

城市化与人口流迁

中国迁移流动人口统计口径比较

周 皓

(北京大学 中国社会与发展研究中心/社会学系, 北京 100871)

摘 要: 统计口径是认识中国人口迁移流动的基础与根本。不同统计口径对应不同的迁移流动人口以及不同的样本结构。以“统计口径—总量结果”为内在逻辑, 首先比较讨论普查视角下的三种口径, 以正确认识各统计口径的含义与特点; 继而利用第五、第六和第七次全国人口普查数据, 重新估计不同统计口径下的迁移流动人口比例, 以真正认识中国人口迁移流动的整体强度, 加深对中国迁移流动人口的认识。文章认为: 户籍口径和出生地口径作为非时期指标具有时间累积效应, 五年前常住地口径存在混合样本属性和未包括县内迁移流动人口的低估问题等; 目前仍然没有一种口径能够合适准确地描述中国迁移流动人口的规模与比例; 有关迁移流动人口的统计口径亟须调整, 需要在明确调查对象、完备登记信息的原则下强调空间范围、时间性以及中国社会制度背景这三个统计口径的调整标准。同时, 文章建议应加强人口负增长条件下的中国人口迁移流动趋势研究。

关键词: 人口迁移; 人口流动; 统计口径; 普查数据

中图分类号: C922 **文献标识码:** A **文章编号:** 1000-4149 (2024) 04-0032-13

DOI: 10.3969/j.issn.1000-4149.2024.00.031

如何认识中国人口迁移流动的规模与强度? 如何认识中国迁移流动人口的内部结构? 如何更准确地使用普查数据描述流动人口强度、结构及其变化? 这些问题涉及对当前乃至未来一段时间我国人口迁移流动的基本认识与发展趋势的判断, 更关涉对人口迁移流动与社会经济发展关系的理解与判断。其中, 统计口径是根本, 样本结构是基础。只有从统计口径出发, 深刻认识不同统计口径之间的区别及各自存在的潜在问题, 真正理解各种统计口径之下的迁移流动人口的内部结构及其差异性, 并根据不同的研究目标选择合适的分析样本, 才有可能真正揭示迁移机制、动因与后果, 才能真正理解中国的人口迁移流动。

收稿日期: 2023-12-13; **修订日期:** 2024-05-13

基金项目: 教育部人文社会科学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: 周皓, 法学(人口学)博士, 北京大学中国社会与发展研究中心研究员, 社会学系教授, 北京大学人口迁移研究中心主任。

2020 年第七次全国人口普查（以下简称“七普”，“四普”、“五普”、“六普”简称下同）数据表明：户籍口径下，我国人户分离人口规模达 49276 万人，其中市内人户分离规模达 11694 万人，跨越乡镇街道以上行政区域的流动人口规模已达 37582 万人（占我国总人口的 26.62%）；其中省际流动人口占全国总人口的 8.8%^[1]；五年前常住地口径之下（根据长表人口计算），“本省其他县市区”和“省外”的人口共为 5761821 人，占长表合计人口（131271202 人）的 8.32%，其中省际迁移流动人口占 3.93%^[2]；出生地口径之下（根据长表人口计算），我国人口终身迁移比例（居住在出生地（县、市、区）以外的人口占全国总人口的比例）为 22.60%，其中省际终身迁移人口占 10.28%^[2]。三种口径之间的差异如此巨大，那么到底哪个口径真正表示我国的人口迁移流动强度？

近期有许多研究关注了我国人口迁移流动的定义问题。如朱宇等曾讨论了“人口流动”在测量人口迁移流动事件时的五大失效和失真问题，并指出应扩大采用五年前常住地口径以便于国际比较^[3]；严善平将三种统计口径分别定义为终身迁移、期间迁移和人户分离迁移，并分别释明各自属性特征^[4]；丁金宏试图结合三种统计口径间接估计中国人口迁移流动的总量问题^[5]。这些研究对澄清和解释目前中国人口迁移流动状况（总量估计与结构分析）提供了基础。虽然周皓曾从人口普查角度讨论了中国迁移流动人口的统计口径问题^[6]，并指出，五年前常住地中户口迁移人口与流动人口、户籍口径下的流量（在流入地居住不满五年）与存量（在流入地居住五年以上）流动人口存在显著的结构差异；但没有对官方发布的迁移人口指标的属性特征、局限性展开研讨^[4]，也未深入比较不同口径定义下不同类别的迁移流动人口的结构差异。

正因统计口径事关中国人口迁移流动的测量与代表性，也事关后续的相关研究（特别是样本结构与因果推断）和政策服务，本文将首先比较与讨论中国迁移流动人口的不同统计口径，以正确认识各统计口径的含义与特点；并利用 2000 年以来的三次人口普查数据，重新估计三种统计口径之下中国迁移流动人口的总体比例及各口径间的相互关系，以真正认识中国人口迁移流动的整体强度，加深对中国迁移流动人口的认识。

一、中国迁移流动人口统计口径的比较

国际通用的人口迁移（migration）统计口径一般包括出生地、五年前常住地和最近一次迁移^[7]。本质上，迁移流动的统计标准是常住地是否改变。只是在户籍制度作为社会基本制度的中国社会情境之下，我国的人口迁移流动概念在国际通用定义的基础上由于加入户籍而使常住地的识别变得相对较为复杂。

