977	0.96	0.971	0.981

注：A组部分包含主要变量的估算结果，B组部分报告1990年以后的年度虚拟变量的系数估算结果；括号内为稳健标准差(SE)；*、**和***分别表示10%、5%和1%显著性水平。

六、稳健性检验

(一) SMS作为信息的另一个代理变量

首先，我们使用发送的短信息(SMS)数量作为信息的另一个代理变量。遗憾的是，ITU数据库中仅包含2000年之后的跨国数据，所以样本体积减小了很多。表8报告了回归结果。与前文一致，年度虚拟变量系数变得不显著

且没有趋势。信息变量的系数估算值为正，虽然在第（1）列中不显著。

表 8 SMS 作为信息代理变量

	全样本	亚洲和非洲	美洲和欧洲
	(1)	(2)	(3)
A 组			
<i>lnsms</i>	0.112 (0.31)	0.241* (0.13)	0.412** (0.20)
<i>lngdp</i>	4.273 (3.83)	7.100*** (2.13)	-5.353 (3.25)
<i>gdpgr</i>	0.039 (0.05)	-0.073 (0.06)	0.051 (0.03)
<i>nagr</i>	0.202 (0.20)	0.06 (0.05)	0.204 (0.19)
<i>sch</i>	3.12 (1.91)	-2.709* (1.59)	3.528 (4.01)
<i>open</i>	-0.004 (0.02)	-0.01 (0.01)	-0.002 (0.02)
<i>density</i>	7.202 (5.40)	3.875 (2.87)	-3.012 (9.63)
<i>primacy</i>	-0.395* (0.22)	0.532 (0.42)	-0.496* (0.27)
<i>popdensity</i>	0.051 (0.06)	-0.137** (0.05)	0.226*** (0.08)
<i>popgr</i>	-1.151* (0.68)	-0.144 (0.92)	-0.87 (0.56)
<i>urban60</i>	0.378 (0.33)	0.618** (0.27)	0.877* (0.49)
Constant	7.826 (22.79)	-45.147** (20.00)	37.507** (17.17)
国家虚拟变量	是	是	是
<i>N</i>	219	112	107
adj. <i>R</i> ²	0.996	0.999	0.995

(续表)

	全样本	亚洲和非洲	美洲和欧洲
	(1)	(2)	(3)
B组			
2001	0.261	0.350	0.000
2002	0.451	0.667	0.000
2003	-0.370	1.360	-0.091
2004	0.128	2.431	-0.232
2005	-0.177	3.279	-0.441
2006	-0.431	3.427	0.301
2007	-0.243	3.513	0.747
2008	0.034	3.834	1.189
2009	0.592	4.168	1.366
2010	0.058	5.018	0.968
2011	0.452	5.806	1.421
2012	0.771	6.282	1.869
2013	0.484	6.552	2.080

注：A组部分包含主要变量的估算结果，B组部分报告2001年以后的年度虚拟变量的系数估算结果；括号内为稳健标准差（SE）；*、**和***分别表示10%、5%和1%显著性水平。

（二）其他稳健性检验

还有一种关于城市化加速的可能解释是人口流动的雪球效应，即某个国家上期移民规模越大，通过个人关系网络，带来本期移民人数的增加（Sahota, 1968）。这也可能是移民成本下降的一种表现（Carrington *et al.*, 1996），但不是本文所要论述的信息带来的影响。为了检测本文主要发现和结论的稳健性，在基准和IV模型中加入了城市人口的年度变化，以控制这种雪球效应，同时，为了去除样本极值的影响，我们对数据进行了缩尾处理（参见Barnett and Lewis, 1994），结果如表9所示。

表9 人口流动雪球效应

	(1)	(2)	(3)
A组			
<i>urbanpopchange</i> (-1)	2.088***	2.887***	2.875***
	(0.41)	(0.56)	(0.56)

(续表)

