

户口、“农转非”与中国城市居民中的收入不平等*

郑冰岛 吴晓刚

提要:导致城乡分割的户口制度对理解当代中国社会的分层和流动有着重要的作用,对中国城市化过程中的户籍制度改革亦具有政策启示意义。本文将中国快速城市化过程中产生的大量“农转非”的城市居民纳入研究视野,利用中国综合社会调查数据(CGSS 2003、2006、2008),考察非农户口对收入的影响效应。研究发现,户口转换所带来的收入优势仅限于通过高度选择性渠道实现“农转非”的群体。倾向分值匹配方法的分析更进一步显示,户口转换带来的收入效应仅限于在国有部门内教育程度和职业地位较高的群体。这些结果共同展示了户口身份对收入影响式微的过程,在市场部门内则尤其如是。

关键词: 户口 收入不平等 倾向分值匹配方法 异质性处理效应模型

户籍制度导致的中国城乡二元结构及其社会不平等长期受到学术界和媒体的广泛关注(Cheng & Selden, 1994)。自1990年代以来,随着农村劳动力跨地区迁移的迅速增长,特别是从内陆农村地区流向发达的沿海城镇地区,由户籍制度这堵“无形之墙”所造成的城市劳动力市场的社会分割和就业歧视,引起了更多的讨论。户口所代表的身份界限不仅影响农民工自身,也不可避免地通过教育机会的获得等结构性因素影响其子女的命运(Liang & Chen, 2007)。以往的研究将农民工相对于城镇居民的不利地位直接归咎于其农业户口身份(Chan, 1994; Cheng & Selden, 1994; Chan, 1996; Solinger, 1999; Wang et al.,

* 本文所使用的数据来自“中国综合社会调查”项目一期(2003-2008),该项目由中国人民大学社会学系与香港科技大学社会科学部联合执行。作者感谢香港研究资助局优配基金对吴晓刚的资助(项目编号:644208)。本文的英文版本曾于2010年在“国际社会学学会社会分层与流动研究委员会春季会议”上宣读,作者感谢与会的美国密歇根大学社会学系谢宇教授、加州大学洛杉矶分校社会学系唐启明(Donald J. Treiman)教授、香港科技大学社会科学部叶华博士的建议和北卡罗莱纳大学社会工作学院郭中阳教授对本文方法的指导,以及匿名评审人的宝贵意见。文责自负。

2002; Wang, 2004, 2005)。这种观点看似合理,却缺乏实证数据的支持,也忽视了户口以外的移民经历、成长环境、文化融合等因素。西方大量的移民研究文献揭示,即使在没有户口分野的社会,从农村到城市的移民在新的居住地也往往处于职业市场、收入,以及心理状态上的弱势地位(Bhugra, 2004; Evans & Kelley, 1991);此外,成长环境的城乡差别也对个体成年后的生活际遇打上深深的烙印(Zhang & Treiman, 2012)。

要真正理解户口对农村移民的影响,我们不妨以一种反事实(Counterfactual)的方式来思考这个问题(Holland, 1986),即有了城市户口又怎样?城市户口对于城镇出身的居民而言是一种先赋性身份,而对于那些出身农村的人来说,却需要通过各种途径获得,即通常所说的“从农业户口转为非农业户口”(简称“农转非”)。基于中国综合社会调查(Chinese General Social Survey, CGSS)2003、2006、2008年数据,本文聚焦城市居民中的农村出身人口,研究在移民经历与成长环境相对同质化的前提下,其“农转非”过程对收入不平等可能产生的因果效应。首先,我们将不同的户口转换机制归纳成选择性“农转非”与政策性“农转非”,比较他们与出身城市的居民,以及出身农村但迁移到城市却没有获得城市户口的农民工之间的收入差异。进而,我们再利用倾向分值匹配(Propensity Score Matching Methods)方法,估计非农户口(获得非农户口倾向)对收入所带来的因果效应,以及这些因果效应如何随着就业部门和获得非农户口倾向的变化而变化(Dehejia & Wahba, 1998, 2002),从而展示人们如何获得户口的过程及其对社会分层的影响。我们发现,最近有一项类似的研究得出了与我们不同的结论(魏万青, 2012)。我们将自己的研究发现与该研究的结果进行对照,藉此希望推动社会科学中的复制性研究和学术批评的发展(陈云松、吴晓刚, 2012)。

一、户口、“农转非”的不同机制及其社会后果

中国的户籍制度自1950年代建立以来,就作为一个重要的标准决定着资源和生活机会的分配,从而成为刻画中国社会结构不平等的最重要参数。通过工作单位制度(Bian, 1994),国家向城市居民(非农业人口)提供一系列社会福利,而占全国人口大多数的农民及其子女却

不被这些社会福利所覆盖,成为再分配体制下的二等公民(Chan, 1994, 1996; Cheng & Selden, 1994)。正因为户口身份所代表的生活际遇的巨大差别,要实现“农转非”这一社会流动过程是非常困难的(Fan, 2001; Wu & Treiman, 2004)。^①一方面,为了控制城市人口的增长,国家对“农转非”实施严格的名额控制,每年的户口变动率约在 1.5 - 2.0‰之间,即使是在改革开放年代亦如此(陆益龙, 2003: 144 - 146)。另一方面,被中专或大专以上学校录取的学生可以转为城市户口,且这些名额并不算在国家的配额之内。实际上,“农转非”通常是通过升学或招工的渠道达成;而获得党员、干部或军人身份,虽然在制度上并不能保证获得非农户口,但却会大大增加招工、转干,进而被城市劳动部门录用而享有“农转非”的机会。

由于上述原因,在严格的户籍管理制度下,“农转非”这一过程非常具有选择性——通常只有农村中能力最强的人才有机会。吴晓刚和唐启明(Wu & Treiman, 2004)利用 1996 年的人口抽样调查数据第一次分析了“农转非”现象,展示了这一过程的高度选择性,说明教育获得(中专和大学)和政治身份(党员和参军经历)这两个因素在其中的关键作用。后来的一项研究进一步表明,直到 1980 年代,由于城乡迁移和户口流动机会的极度稀缺性,那些基于教育程度或其他方面的优势而成功地从农村户口转变为城市户口的人,经历了一个长距离的向上流动过程,其社会经济地位和职业获得往往要高于出身城市的一般居民。不了解由于户口制度障碍导致的这一具有高度选择性的“极度向上流动”,往往会导致对中国城市社会的流动模式的误读(吴晓刚, 2007)。这类“农转非”可谓遵循着精英选择的机制,我们因而称之为选择性“农转非”,以凸显该群体在能力禀赋上与其他群体的差异。这类“农转非”从户籍制度建立之初一直存在至今(Chan & Buckingham, 2008)。

