

中国省际流动人口年龄模式及其特征

——基于省级尺度的类型分析

周皓¹, 陈晓雄²

(1 北京大学 中国社会与发展研究中心, 北京 100871; 2 北京大学 社会学系, 北京 100871)

摘要: 基于全国第七次人口普查微观数据, 以流量流动人口为研究对象, 在描述中国各省区省际流出与流入人口年龄别流动率的基础上, 结合聚类分析方法, 类型化分析了中国 31 个省区的年龄别流出率和年龄别流入率, 并进一步结合各省总省际流入率与流出率综合性地提炼出五种类型。分析结果表明, 各省的省际流动人口年龄别流动率存在较大差异; 四类流出模式和五类流入模式可以概括省际流动人口的年龄别流动率模式; 结合各省总的流出率与流入率则可以将各省的年龄模式综合为五种类型。这些结果可以作为未来区域性人口预测的相关参数。最后还提出了有关人口迁移流动的定义问题、年龄模式的地域差异性反映了我国人口迁移转变的地区阶梯性, 以及中国流动人口问题(甚至中国人口问题)的地域差异性。

关键词: 省际流动人口; 流量流动人口; 年龄别流动率; 类型化; 省级尺度

中图分类号: C922

文献标识码: A

文章编号: 1674-1668(2024)02-0002-13

Age Profiles and Characteristics of Interprovincial Floating Population in China ——Type Analysis Based on Provincial Scale

ZHOU Hao¹, CHEN Xiaoxiong²

(1 China Social and Development Research Center, Peking University, Beijing 100871, China;
2 Department of Sociology, Peking University, Beijing 100871, China)

Abstract: Based on microdata from the seventh national population census, this study focuses on the flow floating population to examine the age profiles of interprovincial migration in China's provinces. Utilizing clustering analysis, the study categorizes the age-specific out-migration rates and in-migration rates of 31 provinces in China, synthesizing them into five distinct types. The analysis reveals substantial variations in age-specific migration rates among the provinces. Four patterns of out-migration and five patterns of in-migration can summarize the age profiles of interprovincial migrants. Furthermore, by considering the overall out-migration and in-migration rates of each province, a comprehensive classification of provincial age

收稿日期: 2023-05-18; 修订日期: 2023-09-22

基金项目: 本文为国家社会科学基金项目“中国流动人口调查的比较研究”(19BRK039)的阶段性成果。

作者简介: 周皓(1972—), 男, 北京大学中国社会与发展研究中心研究员、社会学系教授; 陈晓雄(1995—), 男, 北京大学社会学系博士研究生。

profiles into five types is derived. These findings provide relevant parameters for future regional population projections. Additionally, the paper addresses the issue of defining migration population, highlights the regional disparities in age patterns that reflect the geographical transition of population migration in China, and underscores the regional variations in the issue of floating population (or even the population issue) in China.

Key words: Interprovincial Floating Population; Flow Floating Population; Age - specific Migration Rate; Typification Analysis; Provincial Scale

人口迁移流动是影响一个地域内人口规模、结构与分布的重要人口变动之一。迄今为止,人口学对年龄模式的讨论主要关注于生育、死亡这两种变动,如生育研究通常利用年龄别生育率及由此得到的总和生育率;死亡研究则利用年龄别死亡率构建模型生命表及平均预期寿命。但人口迁移流动的年龄别模式则由于变化与不稳定而导致不确定性,未被给予足够的关注和利用。随着社会经济的发展和社会变迁,中国人口迁移强度在历史进程中呈现出波浪式的上升发展(刘金菊、陈卫,2021),且全国省际流动人口的年龄模式随时间呈现出一定的发展变化(周皓,2023)。由变化带来不确定性的一个重要标志就是各省区、各类迁移流动人口的年龄模式存在差异,既有流出地与流入地选择过程中的年龄模式差异(王桂新,1994,1995),也有不同省区、不同特征(如受教育水平、婚姻状况等)人群的年龄模式差异(杨云彦,1992a,1992b;周爽、黄匡时,2015)。根本上,人口迁移流动既是当前各省区间社会经济发展水平与社会政策等共同影响的产物,同样也深受流出地以流出率为代表的迁移文化之影响(周皓、刘文博,2023):有些省份安土重迁、偏居一隅;有些则更好外出游历闯荡。不同省份的文化属性会烙印于人口迁移流动的特征之上,从而表现出不同的年龄模式。因此,各省区流动人口年龄模式的类型与特征是文化属性的显性化。

人口的年龄别流动率(水平与模式)既是人口预测的基础,亦是其他相关社会问题预测的基础。年龄别流动率的准确性既会影响到人口规模与结构估计的准确性(孟向京、姜凯迪,2018),也会影响到对社会经济发展指标如养老金(刘昌平等,2008;王泽强,2011)、产业结构(逯进等,2019)及公共设施(何晓萍等,2009)等估算结果的准确性。因此,迁移流动人口的年龄模式对后续社会公共政策的制定同样具有重要的社会意义。以往的人口预测通常会结合使用城镇化水平和某区域当前的人口年龄别流动率,其中流动人口的定义和年龄别模式的选择是两个基础问题。由于户口标准下的流动人口包括存量和流量人口,而五年前常住地口径则是不同属性人口之混合体(周皓,2022,2023),因此,人口预测应使用五年内的流量流动人口以反映人口流动的发展趋势(当然,如果可能,还应加上五年内的迁移人口)。而人口预测中迁移流动人口年龄模式的选择无法像死亡模式那样可以从各种模型生命表中挑选,只能以全国或某省区的年龄别模式为基础。虽然 Rogers 等人提出的人口年龄别迁移率参数化是一种理想,但该模型并不一定符合中国的社会现实(郑真真、杨舸,2013;周皓,2023)。此外,迁移流动人口的年龄模式存在的区域间异质性,难以归纳出如生育、死亡那样适用于全国或各区域的标准模式,使得在做较大地理空间尺度的人口预测时,难以将人口迁移因素正确地参数化并纳入预测模型,进而影响预测结果的准确性。那么,退而求其次,在描述我国各省区流动人口年龄模式的地域差异的基础上,是否能归纳为类似于模型生命表不同模式的几种共性类型,或者对我国迁移流动人口的年龄模式进行类型化分析,以供未来人口预测模型的选择使用?

为此,本文将利用2020年第七次全国人口普查数据(以下简称“七普”),以省际流量流动人口为研究对象,在简单描述各省流量流动人口年龄模式之差异性的基础上,运用聚类分析方法对各省省际流出和流入人口的年龄别流动率进行类型化分析,并通过平均化处理各类型年龄别流动率以凸显各类模式的特征,以期对未来人口预测提供类型化的年龄模式。

1 数据与指标说明

本文将使用 2020 年第七次全国人口普查 1‰ 微观样本数据(以下简称“微观数据”)。周皓(2023)曾比较了七普微观数据与汇总数据之间流动人口的年龄性别结构,本文又进一步比较了两种数据来源的总人口、省际流动人口的年龄结构与省区分布(具体请见图 1、图 2)。结果表明,两种数据来源在这些结构上存在着微小的差异。由于普查短表的汇总数据并未提供各类流动人口的相关信息(如分省的、流量与存量流动人口的年龄结构),因此,本文将仍然以微观数据为基础。同时,由于微观数据根据省区、年龄、性别进行细分后各单元内的样本规模过小,因此,本文分析中将采用更为稳健的 5 岁组,而非 1 岁组。此为有关数据的两点说明。

