

户口、职业隔离 与中国城镇的收入不平等^{*}

吴晓刚 张卓妮

摘要：基于中国 2005 年 1% 人口抽样调查大样本数据，比较同工不同酬问题和职业隔离等制度性障碍在导致农民工与城镇工人间的收入不平等中所起的作用。结果表明，人们通常观察到的农民工收入低于城镇当地工人的现象主要可归因于以户口为基础的职业隔离，而非同工不同酬的劳动力市场中的直接歧视。城乡地区间不平衡的机会结构，尤其是教育机会的巨大不平等，是导致农民工与本地城镇工人职业隔离的重要原因。利用倾向评分匹配法把农民工与具有相似背景的城镇工人进行比较，发现在党政机关/事业单位类型内的农民工收入比城镇工人低；在国有/集体企业和私有部门内，农民工的收入比相应的城镇工人的收入反而更高。这说明，要改善农民工在城镇的经济社会地位，在改善城乡间教育机会失衡状况的同时，还要消除针对农民工的职业准入障碍，保证劳动力市场中的机会均等。

关键词：农民工 劳动力市场 职业隔离 收入不平等

作者吴晓刚，香港科技大学社会科学部教授，上海高校“东方学者”讲座教授；张卓妮，香港城市大学应用社会科学系助理教授。

一、引言

近年来关于农民工的不平等待遇和歧视等问题引发各界越来越多的关注。根据 2005 年全国人口抽样调查资料，我国的“流动人口”数量已高达 1 亿 4700 万，占全国总人口的 11% 以上。^① 2010 年第六次人口普查的资料进一步显示，同 2000 年

①* 感谢香港特区政府研究资助局优配基金（GRF644208、GRF646411）、2012—2013 年度“人文学及社会科学杰出学者奖”（HKUST602-HSS-12）对吴晓刚的资助。特别感谢唐启明（Donald J. Treiman）教授、谢宇教授，以及蒋勤、李骏、叶华、郑冰岛等博士和匿名评审专家提出的宝贵建议。

① 国家统计局《2005 年全国 1% 人口抽样调查主要数据公报》，2006 年 3 月 16 日，<http://wenku.baidu.com/view/6ad5bdeef8c75fbfc77db2cf.html>，2013 年 7 月 8 日。

相比,居住地与户口登记地所在的乡镇街道不一致且离开户口登记地半年以上的人口增加1亿1600多万人,增长81.03%。^①他们中的大部分是从内陆农村地区流向发达的沿海城镇地区。尽管由农业户口转为城镇非农户口理论上对流动人口是开放的,实际上获得城镇户口的人非常具有选择性,数量也很少。农民工没有当地(非农)户口,就无法获得当地政府的津贴、福利以及较好的工作机会。在农民工与本地城镇工人间存在的各种差别待遇中,最明显的是收入差距。农民工每周比城镇当地工人平均多工作8个小时,收入却只有当地居民的68%。^②

很多学者将农民工与城镇居民之间的收入差距归因于户籍制度这堵“无形之墙”所造成的社会分割和就业歧视。然而,户籍制度是通过何种机制决定两组群体间的收入差异尚未有定论。一些研究者认为,户籍制度直接影响农民工的收入。农民工在城镇劳动力市场中,不论是否在其他方面与城镇工人有差异,都被贴上了歧视性标签,其直接导致了报酬上的不平等。也就是说,这两组人之间的收入差距是由雇主对农民工身份的区别对待造成的。另一些学者则认为,农民工身份本身并不直接影响他们和城镇工人之间的收入差异,农民工收入低主要归因于与经济效率相关的个人和结构性特征,而非在工资报酬方面对户口身份的区别对待。一方面,农民工在劳动技能、教育水平和工作经验等人力资本特征上比不上城镇工人;另一方面,城镇劳动力市场的结构性分割阻止农民工进入待遇较好的部门和职业。当考虑了这些个人和结构性特征后,农民工和城镇工人间的收入差异将大大减少甚至消失。^③

职业作为一种重要的“报酬机制”,在决定农民工和城镇工人的收入不平等中起了关键性的“中介”作用。以往大部分关于农民工收入获得的研究仅仅把职业当成分析模型中的一个控制变量,而未重视职业在个人特征和收入获得之间的中介效应。在城镇劳动力市场中,农民工和当地工人首先根据先赋的和自致的特征,被筛选进入不同的职业之中;他们在职业分布上的差异接着又与其他有差异的特征一起决定两个群体之间的收入差异。因此,城镇工人和农民工之间的收入不平等是以下两个部分共同作用的结果:(1)不同职业间的报酬差异;(2)职业内部的报酬差异。前者为入职门槛效应,反映的是户籍制度通过职业结构对收入不平等产生的间接影响;后者可以近似地视为“同工不同酬”问题。只有把职业间和职业内的差异分开,才能揭示户籍制度(农民工身份)对收入不平等的作用机制。本文旨在推进

① 国家统计局《2010年全国第六次人口普查主要数据公报》,2012年4月20日, http://www.gov.cn/test/2012-04/20/content_2118413.htm,2014年2月15日。

② 李培林、李炜《农民工在中国转型中的经济地位和社会态度》,《社会学研究》2007年第3期。

③ C. Cindy Fan, “The Elite, the Natives, and the Outsiders: Migration and Labor Market Segmentation in Urban China,” *Annals of the Association of American Geographers*, vol. 92, 2002, pp. 103-124.

这一研究。

二、职业、户籍制度和城镇的收入不平等

社会学家通常将职业视为现代社会不平等产生的核心机制。在收入不平等的产生过程中，职业通过两个不同的机制发挥影响：第一，职业结构决定了不同职业间的经济回报差异。劳动者依据不同的人力资本特征（如教育、工作经验、具体的工作技能）和人口统计学特征（如年龄、性别等）被分配到职业结构中的不同位置上。因为这些不同的职业位置附着着不同的收入回报水平，被分配到不同职业的人之间就产生了收入差异。这种职业间的收入不平等在本质上是“工作准入”机会差异的产物。

第二，在相同的职业内部，对不同类型工人的报酬差异决定了职业内的收入不平等。这种职业内不平等与由职业分布差异所带来的不平等是截然不同的，它反映的是在相同的职业类别中，不同类型的人基于何种因素而获得有差异的经济报酬。一般来说，职业内的报酬差异可能来源于雇主对某类雇员的歧视，比如对相同职业内具有同等教育水平和工作经验的男性雇员和女性雇员区别对待，而且这种区别对待仅仅是由性别身份带来的。迄今为止，职业隔离和同工不同酬在不平等研究中被广泛用于分析不同群体之间的收入差异，如按性别、移民身份等划分的比较分析。^① 本文以变迁中的中国城镇劳动力市场为背景，从一个不同的维度研究职业间和职业内收入不平等，重点关注以户籍制度为基础的农村劳动力迁移（农民工身份）对收入获得的影响。

我国长期以来实行的户籍登记制度实际上是一种行政上的“准入”控制，一方面它控制农村居民进入城镇地区，另一方面对已进入城镇地区的农村劳动力的就业进行管理。随着经济改革的深入，地区“准入”控制已经被大大削弱，但工作“准入”障碍仍然是当今城镇劳动力市场的主要特征。农民工可以在城市工作，但大部分从事的是那些城镇当地工人不愿从事的低技术含量、甚至具潜在危险性的体力工作。户籍制度严重影响了人们的职业获得和地位晋升。事实上，许多地方政府制定了各种就业规定保护当地居民，以免他们在与外来劳动力的竞争中失利；在 20 世纪 90 年代中期城镇下岗和失业问题变得日益严重之后，尤其如此。^② 例如，有些地方

^① Philip N. Cohen and Matt L. Huffman, "Individuals, Jobs, and Labor Markets: The Devaluation of Women's Work," *American Sociological Review*, vol. 68, 2003, pp. 443-463; Trond Petersen and Laurie A. Morgan, "Separate and Unequal: Occupation-Establishment Sex Segregation and the Gender Wage Gap," *The American Journal of Sociology*, vol. 101, 1995, pp. 329-365.

