

跨越省际移民中的文化壁垒： 信息沟通与身份认同

李仲达 林建浩 邓虹*

摘要 中国存在源远流长、各具特色的地域文化，文化差异是否成为影响劳动力跨区域流动的隐形壁垒？本文利用基因、姓氏和方言等群体特征数据，通过主成分方法构建省际文化差异的度量指标。研究表明：在控制地理和经济因素之后，文化差异对 2000—2010 年的省际移民具有显著的阻碍效应，并且这种阻碍效应随着时间的推移逐渐减缓，工具变量估计等稳健性分析均支持这一发现。进一步分析显示，信息沟通与身份认同是文化差异影响省际移民的两个主要机制。

关键词 文化差异，信息沟通，身份认同

DOI: 10.13821/j.cnki.ceq.2021.05.09

一、引 言

这是一个流动的时代，一个迁徙的中国。大规模的劳动力跨地区流动是中国改革开放以来最为重要的经济事件之一，这是在城乡制度障碍、地区市场分割的背景下所发生的，对于中国经济的起飞、区域结构调整以至全球经济的重新布局都有重要影响（陆铭，2011）。据《中国流动人口发展报告 2016》显示，2015 年中国流动人口¹规模达 2.47 亿人，占总人口的 18%，而且越来越多的人选择“离土又离乡”的远距离流动，跨省流动比例逐渐上升，东部地区的流动人口更是以跨省流动为主。从长期来看，随着制度性市场分割问题受到政策制定者的重视，制度问题将逐步得到解决，但是一些非正式

* 李仲达，暨南大学经济学院；林建浩，中山大学岭南学院；邓虹，荷兰鹿特丹伊拉斯姆斯大学经济学院和丁伯根研究所。通信作者及地址：林建浩，广东省广州市海珠区新港西路 135 号中山大学岭南行政中心，510275；电话：(020) 84112104；E-mail: linjh3@mail.sysu.edu.cn。本文得到国家自然科学基金 (71703057、71773147、72073148)、广东省基础与应用基础研究基金 (2020A1515011481、2021A1515011659) 以及中央高校基本科研业务费专项资金 (19JNQM13、19wkzd09) 的资助。作者特别感谢主编与匿名审稿专家的宝贵意见。当然，文责自负。

¹ 在讨论中国城乡和地区间劳动力流动问题时，“流动” (mobility) 和“迁移” (migration) 已经成为相互混用的词汇。国际学术界已经习惯于将中国的移民用以指未获得城市户籍的短期移民（主要指在城市里每年居住达到半年以上的常住人口），相比之下，那些已经获得本地城镇户籍的外来人口反而不被作为“移民”来对待（陆铭，2011）。

制度的因素对于市场整合的影响应该尽早引起关注,那就是文化(陆铭,2017)。

文化作为移民研究中一个历久弥新的话题,一直被认为是影响迁移的重要因素。尤其是迁入地与迁出地之间的文化差异可能增加迁移预期成本进而阻碍两地间人口流动,而迁入地与迁出地之间的文化趋同则有助于降低因语言、习俗等差异造成的迁移风险(Taylor, 1986; Akerlof, 1997)。文化壁垒背后的机制既包括文化差异所带来的信息沟通成本,又涉及由身份认同和群体识别所导致的偏见与信任缺失,例如以色列与阿拉伯国家之间的移民、冷战时期东西方国家之间的人口流动。

中国幅员辽阔,地域文化具有多样性,同时历史又具有连续性和稳定性,为考察文化变迁提供了较好的条件(李楠,2015)。中国地域文化的划分有多个层级,如一级文化区系划分以空间地域而定的“东西”“南北”文化格局:东部是以平原丘陵为主的发达农业区,西部是以沙漠高原为主的游牧区,形成农耕和游牧文明之别;南北文化的划分则更具人文色彩,自古以来南北文化分野便出现在政治、经济、思想、人文、学术等部类中,古时王夫之言“三代以上,淑气聚于北,而南为蛮夷”,发展到东晋时“北强南秀”,而今南北方各自发展出了成熟的经济格局。随着制度文化的发展,地方行政设置使得区域文化越来越依附于这种行政区划,使得文化划分得以在地图上呈现,文化汇聚地与地域联系更为密切,如岭南文化、齐鲁文化这些区域文化对应的地带更为明晰。国内不少文献对比了我国各个地区在文化价值观和文化习俗方面的异同,构造出中国文化地图(王慧,2004;崔向红,2009)。

地域文化差异广泛存在于生活方式、生产方式、思想意识、风俗习惯等多个方面,再加上经济水平发展的差异、社会刻板印象、从众心理等的综合作用,地域歧视成为普遍的社会文化现象。中国的地域歧视具有悠久的历史根源,《晏子春秋》记载晏子使楚时,楚王嘲讽齐国道“齐人固善盗乎”,便带着地域歧视的偏见;东汉时期,中原地区称呼黄河流域以北的游牧民族为“北夷”,南方未开发的地区为“蛮”;到了明清,满族人称与其反抗的汉人为“南蛮子”。现如今,地域群体偏见或是刻板印象不仅在空间上普遍存在,而且长期来看难以改变甚至趋于恶化。

那么,文化差异是否成为阻碍省际劳动力自由流动与优化配置的隐形壁垒?其作用机制是什么?本文致力于探究这些问题。由于文化的准确内涵难以把握,导致文化差异的量化成为制约实证研究的障碍。不少文献认为基因是合适的变量,赵子乐和林建浩(2017)则认为在儒家文明圈中,基因与文化的对应关系由于文化横向传播而被打乱,方言作为区域文化的载体能兼顾垂直传递和横向传播,进而采用方言距离测度文化差异。李楠(2015)和陈永伟(2016)则分别基于古代和现代人口的姓氏构造姓氏距离测度省际文化差异。本文综合利用遗传学和语言学的研究成果,构建一个囊括基因、姓氏

和方言的省际群体特征数据库，并通过主成分方法获得省际群体文化差异的综合测度。基于此的实证研究发现，在控制地理和经济因素之后，文化差异对省际移民仍然具有显著的阻碍效应。