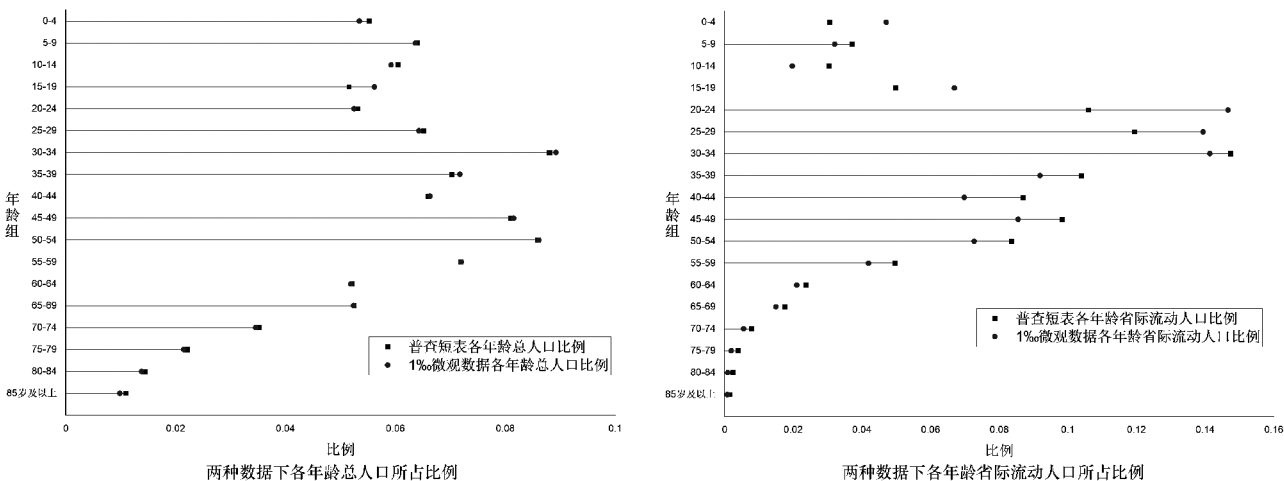


图 1 年龄视角下两种数据总人口和省际流动人口的比较

资料来源:第七次全国人口普查短表数据 1-5、7-2、7-3;2020 年全国第七次人口普查微观数据。

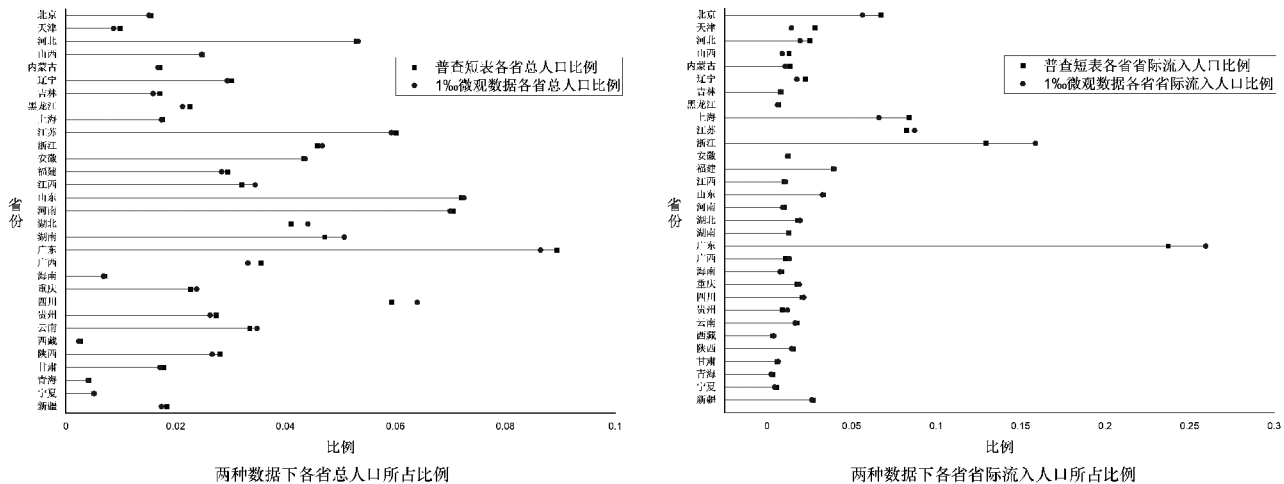


图 2 省区视角下两种数据总人口和省际流动人口的比较

资料来源:第七次全国人口普查短表数据 1-5、7-2、7-3;2020 年全国第七次人口普查微观数据。

1.1 研究对象与指标

中国的人口普查中,流动人口的定义共有三种:户籍口径(第二款人,即户口不在本地,但居住本地半年以上;以及在本地居住未滿半年,但离开户口登记地半年以上)、五年前常住地(或一年前常住地)和

出生地信息。朱宇等(2022)提出以五年前常住地作为反映我国人口迁移流动状况的统计口径。本文认同,五年前常住地口径可以很好地估计我国人口迁移流动的规模与强度。但由于五年前常住地口径下的迁移流动人口既包括五年内的流量流动人口,也包含迁移与返迁人口(普查表题项并无法真正区分迁移人口与返迁人口),而户籍迁移或返迁人口与流量流动人口之间在各种社会经济特征上存在着极大的差异,从而该口径下的样本具有混合属性(周皓,2022),因此,在讨论诸如迁移原因与机制或迁移过程与结果时,应该注意区分这两类人群的不同性质。虽然因为户籍口径下的五年内流量流动人口不包括迁移与返迁人口而低估五年内的人口迁移流动规模,但却能够真正反映人口迁移流动的机制与原因等问题。同时,将人口迁移流动置于中国特定的户籍制度背景之下,流量流动人口的使用将有助于提高人口预测的准确程度,因为流量流动人口反映了近五年内最新流入人口的特征信息,且与存量流动人口在包括年龄别流动率模式的各方面都存在较大差异(周皓,2023)。因此,结合普查定义,本文的研究对象——流量流动人口的定义为:(1)户籍登记地未发生改变,但现居住地跨越户籍登记地所在的乡镇街道并在本地居住半年以上,或者在本地居住不满半年但离开户籍登记地半年以上;(2)五年内离开户口登记地的人口。省际流量流动人口则是五年内离开户籍登记地的省际流动人口。

本文所使用的年龄别流动率实际上是一个比例的概念。严格意义上,年龄别流动率应该是:某时期内,某年龄的流动人口数除以相应风险人口的平均生存人年数。以省际流出率为例,某省的省际流出率所对应的风险人口是由该省某时期内的存活人口+死亡人口+迁出人口-迁入人口。一方面,本文中的“时期”长度跨越五年,另一方面,死亡人口与标准时点的调整相对比较麻烦,因此,本文简单地采用“现有人口”作为风险人口。但为了更加接近“率”定义中的风险人口概念,以更加准确地反映各省的省际流动强度,将流出省流动人口对应的风险人口调整为:现有人口+该省流出人口-该省流入人口。由此,各省年龄别省际流出率即为:

$$m_{oaj} = M_{oaj} / (P_{aj} + M_{oaj} - M_{iaj})$$

省际流入率则完全是比例的概念,直接以省际流入人口与流入省人口之比来计算:

$$m_{iaj} = M_{iaj} / P_{aj}$$

其中, o 表示省际流出, i 表示省际流入, a 表示年龄, j 表示省份; m_{oaj} 表示 j 省 a 岁人口的年龄别流出率, m_{iaj} 是 j 省 a 岁人口的年龄别流入率; M_{oaj} 表示 j 省 a 岁流出人口规模; P_{aj} 表示 j 省 a 岁总人口数, M_{iaj} 是 j 省 a 岁流入人口规模。

本文将采用聚类分析方法对各省区的省际流动人口年龄模式进行类型化分析。

2 我国省际流动人口年龄模式的省际比较

已有研究从空间的角度说明了我国省际流动人口迁出和迁入选择过程的年龄模式及其特征(王桂新,1994,1995),但总体上流动人口的年龄别流动率在近二十年间发生了一定的变化(周皓,2023),因此,在进行类型化分析之前,本文将首先简略描述我国各省省际流动人口的年龄模式。

2.1 省际流出人口的年龄模式差异

“七普”所反映的我国各省省际年龄别流出率请见图3。图示表明,各省省际年龄别流出率模式之间存在很大的差异,且不同于全国省际流动人口,主要体现在:(1)峰值年龄,省际流出的高峰年龄段在各省之间存在差异,如省际流出率最高的贵州、广西、甘肃三省的高峰年龄位于20~24岁,而同样为我国传统人口流出省份的湖北,流出高峰年龄则为25~29岁;(2)在低龄段,各省流出率的上升幅度差异大,部分省份在5~9岁年龄组就表现出显著的上升趋势,而部分省份在10~14岁;(3)在峰值年龄后,各省流出率的下降模式也不同,如高省际流出率的广西在20~24岁后各年龄的下降趋势中表现出4个拐点,而具有低省际流出率特征的广东并未表现出明显的下降拐点;(4)老年段,尽管从全国看,我国尚未表现出Rogers所提出的年龄别迁移率模式中老年人口迁移流动的小高峰,但部分省份则已经表现出这种趋势,如黑龙江、吉林、青海。