	(1)	(2)	(3)
A 组			
<i>lngdp</i>	5.564*** (0.71)	4.649*** (0.79)	4.435*** (0.80)
<i>gdpgr</i>	-0.115*** (0.03)	-0.132*** (0.03)	-0.164*** (0.04)
<i>nagr</i>	-0.024 (0.03)	-0.003 (0.03)	0.001 (0.03)
<i>sch</i>	1.794** (0.77)	1.818** (0.92)	2.033** (0.89)
<i>open</i>	0.008 (0.01)	0.011 (0.01)	0.011 (0.01)
<i>density</i>	-0.015* (0.01)	-0.017** (0.01)	-0.024*** (0.01)
<i>urban60</i>	0.495*** (0.05)	0.472*** (0.07)	0.473*** (0.07)
<i>primacy</i>	-0.063 (0.05)	-0.044 (0.05)	-0.038 (0.05)
<i>popdensity</i>	-0.079*** (0.01)	-0.081*** (0.01)	-0.083*** (0.01)
<i>popgr</i>	-0.284 (0.21)	-1.112*** (0.29)	-1.489*** (0.37)
<i>lninter</i>		0.530*** (0.13)	0.551*** (0.13)
constant	36.233*** (7.71)	36.133*** (9.50)	39.400*** (9.05)
N	1 429	1 154	1 154
adj. R^2	0.98	0.98	0.98
B 组			
1990	-4.481***	0.000	0.000
1991	-3.682***	-2.814	-2.447
1992	-3.021***	-2.983*	-2.511
1993	-2.516***	-2.644	-2.322
1994	-2.138***	-2.671*	-2.262

(续表)

	(1)	(2)	(3)
B 组			
1995	-1.812**	-2.563*	-2.171
1996	-1.213*	-2.209*	-1.857
1997	-0.424	-2.241*	-1.914
1998	-1.076	-3.204**	-2.957**
1999	0	-2.762**	-2.618**
2000	0.182	-2.689**	-2.548**
2001	0.47	-2.923***	-2.808***
2002	0.963	-2.689***	-2.708***
2003	1.883**	-2.033**	-1.915**
2004	2.909***	-0.867	-0.617
2005	2.515***	-1.378	-1.189
2006	2.591***	-1.501	-1.388
2007	3.484***	-0.421	-0.352
2008	3.405***	-0.91	-0.901
2009	3.182***	-1.229	-1.162
2010	3.484***	-1.054	-0.982
2011	4.312***	-0.402	-0.327
2012	4.603***	-0.17	-0.075
2013	4.742***	0.000	0.000

注：第（1）列中不加入信息变量；第（3）列是缩尾样本结果。A组部分包含主要变量的估算结果，B组部分报告1990年以后的年度虚拟变量的系数估算结果；括号内为稳健标准差（SE）；*、**和***分别表示10%、5%和1%显著性水平。

表9中每一列都显示城市人口变化的系数估计值为正，表明人口流动的雪球效应确实存在，并对城市化有正向影响。但是，如B组第（1）列所示，即使控制这一变量，年度虚拟变量的系数估计值仍然显著为正，且有增大的趋势。这说明雪球效应并不是城市化加速的原因。相比之下，在加入了信息变量的第（2）列和第（3）列中，年度虚拟变量系数变得不那么显著。所以，城市化的加速更多地来源于信息的作用而非人口雪球效应。

最近的一项研究认为，目前发展中国家的城市化水平提升较快是因为城市人口的自然增长率远远高于工业时代的欧洲（Jedwab *et al.*，2017）。需要强调的是，我们研究的快速城市化并不等同于城市人口的快速扩张。这是因

为城市化率是一个比值的概念，它取决于城市和农村两方面的人口变化。实际上 Bloom *et al.* (2010) 的研究表明，农村和城市的人口自然增长率相差不多，如果忽略了移民的贡献，城市化率很有可能是不变的。但是，Jedwab *et al.* (2017) 的表 2，恰恰显示了城市化率的变动主要是由移民带来的，而不是城市人口自然增长更快的结果。即便如此，为了增强本文结论的稳健性，我们在 IV 模型中加入城市人口自然增长率这个额外变量。因为数据来源于 Jedwab *et al.* (2017) 使用的十年平均自然增长率 (*nurban*)，我们也将其他变量取十年平均值后再进行模型估算，结果见表 10。

表 10 的第 (1) 列是自然增长率 *nurban* 和信息变量都不加入的情况，第 (2) 列仅加入自然增长率。在这两列中，年度虚拟变量的系数估算值仍然具有明显的时间趋势，说明城市化加速仍然不能被解释。而第 (3)、(4) 列的结果表明，当引入信息变量后，时间趋势就消失了，而且信息变量的系数仍然显著为正。特别重要的是，在第 (2)、(4) 列中，城市人口自然增长率变量并不显著。

