**迟到的婚姻：流动经历、接入城市与流动人口的初婚年龄[[1]](#footnote-1)\***

潘丽群 李静 张少华

**摘要：**“剩男”和“剩女”向城市新增流动人口蔓延的新现象，引发我们重新思考流动对初婚年龄的影响。为此，本文基于2016年全国流动人口动态监测调查数据，系统考察流动经历和流入地属性对流动人口初婚年龄的影响。研究发现：（1）流动经历会显著地推迟流动人口的初婚年龄。婚前具有流动经历个体比婚前没有流动的个体晚婚1.893年，而且以流动时间和流动范围衡量的流动性越大，流动经历的推迟效应越强。（2）接入城市的规模显著影响了流动人口初婚年龄，接入城市的规模每增加10%，初婚年龄提高0.0912%。（3）将流动对婚姻决策的影响过程显性化，分析接入城市规模推迟初婚年龄的影响路径，接入城市会通过改变流动人口的经济地位、婚姻观念和婚配概率这三类婚姻的重要决策因素来推迟其初婚年龄。上述研究结论在运用倾向得分匹配法、工具变量法、更换数据库和变量等检验后依然稳健。

**关键词：**流动人口 初婚年龄 流动经历 城市规模

中图分类号：F063.4 F061.5 F064.1 JEL：J12 J61 R23

**Delayed Marriage: Migration experience, host city and the Age at First Marriage of Migrants**

PAN Liqun1 LI Jing2 3 ZHANG Shaohua1

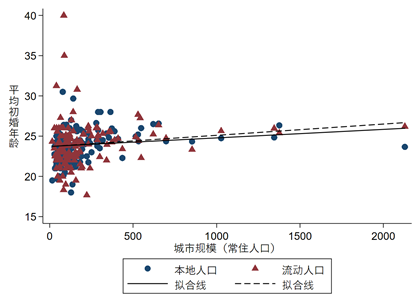
1. Guangzhou University, Guangzhou, China; 2. Xiamen University, Xiamen, China; 3. Xinjiang University of Finance & Economics, Urumqi, China)

**Abstract：**The new phenomenon of "leftover men" and "leftover women" spreading to the new floating population in the city leads us to rethink the impact of migration experience, especially the access to the city on their first marriage age. Based on the data of 2016 China Migrants Dynamic Survey, this paper systematically examines the impact of pre-marital migration experience and the characteristics of host cities on their first marriage age. The empirical findings show that: (1) The pre-marital migration experience significantly delays the age at the first marriage of migrants, with migration experience were married for 1.893 years later than those without migration experience, and the migration duration and migration range further enhance the delay effect of the migration experience. (2) The scale of host cities has a significant impact on the age at their first marriage. For 10% increase in the host city size, the age at first marriage increases by 0.0912%. (3) The mechanisms of host cities delay the migrants’ marriage by changing their economic status, the concept of marriage and the probability of marriage matching. The above research conclusions are still robust after replacing micro-database, using the propensity score matching method and the instrumental variable method to control individual selection bias and endogenous.

**Keywords:** Migrants；First marriage age；Migration experience；City size

一、引言

伴随着快速的工业化和城市化进程，人们的婚育观念与行为激烈转变，被概括为“人口的第二次转变”（於嘉、谢宇，2019），最显著的表现之一就是结婚越来越晚，出现了大量的“剩男”和“剩女”。2010年全国人口普查数据显示，从1990年到2010年的二十年间，总体上女性初婚年龄平均增加了1.87岁，男性则增加了2.29岁。更为重要的是，作为城市新居民的流动人口，相比本地人结婚更晚，无论是农村户籍流动人口还是城市户籍流动人口（图1）。而且，2015年中国家庭金融调查微观数据表明，城市规模越大，晚婚趋势越明显（图2）。不同于西方社会的第二次人口转变过程，我国在这一进程中，婚姻仍然是生育的必要条件（於嘉等，2020），初婚年龄的推迟意味着生育年龄以及此后多次生育的育龄周期的缩短，将会进一步加剧我国面临的人口老龄化问题。流动人口作为推进城市化的主要力量，他们的婚育行为对人口结构和城市发展至关重要。那么从实证层面我们需要了解，流动是如何影响流动人口的婚姻决策？流动影响婚姻决策是一个过程，除了流动经历本身带来的推迟效应，还需要探讨因为流入到新的城市，随新的环境而改变的经济地位，与之相伴甚至“入乡随俗”的婚姻观念，以及新城市婚姻市场的匹配概率又是如何对流动人口的初婚年龄产生影响？考察和探究流动人口的初婚年龄，以及他们初婚年龄受到城市化进程的影响程度以及影响的路径，将有助于我们认清城市内部婚配模式、引导流动人口婚育决策、实现城乡发展高度融合以积极应对人口老龄化提供解决思路。

**图1：本地人和流动人口初婚的差异 图2：初婚年龄与城市规模的关系图**

注：基础数据来源于2015中国家庭金融调查，经作者整理计算而得。

现有文献针对流动人口的初婚年龄问题做了有意义的探索，主要集中在初婚前的流动经历是否影响了初婚年龄（郑真真，2002；靳小怡等，2005；刘厚莲，2014；许琪，2015；曾迪洋，2014；刘利鸽等，2019），相对较少分析流入地属性对初婚年龄的影响及其机制的实证研究。虽然现有研究均认为相比没有流动经历的个体，有流动经历个体的初婚会更晚，但是，流动影响初婚年龄不仅仅是流动本身，还有流动之后，流入地城市所代表的新环境，也可能通过改变婚姻决策的影响因素而作用于初婚年龄。具体而言，第一，流动过程带来收入增长和经济地位的改变（Becker，1973；Bergstrom& Schoeni，1996； Oppenheimer，1997； Loughran，2002；Field & Ambrus，2008 ；Yeung & Mu，2019），这种变化显然会影响流动人口对未来配偶的选择，进而影响其婚期。第二，流入地文化观念的影响（郑真真，2002；Ikamari，2005；李静等，2015），接入地的生活观念、婚姻观念等可能显著异于流动人口流出地，这些新的甚至截然相反的观念会潜移默化地影响流动人口的婚配过程和初婚选择。第三，不同流入地婚姻市场的匹配概率不同（McGoldrick & Robst，1996；Drewianka，2003），流动人口进入流入地生活和工作，一方面，流入地可能提供更多的选择和匹配机会，另一方面，陌生“花花世界”的锦绣繁华也可能导致其“无所适从”，可见，流入地城市对流动人口的匹配概率存在不确定的影响。总之，以流入地城市属性为视角，可以从理论层面更好地揭示流动推迟初婚年龄的过程。在接入城市的众多属性当中，城市规模这种显著影响个体经济地位、婚姻观念以及婚配市场规模的城市属性，无疑是一个非常好的观察角度。

为此，本文基于2016年全国流动人口动态监测调查数据，系统考察流动经历和流入地属性对流动人口初婚年龄的影响。本文的边际贡献主要体现在：（1）研究内容上，以城市规模为视角，实证分析了流入地属性对流动人口初婚年龄的影响，完善了流动影响初婚年龄研究，更全面地揭示了中国城市化进程中，流动人口所承受的婚姻压力。（2）突破以往文献对流动推迟初婚年龄机制分析停留在现象和经验证据的描述，本文从经济地位、婚姻观念和婚配市场的匹配概率实证研究了流入地城市规模强化流动人口初婚年龄推迟效应的三种可能路径，从理论和实证两方面将流动对婚姻决策的影响过程显性化。（3）研究方法上，相较以往的研究更稳健地评估流动对流动人口初婚年龄的影响：运用倾向得分匹配方法弥补研究中被忽视的婚前是否流动的样本自我选择偏误问题；采用城市历史人口和地理平坦面积作为城市规模的工具变量解决内生性问题。

本文其余部分结构安排如下：第二部分是文献综述与研究假说的提出；第三部分是使用数据的说明、变量的处理以及计量分析模型的构建；第四部分是详细的实证分析过程以及解释；第五部分是流入地城市规模影响流动人口初婚年龄的作用机制；第六部分是异质性分析；最后是结论与建议。

二、文献综述与研究假说

随着经济全球化的到来，国家间的迁移行为愈加平常。同时，二十一世纪发展中国家进入快速的城市化进程，农村人口向城市流动以及城市间的流动也成为一个普遍现象。流动对初婚年龄的影响成为研究热点，既有从国家间移民视角的研究（Yeung & Mu，2019），也有从一国内部的乡城、城城人口流动视角的研究（Jampaklay，2006；Mu & Yeung，2019）