尽管已有文献从理论上讨论文化差异对移民阻碍效应的作用机制，但经验证据尚少。本文研究表明，文化差异主要通过信息沟通和身份认同这两个机制产生影响。对于信息沟通而言，语言的使用至关重要，一方面，共同语言的普及有助于消除文化差异所造成的沟通与信息传递障碍，因此迁入地使用普通话比例越高，则文化差异的阻碍越小；另一方面，相比其他方言，官话与普通话更为接近，以官话群体为主体的迁入地具有更高的沟通便利性，文化差异对于跨省移民的阻碍效应更弱。对于身份认同机制，已有文献表明，具有水稻种植传统的地区更具协作与共同利益并存的集体主义心理，户籍政策宽松的地区对外来移民更具开放和包容理念，这些融合的力量有助于克服文化壁垒所带来的身份认同与信任问题，本文的实证结果也表明文化差异在这些地区的阻碍效应更弱。

剩余内容结构如下：第二部分总结讨论文化差异与移民的文献；第三部分介绍数据来源以及关键变量的构造；第四部分验证文化差异对省际移民的阻碍作用；第五部分进一步探讨影响机制；最后是结论性评述。

二、文献综述

文化差异一直是学者们研究社会经济现象的重要视角。Spolaore and Wacziarg (2009) 通过一个简单的理论框架，首次刻画了群体间遗传距离与文化差异的关联，并论证了各个地区与技术前沿的文化差异对于技术扩散的阻碍效应。对于中国而言，无论是省级数据，还是市级数据，都有相似的发现，即族群文化差异对当前地区间经济发展差异有较强的解释力（赵子乐和林建浩，2017；李楠和林友宏，2019）。而且，文化差异与行政分割并存形成了一种特殊的区域文化分割现象，高翔和龙小宁（2016）的研究发现省际行政区划造成的文化分割会造成当地经济损失，并且这种负面影响随着邻省地方保护主义的提高而加重，随着各地区被划分进当前省份的历史延长而减轻。

文化差异影响发展扩散的重要机制在于阻碍创新、技术以及生产率的传播。文化差异导致不同群体之间存在沟通障碍、隔阂、歧视、偏见等，这阻碍了世界范围内的技术传播（Spolaore and Wacziarg, 2011）。对于中国而言，Bai and Kung (2014) 发现文化差异对于十七八世纪玉米作物的传播以及清末民初蒸汽机的传播有显著的阻碍作用。阮建青等（2016）发现省际基因距离并非通过智商上的差异导致了创新差异，而是通过影响人群交流从而影响技术创新差异。林建浩和赵子乐（2017）发现在同等条件下，更高的人力资本

水平、更高的外来移民人口比例、更多的外国直接投资能削弱文化差异的阻碍效应；城市历史上对外开放和采用新技术的经验，同样有助于削弱文化差异这一技术扩散的隐形壁垒。刘毓芸等（2017）使用生产率差距度量资源错配，实证研究发现方言上的不同显著增强了相邻两县间的资源错配。

族群文化差异也可能影响制度的复制和扩散，从而造成经济发展的差异。例如，Lecce and Ogliari（2019）针对法国占领下的普鲁士的研究发现，与法国文化更为接近的地区更容易移植法国的制度，并具有更好的经济发展表现。阮建青和王凌（2017）发现省际语言差异除了会因为交流的障碍增加市场制度差异外，还会通过文化观念差异进而影响市场制度差异。林建浩和赵子乐（2017）将文化差异、制度扩散与技术溢出联系起来，发现文化差异阻碍了技术从前沿地区向其他地区扩散，并且制度是这一过程的重要中介变量。

除了经济、技术、制度以外，人类活动也会受到文化差异的影响。其中最为典型的是人口大规模迁移，族群间文化差异也被发现对人口迁移有着显著影响（Adserà and Pytliková, 2015；李楠，2015）。Falck *et al.*（2012）发现历史形成的德语方言相似性会对当前的人口迁移产生显著影响，随着标准德语的广泛推广，方言差异并不影响沟通，更多体现的是持久的文化差异。同样是基于文化差异的角度，Gong *et al.*（2011）讨论了中国地区之间的异质文化，并研究方言对劳动者与雇主之间匹配质量的影响。Chen *et al.*（2014）基于上海市外来人口调查的研究发现，在人们普遍可以用普通话进行交流的背景下，本地方言能力成为一种身份显示机制，虽然听上海话的能力没有显著作用，但说上海话的能力却显著影响受访者的收入和就业。李秦和孟岭生（2014）以及刘毓芸等（2015）从方言的角度研究中国的劳动力区域流动，前者基于引力模型发现劳动力倾向于流动到普通话沟通障碍小、拥有共同文化背景的地方工作，后者则是基于中国劳动力动态调查（CLDS）微观数据发现方言距离呈现出先促进、后抑制劳动力流动的倒 U 形模式，其中认同效应是导致两者负相关的核心机制。Krieger *et al.*（2018）利用基因距离作为国家的文化距离，研究文化差异如何影响不同技能水平的劳动力的移民决策。李楠（2015）利用宋明两代传记资料中的姓氏信息分别计算各省之间的姓氏距离以及移民率，发现以姓氏距离度量的地区间文化差异对古代的移民行为具有较强的负向影响。

与已有文献相比，本文可能的边际贡献在于：第一，综合利用基因、姓氏和方言这些具有垂直传递与分化特征的文化载体，构建一个刻画省际群体文化差异的综合测度。第二，对于文化差异的信息沟通机制，本文利用迁入地普通话普及比例和官话方言片区地域分布来检验信息沟通机制，则发现普通话的推广能够缓解文化差异对于移民的阻碍效应，而且随着时间推移，这种作用还在显著增强。第三，对于身份认同机制，Falck *et al.*（2012）和 Chen *et al.*（2014）都是先验地认为方言不影响现代经济活动所需的各种沟

通，因此直接推断是认同机制起作用，并没有对认同机制进行严谨的实证验证。文化经济学实证研究的一个趋势就是不直接检验观念，而是使用影响观念形成的外生、久远的地理和历史变量进行间接检验。为此，本文采用粮食历史种植等刻画迁入地居民的认同与包容性，进而验证了身份认同机制。

三、数据来源与描述统计

（一）省际移民率

本文所使用的省际移民率基于2000—2010年的全国人口普查数据计算得出，其中2005年为全国1%人口抽样调查数据，涵盖了中国31个省级行政区。这三次人口普查项目是在我国进入社会主义市场经济之后进行的，能够综合反映我国市场化改革之后所发生的人口变迁状况。在普查数据中，可以将迁移人口划分为省内迁移与省外迁移两种类型。根据普查数据统计显示，跨省迁移人口的比例约为30%，而省内迁移的比重约为70%，说明对于迁移个体而言，跨省流动面临的迁移成本更高，迁移决策所考虑的因素更为复杂。从趋势来看，跨省迁移人口在全国总人口中比例不断上升，跨省迁移的人口规模也从2000年的4 242万人上升到了2010年8 588万人；从迁移方向来看，跨省迁移主要集中流向东部沿海地区。

