

住房公积金与新生代农民工留城意愿*

——基于流动人口动态监测调查的实证分析

祝仲坤

摘要: 本文结合 2014 年流动人口动态监测调查数据,系统考察了住房公积金的缴存对新生代农民工留城意愿的影响。研究表明,缴存住房公积金使新生代农民工的留城意愿显著提升 5.3%,这一结果在加入控制变量、地区虚拟变量后依然成立。进一步,本文利用倾向得分匹配法纠正可能存在的样本选择性偏误,利用 Bioprobit 模型和 CMP 方法控制潜在的内生性问题,所得结论依然稳健。分样本的研究结果表明,缴存住房公积金对“80 后”和在当地居留满 5 年新生代农民工的留城意愿影响更为明显。此外,住房公积金呈现出“嫌贫爱富”特征,其对高收入新生代农民工留城意愿的积极作用远大于对低收入新生代农民工的积极作用。

关键词: 住房公积金 留城意愿 新生代农民工 内生性

中图分类号: F323.6 **文献标识码:** A

一、引言

国家统计局发布的《全国农民工监测调查报告》显示,2016 年,中国农民工总量达到 2.82 亿人,占全国总人口的比例超过 20%,其中,外出农民工接近 1.70 亿人^①。农民工群体为中国经济高速发展做出了无以替代的特殊贡献。然而,广大农民工仍面临着“半城市化”的窘境,仅仅实现了地域转移与非农化,却并未实现身份和地位的转变。2014 年 3 月,李克强总理在《政府工作报告》中明确提出了“三个 1 亿人”的新型城镇化战略,其中,首要的“1 亿人”就是要在 2020 年实现约 1 亿农民工的市民化^②。对于广大农民工而言,市民化进程当中最大的消费支出是住房,最大的困扰还是住房。陈锡文(2013)曾指出,推进农民工市民化的关键之一就在于如何解决好农民工的城镇住

*本文系中国人民大学 2016 年度拔尖创新人才培养资助计划成果,同时,本文研究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,感谢匿名审稿专家提出的意见和建议。作者文责自负。

^①数据来源: http://www.stats.gov.cn/tjsj/zxfb/201704/t20170428_1489334.html。

^②资料来源: <http://lianghui.people.com.cn/2014npc/n/2014/0305/c376646-24535026.html>。

房问题。实际上，早在 2005 年和 2006 年，国务院及多个国家部委就曾连续出台文件，提出要将农民工纳入住房公积金体系。这一举措被认为是国家推动农民工在城市“生根立足”的破题之举，不仅有助于缓解农民工城市住房困难，让农民工共享城市发展果实、实现安居乐业，更是与国家新型城镇化战略相互契合。

目前，该项政策已经推行十年有余，住房公积金到底在农民工市民化进程中发挥着怎样的作用呢？对于这一问题的研究不仅有助于重新审视农民工住房公积金的发展现状，同时也为进一步缓解农民工城市住房困难、推进农民工市民化提供必要的借鉴。基于以上考虑，本文拟利用国家卫生和计划生育委员会（以下简称“国家卫计委”）组织实施的 2014 年流动人口动态监测的调查数据，系统考察住房公积金对农民工留城意愿的影响。需要指出的是，本文关注的农民工群体并非全体农民工，而是出生于 1980 年及以后的新生代农民工。之所以聚焦于新生代农民工，原因有二：其一，新生代农民工已经逐渐成为农民工的主体，总量已经超过 1.40 亿人^①；其二，新生代农民工普遍缺乏基本的农业生产技术和经验，他们脱离农业生产、向城市流动已经成为难以逆转的事实（李培林、田丰，2011）。相比之下，他们外出务工早已不是基于“生存理性”的“候鸟式”迁移，而是希望有朝一日能够融入城市、扎根城市，实现真正意义上的市民化。

二、文献综述

1991 年，中国借鉴新加坡的经验引入了住房公积金制度，并首先在上海开展试点，随后逐步向全国各大城市铺开（陈杰，2010）。随着 1998 年住房福利分配制度的取消，住房公积金逐渐演变为全国性的住房保障制度（顾澄龙等，2016），在居民住房消费中扮演了关键角色，为中国城镇住房体制顺利转轨、缓解居民住房困难发挥了“稳定器”和“助推器”的作用（刘丽巍，2013）。

自住房公积金制度建立以来，相关学术研究成果便纷至沓来。学者们的关注点非常广泛，包括对新加坡发展经验的总结、归纳，对住房公积金管理运作模式、保障水平、融资机制的探讨，以及住房公积金制度的改革方向与顶层设计等（参见 Chen and Deng, 2014；黄燕芬、李怡达，2017）。当然，从过往文献来看，研究焦点是住房公积金对房地产市场的影响以及制度本身的公平性问题。对于住房公积金能够显著增强居民的住房消费能力，已经形成了广泛共识（Tang and Coulson, 2016），但住房公积金很可能导致房价上涨，推动房地产市场“泡沫化”（顾澄龙等，2016）。与此同时，住房公积金还有“劫贫济富”之嫌（周京奎，2012），即高收入群体成为住房公积金制度的最大获益者，低收入群体往往因住房支付能力不足，无法通过住房公积金获得“廉价”贷款，而这一现象与住房公积金制度设立的初衷背道而驰。

随着改革开放进程的不断深入，进入城市的农民工数量持续攀升，业已成为支撑工业发展、促进城市繁荣的主力军。然而，长期以来农民工被排斥在城市住房保障体系之外，面临着严重的住房困难。2005 年 4 月，建设部等三部委发布了《关于住房公积金管理若干具体问题的指导意见》（建

^①数据来源：http://www.stats.gov.cn/tjsj/zxfb/201704/t20170428_1489334.html。

金管[2005]5号),首次提出城镇单位雇佣的进城务工人员可申请缴存住房公积金。2006年3月,国务院发布了《国务院关于解决农民工问题的若干意见》(国发[2006]5号),再次明确指出城镇单位聘用的农民工,可缴存住房公积金用于购买或者租赁自住住房。