中国的流动人口是快速城市化进程的产物，关于中国人口流动对初婚年龄的影响虽然研究重点不同，使用样本也显著差异，但得出的结论较为一致，均认为相比没有流动经历的个体，有流动经历个体的初婚会更晚，同时这种推迟效应呈现明显的异质性。实证方面有基于中国妇女社会地位调查数据，探究农村男女的初婚年龄与流动性的关系，研究表明初婚前有过外出务工经历的农村人口结婚年龄要大于婚前未曾外出的人口，而且这种效应在男性农民工身上表现得更加明显（许琪，2015）；也有基于全国性流动人口调查数据展开研究，发现流动推迟了新生代流动人口的初婚年龄，并且流动人口初婚年龄具有显著差异，表现为男高女低、农业户口低于非农业户口、少数民族高于汉族，且随着自身和配偶教育程度提高而推迟（刘厚莲，2014）。除了一般性的实证分析外，也有文献试图从理论视角来解释流动对初婚年龄的推迟效应，如从生命历程的研究视角出发，指出人生当中的重要事件——迁移是影响人们初婚时间的不可忽视的因素（曾迪洋，2014）；还有基于奥本海默的婚姻搜索理论框架，认为流动经历会提高农民工在婚配市场的匹配标准以及他们的婚姻市场范围，从而造成婚姻延迟，影响渠道分别是年龄差、教育时间差和通婚范围（刘利鸽等，2019）。

进一步地，流动的推迟效应在流动范围和流动时长上是不同的。流动范围方面，基于四川和安徽四个县的调查问卷数据，通过事件史的研究方法得出婚前有流动经历的妇女比未流动的妇女更晚结婚，并且这种流动带来的婚姻推迟效应在流动人口跨省流动时更明显（郑真真，2002）；基于2015年流动人口监测数据的研究也发现流动范围（跨省、省内跨市和市内跨县）对城市流动人口的初婚和初育年龄有显著的差异性影响（杨雪、徐嘉树，2018）。流动时长方面，基于“浦东新区外来人口调查”数据，定量研究社会网络与流动因素对妇女初婚年龄的影响，发现流动性推迟婚姻的同时，流动滞留时间更长的妇女初婚年龄更高（靳小怡等，2005）。与时长类似，在流动频次上，婚前有多次流动经历的人，其初婚年龄显著推迟（曾迪洋，2014）。

综上分析，本文提出假说一：初婚前的流动经历将推迟流动人口的初婚年龄，且流动时长越长、流动范围越广对婚姻的推迟效应越强。

初婚年龄是婚姻决策的结果。婚配过程中的选择偏好主要包括两种解释：地位竞争论和文化匹配论（Kalmijn，1994），地位竞争论侧重强调个人的社会经济地位，文化匹配论强调夫妻间共享的价值、规范和文化资本。文化匹配论较多用于社会学的研究，关于初婚年龄的研究从婚姻观念的角度进行了大量的分析，我国近几十年在婚姻形成过程中，个人因素越来越重要（齐亚强、牛建林，2012），对婚姻陪伴功能的重视会推迟初婚年龄（於嘉等，2020），尤其是婚前性行为以及婚前同居的观念深刻地影响初婚年龄。

在经济学的研究范式下，关于初婚年龄的研究主要从构建家庭效用模型展开（Becker，1973），强调从地位竞争论的视角来解释婚姻决策过程；后续又发展出来婚姻市场的搜索匹配模型（Oppenheimer，1997），着重解释婚姻市场范围及其匹配概率也成为影响婚姻决策的重要因素。具体而言，早期Becker（1973）将婚姻视作夫妻双方提升经济效用的协商结果，基于性别分工（男主外女主内）的设定构建了效用模型，研究发现男性收入的增加会提高初婚年龄，因为积累足够的经济资源需要时间。而Bergstrom& Schoeni（1996）却有截然相反的观点，认为高收入的男性更容易被婚姻市场预期到，从而越早进入婚姻。所以，Oppenheimer（1988）指出男性获取稳定收入的确定性程度决定了男性进入婚姻的时间，并在后续研究中具体提出职业成熟度、职业转换的能力以及未来长期职业的地位来刻画男性获取稳定收入的确定性程度。随着二十世纪末女性劳动参与率的提升，女性经济独立性的提升被认为导致了结婚年龄的普遍推迟（Field & Ambrus，2008）。Becker（1991）强调女性经济独立性的提升降低了其在婚姻中的收益，从而降低了进入婚姻的意愿。但是，Oppenheimer（1997）认为女性经济独立性的提高导致了女性对婚姻贡献的不确定性提高，从而延迟了其进入婚姻的时间，但经济的提高会增强其婚姻收益。基于女性经济地位，研究者进一步探讨了推迟是由于女性婚前搜寻行为回报率的提高使其婚前搜寻的时间延长，从而造成各年龄组女性的结婚率都会降低（Loughran，2002）。与地位提高类似的是关于教育程度的提升对女性婚姻的影响，在教育时间上，接受高等教育推迟了进入婚姻的时间，推迟效应不仅来自接受教育时间的拉长，而且来自于从教育系统进入劳动市场的身份转换和搜寻工作的时间（Cantalini，2017）。

总之，从理论发展脉络来看，经济地位、婚姻观念和婚配市场的匹配概率是婚姻决策最重要的三类影响因素。而且，这三大类影响因素在不同规模的城市是异质的，随着流动人口流入不同属性（不同规模）的接入城市，接入城市可能会通过改变、同化流动人口婚姻决策的影响因素而对其婚配过程产生重大影响。

（1）经济地位。男性收入（Becker，1973；Bergstrom& Schoeni，1996）和女性经济地位的提高（Oppenheimer，1997；Loughran，2002；Field & Ambrus，2008）都将影响各自进入婚姻的时间。在城市劳动力市场中，异质性劳动力获得的城市规模工资溢价或工资收入具有差异性（潘丽群等，2020）；对流动人口而言，女性获得城市规模工资溢价比男性高（潘丽群、张少华，2021）。所以，通过流动，女性因为获得更多的就业机会和收入，从而相对降低了他们的婚姻预期收益，进一步影响其进入婚姻的决策。由于女性向上婚（李静等，2015）的广泛性，且流入地带给男性收入的提升程度可能弱于女性的事实，导致婚配过程中男性和女性经济地位的改变，进而影响其进入婚配状态的时间。

（2）婚姻观念。工作挣钱和职业发展是流动人口流动的目的，但是工作和生活地的变化潜移默化中改变和重塑他们的婚姻观念（许琪，2015）。工业化和现代化进程影响了我国的婚配观念和婚配模式，婚姻匹配的变化趋势印证了工业化理论（齐亚强、牛建林，2012），越大规模的流入地城市，其工业化和现代化程度越高，婚配观念和婚配模式也相异于流出地，更倾向于后现代化的婚姻观念。如果按照婚配观念的变迁路径，可以将传统文化观念与后现代化观念视为婚姻观念变迁的两个端点，传统文化观念更看重婚姻的经济功能和生育功能（於嘉、谢宇，2019），反对婚前性行为，提倡早婚早育；后现代化则强调婚姻的陪伴功能和感情功能，婚姻模式多样化（Mu & Yeung，2019），对婚前性行为和同居行为普遍接受，是否需要进入婚姻不成为人生的必修课。那么按照这样的谱系去刻画不同的流入地，流动人口会面对不同程度的婚姻观念的冲击，大体上来看，城市规模越大，越处于开放前沿，婚姻观念越靠近后现代化婚姻观的端点。

婚姻观念指导行动的结果呈现为婚配模式，传统婚姻观念会导致男高女低的婚配模式；而后现代化婚姻观念往往呈现为随机匹配模式。Qian（2017）利用美国1980人口普查数据和2008-2012年社区调查数据，发现从1980到2008-2012年，虽然女性与比自己收入更高的男性结婚的趋势依旧持续，但教育婚姻匹配的趋势完成了从女性向上婚到男性向上婚的转变。随着教育的普及，我国“男高女低”的现象也在慢慢改变。流动人口的婚姻观念在流入地城市被深刻地影响（郑真真，2002），同时，流出地传统观念下女性向上和男性向下的匹配模式导致的错配效应被大城市放大，导致优质女性和非优质男性都很难进入婚姻（李静等，2015）。