按照目前国内学界的共识，迁移人口一般系指户口登记地随居住地发生改变的人口。当然在研究对象同时包含迁移人口与流动人口时，有些研究也使用迁移人口这一概念来统领，这时需要特别注意该研究中的迁移人口是不是纯粹意义上的户籍迁移人口^①。流动人口则是户口不在本地，但在本地居住半年以上的人口，或者是户口不在本地、在本地居住不足半年，但离开户口登记地半年以上的人口，即对应普查中的第二、三条款人员；它需要符合以

① 例如“返迁”一词则并非指返回原户籍登记地，许多情况下是指返回户籍登记地的流动人口。在中国的情境之下，它并不与迁移对应，而与流动对应。

下三个条件：①行政边界上，常住地至少要跨越乡镇街道；②时间上须在流入地居住满半年以上；③户口登记地未发生变化（且不同于现住地）。

从普查角度看，我国自“五普”以来，人口迁移流动主要包括三种统计口径：出生地、五年前常住地和户籍登记地。但“五普”、“六普”、“七普”之间在某些具体选项上存在一定的差异。其中，①出生地口径是出生地点与现住地之间的比较，其中的选项包括：本县（市、区、旗）、本省其他县市和省外三个选项。②五年前常住地登记了五年前“11月1日常住地”，选项包括“1. 本县（市、区、旗）”和“2. 其他地区”（要求填写具体省市县地址）。③最后是户籍登记地口径（简称户籍口径）：“五普”中，这一口径以“户口登记地”和“最近一次来本地居住时间”为标准；“六普”和“七普”则以“户口登记地”和“离开户口登记地时间”为标准，即：户口登记地在“3. 本县（市、区、旗）其他乡镇街道”或“4. 其他县（市、区、旗）”（登记具体省市县），且离开户口登记地时间选择“半年以上”（对应第3—9选项）；这个定义类似于“四普”中“二三款”人，也是目前使用最为广泛和普遍的定义。有关定义可参考历次普查问卷和相关文献^[6]。

针对上述定义，需要强调以下几点：①迁移人口与流动人口这两个概念在中国情境下不应混用。中国情境下，迁移人口与流动人口的根本区别就在于户口登记地是否随居住地而改变，并由此导致迁移人口与流动人口存在着根本性的特征与结构差异^[8]。②“六普”和“七普”无法真正识别户口迁移人口，这是因为“六普”和“七普”登记的是“离开户口登记地时间”，而非户口迁入或来本地居住的时间。相反，“五普”登记了最近一次来本地居住的时间，从而可以将户口在本地的常住人口区分出一直居住本地和近年来迁入户口的迁移人口^[6]。③正因为出生地口径和五年前常住地口径无法区分迁移与流动人口，并为简化起见，许多研究会以迁移人口来统一称谓这两种口径之下的迁移流动人口。但这两个口径都同时包含了迁移人口与流动人口，因此，不同研究间的比较需要注意迁移人口在研究中的定义及研究对象。④相应的，本文建议三种口径可以对应三种不同的称谓以示区分：出生地口径对应终身迁移人口；五年前常住地口径对应迁移流动人口；而户籍口径则对应流动人口。

此外，这三种口径在时间与空间方面存在一定的差异性特点，并对迁移流动人口总量的估计产生不同的影响。

时间方面，出生地口径和户籍口径都具有时间累积效应，表现为存量概念；而五年前常住地则是固定时间点之间的比较，具有时期性的流量特征。以出生地口径为例，该口径只比较了出生地点与现居住地之间的关系。如对于一位80岁老人而言，他的迁移流动既可能是在60年以前，也可能是在近几年。因此，出生地口径的统计结果汇集了各个历史时期的迁移流动，具有时间累积效应，且带有明显的历史印记^[9]。户籍口径也类似，流动人口可能已在流入地居住了一辈子（如某些无户口的老人），也可能是在改革开放之初就流入本地，或者是普查前某一个较近的时间点流入本地。因此，户籍口径同样具有时间累积效应。正因如此，周皓在讨论年龄别流动率的历史变化时将流动人口按照离开户口登记地时间区分为“流量流动人口”（普查前五年内离开户口登记地的流动人口）和“存量流动人口”，并使用流量流动人口来表示时期性的流动人口规模及相应特征等^[10]。区分流量与存量的意义不仅在于它可以清楚地表示普查前某一段时期内的人口流动状况，而且这种流量流动人口的定

义既为后续的人口预测及相应的人口流动特征的讨论提供相对更为准确的基础数据，也可以解决朱宇等提出的失真与失效的多数问题^{①[3]}。

空间方面，跨越乡镇街道是人口迁移流动的判断标准之一。但目前只有户籍口径可以识别出县内跨乡镇街道、省内县际和省际这三种地域范围的流动人口；出生地口径和五年前常住地口径只能识别出省内县际和省际这两类迁移流动人口。因此，虽然五年前常住地口径是一个时期指标（以五年为时间标准）（2015年“小普查”时还调查了“一年前常住地”），有助于从时期角度了解中国人口迁移流动的情况，但需要注意该口径存在两个缺陷：其一，混合样本属性^[6]；其二，忽略了在我国人口迁移流动中占据重要地位的“县内跨乡镇”的迁移流动，从而会显著低估中国人口迁移流动的规模与强度。正因为五年前常住地口径存在这两个缺陷，在使用时需要特别小心。当然，这两个缺陷可以通过一定的方法来予以弥补：对于混合样本属性问题可以通过识别该案例的户口登记地是否与现居住地相同来解决（但仍然很难有效识别迁移与返迁人口）；而后者则可以通过改变普查选项的设计来解决（即增加县内其他乡镇的选项）。