肇始于 1978 年的经济改革极大地冲击了户籍制度。自 1990 年代以来,大量农民由内陆农村涌入沿海城市,但是由于户籍制度改革滞后,他们无法获得当地城镇户口,从而被称为“流动人口”(Liang & Ma, 2004)。相对于户口迁移者来说,成为城市“流动人口”的门槛较低,因为他们通常从事当地城镇居民(包括出身农村的户口迁移者)不愿从事

^① 吴晓刚和唐启明(Wu & Treiman, 2004: 367)报告说,只有 11% 的原来是农村户口的受访者成功地获得了城市户口,其中通过教育这一途径的约占所有“农转非”人口的一半。

的职业。据 2005 年全国人口抽样调查资料显示,我国的“流动人口”数量已高达 1 亿 4 千 7 百万,占全国总人口的 11% 以上(国家统计局, 2006)。他们大部分没有当地(非农)户口,无法获得政府津贴、福利以及较好的工作机会(蔡昉等, 2002; 李春玲, 2006; Solinger, 1999; Liang, 2004)。对于这部分没有非农户口的城镇居民,我们可以将其视为户口转换(“农转非”)的失败者。

“流动人口”的迅猛增长,使得在 1980 年至 2010 年间,中国的城市化率由不到 20% 激增至近 50%。不少城市出台了法规,给流动人口办理本地户口留下了空间。随着户籍政策的放宽,除了以往很少一部分出身农村的“精英”可以通过升学、招工、参军、转干等方式获得非农户口(Fan, 2001, 2002; Wu & Treiman, 2004)之外,“农转非”的渠道变得多元化起来(Chan & Zhang, 1999),“流动人口”中的一小部分通过各种途径获得了当地户口并定居下来。例如深圳从改革开放前的一个边陲小镇,经过短短 30 年的发展,至 2010 年成为一个常住人口超过千万、没有农村的国际性大都市,户籍人口也超过 250 万,其中不少来自于农村(国家统计局, 2011)。

快速城市化进程中的经济建设与公共事业发展常常需要征用原归农民集体所有的土地,从而导致土地所有权和使用方式的变更,造成农民的劳动力剩余。为了解决征地农民的就业与生计问题,除经济补偿外,在适当情况下政府会通过政策性“农转非”给予符合条件的征地农民以非农业户口,并以转拨补助等方式帮助他们进入征地单位或其他集体所有制单位就业(国务院, 1982)。除城建征地外,一些大型工程,如三峡水利工程的部分移民安置工作也通过政策性“农转非”完成。与选择性“农转非”过程中个体的主动争取不同,此类“农转非”是农民被动接受补偿性政策而实现的,我们称之为政策性“农转非”。

总之,随着户籍制度的调整,非农户口获得机制变得多元化。非农户口对于通过升学、招工、参军、转干等高度选择性的方式实现“农转非”的个体来说是内生性的;其间,在竞争中胜出的佼佼者,其突出的人力资本和政治资本,以及其他未被观察到的禀赋特征,会在城市劳动力市场中给其带来显著的职业优势与收入回报。而通过政策性渠道实现“农转非”者,则很少具有就国家征地计划和移民安置政策与政府“讨价还价”的话语权,他们基本上是独立于农民个体能力和自我选择而被整体关照的,且大多不具备人力资本上的优势,往往需要依靠国家

安排才能从事一些非技术类的职业,或在体制外自谋生路。政策性“农转非”暗示政府政策安排对农民户籍身份的外生性影响。研究这类“农转非”对个体生活境遇的影响,有助于深入理解户籍制度与社会分层的因果关系及其演变。

综上所述,本文关注的焦点是,非农户口获得即“农转非”到底对城市劳动力市场中的收入分层有无因果处理效应(Treatment Effect)?如果有,该效应是否因人而异?户口流动的过程对社会分层又存在着怎样的影响?例如,尽管同属于农村出身的非农户籍人口,选择性与政策性“农转非”群体之间仍然存在着巨大的差异。此外,由于户口是再分配制度下的产物,因而针对农民工的体制性歧视,在再分配特征明显的国有部门越发严重(吴晓刚、宋曦,2010);而市场化却打破了这种体制性障碍,其可能的原因是私有部门的雇主更看重职工的工作技能、效率和个人努力。因此,我们的分析也将考虑到中国城市劳动力市场中所有制的二元结构(Bian,1994;Wu,2002),考察户籍转换效应在国有部门和私有部门之间的差异。

二、数据、变量与分析策略

(一)数据

本文采用CGSS 2003、2006及2008年的调查数据。作为一项连续性的数据搜集项目,CGSS调查采用多阶分层随机抽样方法,广泛了解中国居民的教育、就业、迁移、社会关系、生活方式等信息。调查迄今为止已发布四轮数据(2003、2005、2006、2008)。由于历次调查基本使用相同的抽样框,从而形成有代表性的全国样本,而仅有2005年数据无法确定居民的户口出身,因而我们将其他三轮数据合并,以获取足够的农村出身人口,从而提高研究结果的统计效力(statistical power)。由于在实际操作过程中,各组数据的选择性偏差不可避免,因而在使用合并样本进行分析时,我们控制关于各组数据的年份虚拟变量,以消除数据特殊性所带来的估计误差。由于研究侧重的不同,我们假定2003年到2008年间的中国城市劳动力市场状况是基本平稳的,忽略各组数据之间的历时比较。

除2003年数据仅包含城市样本外,其他两组数据皆为全国性的城乡样本。我们仅保留各组数据中的城市样本,并选取18至60岁的居

民。表 1 显示,中国城市居民样本中仅有约 15% 仍然保持农业户口,他们大部分是进城农民工;而在约占城市居民总数 85% 的非农户口群体中,约 30% (25.48/84.61) 出身农村。由于本文关注的重点是“农转非”的效应,因而我们以农业户口出身的人群为主要研究对象,占城市居民的 41% 左右(见表 1 第 1 行),比较其中的农业户口群体与非农户口群体间的收入差别。相比于以往只从“户口现状”的单一维度进行的比较研究,我们根据“户口出身”界定研究对象的策略从某种程度上控制了移民经历、成长环境、家庭背景、文化融合等多种因素,使得两类户口群体之间更具可比性。

表 1 中国城市居民的户口出身与户口现状* (2003 - 2008)

户口出身	户口现状(%)		总体(%)
	农业户口	非农户口	
农业户口	15.51(2107)	25.38(3447)	40.89(5554)
非农户口	—	59.11(8028)	59.11(8028)
总 体	15.39(2107)	84.61(11475)	100(13582)

