

人口流迁

# 流动人口的流入地选择机制\*

周 皓 刘文博

**【内容摘要】**基于 2017 年全国流动人口动态监测调查数据和相关统计年鉴数据,使用嵌套 Logit 模型,以地级行政区为基本地域单位,以流量流动人口为研究对象,讨论流动人口的流入地选择机制问题。分析结果表明:流动人口中的流量样本与存量样本存在显著的结构性差异,且部分相同变量对两者具有不同作用;流动者个体所能切身感受到的社会经济回报在流入地选择机制中起重要作用;地级特征变量和个体特征变量的交互作用表明地区层级特征对流入地选择机制有异质性影响。结论认为,研究流动人口的不同问题时先要考虑使用流量样本还是存量样本;社会经济“预期回报”而非纯粹的“预期收入”是吸引流动人口的重要因素;流入地选择是流动者基于个体特征、结合流入地特征而做出的综合性的理性决策。

**【关键词】**流动人口;流入地选择;选择机制;嵌套 Logit 模型;地级行政区

**【作者简介】**周皓,北京大学中国社会与发展研究中心研究员、北京大学社会学系教授;刘文博,北京大学社会学系博士研究生。北京:100871

## Destination Selection Mechanism of Migrants in China

Zhou Hao Liu Wenbo

**Abstract:** Based on the 2017 China Migrants Dynamic Survey and data from relevant statistical yearbooks, this paper uses nested logit model (NLGT) to analyse the destination selection mechanism of floating population in China. This study takes prefecture regions as the basic geographical analysis unit and the flow sample as the research object. The results show significant structural differences between flow sample and stock sample, and heterogeneous effects of the same variables on different samples. The socio-economic returns that floating population can personally feel play an important role in the destination selection mechanism. The interaction between regional level and individual characteristics shows that regional level characteristics have a heterogeneous effect on the selection mechanism. The paper suggests that focusing on flow or stock samples should be based on the specific study aim of different research questions. Expected socio-economic return, rather than expected income, is one of the most important factors attracting floating population. The destination selection is a rational and comprehensive decision made by floating population based on individual characteristics and the regional level characteristics of the destination.

**Keywords:** Floating Population, Destination Selection, Selection Mechanism, Nested Logit Model, Prefecture Level Administrative Region

**Authors:** Zhou Hao is Researcher, Center for Sociological Research and Development Studies of China, and Professor, Department of Sociology, Peking University; Liu Wenbo is PhD Candidate, Department of Sociology, Peking University. Email:zhouh@pku.edu.cn

\* 本研究受到国家社会科学基金一般项目“中国流动人口调查的比较研究”(19BRK039)的资助。

## 1 引言

人口流动研究最关注的 3 个问题是:谁在流动?流向哪里?流动结果如何?流动者个体的决策过程总是同时考虑是否流动和流向哪里,不会只关注其中一个方面。但人口流动研究则以总体思维为基础,分别考察“谁在流动”(流动人口的选择性)以及人口的流向与流量问题,而且两者之间通常看似被割裂。这个被忽视的中间连接点就是对流动人口流入地选择机制的考察与讨论。“谁会选择流入到哪里”这一问题的回答既有助于深入了解流动决策的机制,也有助于了解我国规模巨大的流动人口流向分布的成因与过程,还将有助于各地制定相应政策吸引流动人口流入。

已有研究根据历次全国人口普查数据从宏观角度总结出我国人口流迁的空间格局从 1980 年代的“五大迁移圈两个副圈”逐步转变为“三大圈一小圈”的流入地格局(王桂新, 2000),并指出“21 世纪以来我国流动人口流入地分布的整体集中趋势没有发生根本性的变化”(段成荣等, 2020)。有关人口流向的研究“经历了以使用宏观数据进行人口现象的宏观空间格局分析为主导,向越来越重视基于微观数据考察个体人口学过程的转变”(朱宇等, 2017)。但在浩瀚的人口迁移与流动研究中,仅有少量研究结合宏观与微观层次因素讨论了流动人口的流入地选择机制。已有研究尚存在解释变量与流动行为之间的时间问题、样本选择(流量与存量)问题、分析单位与分析方法问题等,这都需要进一步深入思考和讨论。

本文主要关心 3 个问题:第一,对于流动人口流量与存量样本,相同时点的社会经济变量对流入地选择机制的影响是否相同?第二,除了宏观经济水平,流入地的经济发展速度对流入地选择机制是否有影响?第三,地区层级特征对流动者个体的吸引作用是否存在异质性?同时,本文还将从方法的角度讨论流入地选择机制研究中无关选项独立性假设的检验问题。

为此,本文将以重力模型为基础,使用 2017 年中国流动人口动态监测调查数据和相关统计年鉴数据,以地级行政区为基本地域单位,运用嵌套 Logit 模型,以流动人口流量样本为研究对象,以因果变量的时间顺序和流入地经济发展速度为切入点,通过回答上述问题来检验:流入地选择是流动者基于个体特征,结合流入地区域特征而做出的综合性的理性决策。

## 2 文献综述与理论基础

### 2.1 经典迁移决策理论概述

经典的推拉理论(Lee, 1966)将迁移决策视为迁出地与迁入地的推力和拉力以及中间因素的比较过程,是解释迁移行为的经典框架。在经济学框架之下,新古典经济学的宏观理论认为,二元经济结构下现代工业部门对传统农业部门中剩余劳动力的吸纳,以及不同部门在经济发展和工资水平上的差异是影响迁移的主要因素。微观理论则认为,由地区工资差距和迁入地就业机会共同决定的预期收入才是迁移的重要动力,且人们通常会选择能够获取最大预期收益的地点。新迁移经济学以家庭为分析单位,认为迁移的决策过程取决于相对剥夺感和家庭风险最小化。二元劳动力市场理论认为,城市中次级劳动力市场对低端劳动者的结构性需求吸引着外来移民(Massey 等, 1993)。

除经济学解释,围绕迁移决策还有一系列非经济学解释。以血缘、地缘等为基础的社会网络不仅可以减少迁移成本、降低迁移的风险,还会促进迁移流的持续(Massey, 1990)。重

力模型强调人口规模和迁移距离的影响。Tiebout(1956)提出的“用脚投票”理论引发了有关城市公共服务与迁移决策的讨论。

## 2.2 国内迁移决策机制研究回顾

有关迁入地的研究通常利用人口普查或1%人口抽样调查数据,以宏观流向为视角,从空间格局及特征、迁移流动模式两个方面来进行深入考察(朱宇等,2017)。一方面,我国人口流迁的重要特征之一是以乡城流动为主,但城城流动人口的规模和比例正逐步提高(朱宇等,2016),且我国的流动人口以省内流动为主,省际与省内流动的比例约为四六开,省内县际的流动人口规模已接近省际流动人口规模(刘涛等,2015);另一方面,我国人口流迁的空间格局从1980年代的“五大迁移圈两个副圈”逐步转变为“三大圈一小圈”的流入地格局,形成北京、上海两大全国级强势吸引中心和广东、新疆两大地区级强势吸引中心(王桂新,2000)。

