

# 住房状况、社会地位与农民工的城市身份认同\*

## ——基于社会融合调查数据的实证分析

祝仲坤<sup>1</sup> 冷晨昕<sup>2</sup>

**摘要：**本文结合 2014 年国家卫生和计划生育委员会开展的社会融合调查数据，系统考察了住房状况对农民工城市身份认同的影响及其作用机制。研究表明：①自有住房对农民工形成城市身份认同有显著正向影响。在利用倾向得分匹配法和处理效应模型纠正可能的选择性偏误后，所得结果依然稳健。②政策性住房对农民工城市身份认同的影响不显著。合理的解释是，当前以公租房为核心的农民工住房保障政策未能发挥有效作用。③作用机制的分析结果显示，主观社会地位认知是住房状况影响农民工城市身份认同的中间机制之一，即住房状况的差距使得农民工主观社会地位认知产生分化，进而导致其城市身份认同存在差异。

**关键词：**农民工 住房状况 社会地位 城市身份认同 作用机制

**中图分类号：**F323.6 **文献标识码：**A

### 一、引言

尽管数量庞大的农业人口实现了向城市的转移，使中国农民工的总量持续增加，但农民工市民化进程仍面临巨大的挑战。长久以来，大量进入城市的农民工处于“半城市化”的尴尬境地，并未实现地位和身份的转变（王春光，2006）。已有的研究和实践进程表明，市民化并非只是使农民工完成从农业户籍向城镇户籍的转换，更重要的是让农民工真正融入城市（韩俊强，2013；杨菊华，2015）。伴随着公共服务逐步走向均等化，户籍制度本身所发挥的作用正在不断弱化。相对于户籍身份的转变，农民工如何适应心理层面的社会角色转变、形成对城市的归属感，才是当务之急。正如彼得·布劳（1988）所言，“流动的人……如果不接受并建立新的角色属性，那他们就不能适应新的环境”。从农民工角度

---

\*本文系中国人民大学 2016 年度拔尖创新人才培养资助计划、清华大学中国农村研究院博士论文奖学金项目“住房对农民工城市融入的影响机制与经验研究”（项目编号：201720）阶段性成果，同时得到国家自然科学基金项目“农民工的住房保障、在外居住抉择与家乡住房投资行为研究”（项目编号：71373271）的支持。作者感谢“2017 清华农村研究博士生论坛”各位与会者的有益评论，感谢国家卫生和计划生育委员会流动人口服务中心提供的数据支持，感谢匿名审稿专家和编辑老师的意见和建议。作者文责自负。冷晨昕为本文通讯作者。

来讲，若农民工仍存“过客”心理，他们将很难以主人翁的姿态为城市服务，其社会责任感也会因此弱化。反之，若农民工对城市有较高的归属感，则意味着他们接受了特定的行为准则与角色规范，能够依据“新的角色属性”规范自我行动（Stets and Biga, 2003；钱龙等，2015）。因此，在当前和今后一个时期，市民化进程中的关键问题之一就是要助推农民工突破心理壁垒、摆脱心理上的边缘化，逐步建立起对城市的身份认同。

在推动农民工建立城市身份认同的进程中，住房问题不容忽视。俗话说得好，“有房才有家”、“房在哪，家就在哪”，因此，住房是农民工在城市安身立命之所。对于农民工而言，在城市生存发展最大的消费支出是住房，最大的困扰还是住房。那么，住房状况会影响农民工的城市身份认同吗？进一步地，如果住房状况影响农民工的城市身份认同，那其中的作用机制是什么？为了回答以上问题，本文结合2014年国家卫生和计划生育委员会（以下简称“国家卫计委”）开展的社会融合调查数据，系统考察住房状况对农民工城市身份认同产生的影响以及可能的作用机制。本文研究结论不仅能为政府部门制定相关政策提供经验证据，还能为推动农民工市民化进程、实现新型城镇化战略提供决策参考。

本文余下部分的安排如下：第二部分对已有相关文献进行评述；第三部分介绍数据来源及概况，并对主要变量进行描述性统计；第四部分构建计量模型、展示并讨论实证分析结果；第五部分是结论及相关的政策启示。

## 二、文献评述

身份认同通常包含两重含义：一是“身份”，是个体在社会中的定位，是对“我是谁”的自我认知；二是“认同”，是对“我和谁一样”的自我认知（Tajfel et al., 1971）。身份认同的研究发轫于哲学领域，随后进入心理学领域，在被引入社会学领域后身份认同得到了广泛关注，并逐步成为社会学的重要研究范畴（Tajfel et al., 1971；Stets and Biga, 2003）。Akerlof and Kranton（2000）首次把身份认同引入经济学领域，并将其纳入新古典经济学框架，还以该框架为基础提出了身份认同假说，指出身份认同是一种社会规范，能够对经济行为、机会和福利等产生重要影响。

国内关于身份认同的经济学研究也并不少见。例如，李书娟、徐现祥（2016）、李书娟等（2016）通过考察中国省级官员对家乡身份认同产生的经济效应，为Akerlof的身份认同假说提供了来自中国的经验证据。雷震等（2016）同样依据Akerlof的分析框架，探讨了社会身份对腐败行为产生的影响。由于户籍制度这一特殊国情的存在，国内身份认同研究领域中关注农民工群体的研究相当丰富。研究的焦点主要集中在三个方面：

一是农民工身份认同的现实状况。作为中国户籍制度背景下特有的弱势群体（赵立，2017），已有研究表明，在当前阶段，农民工往往面临着进退失据的身份认同困境（郭星华、李飞，2009），成为了“双重边缘人”。一方面，农民工从农村迁徙至城市，希望有朝一日扎根城市，但城市社会中的种种歧视与排斥，导致他们难以融入；另一方面，农民工长期远离家乡故土，成员权利不断淡化，导致他们缺乏社会存在感（陈映芳，2005）。