^② 蔡昉、都阳、王美艳 《户籍制度与劳动力市场保护》，《经济研究》2001 年第 12 期。

政府规定，企业必须遵守“城市工人优先于农村工人”和“本地工人优先于外来工人”的雇佣原则。北京市有关部门曾规定，所有的工作都被划分为三类：（1）可以雇佣农民工但要求被雇佣者具有初中以上学历的工作；（2）不允许雇佣农民工的工作（这类工作的数目从1997年的32个上升到2000年的103个）；（3）在雇佣了某个特定比例的城镇下岗工人后才对农民工开放的工作。不少地方政府还规定，招聘农民工的单位必须缴纳一定费用，而被雇佣的农民工也要花钱办理各种证件。^①

虽然已有大量研究关注与户籍制度相关的制度障碍如何影响农民工与当地城镇工人之间的收入差异。然而，只有少数学者关注两者之间收入差异究竟在多大程度上归因于职业间和职业内的差异，而且他们的分析存在的最大问题是职业分类过于简单，或者干脆忽略之；或者主要以单位类型辅以两类职业为依据检验就门槛和同工不同酬对收入差异的影响。如王美艳仅把就业岗位划为三大类，即公有单位职工、非公有单位职工和非公有单位行政管理及专业技术人员；田丰关注的是就单位而完全没有考虑职业。^② 这些分析无法准确反映就门槛和同工不同酬的作用，原因有二：首先，单位和职业是反映中国城镇劳动力市场分割的两个非常重要却截然不同的维度。相同单位内职业千差万别，只关注单位对农民工的隔离作用过于简单，无法准确测量工作“准入”障碍，亦无法反映同工不同酬问题。其次，由于大部分（约95%）农民工都在非公有单位工作^③，仅以单位类型划分，同一类型单位内部从事不同职业的人员的异质性依然很大。要检验就门槛与同工不同酬对农民工与城镇工人间收入不平等的影响，必须在单位基础上以他们从事的具体职业为分析依据。

孟昕和张俊森^④利用上海1995、1996年的两个调查数据，分析职业隔离和同工不同酬对农民工和城镇工人收入不平等的各自不同作用。他们依据职业而不是单位来分析收入不平等的来源，发现农民工和城镇工人间大部分收入差距是由相同职业内不同的工资待遇造成的，而且这种职业内收入差异无法被可观测到的个体特征差异所解释，他们由此断定对农民工存在着身份歧视。我们认为这篇文章有两个严重的不足。首先，他们仅把职业分为4个大类，这种大分类法并不足以准确测量职业

① 蔡昉 《中国城市限制外地民工就业的政治经济学分析》，《中国人口科学》2000年第4期；谢桂华 《农民工与城市劳动力市场》，《社会学研究》2007年第5期。

② 王美艳 《转轨时期的工资差异：工资歧视的计量分析》，《数量经济和技术经济研究》2003年第5期；田丰 《城市工人与农民工的收入差距研究》，《社会学研究》2010年第2期。

③ 田丰 《城市工人与农民工的收入差距研究》，《社会学研究》2010年第2期。

④ Xin Meng and Junsen Zhang, “The Two-Tier Labor Market in Urban China: Occupational Segregation and Wage Differentials between Urban Residents and Rural Migrants in Shanghai,” *Journal of Comparative Economics*, vol. 29, 2001, pp. 485-504.

隔离和同工不同酬的作用。因为这4个大类的职业中,每一个类别都包含了十几个甚至几十个不同的职业。他们所谓的(大类)职业内的报酬差异可能反映的正是(细分)职业间的差异。职业分类的粗细不同,得到的结论可能大相径庭。此外,孟昕和张俊森的研究忽略了单位的作用。大量实证研究表明,单位类型是城镇劳动力市场中收入获得的另一个重要决定因素。

单位类型与户籍制度联系在一起,是对农民工造成职业隔离之外的另一种“准入”障碍。一方面,不同的单位附着着不同的机会和回报结构。国有部门工作稳定,且提供岗位津贴、住房补贴/住房公积金、医疗保险和养老保险等一系列福利;而非国有部门的工人必须凭工作绩效获取稳定的工作和相应的收入,福利保障相对少很多。随着改革的深入,单位制对城市就业人口社会经济地位获得的影响正在减弱,但单位壁垒效应持续至今。^①另一方面,户籍制度阻碍了农民工进入国有部门。大部分农民工只能在私有部门工作,即使有极少数农民工能够顺利进入国有部门,大多也只是编制外的临时工人。

以往的研究之所以无法同时考虑单位和职业的一个主要原因是所用调查数据的样本量太小(往往只有数千人),且农民工在国有部门的数量非常少,若同时分析单位和职业的作用就无法得到足够有效的样本。小样本数据也使得研究者无法使用细分的职业类别。本文利用2005年1%人口抽样调查数据的特殊性,试图弥补过往研究的不足,以检验农民工和城镇工人间的收入差异,特别关注职业隔离在造成这种不平等过程中的作用。

三、数据、变量及其描述性分析

(一) 数据和变量

由国家统计局负责设计和执行的2005年全国1%人口抽样调查,俗称“小普查”,使用多阶段、分层、整群概率比例的抽样方法,在两次人口普查之间搜集社会人口信息,其设计具有以下两个特点。首先,与以往人口普查和小普查不同,2005年小普查除搜集户口性质、户口登记地、现居住地、教育程度、职业和其他人口特征等方面的信息外,第一次将每月工作收入、雇佣身份、单位类型、工作时间、工作福利等内容纳入调查范围。小普查数据中包括的职业类别尽管不如人口普查中使用三位码的中国标准职业编码(CSCO)那么细致,其使用的CSCO两位码亦涵盖近80种职业。我们将研究对象限定在现居住于城镇地区的16—60岁的一个人口样

^① 边燕杰、李路路、李煜、郝大海《结构壁垒、体制转型与地位资源含量》,《中国社会科学》2006年第5期。

本 (N = 119675), 共有 68 类两位数的中国标准职业编码 (CSCO)。其次, 与学术研究机构设计和搜集的抽样调查数据相比, 2005 年小普查数据样本量大, 职业分类较细。特别值得一提的是, 这项调查设计首次采用了“两头登记”的原则, 即在所抽中的调查小区中, 所有在调查时点 (2005 年 11 月 1 日凌晨) 当晚居住在当地的居民, 无论其户口登记地在何处, 都进行登记; 也包括所有户口登记在本地的人口。调查后再根据不同的口径标准, 对户籍人口、常住人口、流动人口进行定义和统计。这种新的登记方法, 比以往使用的其他方法能涵盖更多的迁移人口。因此, 2005 年小普查数据在研究职业隔离对农民工—城市居民劳动力市场中的收入不平等的影响问题具有不可替代的价值, 是迄今为止研究中国劳动力市场转型的最权威数据来源。

基于 2005 年小普查数据, 表 1 按是否外来人员 (户口登记在外乡 [镇/街道])、户口性质和户口登记地类型, 列出了城镇地区 16—60 岁劳动人口的分布情况。每个单元格内的百分比反映了在城镇总体中的比例, 括号内的数字为该组就业人口的平均月收入。在所有处于工作年龄 (16—60 岁) 且居住于城镇地区的人口, 42.6% 是具有非农户口的本地居民; 27.3% 是具有农业户口的本地居民, 这些人大部分是城镇郊区的农民, 或因城镇化而被划为城镇居民但并未改变户口性质的人。外来人口的比例达 30.2%, 其中约 61% (= 18.3%/30.2%) 户口登记地在农村, 60% (= [17.3% + 0.8%] / 30.2%) 具有农业户口。大部分户口登记地在农村的外来人员持有农业户口 (17.3%/18.3% = 95%), 而大部分户口登记地在城镇的外来人员持有非农户口 (11.1%/11.9% = 93%)。