在研究移民问题的文献中，常将迁移定义为“是否取得迁入地户籍”“迁移后在迁入地的定居时间长达特定时限”等。本文的“流动”与“移民”均指劳动力发生迁移，包括人户分离、现住地与五年前常住地不同、现住地与出生地不同等情况，对“劳动力迁移”与“人口流动”不进行严格区分。本文所使用的移民率指标按五年前常住地的口径进行统计，计算方法是“ i 省迁入 j 省的人数占 i 省总人口的比例”。本文在实证分析部分，也尝试更换了其他口径计算的移民率指标，并不改变文章的结论，实证结果稳健。

（二）省际文化差异

文化作为一种非物质因素，会通过人类的代际传递与演化，对经济与社会发展产生长期影响。这一观点已逐渐成为学界共识，并被许多经验研究所反复验证（Bai and Kung, 2011, 2014; Cavalli-Sforza *et al.*, 1994; Desmet *et al.*, 2011; Krieger *et al.*, 2018; Spolaore and Wacziarg, 2009, 2013）。本文通过基因、姓氏和方言三个维度的信息测度群体特征。

文化差异和基因差异都是垂直传递的，族群的文化或基因特征通过跨代传递得以继承。其中，基于血型的基因距离测度应用最广。人类红细胞血型有近20种系统，最重要的是ABO血型系统。ABO血型位点是由O、A和B三个等位基因所组成，在人体中表现出O、A、B和AB四种血型。ABO血

型在全国分布是不均匀的,存在地区上的差异。本文利用中国科学院的数据 (Du *et al.*, 1992), 共计有效样本 523 749 份, 参考 Cavalli-Sforza and Feldman (1981) 的方法根据血型计算出各省之间的基因距离。同样采用中国科学院的姓氏分布数据, 本文根据 $D(p, q) = -\log\left(\sum_{s=1}^S p_s q_s\right)$ 测度两个群体之间的姓氏距离, 其中 p_s 和 q_s 分别表示群体 p 和群体 q 各自第 s 个姓氏的频率, S 表示姓氏种类的总数。因此, $\sum_s p_s q_s$ 度量了群体 p 和群体 q 之间的同姓率, 指标 $D(p, q)$ 则刻画了群际姓氏距离。方言距离的计算来自刘毓芸等 (2015), 具体度量规则如下: 当两个地级市属于同一个方言片时, 方言距离为 0; 属于同一方言区的不同方言片时, 方言距离为 1; 属于同一方言大区的不同方言区时, 方言距离为 2; 属于不同方言大区时, 方言距离为 3; 在此基础上, 采用各地级市 2000 年的人口占比, 利用加权方法得到省份之间的方言距离。

最后, 为了能够同时囊括基因、姓氏和方言这三个维度的信息, 本文采用主成分方法, 构建出一个度量文化差异的综合指标——文化距离。主成分方法是提炼数据有效信息的常用统计方法, 通过将原始变量转换为变量的线性组合, 在保留数据主要信息的基础上, 达到数据降维的目的。

(三) 其他控制变量

本文还控制了影响移民跨省迁移决策的其他因素, 主要包括: 地理距离、是否接壤、经纬度差异与经济发展水平差异。现有研究移民问题的主流文献普遍认为, 地理距离是阻碍移民的关键因素, 常见的度量指标包括球面距离与公路距离。通过省会城市或直辖市的经度和纬度信息, 可以计算出两地之间的球面距离; 而公路距离, 则主要通过“搜狗地图搜索引擎”获得省会城市的最短公路距离并取对数。除了地理距离以外, 本文还采用省份地区是否接壤与经纬度来刻画其区位特征。对于两个省份是否接壤的信息, 可通过虚拟变量的方式进行定义, 若两个省份交界则记为 1, 否则记为 0 (海南省定义为只与广东省相邻)。经纬度差距可借由 Google 地图数据获得, 利用各省会城市、直辖市政府所在地标进行计算。

许多研究认为, 工资差异或地区间发展水平差异是吸引劳动力流动的首要经济动因。为控制迁入地与迁出地的经济发展水平差异, 本文使用各省份地区五年人均 GDP 的差值进行度量, 以此刻画个体选择跨省移民可能增加的经济收益机会。实证分析中还考虑了迁入地开放程度和教育水平的影响, 分别以每年接待国际游客数量和每十万人中接受高等教育的人口数量进行度量; 此外, 还加入了迁出地不确定性规避程度与迁入地人际关怀导向程度, 这两类指数来源于赵向阳等 (2015), 不确定性规避指数刻画了移民做出迁移决策

时的心理成本，而人际关怀指数度量了移民可以被当地居民接受的程度。表 1 列示了变量的描述性统计。

表 1 描述性统计

变量名称	样本量	均值	标准差	最小值	最大值
移民率 ($\times 10^{-4}$)	2 790	34.62	150.5	0.000	3 754
文化距离	2 268	0.000	1.396	-3.967	3.095
球面距离 (千米)	2 790	1 370	722.3	103.6	3 558
公路距离 (千米)	2 610	1 741	978.3	139.1	4 901
是否接壤	2 790	0.151	0.358	0	1
经度差异	2 883	9.640	7.917	0.000	39.04
纬度差异	2 883	7.546	5.476	0.000	25.72
五年人均 GDP 差值 (千元/人)	2 883	0.000	6.875	-31.10	31.10
迁入地年均接待外国游客数量 (百万人次)	2 883	1.605	3.409	0.010	27.48
迁入地每十万人中高等学校人数 (千人)	1 922	1.944	1.142	0.550	6.410
迁出地不确定性规避指数	2 883	7.335	2.354	0.380	11.80
迁入地人文关怀指数	2 883	4.328	0.089	4.120	4.560

注：基因距离与姓氏距离只包含了 28 个省份的数据，文化距离借助主成分方法对基因、姓氏和方言距离三个维度进行提取，因此数据量为 $(28 \times 27) \times 3 = 2 268$ ；公路距离缺失海南省的数据，因此数据量为 $(30 \times 29) \times 3 = 2 610$ ；高等学校人数仅含 2005、2010 年两年的数据；五年人均 GDP 差值的均值之所以为零，是因为样本中同时存在两地相互作差的数据点，如“北京—天津”与“天津—北京”。

