

人口与社会
Population and Society
ISSN 2095-7963, CN 32-1851/C

《人口与社会》网络首发论文

题目： 基于不同口径的省际迁移流动人口的结构比较
作者： 雷琳旋，周皓
网络首发日期： 2024-11-15
引用格式： 雷琳旋，周皓. 基于不同口径的省际迁移流动人口的结构比较[J/OL]. 人口与社会. <https://link.cnki.net/urlid/32.1851.c.20241114.1012.002>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于不同口径的省际迁移流动人口的结构比较

雷琳旋 周皓

摘要：不同的统计口径对应着不同的人群，并对应着不同的样本结构，进而可能导致分析结果存在结构性偏差。针对户籍口径与5年前常住地口径，本文利用七普数据，描述比较分析了两种口径下的中国迁移流动人口、以及5年前常住地口径下的户口迁移人口与流动人口、户籍口径下的流量与存量流动人口这三组样本间的结构差异。分析结果显示，这三对迁移流动人口存在明显的结构性差异。因此，实证分析应注意研究问题与分析样本之间的对应关系；并且，在强调年龄选择性和空间分布等特征的同时，应重视通过描述结构及其变化，以揭示中国人口迁移流动的新变化与新趋势。

关键词：人口迁移；人口流动；统计口径；结构比较

Abstract: Different statistical definitions correspond to different populations and sample structures, which may lead to structural bias in the analysis results. Using data from the Seventh National Census, this paper conducts a descriptive comparative analysis of the structural differences among three sets of samples: the migrant population under the hukou definition, the five-year residence definition, the migration and mobile population under the five-year residence caliber, and the flow and stock of the floating population under the hukou caliber. The analysis results show that there are significant structural differences among these three pairs of migratory and mobile populations. Therefore, empirical analysis should consider the correspondence between the research question and the analysis sample. Additionally, while emphasizing characteristics such as age selectivity and spatial distribution, it is crucial to focus on describing the structure and its changes to reveal new trends and changes in population migration and mobility in China.

Keywords: Migration; Floating Population; Statistical Caliber; Structural Comparison

一、引言

“谁在流动”是人口迁移流动研究中的基础问题，是了解与认识中国人口迁移流动的根本，也是讨论人口迁移流动研究其他方面（如过程与后果）的基础。已有很多研究讨论了流动人口的定义与测量，并比较了普查中不同统计口径的优劣^[1-6]，及由此反映的中国人口迁移流动的格局和变化趋势^[7]。当然其中仍有可商榷之处。虽然有学者已在“统计口径—总量结果”这一内在逻辑之下估计和比较了普查中三种统计口径下流动人口的总量差异^[2,5]，但不同统计口径之下的迁移流动人口结构却尚未引起足够的重视。本文是已有研究的延续^[1,2]。

流动人口的结构，既是“谁在流动”这一宏观问题的答案，也是后续统计分析中消除样本结构性偏差、保证统计结构更贴近现实的保证。结合人口迁移流动研究，样本结构偏差可以理解为流动人口调查样本中部分变量的分布与总迁移流动人口相应变量分布的偏离（如性别、民族等），从而导致统计结果的偏差，且这种偏差并无法使用统计方法予以消除^[8]。因此，基于重新审视流动人口的定义与测量，并进一步关注其结构，不仅有助于了解和认识中国的迁移/流动人口，还将为政策制定和社会治理提供更贴近现实的参考结果。

相比出生地口径，户籍口径和五年前常住地口径是目前人口迁移流动研究中广泛使用的

基金项目：教育部人文社科重点研究基地重大项目“中国人口长期均衡发展关键问题研究(22JJD840001)”

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

两种统计口径。其中户籍口径样本包括了存量与流量这两类流动人口^①；五年前常住地样本包括了户口迁移人口、返迁人口和流动人口等，具有混合属性的特征^[1]。混合了不同类型人口的样本将因为样本结构性偏差而无法准确回答各种现实问题。比如，在中国的社会环境之中，户口迁移人口和流动人口的流入地选择机制可能是完全不同的；由于流动时间的不同，存量与流量流动人口的流动动机与机制（包括流动时的社会背景与环境等）都存在差异。使用上述具有混合属性的样本会导致分析结果是不同类型的迁移流动人口的均值，而非真正具有因果关系的估计结果。因此，在人口迁移流动研究中，需要根据分析问题选择相对应的分析样本。这是从人口迁移流动研究角度理解样本结构的重要性。