表 1 2005 年按迁移身份和户口特征划分的城镇人口分布 (16—60 岁)

	流动人口		当地居民
	户口登记地类型		
	农村	城镇	
户口性质			
农业户口	17.3% (945)	0.8% (978)	27.3% (550)
非农业户口	1.0% (1182)	11.1% (1609)	42.6% (1134)
小 计	18.3%	11.9%	
总 计	30.2%		69.9%

注: N = 119675。所有的百分数均为城镇总人口 (16—60 岁) 中的比例。括号内为月收入均值 (单位: 元), 从该类别中劳动人口所报告的月收入中计算而得 (第一行数据对应频数从左至右为 16349, 617 和 24988, 第二行数据对应频数从左至右为 826, 8709 和 30206)。

表 1 括号内的数字显示, 按平均月收入来看, 主要差距在于是否拥有农业/非农

户口。具有农业户口的城镇当地居民的平均月收入为 550 元，大约只有非农户口居民月收入（1134 元）的一半。从农村和城镇地区来的农业户口外来人员平均月收入分别为 945 元和 978 元；而相应的具有非农户口的外来人员月收入则分别为 1182 元和 1609 元，这些非农外来人员的平均月收入甚至比当地居民还高。可见，劳动力迁移对收入的影响更多地取决于户口性质而不是户口登记地的类型。

图 1 对表 1 中不同类别的群体分别划出了职业的国际社会经济地位指数（ISEI）分布密度曲线。在分析样本中，一共有 68 类两位数的中国标准职业编码（CSCO），先把这些详细的 CSCO 编码转换成国际标准职业编码（ISCO88），然后再把 ISCO88 转换成 ISEI，即职业的国际社会经济地位指数。^① ISEI 分值越大，对应的职业的社会经济地位越高。具有农业户口的城镇当地居民不在图 1 中，因为其职业分布与其他 5 组人的职业分布非常不同，他们大部分从事农、林、牧、渔、水利业，与城镇劳动力市场关系不大。

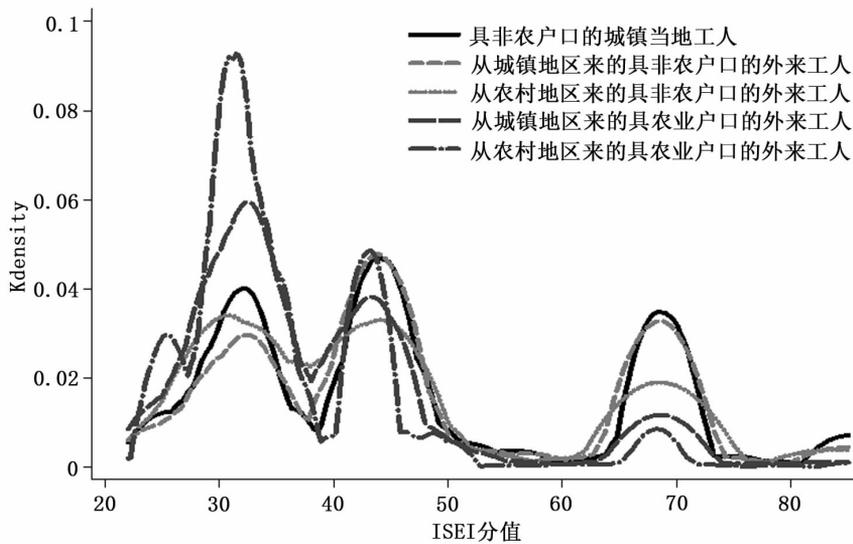


图 1 2005 年城镇地区不同类别工人的 ISEI 密度（Kernel Density）分布曲线

其他 5 组人的 ISEI 分布大致呈现两种模式。具有非农户口的三组工人（包括具有非农户口的城镇当地工人、从城镇地区来的和从农村地区来的具有非农户口的外来工人）的分布模式相似，大约 1/3 ISEI 分值低于 40，1/3 在 40 和 60 之间，最后 1/3 高于 60。相比之下，大部分具有农业户口（包括从城镇地区和农村地区来的）外来工人 ISEI 分值低于 50；其中从农村来的具有农业户口的外来工人高度集中在低端职业。不同的分布模式再次说明，劳动力迁移的影响主要来自户口性质而不是户口登记地。具有非农户口的外来就业者，不管户口登记地在农村还是

^① Harry B. G. Ganzeboom and Donald J. Treiman, “Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations,” *Social Science Research*, vol. 25, 1996, pp. 201-239.

在城镇，其社会经济地位分布情况都与城镇当地非农工人的职业分布类似，而外来的农民工，同样不管户口登记地在农村还是在城镇，与当地工人相比职业分布模式很不一样。

因此，在本文中采用的主要自变量是农民工身份，它是根据户口性质而不是户口登记地来界定的，是一个虚拟变量（农民工 = 1，城镇工人 = 0）。农民工指在城镇居住、户口登记在外乡（镇/街道）、具有农业户口的外来工人，城镇工人指户口登记在本乡（镇/街道）且具有非农户口的城镇当地工人。持有非农户口的外来工人和持有农业户口的本地居民不在本文的分析范围内，因为前者主要是人才/精英流动，而后者主要是城镇郊区农民，这些人因为城镇扩大化而改变了户口登记地类型（由农村转为城镇），但没有改变户口性质。排除了具有缺失值的观测记录后，分析样本中共有 15996 个农民工和 28661 个城镇当地工人。

除农民工身份外，职业和单位类型是另外两个重要的自变量。就职业而言，我们使用了大分类和详细的两位数职业分类方式。由于某些详细的职业类别样本量太小，我们把 68 类 CSCO 职业代码归纳为 38 类。单位分为三类，包括党政机关/事业单位、国有/集体企业和私有部门。

此外，我们还控制了其他因素对收入的影响，包括教育程度、性别、婚姻状况、年龄、雇佣身份、每周工作时间和居住地（县级地区）。教育程度指受访者获得的最高教育水平。婚姻状况是虚拟变量（已婚 = 1，其他 = 0）。雇佣身份包括自雇人士、雇员和雇主。年龄和每周工作时间是连续变量。在分析收入时，我们使用了一系列县级层次的虚拟变量，以控制地区间差异对农民工与城镇工人收入差异造成的影响。

我们主要关注的因变量是收入，将月收入取自然对数作为回归分析中的因变量。

（二）描述性分析

表 2 列出了农民工与城镇当地工人在教育、职业、雇佣身份、单位类型、性别、年龄和工作时间等方面的统计数字。总的来说，农民工与城镇工人之间有很大的差异。首先，农民工的教育水平较低。只有 18.5% 的农民工获得高中及以上的教育，但 66.2% 的城镇工人至少拥有高中学历。这种差异反映了我国城乡之间在教育机会和教育资源上的巨大差别，而且这种差别并没有因为教育扩张而显著缩小。其次，职业和单位分布的差异显示农民工在城镇劳动力市场中的隔离状况。只有不到 10%（1.6% + 3.2% + 3.6%）的农民工是管理者、专业技术人员或办事人员，但将近一半的当地工人从事这些职业。两者间在工作单位类型上的差异更加明显，只有 1.5% 的农民工在党政机关/事业单位，9% 在国有/集体企业；但将近 62%（27.4% + 34.4%）的城镇工人在这两类部门工作。

表 2 2005 年城镇地区农民工与当地工人的描述性统计 (16—60 岁)