二是农民工城市身份认同的外部性。已有研究表明，城市身份认同有助于增加农民工的劳动供给

(卢海阳、梁海兵, 2016), 改变农民工的消费方式、增加文化娱乐等符号性消费(钱龙等, 2015; 金晓彤等, 2017), 也会对农民工的职业规划、发展定位等产生正向激励(刘晓丽、郑晶, 2013)。杨菊华等(2016)还指出, 农民工较高的城市认同度有助于推动流入地政府在劳动就业、权益保障、公共服务等方面进行必要的调整, 从而使农民工获得更为平等的市民待遇。

三是影响农民工城市身份认同的关键因素。已有研究从多个层面探讨了影响农民工城市身份认同的因素, 概括起来主要包括制度约束、乡土记忆、城市生活体验、社会交往状况、人力资本水平、流入地公共服务等, 其中, 住房状况也被认为是影响农民工城市身份认同的重要因素之一(彭远春, 2007; 李荣彬、张丽艳, 2012; 刘晓丽、郑晶, 2013)。然而, 已有对农民工城市身份认同影响因素的研究常常是“眉毛胡子一把抓”, 将多侧面、多角度的多个因素同时放入模型, 这些研究很可能因为关注过多因素、忽视内生性问题而导致结论有所偏差。除此之外, 已有研究所用的数据以小规模、小范围调查为主, 不仅抽样方法可能有所偏差, 而且缺乏全国甚至省级层面的代表性。更为重要的是, 已有研究虽对住房状况影响农民工城市身份认同有所关注, 但多是旁敲侧击、一带而过, 缺少深入的定量研究, 尤其是缺乏对住房状况影响农民工城市身份认同的作用机制的研究。

关于住房状况对农民工城市身份认同的影响及其作用机制, 虽然缺乏深入的定量研究, 但已有研究在住房阶层化等方面的探讨为进一步研究提供了思路。伴随着中国实行住房市场化改革, 相当数量的中国人将目光投向房产(张海东、杨城晨, 2017), 拥有住房产权已经成为衡量个人或家庭经济能力和成就的重要指标(刘祖云、毛小平, 2012; 张文宏、刘琳, 2013)。住房已经不只是满足居住属性的物质实体, 还是一种符号性、地位性消费活动, 成为社会地位的重要象征(Frank, 1985)。正如赵晔琴、梁翠玲(2014)在研究中指出的, 在市场机制下, 住房消费本质上就是一种社会分化现象。李斌、王凯(2010)也指出, 在市场化时代, 遮风挡雨、驱寒保暖的住房正逐步“异化”为一种社会地位的筛选机制。

那么, 住房状况是否会通过这种社会地位的筛选机制来影响农民工的城市身份认同呢? 已有研究没有做出明确的回答, 但也有研究尝试在分析农民工城市身份认同的影响因素时, 同时考虑住房状况和社会地位。例如, 潘泽泉、何倩(2017)指出, 住房状况、社会地位均会直接影响农民工的城市身份认同。但略有遗憾的是, 该研究并未关注住房状况与社会地位之间的内在关联, 即住房状况的差异是否通过影响主观社会地位的认知而对农民工的城市身份认同产生影响。从逻辑上讲, 农民工在住房状况上面临的差异, 以及“参照系”由农村居民变为城镇居民产生的心理落差, 既可能直接作用于农民工对城市的认同感, 也可能率先“反馈”到农民工的主观社会地位认知层面, 然后再作用于他们对城市的身份认同层面。更为具体地说, 住房状况不佳的农民工更有可能感觉自己处在边缘群体、甚至是“二等公民”的地位, 也就更容易(或者说也不得不)认同自己的先赋农民身份; 相反, 住房状况较好的农民工则容易形成心理优势, 主观上感觉自己接近或已经融入城市主流群体, 与城市以及城市居民之间的心理距离不断缩短, 因此也就更容易产生城市身份认同。简而言之, 在住房状况影响农民工城市身份认同的过程中很可能存在“住房状况→社会地位→城市身份认同”的逻辑链条, 即住房状况的差别使得农民工主观社会地位认知出现分化, 进而导致农民工的城市身份认同存在差异。

有别于以往文献,本文可能的创新点在于:①研究视角新颖。本文专门研究住房状况对农民工城市身份认同的影响,在分析过程中,将住房状况对农民工城市身份认同的影响作为核心关注点,同时考虑了政策性住房<sup>①</sup>对农民工城市身份认同的影响。②实证策略严谨。本文使用国家卫计委开展的全国性大样本调查收集的数据,考虑到农民工居住在自有住房很可能是自选择的结果,利用倾向得分匹配法和处理效应模型纠正选择性偏误,并通过补对数—对数模型等进行了一系列稳健性检验,最终得到稳健、可信的实证结果。③本文基于中介效应模型探讨了住房状况影响农民工城市身份认同的作用机制,验证了“住房状况→社会地位→城市身份认同”的逻辑链条。

### 三、数据来源与变量描述

#### (一) 数据来源与说明

本文使用的数据来源于2014年国家卫计委组织的流动人口动态监测调查中的C模块<sup>②</sup>,即基于社会融合试点城市开展的“社会融合与心理调查”(下文简称“社会融合调查”)。样本覆盖北京市朝阳区、山东省青岛市、福建省厦门市、浙江省嘉兴市、广东省深圳市、广东省中山市、河南省郑州市和四川省成都市这8个社会融合试点城市。虽然这8个城市并非随机抽取,但城市内样本均采取PPS抽样方法具有较强的代表性(杨菊华,2015)。调查对象为在流入地居住1个月以上,非本市(区、县)户口的15~59周岁流动人口。由于本文关注的是农民工群体,因此仅考虑因务工经商迁移的农业户籍流动人口,最终获得11841个基准观测值。

#### (二) 变量选择与描述

1.被解释变量——农民工的城市身份认同。社会融合调查中,调查员通过询问“您认同自己是本地人吗?”来衡量农民工的城市身份认同情况,被访者回答的选项为“是”或“否”。调查问卷中同时设置了“您认同自己是老家人吗?”的问题,选项同样为“是”或“否”。在总样本中,认同自己是本地人的农民工有782人;认同自己是老家人的有9329人;还有1730个农民工既认同自己是老家人也认同自己是本地人,即存在“双向认同”(杨菊华等,2016)。考虑到双向认同群体可能与单向认同群体存在系统性偏差,从而影响到实证分析的准确性,故本文的主要分析过程并未将该群体纳入考察范围,即后面的实证分析中,如无特殊说明,本文使用的都是10111个单向认同的样本<sup>③</sup>。