说明: * CGSS 调查中,对现在的农业人口并不询问其户口转换经历。我们只能假设不存在从非农户口到农业户口的流动,尽管在现实情况下这是有可能的,但却极为少见。

注:(1)样本来自于 2003、2006、2008 CGSS 数据中 18 - 60 岁的城市居民;(2)括号中为绝对值。

(二)变量

对农村出身的城市居民根据户口身份分类是本研究进行比较的关键。我们将虽在城市生活却仍保留农业户口身份的群体定义为“农业户口”群体,他们主要是城市农民工。与之相比较的是出身农村的非农户口群体;根据“农转非”机制的不同,我们又将其分为“选择性”与“政策性”农转非群体。2003 及 2008 CGSS 调查数据询问了被访者“户口迁移的原因”,本研究将其中的“升学”、“参军”、“招工”、“转干”视为“选择性”渠道,而将“征地”与“移民安置”视为“政策性”渠道。根据这些界定,两轮数据总共为我们提供了 206 个政策性“农转非”及 846 个选择性“农转非”的有效分析样本,而其他农业户口出身的被访者则未能在城市改变自己的户口身份,其有效样本数为 511 名;其中,在农业户口出身的城市居民中,约 67% 最终实现了“农转非”。由于 2006 年的数据中不包含“农转非”原因这一关键变量,我们在区分选择性“农转非”和政策性“农转非”的分析中,仅使用 2003 及 2008 年的数据。

我们用收入来作为社会分层的测量指标。调查分别搜集了受访者在2002、2005、2007年的“个人全年总收入”，它包括职业性收入，如工资、奖金、其他经营性收入和资本性收入。为了使三轮数据搜集的居民收入信息更具可比性，我们根据城市居民购买力指数对收入进行标准化，使其具有与2002年居民收入同等的购买力意义。^①

个人收入决定模型中一般包括以下预测变量：教育水平、工作年限、性别等。其中教育是最为重要的人力资本指标之一，在研究中我们以一个三分类变量来描述城市居民的受教育程度：初中及以下、高中及同等学历（职高、技校）、大专及以上。党员身份测量的是被访者的政治资本，它是一个虚拟变量（党员为1，非党员为0）。对于农村出身者，参军也是重要的一条向上流动的渠道，也代表着一种政治身份，它也是一个虚拟变量（曾经参军为1，未曾参军为0）；同样是虚拟变量的还有性别（男性为1）、婚姻状况（已婚为1）。工作年限和每周工作时间则为连续变量。

被访者的职业和工作单位直接影响他们在城市劳动力市场的收入，因而也被我们纳入模型。我们将被访者的职业分为六大类：管理人员、专业技术人员、办事人员、商业及服务业人员、体力劳动者，以及自雇佣者。工作单位则分为国有部门及私有部门。

（三）分析策略

农村出身者实现户口流动的过程影响着我们对户籍制度的社会分层后果的估计。通过选择性“农转非”和政策性“农转非”获得的非农户口身份，在城市劳动力市场的涵义是不同的，因而对收入不平等的影响也不一样。对此观点的一般推论是：由于跨越户籍制度障碍而获得非农户口的能力不同，因而个体实现的户口转换所带来的社会分层效应也不同。

虽然我们可以将“农转非”从概念上分解为选择性“农转非”与政策性“农转非”两个群体，与没有成功“农转非”的城市农民工相比较，通过多元线性回归模型估计组间平均数的差异；但严格说来，这种二分法对户籍身份分野过程的过分简化，将不可避免地导致测量误差。例如，通过招工获得非农户口的个体也完全可能不具备人力资本优势而仅仅是凭运气；由于征地政策而被纳入城市居民的也可能包括农村中的一些精英。

^① 标准化指数为：2002 = 100；2005 = 93.87；2007 = 85.03（国家统计局，2007）。

为减少两分法所带来的测量误差,我们采用倾向分值匹配法来测量人们对自己的户口变动的选择性。这种倾向性由各种可以观测到的个体特征所决定。倾向分值匹配法假定:“农转非”倾向性较低但事实上却转了户口的人,比那些“农转非”倾向性较高的人更多地受到政策放宽的影响。利用这种方法,人们获得城市户口的选择性就可以被当作一个连续变量而非两分变量来处理。

要研究人们从农业户口向非农户口的转变这一社会过程究竟如何影响其当前的经济地位,需要将上述两种方法结合使用。根据两分法,我们可以将虚拟变量代入传统的线性回归模型来比较农民工、政策性“农转非”、选择性“农转非”,以及城市出身居民之间的收入差别。然而,利用横截面调查数据,区分选择性“农转非”与政策性“农转非”,没有从根本上解决因果效应估计中遇到的内生性问题。我们可以采用倾向分值匹配法估计“农转非”的平均因果效应。

倾向分值匹配方法先以个体发生某个事件的倾向性或可能性为标准,总结“实验组”(“农转非”)和“控制组”(“农民工”)两个群体之间在各个层次上的全部差异;然后根据倾向分值在每一个层次上对两组人进行匹配,以消弭各层次内部“实验组”与“控制组”的个体特征差异。我们由此可以在一个反事实的框架(Counterfactual Framework)下分析转变户口带来的收入效应(Holland, 1986):在“农转非”倾向分值的各个层次内部,个体特征相似的群体,由非农户口获得的不同结果而构成可比较的“实验组”与“控制组”,他们之间的收入差异即为该层次“农转非”对收入的处理效应(Treatment Effect)。该处理效应在各层次间差别则构成我们所估计的“农转非”的异质性处理效应(Heterogeneous Treatment Effect)。

为保证这种估计是无偏的,我们必须假定所观察到的一系列个体特征是穷尽的,也就是说,没有其他任何重要的协变量存在从而影响到个体发生某个事件的倾向性或可能性,这又被称为“严格可忽略性假设”(Strongly Ignorability Assumption)。“严格可忽略性假设”是倾向分值匹配方法实施的重要基础(Imbens, 2004; Rosenbaum & Rubin, 1984),它要求除了我们用于预测的个体特征之外,没有其他任何因素能够影响个体经历某一事件的倾向性。在本研究中,这个假设要求我们找到一组非常完善的可观测变量,使得没有其他任何的变量可以区分农村出身的非农人口与农业人口,同时又影响收入。

出身农村的人们获得非农户口的倾向性不同,因而他们通过各种机制实际获得的非农身份对其命运的影响也不同。为考察个体“农转非”的倾向到底是否会增加其收入回报水平,我们需要使用一个允许多种影响效应存在的多层线性模型(HLM)(Brand & Xie, 2010; Raudenbush & Bryk, 2002)。具体而言,我们引入一个两层分析模型来量化“农转非”处理效应与其倾向分值层次之间的关联,即:

$$\text{第 1 层: } y_{ij} = \delta_{0j} + \delta_{1j} \text{treat}_{ij} + \mu_{ij} \quad (1)$$