自此之后,学术界一直在关注农民工住房公积金的实践进程,众多学者探究了农民工住房公积金政策面临的发展困境,剖析了其“症结”之所在。首先,从政策本身来讲,涉及农民工住房公积金的相关文件均属于规范性文件,不具备强制性,这导致各地区在政策设计、企业在政策执行上有较大的调整空间(郑小晴、胡章林,2008),给政策推行带来了极大的或然性。其次,在地方政府层面,各地区往往陷入监督与保障的“两难”境地(祝仲坤,2016)。强化监督意味着农民工权益可以得到保障,但可能会增加企业用工成本;反过来,弱化监督又会导致农民工权益受损。再者,在企业层面,住房公积金作为企业对职工的保障性福利之一,短期内会增加企业的用工成本,不符合企业利益最大化原则。不仅如此,农民工的流动性给企业缴存住房公积金带来了现实障碍(郑小晴、胡章林,2008)。当然,也有学者关注的是住房公积金在农民工市民化进程中所发挥的作用(例如王桂新、胡健,2015)。但此类文献往往是在探讨社会保障对农民工市民化影响的过程中^①,将住房公积金作为住房保障的一部分,在研究中一带而过。仅有汪润泉、刘一伟(2017)专门探讨了住房公积金对流动人口市民化意愿的影响。但他们关注的是整个流动人口群体,且所涉及数据仅来自全国7个大城市^②,而且在研究过程中忽视了住房公积金与市民化意愿之间潜在的内生性问题。

相比之下,本文可能的边际贡献在于:①视角比较新颖。与过往研究不同,本文专门探讨住房公积金对农民工市民化的作用,且关注的群体是市民化意愿最为强烈的新生代农民工。②数据有代表性。流动人口动态监测调查是由国家卫计委组织实施的覆盖全国的大样本调查,是研究农民工问题的权威性数据。③实证分析策略相对严谨。本文结合倾向得分匹配法构造反事实框架,以纠正可能存在的样本选择性偏误,还利用 Bioprobit 模型和 CMP 方法比较有效地控制潜在的内生性问题,以期得到稳健、可信的研究结论。

三、数据、变量与模型

(一) 数据来源与说明

本文使用的数据是由国家卫计委组织实施的2014年流动人口动态监测调查数据。该调查基于分层、多阶段、与规模成比例的抽样方法,覆盖了全国31个省、自治区、直辖市及新疆生产建设兵团共32个省级行政单位,调查对象为在流入地居住一个月以上,非本县(区、市)户口的15~59周岁流动人口。考虑到本文关注的群体为新生代农民工^③,故保留样本中出生于1980年及以后的农业

^①从已有研究来看,衡量农民工的市民化意愿一般以“是否愿意转户口”(转户意愿)为依据,但也有不少学者以“是否愿意在城市长久居留”(留城意愿)作为判别指标。两个指标的内涵虽略有差异,但在研究中往往不做进一步区分。

^②该文仅选取天津、上海、广州、武汉、哈尔滨、成都、兰州等7个大城市作为调查地点。

^③如无其他说明,后文中“农民工”一般指“新生代农民工”。

户籍流动人口，最终获得了包括 44642 个观测值的基准样本。

(二) 变量选择与描述性统计

1.被解释变量——留城意愿。问卷中通过询问“您是否打算在本地长期居住(5年以上)?”来衡量农民工的留城意愿。被访者回答的选项分别为“不打算”“没想好”“打算”，并分别赋值为1~3的整数。从问卷调查结果来看，“不打算留城”的农民工有8272人，占比为18.53%；“没想好”的农民工有17093人，占比为38.29%；“打算留城”的农民工有19277人，占比为43.18%。

2.核心解释变量——住房公积金缴存情况。问卷向被访者询问了“您在本地是否拥有住房公积金?”，被访者回答的选项分别为“有”“没有”“不清楚”，考虑到“不清楚”很难归类，且样本范围内“不清楚”的样本只有1101个，占比仅为2.47%，因此做删除处理^①。随后，本文将选项“有”赋值为1，“没有”赋值为0。根据表1中的描述性统计可知，缴存住房公积金的农民工有4651人，占样本总人数的10.68%；不缴存住房公积金的农民工有38890人，占样本总人数的89.32%。

为了更为直观地呈现农民工住房公积金缴存情况与留城意愿的关系，本文绘制了柱状图。如图1所示，缴存住房公积金的农民工中“打算留城”的比例为61.02%，而不缴存住房公积金的农民工中这一比例仅为41.49%。相比之下，不缴存住房公积金的农民工当中“不打算留城”和“没想好”的比例则要明显高于缴存住房公积金的农民工中相应的比例。以上描述性分析粗略地表明，相比于不缴存住房公积金的农民工，缴存住房公积金的农民工留城意愿更加强烈。

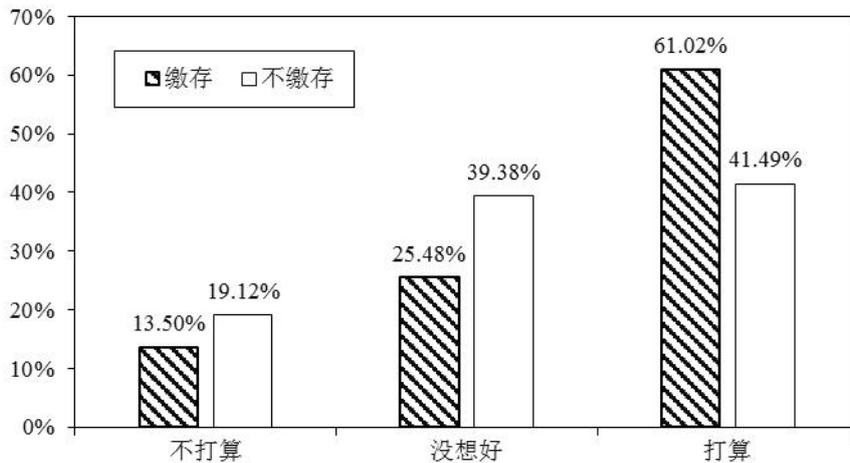


图1 住房公积金缴存与农民工留城意愿

3.控制变量。依据2014年流动人口动态监测调查数据，本文还控制了一些可能影响农民工留城意愿的变量，包括年龄、受教育程度、性别、婚姻状态、收入水平、职业(设置“是否是公务员”“是否是专业技术人员”两个虚拟变量)、单位性质、住房性质、流入时间、流动范围、是否建立健