（3）婚配市场的匹配概率。对于流动人口而言，原本依赖地缘、亲缘的通婚圈被冲击和拓展，面临着新的婚姻市场（刘利鸽等，2019）。虽然婚配市场范围的扩大，会提高匹配概率，但是，大部分男性流动人口相对于本地人口，婚姻的竞争力实际上是下降的，会降低匹配概率；而女性流动人口则相反。因此，在传统的婚配模式下，总体上降低了流动人口匹配概率。McGoldrick & Robst（1996）指出婚配市场的匹配成功率取决于婚配市场上的结构，城市规模的扩大集聚了更多婚配的差异个体，也集聚了对差异婚配对象的需求个体，从而提升匹配的可能性。Drewianka（2003）也认为越大规模的城市集聚了相对更多的单身个体，这些单身个体遇见某一质量的潜在配偶的可能性更高了，从而加快了婚姻个体的匹配速度；但也指出大城市更多吸引了高教育和高技能者，以及那些渴望与高教育或高技能者结婚的个体。除了自致性个人特征如教育背景外，生活风格的匹配和文化相似性是影响流动者婚姻模式的重要因素（李家兴，2020），流动人口作为流入地的新人，往往在文化和生活方式上本地居民有一定的相异性。同时，在城市化进程中流动人口的状态是“半城市化状态”，这种状态呈现出相对封闭的亲缘或地缘性社会网络，降低了他们与本地居民匹配的概率。一方面是城市规模越大，单身个体更多，匹配可能性更高；另一方面流动人口的整体教育水平相对较低，在高等学历人群比例更高的大城市寻找到潜在配偶的可能性降低了。因此，城市规模通过婚姻市场的匹配概率影响流动人口结婚的可能性。

综上，本文提出假说二：流入地属性——城市规模会强化推迟效应。进一步地，城市规模强化推迟效应的路径可能是：经济地位、婚姻观念和婚姻市场的匹配概率。

三、数据、变量和模型设定

（一）数据说明

本文采用2016年中国流动人口动态监测调查数据（CDMS）这一微观数据进行实证研究。该数据是由国家卫生计生委在流动人口的流入地展开的全国性调查，个体为在调查前一个月来本地居住、非本区（县、市）户口且2016年15周岁及以上的流入人口，调查区域包括以各省（区、市）和新疆生产建设兵团共32个省级单位。宏观数据指各地级市的宏观变量数据，来自《中国城市统计年鉴2017》，包括市辖区人口数量、产业结构、实际利用外资情况、人均道路面积和每万人拥有汽车辆等。

基于实证设计，本文做了如下样本筛选。首先，针对流动经历推迟效应的实证研究部分，筛选出所有已婚且有报告初婚时间和初婚前有无流动经历的个体。其次，针对流入地城市规模强化推迟效应的实证研究部分，由于问卷只提供当下流入地的具体城市信息，不能提供此前流动时流入地的信息，因此，将样本限定在仅流动一次的人群，以确保婚前流动城市与当前流入城市一致。同时，由于县规模较小，去掉了县仅保留了地级市，共获得来自106个地级市的7905个样本。最后，需要特别说明的是，流动人口个体的流动时间和初婚时间各不相同，而不同时间段的城市规模不可直接比较。考虑到城市之间的相对规模是比较稳定的，即使处于快速的城市化进程中，其当前城市规模在很大程度上是依赖于以往城市规模，其在全国城市中排名的次序是变化不大的，所以，本文选取2016年的城市截面数据进行比较[[2]](#footnote-2)。

（二）变量说明

因变量是初婚年龄，即首次结婚的年龄。核心解释变量是流动特征和流入地特征，流动特征包括流动经历、流动时长和流动范围。流入地特征主要指接入城市规模。流动经历是指个体在初婚前是否有流动经历，是一个二值虚拟变量，有流动经历则为1，否则为0。流动时间是指初婚前的流动总时长。流动范围是指跨县、跨市还是跨省流动。接入城市规模采用城市市辖区常住人口作为表征。其他的控制变量具体定义如表1所示。

表1 变量定义和统计描述

|  |  |  |  |
| --- | --- | --- | --- |
| 变量 | 变量定义和赋值 | 均值 | 样本量 |
| **个体特征** | | | |
| 初婚年龄 | 首次结婚的年龄=结婚年份-出生年份 | 23.83 | 119764 |
| 流动经历 | 初婚前是否有流动经历，有为1，否为0 | 0.26 | 119764 |
| 流动时间 | 初婚前流动总时长，回归中分为四种程度：无流动时间为0；0-2年赋值为1；2-5年赋值为2，5-8年赋值为3，8年及以上赋值为4 | 1.32 | 119764 |
| 流动范围 | 将本市跨县、本省跨市和跨省赋值为1、2和3，无流动经历则范围为0 | 0.37 | 119764 |
| 性别 | 男性为1，女性为0 | 0.53 | 119764 |
| 教育 | 教育程度分为：未上过学、小学、初中、高中、大学专科、大学本科和研究生7个等级，分别赋值1-7 | 3.31 | 119764 |
| 户口性质 | 分为农业、非农业和其他，分别赋值为0、1、2 | 0.21 | 119764 |
| 民族 | 分为汉族和少数民族，少数民族为1，汉族为0 | 0.08 | 119764 |
| 父母流动 | 流动之前父母是否有流动经历 | 0.16 | 119764 |
| **城市特征** | | | |
| 接入城市规模 | 市辖区常住人口（万人） | 227.60 | 106 |
| 利用外资 | 利用外资比重（%） | 2.02 | 106 |
| 产业结构 | 第三产业与GDP占比（%） | 49.29 | 106 |
| 人均道路  面积 | 人均城市道路面积 (平方米） | 14.02 | 106 |
| 每万人拥有汽车（辆） | 拥有汽车量/每万人 | 10.45 | 106 |

（三）模型设定

本文构建如下几个方程对两个假说进行分别论证。方程（1）是对全样本检验假说一：流动经历影响初婚年龄。是城市*j*中个体*i*的初婚年龄。表示个体*i*在初婚前是否有流动的经历。是*i*的系列个体特征，包含年龄、性别、教育等。是重点考察系数，表示流动经历对初婚年龄的影响程度。

方程（2）和方程（3）是进一步考察流动时长和流动范围对初婚的推迟效应。表示个体*i*在初婚前的流动时长，初婚前无流动经历则流动时长为0，如果有流动经历，则按照时间的长短分为2年以下，2-5年，5-8年，8年及以上赋值1-4等级。表示个体*i*在初婚前的流动范围，初婚前无流动经历则流动范围为0，如果有流动经历，则按照本市跨县、本省跨市、跨省的三个等级赋值1，2，3。这里需要说明的是考察流动时长和流动范围的回归方程是直接用流动时长和流动范围代替流动经历，而不是在流动经历的基础上交乘流动时间和流动范围，是因为流动时长和流动范围中已经包含了流动经历，只有流动经历为1的前提下才会有流动时长和流动范围。如果系数为正，说明流动的时间越长，推迟效应越强；同理，就流动范围而言，跨省流动比本省跨市、本市跨县流动的推迟效应要强。

（1）

（2）

（3）

方程（4）验证假说二，流入地城市规模对初婚年龄的影响。其中 用城市*j*的市辖区常住人口来表征，是关注的系数，如果系数显著为正，说明城市规模对个体的初婚年龄有显著推迟作用。其中是城市特征控制变量，包含产业结构、基础公共设施等。这里需要说明的是学界对城市规模的表征变量暂无统一，既可以是广泛范围的常住人口，也可以是狭隘范围的城区人口，也可以衡量经济发达程度的GDP水平或是建成区面积等。鉴于此，我们将用城区人口、GDP水平和建成区面积作为稳健性检验。

（4）

四、实证结果与分析

（一）基准回归

1.流动经历、流动时长和流动范围对初婚年龄的推迟效应

表2方程（1）的回归结果是对假说一流动经历的推迟效应进行验证。从回归结果可以看出，在控制了个体特征后，婚前具有流动经历个体比婚前没有流动的个体晚婚1.893年，且回归系数通过统计水平上1%的显著性。该回归结果显著为正，证实了假说一中流动经历的推迟效应。

表2方程（2）的回归结果是对流动时长的影响效应进行验证。根据流动时长变量的设定，回归系数是以流动时长为零为基准，回归系数表示每增加一个等级的流动时长，初婚年龄被推迟的效应大小。从第二列回归结果来看，相比无流动经历个体，流动时间为2年以下的流动经历个体晚婚1.081年，且1%统计水平的显著性。流动时间2-5年的回归系数与2年以内的系数接近，相较于无流动经历个体晚婚1.035年。对于流动时长达5-8年的个体，他们的初婚比无流动经历个体晚1.706年。当初婚前的流动时长增加到八年以上时，初婚年龄的推迟达到4.219年，同样系数在1%的统计水平上显著。从回归结果可以看出，随着流动时长的逐渐递增，回归系数逐渐增大，特别是流动时长5年或8年以上，对初婚的推迟效应急剧增加，假说一的流动时长越长推迟初婚的效应越强得到验证。