四、文化差异对于跨省移民的阻碍效应

(一) 基准回归分析

本文使用如下基准回归方程，考察文化差异对移民决策的影响：

$$\log P_{ijt} = d_{ij}\beta + z'_{ijt}\gamma + \mu_i + \lambda_j + \eta_t + e_{ijt},$$

其中，被解释变量是省际移民率 P_{ijt} 的对数， μ_i 、 λ_j 分别表示迁出地和迁入地固定效应， η_t 表示年份固定效应， d_{ij} 是核心解释变量省际文化差异， z_{ijt} 是其他控制变量， e_{ijt} 是随机扰动项。本文首先使用文化距离的综合指标对省际文化差异进行测度，估计结果如表 2 所示。第 (1) 列只加入文化距离作为解释变量，估计结果表明文化距离越大的两个省份，这两个省份之间的人口迁移数越少。第 (2) 和 (3) 列进一步控制区位特征和经济发展状况，文化距离

对跨省移民的阻碍效应依旧显著。此外, 区位特征控制变量的估计结果符合预期, 而经济发展水平差异的影响为负, 但系数绝对值不大²。表 3 进一步报告了 2000、2005 和 2010 年三个截面回归的结果, 文化距离的阻碍效应应在三个时期依旧存在, 但是呈现出随时间逐渐减弱的态势。相比 2000 年, 文化差异对 2010 年跨省移民的阻碍效应下降约 40%。

表 2 基准回归: 文化距离与跨省移民

解释变量	被解释变量: 移民率			
	(1)	(2)	(3)	(4)
文化距离	-0.653*** (0.026)	-0.172*** (0.031)	-0.172*** (0.031)	-0.172*** (0.051)
地理距离		-0.205** (0.092)	-0.205** (0.092)	-0.205** (0.091)
是否相邻		0.650*** (0.071)	0.650*** (0.071)	0.650*** (0.111)
经度差异		-0.023*** (0.005)	-0.023*** (0.005)	-0.023*** (0.008)
纬度差异		-0.082*** (0.007)	-0.082*** (0.007)	-0.082*** (0.013)
经济发展水平差异			-0.006* (0.004)	-0.006 (0.007)
迁出地固定效应	是	是	是	是
迁入地固定效应	是	是	是	是
年份固定效应	是	是	是	是
R ²	0.785	0.845	0.846	0.846
样本量	2 267	2 267	2 267	2 267

注: 括号中报告的是 White 异方差稳健标准误; 第 (4) 列使用了“迁出地—迁入地”双维聚类标准误, 估计结果与第 (3) 列的非聚类稳健标准误相似; **、* 和 * 分别表示结果在 1%、5% 和 10% 的显著性水平下显著。此外地理距离使用球面距离指标, 使用公路距离的估计结果相似, 下同。

² 直观上, 经济发展水平差异越大的省份, 越能够吸引跨省移民。表 2 得到的是负向的估计结果, 在后文各类回归模型中, 经济发展水平差异的系数有正有负, 数值较小且大多不显著。为理解这一点, 本文尝试在不引入地区固定效应的设定下, 发现经济发展水平差异的作用在 1% 显著性水平下显著为正; 一旦引入地区固定效应, 这一正向作用就消失。因此, 可以理解为, 对中国跨省流动的移民群体而言, 其迁移决策带有明显的地域偏好, 而经济差异的作用也体现在这种相对持久的地域偏好中。

表 3 逐年截面回归

解释变量	被解释变量：移民率		
	2000 年	2005 年	2010 年
	(1)	(2)	(3)
文化距离	-0.203***	-0.190***	-0.122***
	(0.053)	(0.058)	(0.044)
控制变量	是	是	是
迁出地固定效应	是	是	是
迁入地固定效应	是	是	是
R ²	0.814	0.802	0.837
样本量	756	755	756

(二) 稳健性分析与内生性问题

为进一步验证文化差异的阻碍效应，本文还进行了全面的稳健性分析，包括改变固定效应设定方式，更换移民率的计算方式，更换文化差异的度量指标，分别利用基因、姓氏和方言距离三个不同变量进行考察，相应的估计结果都是稳健的。³

文化差异对省际移民的阻碍效应识别还可能会受到内生性问题的干扰。一方面，移民会带来文化扩散与融合，即存在移民对文化距离的反向因果作用。然而，本文使用的基因、姓氏、方言等形成时间久远，当前的移民不会影响历史久远的指标，因而反向因果关系所导致的内生性问题比较微弱。另一方面，早期的移民活动与文化差异有关，而移民历史与移民现状通常存在一定相关性，因此可能存在遗漏变量偏误。为此，本文分别利用宋明时期人口迁移（袁义达和张诚，2002）以及 1982—1987 年间人口迁移（杨云彦，1989）以控制早期移民历史的影响。需要说明的是，袁义达和张诚（2002）使用历代姓氏资料对人口迁移历史特征进行测算是一种通用做法（Bodmer and Cavalli-Sforza, 1968; Wijsman *et al.*, 1984; Du *et al.*, 1992）。表 4 结果显示，控制早期移民历史的影响之后，文化距离的阻碍效应仍然显著，而且系数与表 3 的基准估计结果较为接近。

除了早期人口迁移，当代移民活动还可能与其他地区特征相关。我们进一步增加有关迁入地对外开放、教育程度以及文化特征等控制变量，引入产业结构、户籍政策、就业人口分布这三方面因素的水平值，以及迁入地与迁出地的相对差值。其中，产业结构特征使用地区第二、三产业的产值占比来

³ 由于篇幅所限，具体结果留存备索。

衡量, 户籍政策使用外来人口在地区落户的难易程度进行刻画⁴, 就业人口分布使用就业人员在不同产业的分布状况、就业人员总数两方面来反映。结果显示, 文化距离对于跨省移民的阻碍效应均保持显著为负。⁵

表 4 控制早期人口迁移特征

解释变量	被解释变量: 移民率	
	宋明时期人口迁移特征	1982—1987 年间人口迁移特征
	(1)	(2)
文化距离	-0.171*** (0.031)	-0.144*** (0.029)
早期人口迁移	-0.689 (1.030)	0.089*** (0.009)
控制变量	是	是
迁出地固定效应	是	是
迁入地固定效应	是	是
年份固定效应	是	是
R ²	0.846	0.856
样本量	2 267	2 264