^①删除后样本量为43541个，如无特殊说明，后文将主要运用删除后的样本进行分析。

康档案（用以衡量流入地公共服务）等。考虑到不同地区（省份）农民工住房公积金的政策措施存在差异，大部分地区对农民工基本上能做到一视同仁，但也有个别地区只是“纸上谈兵”，因此，本文以虚拟变量形式控制地区（省份）效应，以弱化地区政策不统一引致的回归分析中的“偏误”。表1列出了变量的描述性统计结果。

表1 变量的描述性统计

变量		缴存 (N=4651)		不缴存 (N=38890)		均值 t 检验
		平均值	标准差	平均值	标准差	
被解释 变量	留城意愿（以“不打算”为参照）					
	没想好	0.255	0.436	0.389	0.489	18.545***
	打算	0.610	0.488	0.420	0.493	-25.569***
核心解 释变量	住房公积金缴存情况（缴存=1，不缴存=0）	1.000	0.000	0.000	0.000	—
其他 解释 变量	年龄	27.368	3.556	26.954	3.976	-6.794***
	受教育程度（以“小学及以下”为参照）					
	初中	0.244	0.430	0.556	0.497	41.004***
	高中	0.307	0.461	0.261	0.439	-6.711***
	大学及以上	0.439	0.496	0.124	0.330	-57.791***
	性别（男=1，女=0）	0.645	0.479	0.609	0.488	-4.686***
	婚姻状态（已婚=1，其他=0）	0.599	0.490	0.565	0.496	-4.420***
	收入水平（月收入的自然对数）	8.521	0.556	8.298	0.505	-28.146***
	是否为公务员（是=1，否=0）	0.070	0.255	0.013	0.114	-26.953***
	是否为专业技术人员（是=1，否=0）	0.272	0.445	0.084	0.278	-40.148***
	单位性质（“以机关、事业单位与国有企业”为参照）					
	外资、合资企业	0.315	0.465	0.061	0.239	-60.276***
	私营企业	0.365	0.481	0.531	0.499	21.565***
	个体工商户	0.040	0.195	0.334	0.472	42.078***
	住房性质（以“租房”为参照）					
	保障性住房	0.019	0.138	0.005	0.068	-11.978***
	自有住房	0.176	0.381	0.071	0.256	-25.002***
	流入时间（年）	3.313	3.069	2.798	2.967	-11.142***
	流动范围（以“跨省流动”为参照）					
	省内跨市	0.351	0.477	0.292	0.455	-8.242***
市内跨县	0.147	0.355	0.175	0.380	4.745***	
是否建立健康档案（有=1，无=0）	0.243	0.429	0.182	0.386	-10.041***	

注：为对比缴存住房公积金与不缴存住房公积金的两组农民工样本是否存在显著差异，文中按照缴存与否分组，对被解释变量及控制变量分别进行描述性统计，并通过均值 t 检验分析两组样本的差异性。

(三) 模型设定

1. 基准模型——有序概率模型。留城意愿是离散型排序数据，本文遵循国内外相关文献中通用的处理方式，采用有序概率模型（ordered Probit，简称“Oprobit”）进行回归（参见 Clark et al., 2008）。Oprobit 模型将留城意愿视为排序变量，需使用潜变量推导出极大似然估计量。具体而言，本文估计以下模型：

$$Long_staying_i^* = \alpha + \beta HPF_i + \gamma Z_i + \delta Area_i + \varepsilon_i \quad (1)$$

(1) 式中， $Long_staying_i^*$ 为农民工留城意愿的潜变量， HPF_i （housing provident fund，HPF）表示住房公积金缴存情况， Z_i 为控制变量， $Area_i$ 表示地区虚拟变量， α 、 β 、 γ 、 δ 为待估参数， ε_i 为随机扰动项。尽管 $Long_staying_i^*$ 为不可观测的潜变量，但与农民工的留城意愿存在如下关系：

$$Long_staying_i = \begin{cases} 1, Long_staying_i^* \leq C1 \\ 2, C1 < Long_staying_i^* \leq C2 \\ 3, C2 < Long_staying_i^* \end{cases} \quad (2)$$

如 (2) 式所示， $Long_staying$ 表示农民工的留城意愿， $C1$ 、 $C2$ 为待估参数，称之为切点。当 $Long_staying_i^*$ 低于 $C1$ 时，农民工不打算留城（ $Long_staying = 1$ ），高于 $C1$ 但低于 $C2$ 时，农民工处于没想好状态（ $Long_staying = 2$ ），当 $Long_staying_i^*$ 高于 $C2$ 时，农民工打算留城（ $Long_staying = 3$ ）。进一步地，当 ε_i 服从标准正态分布时，由 (2) 式可推导出样本的似然函数，并得到极大似然估计量，即得到 Oprobit 模型。由于 Oprobit 模型估计出的系数只能从显著性和符号方面给出有限信息，因此，本文在后文分析中汇报的是各解释变量对留城意愿的边际效应。

2. 倾向得分匹配法。如表 1 中的均值 t 检验结果所示，缴存住房公积金的农民工（处理组）与不缴存住房公积金的农民工（控制组）在诸多方面存在显著差异。这意味着，缴存住房公积金很可能是农民工“自我选择”的结果，直接进行回归极有可能存在选择性偏误^①。为此，本文利用 Rosenbaum and Rubin（1983）提出的倾向得分匹配法构造反事实框架来纠正选择性偏误。具体步骤包括：

第一步，估计倾向得分值，即根据可观测到的“混淆”变量，运用 Probit 模型预测农民工缴存住房公积金的概率。第二步，使用不同匹配方法，根据倾向得分值进行匹配，以消除样本的选择性偏误。第三步，基于匹配样本，比较处理组和控制组农民工留城意愿的平均差异，得到因果关系系数，即处理组平均处理效应（average treatment effect on treated，ATT），如 (3) 式所示：

^①倾向得分匹配法主要控制可观测变量的影响，若可观测变量选择不当或过少，容易产生估计偏差。由于本文没有足够的把握证明所选取可观测变量不存在任何遗漏，因此，对于倾向得分匹配法所得结果需采取谨慎态度（陈强，2014）。

$$\begin{aligned}
 ATT &= E[(Y_{1i} - Y_{0i}) | D_i = 1] \\
 &= E\{E[(Y_{1i} - Y_{0i}) | D_i = 1], P(X_i)\} \\
 &= E\{E[(Y_{1i} | D_i = 1, P(X_i))] - E[(Y_{0i} | D_i = 0, P(X_i))] | D_i = 1\}
 \end{aligned} \tag{3}$$