2.2 省际流入人口的年龄模式差异

我国各省省际年龄别流入率请见图 4。图 4 表明,相比于省际年龄别流出率,各省的省际年龄别流入率呈现出更大的差异:(1)低龄段,如上海、浙江和北京三省市的低龄段表现出较高的流入率,说明这些省市有相对较多的流动儿童跟随流动人口进入流入地,而其他省市则相对较低。(2)从峰值角度可以将全部省份划分为三类:北京、上海、浙江和天津呈现出明显的年龄选择性,即 20~24 岁组年龄段的流入率极高,而广东、江苏、福建和吉林这四个省份的峰值处于中间,其他省份的流入率相对而言并未呈现出特别的年龄选择性(或者是特别高的峰值)。(3)从峰值年龄以后的年龄别流动率来看,北京、上海、广东和浙江四省市的年龄别流动率在峰值年龄以后下降过程很有意思,北京、上海两市从峰值快速下降到相对较低(但比其他省份相对更高)的水平,而广东和浙江两省的峰值虽然不如北京、上海两市高,但峰值年龄之后直至 50~54 岁组的年龄别流动率都维持在相对较高的水平上,下降趋势相对比较平缓。而西藏表现出的年龄别流入率在 45~49 岁组仍然会有一个小峰值,这是与其他省份不同的地方。(4)老年段中,海南和北京的年龄别流动率相对较高,广东、上海和新疆则次之。本文仅针对部分有特点的省份作简单的总结,以凸显各省之间的差异性。

上述从省级尺度分别对各省省际流出和省际流入年龄别模式的描述表明,各省年龄模式的差异性与共

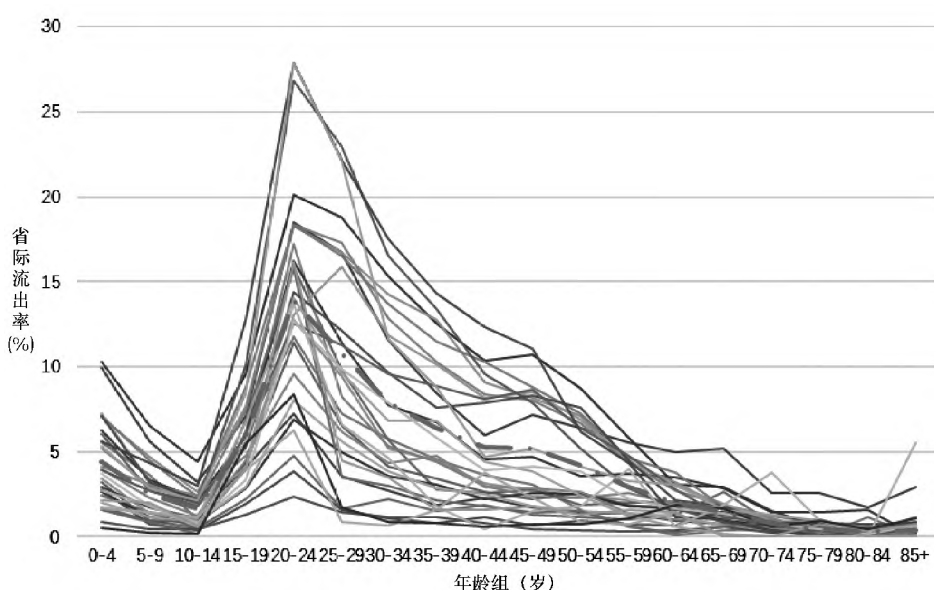


图 3 2020 年我国各省省际流出率的年龄别模式分布

资料来源:2020 年全国第七次人口普查微观数据。

注:图中虚线标记的曲线代表全国五年内省际流出率。

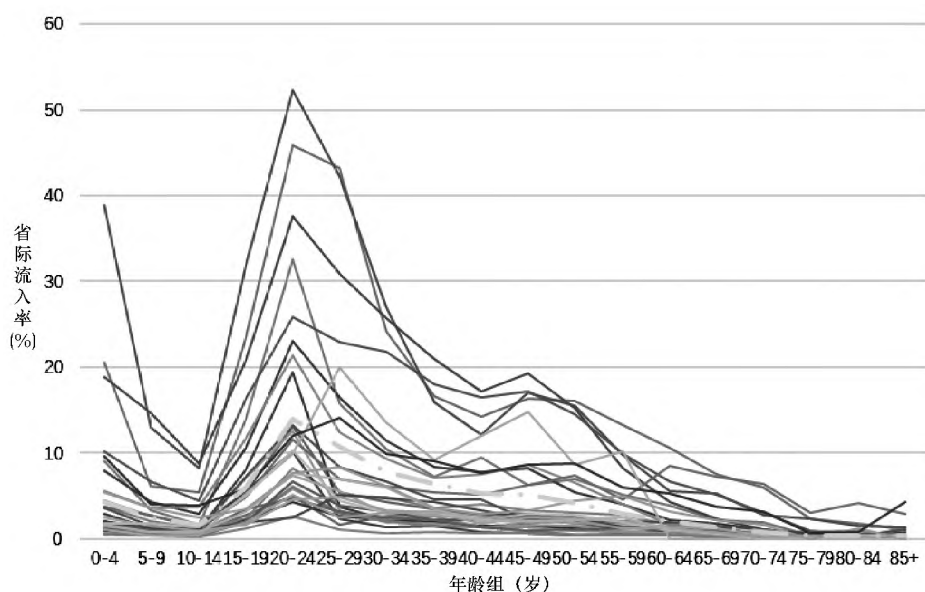


图 4 2020 年我国各省省际流入率的年龄别模式分布

资料来源:2020 年全国第七次人口普查微观数据。

注:图中虚线标记的曲线代表全国五年内省际流入率。

性并存,如年龄选择性和峰值年龄后的下降趋势等都是各省之共性。但要注意的是,如果流出率能够真正体现迁移流动人口的年龄选择性,那么流入率水平则受到流入人口与本地人口年龄结构的共同影响。如果流入省区的人口年龄结构相对老化(如上海、北京),则在峰值年龄段附近,由于流动人口的年龄选择性和流入地人口在该年龄段中的人口相对较少,其年龄别流入率会相对较高;而浙江、广东等省虽然流入人口呈现出相同的年龄选择性,但由于该年龄段的流入地人口相对较多,从而未显现出特别高的年龄别流动率。这是北京、上海两市表现出的年龄别流动率峰值较高的主要原因。正因如此,有必要从标准化角度去进一步探讨流入人口的年龄别流动率模式问题。

3 我国省际流动人口年龄模式的类型分析

省际流动模式的差异性反映我国劳动力人口流动在区域间再分配年龄模式的差异。长久以来,我国的省际人口流动已经具有明显的凝聚性特征,呈现出“非对称双向迁移模式”(豆晓等,2018),即在省际迁移空间分布上表现为劳动力人口的“向海”分布(王桂新等,2012)和向内陆省会等区域性中心城市聚集(刘涛等,2022)。那么,在流动人口总量的空间分布模式下,不同省区的流出或流入人口具有何种年龄结构特征,是否可能对各省份的年龄模式进行类型化分析,并将其归纳为几种典型的年龄结构模式,以便服务于后续的研究与实际工作。为此,本文将在上述分析基础之上,运用聚类分析方法^①对各省年龄模式进行类型化分析,并进一步将同类型下各省的各年龄别流动率平均化处理,得到各类型的特征性年龄别流动率,以揭示各类型间的差异。