2.核心解释变量——农民工住房状况。2014年社会融合调查问卷中,调查员向被访者询问了“您现在的住房属于哪种性质?”,被访者回答的选项包括“租住单位房”“租住私房”“廉租房”“公租房”“其他保障性住房”“已购商品房”“借住房”“就业场所”“自建房”“其他非正规场所”。本文根据分析需要将其分为三大类,其中,“已购商品房”“自建房”归为“自有住房”,“廉租房”

<sup>①</sup>政策性住房也称保障性住房,指的是由国家提供政策支持,各种社会主体通过新建或者其他方式筹集房源、专门面向中低收入群体提供的住房,目前主要包括经济适用房、廉租房、公租房等。

<sup>②</sup>具体信息详见国家卫生计生委流动人口数据平台(<http://www.chinaldrk.org.cn/home/>)。

<sup>③</sup>第四部分尝试将双向认同的样本考虑在内进行稳健性检验,估计结果并无显著差异。

“公租房”“其他保障性住房”归为“政策性住房”，其他各类型归为“租房”。样本中，农民工仍以租房为主，租房的农民工总量达到 9378 人，比重为 92.75%，居住在自有住房的农民工有 631 人，比重为 6.24%，而居住在政策性住房的农民工只有 102 人，比重仅为 1.01%。进一步地，考虑到本文后面部分重点关注的是自有住房对农民工城市身份认同的影响，且倾向得分匹配法、处理效应模型主要适用于核心解释变量为二分类变量的情形。为此，本文将租房和政策性住房两类合并起来，称之为“非自有住房”。

为了较为直观地呈现住房状况与农民工城市身份认同之间可能存在的关系，本文绘制了柱形图。如图 1 所示，总体来看，农民工认同城市身份的比例为 7.73%。其中，租房的农民工认同城市身份的比例最低，仅为 7.07%；居住在政策性住房的农民工认同城市身份的比例略高于租房农民工，为 7.84%；相比之下，居住在自有住房的农民工认同城市身份的比例最高，达到了 17.59%。

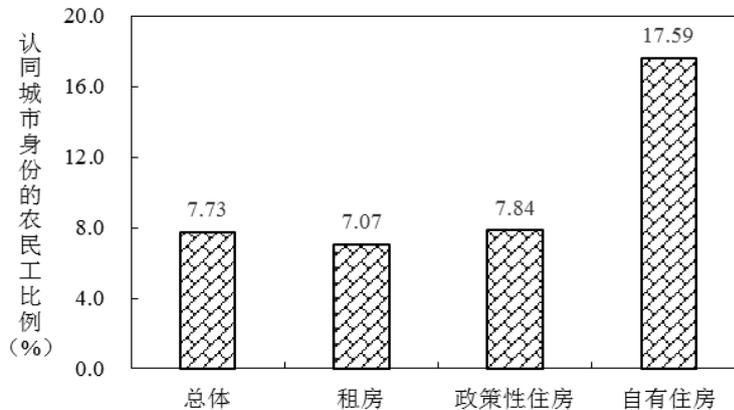


图 1 住房状况与农民工的城市身份认同

3. 中介变量。中介变量指的是解释变量对被解释变量产生影响的途径或机制。本文以“主观社会地位认知”为中介变量，进一步探讨住房是否通过影响农民工的主观社会地位认知（下文简称“社会地位”）进而影响其城市身份认同。问卷中向受访者询问了“与全社会的人们相比，您处在哪个位置？”，受访者回答的选项为 1~10 的分值，代表从低到高的 10 个等级，1 表示处在社会中的最低地位，10 则表示处在社会中的最高地位。样本范围内，农民工社会地位的平均分值为 5.06。其中，自有住房农民工的社会地位分值最高，为 5.70；租房农民工的社会地位分值次之，为 5.02；政策性住房农民工的社会地位分值最低，仅为 4.50。

4. 控制变量。依据 2014 年社会融合调查数据，并借鉴已有文献的做法，本文控制了可能影响农民工城市身份认同的变量，包括年龄、年龄平方项、性别、受教育程度、婚姻状态、健康状况、收入水平、流入本地时间、本地方言掌握情况以及行业类别<sup>①</sup>等。

变量的含义及其描述性统计见表 1。

<sup>①</sup>在 2014 年社会融合调查数据中，行业类别依据《国民经济行业分类》(GB/T4754-2011) 共划分为 20 类。考虑到样本中部分行业的农民工数量偏少，本文进一步将行业类别归为 6 类。

住房状况、社会地位与农民工的城市身份认同

表 1 变量的含义及其描述性统计

变量名称	变量含义和赋值	平均值	标准差	最小值	最大值
城市身份认同	“您认同自己是本地人吗？” 否=0, 是=1	0.077	0.267	0	1
住房状况					
租房	其他=0, 租房=1	0.928	0.259	0	1
政策性住房	其他=0, 政策性住房=1	0.010	0.099	0	1
自有住房	其他=0, 自有住房=1	0.062	0.242	0	1
社会地位	“与全社会的人们相比, 您处在哪个位置?” 1~10分, 最低为1分, 最高为10分	5.056	1.731	1	10
年龄	样本农民工2014年时的年龄(岁)	32.555	8.823	15	59
年龄平方项	年龄平方/100	11.377	6.121	2.250	34.810
性别	女=0, 男=1	0.583	0.493	0	1
受教育程度					
小学及以下	其他=0, 小学及以下=1	0.104	0.305	0	1
初中	其他=0, 初中=1	0.563	0.496	0	1
高中	其他=0, 高中=1	0.237	0.426	0	1
大专及以上	其他=0, 大专及以上=1	0.096	0.295	0	1
婚姻状态					
未婚	其他=0, 未婚=1	0.275	0.446	0	1
在婚	其他=0, 在婚=1	0.713	0.453	0	1
离异或丧偶	其他=0, 离异或丧偶=1	0.012	0.111	0	1
健康状况	差=1, 一般=2, 好=3, 很好=4, 非常好=5	3.756	0.973	1	5
收入水平	样本农民工月收入(元)(取自然对数)	8.513	0.559	5.298	12.044
流入本地时间	样本农民工在本地居留的时间(年)	4.055	4.330	0	33
本地方言掌握情况					
不懂	其他=0, 不懂=1	0.165	0.371	0	1
听得懂但不会讲	其他=0, 听得懂但不会讲=1	0.244	0.430	0	1
听得懂会讲一些	其他=0, 听得懂会讲一些=1	0.226	0.419	0	1
听得懂也会讲	其他=0, 听得懂也会讲=1	0.364	0.481	0	1
行业类别					
制造业	其他=0, 制造业=1	0.312	0.463	0	1
批发零售业	其他=0, 批发零售业=1	0.054	0.226	0	1
建筑业	其他=0, 建筑业=1	0.214	0.410	0	1
居民服务业	其他=0, 居民服务业=1	0.146	0.353	0	1
住宿餐饮业	其他=0, 住宿餐饮业=1	0.145	0.352	0	1
其他行业	其他=0, 其他行业=1	0.129	0.335	0	1