其中, δ_{1j} 表示在“农转非”倾向分值的第 j 层, 非农户口对收入的处理效应, 即该层非农人口与农业人口的平均收入差别。而非农户口的处理效应与“农转非”倾向性相关, 则:

$$\text{第 2 层: } \delta_{1j} = \gamma_{10} + \gamma_{11} \text{stratum}_j + \vartheta_{1j} \quad (2)$$

其中 stratum_j 表示处于倾向分值第 j 层, γ_{11} 表示倾向分值层级的变化对处理效应的影响, 它可以告诉我们, 非农身份在那些以不同倾向性经历农转非的人群中, 到底有什么不同的作用。 μ_{ij} 和 ϑ_{1j} 分别为个体层次和倾向得分层次上的误差项。

这里值得一提的是, 我们的分析策略和研究设计与新近发表的一篇关于户籍分层的研究有许多共同之处(魏万青, 2012, 以下简称“魏文”)。然而, 即使采用时间上非常相近的两组数据, 我们将展示的结论却与该研究不尽相同。魏文结论的可商榷之处, 在于运用倾向分值匹配法的时候, 对该方法的一个重要前提, 即“严格可忽略性假设”缺乏足够的重视, 从而使得其研究中的“永久移民”与“流动人口”实际可能并不具备可比性, 因而不构成严格的反事实分析。此外, 魏文将反映处理效应异质性的参数(γ_{11})当作非农户口的平均处理效应(δ_{1j}), 这一理解也不尽准确。

三、实证分析结果

(一) 描述性统计

表 2 是对样本的描述性分析, 并对我们关注的三类户口身份群体进行比较。我们同时给出非农户口出身群体的描述性统计作为研究的参照。从教育水平看, 政策性“农转非”群体与农业户口群体具有很强的相似性: 他们中有大专及以上学历的人属凤毛麟角(8% 与 5%)。与

此形成对比的是,选择性“农转非”群体中约 40% 的人都具有大专以上学历,该群体的教育程度甚至高于城市出身居民。在选择性“农转非”群体中,党员及参军者比例也最高,都超过了城市出身居民,也高于农业户口群体,表现出他们在政治资本上的显著优势。这也体现了他们是农村中那些能力最强、最聪明的人,是农村中的精英。

表 2 不同户籍身份的中国城市居民之描述性统计

户籍身份	农业户口	政策性“农转非”	选择性“农转非”	非农户口出身
教育				
初中及以下	64.38	66.20	29.91	38.10
高中及同等学历	27.20	28.64	30.14	37.60
大专及以上	8.41	5.34	39.95	24.30
党员	6.07	14.56	36.05	16.80
参军经历	1.76	3.88	12.53	4.72
职业				
管理人员	4.31	5.34	8.39	6.01
专业技术人员	10.96	10.68	33.92	20.86
办事人员	4.11	13.59	12.88	11.42
商业及服务人员	22.90	16.50	12.65	19.89
体力劳动者	42.07	42.72	28.72	36.94
自雇者	15.66	11.17	3.43	4.88
所有制部门				
国有部门	18.40	39.81	75.18	63.31
私有部门	81.60	60.19	24.82	36.69
男性	47.75	50	64.42	50.43
已婚	78.82	90.29	92.43	89.47
每周工作时间 (小时)	55.35 (17.35)	50.10 (15.90)	46.73 (13.23)	47.25 (12.54)
年龄	33.52 (9.36)	40.02 (10.62)	41.33 (10.77)	41.41 (10.39)
工作年限	15.83 (10.29)	21.54 (11.55)	21.69 (11.46)	22.62 (10.89)
N	511	206	846	3625

注:(1)样本来自于 2003、2008 CGSS 数据中 18-60 岁城市居民;(2)对于教育、党员、参军经历、职业、所有制部门、男性、已婚这些分类变量,我们报告其各类别的百分比,而其他连续变量则报告其均值;括号内是其标准差。

从职业和单位类型看,高度选择性的非农户口获得者极少从事体力劳动及自雇佣职业,他们半数以上为管理人员、专业技术人员及办事人员(8.39% + 33.92% + 12.88% = 55.19%),且75%就职于国有部门。相反,那些通过政策性渠道获得非农户口的人们,尽管进入国有部门的比例是农业人口的两倍(39.81%与18.40%),但他们却大体上与农业人口一样主要从事服务业与体力劳动型职业,并更多地成为自雇佣者。“选择性”农转非者成为管理人员或专业技术人员的比例甚至高于城市出身居民,从另外一个角度说明了该群体所经历的“极度向上流动”(吴晓刚,2007)。

从性别分布看,多达64%的选择性“农转非”者都是男性,而其他几组人群的性别比例则较为平均。

选择性非农户口获得者与城市出身居民的工作时间最短,平均每周工作47小时左右;而农民工群体则平均每周工作55小时。在年龄与工作年限上,农业户口群体最为年轻,工作经历也最短;而城市户口群体内部(包括两类“农转非”及城市出身居民)在这两个维度上则没有显著差异。

表3描述了不同户籍身份的中国城市居民在各所有制部门内的标准化全年总收入。总体而言,选择性非农户口获得者收入最高,他们的全年总收入达16315元,比城市出身居民还高将近2000元。有趣的是,城市中收入最低的并不是农民工,而是通过政策性渠道实现“农转非”的群体,他们的平均年收入仅为1万余元,比农民工还低3600余元,且这一差异在统计上是显著的($p = 0.043$)。若分单位类型来看,则中国城市劳动力市场的收入分层在国有部门和私有部门间存在着差异。在体制内,农业户口者收入最低,但虽然其平均收入少于政策性“农转非”群体,该差异在统计上却并不显著($p = 0.428$)。但在私有部门内,政策性“农转非”群体的收入弱势则更为明显:相比农业户口群体,他们的收入少了将近5000元,而相比选择性“农转非”群体则少将近12000元。值得注意的是,在私有部门内部,高度选择性的非农户口获得者的收入甚至高于本地城市居民6000余元。这说明相比于再分配特征更为明显的国有部门,在私有部门,农村出身的精英的知识、技能以及个人能力能够得到更大的收入回报。表3的结果提示我们,“农转非”并不一定意味着收入的优势地位。

表 3 不同户籍身份的中国城市居民在各所有制部门内的全年总收入

户籍身份	农业户口	政策性“农转非”	选择性“农转非”	非农户口出身
国有部门	8740	10337	14367	13298
	(7362)	(8756)	(14428)	(16253)
	94	82	636	2295
私有部门	15152	10379	22213	15834
	(26932)	(11224)	(36848)	(36654)
	417	124	210	1330
总 体	13972	10363	16315	14228
	(24652)	(10290)	(22445)	(25718)
	511	206	846	3625