表 10 城市化和城市人口自然增长率

	(1)	(2)	(3)	(4)
<i>lngdp</i>	4.207** (1.90)	3.792* (2.04)	3.466* (1.88)	3.20 (2.14)
<i>gdpgr</i>	-0.319** (0.15)	(0.24) (0.19)	-0.310** (0.15)	(0.20) (0.18)
<i>nagr</i>	(0.02) (0.07)	0.12 (0.12)	(0.00) (0.07)	0.13 (0.12)
<i>sch</i>	1.48 (1.90)	2.62 (2.28)	2.01 (1.99)	3.56 (2.48)
<i>open</i>	0.04 (0.03)	0.078* (0.05)	0.04 (0.03)	0.07 (0.05)
<i>density</i>	(0.20) (0.21)	(0.40) (0.32)	(0.19) (0.21)	(0.40) (0.32)
<i>urban60</i>	0.372*** (0.13)	(0.60) (0.45)	0.370*** (0.13)	(0.54) (0.44)
<i>primacy</i>	0.07 (0.11)	-0.367** (0.15)	0.07 (0.11)	-0.385** (0.15)
<i>popdensity</i>	-0.070*** (0.02)	-0.063*** (0.02)	-0.068*** (0.02)	-0.059*** (0.02)

(续表)				
	(1)	(2)	(3)	(4)
<i>popgr</i>	0.14 (0.52)	(0.73) (0.76)	0.35 (0.53)	(0.56) (0.76)
<i>nurban</i>		1.41 (1.22)		1.51 (1.17)
<i>lninter</i>			0.784** (0.35)	0.880** (0.42)
1970s	4.434*** (1.58)	-11.113*** (3.68)	4.391*** (1.58)	2.24 (8.07)
1980s	9.764*** (2.08)	-8.587*** (2.43)	9.563*** (2.12)	4.37 (6.97)
1990s	15.121*** (2.59)	-5.687*** (1.58)	8.169* (4.69)	(0.20) (3.33)
2000s	19.653*** (3.13)	-2.612** (1.22)	9.40 (6.34)	(1.25) (1.51)
2010s	22.146*** (3.56)	0.00 (.)	10.68 (7.25)	0.00 (.)
constant	47.462** (18.76)	69.161*** (23.97)	49.530*** (18.47)	55.058** (25.32)
国家虚拟变量	是	是	是	是
<i>N</i>	289	119	289	119
adj. <i>R</i> ²	0.963	0.969	0.963	0.97

注：第(1)列中不加入信息变量和城市人口自然增长率；第(2)、(3)列分别加入信息变量和自然增长率。括号内为稳健标准差；*、**和***分别表示10%、5%和1%显著性水平。

最后，在前文关于工具变量的预测中，我们仅仅考虑了1990年以后的互联网普及率，而将1990年之前的信息变量或其影响假设为0。实际上，在1990年之前，互联网已经被用于军事和研究用途，只是没有大规模地在民间普及。这可能也是为什么由ITU提供的互联网普及率数据，在1990年之前大部分是缺失的。但电话在1990年之前是存在的，它也是传递信息的最为重要的载体。所以，作为稳健性检验的一部分，我们用模型式(12)来倒推1990年之前的信息的工具变量观察值，并使用1963年之后的数据估算IV模型。表11报告了结果，图3描绘了对应的年度虚拟变量的系数估算值。结果都显示，信息变量可以在很大程度上解释城市化的加速。