3.1 省际流出模式的类型及特点

我国各省省际流出人口的年龄别流动率按照流出强度及特点可以被分为四种类型(具体请见图5)。这四种类型的主要差异表现在整体水平、峰值高度及峰值后的下降速度三个方面。

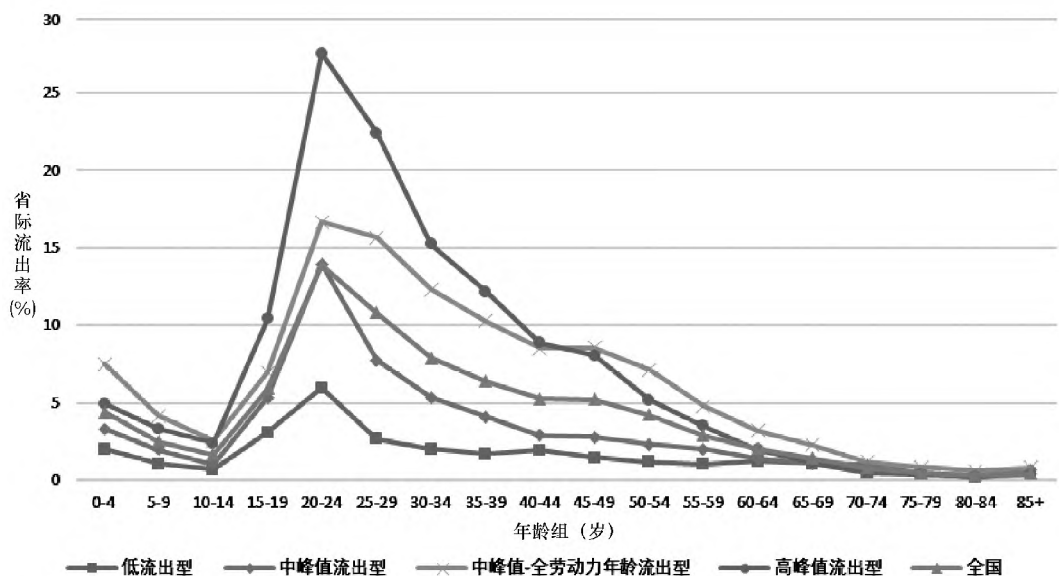


图5 2020年各种类省际流出人口平均年龄别流动率模式

资料来源:2020年全国第七次人口普查微观数据。

第I类的整体水平极低;第II类和第III类之间在峰值高度上尽管相近,但峰值年龄后的下降速度却并不相同;第IV类则以高峰值为特点。

其中,第I类不论是整体水平,还是20~24岁组人口流动最活跃的年龄段,都呈现出极低的流出率。从地域空间来看,主要覆盖了我国三个传统人口迁入中心(北京、长三角、珠三角)和一个人口迁入副中心(新疆),以及西藏、山东两省。根据模式分布特征可以称之为低流出型。

^①本部分针对省际流出和流入人口年龄模式的聚类分析方法采用的是离差平方和法,它倾向于把案例数少的类别聚在一起,发现规模和形状大致相同的类。实际上,本研究同时也采用最长距离法、平均联结法进行聚类比较,结果证明各种方法的聚类结果差别不大。

第Ⅱ类的20~24岁组有中等的流出水平,但其他年龄的流出率都相对较低,表现出一定的年龄选择性,地域空间主要包括华北、东北(黑龙江除外)和西北部分省区,以及南方的福建、海南和云南三省。根据模式分布特征,可以称之为中峰值流出型。

第Ⅲ类的20~24岁组同样具有中等的流出水平,但其后各年龄别流动率也相对较高(即峰值年龄后的下降速度极其缓慢,呈现出劳动力年龄人口外流的普遍性),地域空间主要包括中部省份和西部的四川、重庆,以及东北的黑龙江。这些省份基本上都是我国主要的人口流出省份。根据模式分布特征,可以称之为中峰值-全劳动力年龄流出型。

第Ⅳ类的20~24岁组流出率水平极高,该年龄组中至少有四分之一的人口呈外流状况,但其后各年龄的流出率则相对较低(即峰值年龄以后的下降速度相对较快)。地域空间对应甘肃、贵州和广西三个流出省份,而且这三个省区流出人口的目的地相对比较稳定、集中,如甘肃到新疆、贵州到浙江、广西到广东,形成了“一对一”的定向流动格局。根据模式分布特征,可以称之为高峰值流出型。

以上分析表明我国省际流出人口年龄别流动率模式存在明显的地域性差异,但全国层面的人口年龄别流动率模式并无法反映省际流出模式的这种差异性。在分区域(或分省区)的人口预测中如果使用全国层面的年龄模式,则必然会使预测结果产生偏差(不论是水平还是结构)。而类型化后的人口年龄别流动率则可以作为各区域人口预测的基础参数,从而提高人口预测的精度,使预测结果更贴近于现实。当然,本文仅仅在方法论层面初步提出改进人口预测模型的可能设想,在具体的操作化和技术层面仍需进行更深入的探讨。

上述流出率类型之间的差异也体现区域流动文化的差异。如同从出生地视角分析的结果,部分省市(特别是人口流入大省,如浙江、广东等)表现出迁出的绝对凝固性(周皓、雷琳旋,2023)。上述第Ⅰ类中除西藏以外,其他省份基本上都是吸引人口流入的大省,而其流出率则是极低,即这些省份的人口更愿意留在本省,或者如浙江向上海的流动,即更愿意流入经济发展水平相对更高的临近省区,从某种意义上既体现了人口迁移随经济发展水平(经济阶梯)而呈向上流动的趋向,也表明理论上应该存在的逆向人口迁移流动并不一定会在现实中存在。而第Ⅲ类模式则体现中部省份普遍流出的特点,即除峰值后的各年龄别流动率仍然相对较高(年龄别流出率的下降速度相对较慢),体现了这些省区人口对人口流动的普遍接受且流出意愿相对较高。第Ⅳ类模式体现更强的年龄选择性,但峰值年龄以后的流动率则下降得极快,这说明广西、贵州、甘肃三省仍然是以年轻的劳动力输出为主,尚未进入第Ⅲ类的普遍流出的模式。但不论是从迁移文化扩散的角度,还是从相关迁移理论(如相对剥夺理论、家庭风险最小化理论,乃至机构理论等)的角度,抑或是在极低生育水平条件下从流出地人口年龄结构转变的角度,都预示着第Ⅳ类模式向第Ⅲ类模式的转变。至于未来是否真的能转变为第Ⅲ类模式则需要拭目以待。

3.2 省际流入模式的类型与特点

我国各省省际流入人口的年龄别流动率按照流入强度及特点可以被分为五种类型(具体请见图6)。第Ⅰ类是流入水平相对最高的类型,其特点表现为:具有相对较高比例的流动儿童(0~4岁比例达20%以上);20~24岁组的流入率峰值也极高(将近40%左右,表现出极强的年龄选择性);峰值年龄以后的年龄别流动率虽然也快速下降,但即便到50~54岁组的年龄别流动率仍然高达15%以上,远高于其他类型。这一类型的主要特点是:整体流入率极高、儿童段及迁移最活跃的20~24岁组流入率也极高(特别是儿童段)。主要包括北京、上海、浙江和广东这四个吸引省际流动人口最多的省市。由于经济结构对劳动力的吸引、社会发展对人口的需求(如因隔代抚育而流入的中老年人,或因年轻人口流入而生育的流动儿童),以及高等教育对人才的集聚效应等因素,在考察这些省市人口流入状况时,需要以全年龄的视角去考虑各年龄段流入人口的具体驱动机制以及可能面临的各种社会问题(如流动儿童的教育及长期发展和中老年流动人口的社会适应、医疗保障、社会保障等),而非仅关注流入的劳动力人口问题。根据模式分布特征,可以称之为高流入型。

第2类的流入水平相对较高、有较强年龄选择性(即峰值年龄以外的年龄别流动率相对较低,但峰值相对较高),当然,这种较低或较高都是相对的。相对于第一种流入类型,则其不仅峰值相对较低,而且其他年龄(特别是儿童