注：(1)全年总收入已根据历年城市居民购买力指数做了标准化；(2)我们对不同户籍身份的中国城市居民分别报告其在国有部门与私有部门收入的均值，括号内是标准差，斜体数字为其分布的样本量。

(二)线性回归模型

表 4 是中国城市居民样本收入获得模型的线性回归结果，其因变量为受访者全年总收入的对数。模型 1 表明，在控制受访者的性别、婚姻状况及工作年限后，城市劳动力市场最主要的收入差别并不如同一些研究所认为的那样处于农业户口和非农户口之间。政策性“农转非”群体虽然拥有非农户口，但其收入与农业户口人群(农民工)却不存在显著差异；而选择性“农转非”群体的年收入则几乎达到农业户口群体的两倍($e.626 = 1.870$)，甚至高于非农户口出身的城市居民。

表 4 不同户籍身份的中国城市居民之收入决定模型

变量	模型 1	模型 2	模型 3	模型 4
户籍身份:农业户口(参照组)				
政策性“农转非”	-.118	-.215	-.228	-.424*
	(.205)	(.199)	(.200)	(.246)
选择性“农转非”	.626***	-.082	-.091	-.225
	(.141)	(.143)	(.146)	(.206)
非农户口出身	.243**	-.239**	-.227*	-.436***
	(.119)	(.119)	(.122)	(.139)
男性	.777***	.700***	.757***	.759***
	(.070)	(.069)	(.071)	(.071)

续表 4

变量	模型 1	模型 2	模型 3	模型 4
已婚	-.238* (.142)	-.256* (.138)	-.211 (.139)	-.201 (.139)
工作经验	-.037*** (.014)	-.006 (.014)	-.006 (.014)	-.008 (.014)
工作经验 ² × 100	.001*** (.000)	.001** (.000)	.001** (.000)	.001** (.000)
教育:初中及以下(参照组)				
高中及同等学历	.589***	.501*** (.081)	.508*** (.083)	
大专及以上		1.597*** (.101)	1.333*** (.115)	1.329*** (.115)
党员		.283*** (.098)	.220** (.101)	.216** (.101)
参军经历		-.042 (.154)	-.012 (.155)	-.013 (.155)
职业:体力劳动者(参照组)				
管理人员			.487*** (.156)	.504*** (.156)
专业技术人员			.481*** (.105)	.492*** (.105)
办事人员			.171 (.123)	.177 (.123)
商业及服务业人员			.059 (.096)	.065 (.096)
自雇佣者			-.206 (.156)	-.216 (.156)
国有部门			-.173** (.079)	-.037*** (.276)
国有部门 × 政策性“农转非”				.957** (.437)
国有部门 × 选择性“农转非”				.827** (.334)
国有部门 × 非农户口出身				.942*** (.287)
每周工作时间			-.004 (.003)	-.004 (.003)

续表 4

变量	模型 1	模型 2	模型 3	模型 4
截距	8.403 ***	7.740 ***	7.940 ***	8.100 ***
	(.157)	(.160)	(.223)	(.228)
N	5179	5179	5179	5179
R ²	.035	.093	.099	.101

注：(1)因变量为标准化后个人全年总收入取对数；(2)括号内为标准误；(3) *** p < .001, ** p < .01, * p < .05。

在模型 2 中,我们进一步控制被访者的教育程度、党员身份以及参军经历。此时,在农村出身的城市居民中间,户口身份之间的收入差别消失了。更有趣的是,在其他因素一致时,城市户口出身的居民收入竟然低于农村出身居民而无论其现在的户口身份是什么。当我们在模型 3 中加入职业、单位类型以及工作时间时,我们仍然获得与模型 2 类似的结果。上述结果说明,我们所观察到的与非农户口相关的收入优势,仅仅局限于那些通过高度选择性渠道实现户口流动的群体;而他们的收入优势又在很大程度上来自其教育资本、政治资本,以及与之关联的职业和所在单位的优势。

在模型 4 中,我们加入户口身份与单位类型的交互项,以考察城市劳动力市场中的户籍分层在体制内外是否有所不同。与表 3 类似,模型 4 的结果说明,非农户口与农业户口之间的收入差别仅在国有部门内存在。在体制内,政策性“农转非”群体的收入比农民工高约 70% ($e^{0.957-0.424} = 1.704$),而选择性“农转非”群体的收入则比农民工高 83% ($e^{0.827-0.225} = 1.826$)。非农户口出身者在体制内也具备优势,其收入比农民工多 66% 左右 ($e^{0.942-0.436} = 1.659$);而在体制外,非农户口的收入优势则不复存在,尤其是政策性“农转非”群体与城市出身居民,他们相比于农民工甚至收入更低。

表 4 的线性回归模型提示我们,非农户口并不一定带来收入上的优势,对户口的收入回报在体制内和体制外,根据非农户口获得方式的不同而呈现出很大差异。为了更加严格地考察非农户口对收入的因果效应,以及户口转换机制这一内生性过程在其中的影响,我们使用倾向分值匹配方法进一步探讨这个问题。

(三) 倾向分值匹配方法

倾向分值匹配方法的优越性在于,利用可观测信息消除因果关系推论中的估计偏误(郭申阳、马克·弗雷泽,2012)。这一方法近年来在社会科学研究领域得到了广泛的运用(Brand & Xie,2010;Harding,2003;Xie & Wu,2005)。在倾向分值匹配方法的分析框架下,对于农村出身的城市居民,我们将其中的非农户口获得者(无论通过何种机制)看作“实验组”,将农业户口保持者(农民工)看作“控制组”。我们的操作方法是:首先通过一个二元逻辑斯蒂(Binary Logistic)回归模型,将“实验组”和“控制组”之间的差异以倾向分值的形式总结出来;然后将该倾向分值分层,并结合可观测的各项变量,对“实验组”和“控制组”进行匹配;最后在“实验组”和“控制组”实现可比性的前提下,计算获得非农户口对个人收入的真正影响。

为了对“实验组”和“控制组”完成匹配,以在各倾向分值层次内部最大化地消除组间差别,我们必须首先对“实验组”和“控制组”之间的组间差异做充分了解。我们以下列自变量来估计每个个体的“农转非”倾向分值:受教育年限、党员身份、参军经历、职业、婚姻状况、工作年限、年龄、性别,以及18岁时父亲的工作单位性质(国有部门或私有部门)。表2充分显示了各个户口身份群体之间的职业差别。在最近的一项研究中,张卓妮与吴晓刚更指出,农民工相对于城镇居民的低收入主要不是源于“同工不同酬”,而是因为制度性的职业隔离(Zhang & Wu,2012),因而我们引入职业来预测倾向分值,以使“实验组”和“控制组”在职业分布上尽量均衡。同时,本研究引入父亲的工作单位性质以测量家庭背景的影响,其假设是:若被访者父亲在国有部门工作,则他更有可能通过招工顶替等方式,或利用父亲的城市户口、城市社会网络等实现户口转换(Bian,1994)。