	农民工	当地工人
教育程度		
小学及以下	19.6	4.7
初 中	61.9	29.1
高中及以上	18.5	66.2
职业类别		
管理者	1.6	4.9
专业技术人员	3.2	24.7
办事人员	3.6	15.9
商业及服务业人员	34.9	26.4
体力劳动者	56.7	28.1
雇佣身份		
雇 员	78.8	85.1
雇 主	4.3	3.5
自雇人士	16.9	11.4
单位类型		
党政机关/事业单位	1.5	27.4
国有/集体企业	9.0	34.4
私有部门	89.6	38.2
女 性	41.5	40.9
年 龄	30.2 (9.4)	38.0 (9.1)
每周工作小时数	55.2 (13.2)	45.8 (10.4)
N	15996	28661

注：对于类别变量，数字对应的是百分数；对于连续变量，数字对应的是均值（括号内为标准差）。

表 2 中其他变量的分布情况与以往研究的发现大致相同。比如，农民工从事自雇工作的比例更高，因为他们在城镇劳动力市场中更难找到正式工作；就平均年龄来说更年轻；其工作时间远远高于城镇工人，大约每周多 10 个小时。

表 3 列出了农民工和当地城镇工人的平均月收入，分别比较了不同的单位类型内两者的差异。总体上，农民工平均每月收入为 968 元，而城镇工人的月平均收入

为 1169 元，T 检验结果显示这种差别是显著的。结合前述平均工作时间的差别，农民工与城镇工人之间的收入差距更加显著。虽然农民工每月比城镇工人多工作 40 个小时（10 小时 × 4 周），他们的月收入依然比城镇工人低 17%（ $= 1 - 0.83$ ）。这是因为农民工所从事的是城镇工人不想做的低技术性工作，这些工作需要的工作时间较长且收入回报较低。

表 3 2005 年城镇农民工与当地工人的平均月收入（按单位类型划分）

	农民工 (R)	当地工人 (U)	比率 (R/U)
党政机关/事业单位	846 (500)	1329 (852)	0.64
国有/集体企业	1022 (576)	1172 (910)	0.87
私有部门	965 (754)	1052 (1295)	0.92
总样本	968 (737)	1169 (1066)	0.83
T 检验	$\Pr (T_R < T_U) = 0.0000$		
N	15996	28661	

注：单位为元，括号内为标准差。

农民工与城镇工人间的收入差距随单位类型而变化。两者间的收入差距在党政机关/事业单位最大，他们平均月收入比城镇工人少 36%；在国有/集体企业的差距次之，平均每月少 13%；在私有部门差距最小，平均每月少 8%。表 3 反映的农民工与城镇工人间的收入差距，可能是农民工身份对收入获得的作用，也可能源于其他因素的影响，如教育程度。为排除这些因素的影响，我们首先利用多元线性回归法，检验农民工身份与收入获得之间的关系，以及这种影响如何随单位类型的变化而变化；然后用布朗分解法（Brown Decomposition Method）将农民工和城镇工人间的收入不平等分解为职业间和职业内收入差异，检验这两组差异对总体不平等的相对影响；^①此外，作为对布朗分解法的补充，我们利用倾向性评分匹配法（propensity score matching method），进一步研究在排除所有可观测的决定因素的影响后，农民工身份变量对收入差异的作用。

^① R. S. Brown, M. Moon and B. S. Zoloth, “Incorporating Occupational Attainment in Studies of Male-Female Earnings Differentials,” *Journal of Human Resources*, vol. 15, 1980, pp. 3-28.

四、分析结果

(一) 农民工身份与中国城镇劳动力市场的收入决定模式

表 4 列出了月收入对数的多元线性回归模型的分析结果。这些模型控制了县级层次的效应。在前三个模型中,我们逐步加入不同的控制变量,以比较农民工身份的影响如何随着控制变量的增加而改变;在最后一个模型中,我们加入了农民工身份和单位类型之间的交互项,以检验农民工身份的作用是否随单位类型而变化。

表 4 2005 年城镇地区月收入对数的决定模式

因变量: 月收入对数	模型 1	模型 2	模型 3	模型 4
农民工	-0.119** (0.007)	-0.077** (0.007)	-0.054** (0.007)	-0.450** (0.033)
女 性	-0.205** (0.005)	-0.197** (0.005)	-0.195** (0.005)	-0.195** (0.005)
已 婚	0.083** (0.009)	0.068** (0.009)	0.060** (0.009)	0.059** (0.009)
教育程度 (参照组: 小学及以下)				
初 中	0.176** (0.009)	0.161** (0.008)	0.158** (0.008)	0.158** (0.008)
高中及以上	0.487** (0.009)	0.379** (0.009)	0.360** (0.009)	0.358** (0.009)
年 龄	0.024** (0.002)	0.024** (0.002)	0.023** (0.002)	0.024** (0.002)
年龄 ² × 100	-0.028** (0.003)	-0.029** (0.003)	-0.030** (0.003)	-0.029** (0.003)
职业类别 (参照组: 体力劳动者)				
管理者		0.423** (0.013)	0.390** (0.013)	0.389** (0.013)
专业技术人员		0.273** (0.008)	0.213** (0.008)	0.210** (0.008)
办事人员		0.166** (0.008)	0.104** (0.009)	0.103** (0.009)
商业及服务人员		-0.078** (0.006)	-0.077** (0.006)	-0.073** (0.006)

续表 4

因变量: 月收入对数	模型 1	模型 2	模型 3	模型 4
雇佣身份 (参照组: 自雇人士)				
雇 员		-0.017* (0.008)	-0.067** (0.008)	-0.069** (0.008)
雇 主		0.289** (0.014)	0.297** (0.014)	0.300** (0.014)
单位类型 (参照组: 党政机关/事业单位)				
国有/集体企业			-0.121** (0.008)	-0.130** (0.008)
私有部门			-0.217** (0.009)	-0.247** (0.009)
交互项				
国有/集体企业 × 农民工				0.352** (0.035)
私有部门 × 农民工				0.420** (0.033)
每周工作时间		0.003** (0.001)	0.005** (0.001)	0.006** (0.001)
每周工作时间 ² × 100		-0.003** (0.001)	-0.004** (0.001)	-0.004** (0.001)
截 距	6.089** (0.036)	6.009** (0.048)	6.188** (0.049)	6.194** (0.048)
N	44657	44657	44657	44657
R ²	0.186	0.251	0.262	0.265
县级地区的数量	2418	2418	2418	2418

注: 括号内为标准误差; **p < 0.01, *p < 0.05。

在控制了性别、婚姻状况、教育程度和年龄的效应后, 农民工的月收入依然比城镇工人低 11% ($1 - e^{-0.119}$), 而加入职业类别、雇佣身份和每周工作小时后, 这种差距缩小到 7% ($1 - e^{-0.077}$)。在模型中进一步控制单位类型的影响后, 差距缩小到 5% ($1 - e^{-0.054}$)。这些系数的变化表明, 我们所观察到的农民工与城镇工人间的实际收入差距 (如表 3 所示), 在一定程度上来自两者在人力资本和职业类别等因

素上的差异,但这些因素并不能完全解释为什么农民工比城镇工人收入更低。在城镇劳动力市场上,户口性质对收入依然存在着显著影响。

表 4 模型 4 中农民工身份和单位类型的交互作用显示,农民工身份的影响在不同单位类型间存在显著差别。在党政机关/事业单位中影响最大,农民工的平均月收入比城镇工人低 36% ($1 - e^{-0.450}$); 在国有/集体企业中相应的比例为 9% ($1 - e^{-0.450+0.352}$); 在私有部门为 3% ($1 - e^{-0.450+0.420}$)。这些结果与表 3 的结果一致,但相比表 3,因排除了其他特征的影响,收入差异明显缩小。