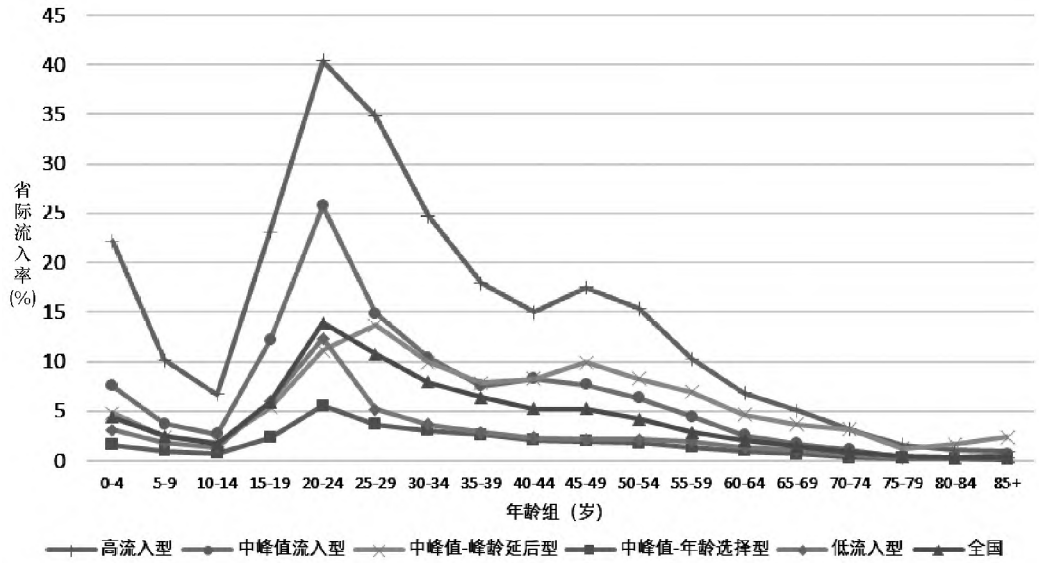


图6 2020年各种类型省际流入人口的平均年龄分布模式

资料来源:2020年全国第七次人口普查微观数据。

年龄段)都相对较低;但如果相对于其他类型,则其峰值和其他年龄段的流入率都相对较高。它与第3类之间的主要区别在于峰值年龄组和峰值水平,但在25~29岁及其后的各年龄组中,两种类型的年龄别流动率都基本相同,基本上维持在5%到10%之间,且在55~59岁组时才低于5%。这说明这一类省份中的流动人口退出流入地劳动力市场相对会较早。如果再回头看第1类,即便是65~69岁组的低龄老年组,其比例仍然在5%以上。当然,在第1类中,这种相对较高的流入率既可能是低龄老年人口延缓退出流入地劳动力市场的原因,也可能是由于社会或家庭原因而流入到流入地的原因。总之,第2类模式的主要特点是,中等流入强度、中等年龄选择性及中等(峰值年龄后的)持续流入,但退出年龄也相对较早。这种类型主要包括天津、江苏、福建三个省市。根据模式分布特征,可以称之为中峰值流入型。

上述两种类型所对应的均为东部沿海省份,但无论是从人口流入水平高低的角度来看,还是从流入人口的年龄别流入率模式来看,尽管我国的省际人口“向海”流动的趋势仍旧保持着相当的稳定性,但沿海地区的各省市也表现出内部分化特征。这说明在研究我国的人口省际流动时,至少应从地域尺度进行异质性考察,而非将全部省际流动人口当成一个整体的研究对象。

第3类的主要特点在于峰值年龄的推迟和峰值年龄后的缓慢下降。与第2类相比,其主要特点是,不仅峰值年龄推迟,而且峰值高度下降(第2类的峰值高达25%左右,但第3类的峰值仅不到15%);与第2类相同的是,峰值年龄后的各年龄别流动率几乎没有下降,甚至在45~49岁组时还有所提高,且其后年龄组均相对高于第2类,甚至在最高年龄段是这几个类别中相对最高的。当然,这一类包括新疆、西藏和海南这三个省区,但这三个省区形成第3类年龄别流动率的原因可能存在一定的差异,西藏可能更多与政策相关联(如援藏等),从而使峰值年龄推后,且不明显;对海南而言,其老年段的提高与其作为养老基地的气候环境相关,而且也应该特别注意“候鸟式”流动老年人对本省的社会经济发展与社会资源配置的影响(韦晓丹、陆杰华,2017)。根据模式分布特征,可以称之为中峰值-峰龄延后型。

第4类与第5类其实有些接近,但由于两类间峰值存在相对较大差异,才进一步区分成两种类型。其中前者的峰值相对较高,达12%左右,而后者则是最底的,仅为5%左右;在其他年龄上,这两种类型的年龄别流动率基本相同。显然,这两种类型涵盖的省区有所不同。第4类主要包括我国的华北地区和西南地区,以及青海、甘肃等,基本上是主要的省际人口流出省区。而第5类则主要包括陕西-鄂渝-东北等省

区,主要分布在我国的中部区域和东北三省,呈现出高度的地域集中特征。根据模式分布特征,分别称之为中峰值-年龄选择型、低流入型。

当然,正如上文所述,年龄别流入率受到流入人口与流入地人口年龄结构的双重影响,还需要进一步的分析与讨论。与流出人口相比,流入人口年龄分布的类型分析结果在地域分布上更集中。理论上,我国省际流入地区表现出集聚中扩散趋势,而省际流出地区表现出漫延式扩散趋势(周皓、雷琳旋,2023),但这并不意味着对人口流出地区只能关注其流出人口的状况,如果同时考虑其流入人口状况,就会发现这些省区的人口流入模式更加复杂。诚然,本研究采用聚类分析得出的五种流入模式可能并不十分完善,但也可以反映出多数省份复杂的人口流入模式及其特征。

总体上五种省际流入模式存在显著的差异,与省际人口流出模式相比,在大多数年龄段表现出与全国的省际流量流动人口年龄别流动率模式完全不同的特征。这再次说明,省际流动人口是内部存在明显异质性的群体,至少需要同时从流量与存量、流入与流出的交互视角予以考察。

具体来看,五种省际流入的平均年龄分布模式的差别在三个地方:(1)5~14岁的下凹槽形状;(2)流入水平的双峰值分布;(3)峰值年龄后的拐点。

首先,与省际流量流动人口相比,不同省际流入人口的年龄别模式在5~14岁表现出完全不同的形状特征。就5~14岁年龄段而言,高流入型的下凹槽形状尤其明显,在10~14岁上流入水平下降至最低点,而15岁后迅速上升;中峰值流入型表现出极为微弱的下凹槽分布;其余三种模式在该年龄段上无明显的下凹形状。上述四种省际人口流出年龄别流动率模式也没有明显的下凹槽分布。与省际人口流出相比,我国的省际人口流入不仅在地域空间上表现出更强的集聚性特征,而且对全国年龄别流动率的塑造也尤为突出。

其次,五种省际人口年龄别流入率模式表现出不同的峰值分布形态,其中高流入型和中峰值-峰龄延后型呈明显的双峰分布形态,而其余三种类型的分布形态表现为单峰型。高流入型的两个高峰年龄段分别位于20~24岁和45~49岁,中峰值-峰龄延后型的两个高峰值则位于25~29岁和45~49岁,其中后者的两个峰值流入率水平相差不大,均在12%左右;前者在两个峰值上的流入率水平差异大,20~24岁的流入水平高达40.4%,而45~49岁也高达17.5%。其余三种类型为典型的单峰分布,高峰年龄段均为20~24岁。

不同的流动人口年龄模式是由各省区的经济结构和社会需求所决定,也从年龄结构角度说明我国省际人口流动空间格局的形成。高流入型主要是由我国的三大人口流入中心省市所构成,强烈地吸引着全国劳动力年龄段的人口,因此,这些传统人口流入省区在吸引流动人口上具有全年龄段的特征(几乎所有年龄段的流入率都处于最高水平,且拥有两个峰值)。与此类似,中峰值-峰龄延后型虽然在年轻劳动力人口的吸引上不如中峰值流入型,但在吸引45岁之后的中老年人口的流入上表现突出,这也可能是拉动新疆成为我国省际人口流入副中心的重要因素。