正因如此，本文希望通过对比分析，揭示不同类别的迁移流动人口的结构性差异。有鉴于户籍口径和五年前常住地口径是目前人口迁移流动研究中广泛使用的两种统计口径，本文将首先比较这两种口径下样本间的结构差异；在此基础上，将5年前常住地样本更具体地细分为户口迁移人口与流动人口两种类别，以揭示其混合样本属性；将户籍口径下的流动人口细分为流量人口与存量人口。这种结构对比可以说明不同统计口径的样本结构之差异。限于篇幅，本文仅使用2020年第七次人口普查微观数据（后文简称“七普”），且只详细呈现年龄结构和教育结构，而其他结构将综合在最后的logistic回归中。

另外需要说明的是有关迁移流动人口的名称。本文中的迁移人口是指近五年内户口登记地发生变化的人口（系5年前常住地口径下样本中的一部分）；“流动人口”与普查中的二三款人相对应；而迁移流动人口则是上述两类人口之统称。同时，本文所讨论的均为省际迁移流动人口。

二、年龄结构

年龄是流动人口特征中最重要的维度之一，但流动人口的年龄结构并没有一个“标准人口”可以作为参考，因此，本文在分析时将直接比较：（1）5年前常住地样本与户籍口径样本、（2）5年前常住地口径下的户口迁移人口与流动人口、（3）户籍口径下的流量人口与存量人口，这三对迁移流动人口的结构，以揭示各样本的结构性差异。

（一）5年前常住地口径样本与户籍口径样本（流动人口）的结构对比

图1给出了七普5年前常住地口径样本与户籍口径样本的年龄金字塔。由图示可见，两个样本的年龄结构存在着较大的差异。5年前常住地样本中，15-30岁的人口相对更多；而户籍口径下的流动人口中，30岁以上年龄组的人口相对较多，但15-30岁的人口所占比例相对较低。

图示中最明显的是，户籍口径下的流动人口中，迁移倾向较强的20岁左右人口的比例远低于5年前常住地口径下的人口。这种差异与两个样本各自的内部年龄结构分布相关联。由于年龄金字塔中有一个前提限定条件：各年龄的比例之和为100%；因此，某些年龄组比例的提高，必然会使其他年龄组的比例有所下降。户籍口径样本中，30岁以上人口所占比例相对较高，从而使得其他年龄组的比例会相应的下降；即便两个样本的15-30岁年龄组的人口规模基本相同，但由于户籍口径样本中的其他年龄组的规模更大，从而使15-30岁组的人口比例会相对降低。其实仔细检查，这两个样本的年龄别流动率基本相同^[9]。因此，上述差异反映了年龄性别结构的内部分布之差异。

^① 有关三种口径的讨论请详见：周皓.中国迁移流动人口统计口径比较[J].人口与经济,2024,(04):32-44.

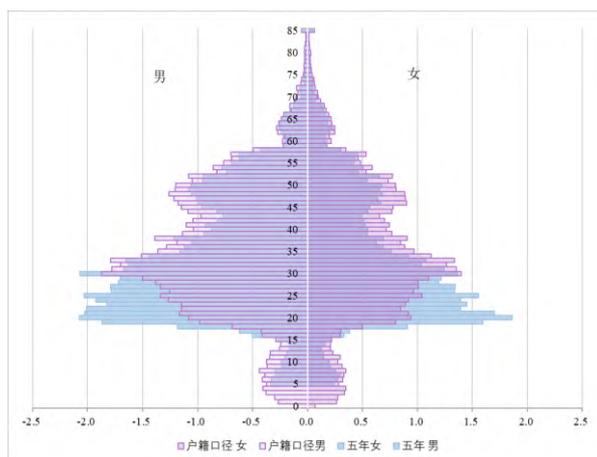


图 1 2020 年七普 5 年前常住地口径样本与户籍口径样本的年龄金字塔

注：根据 2020 年第七次人口普查数据绘制。图 2、图 3 和表 1 资料来源同图 1。

户籍口径样本的年龄分布，与其时间累积效应相关；而两种口径的样本结构也与总人口的年龄结构的变化有关。首先，户籍口径样本中 30 岁以上人口的比例相对较高，这可能是因为在本地居住五年以上的流动人口所形成的累积效应，相对而言年龄越大的，离开户口登记地的时间可能会越长。而且如果观察绝对规模，事实上 35 岁以后，存量流动人口大于流量流动人口。其次，中国人口持续多年的超低生育水平，使得总人口年龄金字塔底部急剧萎缩。则：一方面，每（五）年新增的流动人口规模会由于迁移流动最为活动的 20 岁左右的总人口规模的逐步减少而减少；另一方面，存量流动人口逐步退出流动状态；因而，我国未来流动人口的总体规模将必然会逐步下降。而在人口预测或规划中，应该使用新增的流动人口规模而非流量与存量之和的全部流动人口^[9]。