尽管前文控制了诸多因素, 仍可能受到其他不可观察遗漏因素的影响。为此, 我们参考李楠和林友宏 (2019) 的做法, 使用历史时期的姓氏距离作为工具变量。李楠和林友宏 (2019) 认为明代族群文化差异外生于影响当前经济发展的相关因素, 并且不受改革开放后大规模人口流动的影响, 因而利用明代人物姓氏数据所构造的姓氏距离可以作为工具变量。参考这一思路, 本文使用宋朝姓氏距离与明朝姓氏距离作为工具变量, 数据分别来源于袁义达等 (1999) 与袁义达等 (2000a, 2000b)。为了便于与现代姓氏研究结果相比较, 袁义达等人将宋明时期地名省份的归属按目前的行政区划进行归并, 本文则对合并省份的姓氏数据进行还原, 而其他缺失的省份样本, 则采用接壤省份的平均值进行补充。工具变量估计结果如表 5 所示, 第一阶段回归中历史姓氏距离的系数显著为正, 并且检验结果显示不存在不可识别和弱工具变量问题。工具变量估计结果同样显示存在文化距离对省际移民的阻碍效应, 这与李楠 (2015) 与李楠和林友宏 (2019) 的发现相似, 即潜在的内生性问题并未对估计结果产生实质性影响。

⁴ 本文从各大法律信息网中查找城市户籍管理的相关法律法规, 对户籍政策宽松程度进行评分, 继而获取了各城市在 2000、2005、2010 年的户籍政策得分, 并以相同省份的城市得分平均值, 作为当年该省份户籍政策宽松程度的衡量。指标得分取值高表示当地户籍政策较为宽松, 对外来移民更为友好。

⁵ 由于篇幅所限, 具体结果留存备索。

表 5 IV 估计：使用历史姓氏距离作为工具变量

解释变量	被解释变量：移民率			
	宋代姓氏距离作为工具变量		明代姓氏距离作为工具变量	
	(1)	(2)	(3)	(4)
第一阶段：历史姓氏距离	1.858*** (0.059)	1.101*** (0.058)	1.151*** (0.038)	0.696*** (0.036)
第二阶段：文化距离	-0.670*** (0.042)	-0.117* (0.070)	-0.653*** (0.034)	-0.182*** (0.061)
不可识别检验	219.61***	150.28***	274.13***	188.78***
弱识别检验	1 518.17	438.67	1 826.10	432.64
控制变量	否	是	否	是
迁出地固定效应	是	是	是	是
迁入地固定效应	是	是	是	是
年份固定效应	是	是	是	是
R ²	0.804	0.848	0.812	0.852
样本量	2 093	2 093	2 093	2 093

注：弱识别检验使用的是 Wald 检验，临界值对应的最大 IV 水平为：16.38（10%）、8.96（15%）、6.66（20%）、5.53（25%）。

五、文化差异影响移民的机制分析

既然文化差异对跨省移民有显著的阻碍效应，为了缓解这种负向影响，需要识别出阻碍效应的作用机制。本文将从信息沟通和身份认同两个角度进行分析。

（一）信息沟通机制

信息沟通机制是指迁入地的语言沟通便利程度影响着信息传递的有效性。不同地区使用不同的方言，语音、语调以及词汇使用等方面的差异影响着信息传递，即便不同方言区的人使用了共同的语言文字，使用同一套语音和词汇系统，但是对同一个词语、同一句话的理解却可能不同，因此存在实际上的沟通障碍。

为检验信息沟通机制，本文首先引入迁入地普通话比例⁶这个变量。当移民的迁入地普通话使用比例越高时，移民与迁入地居民的沟通越便利，文化差异的阻碍效应会相对减弱。在表 6 的第（1）—（3）列，我们在模型中引入

⁶ 数据来自中国语言文字使用情况调查领导小组办公室 2006 年编写的《中国语言文字使用情况调查资料》。

迁入地普通话比例与方言距离的交叉项⁷, 无论是否单独引入迁入地普通话比例, 交叉项的系数都是显著为正, 表明普通话的推广能够缓解文化差异对于移民的阻碍效应, 文化差异的信息沟通机制得到印证。进一步地, 我们还引入“方言距离×普通话比例×时间趋势”这一交互项, 反映信息沟通机制随时间的演变, 系数显著为正, 意味着文化差异在信息沟通方面的负面影响正在逐渐下降。

表 6 信息沟通机制: 普通话比例

解释变量	被解释变量: 移民率				
	方言距离与普通话比例的交叉项			引入年份交叉项	
	(1)	(2)	(3)	(4)	(5)
方言距离	-0.902*** (0.184)	-0.882*** (0.277)	-0.884*** (0.277)	-0.566*** (0.157)	-0.802*** (0.183)
迁入地普通话比例			-68.56 (48.37)		200.2*** (71.17)
方言距离×普通话比例	1.445*** (0.336)	1.408*** (0.484)	1.413*** (0.485)		
方言距离×普通话比例×时间趋势				0.391*** (0.085)	0.626*** (0.126)
控制变量	是	是	是	是	是
迁出地固定效应	是	否	否	否	否
迁入地固定效应	是	否	否	否	否
年份固定效应	是	否	否	否	否
迁出地固定效应×年份	否	是	是	是	是
迁入地固定效应×年份	否	是	是	是	是
R ²	0.841	0.673	0.673	0.674	0.675
样本量	2 607	2 607	2 607	2 607	2 607

除了普通话比例, 本文进一步借助是否属于官话方言片区来进一步验证信息沟通机制。官话是汉语诸方言中地域分布最广的一种, 而且与普通话最为接近, 主要分布在中国北方大多数地区, 另外还包括云南、贵州、四川、重庆等西南地区省份。如果信息沟通机制的确存在, 那么在官话方言片区内部的沟通成本相对更低, 意味着文化差异阻碍效应的影响会相对更小。因此, 本文引入一个虚拟变量, 用以表示迁入地是否属于官话方言片区⁸, 并将虚拟

⁷ 由于这里考察的是信息沟通机制, 因此不采用文化距离的指标, 而采用方言距离。

⁸ 这些地区包括: 北京、天津、河北、山西、辽宁、吉林、黑龙江、安徽、山东、河南、湖北、重庆、四川、贵州、云南、陕西、甘肃、宁夏。

变量与方言距离相乘构造交叉项。表 7 的回归结果显示交叉项的系数均显著为正，表明官话方言片区内部的阻碍效应相对更弱，再次论证了文化差异阻碍效应的信息沟通机制。

表 7 信息沟通机制：官话方言片区

解释变量	被解释变量：移民率				
	方言距离与虚拟变量的交叉项			引入年份交叉项	
	(1)	(2)	(3)	(4)	(5)
方言距离	-0.297*** (0.105)	-0.295** (0.138)	-0.294** (0.137)	-0.254* (0.134)	-0.289** (0.134)
官话方言片区			62.89*** (16.60)		83.38*** (17.41)
方言距离×官话方言片区	0.277*** (0.065)	0.276*** (0.090)	0.275*** (0.089)		
方言距离×官话方言片区× 时间趋势				0.092*** (0.032)	0.133*** (0.034)
控制变量	是	是	是	是	是
迁出地固定效应	是	否	否	否	否
迁入地固定效应	是	否	否	否	否
年份固定效应	是	否	否	否	否
迁出地固定效应×年份	否	是	是	是	是
迁入地固定效应×年份	否	是	是	是	是
R ²	0.841	0.673	0.675	0.673	0.675
样本量	2 607	2 607	2 607	2 607	2 607