再次,五种省际人口年龄别流入率模式在峰值年龄后也表现出不同的分布形态,主要体现在流入率下降的拐点的不同:高流入型在峰值年龄后的下降模式最为复杂,共有4个拐点,其中在25~39岁的几个年龄段上下下降速度最快,在40~44岁上速度明显减缓,而在45~49岁则呈现出上升的趋势,由此形成了第二个流入小高峰,在此之后继续下降,并在60岁之后下降速度进一步加快;中峰值流入型在25~29岁上的下降速度最快,在30~39岁上流入率有所减缓,但在40~44岁上略微上扬,此后又稳步下降;中峰值流入-峰龄延后型在30~39岁上缓慢下降,而在40~49岁上又有所回升,形成了人口流入的第二个小高峰,之后则稳步下降;其余两种模式在峰值年龄后稳步下降,没有表现出明显的速度变化特征。

3.3 省际流入与省际流出的交互分析

上述类型化分析是从流入与流出角度分别展开的,那么流入与流出模式在各省区之间的交互分类又会呈现出何种情况,为此本文将上述各自的分类进一步归纳总结。具体可见表1。同时,为了更好地展示这种

交互分类与各省总体的流入/流出水平的比较,本文用图 7、图 8 来予以展示。

表 1 各省流入与流出的交互分类表

流出 流入	I 低流出型	II 中峰值流出型	III 中峰值 - 全劳动力 年龄流出型	IV 高峰值流出型
1 高流入型	北京、上海、广东、浙江			
2 中峰值流入型	江苏	天津、福建		
3 中峰值 - 峰龄延后型	新疆、西藏	海南		
4 中峰值 - 年龄选择型	山东	河北、山西、内蒙古、云南、青海		
5 低流入型	安徽、江西、湖南、河南、四川			广西、贵州、甘肃
	黑龙江、湖北、重庆			

结合图表中的分类可以看到,事实上我国各省区的流出流入模式,除西藏、新疆和山东以外,可以大致分为四个大类:低流出、高流入(即图 7 的右下角的四个省份);低流出、中流入(对应天津、福建、江苏和海南四省,即将流入模式 1 和模式 3 合并成一类);中流出、中流入(对应着坐标轴原点或中心部分的省份,包括吉林、辽宁、河北、山西、内蒙古、云南、青海等);高流出、低流入(对应着坐标轴左上角的省区,包括安徽、江西、湖南、河南、四川、黑龙江、湖北、重庆等,相当于将第 III 类和第 IV 类流出模式和第 4、第 5 类流入模式相结合)。

这四种分类既可以简单地根据总的流出率与流入率来判断,也可以根据表 1 中的分类来进一步细化性的选择。当然,其中高流出、低流入中的三个省份(广西、贵州、甘肃)因其特殊的流入\流出模式(包括省对省的高年龄选择性的流出)也可以单独归为一类,称其为高选择流出、低流入类。

上述四种分类只是对目前我国各省区的省际流动人口年龄别流动率特征类型化的结果,类型化后的模式可以综合地揭示我国省际流动人口年龄模式的地域类型。当然,在以各省区为人口预测单位时最好应该结合各省的实际情况予以调整。

4 总结与讨论

本文利用 2020 年第七次人口普查 1‰原始微观数据,以省区为基本分析单位,描述各省省际流量流动人口的年龄别流动率,并在此基础上,通过类型化分析方法,归纳分类了全国 31 个省区的人口流出和流入的年龄模式,并比较各种分类模式之间的异同。