（二）5 年前常住地口径下迁移人口与流动人口的比较

五年前常住地口径下的户口迁移人口与流动人口在年龄结构上存在着非常显著的差异（请见图 2）。主要体现在：其一，在 5-15 岁人口中，户口迁移人口所占比例远高于流动人口；其二，在 15-30 岁的年龄段中，流动人口的比例远高于户口迁移人口；其三，而在 30-60 岁段中，户口迁移人口的比例都高于流动人口；特别的，在 30-55 岁左右，女性迁移人口所占比例远高于该年龄段的女性流动人口。

从整体年龄结构来看，在 20 岁左右的流动人口占比相对极高，这体现了流动人口的年龄选择性；随着年龄的提高，该比例逐步下降，在 40-45 岁之间形成一个小的凹槽。这个凹槽既与流动人口在该年龄段的生活稳定及降低流动倾向有关，而且也与总人口的年龄结构（对应 40 岁左右的凹槽）有关。

但迁移人口并未表现出在迁移流动最活跃的年龄段（20 岁左右）上的选择性；所占比例相对较高的年龄出现在 30 岁左右。特别是女性迁移人口，在 30 岁左右呈现出最高的比例。这说明：虽然人口迁移的政策性壁垒逐步消除、人口迁移流动的自由权益得到保障，特别是随着新型城镇化政策的落实，迁移人口的规模和比例都在逐步扩大，但户口迁移的制度性壁垒事实上仍然存在，户口迁移只是针对具有特定选择性的群体而开放。当然，不同地域、不同规模等级的城市中，户口迁移的困难程度也不尽相同。但至少从年龄角度看，户籍制度对于迁移倾向相对较大的年龄段人口而言，其难度依然存在。另一方面则可能是，年轻一代的流动人口一方面可能是学业流动，另一方面更可能是他们的观念相对不再保守，在市场经济条件下不再十分在意并刻意强调户口问题。但不论如何，可以看到的是，户籍的制度性壁垒依然存在。

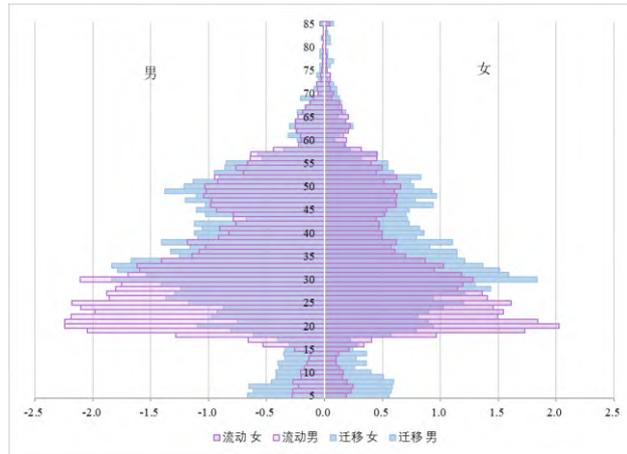


图 2 2020 年七普 5 年前常住地口径下迁移人口与流动人口的年龄金字塔

(三) 户籍口径下存量流动人口与流量流动人口的对比

户籍口径下存量流动人口与流量流动人口的年龄结构同样存在着显著的差异(具体请见图 3)。如果说流量人口明显体现出 16-30 岁的年龄选择性,那么,存量人口更表现出 35 岁时的年龄选择性。

由于存量人口不包括 5 岁以下,因此,5 岁以下部分只有流量人口,当然他们中有很大部分儿童的出生地可能就是在流入地,只是因为他们没有流入地的户籍而被视为流动人口。在 5-15 岁年龄段,流量流动人口的比例在逐步缩小,并在 13 岁时达到最小;这一段与义务教育阶段相对应,为保证其学业,这段年龄的人口会相对比较稳定,从而使流量人口的规模下降。在 16 到 30 岁段,则以流量流动人口为主。正如年龄别流动率在 16-20 岁段的迅速上升^[9],这一年龄中的流量流动人口比例也迅速上升。并在 20 至 30 岁间维持在相对较高的比例上。30 岁以后,流量流动人口的比例逐步下降,在 40 岁时形成一个下凹,然后再在 50 岁有一个小幅凸起。但在 60 岁时则有突然的下凹。整体来看,这种凹凸的形状与中国人口年龄结构的形状相近。