对迁移流动人口总规模的估计方面，出生地口径和户籍口径由于时间累积效应而无法估计时期的迁移流动人口规模；虽然周皓曾提出使用“流量流动人口”的概念，但它只是剔除时间累积效应、使户籍口径更符合时期指标的简单改进，却未从根本上改变由于忽略了迁移人口而导致低估的问题^[10]。而五年前常住地口径虽然是时期指标，却存在混合样本属性和由于未包括县域内的迁移流动人口而导致的低估问题（当然，这一问题可以通过改进普查题项与选项来解决）。可见这三种统计口径都存在这样或那样的问题，均无法真正有效地估计年度性或时期性的中国人口迁移流动规模与强度。因此，在使用人口迁移流动概念时，应该对这些概念及相应的研究对象有清醒的认识与辨别，如此才有可能真正认识中国的人口迁移流动。正如周皓所提出的，普查题项最好能增加离开户口登记地后的流入地及最近一次迁移流动时间^[6]。而最近一次迁移流动时间，既能结合户籍口径体现中国特色，又能有效识别迁移人口与流动人口，并能与国际标准相一致。

由此，三种口径之间的关系其实并非那种纯粹的包含关系，即出生地口径下的终身迁移人口并非包含了其他两种口径，即三种口径既有相互重合部分，又有相互区分部分（见图1）。如果从结果来看（见表1），2020年“七普”中出生地口径下的终身迁移人口比例为22.60%，而户籍口径下的流动人口比例则达到26.62%，显然出生地口径无法完全包括户籍口径，这既可能是因为出生地口径未包括县内的迁移流动人口，同时也是因为它未包括那些出生在本地的流动人口（如出生在流入地的流动儿童等），这一点是需要修正的。

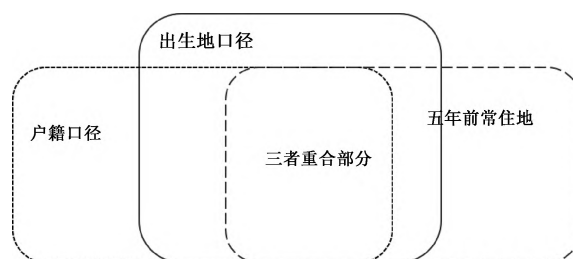


图1 人口普查中三种类型迁移人口的关系

① 当然朱宇文章中提出的第2点“基于‘人户分离’判定的人口流动数据系统地遗漏了首次流动后的后续迁移流动”问题，事实上是三种口径的通病所在。如五年前常住地口径只能测量得到五年前的流出地与现居住地，亦无法了解其中的流动经历。

表 1 历次普查中三种口径下中国迁移流动人口的比例

统计口径		1988 年	2000 年	2010 年	2020 年
出生地	省内	8.9	8.56	8.97	12.32
	省际	6.0	6.21	8.05	10.28
	小计	14.9	14.77	17.02	22.60
户籍	省内	—	7.74	10.14	17.78
	省际	—	2.82	6.44	8.84
	小计	—	10.56	16.58	26.62
五年前常住地	省内	—	7.53	—	4.39
	省际	—	2.74	4.61	3.93
	小计	—	10.27	—	8.32

数据来源：1988 年数据引自参考文献 [11]；2020 年户籍口径的结果引自普查公报；其他数据均引自普查汇总数据。2000 年、2010 年和 2020 年数据分别根据《中国 2000 年人口普查资料》长表 7-1、《中国 2010 年人口普查资料》长表 7-7 以及《中国人口普查年鉴 2020》长表 7-6 中相关数据计算得到。

二、基于不同统计口径的迁移流动人口比例重估

1. 三种口径之下中国迁移流动人口比例的重新估计

表 1 给出了三种口径之下中国迁移流动人口比例的估计结果。之所以给出这些估计结果，主要目的在于说明三种统计口径存在上述讨论中所提及的各种问题。

首先，三种口径所得的结果之间存在量级的差别。出生地口径和户籍口径下所得到的迁移流动人口比例，远高于五年前常住地口径的结果。以“七普”为例，出生地口径的比例达 22.6%，户籍口径的比例为 26.62%；但五年前常住地口径下，中国迁移流动人口的比例仅为 8.32%，相当于每年的比例仅为 1.664%，这一比例相对而言是极低的。而这种量级差异是由出生地口径和户籍口径的时间累积效应引起的，因此，五年前常住地作为时期指标，与这两种口径并不存在可比性。