表2方程（3）的回归结果是对流动范围的影响效应进行验证。根据流动范围变量的设定，回归以无流动经历为基准，然后本市跨县、本省跨市和跨省的流动范围逐步扩大。从第三列回归结果来看，相比无流动个体，本市跨县的回归系数是1.861，且1%的统计水平显著，也就是说本市跨县流动比无流动个体要晚婚1.861年。进一步来看本省跨市和跨省流动的回归系数分别是1.918和2.025，随着流动范围的递增，推迟晚婚的程度越大。因此，回归系数的显著性和回归系数的递增性佐证了假说一中的流动范围越广对婚姻的推迟效应越强。

从控制变量来看，年龄的系数反映的不同出生年代的群体其婚龄状况，结果显示早期出生的流动人口比新生代农民工的婚龄更晚，可能是由于对流动状态的适应以及通信交通的发展使得新生代农民工较少受到流动经历本身的干扰。男性的初婚年龄显著比女性要晚，符合男大女小的主流匹配模式。教育水平越高的个体初婚年龄也会越大。少数民族比汉族结婚要早，这与刘厚莲（2014）的研究结论一致。非农业户口比农业户口晚婚，说明城市的婚配观念更倾向于晚婚。本人流动之前父母有流动经历的系数在-0.36和-0.44之间，且通过1%的显著性，对初婚年龄有负向影响，可能说明父母的流动经历提供了更好的进入婚姻的经济条件和作为子女长期与父母的分离带来的对婚姻的心理需求，有利于其子女在流动状态下尽早进入婚姻。

表2 流动经历、流动时长和流动范围对初婚年龄的影响

|  |  |  |  |
| --- | --- | --- | --- |
|  | 方程（1） | 方程（2） | 方程（3） |
|  | 初婚年龄 | 初婚年龄 | 初婚年龄 |
| 流动经历 | 1.893\*\*\*  (0.0250) |  |  |
| 流动时间（2年以下） |  | 1.081\*\*\*  (0.0538) |  |
| 流动时间（2-5年） |  | 1.035\*\*\*  (0.0346) |  |
| 流动时间（5-8年） |  | 1.706\*\*\*  (0.0395) |  |
| 流动时间（8年及以上） |  | 4.219\*\*\*  (0.0448) |  |
| 流动范围（本市跨县） |  |  | 1.861\*\*\*  (0.0290) |
| 流动范围（本省跨市） |  |  | 1.918\*\*\*  (0.0411) |
| 流动范围（跨省） |  |  | 2.025\*\*\*  (0.0703) |
| 年龄 | 0.0804\*\*\*  (0.00120) | 0.0780\*\*\*  (0.00118) | 0.0803\*\*\*  (0.00120) |
| 性别 | 1.469\*\*\*  (0.0204) | 1.393\*\*\*  (0.0201) | 1.470\*\*\*  (0.0204) |
| 小学 | 0.796\*\*\*  (0.0736) | 0.795\*\*\*  (0.0724) | 0.796\*\*\*  (0.0736) |
| 初中 | 1.298\*\*\*  (0.0723) | 1.301\*\*\*  (0.0711) | 1.298\*\*\*  (0.0723) |
| 高中 | 2.009\*\*\*  (0.0754) | 2.048\*\*\*  (0.0741) | 2.009\*\*\*  (0.0754) |
| 大学专科 | 3.167\*\*\*  (0.0821) | 3.259\*\*\*  (0.0807) | 3.163\*\*\*  (0.0821) |
| 大学本科 | 3.924\*\*\*  (0.0892) | 4.016\*\*\*  (0.0877) | 3.926\*\*\*  (0.0892) |
| 研究生 | 4.965\*\*\*  (0.173) | 5.055\*\*\*  (0.170) | 4.976\*\*\*  (0.173) |
| 少数民族 | -0.242\*\*\*  (0.0439) | -0.237\*\*\*  (0.0431) | -0.241\*\*\*  (0.0439) |
| 非农户籍 | 0.625\*\*\*  (0.0322) | 0.612\*\*\*  (0.0317) | 0.627\*\*\*  (0.0323) |
| 其他户籍类型 | 0.288\*\*\*  (0.0621) | 0.268\*\*\*  (0.0611) | 0.289\*\*\*  (0.0621) |
| 父母流动经历 | -0.359\*\*\*  (0.0284) | -0.434\*\*\*  (0.0279) | -0.358\*\*\*  (0.0284) |
| 户籍来源地 | 控制 | 控制 | 控制 |
| 常数项 | 18.48\*\*\*  (0.386) | 18.69\*\*\*  (0.379) | 18.49\*\*\*  (0.386) |
| R2 | 0.196 | 0.222 | 0.196 |
| 样本量 | 119764 | 119764 | 119764 |

注：括号内数值为标准误；\*\*\*、\*\*和\*分别表示1%、5%和10%的显著性水平。以下表格同此。

2.流入地城市规模对初婚年龄的推迟效应

表3显示了方程（4）的回归结果，是对假说二流入地属性——城市规模强化推迟效应的验证。从回归结果来看，回归系数为0.00912，统计上通过1%的显著性，即城市规模每增加10%，婚前流入到该城市的流动人口的初婚年龄提高0.0912%，从而验证了假说二接入城市规模强化了流动经历的推迟效应。当然，具体的影响路径将在下文中深入分析。

从其他控制变量来看，流动人口男性比女性更晚婚；随着教育等级上升，初婚年龄呈上升趋势；少数民族比汉族流动人口的初婚年龄要早，非农户籍相对农业户籍初婚更晚；父母具有流动经历的流动人口初婚年龄受到流动的影响更小。这些控制变量特征与上表2表现出的影响趋势具有相似性。值得注意的是，城市利用外资水平一直保持显著的负效应，可能是因为外资集中在劳动密集型行业，这些行业的流动人口本地化水平较低，保持着更多传统婚配观念，较少受到现代城市文化观念的冲击，从而在外资越多的城市，流动人口选择较早步入婚姻。

表3 流入地城市规模对初婚年龄的影响

|  |  |  |
| --- | --- | --- |
| Ln(初婚年龄)：方程（4） | | |
| 变量 | 系数 | 标准误 |
| Ln(城市规模) | 0.00912\*\*\* | 0.00262 |
| 年龄 | 0.0126\*\*\* | 0.000232 |
| 性别 | 0.0385\*\*\* | 0.00280 |
| 小学 | -0.00686 | 0.0208 |
| 初中 | 0.00945 | 0.0201 |
| 高中 | 0.0464\*\* | 0.0202 |
| 大学专科 | 0.0947\*\*\* | 0.0203 |
| 大学本科 | 0.114\*\*\* | 0.0205 |
| 研究生 | 0.147\*\*\* | 0.0229 |
| 少数民族 | -0.0146\*\* | 0.00576 |
| 非农户籍 | 0.0124\*\*\* | 0.00389 |
| 其他类型 | 0.00683 | 0.00762 |
| 父母流动经历 | -0.0101\*\*\* | 0.00320 |
| 利用外资 | -0.000587\*\*\* | 0.000130 |
| 产业结构 | -0.000310 | 0.000197 |
| 每万人拥有道路面积 | 0.000113 | 0.000347 |
| 每万人拥有汽车数量 | 0.0000655 | 0.000386 |
| 户籍来源地 | 控制 | |
| 常数项 | 2.883\*\*\* | 0.0521 |
| R2 | 0.415 | |
| 样本量 | 7905 | |

（二）内生性处理和稳健性检验

在识别流动经历效应和接入城市规模效应时，针对可能存在的内生性问题，本文进行了分析和处理，并使用替换变量和更换数据的方法进行稳健性检验。

1.倾向得分匹配法（PSM）处理流动经历对初婚年龄的推迟效应

在识别流动经历效应时，内生性的来源可能有两类：（1）结婚决策的时点会影响流动的概率，即越早结婚的群体，在婚前流动的概率越低，即婚前流动决策会影响婚姻决策，婚姻决策的时点也会影响婚前的流动决策，存在相互因果。（2）婚姻决策和流动决策可能都受到相同变量的影响，例如自身教育水平等。对此，本文采用倾向得分匹配（PSM）方法去寻找相似个体，对于相似个体，婚前有流动和婚前无流动的初婚年龄差异可以被认为是单纯的流动经历的推迟效应。具体而言，使用PSM方法来匹配婚前有流动与无流动的相似个体，然后计算相似个体因婚前有无流动经历而引致的初婚年龄差。处理组婚前有流动经历，对照组婚前无流动经历。根据Caliendo & Kopeinig（2008）的PSM步骤和选择合适的协变量标准，协变量必须是前定变量，影响是否流动且对初婚有影响，但不能是流动后造成的，因此本文确定的协变量为个体特征：性别、年龄、教育、少数民族、户口性质和父母的流动经历。