注: 观测值个数为 10111。

## 四、实证分析

### （一）模型设定

由于农民工的城市身份认同为二分类变量，OLS 估计可能导致异方差等问题，而二元选择模型能够更为准确地估计非线性影响（陈强，2014），因此，本文采用 Probit 模型进行估计。模型的表达式为：

$$P(\text{Identity}_i = 1) = P(\text{Identity}_i^* > 0) = \Phi(\alpha + \beta \text{House}_i + \gamma Z_i + \delta \text{City}_i + \varepsilon_i) \quad (1)$$

如（1）式所示， $\text{Identity}_i^*$  为潜变量。当  $\text{Identity}_i^* > 0$  时， $\text{Identity}_i = 1$ ；当  $\text{Identity}_i^* \leq 0$  时， $\text{Identity}_i = 0$ 。 $\text{Identity}_i$  表示第  $i$  位农民工的城市身份认同。 $\text{House}_i$  表示第  $i$  位农民工的住房状况（租房、政策性住房或自有住房）。 $Z_i$  为控制变量。 $\text{City}_i$  为城市虚变量，用以控制影响农民工城市身份认同的不可观测的地区因素。 $\alpha$ 、 $\beta$ 、 $\gamma$  和  $\delta$  为待估参数， $\varepsilon_i$  为随机扰动项。

### （二）基准回归

表 2 汇报了基于 Probit 模型的回归结果<sup>①</sup>，其中，方程 1 只控制了核心解释变量，方程 2 加入了控制变量，方程 3 进一步加入了城市虚变量。整体来看，模型运行良好，从方程 1 到方程 3，准  $R^2$  值逐步提高，Wald 卡方值均在 1% 的统计水平上通过了显著性检验。并且，核心解释变量的影响方向和显著性水平在各方程之间也没有显著变化，表明模型估计结果具有较强的稳健性。最重要的是，从各方程的估计结果可以看出，以租房为参照，自有住房变量在 1% 的统计水平上显著，且边际效应均为正。从方程 3 来看，与租房的农民工相比，居住在自有住房的农民工认同城市身份的概率显著提升 9.5%。政策性住房变量在统计上并不显著。

控制变量的回归结果与以往研究结论基本一致（例如李荣彬、张丽艳，2012；杨菊华等，2016）。  
 ①受教育程度对农民工城市身份认同有显著的促进作用。从方程 3 的边际效应结果可以看出，相比于小学及以下学历的农民工，高中学历和大专及以上学历的农民工更容易形成城市身份认同，初中学历的农民工次之。  
 ②健康程度越高，农民工认同城市身份的概率越大。  
 ③流入本地的时间越久，农民工越容易形成城市身份认同。  
 ④本地方言的掌握情况越好，农民工越容易产生城市身份认同。  
 ⑤年龄、年龄平方项、性别、婚姻状态及收入水平对农民工城市身份认同的影响在统计上不显著。

表 2 基准回归——住房状况对农民工城市身份认同影响的模型回归结果

变量	方程1		方程2		方程3	
	边际效应	t值	边际效应	t值	边际效应	t值
住房状况（以“租房”为参照）						
政策性住房	0.024	0.736	0.019	0.595	0.020	0.607
自有住房	0.164***	9.328	0.128***	7.779	0.095***	5.461

<sup>①</sup>考虑到 Probit 模型估计结果只包含解释变量显著性和参数符号方面的信息，无法提供各解释变量对农民工城市身份认同影响的大小（参见陈强，2014），因此，如无特殊说明，本文主要汇报变量的边际效应。

住房状况、社会地位与农民工的城市身份认同

年龄	—	—	0.003	1.190	0.002	0.878
年龄平方项	—	—	-0.002	-0.701	-0.002	-0.446
性别	—	—	0.000	0.013	-0.001	-0.147
受教育程度(以“小学及以下”为参照)						
初中	—	—	0.019**	2.345	0.022***	2.733
高中	—	—	0.023**	2.345	0.029***	2.904
大专及以上	—	—	0.019*	1.797	0.028**	2.263
婚姻状态(以“未婚”为参照)						
在婚	—	—	0.012	1.329	0.001	0.070
离异或丧偶	—	—	0.005	0.221	-0.004	-0.189
健康状况	—	—	0.017***	5.887	0.017***	5.389
收入水平	—	—	-0.009	-1.574	0.001	0.242
流入本地时间	—	—	0.002***	2.973	0.002***	3.459
本地方言掌握情况(以“不懂”为参照)						
听得懂但不会讲	—	—	0.018***	2.969	0.012*	1.912
听得懂会讲一些	—	—	0.046***	6.511	0.039***	4.567
听得懂也会讲	—	—	0.087***	12.023	0.092***	8.173
行业类别	未控制		已控制		已控制	
城市虚变量	未控制		未控制		已控制	
Wald卡方值	158.859***		413.831***		528.582***	
准R <sup>2</sup>	0.027		0.080		0.105	
观测值数	10111		10111		10111	