(3) 式中, D_i 是二分类变量, 反映个体 i 是否进入处理组, 当 $D_i = 1$ 时, 个体进入处理组, 当 $D_i = 0$ 时, 个体进入控制组; $P(X_i)$ 为倾向得分; Y_{1i} 和 Y_{0i} 分别表示处理组和控制组的估计结果。

3. Bioprobit 模型和 CMP 方法。除了选择性偏误之外, 本文还可能面临遗漏变量、反向因果等潜在的内生性问题。首先, 一些难以衡量的遗漏变量 (如农民工的个人经历以及对城市的看法和观感等) 可能同时影响农民工住房公积金的缴存和留城意愿, 从而产生遗漏变量问题。其次, 某些农民工可能是由长期留在城市的意愿, 引致了住房公积金需求, 由此便产生了反向因果问题。诸如上述可能出现的内生性问题, 将会导致估计系数不一致。对于这类问题, 常用的解决方法是工具变量法, 但是, 留城意愿和缴存住房公积金都是离散型变量, 基于连续变量的两阶段回归等工具变量方法不再有效 (Angrist, 2001; Sajaia, 2008)。因此, 本文尝试利用 Sajaia (2008) 提出的双变量有序 Probit 模型 (Bioprobit 模型) 和 Roodman (2011) 提出的条件混合过程 (conditional mixed process, CMP) 方法开展回归分析^①。目前, 这两种方法均已得到学术界的认可与较为广泛的应用 (Russo, 2012)。

值得注意的是, Bioprobit 模型和 CMP 方法均是以似不相关回归 (seemingly unrelated regression, SUR) 为基础, 基于极大似然估计法, 通过构建递归方程组而实现的两 (多) 阶段回归模型。其基本过程为: 第一阶段, 寻找核心解释变量的工具变量, 并评估两者的相关性。第二阶段, 将工具变量代入模型进行回归, 并根据内生性检验参数判别其外生性。若内生性检验参数显著异于 0, 则模型存在内生性问题, 即 Bioprobit 模型和 CMP 方法的估计结果优于 Oprobit 模型的估计结果。相反, 若内生性检验参数不显著异于 0, 则参考 Oprobit 模型的估计结果即可。所不同的是, Bioprobit 模型是含内生解释变量的有序选择模型, 采用的是完全信息极大似然估计法; 而 CMP 方法则是可适用于多种模型、多阶段的混合过程估计。本文同时利用上述两种方法, 不仅能够比较有效地控制潜在的内生性问题, 还能够将估计结果相互对照, 以判别实证分析结果的稳健性。

四、模型估计结果与分析

(一) 基准分析

如表2所示, 本文同时采用OLS回归和Oprobit模型进行估计^②。(1)、(4) 列只控制核心解释变

^①Sajaia (2008) 和 Roodman (2011) 分别对 Bioprobit 模型和 CMP 方法的工作原理做了详细的数理推导, 因篇幅所限, 本文仅阐述其核心要点、基本步骤。在此, 感谢匿名审稿专家提出的建议。

^②Ferrer-i-Carbonell and Frijters (2004) 在研究中指出, OLS 回归和排序选择模型的估计结果在显著性和系数符号方面具有较高程度的一致性。为此, 本文在基准分析部分同时汇报了两种方法的估计结果。

量，(2)、(5)列加入控制变量，(3)、(6)列将地区虚拟变量纳入模型。通过比较可以看出，无论是将留城意愿视为连续变量的OLS回归，还是考虑留城意愿内在排序的Oprobit模型，估计结果都显示，缴存住房公积金能够显著增强农民工的留城意愿，即便加入了控制变量以及地区虚拟变量，该结论依然成立。进一步，依据(6)列结果来看，相比于不缴存住房公积金的农民工，缴存住房公积金的农民工“打算留城”的概率显著高出5.3%。

控制变量方面，大多数变量对农民工留城意愿会产生显著影响，且这些结果与过往研究基本一致(例如王桂新、胡健，2015；汪润泉、刘一伟，2017)。年龄对农民工留城意愿有显著正的向作用，女性农民工的留城意愿比男性农民工高，已婚农民工的留城意愿比其他农民工强。受教育程度方面，相比于受教育程度为小学的农民工，受教育程度为初中并未对农民工留城意愿起到显著的积极作用，而受教育程度为高中和大学的强化作用则非常显著，且从边际效应来看，受教育程度为大学的农民工留城意愿最强。就业方面，收入水平对农民工留城意愿也有显著的正向影响，即收入水平越高的农民工，其留城意愿越强。公务员、专业技术人员的留城意愿要明显强于其他职业的农民工。在机关、事业单位与国有企业就业的农民工，其留城意愿也明显强于在其他类型单位就业的农民工。流入时间越长，农民工的留城意愿越强，而且省内流动的农民工的留城意愿要明显强于跨省流动的农民工，且从边际效应来看，市内跨县农民工的留城意愿强于省内跨市的农民工。此外，拥有健康档案的农民工的留城意愿要更强，这在一定程度上展现出公共服务之于农民工的重要作用。当然，需要申明的是，考虑到控制变量可能存在潜在的内生性问题，所以，本文不对这些变量的估计结果进行过多的讨论与引申。

表2 住房公积金缴存对农民工留城意愿的影响

	OLS回归			Oprobit模型		
	(1)	(2)	(3)	(4)	(5)	(6)
住房公积金缴存情况	0.251*** (22.403)	0.081*** (6.855)	0.079*** (6.642)	0.162*** (21.578)	0.055*** (7.102)	0.053*** (6.815)
年龄	—	0.003*** (2.878)	0.004*** (3.504)	—	0.002*** (2.648)	0.002*** (3.410)
受教育程度(以“小学以下”为参照组)						
初中	—	0.030* (1.960)	0.011 (0.717)	—	0.017** (2.002)	0.007 (0.790)
高中	—	0.071*** (4.348)	0.053*** (3.223)	—	0.041*** (4.482)	0.031*** (3.367)
大学及以上	—	0.132*** (7.516)	0.108*** (6.062)	—	0.079*** (7.703)	0.065*** (6.211)
性别	—	-0.031*** (-4.499)	-0.022*** (-3.151)	—	-0.019*** (-4.785)	-0.013*** (-3.328)