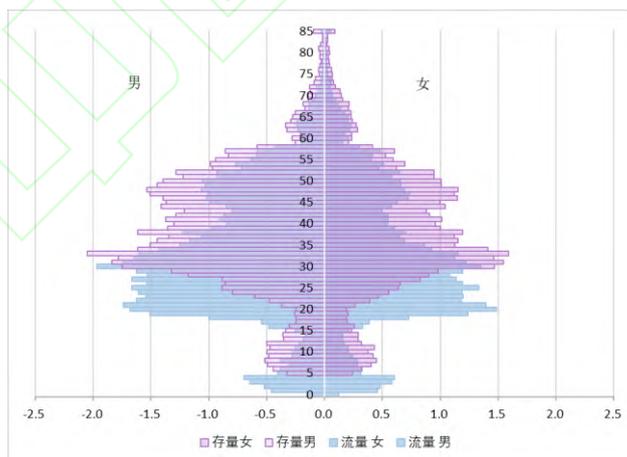


图 3 2020 年七普户籍口径下存量与流量流动人口的年龄金字塔

存量人口的年龄结构变化趋势有些类似于被推迟的流量人口年龄结构。在 5-20 岁阶段,存量人口的比例随年龄而不断缩小;并在 20 岁时处于最小值,即所占比例最低;20 岁至 30 岁之间,存量人口则逐步随年龄上升,并在 32 岁左右达到最大,其后呈现出波折型的下降,其中 45 岁时也形成一个较小的下凹,在 50 岁时又有一个小凸起。其后在 60 岁时同样也有

一个突然的下凹。

如果把存量人口的年龄金字塔向下移动 5 岁，则存量人口在 20 岁处的下凹正好和流量人口在 15 岁处的下凹相重合。这说明了教育阶段对人口流动的影响。另一方面，从流量与存量的角度来看，存量人口在 30—50 岁之间持续保持较大规模是导致流动人口的年龄别流动率在该年龄段的下降速度减缓的主要原因；流量人口的规模在该年龄段则仍然处于持续下降的过程。

三、教育结构

不同口径之下的迁移流动人口的教育结构请见表 1。

5 年前常住地和户籍口径的迁移流动人口，虽然都以初中及以下受教育水平为主，但两者各自内部的教育结构存在较大的差异。从大学和研究生两个层级看，5 年前常住地口径下的迁移流动人口的比例远高于户籍口径。如 5 年前常住地口径下的男性研究生比例达 2.90%，女性更是高达 4.32%，远高于户籍口径下的比例（分别为 1.45% 和 1.90%）。大学受教育水平的比例也是如此，5 年前常住地中的男性大学生比例达 25.95%，女性达 29.06%，远高于户籍口径下的流动人口（分别为 19.52% 和 20.89%）。相反的，5 年前常住地口径下，高中及以下受教育水平的比例则均略低于户籍口径。如果从综合性的平均受教育年限看，5 年前常住地口径的男性人口平均受教育年限为 10.99 年，而女性人口则高达 11.08 年，远高于户籍口径下的 10.23 年和 10.01 年。

那么 5 年前常住地口径下的迁移流动人口具有相对较高的受教育年限是否是因为 5 年前常住地样本中包含了户口迁移人口呢？为此进一步将 5 年前常住地样本分解为户口迁移人口与流动人口。显然，这一解释并不成立。相反，从平均受教育年限来看，5 年前常住地口径下，户口迁移人口的平均受教育年限，不论男女都低于该口径下的流动人口，如男性为 10.67 年，女性为 10.92 年，均低于相应的流动人口（分别为 11.04 和 11.11 年）。从结构上看，户口迁移人口在小学及以下、初中和研究生这三类的比例高于流动人口，而高中和大学这两类的比例却相反低于流动人口。如高中受教育水平上，流动人口的男女各自的比例为 18.84% 和 16.13%，高于户口迁移人口的 13.65% 和 12.35%，相对高出大概五个百分点左右。而在大学受教育水平上，流动人口的比例为 26.70% 和 29.82%，亦比户口迁移人口高出约五个百分点左右。正是因为高中和大学部分的比例相对较高，从而导致五年前常住地口径下，流动人口的平均受教育年限高于户口迁移人口。从某种意义上也说明，户口迁移人口并未表现出教育选择性；相反，初中及以下人口更有户口迁移倾向。或者说，五年前常住地口径所覆盖的流动人口具有更高的受教育水平；或者是 5 年前常住地口径比户籍口径下调查得到更多的高受教育水平人口。