国内有关人口迁移流动空间格局的形成及其演变的研究通常以推拉理论为分析框架,考察3个维度:流出地、流入地、中间因素。以往研究较多关注流动者的个体特征,如年龄、性别、受教育水平、社会网络以及家庭特征等因素(杨云彦,1996;严善平,2004)。21世纪以来,随着分析方法的更新,如条件Logit模型(周皓,2006;夏怡然、陆铭,2015)、泊松重力模型或分层泊松模型(巫锡炜等,2013;Shen, 2015)等,以及数据的丰富和完善,国内研究不仅在分析单位上从省级层次(周皓,2006)扩展到地区或城市层级,还通过结合城市特征与个体特征,深入揭示了流入地选择机制(夏怡然、陆铭,2015;刘涛等,2020)。其中,城市层级面板数据的应用体现出分析优势(赵锋、樊正德,2019)。

在转变分析单位的同时,已有研究表明“中国人口迁移流动的空间格局及其演变特征与区域和城乡间社会经济发展水平(尤其是收入水平和就业机会)的差异、迁出地与迁入地间的距离和历史联系、迁入地的投资水平和环境、全球化、内陆与沿海地区不断变化的区位优势、迁出与迁入地的自然条件和文化特征……等因素密切相关……近年来的研究进一步表明,迁入地城镇收入水平等经济因素的吸引作用趋于增强……”(朱宇等,2017)。尽管已有研究在理论分析中涉及推拉理论的3个维度,但在实证分析中,除少数文献提到流出地与流入地之间的差异外(巫锡炜等,2013),多数研究只强调了流入地的社会经济发展水平,并认为人口流动主要是为了谋求更多经济机会、寻找更好的发展空间和更高的经济收入回报(王桂新,2000;Liang和Ma, 2004);不同地区在经济规模、产业结构、固定资产投资和外商投资、工资水平、就业机会等经济因素上的差异(王桂新,1996;周皓,2006;夏怡然、陆铭,2015)是影响流入地选择的重要因素。

“此外,流动人口流入地的选择受政府和市场双重力量的影响,但后者的影响力更强;中西部地区政府力量在引导人口流动中仍起着重要作用”(朱宇等,2017)。社会因素,特别是城市公共服务或社会环境因素等逐步被纳入流入地选择机制的分析框架中,并与流入地城市的公共政策相联系。除就业与工资等经济因素外,劳动力还会为了更好的公共服务而流动(夏怡然、陆铭,2015),因而地区公共财政支出、基础教育资源、医疗服务和基础设施建设等也是重要的影响因素(夏怡然、陆铭,2015;刘涛等,2020)。房价(赵锋、樊正德,2019)和居住环境(刘涛等,2020)等因素亦有影响。这些研究都体现了对我国流动人口的流入地选择机制的探讨正从纯粹以经济为主导、忽略其他因素,转变为以经济为主导、关注多重社会属性的

共同作用。

尽管空间距离等恒定因素的影响相对弱化(朱宇等,2017),但基于重力模型的流出地与流入地之间的空间距离仍被认为是最重要的中间障碍因素之一(王新贤、高向东,2019;刘涛等,2020)。近期,在流入地选择机制的研究中还出现了研究对象的扩展,如刘涛等(2020)针对再次流动人口构建双重多维邻近性理论框架、原新等(2021)对青年流动人才的讨论,以及洪俊杰和倪超军(2020)对农民工群体的讨论等,都有助于加深对流动人口流入地选择机制的了解。

### 2.3 研究评述

对已有研究的回顾表明,随着数据的丰富和方法的发展,对人口流向问题的研究正逐渐从宏观流向的描述性分析转向更为深入的机制性分析,且处于不断深化和扩展之中。但目前研究中存在的以下一些问题尚需引起足够的重视。

第一是基础性理论框架及其操作化定义。推拉理论一直是人口流迁研究中的基本理论框架。由于其中的影响因素纷繁复杂,相应的操作化定义会随数据的可得性而改变,从而不仅使各时期的研究之间不具有可比性,甚至同时期的类似研究之间也无法对话。在相对剧烈的社会变迁过程中,无论是宏观层次还是微观层次的重要影响因素,想要明确其影响机制是否会随时间而发生变化,都需要在相同的操作化定义下,从历时性的视野去考察。从截面分析来看,尽管研究视角、研究问题以及调查数据不尽相同,但揭示人口流向的机制却是各研究之共性。因此,各研究应该以基础性理论框架及相应的操作化定义为基础,增加所关心的自变量,并最终形成各自的研究特色。基于共同框架的具有可比性的各项相关研究,既有助于深化对问题的认识,更有助于形成适用于中国社会情境的人口流向机制理论。这需要学界的共同努力。

第二是流动人口的流量与存量以及研究变量的时点选择。这一点包含两个方面:

一方面,流动人口可以被划分为流量样本和存量样本。流动人口内部存在着循环替代的过程:如果将流动人口看成是蓄水池,则一段时间之内必然会有部分流动人口返回原流出地或再次流动到其他地方,同时又补充进新的流动人口。这种替代过程犹如总人口与出生人口、死亡人口之间的变动关系。如果将截面调查所得的流动人口看成“总人口”,将在流入地的居住时间视作“年龄”,则对应于“出生人口”的是流量人口,即在调查时点前一定时间段(如一年)内新进入的流动人口,且其总是处于“年龄结构”的底部;“年龄”大于该时间段的流动人口则为存量人口,其只会随着居住时间的延长而逐渐减少。尽管流量或存量样本似乎与在流入地居住时长有关,但流量或存量样本的区分更关心流动人口的内部替代过程,以及新增人口与原有流动人口之间的区别。其中,流量样本所对应的时间段是离调查时点相对较近且较短的时期(如一年),而存量样本可以不关心甚至忽略其内部成员间居住时长的差异。某种意义上,不同时点截面调查得到的存量样本会由于流动人口的退出而随时间不断发生变化。

另一方面,要考虑研究变量的时点选择与流量或存量样本的关联。人口流动研究中普遍存在解释变量(特别是宏观层面的社会经济指标)滞后于流动行为,从而违反因果推断中因与果的时间顺序问题。因此,通常会使用调查时点前一年或人口普查覆盖的5年期间(以5

年前常住地为标准识别人口流动)的中间时点对应的社会经济指标以提高外生性。但即便如此,该时点仍然滞后于部分存量人口的流动决策时点。最理想的情况是,在分析存量人口时,解释变量对应的时间点是每位流动者个体决定流动的时间点或至少是其流动开始的时间点。然而,这一时间要求因导致数据处理过程极为复杂而被忽略。因此,在讨论流入地选择机制时,简化的可行方法是直接选取流量人口(而非存量人口或包括其在内的全部流动人口)以及调查时点前一年的社会经济指标,以真正解决因果推断中变量间的时间顺序问题。而且,从统计分析角度看,流量或存量样本间可能存在的某些结构性(如性别年龄结构等)显著差异同样会影响分析结果。再者,从方法上讲,不同的研究问题应该对应不同的研究总体,流动人口研究应根据研究问题的不同而选择合适的流量或存量样本。如针对流动人口选择性、流入地选择性、流动行为的(实验)效应等问题时,研究对象应是流量人口,因为包括存量人口后还将包含更多混淆因素;在讨论流动后果(如社会融合问题)时,则应更偏重于存量人口。