住房公积金与新生代农民工留城意愿

婚姻状态	—	0.110***	0.103***	—	0.056***	0.052***
	—	(11.794)	(10.950)	—	(10.719)	(9.745)
收入水平	—	0.183***	0.166***	—	0.109***	0.099***
	—	(22.770)	(20.180)	—	(22.346)	(19.884)
是否为公务员	—	0.067***	0.063***	—	0.053***	0.051***
	—	(2.870)	(2.710)	—	(3.127)	(3.050)
是否为专业技术人员	—	0.082***	0.082***	—	0.051***	0.051***
	—	(7.335)	(7.349)	—	(7.175)	(7.150)
单位性质 (“以机关、事业单位与国有企业”为参照组)						
外资、合资单位	—	-0.091***	-0.115***	—	-0.055***	-0.071***
	—	(-5.794)	(-7.123)	—	(-5.736)	(-7.262)
私营企业	—	-0.028**	-0.041***	—	-0.019**	-0.028***
	—	(-2.342)	(-3.422)	—	(-2.429)	(-3.595)
个体工商户	—	-0.020	-0.021	—	-0.014*	-0.016**
	—	(-1.521)	(-1.608)	—	(-1.771)	(-2.011)
住房性质 (以“租房”为参照组)						
保障性住房	—	0.380***	0.382***	—	0.276***	0.281***
	—	(11.360)	(11.419)	—	(9.838)	(10.140)
自有住房	—	0.486***	0.470***	—	0.452***	0.447***
	—	(58.200)	(52.041)	—	(55.708)	(49.030)
流入时间	—	0.039***	0.038***	—	0.025***	0.024***
	—	(34.447)	(33.020)	—	(32.265)	(30.892)
流动范围 (以“跨省流动”为参照组)						
省内跨市	—	0.086***	0.097***	—	0.049***	0.054***
	—	(10.981)	(10.381)	—	(10.649)	(9.888)
市内跨县	—	0.089***	0.134***	—	0.049***	0.077***
	—	(9.433)	(11.797)	—	(8.743)	(11.254)
是否建立健康档案	—	0.070***	0.086***	—	0.041***	0.052***
	—	(8.280)	(10.053)	—	(8.040)	(9.877)
常数项	2.224***	0.348***	0.575***	—	—	—
	(588.031)	(4.904)	(7.737)	—	—	—
地区效应	不控制	不控制	已控制	不控制	不控制	已控制
R ²	0.011	0.151	0.166	—	—	—
Pseudo R ²	—	—	—	0.006	0.092	0.101
Wald卡方值	—	—	—	451.245***	5662.571***	6393.697***
样本量	43541	43541	43541	43541	43541	43541

注：括号内为经稳健标准误差校正过的 t 统计量；***、**、*分别表示在 1%、5%、10%的水平上显著。

(二) 纠正选择性偏误——倾向得分匹配法

本文首先对缴存住房公积金的农民工与不缴存住房公积金的农民工进行倾向得分值匹配，结合一系列影响农民工留城意愿的因素建立 Probit 模型^①，并根据模型结果估计出的倾向得分值进行匹配。图 2 展示了匹配前后处理组与控制组倾向得分值的概率分布情况。如图所示，匹配前两组样本差异极为显著，匹配后两组样本的差异明显减弱。