其次，最能体现出生地与户籍两种口径的时间累积效应的，是历次普查中这两种口径的结果在逐步递增。不论是历次普查中各口径下合计的比例，还是分省内与省际的比例，在历次普查中都呈现出逐步上升的趋势。如出生地口径下，终身迁移人口的比例从 2000 年的 14.77%，上升到 2010 年的 17.02%，继而上升到 2020 年的 22.6%。户籍口径下，流动人口比例的提高幅度更大，从 2000 年的 10.56% 提高到 2020 年的 26.62%。虽然户籍口径下的流动人口内部存在着替代性（即老一批的流动人口退出流动状态的同时，有更多新的流动人口加入流动人口中），但由于是时间累积而成的存量，因此，其规模或比例都在不断扩大。其实这一点也可以从年龄别流动率的变动中得以发现^[10]。

再次，某种意义上，到目前为止，中国人口仍然具有相对的凝固性^[9,12]。按照出生地口径，仅有 22.6% 的人离开了出生地而成为终身迁移人口；按照户籍口径，也仅有 26% 左右的人处于流动状态。这种状况说明绝大部分人口并未离开过本乡镇街道（当然其中有一部分可能是流出或迁出后的重新返回的返迁人口）。但无论如何，这一比例相对而言仍然是较低的，而且上述两个比例均为时间累积结果。从户籍口径下流量流动人口平均每年的流动率来看，全国平均每年的流动率仅在 1% 左右。同时，如果从分省的角度来看，如浙江、广东等省份的终身迁出人口比例极低（如广东的终身迁出比例仅为 1.6%）^[9]，极低的终身迁移比例表明这些省份的人口相对更具凝固性。因此，虽然人户分离人口规模和流动人口规模都在不断增加，有学者认为迁移时代到来，但从根本上讲，中国人口的迁移率相对仍然是较低

的，中国人口仍然具有相对凝固性。

最后，虽然从出生地和户籍两种口径看，从 2000 年到 2020 年的 20 年间，省际迁移流动人口的比例在不断提高；但五年前常住地口径下，2020 年“七普”的这一比例仅为 3.93%（“七普”中对应的为 5160212 人/131271202 人），低于 2010 年“六普”的 4.61%（省际迁移流动人口为 5522798 人/长表总人口为 119893223 人）。显然，由于五年前常住地为省外的人口规模的下降，以及总人口的增加，导致五年前常住地的总迁移率呈下降趋势。但除了人口规模减小外，其背后还有年龄结构的影响：由于长期处于低生育水平，总人口的年龄结构已经发生根本性变化。虽然年龄别迁移率（流动率）仍在上升（特别是 20 岁左右的年龄别流动率在 2010 年至 2020 年间快速提高）^[10]，但年龄别流动率相对较高的年龄段所对应的人口规模却在持续下降，从而使粗流动率也呈现出下降趋势。即该比例的下降包含着多种因素，至少是由本身迁移流动人口规模的下降、总人口规模的增加，以及年龄结构的变化这三种因素共同引起的。这类似于出生率或死亡率的变化（即受年龄结构的影响）。因此，我们并不能仅仅根据总迁移率的下降而得到迁移强度下降的结论，也不能认为“2015 年作为省际迁移规模下降的拐点”^[4]。这种趋势性的判断结论应该予以更加严谨地深入考察。

2. 流动人口的时间累积效应

为了体现上述流动人口的时间累积效应，本文还将户籍口径下的流动人口区分为流量与存量两个部分。其中，流量是指在五年内离开户籍登记地的流动人口；存量是指其流出时间在五年之前。

如表 2 所示，流量与存量的结构随时间发生变化，表明其具有时间累积效应。2000 年“五普”时省际流动人口中的存量人口仅占全部省际流动人口的

18.68%；到 2020 年“七普”存量人口的比例已达 41.95%，即超过四成的流动人口是在五年前离开户口登记地的。相应地，流量流动人口的占比却从 2000 年的 81.32% 下降到 2020 年的 58.05%，仅有六成不到的流动人口是在普查前五年内流出户口登记地的。这个比例的变化既说明户籍口径下的流动人口是多年累积下来的总量，具有时间累积效应；也说明随着时间的推移，存量流动人口的比例将会越来越大。当然，这个比例的变化可能受到普查问卷的影响，毕竟“六普”和“七普”中关于时间的定义是“离开户口登记地的时间”，而“五普”的相应定义则是来本地居住的时间，因此其中有一部分人可能是五年之前离开户口登记地，但近五年内才来本地居住。这批人在“五普”中被识别为流量人口，而在“六普”和“七普”中则被识别为存量人口，从而可能低估了“五普”中的存量人口。但至少“六普”和“七普”在同一口径之下的比较还是能够说明存量人口比例的提高；存量人口的比

表 2 不同口径下各类省际迁移流动人口的比例估计

流动类型		2000 年	2010 年	2020 年
五年前常住地	有本地户口	13.79	6.00	14.92
	无户口	86.21	94.00	85.08
	小计	100	100	100
户口标准	流量（五年内）	81.32	68.36	58.05
	存量	18.68	31.64	41.95
	小计	100	100	100
流量流动人口占总人口比例		2.4608	4.1513	4.9896

数据来源：根据历次普查原始数据计算。

注：1. 2010 年五年前常住地口径下，有本地户口的比例相对较低，可能的原因是无法在本地人口中区分出省内/县内户口迁移人口。2. 2020 年流量流动人口占总人口的比例，按照微观数据计算的结果为 4.9896%，按照普查汇总资料数据计算得到的结果为 5.13162%，两个结果略有差异。本文中由于流量与存量均按照“七普”微观数据计算，因此，此处以微观数据计算结果为准。