注：\*\*\*、\*\*、\*分别代表在 1%、5%、10%的统计水平上显著；篇幅所限，行业类别、城市虚变量的估计结果本表未列出，有兴趣的读者可向作者索取。

### (三) 内生性讨论

农民工是否居住在自有住房可能是自选择的结果，即自有住房变量可能并不满足随机抽样，用它直接回归可能因非随机抽样而使估计结果产生选择性偏误。但由于数据限制，本文无法获得可靠的工具变量。为了纠正可能的选择性偏误，本文运用了倾向得分匹配法(参见 Rosenbaum and Rubin, 1983)和处理效应模型(参见 Maddala, 1983)两种方法。之所以这样做，是因为这两种方法具有一定的互补性，同时运用便于结果之间的比较，也有助于判断估计结果的稳健性。倾向得分匹配法是依据可观测变量构建反事实框架来纠正选择性偏误的方法。正如陈强(2014)所述，该方法过于依赖可忽略性假定，对可观测变量的选取非常敏感，一旦变量选取过少或相关性不强，仍然难以获得一致性估计。相比之下，处理效应模型则是遵循 Heckman(1979)样本选择模型的传统，基于不可观测变量纠正选择性偏误的方法。当然，处理效应模型依赖于对方程的正确设定，因此，在没有足够把握证明处

理方程不存在模型误设或遗漏变量等问题时，对于分析结果的讨论也需要保持谨慎（王智波、李长洪，2016）。

1.倾向得分匹配法。本文首先对居住在自有住房的农民工（处理组）和居住在非自有住房的农民工（控制组）进行倾向值匹配，通过一系列影响农民工城市身份认同的因素建立 Probit 模型<sup>①</sup>，然后根据模型结果估计出农民工居住在自有住房的倾向值，并在此基础上运用多种方法进行匹配。方法一为最小邻近匹配，即寻找倾向得分最近的 k 个不同组个体，本文分别选取 k=1 和 k=4；方法二为局部线性匹配，即利用局部线性回归方法计算个体的权重，进而根据权重大小确定匹配个体；方法三为核匹配，该方法与局部线性匹配比较接近，所不同的是核匹配运用核函数计算权重。图 2 展示了匹配前后处理组与控制组倾向值的概率分布情况。从图 2 可以看出，匹配前两组样本差异极为显著，而匹配后，两组样本的差异明显缩小，接近了随机试验的效果。

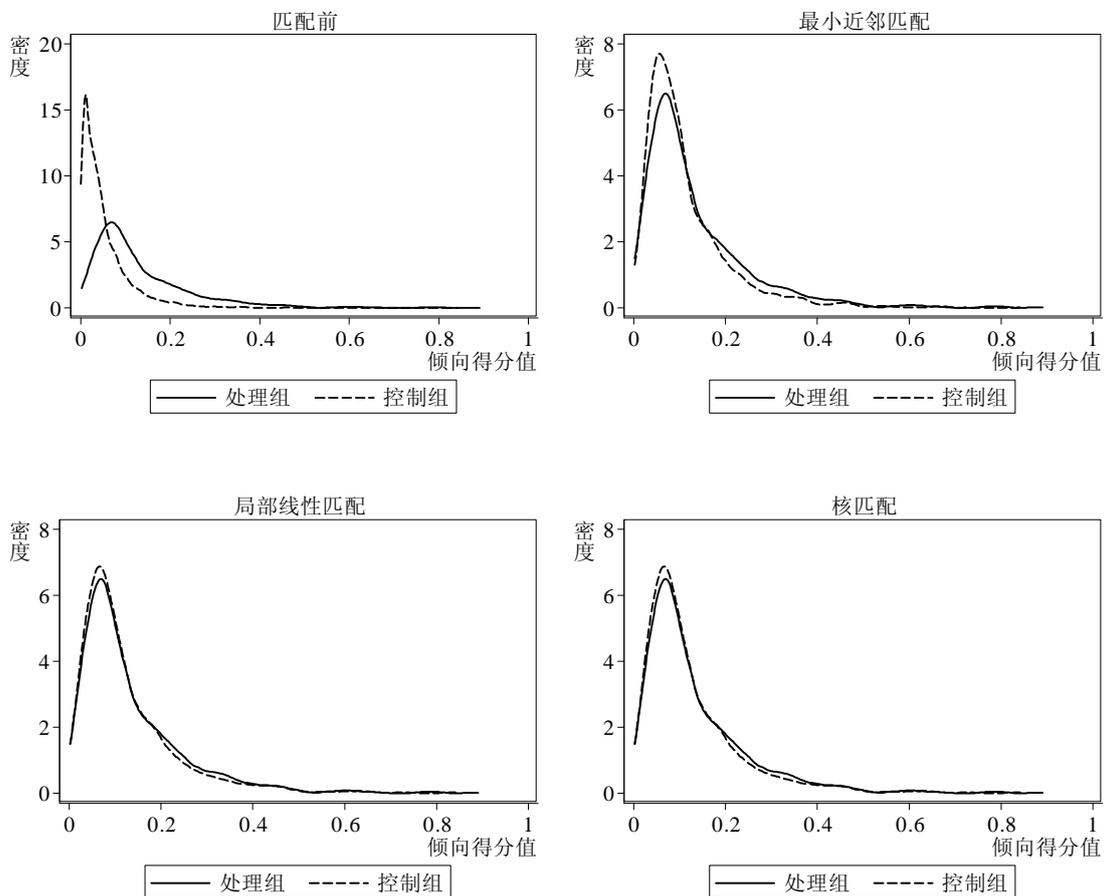


图 2 匹配前后处理组与控制组的倾向值概率分布

<sup>①</sup>陈强（2014）对倾向得分匹配法做出了细致的描述，本文不再赘述。考虑到倾向得分匹配法对可观测变量比较敏感，本文尝试对控制变量进行调整，并尝试建立 Logit 模型进行匹配，最终结果并无显著差异。