综上,流动人口研究中应该注意区分流量与存量样本,并根据研究问题选用合适的样本,以解决因果变量之间的时间顺序和结构性偏差等方法论问题。当然,研究视角、指标选择等问题亦需要更多的讨论。

第三是地域单位的层级与样本选择性。以往研究通常以省份为地域单位,而忽略省内各地区(城市)间的异质性。由于省内不同城市的社会经济发展存在不同程度的差异,省级层面的分析将减少变量的变异程度,降低估计的有效性(夏怡然、陆铭,2015)。当然,降低宏观地域单位的层级已在近期研究中得到重视和改进。但在降低地域单位层级的同时,却出现另一种由于数据可得性而导致的样本选择性问题。如2017年全国共有334个地级行政区<sup>①</sup>,各省的统计年鉴未能提供全部地级行政区的数据,研究者只能以城市统计年鉴为基础,再辅以各省的统计年鉴来获得地区一级的数据。这种情况会导致两方面问题:测量问题与地级行政区样本的代表性问题。在测量方面,各指标的定义在不同的统计年鉴中存在差异,且对地区或城市的定义各不相同,导致各种统计资料的各指标之间缺乏可比性。例如从行政区划来看,地级市应该包括其下辖的区和县,而城市则仅包括建成区在内的市,不包括下辖的县等。这种不同的地域范围划分导致即使针对具有相同定义的指标时也可能出现差别,从而导致测量问题。在样本代表性方面,某些地级行政区由于没有相关数据,在分析中只能由于变量缺失而被剔除,从而无可避免地导致地区层级样本存在选择性偏差。各研究在运用这种带有选择性的样本进行分析之前,亦未予以真正的检验,而是直接推断到全部地级行政区。而且,流入地选择机制的分析方法(如条件或嵌套Logit模型)无法使用统计方法(如固定或随机效应模型)来解决这种样本选择性问题。事实上在已有研究中,地区或城市样本规模有200多个,也有300多个。样本规模的不同所可能带来的偏差到底有多大,尚需予以关注。总之,地区层级数据的测量与代表性问题尚需在今后的研究中加以注意,并予以重视和改进。

第四是分析过程通常忽略流出地的情况。尽管流入地选择机制的研究通常会以推拉理论为基础,但已有的许多研究会由于数据的可得性(如调查中未涉及流出地的相关信息)而

<sup>①</sup> 数据来源为中华人民共和国民政部网站发布的《中华人民共和国二〇一七年行政区划统计表》,详见 <http://xzqh.mca.gov.cn/statistics/2017.html>。

忽略流出地的信息。这就无法具体比较分析流出地的推力因素、流入地的拉力因素、两地间的差异(如两地间社会经济发展水平的差异)以及中间因素对流向选择机制的贡献,更无法使当前的研究与相应的理论(如经典的宏观经济理论)进行对话。因此,未来的流动人口调查特别应该加入流出地的地级行政单位信息。

第五是忽略异质性作用问题。异质性包含两个层面:一是个体间的异质性;二是相同的实验或刺激对不同的个体具有不同的作用,即异质性作用。社会科学研究和公共政策研究都需要强调这种异质性作用。以流入地选择机制为例,个人和家庭特征不同的流动者面对相同的城市可能会有不同的选择结果。这就需要通过社会分组及交互的方式去讨论。特别是在当前流动人口内部结构发生重大变化的条件下,如流动人口的流向构成从早期以乡城流动为主转变为现在的乡城流动与城城流动并重;又如教育结构的变化,特别是大专及以上学历人口所占比重在历次全国人口普查和1%人口抽样调查中都有明显提高。这种结构性变化带来的不仅仅是流入地选择机制的变化,而且不同结构的流动人口在流入地对公共服务的需求和对社会环境及相应政策的理想要求都可能发生根本性变化,甚至提出比以往更高的要求。流入地的选择机制问题是后续各类问题的基础。因此,流入地选择机制的异质性作用问题及其政策含义需要得到更多关注和重视。

第六是方法问题。近20年来,传统人口统计学和人口地理学对人口流向的描述性分析,由于统计分析方法的快速发展而得到充分的改进,最大的特点是不再满足于宏观流向的描述,而将宏观流向的描述与统计分析相结合,将宏观与微观两个层级相结合,以期深入刻画流动人口的流入地选择机制。例如,泊松重力模型是对以流量对数为基础的重力模型的改进;泊松分层模型则是将分层与泊松模型相结合。这两种方法主要用于分析不同地域之间的流量与流出地或流入地特征之间的关系。而从二分类Logit模型向条件或嵌套Logit模型的转变,包括分层线性模型的应用,则体现了由微观个体层次的选择机制研究向宏观流出地或流入地特征与微观个体特征相结合的选择机制研究的转变。但条件Logit模型有比较严格的假设条件——无关选项独立性假设(Irrelevant Alternatives Assumption, IIA),“即任何两个选项被选择的概率之比与其他选项无关。而人口流动中各潜在目的地在社会经济条件、生活环境等方面可能存在一些未被观察到的相似性,不一定相互独立。”(刘涛等,2020)。现实中,流动人口对流入地的选择或决策会首先从某些特征(如距离、省内或省际、迁移网络等)出发,这就使作为选择项的流入地被视为具有一定的相关性或分组性而不再相互独立。因此,在进行相关的定量研究时,在关注样本的测量与代表性的同时,还必须注意统计方法的适用性及相关假设条件。

当然,上述6点问题并不是本文所能够完全解决的,只是由文献回顾而引发的部分思考,希望能够引起人口学界相关学者的高度重视。本文将仅讨论上述部分问题。

## 2.4 研究假设

如前文所述,已有研究在考察城市特征时,已经由纯粹以经济总量指标(如GDP、实际收入等)为主扩展到包括诸如公共服务政策与供给在内的社会发展指标,且验证了社会发展因素对流入地的选择有显著的正向影响,尽管其作用仍然小于经济指标(夏怡然、陆铭,2015)。但该研究使用的是2005年1%人口抽样调查数据,且包含所有流动人口,这导致其存在两方

面问题:第一,正如前文所述,在考察流入地选择机制时应使用流量样本,而非存量样本或所有流动人口;第二,尽管流动人口的流动目的仍然以经济因素为主导,但近年来逐步呈现出多元化趋势,且随着各地公共服务均等化政策的实施,社会政策更加开放和完善,个体流动目的与社会政策的共同作用将会使作为流入地地区特征的社会公共服务政策(包括教育、医疗等)对流动人口的流入地选择显现出更强的影响。因此,基于历时维度,本文提出假设 1:流入地的社会发展特征对流入地选择有显著的正向作用。

已有研究在讨论经济因素对流动决策的影响时,更多关注经济发展水平或绝对量的作用,却鲜有关注经济发展速度的。考虑到经济增速是反映城市发展潜力的重要指标,亦可能通过影响流动人口对未来个体发展的预期或判断而影响到流入地的选择,本文提出假设 2:地区经济增速对流入地选择有显著的正向作用。