例从2010年“六普”的31.64%增加到“七普”的41.95%。

基于上述比例的变化,有如下两点需要说明:其一,从根本上讲,这涉及是从全国还是从地方的角度去看待流动人口。从全国角度看,不论是在五年前还是五年之内离开户口登记的,都属于流动人口(亦可识别流量与存量);但从地方的角度看,虽然他可能是五年前离开户口登记地(但却是在外地或外省居住),只是在近五年内流入的人口,但在居住时长上却把近五年内本地的新增流量流动人口视为已在本地居住五年以上的存量流动人口。因此,这个统计口径对地方的统计结果可能存在影响。其二,不论其实际来本地居住时长,这些离开户口登记地五年以上的流动人口(当然大部分可能就在一个流入地居住五年以上)是否真的是流动人口?他们的各种权益与未来去向该如何保证与安排?等等。这些社会问题已经被提出了很久,但从根本上并未予以真正解决。从学理上,流量与存量概念的提出也正好可以解决越来越多的“流而不动”的流动人口测量问题。

再者,2020年“七普”中流量省际流动人口占总人口的比例为4.99%,相当于在2015—2020年这五年间平均每年新增省际流量流动人口的比例在1%左右,规模约在1300万左右。如果再考虑到流动人口内部的替代循环,则其规模可能稍微再大一些。某种意义上,这种每年的省际流量流动人口比例(及其相应的年龄别模式)可以作为未来人口预测中的迁移参数,而不是以全部流动人口作为预测参数,这一点比较重要。

同时,这一比例比五年前常住地口径下的省际迁移流动人口比例(3.93%)还略高一个百分点;按理五年前常住地口径既包括流量流动人口,也包括流动人口未能包括的户口迁移人口,因此,五年前常住地口径之下的规模按理应该大于户籍口径下的省际流量流动人口的规模。这还需要再进一步讨论,但这一结果至少从某种意义上说明五年前常住地口径可能存在较大的低估或者偏差,因此,对五年前常住地口径的使用仍然需要十分小心,而不是扩大使用。

除此以外,表2还列出了五年前常住地口径下有户口的迁移人口和无户口的流动人口各自所占比例,可以看到,除2010年“六普”中有户口的迁移人口比例相对较低以外,2000年“五普”和2020年“七普”中分别有13.79%和14.92%的人口属于户籍迁移人口。或者大致而言,有将近15%的迁移人口。粗略地估计,如果按每年1300万流量流动人口计算,则户籍迁移人口规模大致在230万左右($1300 \text{ 万} \times 0.15 / 0.85$)。因此,在2015—2020年的五年间,我国每年平均迁移流动人口规模在1530($1300 + 230 = 1530$)万人左右。当然,这只是大致的匡算,但至少可以得到一个大概的“时期”流量性质的比例与规模。

3. 五年前常住地口径下样本混合属性的再检验

周皓曾利用“五普”和“六普”数据讨论了五年前常住地口径下样本的混合属性^[6]。本文在此基础上,计算了“七普”的五年前常住地口径与户籍口径之下的交互分析结果。

由于2000年普查中调查的迁移时间是“最近一次来本乡镇街道居住的时间”,因此,根据这个时间和“从何地来本乡镇街道居住”,以及户口登记状况,可以识别出流动人口和户口迁移人口,特别是将户口迁移人口从本地人口中区分出来,从而形成县内、省内县际和省际三种空间边界的迁移与流动人口(见表3)。

五年前常住地为省内(跨县)的89200人中,按照户籍登记地口径看,仅有19907人仍

表3 2000年普查中五年前常住地口径与户籍口径的交互分析结果

人,%

五年前常住地		户口标准							合计
		本地人口	县内迁移	县内流动	省内县际迁移	省内县际流动	省际迁移	省际流动	
数量	省内	835	24348	28842	13063	19907	696	1509	89200
	省际	199	124	500	86	248	3987	26735	31879
	本地	1055810	464	946	150	785	81	796	1059032
	合计	1056844	24936	30288	13299	20940	4764	29040	1180111
占比	省内	0.94	27.30	32.33	14.64	22.32	0.78	1.69	100
	省际	0.62	0.39	1.57	0.27	0.78	12.51	83.86	100
	本地	99.70	0.04	0.09	0.01	0.07	0.01	0.08	100

为省内流动人口，仅占 22.32%；有 27.30%（24348 人）为县内迁移人口，32.33%（28842 人）为县内流动人口，两类合计占 59.63%。即五年前常住地为省内县际的人口，有近 60% 的人返回到其户口所在的县内，仅有不到 1% 的人是回到原乡镇街道。这说明，五年前常住地在本省其他县的人口并不是返回原户籍登记地，而是户籍登记地所在的县域其他地方，很可能是县城或其他乡镇街道，同时也验证了以往有关返迁人口的一些研究结果。