如表 3 所示, 本文利用自抽样法迭代 500 次, 测算出匹配后居住在自有住房的农民工和居住在非自有住房的农民工的 ATT、ATE 和 ATU<sup>①</sup>。在不同匹配方法之下, ATT 的结果均显示, 在消除了样本间可观测的系统性差异后, 自有住房对农民工城市身份认同在 1% 的统计水平上存在显著的正向作用。具体来看, 核匹配得出的 ATT 为 0.137, 最小近邻匹配 (k=4) 得出的 ATT 为 0.134, 局部线性匹配和最小近邻匹配 (k=1) 得出的 ATT 均为 0.133。虽然在不同匹配方法之下 ATT 的数值略有差异, 但与基准回归所得结论一致, 进一步证实了自有住房有助于农民工形成城市身份认同。

表 3 内生性讨论——不同匹配方法的估计结果

	方程 4		方程 5		方程 6		方程 7	
	最小近邻匹配 (k=1)		最小近邻匹配 (k=4)		局部线性匹配		核匹配	
	系数	t 值	系数	t 值	系数	t 值	系数	t 值
ATT	0.133***	5.281	0.134***	6.548	0.133***	7.621	0.137***	7.628
ATE	0.118***	5.796	0.109***	4.083	0.146***	5.092	0.155***	5.342
ATU	0.108***	3.862	0.108***	3.858	0.147***	4.948	0.154***	4.738

注: \*\*\*、\*\*、\*分别代表在 1%、5%、10% 的统计水平上显著。

2. 处理效应模型。表 4 汇报了处理效应模型的估计结果<sup>②</sup>, 本文同时运用两步法和极大似然法来估计模型。如表 4 所示, 两步法中的内生性参数  $\lambda$  和极大似然法中的内生性参数  $\rho$  在统计上均不显著, 这说明基准回归结果并不存在严重的选择性偏误。进一步看, 两步法和极大似然法的估计结果均显示, 自有住房对农民工形成城市身份认同有显著的正向作用, 与基准回归和倾向得分匹配法所得结果保持一致。

表 4 内生性讨论——处理效应模型的估计结果

变量	方程 8: 两步法		方程 9: 极大似然法	
	系数	t 值	系数	t 值
住房状况 (以“非自有住房”为参照)				
自有住房	0.243***	4.504	0.127***	6.231
常数项	-0.185***	-2.763	-0.193***	-3.142
控制变量	已控制		已控制	
城市虚变量	已控制		已控制	
$\lambda$	-0.009	-0.312	—	—
$\rho$	—	—	-0.032	-0.793
Wald 卡方值	999.796***		563.940***	
观测值	10111		10111	

注: \*\*\*、\*\*、\*分别代表在 1%、5%、10% 的统计水平上显著。

<sup>①</sup>ATT 是本文最关注的, 指的是处理组平均处理效应 (average treatment effect on treated), ATE 指的是平均处理效应 (average treatment effect), ATU 指的是控制组平均处理效应 (average treatment effect on untreated)。

<sup>②</sup>王智波、李长洪 (2016) 对处理效应模型的设定及检验做了非常细致的描述, 本文不再赘述。

#### （四）稳健性检验

除了利用倾向得分匹配法和处理效应模型纠正可能存在的选择性偏误以外，本文还通过补对数—对数模型控制稀有事件偏差，并通过替换模型、替换样本和替换变量等展开一系列稳健性检验。

首先，样本中认同城市身份的农民工比例偏低（小于 10%），直接运用 Probit 模型进行回归可能导致估计结果有偏差，一般称之为“稀有事件偏差”（陈强，2014）。鉴于此，本文参考普冀喆、郑风田（2016）的做法，采取补对数—对数模型修正可能的偏差；随后，本文将基准回归中的 Probit 模型替换为 Logit 模型进行回归；再者，考虑到部分农民工对本地和老家具有双向认同，为此，本文尝试将双向认同的农民工纳入城市身份认同<sup>①</sup>行列重新进行回归；最后，本文利用两个与城市身份认同相关性较强的变量（农民工是否愿意在本地长期居住、农民工是否愿意转为本地户口），分别替换原被解释变量进行回归。如表 5 所示，上述各类稳健性检验的回归结果均显示，自有住房对农民工城市身份认同存在显著的正向影响，而政策性住房的影响在统计上不显著，结果与前文一致。

表 5 稳健性检验——替换模型、替换样本与替换变量

变量	方程10	方程11	方程12	方程13	方程14
	补对数—对数模型	Logit模型	替换样本：加入双向认同的农民工样本	被解释变量：农民工是否愿意在本地长期居住	被解释变量：农民工是否愿意转为本地户口
住房状况（以“租房”为参照）					
政策性住房	0.192 (0.484)	0.016 (0.491)	-0.038 (-0.898)	0.047 (0.935)	0.081 (1.486)
自有住房	0.842*** (7.515)	0.087*** (4.864)	0.115*** (9.705)	0.360*** (20.138)	0.123*** (5.836)
控制变量	已控制	已控制	已控制	已控制	已控制
城市虚变量	已控制	已控制	已控制	已控制	已控制
Wald卡方值	604.882***	541.730***	950.993***	1583.483***	1315.793***
准R <sup>2</sup>	—	0.105	0.084	0.151	0.104
观测值数	10111	10111	11841	10111	10111

注：括号内为经稳健标准误校准过的 t 值；\*\*\*、\*\*、\* 分别代表在 1%、5%、10% 的统计水平上显著。

#### （五）住房状况影响农民工城市身份认同的作用机制

通过前面的分析，本文已经得出自有住房有助于农民工形成城市身份认同的结论。接下来，本文基于中介效应模型探讨住房状况是否通过影响农民工的社会地位，进而影响其城市身份认同。Baron and Kenny（1986）提出了检验中介变量的因果效应逐步回归法（causal step regression），目前该方法已经得到了广泛的应用（温忠麟、叶宝娟，2014；林文声等，2017）。依据 Baron and Kenny（1986）的做法，本文设定如下中介效应模型：

<sup>①</sup>此处将双向认同的样本一并考虑在内，样本观测值数由 10111 增至 11841，认同城市身份的农民工比例为 21.21%。

$$Identity_i = \alpha + \beta House_i + \gamma Z_i + \delta City_i + \varepsilon_i \quad (2)$$

$$Social\_status_i = \alpha + \beta House_i + \gamma Z_i + \delta City_i + \varepsilon_i \quad (3)$$

$$Identity_i = \alpha + \beta House_i + \psi Social\_status_i + \gamma Z_i + \delta City_i + \varepsilon_i \quad (4)$$