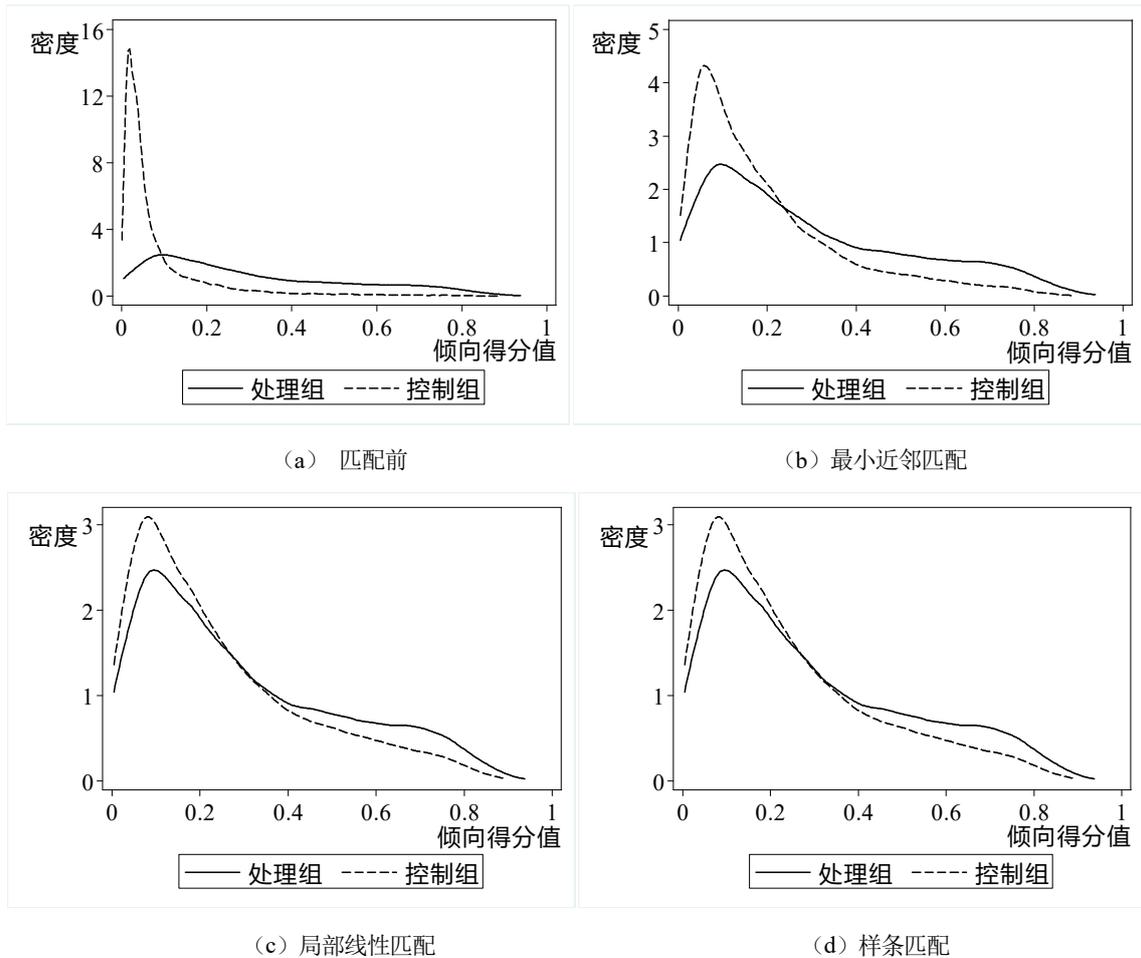


图 2 匹配前后处理组与控制组的倾向得分值概率分布

接下来，本文测算了匹配后缴存住房公积金和不缴存住房公积金两组样本的 ATT。如表 3 所示，无论是采用最小近邻匹配，还是局部线性匹配、半径匹配，亦或是采用自助抽样法（bootstrapping）迭代 500 次的样条匹配，ATT 的测算结果都显示，在消除了样本间可观测的系统性差异后，缴存住房公积金对农民工的留城意愿具有显著正向影响。

^①本文也尝试建立 Logit 模型并对解释变量进行微调，结果并无明显差异，篇幅所限，本文并未汇报这些结果。

进一步看，最小近邻匹配、局部线性匹配、半径匹配、样条匹配后测算出来的 ATT 均在 1% 的统计水平上显著。其中，局部线性匹配与样条匹配后得出的 ATT 数值最大，为 0.081，半径匹配后得出的 ATT 为 0.076，最小近邻匹配后得出的 ATT 相对较小，分别为 0.073（1:1 匹配）和 0.064（1:4 匹配）。虽然在不同匹配方法之下，ATT 的显著性与数值略有差异，但这已经足以证明，对于农民工而言，缴存住房公积金有助于提升其留城意愿。

表 3 不同倾向得分匹配的结果

匹配方法	处理组	匹配成功的控制组	ATT	标准误	t 统计量
最小近邻匹配 (1:1)	2.474	2.401	0.073***	0.018	4.04
最小近邻匹配 (1:4)	2.474	2.410	0.064***	0.015	4.34
局部线性匹配	2.474	2.394	0.081***	0.018	4.46
半径匹配	2.474	2.398	0.076***	0.013	5.61
样条匹配	—	—	0.081***	0.011	7.42

注：最小近邻匹配均采用有放回的方式；半径匹配中，半径选取 0.005；***、**、* 分别表示在 1%、5%、10% 的水平上显著。

（三）内生性讨论——Bioprobit 模型和 CMP 方法

上文分析初步表明，缴存住房公积金对农民工留城意愿有显著的正向效应，但考虑到潜在的内生性问题，本文采用地区（省份）层面农民工住房公积金的平均缴存水平（简称“地区缴存水平”）作为工具变量。从理论上讲，地区缴存水平满足工具变量的相关性与外生性条件。一方面，地区缴存水平是辖区范围内所有农民工住房公积金缴存的总体情况，与个体农民工的缴存情况息息相关；另一方面，地区缴存水平并不会直接影响个体农民工的留城意愿，所以也满足外生性条件。

根据表 4 可知，Bioprobit 模型和 CMP 方法的第一阶段回归结果均显示，地区缴存水平对农民工缴存住房公积金的影响在 1% 的统计水平上具有统计显著性，的确满足工具变量相关性条件。进一步来看，Bioprobit 模型第二阶段回归结果显示，在控制了可能的内生性偏误后，缴存住房公积金对农民工留城意愿在 10% 的统计水平上有显著的正向影响。采用 CMP 方法呈现出相类似的估计结果，再次证实了缴存住房公积金有助于提高农民工留城意愿的结论。本文还发现，采用 Bioprobit 模型和 CMP 方法时得出的内生性检验参数无法拒绝住房公积金缴存为外生变量的假设。这表明，本文的基准分析并不存在严重的内生性问题，同时表明，缴存住房公积金有助于提高农民工留城意愿的结论是真实可信的。

表 4 住房公积金缴存对农民工留城意愿的影响：基于 Bioprobit 模型和 CMP 方法

	Bioprobit 模型		CMP 方法	
	第一阶段	第二阶段	第一阶段	第二阶段

住房公积金与新生代农民工留城意愿

住房公积金缴存情况	—	0.103*	—	0.273**
	—	(1.931)	—	(1.968)
地区缴存水平	3.992***	—	1.000***	—
	(21.117)	—	(33.523)	—
控制变量	已控制	已控制	不控制	已控制
地区效应	不控制 ^a	已控制	不控制	已控制
athrho	0.030	—	—	—
	(1.111)	—	—	—
atanrho_12	—	—	-0.036	—
	—	—	(-0.542)	—
Wald 卡方值	5696.028***		7631.84***	
样本量	43541		43541	

注：括号内为经稳健标准误差校正过的 t 统计量；***、**、* 分别表示在 1%、5%、10% 的水平上显著。

^a 值得注意的是，本文选取的工具变量是地区（省份）缴存水平，将地区（省份）虚拟变量纳入模型，可能导致多重共线性问题。为避免上述问题，此处将地区（省份）虚拟变量调整为包括东部、中部和西部三大地区的虚拟变量。

（四）扩展性分析

前文分析已经证实，缴存住房公积金对农民工留城意愿有显著的正向影响。不过，以上结论只是全样本的平均效应，并未考虑住房公积金对农民工留城意愿影响的群体异质性。接下来，本文将从年龄、流入时间和收入水平三个维度对农民工群体进行分组，以期得到更为细致、深入的研究结论。其中，年龄方面划分为“80后”和“90后”两组；流入时间方面划分为“5年以下”和“5年及以上”两组；收入水平方面则以平均值作为分割点，划分为“高收入组”和“低收入组”。表5展示了基于 Oprobit 模型的估计结果。总体来看，无论是按年龄分组、流入时间分组还是按收入水平分组，缴存住房公积金对农民工的留城意愿存在正向影响，且结果均在 1% 的统计水平上显著。