地域特征对不同流动人口影响的异质性是本文关注的焦点之一。近年来,流动人口内部构成发生极大变化,人口流动在普遍化的过程中亦呈现出多元化趋势,这就必然导致流入地特征的异质性作用机制,而且它是后期流入地各项社会政策制定的基础。为此,本文提出假设 3:地域特征对流入地选择的影响会因流动者特征而异。其中,流动者特征主要考察性别、年龄、受教育年限、户口类型这 4 个变量。

### 3 数据、变量与方法

#### 3.1 数据来源

本研究所用数据涉及个人与地级行政区两个层次。个人层次的数据为 2017 年中国流动人口动态监测调查数据<sup>①</sup>,该数据共包含 351 个地级行政区的 169989 个流动样本。其中,流入时间在一年及以内的流量样本为本文的分析对象。同时,本文将流动人口限定为因务工经商原因而流动的劳动力,因为因其他原因外出的劳动力的流向决策可能受到更多特殊的个人和家庭因素的影响(夏怡然、陆铭,2015)。剔除包含缺失变量的案例后,纳入分析的样本为 24358 个。

为解决因果变量的时间顺序问题,本文的自变量将选用 2016 年数据。地级行政区层次的数据主要来源于《中国城市统计年鉴-2017》以及各省市 2017 年的统计年鉴,房价数据来源于 ECIC 全球数据库<sup>②</sup>。数据采用众多来源的目的是为了尽可能多地包含地级行政区,尽管可能由于指标定义与含义不同而存在测量误差,但亦只能以此替代。最终数据共包含 282 个地级行政区。某些没有 2016 年数据的变量(如平均房价)以 2015 年或 2017 年数据替代。

#### 3.2 变量选择

本研究的地级特征变量包括基本特征、经济发展和社会发展 3 个维度。已有研究在讨论社会发展因素时通常使用因子分析方法将多个相关变量合成一个变量(因子)。这种方法在降维处理和节省自由度方面是有益的,但却可能存在信息丢失的问题,且最后的因子只能反映地区间的相对位置,而非实际发展水平。因此,本文用生均教师数、医院病床数、医疗保险

① 数据详情请见流动人口数据平台:<https://chinaldrk.org.cn/wjw/#/home>。

② ECIC 全球数据库网址为 <https://www.ceicdata.com/zh-hans/products/global-economic-database>。

参保率 3 个变量来表示社会发展因素,并未将其进行合成。主要变量的描述性统计详见表 1。

表 1 主要变量描述性统计

Table 1 Descriptive Statistics of the Variables

	变量	均值/比例	标准差	样本量
个体特征	性别(男性=1)	0.566	0.496	24358
	年龄(岁)	31.971	9.557	24358
	受教育年限(年)	10.423	3.261	24358
	户口类型(非农业户口=1)	0.180	0.384	24358
	基本特征			
	人口规模对数	15.108	0.696	282
	省内城市(是=1)	0.042	0.202	79242
	地区间距离(千米)	1194.983	666.106	79242
	经济发展			
地级特征	地区 GDP(百亿元)	27.034	36.219	282
	产业结构(%)	1.055	0.517	282
	在岗人员平均工资(万元)	5.884	1.176	282
	平均房价(万元/平方米)	0.562	0.397	282
	登记失业率(%)	5.150	3.185	282
	地区 GDP 增速(%)	7.180	2.578	282
	平均工资增速(%)	10.580	7.055	282
	社会发展			
	生均教师数(人)	0.070	0.013	282
	医院病床数(万张)	2.284	1.979	282
	医疗保险参保率(%)	22.013	29.049	282

资料来源:根据 2017 年中国流动人口动态监测调查数据、《中国城市统计年鉴-2017》、各省市 2017 年的统计年鉴以及 ECIC 全球数据库中相关数据计算得到。

注:①“省内城市”用于表示流出地和潜在流入地是否在同一省份,取值为 1 时表示两地在同一省份。②“产业结构”是指第三产业与第二产业的比值。③“登记失业率”有两个来源:一是根据《中国城市统计年鉴-2017》中相关数据计算得到,即登记失业率=登记失业人数/(年末从业人数+登记失业人数);二是来源于各省市统计年鉴中的上报指标。④“医疗保险参保率”是指参保人数在总人口数中所占比重,参保人数为城镇职工医疗保险和城镇居民医疗保险参保人数之和。⑤表 1 中的“人口规模对数”是指各地级行政区年末户籍人口总数的对数,而后文回归模型中的“人口规模对数”则是指流出地的人口规模对数与潜在流入地的人口规模对数之和。

### 3.3 分析方法

本文采用嵌套 Logit 模型进行分析,它是从条件 Logit 模型发展而来。条件 Logit 模型将选项的决策看成选项本身特征的函数,其优点在于可以同时考虑个体特征和选项特征对选择结果的影响,但它需要满足 IIA 假设,即任何两个选项被选择的概率之比与其他选项

无关(Powers 和 Xie,2008;陈强,2014)。嵌套 Logit 模型则认为多个选项首先呈现为类别特征,即个体在选择时会首先考虑类别,然后再考虑具体选项。从现实看,对人口流动而言,各潜在目的地由于地理邻近性、语言相近性等社会经济环境而存在类别化的差异;从方法看,如果违反 IIA 假设,就需要将条件 Logit 模型改成嵌套 Logit 模型。本文后续分析结果表明 IIA 假设被违反,因此嵌套 Logit 模型更适合本文的研究目标。为此,分析过程中需要先将潜在目的地进行分类,但目前这种分类并没有明确的标准。刘涛等(2020)将目的地分为两类,一类为直辖市、省会城市和计划单列市,另一类为其他地级市。这种分类方法可能适用于该文对再次流动人口的分析。鉴于本研究的目的,本文将按照东、中、西 3 个地带进行分类。当然,本文的分类仍有待进一步检验。经过转换处理,最终纳入分析的总样本为 6844958 条记录。