而五年前常住地为省际的 31879 人中，有 26735 人（占 83.86%）在户籍口径下仍然是省际流动人口，仅有 2.58%（ $199 + 124 + 500 = 823$ 人）返回到户籍登记地；也仅有 1.05%（ $86 + 248 = 334$ 人）返回户籍所在省份，其中 0.27% 的人是省内迁移人口（即返回原省，并将户籍迁入本省其他地方）；还有近 12.51% 的人则是属于省际迁移人口。这些结果说明，相比于五年前常住地为省内的人口而言，省际迁移流动人口持续在省外居住的可能性相对较大，而返回原籍地的比例相对较低。但这批人却是返迁人口，在考虑迁移流动机制时是不应该被包括在内的。

上述分析表明五年前常住地口径下的样本中，既包括了迁移人口，也包括了流动人口，检验了其样本的混合属性。同时该样本既包含了迁出或流出的人口，也包括了返迁的人口。某种意义上，返迁人口的估计结果有助于从宏观角度考察中国人口的返迁状况，并回答在实际的定性访谈中得到的有关返迁后居住地安排的问题。这一结果极富社会含义与政策意涵。

由于 2010 年“六普”和 2020 年“七普”中关于迁移流动时间的问题被修改为离开户口登记地的时间，因此，无法再有效地将迁移人口从本地人口中予以区分与识别。本文只能将“六普”与“七普”的交互结果按照空间范围予以分类（见表 4）。其中 2010 年“六普”中五年前常住地为省内的人口由于是将省内与本地户籍人口合并在一起，无法区分，从而导致其在户籍口径下的省内流动人口的比例相对较低。因此，对于 2010 年的五年前常住地口径为省内的数据只为了表格的完整性而呈现，仅供参考。

表 4 的结果同样表明五年前常住地口径下样本的混合属性，同时也体现了省内流动与省际流动人口的区别。五年前常住地为省内的人口，以 2020 年“七普”为例，仅有 65% 的人仍然是省内流动人口；28.35% 的人则被视为本地人口，正如表 3 所示，它既包括了返回原户口登记地的返迁人口，也包括了户口迁移人口。同时，还应该注意，五年前常住地为省内的、户口标准之下被识别为省际流动人口的比例，从 2000 年的 1.69% 上升到现在的 6.71%，这批流动人口事实上应该是在五年前就迁入该省，但在五年内在该省内部进行了流

动。这一比例的提高，既是时间累积效应的体现，也是流动人口在流入地不断扩散的表现。当然，五年前常住地为省内的迁移流动人口，仍然以户籍所在省的省内流动为主。

表 4 五年前常住地口径与户籍口径的交互分析结果 %

五年前常住地		户籍口径			合计
		本地人口	省内流动	省际流动	
省内	2000 年“五普”	43.66	54.65	1.69	100.00
	2010 年“六普”	87.42	10.18	2.40	100.00
	2020 年“七普”	28.35	64.94	6.71	100.00
省际	2000 年“五普”	13.79	2.35	83.86	100.00
	2010 年“六普”	22.70	1.25	76.05	100.00
	2020 年“七普”	17.27	4.52	78.21	100.00

数据来源：2000 年“五普”和 2010 年“六普”数据引自周皓的文章^[6]；2020 年“七普”的结果根据微观原始数据计算得到。

五年前常住地为省际的迁移流动人口与户籍口径下的省际流动人口之间表现出较强的一致性，虽然均为省际流动人口的比例在 2020 年时相对低了五个百分点，但比例仍然达到 78.21%。这也正是三种统计口径的省际流动人口具有较强相关性的原因所在^[13]。除少量省际迁移流动人口转换为省内跨县的流动以外，还有一部分人口转为本地人口，同样这部分本地人口包括了返迁人口和户籍迁移人口，但却无法判断与识别。

同时也可以看到目前普查中五年前常住地的选项设置忽略了县内迁移流动人口。

三、重新认识中国人口迁移流动

1. 本文结论

本文的主要目的就是厘清统计口径，重新估计三种统计口径之下中国迁移流动人口的总体比例，以正确认识中国人口迁移流动的整体强度。通过上述分析与讨论可以得到以下几点主要结论：其一，出生地口径和户籍口径由于时间累积效应而无法表示时期性的迁移流动人口规模。其二，户籍口径下流量与存量流动人口的结构随时间（在三次普查间的）变化，特别是存量人口占比的不断扩大，说明其具有时间累积效应。2020 年“七普”中流量省际流动人口占总人口的比例为 4.99%，相当于 2015—2020 年这五年间平均每年新增省际流量流动人口的比例在 1% 左右；这一比例比五年前常住地口径下的省际迁移流动人口比例（3.93%）还略高一个百分点。其三，几次普查数据的分析再次揭示了五年前常住地口径的“混合样本属性”（既包含迁移、流动，还包含着返迁的流动人口和返迁人口等），以及其忽略了在我国人口迁移流动中占据重要地位的“县内跨乡镇”的迁移流动。后者说明五年前常住地口径为何不能作为对全国迁移流动人口规模估计的标准。