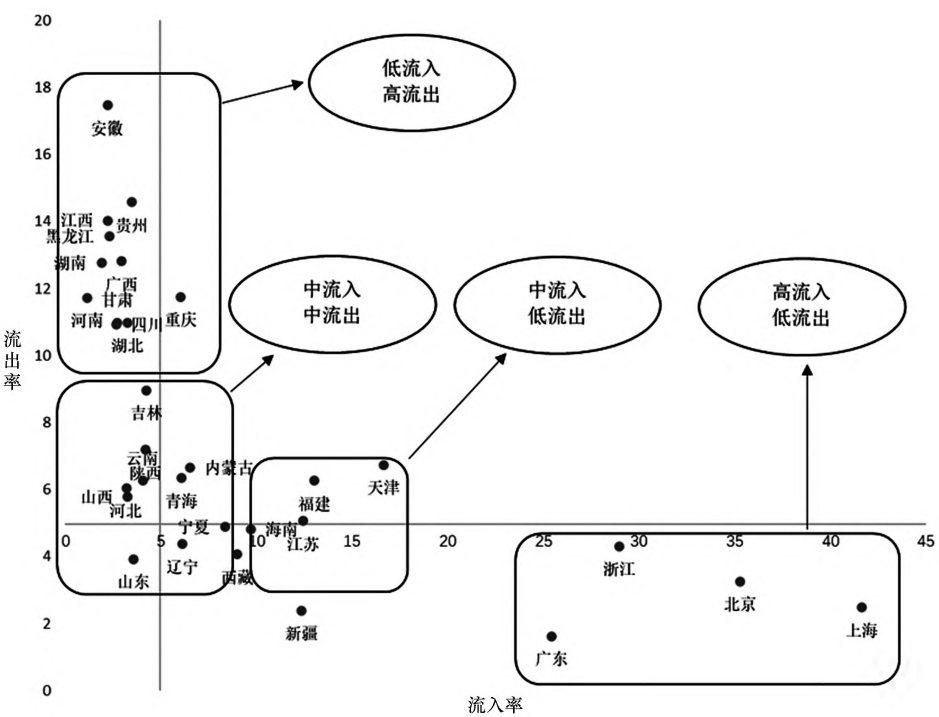


图 7 省际流出与流入年龄模式的交互分类
资料来源:2020 年全国第七次人口普查微观数据。

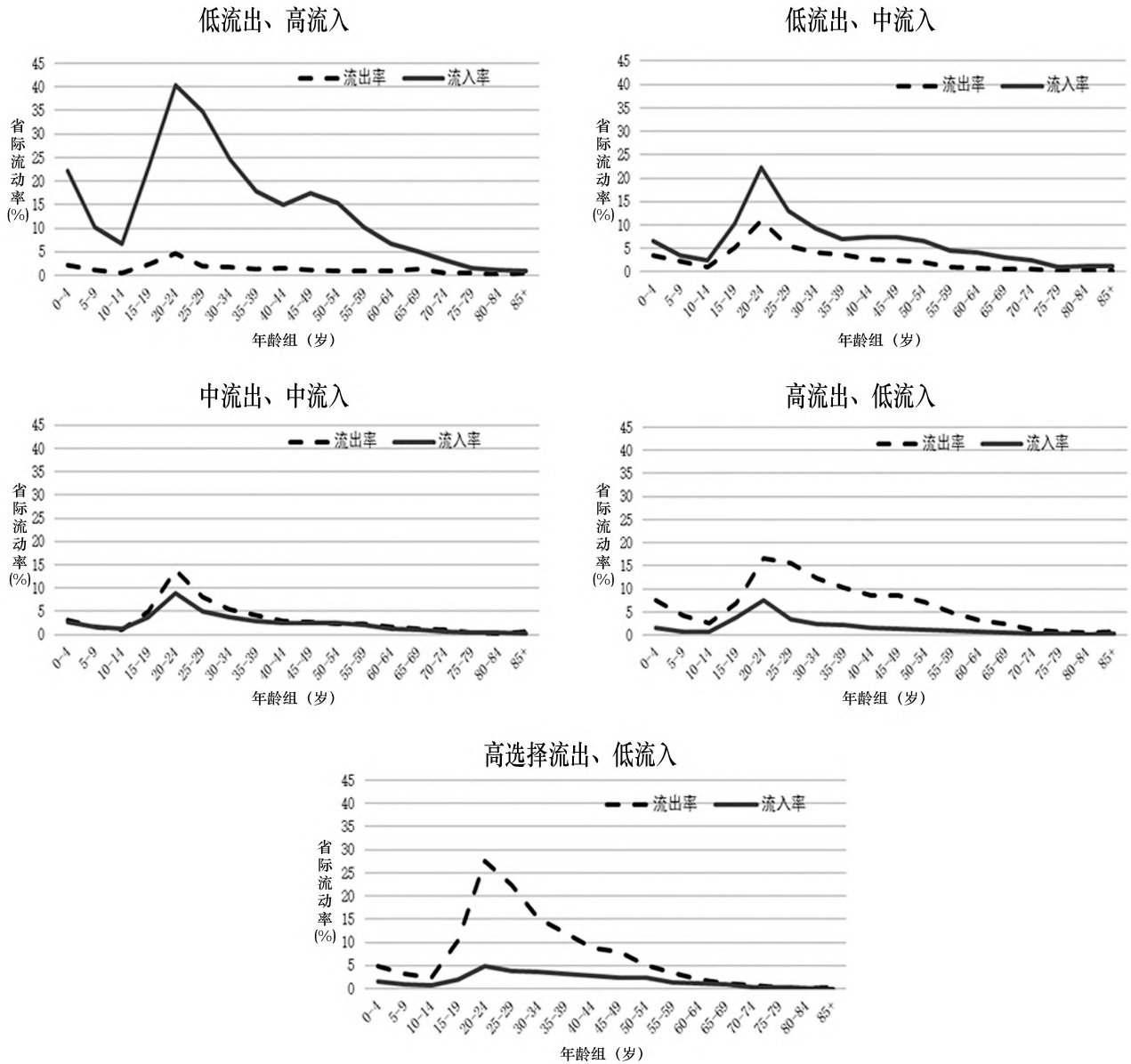


图 8 省际流出与省际流入的交互分类模式结果

资料来源:2020 年全国第七次人口普查微观数据。

4.1 主要结论

上述分析可以得到以下几点主要结论。

第一,各省的年龄别省际流出与流入率模式存在明显的差异,因此,将全国省际人口年龄别流动率模式应用到区域性人口预测时结果很可能有偏差,有必要对各省流入和流出的年龄模式进行类型分析。

第二,类型分析结果表明,我国省际流出人口年龄别流出率可以按照省际流出水平由低到高被划分为四种类型:低流出型、中峰值流出型、中峰值-全劳动力年龄流出型、高峰值流出型;类型间差异主要体现在峰值高度和峰值年龄后的下降速度。四种类型呈现出明显的地带性分异和空间聚散特征。

第三,各省区的省际流动人口年龄别流入率模式按照年龄别流动率的峰值高低、峰值年龄后的年龄别流入率的下降速度等,可以被分为五种类型:高流入型、中峰值流入型、中峰值-峰龄延后型、中峰值-年龄选择型和低流入型。类型间的主要差异在于:高流入型中低龄段(0~4岁)有更高的流入率,形式更接近于

Rogers 模型;后三种则主要是峰值高低、峰值年龄后的下降速度。

第四,根据流入与流出的年龄别模式类型交互,以及各省总体流入率与流出率的交互,可以进一步综合性地提炼出五个大类:低流出+高流入、低流出+中流入、中流出+中流入、高流出+低流入及高选择流出+低流入。这五种分类方式可以作为未来各省区人口预测中人口流动的基础参数。当然,各省的实际参数可以在此基础上进行微调。

4.2 启示与讨论

上述结论在研究方法层面上为我国未来区域人口预测研究(特别是流动人口年龄模式)提供了现实与理论基础。正如上文所述,由于区域(省区)间的人口流入或流出的年龄别模式存在着显著差异,因此,如果利用全国流动人口的年龄结构或年龄别流动率进行区域人口预测,势必会导致人口预测结果在水平和结构上的偏差。本文对各省区省际流动人口年龄别流动率模式的类型化处理,类似于模型生命表,可以为不同区域的人口预测提供相对更为贴近现实的不同模式和基础参数,以提高人口预测精度,进而服务于未来各区域社会经济政策的制定。

当然,预测的准确性还与地域层级、参数设置和预测内容等更多问题相关联。首先,本文仅关注于省区层级,对于更小区域的人口预测由于不确定性而显得更为复杂,其整体思路亦不同于传统人口预测方法。其次,参数设置过程中,区域的宏观社会经济因素和人口因素往往具有空间溢出性质,邻近省区的相关因素也可能对该区域的人口流动趋势产生影响,并导致预测模型的假定失真,降低预测结果的准确性。而本文提出的区域人口流动模式类型可以在一定程度上克服这种缺陷:上述各种年龄模式的类型通过将多个省区绑定,使得该模式类型内部已经包含相关“邻近区域”的宏观社会经济状况和人口流动状况,从而使迁移模式的假定更加稳健,有利于预测精度的提高。再者,人口预测的内容不仅可以是某些省区的流入、流出人口的规模与结构,及其对流出地或流入地人口规模与结构的影响,而且还可以进一步预测某一条具体的流动链条下流入或流出人口的未来特征。这种更小尺度的预测需要依赖相关的理论模型,一般常用的模型有引力模型、机会模型、罗吉特模型、马尔科夫链模型(王桂新,2000)。但是,这种类型的预测对于数据信息的要求更加严格,它需要对两个微观区域的社会经济环境和以往的人口迁移流有足够的了解,同时也需要对宏观迁移流(如某区域所有流入人口状况)有相应的预判。而本文的类型化分析同时考虑了人口流动因素的区域差异和空间溢出性质,一定程度上可以避免忽略某些重要信息而使模型假定失真。

除了对人口预测有意义之外,本文对各省区省际流量流动人口的年龄别流动率模式之差异的讨论还可以得到以下几点启示:

其一,定义问题。近期有关中国人口迁移流动的定义问题受到关注与讨论(朱宇等,2022;周皓,2022;王婧雯等,2023),并分别提出户口统计口径之下的五年内流量流动人口和五年前常住地口径为标准来测量与讨论中国的人口迁移流动问题,并同时指出户口统计口径之下的五年内流量流动人口对中国人口流动规模的低估问题和五年前常住地口径下的样本混合属性问题(周皓,2022)。当然,任何一种统计口径都可能会存在内在的缺陷,因此在使用人口迁移流动数据时需要注意不同统计口径的应用场合及对分析结果的可能影响。由于流量定义存在对中国人口流动规模的低估,因此,对总规模的描述性分析可以使用五年前常住地标准。但在讨论与分析人口迁移流动的机制问题时,正如以往所指出的,由于五年前常住地口径下的样本混合属性问题,就不能再使用五年前常住地口径下的“流动人口”样本(因为它不再是纯粹的流动人口样本),而应该使用户口统计口径之下的五年内流量流动人口,以揭示真正影响人口流动的动因与机制。未来的相关研究应该注意定义问题及不同定义的不同应用场合。这也是本文仍然强调使用户口标准下的五年内流量流动人口的原因。

其二,年龄模式的地域差异性反映了我国人口迁移转变的地区阶梯性。近年来学界一直在讨论中国人口发展是否正面临着第二次人口转变,以及中国人口迁移转型问题(刘金菊、陈卫,2021;段成荣等,2022)。本文的类型化分析揭示了年龄模式的地域异质性。年龄别流动率的变化既是人口发展与社会变迁对人口迁

移流动的作用结果,也是人口与社会发展变化的社会表象特征之一,与中国近三十年来的经济社会发展、经济结构调整、经济转型及区域经济格局变化有很大关系(刘金菊、陈卫,2021)。年龄模式的地域异质性从另一侧面说明各区域社会经济发展及经济格局的异质性。而年龄模式的地域性差异只是中国人口迁移流动转型过程的一个缩影,其背后意含着中国人口迁移转型的区域性差异。不同区域的人口转型事实上并不是同步的,而不能仅以全国性的模式来理解与解释。正如段成荣等(2022)所指出,我国的人口迁移转变并不能简单地用转型前期、转型后期或先进工业社会等阶段来断然划分,而是需要进行分地区的细致考量。