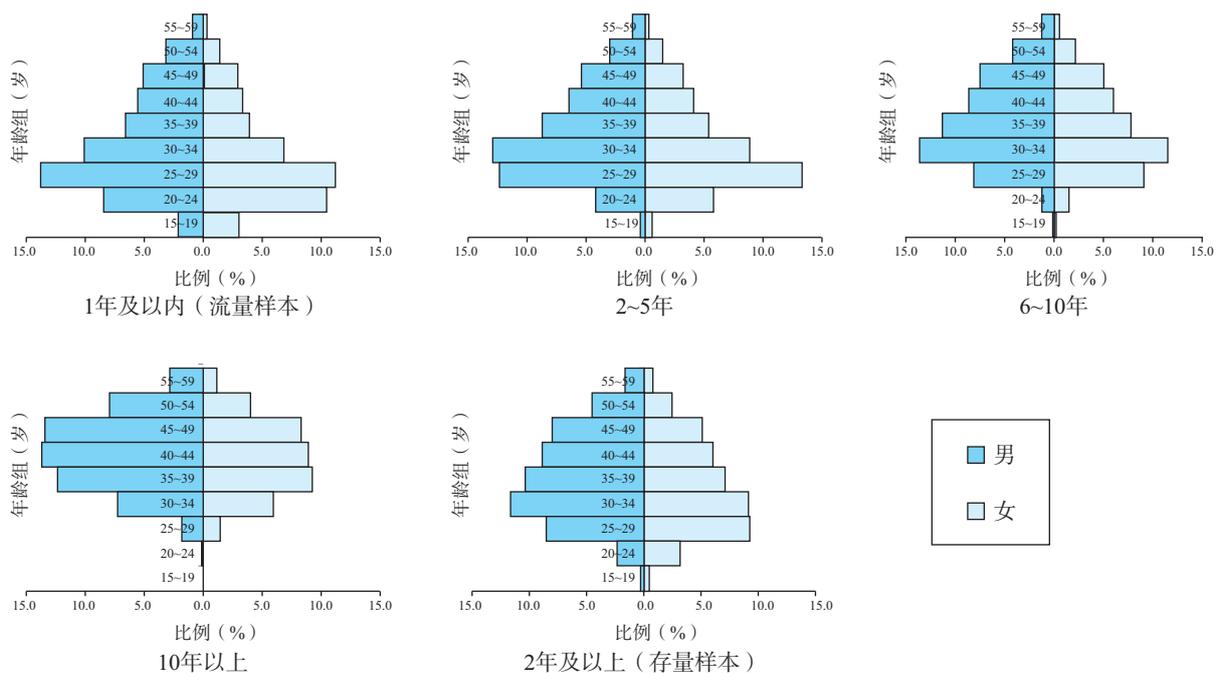
## 4 分析结果

### 4.1 流量样本与存量样本的结构性差异

首先,按照在流入地居住时长将流动人口划分为 4 组:1 年及以下、2~5 年、6~10 年和 10 年以上。将在流入地居住时长为 1 年及以下的流动人口视为流量样本,其他各组则都视为存量样本。图 1 和表 2 对比呈现了各组样本的不同结构特征。需要注意的是,基于调查时点得到的 1 年及以下的流动人口可能是再次流动的人口,而不一定是全国范围内新增的流量样本。

图 1 流量样本与存量样本性别年龄结构的比较

Figure 1 Comparison of the Age-Sex Structure between Flow and Stock Samples



资料来源:根据 2017 年中国流动人口动态监测调查数据计算绘制而成。

表 2 流量样本与存量样本其他人口结构的比较

Table 2 Comparison of Other Demographic Structures between Flow and Stock Samples

结构特征	居住时长	1 年及以内 (流量样本)	2~5 年	6~10 年	10 年以上	2 年及以上 (存量样本)
少数民族比例 (%)		9.44	7.78 ***	7.53 ***	7.06 ***	7.53 ***
不在婚比例 (%)		35.57	20.44 ***	9.55 ***	4.62 ***	13.27 ***
非农业户口比例 (%)		18.03	21.77 ***	22.48 ***	20.60 ***	21.70 ***
跨省流动比例 (%)		63.65	58.29 ***	61.28 ***	66.20 ***	61.13 ***
初中及以下比例 (%)		57.93	55.95 ***	60.84 ***	70.03 ***	60.88 ***
高中比例 (%)		23.04	22.49	20.92 ***	18.96 ***	19.30 ***
大专及以上比例 (%)		19.03	21.56 ***	18.24 **	11.01 ***	17.79 ***

资料来源:同图 1。

注:星号表示各组与 1 年及以内流量样本的 t 检验结果的显著性,\* p<0.05, \*\* p<0.01, \*\*\* p<0.001。

从性别年龄结构来看,流量样本与存量样本之间存在显著差异。在流量样本中,无论男性还是女性,年轻人口的比例都相对较高,而存量样本中年轻人口的比例则相对较低。如果只考察 25 岁及以上人口,那么流量样本的性别年龄结构呈金字塔形,而存量样本呈锥形。此外,流量样本中的少数民族比例、不在婚比例以及跨省流动比例都显著高于存量样本,非农业户口比例则显著低于存量样本。流动人口流量样本和存量样本在各种结构上存在的显著差异可能带来统计结果的结构性偏差。因此,在分析过程中需要特别注意根据不同的研究目的选择不同的流动人群。当然,流动人口动态监测调查数据本身也可能存在结构性偏差,这就需要利用其他数据做进一步检验。

#### 4.2 流量样本与存量样本的流入地选择机制差异

为检验同一变量对流量样本和存量样本的影响作用是否相同,以及 IIA 假设是否成立,本文将在相同变量设置条件下(仅包括地区层级的变量),分别对两个样本进行分析。如表 3 所示,模型 1 对应流量样本,模型 2 对应存量样本,模型 3 对应全样本,模型 4 为包含交互项的流量样本模型。

首先,表 3 最后一行给出了 IIA 检验的卡方值 (IIA chi2\_c),结果拒绝无关假设,表明流入地分析中不能假设地级行政区之间是相互独立的,因此在按地级行政区进行分析时,需要考虑流入地可能存在的选项相关性,应该采用嵌套 Logit 模型,而非条件 Logit 模型。

其次,模型分析结果表明:(1)人口规模对数在流量样本模型中呈现出显著的正向作用,而在存量样本模型中则表现为显著的负向作用;(2)产业结构在流量样本模型中呈现出不显著的负向作用,但在存量样本模型中却呈现出显著的正向作用,而医院病床数虽然在两个模型中都呈现出负向作用,但在流量样本模型中并不显著,在存量样本模型中却是显著的;(3)从增速变量的影响来看,地区 GDP 增速在流量样本模型中呈现出不显著的负向作用,但在存量样本模型中却是显著的,而平均工资增速在流量样本模型中呈现出显著的正向作用,但在存量样本模型中却呈现出不显著的负向作用,两个增速变量对流量样本和存量样本的作用正

好相反;(4)与全样本模型分析结果相比,上述影响差异依然存在,事实上,全样本模型的分析结果更多地体现了存量样本的特征,毕竟相比而言,存量样本规模足够大,以致可以忽略流量样本。

表 3 流入地选择机制的嵌套 Logit 模型回归结果  
Table 3 NLGT Results of Destination Selection Mechanism