表 1 不同口径的迁移流动人口的教育结构（%）

受教育程度	五年前常住地口径		户籍口径		五年前常住地口径样本				户籍口径下的流动人口			
	男	女	男	女	迁移男	迁移女	流动男	流动女	流量男	流量女	存量男	存量女
未上过学	0.35	1.44	0.66	2.05	0.36	0.89	0.35	1.55	0.71	1.98	0.58	2.16
学前教育	0.77	0.88	2.05	2.35	1.56	1.57	0.65	0.74	2.96	3.51	0.78	0.84
小学	11.88	14.62	14.39	18.09	15.44	18.73	11.31	13.81	12.44	15.15	17.11	21.92
初中	40.01	34.17	42.70	38.35	42.52	34.32	39.61	34.14	41.99	36.95	43.69	40.17
高中	18.12	15.51	19.24	16.36	13.65	12.35	18.84	16.13	19.20	16.53	19.30	16.15
大学专科	10.92	11.11	9.61	9.95	8.14	8.52	11.37	11.61	9.87	10.75	9.24	8.91
大学本科	15.03	17.95	9.91	10.94	13.17	16.61	15.33	18.21	11.36	13.13	7.88	8.10
硕士	2.39	3.79	1.25	1.72	3.57	5.66	2.20	3.43	1.27	1.83	1.22	1.58

博士	0.51	0.53	0.20	0.18	1.60	1.34	0.34	0.38	0.20	0.18	0.20	0.18
合计	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
平均受教育年限	10.99	11.08	10.23	10.01	10.67	10.92	11.04	11.11	10.31	10.21	10.10	9.75

再者，比较户籍口径下流量流动人口与存量流动人口可以看到，存量人口中小学和初中受教育水平的比例高于流量人口；高中两者基本相同，但在大学和研究生受教育水平上，流量流动人口则高于存量人口。如流量人口中，男性和女性的大学比例分别为 21.23%和 23.88%，分别高于存量流动人口的 17.12%和 17.01%。当然，这种受教育结构与流量/存量人口的年龄结构存在一定关联。如果分年龄别考察流动人口的受教育水平，则可以发现，在年轻一代的流动人口中，大学及以上比例逐步占据了主流；而随着年龄的递增，初中及小学部分的流动人口比例逐步上升（见图 4）。虽然这是省际流动人口的教育结构，但也可以表现出存量人口与流量人口在受教育水平上的差异。

上述有关教育结构的对比分析可以看出，5 年前常住地口径调查得到高受教育水平的人口，且户口迁移人口未体现出明显的教育选择性（仍以初中及以上受教育水平为主）；户籍口径下，流量与存量流动人口在教育结构上并不存在显著差异；更关键的是年龄别或代际之间的差异，即年轻队列的流动人口中以大学及以上的高受教育水平为主。这既体现了代际或年龄之间的差异，也体现了人口迁移流动行为与个体生命周期之间的相关性^[9-12]。

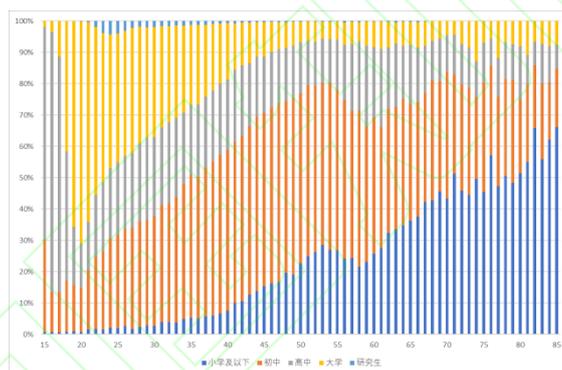


图 4 省际流动人口 15 岁及以上年龄别的教育结构^[13]

本文在分析过程中还考察了户籍、婚姻与户类型等方面的特征。为节约篇幅，本文不再报告这些特征差异性比较结果。如果需要，可以与作者联系。

四、综合对比-logistic 回归

为能够综合体现各类迁移流动人口之间在结构上的差异，本文选用 logistic 回归的形式来予以呈现。两个模型分别对应 5 年前常住地口径下户口迁移人口和流动人口（参照组）、户籍口径下存量流动人口和流量流动人口（参照组）的比较。需要说明的是，(1) 本文无意讨论因果关系，且下文的 logistic 回归也不反映任何因果关系，甚至可能是倒置的因果关系；它只是呈现类别之间的结构性差异。(2) 由于五年前常住地口径下的迁移流动人口与户籍口径下的流量样本之间存在较大的重合，导致无法利用回归分析比较这两个样本；因此，此处综合性对比只针对两种口径内各自的不同类型。在解释表中结果时，应该注意虚拟变量所对应的分类变量的整体性，即如果系数显著，说明该分类变量在两类迁移流动人口之间存在显著差异。(3) 变量的选择：人口普查中所能选用的变量本来就不多，模型尽可能地加入更多的变量；且尽量使用以往研究中曾涉及的、能够说明样本结构差异的变量。