其他变量对收入获得的影响与预期的一致。女性相比男性收入更低。已婚人士比未婚人士收入更高。教育与收入正相关,年龄和每周工作时间对收入的作用是非线性的,我们在模型中加入了他们的平方项。相比体力劳动者,管理者、专业技术人员和办事人员的收入更高,但商业与服务业人员收入更低。相比自雇佣者,雇主收入更高,而雇员则收入更低。相比党政机关/事业单位,国有/集体企业和私有部门的工人收入更低。

表 4 中农民工身份对收入的影响随着不同单位类型而变化,可能因为对农民工身份的回报机制在不同单位类型间存在差异: 农民工在以国家为导向的单位内回报最低,因为户口身份以及其他制度性限制条件的影响更大; 在以市场为导向的企业内回报高些,可能的原因是,这里工作能力和绩效相对更加重要,而面临的身份歧视更少一些。但也有另一种可能,即这是农民工和城镇工人在职业分布上存在的差异随单位部门变化而导致的结果。

首先,不同单位类型就业的比例在农民工和城镇工人间存在显著差异。从描述性分析中已知,只有 2.4% 的农民工在党政机关/事业单位和国有/集体企业中,而超过 60% 的城镇工人在这两类部门。其次,大部分在这两类部门工作的农民工所从事的是低端工作,如在样本中,大多是安全保卫和消防人员、邮政和电信业人员、仓储人员、餐饮服务人员,等等。因此,当用 OLS 回归法估计农民工身份对收入的影响时,其实比较的是在职业分布上存在巨大差异的两组人之间的平均收入,而两者间的职业分布差异在不同的单位类型中是不同的。可以利用适当的统计方法,检验农民工与城镇工人在职业分布等特征方面存在的差异如何影响他们之间的收入差异。布朗分解法和倾向评分匹配法就是其中两种可用的方法。

(二) 职业隔离: 农民工与城镇工人收入不平等的来源

在分析农民工与城镇工人间的职业分布差异是否影响两者间收入不平等前,我们先对样本中的原始职业类别计算邓肯“隔离指数”(Duncan Dissimilarity Index),检验这两组人的职业分布差异程度。邓肯“隔离指数”反映的是,要使农民工与城镇工人具有相同的职业分布模式,那么将有多大比例的农民工需要被重新分配到新

的职业类别中去。^① 据此，总样本中大约 45% 的农民工需要被重新分配到另一种职业类别中去；按单位类型划分，在党政机关/事业单位对应的数字为 57%，在国有/集体企业为 42%，在私有部门为 32%。这些数字充分说明，农民工与城镇工人在职业分布上确实存在巨大差异；而且这种差异随单位类型变化，在党政机关/事业单位差异最大，国有/集体企业次之，在私有部门相对最小但比例仍高于 3/10。

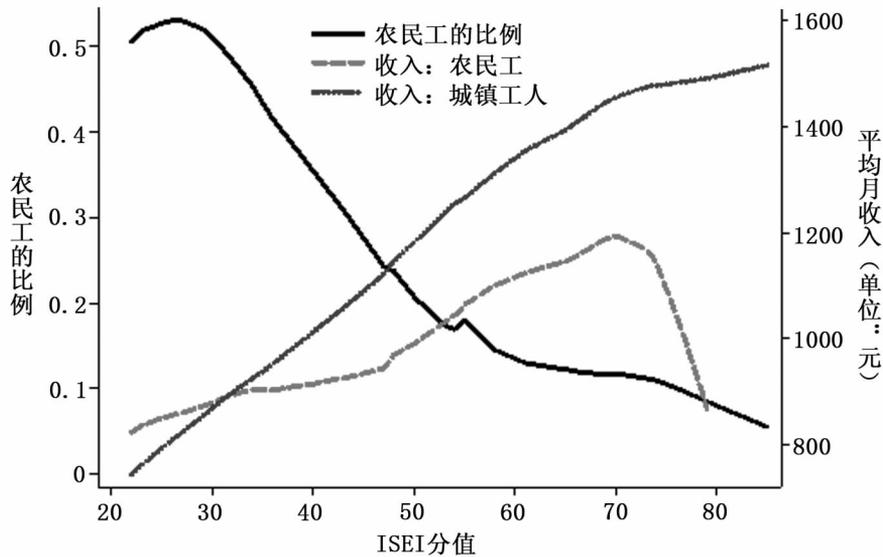


图2 按 ISEI 划分的农民工比例、农民工和城镇工人平均月收入曲线

为比较农民工的职业隔离情况以及他们与城镇工人间的收入差异，在图 2 中画出了按 ISEI 划分的农民工比例、农民工和城镇工人月平均收入曲线。对应不同的 ISEI 分值，实线表示农民工在对应的职业中的比重，虚线和点线对应的是农民工和城镇工人各自月平均收入。

从图 2 可见，随着职业社会经济地位从低到高，农民工的比例显著地从高于 50% 下降到低于 10%。即社会经济地位低的职业明显由农民工占主导，而社会经济地位高的职业则主要被城镇工人所占据。这说明在我国的城镇劳动力市场中，确实存在基于户口身份的职业隔离状况。值得注意的是，农民工与城镇工人的平均月收入曲线存在交叉。当 ISEI 分值高于 30 时，农民工的平均收入明显低于城镇工人的收入；而且这种差异随着 ISEI 分值的增加而越来越大，尤其是当 ISEI 分值高于 70 时，城镇工人的月收入仍平缓增长，但农民工的月收入却急剧下降，说明即使一部分农民工真的进入了高级别的职业类别，他们所获得的报酬却很低。在样本中，

① 邓肯“隔离指数”的计算公式： $DI = \frac{1}{2} \sum_{j=1}^N \left| \frac{r_j}{R} - \frac{u_j}{U} \right|$ ，其中， r_j = 农民工在第 j 个职业类别中的人数； R = 农民工的总人数； u_j = 城镇工人在第 j 个职业类别中的人数； U = 城镇工人的总人数； $N = 68$ ，在本文的分析样本中共有 68 类原始职业类别。

ISEI 分值高于 70 的农民工中 80% 教育程度为高中或以下; 有 2% 为单位负责人、19% 为工程技术人员、20% 为农业技术人员、49% 为卫生专业技术人员, 但对应的月平均收入却分别只有 778 元、1359 元、777 元、879 元; 这很可能是因为他们所从事的具体工作是临时性的、边缘化的。相反, 当 ISEI 分值低于 30 时, 农民工的平均收入却高于城镇工人的收入。这两条交叉的收入曲线再次说明对农民工和城镇工人间的收入差异进行分解的必要性。之后, 使用分解法把这两组人之间的收入差异分解成职业间收入差异和职业内收入差异。

(三) 来自分解法的发现

布朗分解法把职业当成中介变量, 并且允许其他变量解释收入的影响随职业类型发生的变化而变化。^①

利用布朗分解法, 我们把城镇工人和农民工之间总的收入差异分解成 4 个部分, 如公式 (1) 所示:

$$\begin{aligned}
 \bar{W}^U - \bar{W}^R &= \sum_j (P_j^U \bar{W}_j^U - P_j^R \bar{W}_j^R) \\
 &= \sum_j (P_j^R \bar{W}_j^U - P_j^R \bar{W}_j^R + P_j^U \bar{W}_j^U - P_j^R \bar{W}_j^U) \\
 &= \sum_j P_j^R (\bar{W}_j^U - \bar{W}_j^R) + \sum_j \bar{W}_j^U (P_j^U - P_j^R) \\
 &= \underbrace{\sum_j P_j^R \hat{\beta}_j^U (\bar{X}_j^U - \bar{X}_j^R)}_{(1)} + \underbrace{\sum_j P_j^R \bar{X}_j^R (\hat{\beta}_j^U - \hat{\beta}_j^R)}_{(2)} \\
 &\quad + \underbrace{\sum_j \bar{W}_j^U (P_j^U - \hat{P}_j^R)}_{(3)} + \underbrace{\sum_j \bar{W}_j^U (\hat{P}_j^R - P_j^R)}_{(4)} \tag{1}
 \end{aligned}$$