表4的A和B分别为匹配平衡检验结果和估计效应[[3]](#footnote-3)。A部分显示匹配后变量的标准化偏差小于10%，而且t检验结果不拒绝处理组和控制组无系统差异的原假设，匹配结果是有效的。B部分显示了采用近邻匹配方法（1:1）的估计结果，流动经历对初婚年龄的推迟效应为1.733年，且在1%的统计水平下显著。也就是说，相比婚前没有流动经历的个体，有流动经历个体的初婚年龄晚1.733年。

表4 匹配平衡检验和估计结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| A．PSM匹配平衡检验 | | | | | |
|  | 均值 | | 标准偏差% | 均值T检验 | |
|  | 处理组 | 对照组 | t值 | 相伴概率 |
| 性别 | 0.570 | 0.573 | -0.6 | -0.75 | 0.454 |
| 年龄 | 31.577 | 32.599 | -0.2 | -0.39 | 0.696 |
| 教育 | 3.769 | 3.774 | -0.4 | -0.52 | 0.604 |
| 少数民族 | 0.069 | 0.065 | 1.2 | 1.56 | 0.118 |
| 户口性质 | 0.239 | 0.235 | 0.9 | 1.13 | 0.260 |
| 父母流动经历 | 0.272 | 0.271 | 0.2 | 0.23 | 0.817 |
| B. PSM估计结果 | | | | | |
|  | 估计效应 | | 标准误 | 匹配模式 | |
| 流动经历 | 1.733\*\*\* | | 0.208 | 近邻匹配（1:1） | |

2.工具变量法处理城市规模对流动人口初婚年龄的推迟效应

在识别接入城市规模效应时，可能存在自我选择的样本偏误，即偏好或流动能力不同的个体会选择不同规模城市，而且，城市规模与流动人口可能同时受到一些宏观变量（例如城市GDP水平）的影响，这些内生性问题可以采用工具变量法进行解决。本文采用的工具变量分别为历史人口数据和自然地理特征（Combes et al，2011），本文对城市规模选择的工具变量是1953年的城市人口和城市地理特征——城市平坦土地面积[[4]](#footnote-4)，此工具变量满足与城市规模密切相关，但又不会直接影响随机误差项。采用二阶段最小二乘法（2SLS）进行回归的结果如表5。在第一步回归中工具变量——历史人口和城市平坦面积对城市规模有显著正相关关系，并且通过了相关性检验，第一阶段F值为1593.73，且在1%统计水平下显著。在第二步回归中，回归系数为0.0072，在1%统计水平下显著。因此，通过工具变量法[[5]](#footnote-5)进一步佐证了城市规模对初婚年龄的推迟效应。

表5 工具变量法检验城市规模对初婚年龄的影响

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 1st stage | | 2nd stage | |
|  | 城市人口规模 | | Ln(初婚年龄) | |
|  | 系数 | 标准误 | 系数 | 标准误 |
| 城市人口规模 |  |  | 0.0072\*\* | 0.0034 |
| 历史人口 | 0.534\*\*\* | 0.0052 |  |  |
| 城市平坦面积 | 0.0139\*\*\* | 0.0045 |  |  |
| 个体特征 | 控制 | | 控制 | |
| 城市特征 | 控制 | | 控制 | |
| R2 | 0.851 | | 0.130 | |
| 样本量 | 7905 | | 7905 | |

注：控制变量的设定同表三；表6-表11的个体特征和城市特征同此。

3.城市规模更换变量的稳健性检验

鉴于城市规模的表征变量具有不同的衡量方式，基准回归中采用城市常住人口进行表示，这里用城市城区人口、GDP和建成区面积对城市规模进行人口维度和经济维度的衡量。从表6的回归结果来看，城区人口和GDP回归系数比常住人口基准回归系数稍小，数量级保持不变，统计上通过1%的显著性。建成区面积回归系数与常住人口回归系数几乎保持相同大小。更换不同维度衡量的城市规模指标验证了基准回归的稳健性。

表6 城市规模更换变量的稳健性检验

|  |  |  |  |
| --- | --- | --- | --- |
|  | 城区人口 | GDP | 建成区面积 |
| 城市规模 | 0.00549\*\*\*  (0.00151) | 0.00606\*\*\*  (0.00199) | 0.00969\*\*\*  (0.00279) |
| 个体特征 | 控制 | 控制 | 控制 |
| 城市特征 | 控制 | 控制 | 控制 |
| 常数项 | 2.904\*\*\*  (0.0516) | 2.835\*\*\* (0.0580) | 2.885\*\*\*  (0.0520) |
| R2 | 0.415 | 0.415 | 0.415 |
| 样本量 | 7905 | 7905 | 7905 |

4.使用2015年CGSS数据进行再检验

仅使用正在流动的人口样本可能存在样本覆盖面不足，那些“曾经的流动人口”没有纳入到流动人口监测数据当中，即目前在户籍地并没有流动，但是在结婚前是有流动经历的人群。为此，本文使用2015年中国综合社会调查（CGSS）数据对全样本、流动人口样本和本地样本进行稳健性检验。从表7可以看出，针对流动人口和本地户籍人口的流动经历效应分别为1.767和0.745，全样本的流动效应为0.835，验证了流动经历会对初婚年龄产生显著的影响，这也说明本文针对流动人口的前提条件是被满足的。进一步发现全样本回归系数会因各组群体的占比不同而变化，此处 CGSS数据中流动人口比例仅占8%，与实际偏离较大，全样本回归可提供一个流动效应的区间值，但按不同群体来分类测算流动效应是较严谨的。流动效应1.767与本文基准回归1.893接近，本文的估计结果具有稳健性。

表7 采用CGSS数据的检验

|  |  |  |  |
| --- | --- | --- | --- |
|  | 全样本 | 正在流动人口样本 | 本地户籍样本 |
| 流动经历 | 0.835\*\*\*  (0.109) | 1.767\*\*\*  (0.351) | 0.745\*\*\*  (0.117) |
| 个体特征 | 控制 | 控制 | 控制 |
| 常数项 | 21.621\*\*\*  (0.344) | 15.051\*\*\*  (3.988) | 19.17\*\*\*  (0.230) |
| R2 | 0.197 | 0.278 | 0.193 |
| 样本量 | 8914 | 708 | 8206 |

五、城市规模强化推迟效应的机制分析

既然流入地城市规模可以强化流动经历对初婚年龄的推迟效应，那么这种强化效应是如何实现的就成为实证研究的应有之义。本文通过将流动经历对婚姻决策的影响过程显性化来分析接入城市规模推迟初婚年龄的影响路径，具体包括经济地位、婚姻观念和婚配市场的匹配概率。

（一）经济地位

根据传统婚姻观念构建的家庭效用模型的框架，男性的收入或者获得稳定收入能力的确定性成为进入婚姻的重要因素，预期稳定收入越高，其在婚姻市场的价值越高，较早进入婚姻的概率就越大。女性的经济独立性也会影响其婚姻决策，女性经济独立性的提升降低了其在婚姻中的收益，从而降低了进入婚姻的意愿（Becker，1991）。总体来看，男性相对女性的经济地位越高，初婚年龄就会越低。

流动人口往往为了获得更多的收入或者更确定性的预期收入而来到流入地，其收入水平除了受到个人能力影响外，还会受到接入城市规模的影响。根据相关研究，由于城市规模的集聚效应带来工资溢价（潘丽群等，2020），即规模越大的城市，工资水平会更高，然而不同性别的劳动力在城市中获得的收入溢价具有异质性（潘丽群、张少华，2021）。表8显示接入城市规模对流动人口收入的影响，从回归系数来看，接入城市规模每扩大10%，收入会增加0.709%，而且男性增加的幅度（0.766）显著低于女性（0.910）。相比男性，女性由于城市规模扩大获得的工资溢价更高，这可能是由于城市化过程中增加了对服务业人员的需求，女性更容易获得岗位，而且在发达的市场化地区，男女同工同酬也比较完善。