表 11 加入 1990 年前信息工具变量观察值的模型结果

	全样本	亚洲和非洲	美洲和欧洲	全样本	亚洲和非洲	美洲和欧洲
	(1)	(2)	(3)	(4)	(5)	(6)
A 组						
<i>lngdp</i>	4.115*** (0.61)	5.340*** (0.78)	3.277*** (0.72)	4.115*** (0.61)	5.340*** (0.78)	3.277*** (0.72)
<i>gdpgr</i>	-0.117*** (0.02)	-0.115*** (0.03)	-0.033** (0.02)	-0.117*** (0.02)	-0.115*** (0.03)	-0.033** (0.02)
<i>nagr</i>	-0.002 (0.02)	-0.033 (0.02)	0.060* (0.03)	-0.002 (0.02)	-0.033 (0.02)	0.060* (0.03)
<i>sch</i>	2.588*** (0.72)	1.660* (0.86)	5.084*** (1.29)	2.588*** (0.72)	1.660* (0.86)	5.084*** (1.29)
<i>open</i>	0.016** (0.01)	0.007 (0.01)	-0.016** (0.01)	0.016** (0.01)	0.007 (0.01)	-0.016** (0.01)
<i>density</i>	0.003 (0.01)	-0.191*** (0.04)	0.007** (0.00)	0.003 (0.01)	-0.191*** (0.04)	0.007** (0.00)
<i>primacy</i>	-0.02 (0.04)	0.028 (0.05)	-0.273*** (0.07)	-0.02 (0.04)	0.028 (0.05)	-0.273*** (0.07)
<i>popden- sity</i>	-0.071*** (0.01)	-0.083*** (0.01)	0.128*** (0.02)	-0.071*** (0.01)	-0.083*** (0.01)	0.128*** (0.02)
<i>popgr</i>	-0.157 (0.15)	0.171 (0.14)	-1.644*** (0.25)	-0.157 (0.15)	0.171 (0.14)	-1.644*** (0.25)
<i>urban60</i>	0.440*** (0.04)	0.391*** (0.05)	1.113*** (0.04)	0.408*** (0.04)	0.356*** (0.05)	1.118*** (0.04)
<i>lnInter</i>				2.450*** (0.21)	2.681*** (0.26)	1.661*** (0.24)
constant	42.048*** (5.76)	47.070*** (6.76)	-41.937*** (6.21)	58.742*** (5.90)	66.160*** (7.46)	-31.880*** (6.69)
<i>N</i>	1 710	1 113	597	1 710	1 113	597
adj. <i>R</i> ²	0.977	0.966	0.976	0.977	0.966	0.976

(续表)

	全样本	亚洲和非洲	美洲和欧洲	全样本	亚洲和非洲	美洲和欧洲
	(1)	(2)	(3)	(4)	(5)	(6)
B 组						
1990	15.076***	16.043***	13.293***	0.709	-0.500	3.046***
1991	15.747***	16.692***	13.400***	0.829	-0.454	2.780***
1992	16.141***	17.167***	13.762***	0.676	-0.578	2.770***
1993	16.721***	17.757***	13.946***	0.711	-0.584	2.586***
1994	17.195***	18.312***	14.417***	0.645	-0.619	2.691***
1995	17.642***	18.851***	14.440***	0.558	-0.666	2.351***
1996	18.082***	19.211***	14.683***	0.469	-0.885	2.236***
1997	18.779***	19.455***	15.014***	0.646	-1.210	2.215***
1998	18.841***	18.984***	14.810***	0.197	-2.241	1.664**
1999	19.719***	20.569***	15.469***	0.576	-1.202	1.985**
2000	19.348***	21.062***	14.193***	-0.28	-1.241	0.380
2001	19.916***	21.364***	15.175***	-0.182	-1.452	1.043
2002	19.993***	21.609***	15.108***	-0.555	-1.700	0.672
2003	20.918***	22.917***	15.673***	-0.058	-0.860	0.946
2004	22.016***	24.458***	15.909***	0.636	0.239	0.908
2005	21.258***	23.820***	15.225***	-0.498	-0.810	-0.031
2006	21.599***	23.963***	15.286***	-0.504	-1.048	-0.205
2007	22.314***	26.054***	15.248***	-0.105	0.698	-0.457
2008	22.252***	25.135***	15.711***	-0.451	-0.532	-0.186
2009	22.027***	24.663***	15.713***	-0.928	-1.280	-0.355
2010	22.454***	24.815***	15.768***	-0.722	-1.369	-0.450
2011	23.256***	26.583***	16.059***	-0.11	0.190	-0.289
2012	23.698***	26.953***	16.559***	0.168	0.381	0.101
2013	23.668***	26.723***	16.552***	0.000	0.000	0.000

注：列(1)—(3)中不加入信息变量；列(4)—(6)加入信息变量。A组数据说明主变量的系数，B组数据说明1990年以后的年度虚拟变量系数；括号内为稳健标准差；*、**和***分别表示10%、5%和1%显著性水平。