	模型 1 流量样本	模型 2 存量样本	模型 3 全样本	模型 4 流量样本+交互项
人口规模对数	0.0581*** (0.0159)	-0.1443*** (0.0081)	-0.1022*** (0.0071)	0.0688*** (0.0165)
省内城市	1.5843*** (0.0250)	1.4638*** (0.0125)	1.4743*** (0.0111)	1.6348*** (0.0261)
地区间距离	-0.0014*** (0.0000)	-0.0016*** (0.0000)	-0.0016*** (0.0000)	-0.0015*** (0.0000)
地区 GDP	0.0008* (0.0004)	0.0062*** (0.0002)	0.0049*** (0.0002)	0.0005 (0.0012)
产业结构	-0.0080 (0.0115)	0.0882*** (0.0057)	0.0684*** (0.0050)	-0.8805*** (0.0733)
在岗人员平均工资	0.4272*** (0.0094)	0.3054*** (0.0046)	0.3288*** (0.0041)	0.4815*** (0.0407)
平均房价	-0.4485*** (0.0207)	-0.3871*** (0.0119)	-0.4015*** (0.0102)	-0.2587** (0.0800)
登记失业率	-0.0690*** (0.0028)	-0.0100*** (0.0011)	-0.0202*** (0.0010)	-0.2353*** (0.0170)
地区 GDP 增速	-0.0040 (0.0032)	-0.0259*** (0.0016)	-0.0214*** (0.0014)	-0.1062*** (0.0187)
平均工资增速	0.0060*** (0.0009)	-0.0006 (0.0006)	0.0007 (0.0005)	0.0080 (0.0062)
生均教师数	-20.3514*** (0.8282)	-13.6378*** (0.3878)	-14.7670*** (0.3482)	-21.1099*** (0.8624)
医院病床数	-0.0032 (0.0059)	-0.0220*** (0.0032)	-0.0167*** (0.0028)	-0.0032 (0.0061)
医疗保险参保率	0.0071*** (0.0002)	0.0048*** (0.0001)	0.0054*** (0.0001)	0.0073*** (0.0002)
交互项				是
样本量	6367442	20718195	27085637	6367442
chi2	7210.5674	26298.2853	33415.6926	7145.0490
IIA chi2_c	8671.9599	24351.2453	32861.0923	8577.4457

资料来源:同表 1。

注:①括号内数据为标准误。②\* p<0.05, \*\* p<0.01, \*\*\* p<0.001。

由此可见,尽管部分变量在流量样本模型与存量样本模型中有相同表现,但仍有部分变

量在两个样本间具有不同的显著性和作用方向。造成这一差异的原因既可能有 MAUP 问题 (Modifiable Areal Unit Problem) 和样本结构性差异问题,也可能是存量样本中流动时间早于解释变量时点导致因果变量存在时间顺序问题。但至少表明,流量样本与存量样本的流入地选择机制可能是不同的,或者至少不能用同一时点的变量来解释。而且,如果将两个样本混合,以全体流动人口作为分析样本,那么不仅可能违反因果间的时间顺序,而且变量的作用也将是不同方向的混合,从而无法真正揭示流动人口的流入地选择机制。因此,流动人口流入地选择机制研究应该使用流量样本,后文分析将主要围绕模型 1 和模型 4 展开。

### 4.3 流入地特征的影响分析

流入地基本特征方面,省内城市表现出显著的正向作用,体现了中国人口流动以省内流动为主的特点,与第七次全国人口普查的结果一致(周皓,2021);地区间距离的影响显著为负,表明流入地的距离越远,被选择的可能性越小;人口规模对数则呈现出显著的正向作用,即流入地人口规模仍具有显著的正向吸引力。

经济因素方面,地区 GDP 和在岗人员平均工资均呈现出显著的正向作用,即经济规模越大,平均工资越高,流动人口越倾向于流入;登记失业率表现为显著的负向作用,即流动人口更倾向于流入失业率较低的地区;产业结构的影响不显著;平均房价表现为显著的负向作用,即平均房价对流动人口的流入地选择呈抑制作用。这些结果所反映的地区经济发展水平的影响与新古典经济学微观理论的解释一致。

经济增速因素方面,地区 GDP 增速呈现出不显著的负向作用,而平均工资增速则表现出显著的正向作用。这表明在不考虑个体特征的情况下,流动人口在选择流入地时会更多地考虑平均工资及其增长率,这涉及流动人口在流入地能够预期和可感知的经济收入,即“经济机会的多寡固然有影响,但更重要的可能是对经济回报高低的预期……高收入回报对流动人口具有显著的强大吸引力”(巫锡炜等,2013)。

社会发展因素方面,医院病床数的影响不显著,只有医疗保险参保率表现出显著的正向作用。义务教育阶段生均教师数则呈现出奇怪的显著负向作用,这意味着义务教育阶段生均教师数量越多,人口越不会倾向流入。可能的解释是:(1)生均教师数与人口规模相对应,使两个变量之间具有一定的相关性,从而导致解释能力的变化;(2)本文将研究对象限定为流动目的为务工经商的流动人口,其中包含部分未婚人群,而他们不一定很在意流入地的教育问题,除收入回报以外,他们更在意与自身利益相关的医疗保险参保率或相关的公共服务。

尽管少数变量在各模型间有细微变化,但总体来看,流动人口一方面更关注与自身利益密切相关的经济收入回报,而非地区经济发展规模与速度;另一方面,从民生的角度则仅关注与自身相关的社会发展因素,而非地区的整体社会发展水平。不过,也有可能是处于不同家庭生命周期阶段的流动人口所关注的焦点不同。

### 4.4 异质性分析:流入地选择的群体差异

模型 4 在模型 1 的基础上加入地级特征变量与个体特征变量的交互项,以考察个体特征的选择性差异,同时也能进一步考察在控制个体特征后,地级特征的偏作用。由于加入全部地级特征变量与个体特征变量的交互项后,模型无法收敛,故本文仅加入地级特征变量中的经济发展变量与个体特征变量的交互项(见表 4)。

表 4 模型 4 中地级特征变量与个体特征变量交互项的回归结果

Table 4 Regression Results with Interaction of Prefecture-level and Individual-level Characteristics Based on Model 4

	地区 GDP	产业结构	在岗人员 平均工资	平均 房价	登记 失业率	地区 GDP 增速	平均工资 增速
性别	-0.0011** (0.0004)	0.0133 (0.0235)	-0.0019 (0.0135)	0.0323 (0.0259)	-0.0049 (0.0054)	0.0090 (0.0063)	-0.0018 (0.0019)
年龄	-0.0001*** (0.0000)	0.0097*** (0.0013)	0.0042*** (0.0007)	-0.0058*** (0.0016)	0.0022*** (0.0003)	0.0001 (0.0003)	-0.0000 (0.0001)
受教育年限	0.0005*** (0.0001)	0.0518*** (0.0042)	-0.0185*** (0.0023)	-0.0035 (0.0046)	0.0091*** (0.0009)	0.0092*** (0.0011)	0.0001 (0.0004)
户口类型	-0.0015** (0.0005)	0.0070 (0.0294)	0.0892*** (0.0179)	-0.1127** (0.0387)	-0.0046 (0.0071)	-0.0035 (0.0085)	-0.0024 (0.0022)

资料来源:同表 1。

注:①括号内数据为标准误。②\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ 。

从性别来看,除地区 GDP 以外,其他经济发展变量与性别的交互作用都不显著。具体来说,女性相对而言更倾向于流入 GDP 高的地区,但其他经济发展变量对男性和女性的吸引力并没有显著差异。这可能是全民流动时代的一个新特点。

从年龄来看,年龄与增速变量(地区 GDP 增速和平均工资增速)的交互项系数均不显著,其他经济发展变量与年龄的交互项系数则都显著。具体来说,地区 GDP 和平均房价与年龄的交互项系数均显著为负,而产业结构、在岗人员平均工资和登记失业率与年龄的交互项系数则均显著为正。这表明年龄越大的流动者越看重工资收益与产业结构,却相对不在意城市的高失业率和经济增速,而地区经济总量和平均房价对年龄越大的流动者有越强的抑制作用。某种意义上,年龄越大的流动者越关注涉及自身生活的经济条件。