上述分析表明，三种口径对中国迁移流动人口都存在低估或错误估计的可能。如出生地口径由于忽略了县内迁移流动而低估了总量；户籍口径存在着时间累积效应而不具有时期指标的特征，且忽略了户口迁移人口而同样低估了中国的迁移流动人口总量，虽然它调查得到的省际流动人口比例比五年前常住地口径的省际迁移流动人口比例更高。五年前常住地口径虽然包括了户口迁移人口，却忽略了县内迁移流动，从而形成低估问题；并且五年前常住地口径下的省际迁移流动人口比例（3.93%）还低于户籍口径下的流量省际流动人口的比例（4.99%），这说明五年前常住地口径存在更大的低估可能性。通过具体比较这三种口径后可

以发现，到目前为止，仍然没有一种口径能够合适准确地描述中国迁移流动人口的规模与比例，这是本文的重要结论之一。

2. 有关讨论

本文讨论统计口径及相应强度与规模的目的是提请注意由统计口径带来的分析结果以及由此导致的结构性偏差，一方面希望能够注意到统计口径问题以及由此带来的对中国人口迁移流动的认识问题，特别是不同统计口径对应的不同总体强度问题；另一方面也希望能够更准确地估计时期性的中国人口迁移流动状况。与此相关的有几个问题值得思考与讨论。

(1) 解决估计准确性的三个标准。既然到目前为止，三种口径事实上都存在低估中国人口迁移流动的可能，那么如何解决，或者如何准确调查与估计中国迁移流动人口的规模或比例已成为一个急切的问题。不论是丁金宏提出的通过间接估计方法来估计中国迁移流动人口的规模^[5]，还是周皓提出的修改普查问卷设计以更准确地调查得到中国迁移流动人口的规模或比例^[6,10]，都是朝着这个方向在不断努力。但不论哪种方案，都需要在明确调查对象、完备登记信息的基础上^[6]强调空间范围、时间性以及中国的社会制度背景这三个标准。首先，既然户籍制度作为背景性的中国社会经济制度的基础，笔者认为在考察与估计中国迁移流动人口规模时必须与之相结合，并将迁移流动人口按照户籍登记地分为迁移人口与流动人口两类。这既是中国特色的体现，也是现实与学术研究充分结合的表现，也能符合国际比较要求。其次，时间性即时期指标，意指迁移流动人口的估计应该是尽量排除了时间累积效应的时期指标，哪怕是五年期间的迁移流动人口总规模也是可以接受的。再次，调查或估计的空间范围应该是全覆盖的，即包括县内跨乡镇街道的（除市内人户分离人口以外）、省内县际和省际这三类迁移流动人口。坚持这三个原则是准确估计时期性中国人口迁移流动情况的基础。从这个意义上看，目前的五年前常住地口径并不适合作为中国迁移流动人口总规模估计的统计口径，尽管它具有国际可比性。因此，本文建议普查设计中，五年前常住地口径加入县内部分，以便从空间角度满足对总量估计的要求；坚持户籍口径，加入“最近一次迁移”的时间与地点以满足登记信息的完备性^[6]。

(2) 中国人口迁移流动的发展与转变。近年来人口转变已成为人口学研究中重要的研究内容之一，特别是人口生育转变问题。但对人口迁移流动的转变问题，除刘金菊、陈卫利用大型社会调查数据进行分析与讨论以外^[14]，尚未见有更好的研究。严善平在综合利用多次人口普查与小普查数据分析后认为“2015年为跨省迁移的拐点，此前的增加趋势随后出现了逆转”^[4]，这一论断是否为真仍然需要进一步的讨论与深入考察。一方面，2015年小普查数据可能会由于抽样问题而导致低估。最好的例证是2015年小普查的年龄别流动率并未处于2010年“六普”和2020年“七普”两条年龄别流动率曲线之间，相反与2005年小普查的年龄别流动率相接近^[10]。另一方面，上述分析表明户籍口径下的流量省际流动人口规模在近五年中仍然持续增长，尚未到达拐点。更何况，人口转变研究中通常以“率”的转变为基础，而非总量的概念。因此，本文认为省际人口流动是否出现了逆转尚需要进一步严谨的分析与讨论。

更重要的是，随着中国人口结构（包括分城乡）的不断变化，特别是多年以来中国人口生育水平处于相对较低状态，总人口呈现出少子化、老龄化的发展趋势，这种结构趋势也

会影响到流动人口年龄结构和总规模、总比例的变化。简单地看,迁移流动最为活跃的20岁左右人口的年龄别流动率尽管在持续提高,但由于总人口中20岁左右人口的规模急剧下降,从而导致相应的迁移流动人口规模下降。假设以2020年“七普”时的年龄别流动率为标准流动率,如果按照2000年“五普”时的总人口年龄结构,那么2020年“七普”时的“预期”流动人口总规模达4.183亿人,比2020年实际的全国流动人口规模3.758亿人要多出4246万人;如果按照2010年“六普”时的总人口年龄结构,那么,2020年“七普”时的“预期”流动人口总规模应为4.1265亿人,比实际人口规模多3683万人(注:根据2020年“七普”年龄别流动率和总人口数、2000年“五普”与2010年“六普”总人口年龄结构估算得到)。由此可以说明,即便自2015年以来中国迁移流动人口总规模开始下降,但这种下降并不一定是真正源自其内在的迁移流动率的下降,而是在于总人口年龄结构的改变。