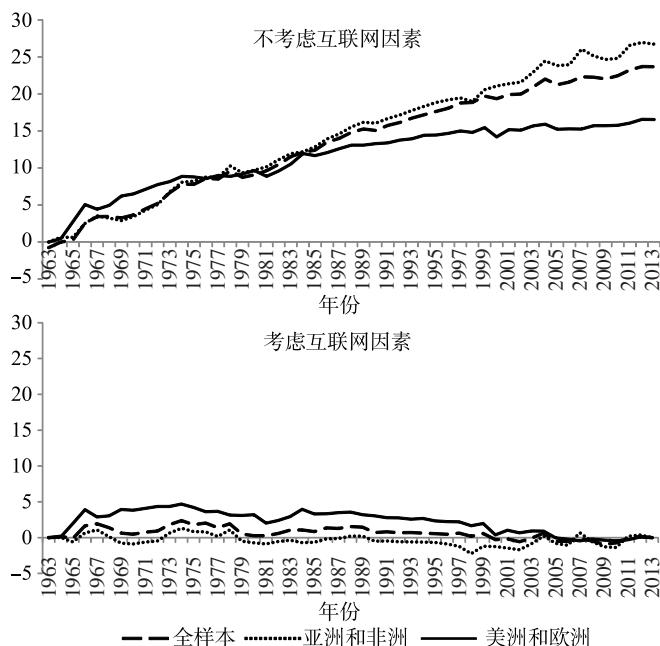


图 3 年度虚拟变量系数估算值

注：对应表 11 中第 (1) — (6) 列的时间虚拟变量系数估算值。

七、结论及政策建议

本文揭示了过去几十年来全球城市化进程加速的重要现象，印证了不少国家抱怨城市化浪潮席卷而来的经典事实，也为万广华（2011）提出的积极推动城市化的主张提供了实证依据。即使控制了传统的“推拉”因素，非 OECD 国家的城市化率在 21 世纪比 20 世纪 60 年代仍平均要高出 16 个百分点。这一加速现象，迄今为止被学术界和政策界普遍忽视，由此可能带来了诸如交通拥挤、污染、城市贫民窟等一系列的严重后果。因此，揭示并厘清这一现象对发展中国家的政府、企业 and 个人的影响，都具有深远的理论和现实意义。

本文的另一更为重要的贡献，是解释这个至今无人注意到的城市化加速现象。我们基于 Sjaastad (1962) 的研究，从移居的心理和货币成本着手，开创性地将信息引进城市化研究。更为具体地，本文首次构建了一个包含信息成本的两部门一般均衡模型，推导出了信息与城市化进程的理论关系。最后，我们用 1960—2013 年 109 个非 OECD 国家或地区的数据，验证了理论模型的可靠性，从而得出了信息在很大程度上可以解释城市化加速的发现和结论。这些基于基准回归和工具变量模型的发现和结论，对于不同地区，不同控制变量组合以及不同的代理变量都具有稳健性。

本文的发现和结论对于政府和社会各界都非常重要。本文分析的是 1960 年到 2013 年的数据，但未来信息技术的发展还将不断降低信息成本，并

进一步推动人口迁移, 所以城市化进程加速的长期趋势不可避免。政策部门在制定发展战略、城乡发展规划、资源分配政策等方面要作长期的准备。单纯根据历史数据主观臆断城市化速度(如中国流行的把城市化率提高控制在每年 0.8% 以下的举措和观念), 很可能使经济陷入无序城市化的泥沼, 从而危害可持续发展, 正如 20 世纪 70 年代以来非洲所经历的情况所证明的 (Fay and Opal, 2000)。显然, 各国政府应该接受城市化加速的事实, 把解决农村农民农业问题的视野扩展到城市里, 积极主动地消除种种移民障碍, 并通过改善公共服务(如为移民提供住房和医疗教育)和强化基础设施建设(如修建城乡和城市道路, 补贴公共交通, 普及 Wifi 等)来应对移民潮。