表 2 2020 年七普数据 logistic 回归结果

	五年前常住地 (0=无户口)		户口标准: (0=流量)	
	OR	se	OR	se
年龄	1.0087***	(0.0019)	1.0208***	(0.0007)
性别 (0=男)	1.0953**	(0.0405)	1.0729***	(0.0144)
婚姻状况 (0=未婚)				
有配偶	1.7716***	(0.0957)	2.1315***	(0.0406)
离婚	3.3644***	(0.3475)	1.9885***	(0.0944)
丧偶	2.5507***	(0.3998)	1.7773***	(0.1111)
教育 (0=未上过学)				
学前教育	1.4155	(1.5124)	0.6708	(0.2087)
小学	2.5652***	(0.6253)	0.9808	(0.0615)
初中	3.2525***	(0.789)	0.9742	(0.0604)
高中	2.7360***	(0.673)	1.1467*	(0.0726)
大学专科	2.3894***	(0.6008)	1.1874**	(0.0775)
大学本科	2.5918***	(0.6531)	1.0881	(0.0719)
硕士	5.6529***	(1.4971)	1.5720***	(0.1278)
博士	19.0680***	(5.636)	1.6579***	(0.2539)
是否有农村土地承包经营权 (0=有)	0.2461***	(0.0114)	0.9912	(0.0138)
户口登记地类型 (0=乡)				
镇的村委会	1.3606***	(0.0617)	1.0696***	(0.0167)
镇的居委会	2.5267***	(0.1903)	0.9961	(0.0274)
街道	2.2505***	(0.1603)	0.9724	(0.0233)
工作情况 (1=工作)				
在职休假	1.8307***	(0.2600)	0.7715***	(0.0459)
未工作	.8670***	(0.0411)	0.8318***	(0.0137)
常数	0.0175***	(0.0045)	0.1796***	(0.0125)
N	45884		108483	
LL	-12190.342		-70607.958	
Pseudo R2	0.0808		0.0453	
LR chi2(19)	2144.19		6698.94	

*** p<0.001, ** p<0.01, * p<0.05

数据来源: 根据七普 0.96‰原始数据计算得到。

从五年前常住地来看,以无本地户口的流动人口为参照组,户口迁移者在诸多因素上与参照组存在着显著的差异。年龄变量每提高1岁,成为户口迁移人口的可能性会提高0.87%;相对于男性而言,女性更可能成为户口迁移者;相对于未婚人口,已婚人口或离异丧偶者更有可能成为户口迁移者;受教育水平越高,成为户口迁移者的可能性越大;相对而言,没有

农村土地承包经营权的人口（类似于原有的非农业户口的人口）更不可能成为户口迁移者，这一点说明，在户口迁移中很大部分仍然是由农业户口向非农户口转移的人口，也同样说明低受教育水平人口在户口迁移人口中所占比例相对较高的原因（或者某种意义上低受教育水平和高受教育水平的两类人群）。如果再结合户口登记地类型，可以发现，这部分人口更能更多的来自于镇的村/居委会。以有工作的为参照组，在职休假的更可能是户口迁移者；而未工作的成为户口迁移者的可能性却越低。这些特征性描述综合性地说明了五年前常住地口径下户口迁移人口与无户口的流动人口在整体结构上的显著差异。

户籍口径下流量人口与存量人口的结构综合对比结果可见表 2 右侧两列。其结果与五年前常住地口径下的结果基本相同。年龄越大越可能成为存量人口，这与存量人口的年龄分布相对应；存量流动人口与流量人口在性别结构、婚姻结构等方面同样存在显著差异；在教育结构上，部分虚拟变量并不显著，但在高中、大专、硕士、博士阶段存在显著差异；是否有农村土地承包经营权这一变量并不显著，这说明在流量与存量人口中，农业人口与非农人口这一结构上并不存在显著的差异，特别是在控制了其他变量的条件下。户口登记地类型亦只有镇的村委会呈显著，而其他两类均不显著，这表明存量与流量人口中城镇人口基本相同，只是在镇的村委会上略有区别。总体来看，户籍口径下的流量流动人口与存量流动人口在各种结构特征上存在着显著的差异。

上述结构性综合对比分析表明，不论是五年前常住地口径下的户口迁移人口和流动人口，还是户籍口径下流量与存量的流动人口，都明显地存在结构性的差异。因此，在利用各种口径下的样本进行分析时，应注意研究问题与分析样本之间的对应关系，以期得到相对一致可信的研究结果。