其中上标 U 和 R 分别表示城镇工人和农民工, 下标 j 表示第 j 个职业类别。 \bar{W}^U 和 \bar{W}^R 分别代表总样本中城镇工人和农民工的平均月收入的对数值, \bar{W}_j^U 和 \bar{W}_j^R 分别代表在第 j 个职业类别城镇工人和农民工的平均月收入的对数值。 \bar{X}_j^U 和 \bar{X}_j^R 分别是第 j 个职业类别两组人在各解释特征上的均值。 $\hat{\beta}_j^U$ 和 $\hat{\beta}_j^R$ 分别是第 j 个职业类别两组人各自的收入 (对数) 决定模式中各解释特征所对应的系数。 P_j^U 和 P_j^R 分别对应两组人的实际职业分布模式, 即在第 j 个职业类别上的比例。 \hat{P}_j^R 是把农民工放到城镇工人的职业获得模式中所得到的“假设性”职业分布, 即假定的在第 j 个职业类别上的比例。^②

由于某些详细的职业类别样本量太小, 这个分解法分析中, 职业一共分为 38 个类别 ($j = 1, 2, 3, \dots, 38$)。对每个职业类别, 我们都把城镇工人和农民工分成两个

① R. S. Brown, M. Moon and B. S. Zoloth, “Incorporating Occupational Attainment in Studies of Male-Female Earnings Differentials.”

② 在本文中, 我们把城镇工人当成参照组, 即假定在城镇劳动力市场中农民工受到了不平等待遇, 而城镇工人则受到正常待遇。

单独的样本，并分别估计具有相同解释变量的平均月收入（对数）普通最小二乘法（OLS）回归方程，以得到对应于方程中各解释变量的系数 $\hat{\beta}_j^U$ 和 $\hat{\beta}_j^R$ 。该收入决定方程中的解释变量包括性别、婚姻状况、教育程度、雇佣身份、就业单位、年龄（及其平方）、工作时间（及其平方）和县级层次的虚拟变量。由于总共有 38 组分析结果，在此不一一报告。此外，在计算农民工的“假设性”职业分布 \hat{P}_j^R 时，我们首先对城镇工人估计一个多项 logit 职业获得模型，其中自变量包括性别、婚姻状况、教育和年龄（及其平方）；然后用这个城镇工人的职业获得分析结果预测农民工的“假设性”职业分布。因为分析结果太长，这里不再赘述。^①

通过公式（1）的分解方法，我们把城镇工人与农民工之间的月收入（对数）差异（即 $\bar{W}^U - \bar{W}^R$ ）分解成以下 4 个组成部分（对应于该公式的最后 4 项）：（1）由两组人间的特征差异（ $\bar{X}_j^U - \bar{X}_j^R$ ）带来的职业内收入差异；（2）由两组人的收入模式中的系数差异（ $\hat{\beta}_j^U - \hat{\beta}_j^R$ ）带来的职业内收入差异；（3）由两组人特征差异所导致的职业分布差异（ $P_j^U - P_j^R$ ）带来的收入差异；（4）由两组人职业获得模式中系数差异所导致的职业分布差异（ $\hat{P}_j^R - P_j^R$ ）带来的收入差异。总的来说，（1）和（2）之和代表职业内收入差异，（3）和（4）之和代表职业间收入差异；而（1）和（3）之和代表可以用特征差异所解释的收入差异，（2）和（4）之和代表不能用特征差异所解释的收入差异。

表 5 列出了布朗分解法的分析结果，包括总的收入差异和对应公式（1）分解后的各部分收入差异，以及可观测特征所解释/不能解释部分在职业内和职业间收入差异中所占的比例。因为公式（1）中，收入差异等于城镇工人的月收入自然对数减去农民工的月收入自然对数，所以表 5 中的正数值表示城镇工人的收入高于农民工的收入，负数值则表示城镇工人的收入低于农民工的收入。

表 5 2005 年城镇地区农民工与当地工人月收入差异的布朗分解法结果

	月收入自然对数差异	占职业内/职业间差异的比例 (%)
总差异 (城镇工人—农民工)	0.1174	
职业内差异	-0.0826	100
可观测特征所解释的部分	0.0528	-64
可观测特征不能解释的部分	-0.1355	164
职业间差异	0.2000	100
可观测特征所解释的部分	0.1534	77

① 数据省略部分如有需要请联系: sowu@ust.hk 或 zhuoni.zhang@cityu.edu.hk.

续表 5

	月收入自然对数差异	占职业内/职业间差异的比例 (%)
可观测特征不能解释的部分	0.0466	23
总体被解释的部分	0.2062	
总体未解释的部分	-0.0889	

总的来说, 城镇工人的平均月收入比农民工高 12% ($e^{0.1174} - 1$)。把两者间总的收入差异分解后可见, 职业内的收入差异是负数 (-0.0826), 而职业间的收入差异是正数 (0.200)。这说明, 总体上城镇工人比农民工收入更高, 完全来自于职业间的收入差异: 相比农民工, 城镇工人主要从事具有更高地位和收入的职业。此外, 有意思的是, 城镇工人由职业优势所带来的收入优势, 在一定程度上被一些职业内的低收入所抵消, 总的来说, 在相同的职业内部, 城镇工人的平均月收入比农民工更低。

数值为负的职业内部收入差异不能被城镇工人和农民工之间的个体特征差异所解释, 反而来自于未解释部分 (即 -0.1355)。在相同的职业内部, 城镇工人比农民工收入更低, 这种差距并不能由性别、婚姻状况、教育、雇佣身份、就业单位和年龄等特征所解释, 而是由其他没有被观察到的因素所带来的。相反, 数值为正的职业间收入差异主要是由两组人在性别、婚姻状况、教育和年龄等特征上的差异带来的。具体来说, 77% 的职业间收入差异可被这些特征所解释。

此外, 总体被解释部分的收入差异是正值 (0.2062), 未解释部分的收入差异为负值 (-0.0889)。这说明, 总的来说, 城镇工人比农民工收入更高是因为他们更可能拥有与高收入相关的一系列特征。我们的发现与孟昕、张俊森的研究结果正好相反, 很可能是因为他们的职业分类太宽泛 (他们仅仅把职业分成 4 大类), 以致他们误认的 (大类) 职业内差异其实在很大程度上是 (小类) 职业间的收入差异。

如上所述, 职业隔离是导致农民工与城镇工人之间收入差异的主要原因, 而职业隔离对收入差异的影响与农民工在性别、婚姻状况、教育和年龄等方面的特征分布密切相关。在这几个特征上, 我们认为教育是最重要的影响因素, 因为教育水平的差别几乎是所有不平等的根源。为了检验这一假设, 我们应用方差分析方法, 检验农民工身份对职业获得的影响, 以及这种影响在多大程度上来自于教育等其他因素的作用, 结果列于表 6。其中, 因变量是职业 ISEI 分值, 自变量包括农民工身份、教育程度、单位类型、性别、婚姻状况、年龄及其平方。自变量的编码方式与表 4 一样。与每一个自变量相对应, DF 表示自由度, R^2 表示单变量模型中职业 ISEI 分值已解释的方差百分比; $\Delta R^2 (1)$ 表示在控制了某个其他自变量后, 加入此变量所增加的已解释方差百分比; $\Delta R^2 (2)$ 表示在控制了所有其他自变量后, 加入此