1. 按年龄分组。从表 5 中（1）、（2）列可以看出，缴存住房公积金对“80后”留城意愿的影响要比对“90后”的影响更为明显。缴存住房公积金会使“80后”打算在城市长期居留的概率显著提升 5.9%，而对于“90后”而言，这一提升幅度仅为 3.0%。对这一结果相对合理的解释是，“90后”往往刚进入职场，处于适应期与学习期，职业发展仍然存在较大可塑性与不确定性，未来规划可能尚不明朗，而多数“90后”尚未结婚，对于未来将在何处定居仍未考虑成熟。因此，住房公积金缴存对其留城意愿的影响偏弱。相比之下，“80后”农民工则有较大不同。他们往往已经拥有相对丰富的工作经验，对自身职业发展有了明确规划，而且已经步入婚姻或即将步入婚姻，对于未来定居于何处已经有了长久的考虑。对于他们而言，缴存住房公积金可能就是其深思熟虑之后所做的决策。

2. 按流入时间分组。对于流入时间 5 年及以上的农民工而言，缴存住房公积金会使“打算留城”的概率显著提升 6.4%；而对于流入时间低于 5 年的农民工，缴存住房公积金会使“打算留城”的概

率显著提升 4.7%。这意味着缴存住房公积金对流入时间 5 年及以上的农民工的留城意愿影响更大。

《国务院关于深入推进新型城镇化建设的若干意见》(国发[2016]8 号)^①明确指出,要优先解决在城镇就业居住满 5 年的农民工群体的落户问题。一方面,居住满 5 年大多意味着农民工已经在此“沉淀”下来,无论是在就业方面还是在日常生活、社会交往等方面,农民工都已经比较充分的适应。另一方面,就业居住满 5 年,说明农民工对此处已经有比较浓厚的“感情”,希望能够有朝一日融入此地、扎根于此。再者说,就业居住满 5 年,也意味着农民工已经为城市发展做出了应有的贡献。

3.按收入水平分组。对于高收入群体而言,缴存住房公积金会使“打算留城”的概率显著提升 6.3%,而对于低收入群体而言,这一提升幅度仅为 3.6%。这一发现与过往研究结论基本一致,从侧面印证了当前住房公积金存在“逆向补贴”“劫贫济富”的特征(周先奎,2012;祝仲坤,2016)。住房公积金作为中国重要的住房保障制度之一,其设立的初衷是增强居民的住房购买能力,缓解居民的住房难问题。但是,随着时间的推移,尤其是近年来房价不断高企,对于低收入群体而言,即便拥有住房公积金,他们往往也无力在城市购买住房。这不仅导致其“近在眼前”的现金利益受损,长此以往还可能弱化其融入城市的意愿。而对于高收入群体而言,住房公积金恰恰为其购买住房提供了相对“廉价”的贷款,成为部分高收入农民工在城市“生根立足”的“桥梁”。

表 5 扩展性分析结果

	按年龄分组		按流入时间分组		按收入水平分组	
	“90后”	“80后”	5年及以上	5年以下	高收入	低收入
	(1)	(2)	(3)	(4)	(5)	(6)
住房公积金缴存情况	0.030** (2.130)	0.059*** (6.410)	0.064*** (3.706)	0.047*** (5.364)	0.063*** (6.167)	0.036*** (3.094)
控制变量	已控制	已控制	已控制	已控制	已控制	已控制
地区效应	已控制	已控制	已控制	已控制	已控制	已控制
Pseudo R ²	0.055	0.108	0.107	0.080	0.107	0.052
Wald卡方值	1224.572***	4613.100***	1253.477***	4188.390***	3398.698***	1619.865***
样本量	13001	30540	9069	34472	23956	19585

注:括号内为经稳健标准误差校正过的 t 统计量;***、**、*分别代表在 1%、5%、10%的水平上显著。

(五) 稳健性检验

为了检验估计结果的稳健性,本文对实证分析过程进行了一系列调整。首先,本文将 Oprobit 模型替换为 Ologit 模型重新回归。其次,本文参考郭云南、王春飞(2016)的做法,利用匹配后的样本进行回归,具体而言包括两种方式:第一种建立在处理组与控制组共同支持的区间上,第二种建立在处理组与控制组成功匹配的样本上。再者,本文尝试将留城意愿中“不打算”和“没想好”

^①资料来源: http://www.gov.cn/gongbao/2016-02/29/content_5045946.htm。

两个选项合并起来，与“打算”组成二分类变量，并利用 Probit 模型和 Logit 模型进行回归。如表 6 所示，多种稳健性检验的结果均支持住房公积金缴存对留城意愿的正向效应。

表 6 稳健性检验结果

	替换模型	样条匹配后回归 ^a		变量调整	
	Ologit模型	共同区间	成功匹配	Logit模型	Probit模型
	(1)	(2)	(3)	(4)	(5)
住房公积金缴存情况	0.059*** (7.459)	0.156*** (6.757)	0.157*** (6.798)	0.063*** (7.973)	0.062*** (7.779)
控制变量	已控制	已控制	已控制	已控制	已控制
地区效应	已控制	已控制	已控制	已控制	已控制
Pseudo R ²	0.104	0.101	0.101	0.163	0.163
Wald卡方值	6173.607***	6369.74***	6393.70***	5945.554***	6684.127***
样本量	43541	43293	43541	43541	43541

注：括号内为经稳健标准误校正过的 t 统计量；***、**、*分别表示在 1%、5%、10%的水平上显著。

^a本文尝试利用了多种方法匹配后的样本进行回归，此处以匹配效果较好的样条匹配为例证。

五、结论与启示

本文利用 2014 年流动人口动态监测的调查数据，系统考察了住房公积金缴存对新生代农民工留城意愿的影响。研究表明，缴存住房公积金会使农民工的留城意愿显著提升 5.3%，这一结果在加入控制变量、地区虚拟变量后依然成立。进一步，本文利用倾向得分匹配法纠正选择性偏误、利用 Bioprobit 模型和 CMP 方法控制潜在的内生性问题，所得结论依然稳健。分样本估计结果表明，相比于“90 后”农民工，缴存住房公积金对“80 后”农民工留城意愿影响更为明显；同时，缴存住房公积金对于流入当地 5 年及以上农民工的留城意愿影响要更大。此外，住房公积金呈现出鲜明的“嫌贫爱富”特征，即缴存住房公积金对高收入农民工留城意愿的正向效应更为明显。住房公积金通过提供“廉价”的贷款，能够增强农民工住房消费能力，理应成为造福农民工，助力其在城市“生根立足”的重要抓手。为此，应当充分发挥住房公积金在推进农民工市民化进程中的作用。具体包括：