表4中收入的线性决定模型已经显示了不同所有制部门对户口身份的不同回报,因而我们的倾向分值匹配方法也在国有部门与私有部门分别进行。使用前文所述来自三轮数据的4125个有效样本,图1描述了两组倾向分值匹配方法后的样本分布结果。在国有部门和私有部门,根据个体“农转非”的倾向分值,样本被分别归类于五个层次(Stratum)。在图1中,两组直方图的X轴分别标注了各层次倾向分值的下限。图1清晰地显示出,无论是在国有部门还是在私有部门,“实验组”总是集中在倾向分值较高的层次,而“控制组”则多聚集在倾向分

值的较低端。这形象地说明了在农村出身者中,非农户口获得者与农业户口保持者之间有多么不同。正是在这一意义上,以往的回归模型常常是在比较实际不具可比性的两组群体,因而得出有偏误的结论。

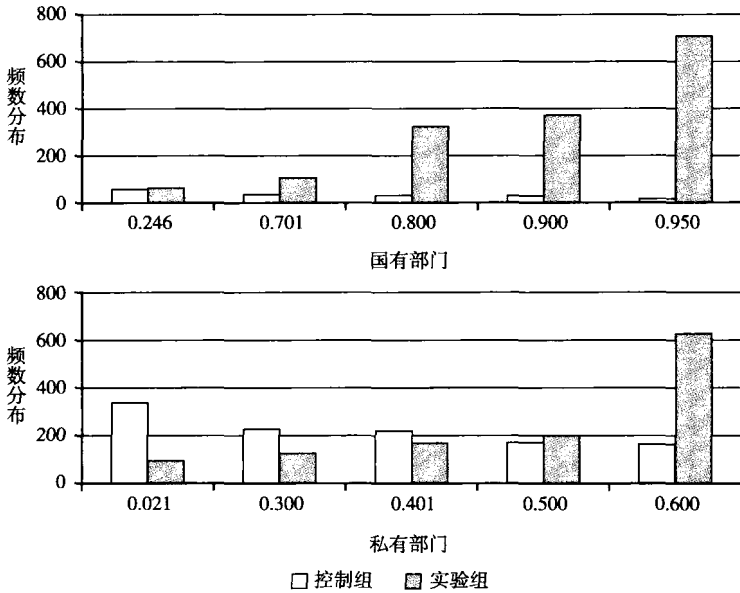


图 1 各所有制部门内农业户口群体与非农业户口群体的匹配结果

我们用每周工作时间和全年个人总收入计算被访者的每小时工作收入,以作为倾向分值匹配方法的因变量。首先,我们在倾向分值的每一个层次,估计非农户口对收入的影响效应。经过匹配,在倾向分值各层次内部,“实验组”与“控制组”之间不再存在系统性的显著差异,因而我们可以直接计算某层次内两组人口的平均收入差别,并将其理解为非农户口在该层次内对收入的因果效应。假设户口对收入的因果效应在各层次间不具有显著性差异,则通过加权平均我们可以得到非农户口对收入的平均因果效应,如表 5 所示。我们发现,非农户口对收入的因果效应只在国有部门存在,如同在体制内工作 1 小时,非农户口群体比农业户口群体平均多收入 1.656 元;但在私有部门,户口身份则并不影响收入。从总体样本的匹配结果看,我们没有发现非农户口对收入不平等的显著的因果效应。

表 5 非农户口对收入的平均处理效应

	非农户口 (T)	农业户口 (C)	平均处理效应 (T)-(C)	N(T)	N(C)	匹配率 (%)
国有部门	5. 869	4. 213	1. 656 ***	1480	170	93. 964
私有部门	6. 043	6. 702	- . 660	1077	1008	88. 012
总体样本	5. 942	6. 343	- . 401	2557	1178	90. 545

注:(1)收入为标准化后的每小时工作收入;(2) *** p < . 001。

然而,我们并不能确定户口对收入的因果效应在各倾向分值层次间没有差异。因此,接下来我们分别估计各个倾向分值层次上非农户口对收入的影响效应,并以多层线性模型拟合其变化趋势,结果如图 2 所示。

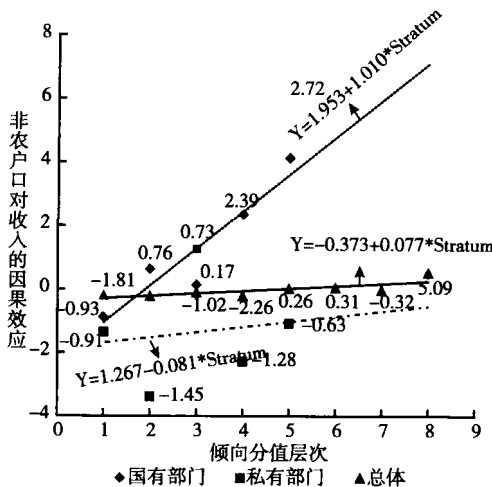


图 2 非农户口对收入的异质性处理效应模型

图 2 中的菱形点表示对国有部门样本的各层次因果效应的点估计值,而方形点则代表对私有部门样本的估计,点旁边的数字表示假设因果效应为零的统计检验值。图 2 说明,在国有部门内,户口对收入的影响效应仅在最高一级的倾向分值层次内部显著。对该倾向分值层次的样本进行描述统计我们发现,其平均受教育年限为 12 年,党员比例高达 48%,且有将近一半的人从事管理或专业技术职业。这说明,在国有部门内,户口的收入回报仅限于教育程度和职业地位很高的群体。

图2 中国有部门各层次点估计的拟合线说明,因果效应的同质性假设在国有部门内不成立(斜率为1.010,标准误0.365)。相反地,在私有部门内,户口对收入的影响在各倾向分值层次内部皆不显著,也不存在层间差异。除了分所有制部门进行分析外,我们也对总体样本进行了各个层次的点估计(总体样本匹配为8层),以三角形点表示,并进行线性拟合。总体样本的分析结果表明,户口对收入的因果效应仅在倾向分值的最高层次显著,且因果效应的大小与倾向分值呈现正相关关系(斜率为0.077,标准误为0.015,统计上显著)。因而,总的说来,户口对收入的因果效应与“农转非”倾向性正相关,户口流动倾向性越高的人,其获得非农户口后的收入回报越大。