正因如此,中国的人口迁移流动是否处于转变过程,是否与人口转变理论相对应,还是中国人口迁移流动在规模增加的同时,内部结构也在经历变化(如朱宇等提出的以城城流动为主^[15])的过程,仍需更进一步地深入分析与讨论,并利用实际数据来予以检验与支持。

(3) 其他相关问题。严善平的文章^[4]中还有几处值得商榷或讨论的地方:①中国的人口迁移流动方向是空间集聚还是多极分散问题。中国人口迁移流动的方向并非是多极分散趋势,而是“多中心”的集中趋势;而且从出生地角度考察的终身迁移,其方向仍然是“三大一副”;空间分布趋势表现为迁出呈普遍提高和地域的漫延式扩散化、迁入呈集聚中扩散,即空间集聚过程并不相同,同时亦存在空间邻近性^[9]。②总和迁移率的计算问题。总和生育率是按照年龄别生育率的总和来计算的,而年龄别迁移率事实上并不是严格意义上的迁移率,特别是对迁入率而言,至多只能算是比例而已,因为年龄别迁入率的分母并不是通常意义下的平均生存人年数。事实上,不论是哪种口径(包括五年前常住地口径),在利用普查数据时都很难计算严格意义上的年龄别流动率(其复杂性不在于五岁年龄组的平均,而在于年龄别死亡人口及平均生存人年数的估算)。③罗杰斯模型对中国现实的适用性问题。尽管中国数据中劳动年龄段的拟合似乎接近于罗杰斯模型,但其所对应的研究对象应该是流量流动人口(或者至少是时期性的五年前常住地口径下的迁移流动人口),而并不能使用具有时间累积效应的流动人口来计算年龄别流动率;同时,中国流量流动人口的年龄别流动率模式完全有别于西方国家的年龄别迁移率模式^[10],特别是在少儿段和老年段。因此,我们仍然认为罗杰斯模型并不适用于中国的情境,至多也只是一种分析思路。

基于上述讨论,本文认为中国人口迁移流动的研究仍有漫长的道路要走。从对人口迁移流动的认识来看,需要清楚地认识各种统计口径及其可能存在的问题,如此才有可能正确地认识与理解中国的人口迁移流动;从研究目标来看,需要针对具体研究问题选择合适的研究样本,如此才有可能真正地进行因果关系推断,进而真正理解中国人口迁移流动的成因、过程与结果。

参考文献:

[1] 国家统计局. 第七次全国人口普查公报(第七号)[EB/OL]. (2021-05-11)[2023-12-01]. <http://www.stats.gov.cn/sj/>

- tjgb/rkpcgb/qgrkpcgb/202302/t20230206_1902007.html.
- [2] 国务院第七次全国人口普查领导小组办公室. 中国人口普查年鉴 2020 [M]. 北京：中国统计出版社，2022：1709-1716.
- [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 (12)：2991-3005.
- [4] 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 (5)：112-125.
- [5] 丁金宏. 人口普查地理信息与人口迁移多维流场分析 [C]. 人口高质量发展与中国式现代化建设全国学术研讨会，2023.
- [6] 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.
- [7] SIEGEL J S, SWANSON D A. The methods and materials of demography (second edition) [M]. San Diego, California: Elsevier Academic Press, 2004: 30-31.
- [8] 杨云彦. 改革开放以来中国人口“非正式迁移”的状况 [J]. 中国社会科学，1996 (6)：59-73.
- [9] 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 (6)：107-125.
- [10] 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.
- [11] 段成荣. 影响我国省际人口迁移的个人特征分析 [J]. 人口研究，2000 (4)：14-22.
- [12] 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 [J]. 中国人口科学，2021 (3)：28-41.
- [13] 王桂新. 中国省际人口迁移变化特征——基于第七次全国人口普查数据的分析 [J]. 中国人口科学，2022 (3)：2-16.
- [14] 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.
- [15] 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 [J]. 人口研究，2016 (5)：50-60.

Comparison of Statistical Caliber for Migrant and Floating Population in China

ZHOU Hao

(Center for Sociological Research and Development Studies of China/
Department of Sociology, Peking University, Beijing 100871, China)

Abstract: The statistical caliber is the foundation for understanding China's recent population migration and mobility. Different statistical calibers correspond to various migrant and floating populations and sample structures. Based on the internal logic of "statistical caliber-total result", this paper firstly compares and discusses the three calibers from the perspective of census, in order to correctly understand the meaning and characteristics of each statistical calibers. Then, using the data from the fifth, sixth, and seventh national population censuses, we will re-estimate the proportion of the migrant and floating population under different statistical calibers, so as to truly understand the overall intensity of China's population migration and mobility as well as deepen our understanding of China's migrant and floating population. The article believes that as non-period indicators, the household registration and birthplace indexes have time cumulative effect, and the permanent residence indexes five years ago have mixed sample attributes and the underestimation problems of not including intra-county migrant and

floating population. At present, there is still no suitable and accurate description of the scale and proportion of China's migrant and floating population. It is urgent to adjust the statistical caliber of the migrant and floating population, among which spatial scope, temporality, and China's social system background are the three standards for adjusting the statistical calibers. At the same time, the article suggests strengthening research on the trend of population migration and mobility in China under the condition of negative population growth.

Keywords: population migration; floating population; statistical caliber; census data

[责任编辑 刘爱华]