结合实证，一方面，女性流动人口因为流动，经济地位的提升会推迟其进入婚姻的年龄，而且，接入城市规模越大，女性流动人口获得的城市规模工资溢价也更高，所以其经济能力的提高相对更多，也相对更晚婚。另一方面，男性流动人口虽然也因为流动获得了更多的收入，然而，相较在同样接入城市的女性流动人口，其获得的溢价相对较少，造成在同样接入城市的流动人口中，男女相对收入随着接入城市规模越大而变小。在以男性为主要家庭经济支柱的观念下，收入水平的相对下降，降低了男性的婚配市场价值，从而推迟其初婚年龄。因此，接入城市规模会通过改变流动人口的男性和女性的相对经济地位，而强化流动经历的推迟效应，佐证了假说二的路径之一。

表8 经济地位的差异性

|  |  |  |  |
| --- | --- | --- | --- |
|  | Ln(收入) | Ln(男性收入) | Ln(女性收入) |
| Ln(城市规模) | 0.0709\*\*\*  （0.0120） | 0.0766\*\*\*  (0.0183) | 0.0910\*\*\*  (0.0229) |
| 个体特征 | 控制 | 控制 | 控制 |
| 城市特征 | 控制 | 控制 | 控制 |
| 常数项 | 6.294\*\*\*  （0.0905） | 6.395\*\*\*  (0.110) | 6.220\*\*\*  (0.146) |
| R2 | 0.183 | 0.193 | 0.224 |
| 样本量 | 6657 | 4229 | 2428 |

注：（1）对于男性和女性分组样本的回归，本文进行了回归系数显著性的SUEST检验，检验表明系数差异性显著。（2）此表采用城市规模IV回归，后续表格皆为城市规模IV回归。

（二）婚姻观念

婚姻观念在近现代以来发生了明显的变化，从传统婚姻观向后现代婚姻观进行转变。按初婚年龄来刻画婚姻观念的话，越传统的婚姻观念越提倡早婚；越现代的婚姻观念，强调个体自致性或个体价值的自我实现，往往更晚婚（齐亚强、牛建林，2012）。当然，在不同时期不同地点，婚姻观念仍然存在很大的差异。总体而言，越大规模的城市，其婚姻观念越靠近后现代化（Mu & Yeung，2019）。

流动人口的婚姻观念在流动过程中受到两方面的冲击：一方面要面对家庭环境所倡导的早婚早育的传统婚姻观，另一方面，又将面对接入城市相对更开放、更倡导个人因素的婚姻观念的吸引。虽然观念是难以衡量的，但越远离传统观念，越长停留在现代观念的环境中，观念越容易被同化。

本文从问卷中筛选出“是否独自流动”这个问题作为远离传统观念程度的一种衡量。如果独自流动，相对远离原生的社会关系网络的婚配舆论压力，来源地的观念更少地被保留，流动人口的婚姻观念更容易被流入地同化（Ikamari，2005），从而现代化的婚姻观念对其初婚推迟更严重；如果伴随着家人流动，较大程度地保留来源地婚姻观念，流入地观念对其影响较小，推迟效应相对减弱。另外，本文也认为停留时间长则流入地观念影响较深，而停留时间短则影响相对较弱。针对这两个衡量指标，将独自流动赋值为1，家人伴随流动赋值为0；将流动时间5年以上赋值为1，5年及以下赋值为0。表9的A部分第一列显示了城市规模越大，越容易独自流动，即城市规模越大影响其观念的可能性越大。第二列城市规模越大，个体流动的停留时长越长，从而越容易受到城市观念的同化。根据两个层面的观念指标，进一步按照流动状态和流动时长分组回归（表9的B部分），结果表明独自流动的个体在越大的城市，停留时间越长，其初婚年龄越晚。因此，接入地城市规模通过更后现代化的婚姻观念的同化强化了流动经历对初婚年龄的推迟效应，佐证了假说二的路径之一。

表9 城市规模对观念和初婚年龄的影响

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A.城市规模对观念的影响 | | | | |
|  | 是否独自流动 | | 流动时间长短 | |
| Ln(城市规模) | 0.0168\*\*\*  （0.0063） | | 0.0325\*\*\*  （0.0105） | |
| 个体特征 | 控制 | | 控制 | |
| 城市特征 | 控制 | | 控制 | |
| R2 | 0.088 | | 0.038 | |
| 样本量 | 7905 | | 7905 | |
| B. 不同程度城市观念影响下的初婚年龄 | | | | |
|  | 独自流动 | 家人伴随流动 | 流动时间长  （>5年） | 流动时间短  （<=5年） |
| Ln(城市规模) | 0.0082\*\*\*  (0.0024) | 0.0036\*\*\*  (0.0011) | 0.0089\*\*  (0.0032) | 0.0051\*\*\*  (0.0013) |
| 个体特征 | 控制 | 控制 | 控制 | 控制 |
| 城市特征 | 控制 | 控制 | 控制 | 控制 |
| R2 | 0.201 | 0.237 | 0.361 | 0.434 |
| 样本量 | 5447 | 2458 | 2401 | 5504 |

注：对于B部分分组样本的回归，本文进行了回归系数显著性的SUEST检验，检验表明系数差异性显著。

（三）匹配概率

从婚姻搜索匹配模型的理论框架来看，婚配成功与否取决于婚姻市场的搜寻范围和搜索时间、成本（McGoldrick & Robst，1996；Drewianka，2003）。一般而言，男高女低被视为主流匹配模式，称为正向匹配，正向匹配模式下的初婚年龄较低，而其它方向的匹配模式则初婚年龄较高。流动人口进入到流入地婚姻市场时，对男性流动人口而言，一方面是城市规模越大，扩大了婚姻市场择偶的范围；但另一方面，男性流动人口的教育水平相对较低，在高等学历人群比例更高的大城市寻找到潜在配偶的可能性反而降低了，即适配对象减少，婚配范围反而缩小了；第三，流动人口由于既游离于接入城市的社会网络，又远离流出地区的社会网络，搜索成本实际上是增加的。这些因素的综合效应更可能降低男性流动人口的正向匹配概率，从而推迟其初婚年龄。对女性流动人口而言，一方面其婚配范围因为进入接入城市而扩大，另一方面在正向婚配模式下，女性可以将婚配范围扩展到接入城市的更多优质男性群体，进而拉长她们的搜寻时间，降低了匹配成功率。所以，在越大的接入城市，虽然男性和女性受到婚姻市场变化的影响不同，但均产生了降低正向匹配概率的结果，正向匹配概率下降导致初婚年龄的提高。

本文对匹配变量的设定为：当丈夫教育程度高于或等于妻子教育程度定义为正向，赋值为1，反之则为0。用城市规模对匹配进行回归，如果城市规模系数为负，则说明城市规模降低了流动人口的匹配概率；如果为正则提升匹配概率。表10显示了城市规模对正向匹配的回归结果，OLS的回归系数是-0.0250，随着城市规模每扩大10%，正向匹配降低了0.00250个单位。对于Probit回归，平均边际效应系数为-0.0553，统计上显著地降低正向匹配概率。两种回归稳健地证明了接入城市规模对流动人口匹配概率的影响，通过降低正向匹配概率，推迟初婚年龄，佐证了假说二的路径之一。

表10 城市规模对正向匹配的影响

|  |  |  |
| --- | --- | --- |
|  | OLS回归  （线性概率模型） | Probit回归  (平均边际效应) |
|  | 正向匹配 | 正向匹配 |
| Ln(城市规模) | -0.0250\*\*\*  (0.0088) | -0.0553\*\*\*  (0.0158) |
| 个体特征 | 控制 | 控制 |
| 城市特征 | 控制 | 控制 |
| R2 (Pseudo R2) | 0.048 | 0.052 |
| 样本量 | 7905 | 7905 |

六、异质性分析

(一)流动经历推迟效应的城市异质性分析

将城市规模分为中小、大、超大和特大城市[[6]](#footnote-6)，用交互项来表示针对有流动经历的个体，不同等级的城市规模对初婚年龄的影响。从回归结果来看，相比中小城市而言，城市规模与流动人口的初婚年龄正相关，随着流入地城市规模越大，初婚年龄越大，晚婚现象越普遍。具体而言，流入地是大城市的流动人口初婚年龄比流入中小城市的晚婚0.0235，回归系数在5%统计水平显著；当流入地是特大城市时，交互项系数增大至0.117；当流入地是超大城市时，交互项系数进一步增大至0.557，且在1%统计水平上显著。因此，回归系数表明城市规模对初婚年龄的推迟强化作用随着城市规模的增大而递增。