(二) 身份认同机制

文化差异除了影响信息沟通，另一个重要的机制是影响身份认同，而身份认同的差异则会导致偏见、隔阂、歧视等，最终阻碍技术扩散、国际贸易等经济交往以及群际信任等社会融合。身份认同理论认为个体通过社会分类，对自己的群体产生认同，并产生内群体偏好和外群体偏见（张莹瑞和佐斌，2006）。社会群体身份及其表征有些是与生俱来的，如种族、民族、性别等，有些是后天形成的，例如对文化的认同、某一群体的社会经济地位等（辛素飞等，2013）。Kramer（1999）以及 Tanis and Postmes（2005）等研究认为社会群体身份信息是一个重要的线索，影响着人们对其他群体可信性的感知和判断，进而诱发信任或不信任的倾向。林建浩等（2018）以基因距离和姓

氏距离测度省际文化差异,发现作为影响群际接触的深层因素,文化差异对省际双边信任的建立起阻碍作用。对于身份认同机制,我们无法直接测度各个省份对其他省份的认同度,但对于那些更具合作和包容精神的地区而言,文化差异对身份认同的阻碍效应更低,更可能接纳和包容外省移民。

“一方水土养一方人”,长期生活的地域不同,会导致人们形成差异化的行为特征。Talhelm *et al.* (2014) 的研究发现,相比起小麦种植地区,以水稻种植为主的地区,其居民更容易与他人开展合作,包容性更强。借鉴他们的做法,本文采用地区的水稻历史亩产量以及水稻历史种植面积比例两个指标来测算粮食种植所反映的居民包容性特征,数据来源于 Perkins (1969) 整理的 1931—1937 年各省区农业发展状况的资料⁹。同时,本文还利用户籍政策的宽松程度来反映当地政府对外来移民的户籍管制程度,如果某地区的落户限制较为宽松,则移民在当地的融入感更强,身份认同度也更高。估计结果如表 8 至表 10 所示,粮食历史种植与户籍管制状况的交叉项均显著为正(表 10 的个别交叉项系数估计值不显著,但其系数仍为正向),一致地支持了阻碍效应存在身份认同机制。由此我们可以下结论认为,迁入地对外来人口的包容性越强,文化差异对跨省移民的阻碍作用则会越低。¹⁰

表 8 身份认同机制:基于水稻历史亩产量的水稻文化测度

解释变量	被解释变量:移民率				
	文化距离与水稻亩产量交叉项			引入年份交叉项	
	(1)	(2)	(3)	(4)	(5)
文化距离	-0.370*** (0.064)	-0.371*** (0.091)	-0.371*** (0.091)	-0.270*** (0.054)	-0.261*** (0.054)
迁入地水稻亩产量			-24.91** (10.22)		-14.03 (10.42)
文化距离×水稻亩产量	0.089*** (0.020)	0.089*** (0.028)	0.089*** (0.028)		

⁹ 除此之外,我们还从赵冈等(1995)与史志宏(2012)两份资料中获取各省在清代时期的水稻亩产量;同时使用 Perkins (1969) 整理的 1914—1918 年、1957 年两个时期的水稻历史种植面积比例数据。借助这些数据,我们对表 8 和表 9 的回归结果进行重新检验,发现交叉项的系数仍然显著为正,支持结论的稳健性。

¹⁰ Talhelm *et al.* (2014) 所提出的“水稻理论”,认为水稻种植越普遍的地区,群体之间的社会合作活动越频繁,其所表现出的包容性越强。一些研究者对此表示质疑,他们发现,在考虑了样本偏差、测量误差与模型误设的问题后,“水稻理论”并不存在,地区是否种植水稻,对于塑造群体文化心理特征的影响是模棱两可的。考虑到以水稻文化测度集体主义和包容性存在争议,我们进一步从五个方面提供了身份认同机制的稳健性证据,包括迁入地的群际信任水平、人际关怀导向、集体主义文化、人格社会化特征、群体居住模式。这些指标与文化距离的交叉项系数均为正数,一致地支持了阻碍效应存在身份认同机制。

(续表)

解释变量	被解释变量：移民率				
	文化距离与水稻亩产量交叉项			引入年份交叉项	
	(1)	(2)	(3)	(4)	(5)
文化距离×水稻亩产量×时间趋势				0.026*** (0.005)	0.024*** (0.005)
控制变量	是	是	是	是	是
迁出地固定效应	是	否	否	否	否
迁入地固定效应	是	否	否	否	否
年份固定效应	是	否	否	否	否
迁出地固定效应×年份	否	是	是	是	是
迁入地固定效应×年份	否	是	是	是	是
R ²	0.845	0.664	0.665	0.666	0.666
样本量	1 781	1 781	1 781	1 781	1 781

表 9 身份认同机制：基于水稻历史种植面积的水稻文化测度

解释变量	被解释变量：移民率				
	文化距离与水稻面积比例交叉项			引入年份交叉项	
	(1)	(2)	(3)	(4)	(5)
文化距离	-0.192*** (0.035)	-0.193*** (0.050)	-0.193*** (0.050)	-0.192*** (0.046)	-0.186*** (0.046)
迁入地水稻面积比例			-62.80** (26.68)		-40.35 (27.27)
文化距离×水稻面积比例	0.158*** (0.046)	0.159** (0.067)	0.159** (0.067)		
文化距离×水稻面积比例×时间趋势				0.079*** (0.021)	0.072*** (0.022)
控制变量	是	是	是	是	是
迁出地固定效应	是	否	否	否	否
迁入地固定效应	是	否	否	否	否
年份固定效应	是	否	否	否	否
迁出地固定效应×年份	否	是	是	是	是
迁入地固定效应×年份	否	是	是	是	是

(续表)