以上分析结果是建立在“严格可忽略假设”基础上的。在附录中,我们执行了严格的罗森鲍姆敏感界限检验(Rosenbaum sensitivity check),更进一步显示了本文研究的稳健性。

四、结论和讨论

1990年代以来的大规模农业人口迁徙、快速城市化过程,以及户籍制度的相应调整,正在深刻地改变我国社会的分层格局。其中一个重要的方面是,农村出身的人获得非农户口的方式变得多样化起来,从而使得非农户口群体乃至整个城市居民群体,都呈现出比以往更强的异质性。就此,本研究总结了两种具有代表性的户口流动方式。其一是自户籍制度建立至今,农村出身的佼佼者们通过在城市升学、就业等方式获得非农户口。这是一种具有高度选择性的流动方式,只有农村中少数的精英才得以抓住机会。其二是自快速城市化进程以来,被征地农民整体获得非农户口。这是一种政策安置性的户口流动,与农民的个人能力并没有关联。

我们的研究发现,非农户口对收入的影响效应在这两种户口流动方式中有所不同。高度选择性的户口转换方式,带来“极度的向上流动”(吴晓刚,2007),其收入回报甚至高于出身城市的居民;而政策补偿式的户口转换则并不能够带来收入回报。可见,只有与个人能力相关联的、主动的选择性户口流动,才能对收入产生影响。我们的倾向分值匹配分析更进一步地说明了这一点:只有对那些户口流动倾向性很

强的人,非农户口才对其收入产生因果效应;且倾向性越强,“农转非”的收益越高。正如图3所示,由于“政策性”农转非群体的户口流动倾向大多很低,即使他们已经获得了非农户口,但却因不具备较强的个人能力和主动性也就无法得到非农户口的收入回报。

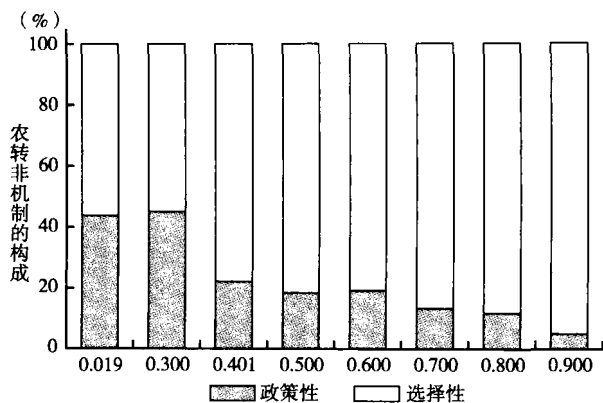


图3 各倾向分值层次上“农转非”机制的构成

以上发现说明,户口身份是否真正起作用,实际取决于户口流动机制中的选择性。以往研究仅仅比较不同户口身份群体间收入的均值差异,并将该差异直接归咎于农业户口身份,实际忽略了户口流动过程中的选择性。在研究户口与收入之间的因果机制问题时,横亘在研究者面前的先决性问题是:同样出身于农村,为什么有的人能够突破户籍制度障碍获得非农户口,而有的人不能?人们获得非农户口的机制有无差异?是否代表了不同的选择性?而这些问题在以往研究中往往被忽略了。那么,我们应当怎样理解户口身份回报与其流动过程中的选择性的关联呢?这是一个典型的内生性问题。对于那些通过高度选择性渠道获得非农户口的人,即对于那些户口流动倾向性非常强的人,我们相信其潜在的个体能力优势是其获得高收入的根本原因。而这类人在教育、入党、参军、就业等一系列流动过程中的突出表现,也从侧面说明了其个体驱动力和能力优势。这种个体差异导致他们更加容易突破体制壁垒,实现向上流动。

正是基于这一点,布莱洛克(Blalock, 1991)建议研究者们应该将理论的中心放在被他称之为“归类”过程的分析上,即关注社会个体如何通过一系列微观决定与选择将自己归入某一社会群体。如果我们仅

仅只关注分层的结果,即不同社会群体之间在教育、职业和收入上的差异,而忽略这些不平等产生的社会过程,则很有可能得出有偏或缺乏证实的结论。事实上,流动的选择性过程与其分层结果之间的理论关联,不仅适用于中国的户籍制度,也同样有助于我们对其他社会不平等结果的理解。例如,高校扩招后教育回报的降低,有可能是教育获得过程的选择性下降造成的(Ye,2012)。在户籍制度调整的今天,由于非农业户口的获得相对越来越容易,因而户口流动的选择性下降了,所以必然导致户口身份的“贬值”。

本研究的另一项发现是,非农业户口的收入效应仅限于国有部门内教育程度和职业地位较高的群体,而在更为看重工作技能以及个人努力的市场部门,户口并不对收入产生显著的因果效应。这一发现揭示了我国城市化过程中户籍制度变迁的内在过程:户口,作为再分配制度下的产物,其式微过程是从市场化部门开始的;而因户口身份的报酬歧视,其作用范围已越来越小,仅局限于体制内壁垒最严重的区域。

以往许多研究将农民工相对于城镇居民的不利地位直接归咎于其农业户口身份,其隐含的政策意义似乎是:只要改革户籍制度,给予农民工当地户口和市民待遇,则他们所面临的种种不公现象就会自然得到解决。然而我们的研究发现,在户籍制度变迁的今天,农民工的不利地位不仅仅在于其农业户口身份,更在于其社会流动过程中的竞争力弱势。因而,我们建议政策制定者们将注意力集中于改善农村人口的教育机会,并帮助他们更有效地融入城市劳动力市场。同时,鉴于“政策性”农转非人群在城市劳动力市场中的不利地位,政策制定者们应给予被征地农民更多的关注和帮扶。

参考文献:

- 蔡昉、张车伟、都阳,2002,《城乡就业问题与对策》,北京:社会科学文献出版社。
- 陈云松、吴晓刚,2012,《走向开源的社会学——定量分析中的复制性研究》,《社会》第3期。
- 国家统计局,2006,2007,2011,《中国统计年鉴》,北京:中国统计出版社。
- 郭申阳、马克·弗雷泽,2012,《倾向值分析:统计方法及应用》,重庆:重庆大学出版社。
- 国务院,1982,《国家建设征用土地条例》,北京:法律出版社。
- 李春玲,2006,《流动人口地位获得的非制度途径——流动劳动力与非流动劳动力之比较》,《社会学研究》第5期。
- 陆益龙,2003,《户籍制度:控制与社会差别》,北京:商务印书馆。
- 魏万青,2012,《户籍制度改革对流动人口收入的影响研究》,《社会学研究》第1期。