事实上,本文所揭示的年龄模式的地域异质性从某种意义上还说明了中国人口迁移流动转型的阶梯性。不论是流出人口年龄别模式还是流入人口年龄别模式都体现着这种转型。从流出模式来看,由整体较低水平、相对较低的年龄选择性,转变为较强的以青壮年为主的年龄选择性,再到劳动力年龄人口的普遍流动和整体水平相对较高的普遍流动,这四种模式之间存在着不断演进的阶梯形式。流入模式同样呈现出这种阶梯形式的转型过程。同时,五种流入模式中仅有第一种高流入型呈现出西方人口迁移年龄模式 Rogers 模型的部分特点,即 0~4 岁婴幼儿年龄段较高的流动率,而其他模式中的婴幼儿年龄段的流动率仍然极低。如果将 Rogers 模型作为人口转型的标志或特点,那么,不同模式反映的地域差异同样呈现出阶梯式的转型过程。而且,即便是在第 I 类流入模式中,亦未出现老年段的年龄别流动率小高峰。因此,一方面,中国的人口迁移转型必然会有中国的特点,西方理论的适用性仍然有待检验;另一方面,中国人口迁移流动转型存在地域阶梯性特征,而普遍的全年龄段的人口流动将会是未来中国人口迁移流动的必然趋势。

其三,年龄模式的地域差异性进一步强调流动人口问题,甚至人口问题的地域差异性。虽然社会公共政策的制定应该面向全年龄段的流动人口,但各地年龄模式的差异表明各省区的流入人口或流出人口的年龄结构并不完全相同,并由此带来的政策需求也可能是不同的。从流入地来看,高流入型的北京、上海、浙江和广东四个省市所面临的是全年龄段的,其中与流动儿童相关的问题尤其重要;而对中峰值流入型的天津、江苏、福建三个省市则可能会更偏重于劳动力年龄人口。从流出地来看,第 III 类(中部省份和四川、重庆、黑龙江)都已形成了全年龄段的流出,不仅峰值相对较高,且低龄段和除峰值年龄外的劳动力年龄段流出率也相对较高,而且,长期以来一直作为重要的人口流出地,这些省份的劳动年龄人口将是否及如何维持必要的地区发展(及乡村振兴等话题),必将成为这些省份所面临的重要问题。可见,由于人口流动的不同特点,不同省区所面临的由人口流动而带来的各种问题亦是完全不同的。本文仅从年龄模式角度予以讨论,更多现实问题需要予以更多深入研究。

参考文献:

- 豆晓, Blanca ARELLANO, Josep ROCA. 基于相互作用关系的中国省际人口流动研究[J]. 地理研究, 2018, 37(9): 1848-1861.
- 段成荣, 邱玉鼎, 黄凡, 谢东虹. 从 657 万到 3.76 亿: 四论中国人口迁移转变[J]. 人口研究, 2022, 46(6): 41-58.
- 何晓萍, 刘希颖, 林艳苹. 中国城市化进程中的电力需求预测[J]. 经济研究, 2009, 44(1): 118-130.
- 刘昌平, 邓大松, 殷宝明. “乡-城”人口迁移对中国城乡人口老龄化及养老保障的影响分析[J]. 经济评论, 2008, (6): 31-38.
- 刘金菊, 陈卫. 中国人口的迁移转变——基于迁移率的考察[J]. 人口与经济, 2021, (1): 37-49.
- 刘涛, 张家瑞, 曹广忠. 人口流动对区域老龄化进程的影响——一个方法论探讨[J]. 地理研究, 2022, 41(10): 2680-2696.
- 逯进, 王晓飞, 刘璐, 陈阳. 中国省域老龄化、人口流动与产业结构的协同效应[J]. 经济地理, 2019, 39(9): 39-47.
- 孟向京, 姜凯迪. 城镇化和乡城转移对未来中国城乡人口年龄结构的影响[J]. 人口研究, 2018, 42(2): 39-53.
- 王桂新. 我国省际人口迁移迁出选择过程的年龄模式及其特征[J]. 人口研究, 1994, (2): 9-17.
- 王桂新. 我国省际人口迁移迁入目的地选择过程的年龄模式及其特征[J]. 人口与经济, 1995, (6): 40-49.
- 王桂新. 区域人口预测方法及应用[M]. 上海: 华东师范大学出版社, 2000.
- 王桂新, 潘泽瀚, 陆燕秋. 中国省际人口迁移区域模式变化及其影响因素——基于 2000 和 2010 年人口普查资料的分析[J]. 中国人口科学, 2012, (5): 2-13.

(下转第 125 页)

- Loh E S. Productivity Differences and the Marriage Wage Premium for White Males[J]. *The Journal of Human Resources*, 1996, 31(3): 566 – 589.
- Ludwig V, Brüderl J. Is There a Male Marital Wage Premium? New Evidence from the United States[J]. *American Sociological Review*, 2018, 83(4): 744 – 770.
- McDonald P. The Male Marriage Premium: Selection, Productivity, or Employer Preferences? [J]. *Journal of Marriage and Family*, 2020, 82(5): 1553 – 1570.
- Nakosteen R A, Zimmer M A. Marital Status and Earnings of Young Men: A Model with Endogenous Selection[J]. *The Journal of Human Resources*, 1987, 22(2): 248 – 268.
- Nakosteen R A, Zimmer M A. Men, Money, and Marriage: Are High Earners More Prone Than Low Earners to Marry? [J]. *Social Science Quarterly*, 1997, 78(1): 66 – 82.
- Pfeffer J, Ross J. The Effects of Marriage and a Working Wife on Occupational and Wage Attainment[J]. *Administrative Science Quarterly*, 1982, 27(1): 66 – 80.
- Pilossoph L, Wee S L. Household Search and the Marital Wage Premium[J]. *American Economic Journal: Macroeconomics*, 2021, 13(4): 55 – 109.
- Pollmann – Schult M. Marriage and Earnings: Why Do Married Men Earn More than Single Men? [J]. *European Sociological Review*, 2011, 27(2): 147 – 163.
- Schoeni R F. Marital status and earnings in developed Countries[J]. *Journal of Population Economics*, 1995, 8(4): 351 – 359.
- Stevenson B, Wolfers J. Marriage and Divorce: Changes and their Driving Forces[J]. *Journal of Economic Perspectives*, 2007, 21(2): 27 – 52.
- Wang S – Y. Marriage Networks, Nepotism, and Labor Market Outcomes in China[J]. *American Economic Journal: Applied Economics*, 2013, 5(3): 91 – 112. ▲

(上接第 14 页)

- 王婧雯, 朱宇, 林李月, 柯文前, 肖宝玉. 城—城流动人口统计口径的改进与新口径下的流动特征和影响因素[J]. *地理科学进展*, 2023, 42(3): 464 – 477.
- 王泽强. 乡—城人口迁移对农村人口老龄化的影响——基于“年龄—迁移率”的定量分析[J]. *西部论坛*, 2011, 21(6): 27 – 33.
- 韦晓丹, 陆杰华. 季节性候鸟老人自评健康影响因素的实证分析——以海南省为例[J]. *北京社会科学*, 2017, (5): 99 – 107.
- 杨云彦. 人口迁移年龄模式的分解与合成[J]. *人口研究*, 1992a, (4): 15 – 22.
- 杨云彦. 中国人口迁移年龄模式及其应用[J]. *人口学刊*, 1992b, (4): 7 – 11.
- 郑真真, 杨舸. 中国人口流动现状及未来趋势[J]. *人民论坛*, 2013, (11): 6 – 9.
- 周皓. 中国迁移流动人口的统计定义——人口普查视角下的分析[J]. *中国人口科学*, 2022, (3): 17 – 30.
- 周皓, 刘文博. 中国省际流动人口流入地宏观选择机制的稳定性——兼论重力模型的理论扩展与实证检验[J]. *人口学刊*, 2023, 45(2): 80 – 98.
- 周皓. 我国流动人口年龄别流动率模式及其演变[J]. *华东师范大学学报(哲学社会科学版)*, 2023, (1): 185 – 201.
- 周皓, 雷琳旋. 中国人口省际终身迁移的时期—队列分析[J]. *人口研究*, 2023, 47(6): 107 – 125.
- 周爽, 黄匡时. 中国流动人口年龄别迁移模式研究——基于 2014 年全国流动人口动态监测调查数据的分析[J]. *西北人口*, 2015, 36(6): 23 – 28.
- 朱宇, 林李月, 李亭亭, 董雅静. 中国流动人口概念和数据的有效性与国际可比性[J]. *地理学报*, 2022, 77(12): 2991 – 3005. ▲