解释变量	被解释变量：移民率				
	文化距离与水稻面积比例交叉项			引入年份交叉项	
	(1)	(2)	(3)	(4)	(5)
R^2	0.850	0.676	0.677	0.677	0.678
样本量	1 862	1 862	1 862	1 862	1 862

表 10 身份认同机制：户籍政策管制放松的视角

解释变量	被解释变量：移民率				
	文化距离与政策得分交叉项			引入年份交叉项	
	(1)	(2)	(3)	(4)	(5)
文化距离	-0.215*** (0.037)	-0.206*** (0.046)	-0.205*** (0.046)	-0.216*** (0.043)	-0.215*** (0.043)
迁入地户籍政策得分 指标			0.010 (0.017)		0.010 (0.017)
文化距离×户籍政策 得分	0.020*** (0.007)	0.015 (0.010)	0.015 (0.010)		
文化距离×户籍政策 得分×时间趋势				0.009*** (0.003)	0.009*** (0.003)
控制变量	是	是	是	是	是
迁出地固定效应	是	否	否	否	否
迁入地固定效应	是	否	否	否	否
年份固定效应	是	否	否	否	否
迁出地固定效应×年份	否	是	是	是	是
迁入地固定效应×年份	否	是	是	是	是
R^2	0.846	0.681	0.681	0.682	0.682
样本量	2 267	2 267	2 267	2 267	2 267

六、结论与启示

文化因素如何作用于区域间经济活动，是文化经济学最重要的前沿议题之一。识别文化差异的因果效应，一个首要前提是剔除制度差异对文化多样性产生的复杂影响。中国幅员辽阔、地理环境多变，即使在大一统的政治体制下仍然形成了源远流长、各具特色的地域文化，为研究文化差异提供了一

个良好的样本。本文基于2000—2010年中国省际移民的数据，系统考察了文化差异对于劳动力跨区域流动的影响及其背后的作用机制。

为了对文化差异进行合理的量化，我们综合利用遗传学和语言学的研究成果，构造一个囊括基因、姓氏和方言的省际群体特征数据库，首次通过主成分方法获得省际群体文化差异的综合测度。基于此的实证结果表明，在控制了地理和经济因素的情况下，文化差异对省际移民存在显著的阻碍效应，并且这一阻碍效应对于不同的模型设定、不同的移民率与文化差异度量指标均非常稳健。

进一步地，本文探讨了文化差异对移民阻碍效应的潜在作用机制，发现文化距离主要是通过信息沟通和身份认同两个渠道产生影响。具体而言，我们发现：第一，迁入地如果属于官话区，或者使用普通话比例越高，则文化差异的阻碍越小，这是由于共同语言沟通便利性有助于消除文化差异所造成的沟通与信息传递障碍；第二，迁入地对外来人口的认同度越高、包容性越强，文化差异对跨省移民的阻碍作用则会越低，这些融合的力量有助于克服文化壁垒所带来的身份认同与信任问题。本文的政策启示非常明显：一方面，通过提高教育水平、加大推广普通话力度等措施减少文化差异所带来的沟通障碍；另一方面，各地区应全面推进对外开放，增加文化交流与联系，形成良好的商贸合作精神。上述措施将有助于提升各区域对外来人口的包容度和认可度，从根本上消除地域偏见和隔阂，实现劳动力要素跨区域的有效配置，促进移民融合。

参 考 文 献

- [1] Adserà, A., and M. Pytliková, "The Role of Language in Shaping International Migration", *Economic Journal*, 2015, 125 (586), 49-81.
- [2] Akerlof, G., "Social Distance and Social Decisions", *Econometrica*, 1997, 65 (5), 1005-1028.
- [3] Bai, Y., and J. Kung, "Genetic Distance and Income Difference: Evidence from Changes in China's Cross-strait Relations", *Economics Letters*, 2011, 110 (3), 255-258.
- [4] Bai, Y., and J. Kung, "Does Genetic Distance Have a Barrier Effect on Technology Diffusion? Evidence from Historical China", Hong Kong University of Science and Technology, 2014, *Working Paper*.
- [5] Bodmer, W., and L. Cavalli-Sforza, "A Migration Matrix Model for the Study of Random Genetic Drift", *Genetics*, 1968, 59 (4), 565-592.
- [6] Cavalli-Sforza, L., and M. Feldman, *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, NJ: Princeton University Press, 1981.
- [7] Cavalli-Sforza, L., P. Menozzi, and A. Piazza, *The History and Geography of Human Genes*. Princeton, NJ: Princeton University Press, 1994.
- [8] 陈永伟, "文化差异对省际贸易的影响及其作用机制研究——基于姓氏距离的考察", 《经济学报》, 2016年第3期, 第1—25页。