表11 基于城市规模等级的异质性

|  |  |  |
| --- | --- | --- |
| 变量 | 系数 | 标准误 |
| 流动经历 | 1.715\*\*\* | 0.118 |
| 大城市 | 0.277\*\*\* | 0.0645 |
| 特大城市 | 0.182\*\* | 0.0806 |
| 超大城市 | -0.0844 | 0.0935 |
| 流动经历\*大城市 | 0.0235\*\* | 0.0113 |
| 流动经历\*特大城市 | 0.117\*\* | 0.0532 |
| 流动经历\*超大城市 | 0.557\*\*\* | 0.144 |
| 个体特征 | 控制 | |
| 城市特征 | 控制 | |
| R2 | 0.198 | |
| 样本量 | 66352 | |

（二）推迟效应的个体异质性分析

从性别、户籍、年龄段和教育程度四个方面考察婚前的流动经历及其接入城市规模对异质性个体的初婚年龄推迟效应。根据表12各交互项的系数，婚前有流动经历对男性初婚年龄的推迟效应更大，晚1.686年，这与许琪（2015）外出经历对男性的推迟效应更大相印佐证。而且这一推迟效应在更大的城市作用越强，可能原因是城市对女性收入提升作用更大，而城市婚配市场中，正向匹配的传统婚配观念导致较低竞争力男性匹配概率下降。

在教育水平上，随着教育水平提高，流动经历对初婚年龄的推迟效应逐渐增大，尤其是对大专及大学以上的流动者推迟效应更大。接受更多的教育本身就会受到更大的现代婚姻观念影响，同时城市规模对高教育劳动力收入提升作用更大，所以，随着教育程度的提升所受到城市规模推迟效应更强。

从户籍来看，对非农业比农业户籍的推迟作用大，推迟效应0.952，其他户籍类型（居民户籍）也比农业户籍稍大。由于城乡二元结构，流动人口很难融入本地城市，所以婚配模式主要依赖于流出地的社会网络，相比于非农业人口，农村的家庭婚配观念比较传统，更倾向于较早进入婚姻。而且，越大规模的城市中低技能的流动人口本地化越困难。

从年龄段来看，流动经历对于60年代至90年代的初婚影响逐渐递减，对于六七十年代的个体影响最大，而对80-90年代的新生代移民影响相对小一些。这可能的原因：一方面是在相对年长的流动人口中，处于底层的被挤压的男性流动人口占据了部分比例，这部分人很难实现向下婚配，从而难结婚晚结婚；另一方面是年轻的新生代流动人口，成长于市场化程度深入和旧体制烙印褪去的社会环境中，他们的资源、能力与价值观念也更开放，新生代流动人口更适应流动状态，相对而言，更容易进入婚姻。

表12 基于个体特征的异质性

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A.流动经历对初婚年龄推迟效应的异质性 | | | | |
|  | 初婚年龄 | 初婚年龄 | 初婚年龄 | 初婚年龄 |
| 流动经历 | 1.063\*\*\*  (0.0348) | 1.604\*\*\*  (0.0314) | 1.743\*\*\*  (0.0268) | -1.932\*\*\*  (0.100) |
| 流动经历\*男性 | 1.686\*\*\*  (0.0400) |  |  |  |
| 流动经历\*高中 |  | 0.528\*\*\*  (0.0491) |  |  |
| 流动经历\*大专 |  | 1.501\*\*\*  (0.0605) |  |  |
| 流动经历\*大学及以上 |  | 2.314\*\*\*  (0.0678) |  |  |
| 流动经历\*非农业户口 |  |  | 0.952\*\*\*  (0.0491) |  |
| 流动经历\*居民户口 |  |  | 0.614\*\*\*  (0.110) |  |
| 流动经历\*25-35岁 |  |  |  | 2.637\*\*\*  (0.0763) |
| 流动经历\*35-45岁 |  |  |  | 4.958\*\*\*  (0.0816) |
| 流动经历\*45岁以上 |  |  |  | 6.984\*\*\*  (0.108) |
| 个体特征 | 控制 | 控制 | 控制 | 控制 |
| R2 | 0.164 | 0.167 | 0.191 | 0.207 |
| 样本量 | 119764 | 119764 | 119764 | 119764 |
| B.接入城市规模对初婚年龄推迟效应的异质性 | | | | |
|  | Ln初婚年龄 | Ln初婚年龄 | Ln初婚年龄 | Ln初婚年龄 |
| Ln城市规模 | 0.0016\*\*  (0.0007) | 0.0730e-2  (0.0007) | 0.0038\*\*\*  (0.0007) | -0.0179\*\*\*  (0.0025) |
| 城市规模\*男性 | 0.0117\*\*\*  (0.0002) |  |  |  |
| 城市规模\*高中 |  | 0.0064\*\*\*  (0.0002) |  |  |
| 城市规模\*大专 |  | 0.0157\*\*\*  (0.0003) |  |  |
| 城市规模\*本科及以上 |  | 0.0215\*\*\*  (0.0004) |  |  |
| 城市规模\*非农业户口 |  |  | 0.0048\*\*\*  (0.0003) |  |
| 城市规模\*居民户口 |  |  | 0.0021\*\*\*  (0.0005) |  |
| 城市规模\*25-35岁 |  |  |  | 0.0200\*\*\*  (0.0005) |
| 城市规模\*35-45岁 |  |  |  | 0.0278\*\*\*  (0.0005) |
| 城市规模\*45岁以上 |  |  |  | 0.0339\*\*\*  (0.0005) |
| 个体特征 | 控制 | 控制 | 控制 | 控制 |
| 城市特征 | 控制 | 控制 | 控制 | 控制 |
| R2 | 0.165 | 0.167 | 0.169 | 0.195 |
| 样本量 | 7905 | 7905 | 7905 | 7905 |

七、结论与政策建议

本文运用2016年流动人口动态监测调查数据，考察分析了城市化过程中，流动人口的流动经历和流入地城市特征对他们初婚年龄的影响。研究发现：首先，流动经历显著地推迟了流动人口的初婚年龄1.893年。且初婚前的流动时长越长，初婚越晚，时长5年之内和5年以上具有较大推迟效应差异；从本市跨县、本省跨市到跨省流动，流动范围的扩张对初婚有递增的推迟效应。在使用倾向得分匹配方法处理流动和婚姻的内生性问题后，结果依然稳健。其次，随着流入地城市规模的扩大，初婚年龄被推迟的效应进一步加剧，城市规模每增加10%，流动人口初婚年龄推迟0.0912%。在运用城市历史人口和城市平坦土地面积作为城市规模工具变量控制内生性后，依然稳健。进一步地，将流动经历对婚姻决策的影响过程显性化，从经济地位、婚姻观念和婚配市场的匹配概率三个角度探讨了流入地城市规模推迟初婚年龄的机制。具体而言，个体获得流入地城市规模的工资溢价程度不同，从而导致其在婚配市场的不同经济竞争力而晚婚，特别造成了对男性流动人口的婚姻挤压；接入地越大规模越现代化的婚姻观冲击了流动人口来源地的婚姻观，导致晚婚；接入地城市规模降低了流动人口正向匹配的可能性，从而推迟初婚年龄。最后，流动经历和城市规模对初婚的效应表现出城市和个体的异质性。城市异质性，城市规模对初婚年龄的推迟强化作用随着城市规模的增大而递增；个体异质性，相比女性，男性受到推迟效应更为严重；相比低教育水平，越高教育水平个体，越晚婚；相比农业户籍，非农业户籍受到的推迟效应更严重；六七十年代的早期移民比八九十年代的新生代移民受到影响程度更大。

当然本文的研究是有限的，一个可能的不足之处在于仅通过分析已婚人群的数据来分析初婚年龄的推迟，缺乏对那些八零后、九零后仍未婚，如果结婚就属于晚婚人群的分析。拓展对这类群体的婚姻决策行为的分析，将是对流动与婚姻研究主题的进一步延伸。

流动人口的晚婚问题背后有深刻的社会原因和经济原因，既要看到其中社会发展规律的必然性，也要看到流动人口所承受的，因为快速的城市化进程中积累下来的问题而产生的婚配压力。从长远来看，让流动人口自由流动的同时，完全能够享有本地的公共服务，系统和持续地建设城乡融合是从根本上缓解婚配压力的手段，这也是我们实现社会主义现代化强国的重要措施。从短期来看，在现阶段可以有针对性地采取一些对策和手段，来缓解其中的不合理压力。