- 吴晓刚,2007,《中国的户籍制度与代际职业流动》,《社会学研究》第6期。
- 吴晓刚、宋曦,2010,《中国劳动力市场中的民族分层:对新疆维吾尔自治区(2005)的实证研究》,《中国社会科学(内刊)》第6期。
- Bhugra, D. 2004, "Migration, Distress and Cultural Identity." *British Medical Bulletin* 69(1).
- Bian, Y. 1994, *Work and Inequality in Urban China*. Albany, NY: State University of New York Press.
- Blalock, H. M. 1991, *Understanding Social Inequality: Modeling Allocation Processes*. Newbury Park, CA: Sage Publications.
- Brand, J. E. & Y. Xie 2010, "Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." *American Sociological Review* 75(2).
- Chan, K. W. 1994, *Cities With Invisible Walls; Reinterpreting Urbanization in Post-1949 China*. Hong Kong: Oxford University Press.
- 1996, "Post-Mao China: A Two-Class Urban Society in the Making." *International Journal of Urban and Regional Research* 20(1).
- Chan, K. W. & W. Buckingham 2008, "Is China Abolishing the Hukou System?" *The China Quarterly* 195.
- Chan, K. W. & L. Zhang 1999, "The Hukou System and Rural-Urban Migration in China: Processes and Changes." *The China Quarterly* 160.
- Cheng, T. & M. Selden 1994, "The Origins and Social Consequences of China's Hukou System." *The China Quarterly* 139.
- Dehejia, R. H. & S. Wahba 1998, "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs." *National Bureau of Economic Research*. Working Paper Series No. 6586.
- 2002, "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *Review of Economics and Statistics* 84(1).
- Evans, M. D. R. & J. Kelley 1991, "Prejudice, Discrimination, and the Labor Market: Attainments of Immigrants in Australia." *American Journal of Sociology* 97(3).
- Fan, C. C. 2001, "Migration and Labor-market Returns in Urban China: Results from a Recent Survey in Guangzhou." *Environment and Planning* 33.
- 2002, "The Elite, the Natives, and the Outsiders: Migration and Labor Market Segmentation in Urban China." *Annals of the Association of American Geographers* 92(1).
- Harding, D. 2003, "Counterfactual Model of Neighborhood Effects: The Effect of Neighborhood Poverty on High School Dropout and Teenage Pregnancy." *American Journal of Sociology* 109(3).
- Holland, P. W. 1986, "Statistics and Causal Inference." *Journal of the American Statistical Association* 81(396).
- Imbens, G. W. 2004, "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review." *Review of Economics & Statistics* 86(1).
- Liang, Z. 2004, "Patterns of Migration and Occupational Attainment in Contemporary China: 1985 - 1990." *Development and Society* 33.

- Liang, Z. & Y. P. Chen 2007, "The Educational Consequences of Migration for Children in China." *Social Science Research* 36(1).
- Liang, Z. & Z. Ma 2004, "China's Floating Population: New Evidence from the 2000 Census." *Population and Development Review* 30(3).
- Raudenbush, S. W. & A. S. Bryk 2002, *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks: Sage Publications.
- Rosenbaum, P. R. & D. B. Rubin 1984, "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *Journal of the American Statistical Association* 79(387).
- Solinger, D. 1999, "China's Floating Population." In Merle Goldman & Roderick Macearquhar (eds.), *The Paradox of China's Post-Mao Reforms*. Cambridge, MA: Harvard University Press.
- Wang, F. L. 2004, "Reformed Migration Control and New Targeted People: China's Hukou System in the 2000s." *The China Quarterly* 177.
- 2005, *Organizing Through Division and Exclusion: China's Hukou System*. California: Stanford University Press.
- Wang, F., X. Zuo & D. Ruan 2002, "Rural Migrants in Shanghai: Living Under the Shadow of Socialism." *The International Migration Review* 36(2).
- Wu, X. 2002, "Work Units and Income Inequality: The Effect of Market Transition in Urban China." *Social Forces* 80(3).
- Wu, X. & D. Treiman 2004, "The Household Registration System and Social Stratification in China: 1955 - 1996." *Demography* 41(2).
- Xie, Yu & Wu Xiaogang 2005, "Market Premium, Social Process, and Statisticism." *American Sociological Review* 70.
- Ye, Hua 2012, "College Expansion and School-to-Work Transition in China." Ph. D. Dissertation, Division of Social Science, The Hong Kong University of Science and Technology, Hong Kong.
- Zhang, Z. & D. J. Treiman 2012, "Social Origins, Hukou Conversion, and the Wellbeing of Urban Residents in Contemporary China." *Social Science Research*. Forthcoming.
- Zhang, Z. & X. Wu 2012, "Registration Status, Occupational Segregation, and Rural Migrants in Urban China." Working paper.

作者单位:香港科技大学社会科学部
责任编辑:张宛丽

tional migration behavior of African foreigners in China. Since the African immigrant population in Guangzhou is a hard-to-reach group, so respondent driven sampling (RDS) method was employed and 648 samples was collected in two immigrant communities. Quantitative analysis shows that the migration behavior of African foreigners in China between 2005 and 2011 has characteristics of early stage of international migration. The study also illustrates that the cumulative causation of migration is an important dynamic in migration, and its effect has been increasingly enhanced. Therefore, as long as China's economy continues to develop and the society keeps stable, the next wave of illegal immigrants would continue to strengthen itself and expand rapidly. The policy implication of this study is that the government should uplift the social strata of the African foreigners coming to China by restraining the speed and the scale of the flowing-in immigrants from the lower strata.

Household Registration, Urban Status Attainment, and Social Stratification in Urban China *Zheng Bingdao & Wu Xiaogang* 160

Abstract: The household registration system (*hukou*) is one of the most important determinants of social stratification and mobility in China. However, the existing empirical literature is fragmented, dealing with either the economic status and social wellbeing of rural migrant workers or the impact of the selective mobility of *hukou* status on social stratification outcomes. Pooling the data from three CGSS surveys (2003, 2006, 2008), we integrate the study of rural-origin people's socioeconomic mobility and the role played by successful and unsuccessful *hukou* changes of rural status. Specifically, we differentiate those who gain urban *hukou* status due to the incorporation of their villages into cities from other successful urban *hukou* obtainers, and compare their earnings with rural migrant workers who fail to obtain urban *hukou*. Linear regression results show that the commonly observed earnings premium of urban *hukou* is limited only to a subgroup of rural-origin people who obtain their urban *hukou* through a highly selective process. Propensity-score matching analyses also demonstrate that the effect of urban *hukou* on earnings is positively associated with the propensity of *hukou* conversion. The findings further show that urban *hukou* only pays off among people with high education and advantaged occupations in the state sector. Our research findings shed light on the relationship between mobility processes and stratification outcomes, and bear important policy implications for the reform of *hukou* system in the process of urbanization in China.