从教育来看,流动人口的受教育年限与地区 GDP、产业结构、登记失业率、地区 GDP 增速的交互项系数均显著为正,而与在岗人员平均工资的交互项系数显著为负,与平均房价和平均工资增速的交互项系数则不显著。这表明学历越高的流动者越看重地区经济总量及其增长速度,以及地区的产业结构,但不会很在意工资收益,也相对不会被高失业率阻碍,而平均房价的阻碍作用对各受教育程度的人群都是相同的。可见,学历越高的流动者越看重城市产业结构和整体经济发展,而学历较低的流动者更关心实际的工资收益和失业率等因素。

从户口类型来看,其与产业结构、登记失业率、地区 GDP 增速、平均工资增速 4 个变量的交互项系数均不显著。这表明农业户口和非农业户口的流动人口在这 4 个指标上没有表现出明显的偏好。地区 GDP 和平均房价与户口类型的交互项系数显著为负,而在岗人员平均工资与户口类型的交互项系数显著为正。这说明非农业户口的流动者可能更倾向于选择虽然 GDP 相对较低但平均工资相对较高且平均房价相对较低的地区。这一点仍然体现了流动者,特别是非农业户口的流动者是基于自身特征而对流入地做出理性选择的。

反过来再看主模型,即模型 4,可以发现,地区 GDP 变量在加入交互项后变得不显著,即

该变量对参照组(农业户口的女性)的作用不显著。这说明反映地区经济总量的 GDP 对不同人群有不同的影响。主模型中其他大部分变量的影响基本没有发生根本性变化。其中,在岗人员平均工资仍然呈现出显著的正向作用,但平均工资与个体特征变量交互时其作用会随个体特征而发生变化。平均工资增速与各个体特征变量的交互项系数均不显著,这说明流入地平均工资增速对全体流动人口具有普遍的强大吸引作用。这也再次表明,一方面,流入地经济总量或经济机会固然重要,但是涉及个体利益的高收入回报具有显著的强大吸引力;另一方面,流入地地级特征变量与个体特征变量的交互项系数所表现出来的不同影响,反映出流入地特征对不同人群具有异质性的吸引作用,并且更说明流动人口的流入地选择机制是流动者基于自身特征、结合流入地社会经济特征而做出的综合性的理性决策。

## 5 结论与讨论

本文利用 2017 年流动人口动态监测调查数据和相关统计年鉴数据,借助嵌套 Logit 模型和扩展的重力模型,讨论了地区经济发展水平、经济增速以及社会发展水平对流动人口流入地选择的影响,并探讨了样本选择和地区层级特征影响的异质性作用问题。主要结论如下:

第一,流动人口中的流量人口与存量人口存在各种结构性差异,且部分相同变量对流量样本与存量样本具有不同的作用。造成差异的原因既可能是样本结构性差异,也可能是其本身因果推断机制上存在时间顺序问题。这表明流量样本与存量样本的流入地选择可能存在不同的影响机制。因此,流入地选择机制研究应选用流量样本。同时,流动人口研究中应根据研究问题选用合适的流量或存量样本,以解决研究中可能存在的方法论问题。

第二,应注意方法的假设条件。在以地级行政区为选择项单位的流入地选择机制研究中,必须注意选择项相互独立的假设,即 IIA 假设,并选择适用的模型。

第三,从地区层级变量来看,流动者不仅关注流入地当前的发展状况,还会关注其未来的发展潜力。流入地的经济发展水平或高收入回报及其增速固然重要,但流动者个体在流入地切身感受到的整体社会经济回报而非纯粹的经济回报可能起着更重要的作用。

第四,地级特征变量和个体特征变量的交互作用表明,地区层级特征对流入地选择机制的影响存在个体差异性。这说明流动人口对流入地的选择本质上是流动者个体结合流入地特征而做出的综合性的理性决策。

根据上述结论,还需讨论以下 3 点问题:

第一,定量研究应该对方法予以足够的重视。此处所讨论的方法是广义的,既包括定量研究中数据的两个重要维度——测量与代表性,也包括实际研究中所运用的具体方法。从样本来看,尽管许多研究都注意到变量的时间顺序问题,但却忽略了人口流动及其调查所固有的弱点,即由不同的流动时间而产生的流量与存量样本问题。从具体方法来看,在定量方法得以普遍运用的同时,却极易忽视每种方法本身所要求的各种假设条件。尽管忽略假设后对统计结果的危害性或许不大,但却无法真正保证结果的可靠性与可复制性。当然,本文的结果同样有待再检验。但无论如何,定量研究必须重视研究问题与研究总体的对应、样本的代表性,以及具体统计方法的假设条件,并怀着敬畏之心。

第二,注重理论总结以及对“预期收入”内涵的扩展。本研究检验了部分相关理论在中国的适用性,但中国的现实国情和发展路径与西方国家存在根本性差异,而且 40 多年的改革

开放使中国人口迁移与流动的实际经验为迁移理论提供了丰富的现实素材,这就迫切需要进行理论总结与提炼。例如,新古典经济学强调的“预期收入”在中国仍然具有较强的解释能力,但当前,“预期收入”不应是简单的经济回报,而应扩展到流动人口能够切身感受到或者可预期的各种社会经济回报,如医疗保险参保率等社会因素。“预期收入”的含义应该扩展为包括社会因素与经济收入在内的“预期回报”。对“预期收入”内涵的扩展,要求流入地政府不仅提供充分的可预期的收入回报,更应该加强社会公共服务的供给,更加强调所提供的均等化社会公共服务的可感受性,需要让流动人口能够切身感受到公共服务的均等化,特别是加强民生建设,而不能只停留在政策层面。此外,理论总结还应该注意原有理论无法解释的个体差异性问题的。

第三,研究表明,流入地的社会经济发展水平,特别是流动人口个体能够切身感受到的经济回报与社会条件固然重要,但异质性分析表明,流动人口的流入地选择机制本质上是流动者结合流入地特征与个体特征而做出的综合性的理性决策。这种异质性作用对流入地政府制定相关政策提出了更高的要求。相关政策应充分考虑不同流动人口的选择偏好,并有针对性地提供各种社会公共服务,以有力地吸引流入地所需要的流动人口。而且,部分流动人口更关注发展,而非只是当前的经济发展水平,这对许多经济发展水平相对滞后的中小城市而言具有重要意义:尽管其目前的经济发展水平对流动人口的吸引能力有限,但由于其提升潜力较大,可以通过加快发展速度、提高就业率等手段提升自身竞争力,并通过提供个性化的社会公共服务,强力吸引不同属性的劳动力流入。

当然,本研究仍然存在一定的不足。例如,本文未能将社会发展变量与个体特征变量的交互项纳入模型,以检验社会发展因素的异质性作用;由于数据可得性问题,本文亦未能加入流动人口的社会网络状况;对于生均教师数这一变量在模型中的异常表现并未予以充分的解释。这些都需要在今后的研究中进一步检验和深入讨论。