参 考 文 献

- [1] Amrhein, C. G., "Interregional Labor Migration and Information Flows", *Environment and Planning A*, 1985, 17 (8), 1111-1126.
- [2] Banerjee, B., "Information Flow, Expectations and Job Search. Rural-to-Urban Migration Process in India", *Journal of Development Economics*, 1984, 15 (1-3), 239-257.
- [3] Barnett, V., and T. Lewis, *Outliers in Statistical Data*. Chichester: John Wiley, 1994.
- [4] Bairoch, P., *Cities and Economic Development: From the Dawn of History to the Present*. Chicago: University of Chicago Press, 1988.
- [5] Barrios, S., L. Bertinelli, and E. Strobl, "Climatic Change and Rural-Urban Migration: The Case of Sub-Saharan Africa", *Journal of Urban Economics*, 2006, 60 (3), 357-371.
- [6] Beals, R. E., M. B. Levy, and L. N. Moses, "Rationality and Migration in Ghana", *Review of Economics and Statistics*, 1967, 49 (4), 480-486.
- [7] Bhuller, M., T. Havnes, E. Leuven, and M. Mogstad, "Broadband Internet: An Information Superhighway to Sex Crime?", *Review of Economic Studies*, 2013, 80 (4), 1237-1266.
- [8] Black, D., and V. Henderson, "A Theory of Urban Growth", *Journal of Political Economy*, 1999, 107 (2), 252-284.
- [9] Bloom, D. E., D. Canning, G. Fink, T. Khanna, and P. Salyer, "Urban Settlement: Data, Measures, and Trends", In: Beall, J., B. Guha-Khasnobis and R. Kanbur (eds.), *Urbanization and Development*. Oxford: Oxford University Press, 2010.
- [10] Carrington, W. J., E. Detragiache, and T. Vishwanath, "Migration with Endogenous Moving Costs", *American Economic Review*, 1996, 86 (4), 909-930.
- [11] Comin, D., B. Hobijn, and E. Rovito, "Five Facts You Need to Know about the Technology Diffusion", 2006, NBER Working Paper. URL: <http://www.nber.org/papers/w11928>, 访问日期: 2018 年 10 月 1 日。
- [12] Czernich, N., O. Falck, T. Kretschmer, and L. Woessmann, "Broadband Infrastructure and Economic Growth", *The Economic Journal*, 2011, 121, 505-532.
- [13] Davis, J. C., and V. Henderson, "Evidence on the Political Economy of the Urbanization Process", *Journal of Urban Economics*, 2003, 53 (1), 98-125.
- [14] Fay, M., and C. Opal, "Urbanization without Growth: A Not-So-Uncommon Phenomenon", *The World Bank Policy Research Working Paper*. <https://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-2412>, 访问日期: 2018 年 10 月 1 日。
- [15] Field, A., *Discovering Statistics Using SPSS*. London: SAGE, 2009.

- [16] Gaspar, J., and E. Glaeser, "Information Technology and the Future of Cities", *Journal of Urban Economics*, 1998, 43 (1), 136-156.
- [17] Geroski, P., "Models of Technology Diffusion", *Research Policy*, 2000, 29 (4-5), 603-625.
- [18] Gollin, D., R. Jedwab, and D. Vollrath, "Urbanization with and without Industrialization", *Journal of Economic Growth*, 2016, 21 (1), 35-70.
- [19] Greenwood, M. J., "Internal Migration in Developed Countries", *Handbook of Population and Family Economics*, 1997. Vol. 1B. 647-720.
- [20] Griliches, Z., "Hybrid Corn: An Exploration in the Economics of Technological Change", *Econometrica*, 1957, 25 (4), 501-522.
- [21] Harris, J. R., and M. P. Todaro, "Migration, Unemployment and Development: A Two-Sector Analysis", *The American Economic Review*, 1970, 60 (1), 126-142.
- [22] Hofmann, A., and G. Wan, "Determinants of Urbanization", 2013, *ADB Economics Working Paper*. <http://hdl.handle.net/11540/2310>, 访问日期: 2019年10月1日。
- [23] Jedwab, R., L. Christiaensen, and M. Gindelsky, "Demography, Urbanization and Development: Rural Push, Urban Pull and... Urban Push?", *Journal of Urban Economics*, 2017, 98, 6-16.
- [24] Keyfitz, N., "Do Cities Grow by Natural Increase or by Migration?", *Geographical Analysis*, 1980, 12 (2), 142-156.
- [25] Krugman, P., *Geography and Trade*. Cambridge: MIT Press, 1991.
- [26] Krugman, P., and R. L. Elizondo, "Trade Policy and the Third World Metropolis", *Journal of Development Economics*, 1996, 49 (1), 137-50.
- [27] Lucas, R. E., "On the Mechanics of Economic Development", *Journal of Monetary Economics*, 1988, 22 (February), 3-42.
- [28] Lucas, R. E., "Life Earnings and Rural-Urban Migration", *Journal of Political Economy*, 2004, 112 (1), 29-59.
- [29] 罗知、万广华、张勋、李敬, "兼顾效率与公平的城镇化: 理论模型与中国实证", 《经济研究》, 2018年第7期, 第89—105页。
- [30] Marshall, A., *Principles of Economics*. London: Macmillan, 1890.
- [31] Massey, D. S., J. Arango, G. Hugo, A. Kouaouci, A. Pellegrino, and J. E. Taylor, "Theories of International Migration: A Review and Appraisal", *Population and Development Review*, 1993, 19 (3), 431-466.
- [32] Munshi, K., "Networks in the Modern Economy: Mexican Migrants in the U. S. Labor Market", *The Quarterly Journal of Economics*, 2003, May, 549-599.
- [33] Moomaw, R. L., and A. M. Shatter, "Urbanization and Economic Development: A Bias toward Large Cities?", *Journal of Urban Economics*, 1996, 40 (1), 13-37.
- [34] OECD, "Appendix B: World Population, GDP and GDP Per Capita Before 1820", in *The World Economy: Volume 1: A Millennial Perspective and Volume 2: Historical Statistics*, Paris: OECD Publishing, 2006.
- [35] Preston, S. H., "Urban Growth in Developing Countries: A Demographic Reappraisal", *Population and Development Review*, 1979, 5 (2), 195-215.
- [36] Rauch, J. E., "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities", *Journal of Urban Economics*, 1993, 34 (3), 380-400.
- [37] Sahota, G. S., "An Economic Analysis of Internal Migration in Brazil", *Journal of Political Economy*. 1968, 76 (2), 218-245.
- [38] Schwartz, A., "Interpreting the Effect of Distance on Migration", *Journal of Political Economy*, 1973, 81 (5), 1153-1169.