五、结论与讨论

（一）结论

本文采用 2020 年第七次人口普查数据，通过比较五年前常住地样本与户籍口径下的流动人口、5 年前常住地口径下的户口迁移人口与流动人口之间、户籍口径下的流量人口与存量人口这三对迁移流动人口在年龄和教育结构下的差异以及样本回归结果的差异。主要结论如下：

1. 户籍口径与 5 年前常住地口径相比：（1）5 年前常住地样本中 20 岁左右人口的比例相对较高；（2）5 年前常住地口径下的迁移流动人口平均受教育年限更高，且迁移流动人口在高中和大学的比例高于户籍口径，显示出较强的教育选择性；（3）中国人口的超低生育水平导致总人口年龄金字塔底部缩小和五年内新增流动人口减少，并且随着存量流动人口退出，未来流动人口总规模将逐步下降。

2. 五年前常住地口径下的户口迁移人口与流动人口相比：（1）迁移人口主要突出在 5-15 岁和 30-60 岁这两个年龄段，而流动人口的突出年龄段集中在 15-30 岁；（2）户口迁移人口的平均受教育年限低于同口径下的流动人口；从结构上看，户口迁移人口在小学及以下、初中和研究生这三类的比例高于流动人口，而高中和大学这两类的比例却相反更低。从某种意义上也说明，户口迁移人口并未表现出教育选择性；相反，初中及以下人口更有户口迁移倾向。其背后的原因尚需进一步讨论。

3. 户籍口径下存量与流量流动人口：（1）流量人口突出集中在 16-30 岁，存量人口则主要集中在 35 岁以后。（2）存量人口的年龄结构类似于被推迟 5 年的流量人口年龄结构。

（3）户籍口径下，流量与存量流动人口在教育结构上并不存在显著差异；更关键的是年龄别或代际之间的差异，即年轻队列的流动人口中以大学及以上的高受教育水平为主。

4. 综合性的回归分析表明，5 年前常住地口径下的户口迁移人口在年龄、性别、受教育水平等诸多因素上与参照组（无本地户口的流动人口）存在着显著的差异；户籍口径下流

量与存量人口在是否有农村土地承包经营权这一变量并不显著,这说明流量与存量人口中农业和非农业人口结构无显著差异,但在年龄、性别、婚姻结构、受教育水平等方面却存在显著差异。

(二) 讨论

上述分析说明各类不同的子群体内部存在着较大的结构性差异,因此在针对具体研究问题时需要选择合适的对应样本才可能真正回答研究问题。基于上述结果,以下几点可能需要进一步讨论。

1. 结构差异的可能原因

如果说5年前常住地口径与户籍口径这两个样本的结构差异反映的是不同人群,那么,同一口径之下的不同样本既是两类不同的群体,也反映了其背后的社会经济背景之不同。流量与存量流动人口之年龄差异与户籍口径的时间累积效应有关;流动人口教育结构的代际差异不仅体现了中国人口整体受教育结构的变化,而且和5年前常住地口径之下的迁移人口与流动人口在年龄结构上的差异类似,都表明虽然人口迁移政策逐步放宽,但户口迁移的制度性壁垒依然存在;年轻一代的流动人口也可能是因学业流动或更关注市场经济下的生活,不再过分强调户口问题。同时,年龄结构(特别是户籍口径下流动人口的年龄结构)不仅与流动人口整体年龄结构的改变有关,而且也与多年来总人口生育率持续下降而导致的底部年龄人口规模缩减有关。因此,这些结构的比较分析可以从一定程度上揭示中国社会经济与人口的发展历程对中国人口迁移流动的影响。

2. 重视人口迁移流动中的基本问题。

谁在流动?这是一个与流动人口定义(统计口径)、测量相关的问题,也是需要通过流动人口结构描述来体现与回答的问题,更是了解与认识中国迁移流动人口的基本问题。只有回答了“谁在流动”,才有可能真正回答“为何流动”、流动后果及政策需求等一系列后续的问题。要回答谁在流动,首要的就是统计口径和测量问题;其次是清晰描述结构。已有许多文献讨论了流动人口的统计口径与测量问题,以期解决总量估计;而本文虽然描述性分析了流动人口各类子总体的结构,但目前已有的多数研究或多或少都存在着忽略包括测量和结构在内的这些基本问题,而只是更关注于后续的分析。这可能和论文发表压力、研究风气与研究基础等都有关系。但回答这些基础性问题的的重要性并不亚于各种所谓的“分析”,本文上述的结构分析也表明,结构不同,谁在流动这一问题的答案是不同的。