---

#### 参考文献/References:

- 1 陈强.高级计量经济学及 Stata 应用(第 2 版).北京:高等教育出版社,2014:205-208  
Chen Qiang. 2014. *Advanced Econometrics and Stata Applications (Second Edition)*. Beijing: Higher Education Press:205-208.
- 2 段成荣,赵畅,吕利丹.中国流动人口流入地分布变动特征(2000~2015).人口与经济,2020;1:89-99  
Duan Chengrong, Zhao Chang, and Lyu Lidan. 2020. Changes of Destination Distribution of Floating Population in China (2000-2015). *Population & Economics* 1:89-99.
- 3 洪俊杰,倪超军.城市公共服务供给质量与农民工定居选址行为.中国人口科学,2020;6:54-65+127  
Hong Junjie and Ni Chaojun. 2020. The Quality of Urban Public Service Supply and Migrant Workers' Settlement Choice. *Chinese Journal of Population Science* 6:54-65+127.
- 4 刘涛,齐元静,曹广忠.中国流动人口空间格局演变机制及城镇化效应——基于 2000 和 2010 年人口普查分县数据的分析.地理学报,2015;4:567-581  
Liu Tao, Qi Yuanjing, and Cao Guangzhong. 2015. China's Floating Population in the 21st Century: Une-

- ven Landscape, Influencing Factors, and Effects on Urbanization. *Acta Geographica Sinica* 4:567-581.
- 5 刘涛,卓云霞,王洁晶.邻近性对人口再流动目的地选择的影响.地理学报,2020;12:2716-2729  
Liu Tao, Zhuo Yunxia, and Wang Jiejing. 2020. How Multi-proximity Affects Destination Choice in Onward Migration: A Nested Logit Model. *Acta Geographica Sinica* 12:2716-2729.
- 6 王桂新.中国人口迁移与区域经济发展关系之分析.人口研究,1996;6:9-16  
Wang Guixin. 1996 Population Migration and Regional Economic Development in China. *Population Research* 6:9-16.
- 7 王桂新.中国经济体制改革以来省际人口迁移的区域模式.人口与经济,2000;3:8-16  
Wang Guixin. 2000. The Regional Model of Inter-provincial Migration and Its Changes since China's Economic Reform. *Population & Economics* 3:8-16.
- 8 王新贤,高向东.中国流动人口分布演变及其对城镇化的影响——基于省际、省内流动的对比分析.地理科学,2019;12:1866-1874  
Wang Xinxian and Gao Xiangdong. 2019. The Evolution of China's Floating Population and Its Impact on Urbanization: A Comparative Analysis Based on Inter- and Intra-provincial Perspectives. *Scientia Geographica Sinica* 12:1866-1874.
- 9 巫锡炜,郭静,段成荣.地区发展、经济机会、收入回报与省际人口流动.南方人口,2013;6:54-61  
Wu Xiwei, Guo Jing, and Duan Chengrong. 2013. Regional Development, Economic Opportunities, Income Returns and Interprovincial Migration. *South China Population* 6:54-61.
- 10 夏怡然,陆铭.城市间的“孟母三迁”——公共服务影响劳动力流向的经验研究.管理世界,2015;10:78-90  
Xia Yiran and Lu Ming. 2015. The “Three Immigrants of Mencius' Mother” in Cities: An Empirical Study on the Influence of Public Services on the Flow of Labor Force. *Management World* 10:78-90.
- 11 严善平.地区间人口流动的年龄模型及选择性.中国人口科学,2004;3:32-41+81-82  
Yan Shanping. 2004. The Model Migration Schedule and Selection of China in the Late 1990s. *Chinese Journal of Population Science* 3:32-41+81-82.
- 12 杨云彦.改革开放以来中国人口“非正式迁移”的状况.中国社会科学,1996;6:59-73  
Yang Yunyan. 1996. The “Informal Migration” of Chinese Population since the Reform and Opening Up. *Social Sciences in China* 6:59-73.
- 13 原新,刘旭阳,赵玮.青年流动人才城市选择的影响因素——基于不同规模城市的比较研究.人口学刊,2021;2:48-60  
Yuan Xin, Liu Xuyang, and Zhao Wei. 2021. What Factors Influence the Choice of Cities in Young Floating Talents? A Comparative Study Based on Cities of Different Sizes. *Population Journal* 2:48-60.
- 14 赵锋,樊正德.高房价对大城市人口流入的抑制效应——来自北上广深四城市的实证证据.城市发展研究,2019;3:41-48  
Zhao Feng and Fan Zhengde. 2019. The Inhibitory Effect of High Housing Prices on Population Inflows in the Megacities: Based on the Empirical Evidence from the Four Cities of Beijing, Shanghai, Guangzhou, and Shenzhen. *Urban Development Studies* 3:41-48.
- 15 周皓.资本形式、国家政策与省际人口迁移.中国人口科学,2006;1:42-51+95-96  
Zhou Hao. 2006. Capital Forms, Governmental Policies, and Inter-Provincial Migration. *Chinese*

- Journal of Population Science 1:42-51+95-96.
- 16 周皓. 中国人口流动模式的稳定性及启示——基于第七次全国人口普查公报数据的思考. 中国人口科学, 2021; 3: 28-41+126-127  
Zhou Hao. 2021. The Stability of Migration Pattern in China and Related Issues; Consideration Based on the Data of Seventh National Census Bulletin. Chinese Journal of Population Science 3: 28-41+126-127.
  - 17 朱宇, 林李月, 柯文前. 国内人口迁移流动的演变趋势: 国际经验及其对中国的启示. 人口研究, 2016; 5: 50-60  
Zhu Yu, Lin Liyue, and Ke Wenqian. 2016. Trends in Internal Migration and Mobility: International Experiences and Their Implications for China. Population Research 5: 50-60.
  - 18 朱宇, 丁金宏, 王桂新, 沈建法, 林李月, 柯文前. 近 40 年来的中国人口地理学——一个跨学科研究领域的进展. 地理科学进展, 2017; 4: 466-482  
Zhu Yu, Ding Jinhong, Wang Guixin, Shen Jianfa, Lin Liyue, and Ke Wenqian. 2017. Research Progress on Population Geography in China in Recent 40 Years: A Cross-disciplinary Perspective. Progress in Geography 4: 466-482.
  - 19 Lee E. S. 1966. A Theory of Migration. Demography 1: 47-57.
  - 20 Liang Z. and Ma Z. D. 2004. China's Floating Population: New Evidence from the 2000 Census. Population and Development Review 3: 467-488.
  - 21 Massey D. S. 1990. The Social and Economic Origins of Immigration. The Annals of the American Academy of Political and Social Science 1: 60-72.
  - 22 Massey D. S., Arango J., Hugo G., Kouaouci A., Pellegrino A., and Taylor J. E. 1993. Theories of International Migration: A Review and Appraisal. Population and Development Review 3: 431-466.
  - 23 Powers D. A. and Xie Y. 2008. Statistical Methods for Categorical Data Analysis (Second Edition). Bingley: Emerald Group Publishing Limited: 256-263.
  - 24 Shen J. 2015. Explaining Interregional Migration Changes in China, 1985-2000: Using a Decomposition Approach. Regional Studies 7: 1176-1192.
  - 25 Tiebout C. M. 1956. A Pure Theory of Local Expenditures. Journal of Political Economy 5: 416-424.

(责任编辑: 陈佳鞠 收稿时间: 2021-08)