(2) ~ (4) 式中,  $House_i$  仍然表示农民工的住房状况, 但与 (1) 式不同的是, 此处  $House_i$  是包括“自有住房”和“非自有住房”的二分类变量;  $Social\_status_i$  表示农民工的社会地位; 控制变量、待估参数、随机干扰项等与 (1) 式基本保持一致。考虑到社会地位是典型的排序数据, 故 (3) 式运用排序选择模型 (ordered probit 模型) 进行回归。(2) 式与 (4) 式中被解释变量为农民工的城市身份认同, 运用 Probit 模型进行回归。

中介效应模型的回归结果见表 6。方程 15 的结果显示, 自有住房对农民工城市身份认同有显著的正向影响, 与前文估计结果一致。方程 16 的结果显示, 自有住房对农民工的社会地位在 1% 的统计水平上有显著的正向影响, 说明居住在自有住房的农民工主观社会地位的认知水平更高。由方程 17 可知, 在同时控制了自有住房和社会地位后, 自有住房与社会地位均表现出对农民工城市身份认同的正向影响, 且在 1% 的统计水平上显著。本文进一步观察到, 自有住房对城市身份认同的边际效应从方程 15 中的 0.071 下降到方程 17 中的 0.067。这表明, 社会地位在住房状况影响农民工城市身份认同的过程中发挥着中介作用<sup>①</sup>。温忠麟、叶宝娟 (2014) 提出, 论证中介效应的存在性还需进行 Sobel 检验。本文利用自抽样法迭代 500 次得出的 Sobel 检验的 z 值为 2.512, 在 5% 的统计水平上显著。这进一步验证了住房状况通过影响农民工的社会地位进而影响其城市身份认同, “住房状况→社会地位→城市身份认同”的逻辑链条也得以证实。

表 6 中介效应模型——住房状况影响农民工城市身份认同的作用机制分析

变量	方程 15		方程 16		方程 17	
	被解释变量: 城市身份认同		被解释变量: 社会地位		被解释变量: 城市身份认同	
	边际效应	t值	边际效应	t值	边际效应	t值
住房状况 (以“非自有住房”为参照)						
自有住房	0.071***	6.168	0.009***	5.168	0.067***	5.983
社会地位	—	—	—	—	0.007***	4.203
控制变量	已控制		已控制		已控制	
城市虚变量	已控制		已控制		已控制	
Wald卡方值	527.403***		527.118***		526.517***	
准R <sup>2</sup>	0.105		0.074		0.109	

<sup>①</sup>根据中介效应模型的基本原理可知, 将核心解释变量与中介变量均纳入回归模型后, 若核心解释变量与中介变量的估计结果都显著, 且核心解释变量的回归系数下降, 则表明存在“部分中介效应”, 即中介变量是核心解释变量影响被解释变量的渠道之一。

观测值数	10111	10111	10111
------	-------	-------	-------

注：\*\*\*、\*\*、\*分别代表在 1%、5%、10% 的统计水平上显著。

## 五、结论与政策启示

本文结合 2014 年国家卫计委开展的社会融合调查收集的数据，系统考察了住房状况对农民工城市身份认同的影响及其作用机制。研究表明：①自有住房对农民工形成城市身份认同有显著的正向影响，在加入控制变量和城市虚变量后，该结论保持一致。②考虑到基准回归中潜在的内生性问题，本文运用倾向得分匹配法和处理效应模型纠正可能存在的选择性偏误，并通过替换模型、替换样本和替换变量等进行稳健性检验，结论依然成立。③政策性住房对农民工城市身份认同的影响不显著。最可能的解释是，当前以公租房为核心的农民工住房保障政策并未发挥应有的功能。④基于中介效应模型的作用机制分析表明，在住房状况影响农民工城市身份认同过程中存在着“住房状况→社会地位→城市身份认同”的逻辑链条，即住房状况的差距使得农民工主观社会地位认知出现分化，进而导致其城市身份认同存在差异。

本文研究结论可为政策制定者提供以下参考：在推进农民工市民化的进程中，各级政府要着重考虑住房因素对于农民工扎根城市、融入城市的关键作用。首先，可以结合城市商品房去库存政策，着力降低农民工进城购房“门槛”。可以考虑进一步加大对农民工购房的支持力度，如为农民工购房提供货币化补贴、降低首付比例、提高税费减免幅度等。还可以探索为农民工提供首套房（尤其是中低位住房、小户型住房）和二手房的专项购房优惠。其次，为发挥政策性住房在农民工融入城市进程中的作用，应进一步调整、优化当前的住房保障政策。可以尝试将公租房建设指标纳入地方政府的考核范围，同时强化中央政府的监督与问责。还可以尝试变“补砖头”为“补人头”，优先采用货币化方式提供住房保障，同时整合现有政策性住房建设资金用于住房租金补贴和购房贷款贴息，将更多符合条件的农民工纳入住房保障范围。

最后需要指出的是，本文还存在一些不足之处。第一，本文只关注了农民工的城市身份认同，而城市认同还包括文化认同、地域认同等，对这些问题还需要做进一步研究。第二，由于数据限制，本文主要是针对社会融合试点城市进行分析，这些城市均属于大城市，而中小城市的情况可能与大城市有所不同，也需要做进一步研究。第三，在当前阶段，农民工在大城市中仍以租房为主，居住在政策性住房的农民工比例明显偏低，回归结果可能存在偏误。今后如能进行专项调查，获取更加完备的调查数据，分析结果的稳健性和可信度将得到进一步提升。

### 参考文献

- 1.彼得·布劳，1998：《社会生活中的交换与权力》，张黎勒译，北京：华夏出版社。
- 2.陈强，2014：《高级计量经济学及 Stata 应用（第二版）》，北京：高等教育出版社。
- 3.陈映芳，2005：《“农民工”：制度安排与身份认同》，《社会学研究》第 3 期。
- 4.郭星华、李飞，2009：《漂泊与寻根：农民工社会认同的二重性》，《人口研究》第 6 期。