再进一步,本文只是描述性分析了省际流动人口在不同统计口径下的不同子总体的结构,但流动人口可以按照不同的标准来划分,如城乡标准、是否跨省、以及市内人户分离与流动人口等,这些不同划分标准之下的样本结构差异同样是在回答谁在回答这一问题,同时还能部分揭示这种分类标准背后不同人群的可能的迁移流动原因。

因此,可以说,对迁移流动人口结构的分析既是帮助我们了解当下中国人口迁移流动特征的重要“窗口”,也是进一步探索中国人口迁移流动的转变历程的基础“抓手”。

3. 关注中国人口迁移流动的新变化与新趋势。

上述分析结果表明:30岁以上的存量流动人口所占比例相对较高,这种由于时间累积效应而导致的相对规模较大的存量流动人口是完全不同于以往历次普查的新趋势。这说明,一方面从研究看反过来说明需要考虑何谓流动人口;另一方面从流入地政策供给角度,应该更加关注这些长期“流而不动”的流动人口的政策需求。相对应的,如果将流量流动人口的结果与以前历次普查相比较,可以发现,虽然年龄别流动率仍然体现了年龄选择性,但从年龄金字塔来看,20岁左右迁移流动最活跃的年龄段的流动人口规模已经有较大比例的减少。这个新变化主要源自于流出地人口结构的变化,特别是20岁左右农村人口的“内陷”^[14]。而流出地(农村)人口的这种“内陷”反过来还将对未来的中国迁移流动人口的结构产生重要而深远的影响作用。因此,当前的结构既是以往人口发展历史的结果,又将成为未来人口

发展的基础；既能体现当前的状况，又可以通过对比而显示新时代的新变化，并进而揭示未来可能的发展趋势。

教育结构的变化提示我们，流动人口内部的结构变化，这是新变化，形成了新形势，需要予以关注。比如，五年前常住地口径下的迁移流动人口显示出较强的教育选择性，而户口迁移人口则以初中及以下教育水平为主；年轻一代的流动人口中，大学及以上受教育水平逐步成为主流，而随着年龄的增加，初中及小学教育水平的比例逐步上升，反映了代际和年龄之间的差异，即表明人口迁移流动与个体生命周期的相关性，也反映了社会变迁过程中迁移流动内部的结构变化。而这种不同结构的人群对于各类社会政策的需求与意愿是各不相同的。

总之，需要在强调年龄选择性等部分特征与空间分布等稳定性的同时，更应该通过描述结构及其变化，揭示中国人口迁移流动的新变化与新趋势。中国的人口迁移流动本身就是极其复杂的社会现象，与社会发展、个体发展之间存在着强烈的相互作用关系。重视基本问题，在强调稳定的同时，关注新变化与新趋势，才有可能更加全面地认识中国的人口迁移流动。

参考文献

- [1] 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, (03): 17-30+126.
- [2] 周皓. 中国迁移流动人口统计口径比较[J]. 人口与经济, 2024, (04): 32-44.
- [3] 朱宇, 林李月, 李亭亭, 等. 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.
- [4] 王婧雯, 朱宇, 林李月, 等. 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(03): 464-477.
- [5] 丁金宏, 常亮, 陈益豪, 等. 人口流动分析范式及中国人口省际迁移流动分量辨识[J/OL]. 地理学报, 1-14 [2024-07-08].
- [6] 段成荣, 孙玉晶. 我国流动人口统计口径的历史变动[J]. 人口研究, 2006, (04): 70-76.
- [7] 严善平. 当代中国地区间人口迁移的基本格局及变化趋势[J]. 人口与经济, 2023, (05): 112-125.
- [8] 周皓. 样本结构性偏差与因果推论——基于实验数据的分析[J]. 社会研究方法评论, 2023, 4(02): 126-188.
- [9] 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, 55(01): 185-201+206.
- [10] Bernard, Aude, Bell, Martin, Charles-Edwards, Elin(2014), Life-Course Transitions and the Age Profile of Internal Migration. *Population and Development Review*, Vol.40(2), 213-239.
- [11] Bernard, Aude, Bell, Martin, Charles-Edwards, Elin(2016), Internal Migration Age Patterns and the Transition to Adulthood: Australia and Great Britain Compared. *Journal of Population Research*, Vol.33(2), 123-146.

- [12]周皓,陈晓雄.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[J].人口与发展,2024,30(02):2-14+125.
- [13]周皓,陈晓雄.21世纪以来我国流动人口教育结构的历时演变.载于韩嘉玲主编.中国流动儿童教育发展报告(2021~2022).[M].社会科学文献出版社,2023:26-50.
- [14]王广州,刘旭阳.城乡人口迁移流动新趋势——基于人口普查数据分析[J].中国特色社会主义研究,2023,(06):76-87+2.