首先，通过城乡融合、公共服务均等化来保障流动人口的基本社会福利水平，降低经济地位在婚姻决策过程中的重要性。例如，开展技能培训、创业培训，保障同工同酬，享有住房公积金制度，将稳定就业的流动人口纳入公共住房保障，从而给流动人口市民待遇。同时，有针对性地提高农村户籍男性所能得到的社会福利范围和水平，例如养老保障、医疗保障以及子女教育保障水平，降低他们因流动到城市婚姻市场面临的婚姻挤压。

其次，在婚姻观念的引导方面，提倡多元化的婚姻理念，不唯经济地位的择偶标准。相比女性，男性流动人口在城市化扩张中得到的经济回报偏低，过度强调男性在婚姻中的经济支柱地位会导致“剩男”、“剩女”现象严重。新婚姻理念的倡导和推广将缓解目前的困境，尤其通过推广男女平等的观念，强调在家务育儿方面要共同分担的新家庭观念，来提高女性流动人口的婚姻预期收益，从而鼓励她们尽早进入婚姻。

第三，在青年婚介服务方面，城市管理部门举办各种公益联谊活动，鼓励本地和外地人共同参与，从而拓宽流动人口的社交网络，打破流动人口来源地通婚圈的局限性，拓展通婚圈，提升流动人口和本地人匹配的概率。当前要充分发挥共青团、妇联等群团组织作用，搭建青年婚恋交友、婚姻服务平台。

综上，通过提升公共服务均等化，促进流动人口本地化融入，淡化经济考量的婚姻理念，提高婚配市场的匹配概率，从而有效缓解初婚年龄的推迟和生育率的进一步下降，积极应对人口老龄化、少子化问题。

参考文献：

靳小怡 彭希哲 李树茁 郭有德 杨续松，2005：《社会网络与社会融合对农村流动妇女初婚的影响——来自上海浦东的调查发现》，《人口与经济》第5期。

李家兴，2020：《家庭背景、职业流动与婚姻匹配》，《社会发展研究》第1期。

李静 潘丽群 踪家峰，2015：《“门当户对”加剧收入不平等吗》，《统计研究》第11期。

刘厚莲，2014：《新生代流动人口初婚年龄及其影响因素分析——基于全国流动人口动态监测调查数据》，《人口与发展》第5期。

刘利鸽 刘红升 靳小怡，2019：《外出务工如何影响农村流动人口的初婚年龄?》，《人口与经济》第3期。

潘丽群 陈坤贤 李静，2020：《城市规模工资溢价视角下流动人口工资差异及其影响路径研究》，《经济学动态》第9期。

潘丽群 张少华，2021：《城市规模对性别工资差距影响的研究》，《贵州财经大学学报》2021第4期。

齐亚强 牛建林，2012：《新中国成立以来我国婚姻匹配模式的变迁》，《社会学研究》第1期。

许琪，2015：《外出务工对农村男女初婚年龄的影响》，《人口与经济》第4期。

杨雪 徐嘉树，2018：《城市流动人口婚育行为的实证研究》，《人口学刊》第4期。

於嘉 谢宇，2019：《中国的第二次人口转变》，《人口研究》第5期。

於嘉 赵晓航 谢宇，2020：《当代中国婚姻的形成与解体：趋势与国际比较》，《人口研究》第5期。

郑真真，2002：《外出经历对农村妇女初婚年龄的影响》，《中国人口科学》第2期。

曾迪洋，2014：《生命历程理论视角下劳动力迁移对初婚年龄的影响》，《社会》第5期。

Becker, G.S. (1973), “A theory of marriage: Part I”, *Journal of Political Economy* 81(4):813-846.

Becker, G.S. (1991), *A Treatise on the Family*, Cambridge: Harvard University Press.

Bergstrom, T. & R.F. Schoeni (1996), “Income prospects and age-at-marriage”, *Journal of Population Economics* 9(2):115-130.

Caliendo, M. & S. Kopeinig (2008), “Some practical guidance for the implementation of propensity score matching”, *Journal of Economic Surveys* 22(1):31-72.

Cantalini, S. (2017), “Does education affect the timing or probability of family formation? : An analysis of educational attainment and first union in Italy”, *Research in Social Stratification and Mobility* 49:1-10.

Combes, P. et al (2011), “The identification of agglomeration economies”, *Journal of Economic Geography* 11(2):253-266.

Drewianka, S. (2003), “Estimating social effects in matching markets: Externalities in spousal search”, *The Review of Economics and Statistics* 85(2):409-423.

Field, E. & A. Ambrus (2008), “Early marriage, age of menarche, and female schooling attainment in Bangladesh”, *Journal of Political Economy* 116(5):881-930.

Ikamari, L.D. (2005), “The effect of education on the timing of marriage in Kenya”, *Demographic Research* 12(1):1-28.

Jampaklay, A. (2006), “How does leaving home affect marital timing? An event-history analysis of migration and marriage in Nang Rong, Thailand”, *Demography* 43(4):711-725.

Kalmijn, M. (1994), “Assortative Mating by Cultural and Economic Occupational Status”, *American Journal of Sociology* 100 (2)：422-452.

Loughran, D.S.(2002), “The effect of male wage inequality on female age at first marriage”, *The Review of Economics and Statistic*s 84(2):237-250.

McGoldrick, K.M. & J. Robst (1996), “Gender differences in overeducation: A test of the theory of differential overqualification”, *American Economic Review* 86(2):280–284.

Mu, Z. & W.J. Yeung, (2019), “Internal migration, marriage timing and assortative mating: a mixed-method study in China”, *Journal of Ethnic and Migration Studies* 46(14):2914-2936.

Oppenheimer, V.K. (1988), “A theory of marriage timing”, *American Journal of Sociology* 94(3):563-591.

Oppenheimer, V.K. (1997), “Women's employment and the gain to marriage: The specialization and trading model”, *Review of Sociology* 23(1):431-453.

Qian, Y. (2017), “Gender asymmetry in educational and income assortative marriage”, *Journal of Marriage and Family* 79(2):318-336.

Yeung, W. J. & Z.Mu (2019), “Migration and marriage in Asian contexts”. *Journal of Ethnic and Migration Studies* 46 (14): 2863-2879.

1. \* 潘丽群，广州大学经济与统计学院，邮政编码：510006，电子邮箱：liqun\_pan@gzhu.edu.cn；李静，厦门大学经济学院、宏观经济研究中心，邮政编码：361005，新疆财经大学经济学院，邮政编码：830012，电子邮箱：lj-bingo@xmu.edu.cn；张少华，广州大学经济与统计学院，邮政编码：510006，电子邮箱：ahua1688@126.com。本文受教育部人文社会科学研究青年基金项目（20YJC790063），国家自然科学基金青年项目（72003052）和广州市宣传文化人才培养专项经费（穗宣通【2021】11号）的资助。感谢国家卫生计生委提供的全国流动人口动态监测调查数据，感谢匿审专家和编辑部的建设性意见，文责自负。 [↑](#footnote-ref-1)
2. 本文搜集了样本城市普查年份2000年和2010年的市辖区常住人口，并分析了两者与2016年市辖区常住人口的相关性，结果显示显著相关且相关系数分别为0.8778和0.9797。 [↑](#footnote-ref-2)
3. 本文采用Rosenbaum sensitivity analysis对倾向得分匹配方法的估计效应进行不可观测敏感性检验，临界值Gamma2.5在1%的统计水平下显著，2.5大于一般性要求的临界值2.2，可以认为表4中的PSM结果对遗漏的不可观测特征不敏感。 [↑](#footnote-ref-3)
4. 城市平坦度的基础数据来源于国家基础地理信息中心网站，本文最终使用城市平坦度数据经作者整理计算所得；城市历史人口来源于1953年中国国家统计局第一次全国人口普查数据。 [↑](#footnote-ref-4)
5. 针对使用工具变量法，我们进行了过度识别检验和弱工具变量检验。过度识别中Chi2(1)=4.042(P=0.444)，由于P=0.444接受原假说，两个工具变量外生合格。弱工具变量检验中Kleibergen-Paap rk Wald F统计量为5279.42，Stock-Yogo 10%的临界值为19.93，所以不存在弱工具变量问题。 [↑](#footnote-ref-5)
6. 根据国务院公布的城市规模划分标准，市辖区人口100万以下的为中小城市，市辖区人口100-500万为大城市，市辖区人口500-1000万为特大城市，市辖区人口1000万以上的为超大城市。 [↑](#footnote-ref-6)