- [39] Sjaastad, L. A., “The Costs and Returns of Human Migration”, *Journal of Political Economy*, 1962, 70 (5), 80-93.
- [40] Todaro, M. P., “A Model of Labor Migration and Urban Unemployment in Less Developed Countries”, *The American Economic Review*, 1969, 59 (1), 138-148.
- [41] United Nations, *World Urbanization Prospects: The 2014 Revision, Highlights*. 2014. <https://www.un.org/en/development/desa/publications/2014-revision-world-urbanization-prospects.html>, 访问日期: 2019年10月1日。
- [42] Vanderkamp, J., “Migration Flows, Their Determinants and the Effects of Return Migration”, *Journal of Political Economy*, 1971, 79 (5), 1012-1031.
- [43] 万广华, “城市化与中国的减贫和不平等”, 《比较》, 2008年第36辑, 第160—172页。
- [44] 万广华, “2030年: 中国城镇化率达到80%”, 《国际经济评论》, 2011年第6期, 第99—111页。
- [45] 万广华, “城镇化与不均等: 分析方法和中国案例”, 《经济研究》, 2013年第5期, 第73—86页。
- [46] WHO, “Unmasking and Overcoming Health Inequities in Urban Settings”, *Evolutionary Ecology*, 2010, 10, 1-145.
- [47] Winters, P., E. Sadoulet, and A. de. Janvry, “Family and Community Networks in Mexico-U. S. Migration”, *The Journal of Human Resources*, 2001, 36 (1), 159-184.
- [48] 章元、万广华, “国际贸易与发展中国家的城市化: 来自亚洲的证据”, 《中国社会科学》, 2013年第11期, 第65—84页。

Accelerating Urbanization Explained —The Role of Information

GUANGHUA WAN

(*Fudan University*)

YAN ZHANG*

(*Chongqing Technology and Business University*)

Abstract In this paper, we reveal and substantiate that urbanization has accelerated in recent decades. We propose a simple two-sector general equilibrium model, demonstrating the potential effect of information on the urbanization rate. The theoretical prediction is then verified by employing cross-country panel regressions using 1960 - 2013 data from 109 non-OECD countries. Results confirm that: (1) the conventional drivers of urbanization cannot account for this newly discovered phenomenon; (2) Internet penetration as a proxy of information can explain much of the acceleration trend in urbanization; and (3) our findings are robust to different control variables, different regions and different information proxies.

Keywords urbanization acceleration, information, migration Cost

JEL Classification R11, R23, R58

* Corresponding Author: Yan Zhang, School of Economics, Chongqing Technology and Business University, Institution Avenue 19, Nan'an District, Chongqing, 400067, China; Tel: 86-13817512257; E-mail: phd09010@grips.ac.jp.