第一，要增强住房公积金制度的约束力，着力提升政策执行力。在后续《住房公积金管理条例》的修订、完善过程中，应当明确将农民工纳入住房公积金保障体系，摆脱当前各地区各行其是的情况，消除农民工缴存住房公积金的或然性，使该项制度真正落地生效。

第二，扭转住房公积金“嫌贫爱富”的尴尬局面，避免住房公积金成为“逆向补贴”。可尝试扩大住房公积金的使用范围。农民工收入偏低、流动性较强，多数农民工在城市短期内难以购房，租房往往是他们居住的首选。因此，可尝试调整现行《住房公积金管理条例》的规定，将租房全面纳

入住房公积金使用范围^①，即只要农民工缴存住房公积金，就可将其用于租房支出。同时，力争使住房公积金在农民工家庭遭遇重大变故，或者面临子女升学等重大问题时有效发挥作用。

第三，不仅要“五年”作为基准线，在住房公积金缴存方面“开绿灯”、做到“应保尽保”，还要以“五年”作为落户的重要依据，优先解决就业居住满“五年”农民工群体的落户问题。对于实施“积分入户”的地区，可尝试将缴存住房公积金纳入积分范畴，同时大幅提高居住年限以及缴纳社保年限的积分权重。

第四，要将住房公积金制度作为推进农民工市民化的长久之计来实施。住房公积金并非短期内能够立竿见影的一剂“猛药”，但作为政府、企业与农民工成本共担、有效互动的重要举措，有助于从根本上提升农民工的收入水平，进而提升其购房能力，是推进他们摆脱“半城市化”、真正实现在城市“安居乐业”的长久之计。

参考文献

- 1.陈杰，2010：《中国住房公积金的制度困境与改革出路分析》，《公共行政评论》第3期。
- 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.陈锡文，2013：《当前我国农村改革发展面临的几个重大问题》，《农业经济问题》第1期。
- 4.顾澄龙、周应恒、严斌剑，2016：《住房公积金制度、房价与住房福利》，《经济学（季刊）》第1期。
- 5.郭云南、王春飞，2016：《新型农村合作医疗保险与自主创业》，《经济学（季刊）》第3期。
- 6.黄燕芬、李怡达，2017：《关于我国住房公积金制度改革顶层设计的探讨》，《国家行政学院学报》第2期。
- 7.李培林、田丰，2011：《中国新生代农民工：社会态度和行为选择》，《社会》第3期。
- 8.刘丽巍，2013：《我国住房公积金制度的现实挑战和发展方向》，《宏观经济研究》第11期。
- 9.汪润泉、刘一伟，2017：《住房公积金能留住进城流动人口吗？——基于户籍差异视角的比较分析》，《人口与经济》第1期。
- 10.王桂新、胡健，2015：《上海市农民工社会保障与市民化意愿》，《人口学刊》第6期。
- 11.郑小晴、胡章林，2008：《将农民工纳入住房公积金制度保障体系的探讨》，《重庆大学学报（社会科学版）》第6期。
- 12.周京奎，2012：《公积金约束、家庭类型与住宅特征需求——来自中国的经验分析》，《金融研究》第7期。
- 13.祝仲坤，2016：《农民工住房公积金制度的“困境摆脱”》，《改革》第7期。
- 14.Angrist, J. D., 2001, “Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice”, *Journal of Business & Economic Statistics*, 19(1): 2-16.
- 15.Chen, J., and L. Deng, 2014, “Financing Affordable Housing through Compulsory Saving: The Two-decade Experience of Housing Provident Fund in China”, *Housing Studies*, 29(7): 937-958.

^①根据2002年修订的《住房公积金管理条例》（国发[2002]350号）第四章第二十条第六款，若职工房租超出家庭工资收入的规定比例，职工可提取住房公积金账户内的存储余额用于交房租。

16. Clark, A. E., P. Frijters, and M. A. Shields, 2008, “Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles”, *Journal of Economic Literature*, 46(1): 95-144.
17. Ferrer-I-Carbonell, A., and P. Frijters, 2004, “How Important is Methodology for the Estimates of the Determinants of Happiness?”, *Economic Journal*, 114(497): 641–659.
18. Roodman, D., 2011, “Fitting Fully Observed Recursive Mixed-process Model with CMP”, *The Stata Journal*, 11(2): 159-206.
19. Rosenbaum, P. R., and D. B. Rubin, 1983, “Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome”, *Journal of the Royal Statistical Society*, 45(2): 212-218.
20. Russo, G., 2012, “Job and Life Satisfaction among Part-time and Full-time Workers: The “Identity” Approach”, *Review of Social Economy*, 70(3): 315-343.
21. Sajaia, Z., 2008, “Maximum Likelihood Estimation of a Bivariate Ordered Probit Model: Implementation and Monte Carlo Simulations”, *The Stata Journal*, 8(2): 1-18.
22. Tang, M., and N. E. Coulson, 2016, “The Impact of China's Housing Provident Fund on Homeownership, Housing Consumption and Housing Investment”, *Regional Science & Urban Economics*, 46(6): 25-37.

（作者单位：中国人民大学农业与农村发展学院）

（责任编辑：午 言）

Housing Provident Fund and Migrant Workers' Willingness to Stay in Cities: An Empirical Analysis with a Dynamic Monitoring Survey of Floating Population

Zhu Zhongkun

Abstract: Using the dynamic monitoring survey data of floating population in 2014, this article systematically examines the impact of housing provident fund on the new generation of migrant workers' willingness to stay in cities. The study shows that housing provident fund significantly increases migrant workers' willingness by 5.3%. This result is still valid when control variables and regional dummy variables are added. Furthermore, the study corrects the possible selection bias by using the propensity score matching method, controls the potential endogenous problem with a Bioprobit model and the CMP method, and its conclusion remains robust. The sub sample results also show that housing provident fund has more obvious influence on the willingness of the post-80s migrant workers and those who stay in cities for more than 5 years. Besides, housing provident fund shows the characteristics of “despising the poor and currying favour with the rich”, because its positive effect on the willingness of high-income migrant workers is far greater than that of the low-income.

Key Words: Housing Provident Fund; Willingness to Stay; New Generation of Migrant Worker; Endogeneity