变量所增加的已解释方差百分比。

表 6 职业 ISEI 被解释方差的百分比

变 量	DF	R ²	ΔR ² (1)		ΔR ² (2)
			控制以下单个变量		控制所有其他自变量
农民工	1	12.49 ^{***}	教育程度	2.98 ^{***}	0.72 ^{***}
			单位类型	3.05 ^{***}	
			性 别	12.51 ^{***}	
			婚姻状况	11.38 ^{***}	
			年龄及其平方	10.55 ^{***}	
教育程度	2	18.94 ^{***}			4.65 ^{***}
单位类型	2	21.83 ^{***}			7.32 ^{***}
性 别	1	0.46 ^{***}			0.45 ^{***}
婚姻状况	1	1.11 ^{***}			0.01 [*]
年龄及其平方	2	1.94 ^{***}			0.04 ^{***}

注: ^{***}p < 0.001, ^{**}p < 0.01, ^{*}p < 0.05, 基于 F 检验。

从表 6 可见, 在单变量模型中 (如 R² 所示), 农民工身份解释了整个样本中 ISEI 分值差异的 12.49%; 教育程度和单位类型分别解释了 18.94% 和 21.83%; 性别、婚姻状况和年龄 (及其平方) 的解释度都较低, 分别为 0.46%、1.11% 和 1.94%。当控制了所有其他自变量之后 (如 ΔR² (2) 所示), 每个自变量的净影响均显著下降; 但教育程度和单位类型的解释力远远大于其他自变量, 它们分别解释了 4.65% 和 7.32% 的 ISEI 分值差异, 而其他自变量的解释度都不足 1%。

农民工身份的解释度从单变量模型的 12.49% 下降到 0.72%。这说明, 农民工与城镇工人之间在职业获得上的差异, 主要是由教育等其他特征所带来的。但具体受哪些特征影响呢? 表 6 中的 ΔR² (1) 报告了在分别控制教育、单位、性别、婚姻状况和年龄后, 农民工身份对职业 ISEI 分值差异的解释力。结果显示, 教育程度和单位类型是导致农民工和城镇工人之间职业差异的主要因素。具体说来, 当控制了教育程度后, 农民工身份的解释度由单变量模型中的 12.49% 下降到 2.98%; 当控制了单位类型后, 农民工身份的解释度下降到 3.05%; 而分别控制性别、婚姻状况和年龄后, 农民工身份的解释度仍然超过 10%。在教育和单位两者之间, 前者对农民工与城镇工人之间的职业差异起着更加重要的作用, 因为一般情况下人们都是

先接受教育再进入某个单位从事某种职业。^①可以说,我国城乡之间不平衡的机会结构,尤其是教育机会的巨大不平等,是导致农民工在城镇劳动力市场上被隔离在低端职业的重要原因,这种职业隔离又成为农民工与城镇工人之间收入差异的主要原因。

(四) 户籍对城镇劳动力市场收入的因果效应: 倾向性评分匹配法分析

许多研究已经表明,倾向评分匹配法可以在因果推论中排除由其他可观测特征的影响所带来的偏差。倾向评分匹配法的主要优点是,研究者可以用单一的维度概括两个不同的群体(如农民工和城镇工人)在所有其他可观测特征上的差异,这个单一的维度就叫倾向评分。倾向评分指在给定可观测特征条件下,成为其中一个群体(实验组,如农民工)而不是另一个群体(控制组,如城镇工人)的条件概率。倾向评分匹配法的基本原理:首先,对两个研究组别分别估计倾向评分;其次,按照倾向评分的高低对这两个研究组别分别进行分层(即高低分组),并按倾向评分的分层进行匹配;最后,对每个层级内匹配好的样本计算这两组人在结果变量(如收入)上的差异,最终得到匹配后的平均差异,即实验效应。^②

在分析中,农民工相当于实验组,城镇当地工人相当于控制组。^③我们用以估计倾向评分的可观测特征包括教育年限、职业 ISEI 分值、年龄、婚姻状况和(县级层次的)居住地。由于农民工与城镇工人的工作时间存在很大差别,为了解释方便,这里用每小时工资收入作为结果变量,用月平均收入除以每月总的工作小时数而得。由前述分析得知,农民工身份的作用随不同的单位类型而变化,我们按党政机关/事业单位、国有/集体企业、私有部门分别进行倾向评分匹配分析。

首先,用上述那些可观测自变量估计一个二分类 logit 模型以得到倾向评分值,这个模型的因变量为农民工身份(农民工=1,城镇当地工人=0)。其次,按照倾向评分值对农民工和城镇工人进行一对一最近卡尺替换匹配(one-to-one nearest-neighbor caliper matching with replacement)。^④最后,通过比较农民工与可匹配的城镇工人间的小时工资收入差异,得到农民工身份对收入获得的因果效应。

表 7 列出了农民工与城镇工人小时工资收入差异的倾向评分匹配法分析结果。

① 当控制了教育后,单位类型对职业 ISEI 的影响显著下降,由单变量模型中的 21.83% 变为 9.55%。

② Paul R. Rosenbaum and Donald B. Rubin, "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score," *Journal of the American Statistical Association*, vol. 79, 1984, pp. 516-524.

③ Shenyang Guo and Mark W. Fraser, *Propensity Score Analysis: Statistical Methods and Applications*, Los Angeles: Sage, 2010.

④ Stephen L. Morgan and Christopher Winship, *Counterfactuals and Causal Inference*, New York: Cambridge University Press, 2007, pp. 107-109.

ATT (the average treatment effect on the treated) 是对实验组的平均处理效应, 等于农民工的实际的小时工资收入 (T) 减去他们被当成城镇工人时 (C) 的小时工资收入。^①

表 7 2005 年城镇地区农民工与当地工人收入差异的倾向评分匹配法结果 (按单位类型划分, 单位: 元/小时)

	农民工 (T)	城镇工人 (C)	ATT: (T) - (C)
总样本	4. 66	4. 38	0. 28 ^{***}
党政机关/事业单位	4. 67	5. 73	- 1. 06 ^{***}
国有/集体企业	5. 42	4. 92	0. 50 [*]
私有部门	4. 59	4. 10	0. 49 ^{***}

注: 显著性水平: *** $p < 0. 01$, ** $p < 0. 05$, * $p < 0. 10$ 。

在总样本中, 对实验组的平均处理效应 (ATT) 是正值, 且统计上显著。这表示, 给定两组人, 他们在基于可观测特征估计到的倾向评分上具可比性, 但一组是农民工, 另一组是城镇工人, 农民工的小时工资收入均值比城镇工人多 0. 28 元。按前述的描述性统计, 假定这两组人每个星期都工作 55 个小时, 那么农民工的月收入均值将比城镇工人多 62 ($0. 28 \times 55 \times 4$) 元。当把样本分成不同的单位类型时, 分析结果出现了有趣的变化。

在党政机关/事业单位中, 对实验组的平均处理效应 (ATT) 是负值。农民工每小时的工作收入比城镇工人平均低 1. 06 元。如前所述, 在党政机关/事业单位中农民工从事的大部分是低端的非技术工作。这里的结果表示, 如果把农民工与具有相似背景特征的城镇工人进行比较, 他们的收入仍然比城镇工人低, 但是收入差异比没有进行匹配时小很多, 因为我们排除了那些不可比的高端人群。匹配后农民工的收入比城镇工人少 18% ($-1. 06/5. 73$), 而 OLS 估计结果是 36%。这说明传统的回归模型估计结果确实会受样本分布不平衡问题的影响。

在国有/集体企业和私有部门中, 对实验组的平均处理效应 (ATT) 是正值且统计上显著。在这两类部门, 农民工的小时工资收入均值比城镇工人分别高 0. 50 元和 0. 49 元。假定这两组人每个星期都工作 55 个小时, 那么农民工的月收入均值将比城镇工人分别多 110 元 ($=0. 50 \times 55 \times 4$) 和 108 元 ($=0. 49 \times 55 \times 4$)。前述 OLS