- [9] Chen, Z., M. Lu, and L. Xu, "Returns to Dialect Identity Exposure through Language in the Chinese Labor Market", *China Economic Review*, 2014, 30, 27-43.
- [10] 崔向红,《1978—2009 中国文化地图》。广州:花城出版社,2009年。
- [11] Desmet, K., M. Breton, I. Ortuno-Ortin, and S. Weber, "The Stability and Breakup of Nations: A Quantitative Analysis", *Journal of Economic Growth*, 2011, 16 (3), 183-213.
- [12] Du, R., Y. Yuan, J. Hwang, J. Mountain, and L. Cavalli-Sforza, "Chinese Surnames and the Genetic Differences between North and South China", *Journal of Chinese Linguistics Monograph Series*, 1992, 5, 1-93.
- [13] Falck, O., S. Heblich, A. Lameli, and J. Südekum, "Dialects, Cultural Identity, and Economic Exchange", *Journal of Urban Economics*, 2012, 72 (2-3), 225-239.
- [14] 高翔、尤小宁,“省级行政区划造成的文化分割会影响区域经济吗?”,《经济学》(季刊),2016年第15卷第2期,第647—674页。
- [15] Gong, Y., I. Chow, and D. Ahlstrom, "Cultural Diversity in China: Dialect, Job Embeddedness and Turnover", *Asia Pacific Journal of Management*, 2011, 28 (2), 221-238.
- [16] Kramer, R., "Trust and Distrust in Organizations: Emerging Perspectives, Enduring Questions", *Annual Review of Psychology*, 1999, 50, 569-598.
- [17] Krieger, T., L. Renner, and J. Ruhose, "Long-Term Relatedness between Countries and International Migrant Selection", *Journal of International Economics*, 2018, 113, 35-54.
- [18] Lecce, G., and L. Ogliari, "Institutional Transplant and Cultural Proximity: Evidence from Nineteenth-Century Prussia", *Journal of Economic History*, 2019, 79 (4), 1060-1093.
- [19] 李楠,“文化因素对人口流动的长期影响:基于中国历史经验的实证分析”,《社会》,2015年第4期,第159—176页。
- [20] 李楠、林友宏,“族群文化差异与经济发展——基于基因和姓氏的实证研究”,《经济学动态》,2019年第3期,第44—58页。
- [21] 李秦、孟岭生,“方言、普通话与中国劳动力区域流动”,《经济学报》,2014年第4期,第68—84页。
- [22] 林建浩、辛自强、范佳琳、周先波,“中国省际双边信任模式及其形成机制”,《经济学》(季刊),2018年第17卷第3期,第1127—1148页。
- [23] 林建浩、赵子乐,“均衡发展的隐形壁垒:方言、制度与技术扩散”,《经济研究》,2017年第9期,第182—197页。
- [24] 刘毓芸、徐现祥、肖泽凯,“劳动力跨方言流动的倒U型模式”,《经济研究》,2015年第10期,第134—146页。
- [25] 刘毓芸、戴天仕、徐现祥,“汉语方言、市场分割与资源错配”,《经济学》(季刊),2017年第16卷第4期,第1583—1600页。
- [26] 陆铭,“玻璃幕墙下的劳动力流动——制度约束、社会互动与滞后的城市化”,《南方经济》,2011年第6期,第23—37页。
- [27] 陆铭,“城市、区域和国家发展——空间政治经济学的现在与未来”,《经济学》(季刊),2017年第16卷第4期,第1499—1532页。
- [28] Perkins, D., *Agricultural Development in China: 1368-1968*. Chicago, IL: Aldine Publishing Company, 1969.
- [29] 阮建青、王凌,“语言差异与市场制度发展”,《管理世界》,2017年第4期,第80—91页。
- [30] 阮建青、王凌、李焱,“创新差异的基因解释”,《管理世界》,2016年第6期,第107—117页。
- [31] 史志宏,“十九世纪上半期的中国粮食亩产量及总产量再估计”,《中国经济史研究》,2012年第3期,第52—66页。

- [32] Spolaore, E., and R. Wacziarg, "The Diffusion of Development", *Quarterly Journal of Economics*, 2009, 124 (2), 469-529.
- [33] Spolaore, E., and R. Wacziarg, "Long-term Barriers to the International Diffusion of Innovations", 2011, *NBER Working Paper*.
- [34] Spolaore, E., and R. Wacziarg, "How Deep Are the Roots of Economic Development", *Journal of Economic Literature*, 2013, 51 (2), 325-369.
- [35] Talhelm, T., X. Zhang, S. Oishi, C. Shimin, D. Duan, X. Lan, and S. Kitayama, "Large-Scale Psychological Differences Within China Explained by Rice Versus Wheat Agriculture", *Science*, 2014, 344 (6184), 603-608.
- [36] Tanis, M., and T. Postmes, "A Social Identity Approach to Trust: Interpersonal Perception, Group Membership and Trusting Behavior", *European Journal of Social Psychology*, 2005, 35 (3), 413-424.
- [37] Taylor, J., "Differential Migration, Networks, Information and Risk", In: Stark, O. (ed.), *Migration, Human Capital and Development*. Greenwich, CT: JAI Press, 1986, 147-171.
- [38] 王慧,《中国文化地图》。北京:中国长安出版社,2004年。
- [39] 辛素飞、明朗、辛自强,“群际信任的增进:社会认同与群际接触的方法”,《心理科学进展》,2013年第2期,第290—299页。
- [40] 杨云彦,“1982—1987年中国人口迁移矩阵”,《西北人口》,1989年第3期,第47—51页。
- [41] 袁义达、张诚,《中国姓氏——群体遗传和人口分布》。上海:华东师范大学出版社,2002年。
- [42] 袁义达、金锋、张诚、斋藤成也,“宋朝中国人的姓氏分布与群体结构分化”,《遗传学报》,1999年第3期,第187—197页。
- [43] 袁义达、张诚、马秋云、杨焕明,“中国人姓氏群体遗传——I. 姓氏频率分布与人群遗传分化”,《遗传学报》,2000年第6期, a, 第471—476页。
- [44] 袁义达、张诚、杨焕明,“中国人姓氏群体遗传——II. 姓氏传递的稳定性与地域人群的亲缘关系”,《遗传学报》,2000年第7期, b, 第565—572页。
- [45] 赵冈、刘永成、吴慧、朱金甫、陈慈玉、陈秋坤,《清代粮食亩产量研究》。北京:中国农业出版社,1995年。
- [46] 赵向阳、李海、孙川,“中国区域文化地图:‘大一统’抑或‘多元化’?”,《管理世界》,2015年第2期,第101—119页。
- [47] 赵子乐、林建浩,“经济发展差距的文化假说:从基因到语言”,《管理世界》,2017年第1期,第65—77页。
- [48] 张莹瑞、佐斌,“社会认同理论及其发展”,《心理科学进展》,2006年第3期,第475—480页。
- [49] Wijsman, E., G. Zei, A. Moroni, and L. Cavalli-Sforza, "Surnames in Sardinia II. Computation of Migration Matrices from Surname Distributions in Different Periods", *Annual of Human Genetics*, 1984, 48 (1), 65-78.

Beyond Cultural Barriers in Interprovincial Migration: Information Communication and Identity Recognition

ZHONGDA LI

(*Jinan University*)

JIANHAO LIN*

(*Sun Yat-sen University*)

HONG DENG

(*Erasmus University Rotterdam and Tinbergen Institute*)

Abstract Using the principal component approach to construct measures for interprovincial cultural differences based on various interprovincial group characteristics including genes, surnames and dialects, we aim to empirically explore whether cultural differences are invisible barriers that hinder the labor migration across areas. We find that after controlling for geographical and economic factors, cultural differences have a significantly negative effect on interprovincial migration during the period from 2000 to 2010, and this effect fades over time. Our findings are robust to a series of robustness checks including instrumental variables estimation. Further analysis reveals that information communication and identity recognition are the two major mechanisms through which cultural differences affect interprovincial migration.

Keywords cultural differences, information communication, identity recognition

JEL Classification J61, R23, Z10

* Corresponding Author: Jianhao Lin, Lingnan College, Sun Yat-sen University, No. 135 Xingangxi Road, Guangzhou, Guangdong, 510275, China; Tel: 86-20-84112104; E-mail: linjh3@mail.sysu.edu.cn.