- 5.金晓彤、韩成、聂盼盼, 2017:《新生代农民工缘何进行地位消费?——基于城市认同视角的分析》,《中国农村经济》第3期。
- 6.韩俊强, 2013:《农民工住房与城市融合——来自武汉市的调查》,《中国人口科学》第2期。
- 7.雷震、田森、凌晨、张安全、李任玉, 2016:《社会身份与腐败行为:一个实验研究》,《经济学(季刊)》第3期。
- 8.李斌、王凯, 2010:《中国社会分层研究的新视角——城市住房权利的转移》,《探索与争鸣》第4期。
- 9.李荣彬、张丽艳, 2012:《流动人口身份认同的现状及其影响因素研究——基于我国106个城市的调查数据》,《人口与经济》第4期。
- 10.李书娟、徐现祥, 2016:《身份认同与经济增长》,《经济学(季刊)》第3期。
- 11.李书娟、徐现祥、戴天仕, 2016:《身份认同与夜间灯光亮度》,《世界经济》第8期。
- 12.林文声、秦明、苏毅清、王志刚, 2017:《新一轮农地确权何以影响农地流转?——来自中国健康与养老追踪调查的证据》,《中国农村经济》第7期。
- 13.刘晓丽、郑晶, 2013:《新生代农民工身份认同及其影响因素研究》,《华南农业大学学报(社会科学版)》第1期。
- 14.刘祖云、毛小平, 2012:《中国城市住房分层:基于2010年广州市千户问卷调查》,《中国社会科学》第2期。
- 15.卢海阳、梁海兵, 2016:《“城市人”身份认同对农民工劳动供给的影响——基于身份经济学视角》,《南京农业大学学报(社会科学版)》第3期。
- 16.潘泽泉、何倩, 2017:《居住空间、社会交往和主观地位认知:农民工身份认同研究》,《湖南社会科学》第1期。
- 17.彭远春, 2007:《论农民工身份认同及其影响因素——对武汉市杨园社区餐饮服务员的调查分析》,《人口研究》第2期。
- 18.普冀喆、郑风田, 2016:《高房价与城镇居民创业——基于CHIP微观数据的实证分析》,《经济理论与经济管理》第3期。
- 19.钱龙、卢海阳、钱文荣, 2015:《身份认同影响个体消费吗?——以农民工在城文娱消费为例》,《南京农业大学学报(社会科学版)》第6期。
- 20.王春光, 2006:《农村流动人口的“半城市化”问题研究》,《社会学研究》第5期。
- 21.王智波、李长洪, 2016:《好男人都结婚了吗?——探究我国男性工资婚姻溢价的形成机制》,《经济学(季刊)》第2期。
- 22.温忠麟、叶宝娟, 2014:《中介效应分析:方法和模型发展》,《心理科学进展》第5期。
- 23.杨菊华, 2015:《中国流动人口的社会融入研究》,《中国社会科学》第2期。
- 24.杨菊华、张娇娇、吴敏, 2016:《此心安处是吾乡——流动人口身份认同的区域差异研究》,《人口与经济》第4期。
- 25.张海东、杨城晨, 2017:《住房与城市居民的阶层认同——基于北京、上海、广州的研究》,《社会学研究》第5期。
- 26.张文宏、刘琳, 2013:《住房问题与阶层认同研究》,《江海学刊》第4期。
- 27.赵立, 2017:《农民工群体身份认同与群体偏好研究——自我认同的中介作用》,《浙江社会科学》第2期。
- 28.赵晔琴、梁翠玲, 2014:《融入与区隔:农民工的住房消费与阶层认同——基于CGSS2010的数据分析》,《人口与发展》第2期。
- 29.Akerlof, G. A., and R. E. Kranton, 2000, “Economics and Identity”, *Quarterly Journal of Economics*, 115(3): 715-753.

30. Baron, R. M., and D. A. Kenny, 1986, "The Moderator-mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations", *Journal of Personality & Social Psychology*, 51(6): 1173-1182.
31. Frank, R. H., 1985, "The Demand for Unobservable and Other Non-positional Goods", *American Economic Review*, 75(1): 101-116.
32. Heckman, J., 1979, "Sample Selection Bias as a Specification Error", *Econometrica*, 47(1): 153-161.
33. Maddala, G. S., 1983, *Limited-dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.
34. Rosenbaum, P. R., and D. B. Rubin, 1983, "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, 70(1): 41-55.
35. Stets, J. E., and C. F. Biga, 2003, "Bringing Identity Theory into Environmental Sociology", *Sociological Theory*, 21(4): 398-423.
36. Tajfel, H., M. G. Billig, R. P. Bundy, and C. Flament, 1971, "Social Categorization and Intergroup Behavior", *European Journal of Social Psychology*, 1(2): 149-178.

(作者单位: <sup>1</sup> 中国人民大学农业与农村发展学院;

<sup>2</sup> 中国农业大学经济管理学院)

(责任编辑: 张丽娟)

## **Housing Conditions, Social Status and Urban Identity of Migrant Workers: An Empirical Analysis Based on Social Integration Survey Data**

Zhu Zhongkun Leng Chenxin

**Abstract:** Based on the data from social integration survey organized by the National Health and Family Planning Commission in 2014, this article systematically examines the impact of housing conditions on the migrant workers' urban identity and its mechanisms. The study shows that, first of all, housing ownership has a significant positive effect on the formation of the migrant workers' urban identity. The results still appear robust after the study corrects the possible selective bias using the propensity score matching method and the treatment effect model. Secondly, the influence of policy-based housing on the urban identity of migrant workers is not significant. One reasonable explanation is that the current housing security policy for migrant workers, particularly public rental housing policy, has failed to play an effective role. Thirdly, the mechanism analysis shows that the subjective social status cognition is one of the intermediate mechanisms that affect migrant workers' urban identity. Therefore, the gap found in housing conditions produces differentiations in the cognition of migrant workers' subjective social status, which leads to a variation in their urban identity.

**Key Words:** Migrant Worker; Housing Condition; Social Status; Urban Identity; Mechanism