① 倾向评分匹配法还能估计其他效应, 如平均实验效应 (ATE, the average treatment effect) 和对控制组的平均处理效应 (ATU, the average treatment effect on the untreated)。ATT、ATU 和 ATE 分别针对不同的研究假设。以本文的分析为例, ATT 是指, 假如农民工 (实验组) 是城镇工人 (控制组), 这两组人在平均收入上存在的差别; ATU 是指, 假如城镇工人是农民工时的情况。通常情况下, 倾向评分匹配法的估计结果指的是 ATT (Thomas A. DiPrete and Markus Gangl, “Assessing Bias in the Estimation of Causal Effects: Rosenbaum Bounds on Matching Estimators and Instrumental Variables Estimation with Imperfect Instruments,” *Sociological Methodology*, vol. 34, 2004, pp. 271-310)。

分析结果显示,在控制其他因素的作用下,农民工在这两类部门的平均月收入显著地低于城镇工人。这里的结果截然相反,假如把农民工与匹配后的城镇工人进行比较,农民工的收入反而更高。

倾向评分匹配法的分析结果再次证实了解析法的分析结果。当把可观测特征(在分解法中为职业获得)的显著差异排除后,农民工并不比城镇工人收入更低(党政机关/事业单位除外)。其实,当进一步把职业以外的其他特征差异排除后,倾向评分匹配法的分析结果提供了更加有力的证据,显示实际观察到的农民工相对城镇当地工人的低收入是由职业“准入”的结构性障碍和其他机会结构的不平衡情况造成的,而不是同工不同酬。

五、总结和讨论

由农村流向城镇地区的人口迁移大潮已逐渐改变了中国传统的城乡二元社会结构,城镇的农民工构成了一个介于农民和城里人之间的新阶层。虽然农民工获得了从农村到城市的向上流动,但他们在城镇地区仍然面临着以户籍制度为基础的城乡身份(城镇当地户口与农村外来户口)隔离和社会排斥,导致其在经济地位上与城镇当地人口还有相当的差距。在这一背景下,本研究的一个重要关注是,农民工与城镇当地工人在劳动力市场中的收入不平等。

对 2005 年全国 1% 人口抽样调查数据的分析结果显示,农民工与城镇工人之间存在很大的收入不平等;而且农民工确实在城镇劳动力市场上遭遇职业隔离,这使他们很难进入经济回报较高的职业。为了比较职业隔离和同工不同酬对收入不平等的影响,我们利用布朗分解法,把农民工与城镇工人间的收入差异分解成职业间和职业内的差异。结果发现,这两组人的收入差异主要来自以户口性质为基础的职业隔离,而非同工不同酬。

城镇地区内部的职业结构障碍是导致农民工低收入的主要原因,而职业隔离对收入差异的影响与农民工在性别、婚姻状况、教育和年龄等方面的特征分布密切相关。在以上特征中,我们认为教育是最重要的影响因素,农民工普遍在农村接受教育且近 80% 的人教育程度在初中及以下,城乡地区间不平衡的机会结构,尤其是教育机会的巨大不平等,是导致农民工被隔离在低端职业从而影响其收入的重要原因。

由于农民工和城镇工人不仅在职业分布上,而且在教育等其他特征上也存在显著差异,进一步利用倾向评分匹配法检验户籍身份(农民工相对城镇工人)对每小时工资收入均值的影响。结果显示,当依据教育、职业 ISEI 分值、雇佣身份、居住地和其他人口特征把具有相似倾向评分的农民工和城镇工人进行匹配后,农民工只在党政机关/事业单位类型内收入比城镇工人低;在国有/集体企业和私有部门,农民工的收入比可匹配的城镇工人反而更高。

农民工在国有/集体企业和私有部门中比城镇工人收入更高可能是因为留在城里的农民工具有选择性。如当进城打工的农民工面临严重的健康问题时，他们通常会选择返回农村老家。因此，经过选择且留在城镇地区的农民工很可能比城镇当地居民更加健康，而健康与劳动力市场中的表现如职业收入密切相关。此外，其他未被测量的因素也可能导致农民工在相同条件下比城镇工人赚得更多，如农民工相对更具有雄心壮志，更愿意努力工作，因为他们不满于在农村生活的状况，并且竭力改变目前的处境，希望过上更好的生活。他们也可能把自己与城镇当地人进行比较，而比较的结果会使他们更加努力地工作以缩小与当地人的差异。

此外，是否进入党政机关/事业单位，对农民工和城镇工人来说具有不同的意义，这可能是导致农民工在国有/集体企业和私有部门中比城镇工人收入更高的另一个原因。在党政机关/事业单位，正式的职位（如办公室文员）与低端的体制外工作（如门卫或清洁工）可以看作是一个分割的“劳动力市场”；前者相当于正规部门，除工资较高还提供就业保障和福利待遇，后者相当于非正规部门，没有高工资、就业保障或福利待遇。这是由职业准入障碍和教育机会结构的不平等，而不是同工不同酬导致的。我们揭示的农民工与城镇工人在城镇劳动力市场中的职业隔离模式，虽然是基于2005年的全国抽样调查数据的分析结果，但是，对消除劳动力市场对农民工的歧视，促进农民工及其随迁家庭成员在流入地城市的社会经济融入，至今仍然具有政策启示意义。

事实上，近十年来，中国的城乡和区域劳动力流动的模式发生了许多新变化^①，然而农村劳动力人口向城市、特别是向沿海发达地区的城市流动的总态势没有改变，而户籍制度改革的滞后和流入地城市对农民工的社会排斥，使得农民工的边缘化问题依然突出，如何消除制度障碍，促进农民工的经济融合和市民化，依然是学者和社会各界的重要关注。^② 尽管中央和不少地方政府早在2003年就开始明确提出要取消对进城务工人员的就业限制和歧视，但在具体的操作程序和制度保障方面仍

① 如2003年6月20日国务院颁令废止《城市流浪乞讨人员收容遣送办法》，随后一些城市的收容遣送相关条例和制度也陆续被废止；2006年1月1日国家取消农业税政策正式实施。

② 张永丽、黄祖辉 《中国农村劳动力流动研究述评》，《中国农村观察》2008年第1期。

存在不少问题,许多城市现在仍然限制农民工的就业范围。^① 2011年国务院颁布的《关于积极稳妥推进户籍管理制度改革的通知》,进一步明确要求今后出台有关就业、义务教育、技能培训等政策措施,不与户口性质挂钩。^② 可见,实现对农民工的市民待遇道路依旧漫长。^③ 促进农民工在城镇劳动力市场上的经济与社会融入,不仅需要国家层面制定长远的相关政策,改善城乡间在教育机会等方面存在的结构不平衡状况,更要集中梳理城镇地区对外来人口/劳动力相关的歧视性政策措施,完善相关制度,消除针对农民工的准入障碍,打破劳动力市场的职业隔离,建立公平、开放、统一的劳动力市场,使农民工获得平等的就业权。

(责任编辑:刘亚秋 责任编审:冯小双)

-
- ① 如北京市在2003年宣布取消用人单位使用外地来京务工人员计划审批和岗位(工种)限制,但仍然要求办理《北京市外来人员就业证》。办理《就业证》需要提交各种证明材料,包括外地来京务工人员身份证、暂住证、婚育证、外埠外出人员就业登记卡、用人单位法人营业执照、社会保险缴纳证明、承办人员身份证等,实际上是对农村流动人口进城务工的准入限制(北京市人力资源和社会保障局《关于加强外地来京务工人员就业服务工作的通知》,2010年1月20日, <http://www.bjld.gov.cn/LDJAPP/search/searchdetail.jsp?no=14057>,2013年7月8日)。
- ② 《国务院办公厅关于积极稳妥推进户籍管理制度改革的通知》(国办发〔2011〕9号),2012年2月23日, http://www.gov.cn/zwggk/2012-02/23/content_2075082.htm,2014年2月19日。
- ③ 王春光《农村流动人口的国民待遇与社会公正问题》,《乡音》2004年第6期。