

# 流动人口享受工资溢价了吗？<sup>\*</sup>

## ——对户籍来源地“反向歧视之谜”的再检验

刘金东 秦子洋 孔培嘉

**摘要：**近年来劳动力市场出现了外地人收入显著高于本地人的现象，这与传统的户籍地歧视观点恰好相反，被研究者称为“反向歧视之谜”。该谜题在中国是否真实存在，以及劳动力市场为什么会出现外地人工资溢价的现象，都亟待得到有效检验。本文利用2012年和2013年中国综合社会调查(CGSS)数据，将人口流动划分为非向上流动和向上流动两种类型，在控制外地人和本地人特征变量一致的情况下，基于处理效应模型，两种流动类型均不存在显著的“反向歧视”问题，考虑内生性和使用国家卫健委流动人口动态监测调查(CMDS)数据嵌套等的稳健性检验也支持了本文观点。进一步利用Oaxaca-Blinder分解和Theil嵌套分解等的结果表明，外地人工资溢价假象不仅仅是自选择效应所致，其中还有被选择的因素，特别是针对外地人的所有制歧视迫使生产力偏高的外地人进入非国有单位与生产力较低的本地人竞争，造成了这一被动性结果。本文结论不仅没有支持外地人在工资收入上被优待的观点，反而证明了外地人仍然受歧视的客观事实。

**关键词：**流动人口 工资溢价 户籍来源地 反向歧视

### 一、引言与文献综述

随着市场化改革的推进，劳动人口的跨区域流动成为常态。国家统计局第六次人口普查数据显示，全国范围内现住地和户口登记地相分离的人口约为2.61亿，剔除市辖区内人户分离的部分，仍有2.21亿流动人口。从微观来看，跨地区迁移流动是劳动者寻求更好就业机会、实现人的全面发展的重要途径(孙三百等,2012)。从宏观而言，人口流动有助于优化资源配置效率，进而提升全要素生产率(王金营、李竞博,2016)。但由于体制问题，现实中仍然存在很多固有的障碍因素限制了外地人口的市民化进程，使得外来劳动者无法在本地市场获得公平的就业机会和收入。在2011年之前的研究中，国内外学者一般认为存在对外地户籍就业者的双重歧视：一是特定部门的就业机会歧视；二是“同工不同酬”的收入歧视，造成外地人的平均收入偏低(王美艳,2005；邓曲恒,2007)。但以2011年为界，关于中国流动人口相对收入差距的问题出现了明显的研究拐点。在此之前，针对流动人口相对收入差距的研究大多围绕农民工群体，故而容易混淆针对户籍性质的歧视和针对户籍来源地的歧视。随着研究的深入，研究者开始将“户籍性质歧视”和“户籍来源地歧视”相剥离，认为以往的研究没有有效区分两种不同类型的户籍歧视，对外来农民工的收入歧视到底是因为农村户口性质还是因为他们来自外地，并没有得到直接的验证。从2011年开始，当研究者控制了户籍性质变量后，发现户籍来源地歧视不仅消失了，反而出现了“外地人平均收入偏高”的颠覆性结论，例如，章元和王昊

<sup>\*</sup> 刘金东，山东财经大学财政税务学院，邮政编码：250014，电子邮箱：kuangzhu1990@163.com；秦子洋、孔培嘉，山东大学经济学院，邮政编码：250100，电子邮箱：q19950906@sina.com，kongpeijia@163.com。本文受国家自然科学基金青年项目“家庭债务，需求不足与经济增长——基于家庭流动性约束机制探讨”(71803118)和教育部人文社科青年基金“体制内就业偏好的形成、影响及干预研究”(19YJC790081)资助。感谢匿名审稿人的修改建议，文责自负。

(2011)基于2005年上海市人口普查数据的研究发现,剔除对农村户籍性质的歧视以外,外地工人获得了10%的工资溢价;徐凤辉和赵忠(2014)利用2012年度涵盖北京、齐齐哈尔、长春、济南等10个城市的“中国雇员—雇主匹配数据”的实证结果显示,外地人比本地人工资高出16%;Kuhn & Shen(2015)基于厦门人才网的微观招聘数据发现,外地求职者相比本地求职者总能得到更高的雇主应答率,即相比本地人,雇主更加青睐外地人。以上研究只是基于部分大中城市的研究结果,陈昊等(2017)基于范围更大的中国综合社会调查(CGSS)微观数据也验证了外地人收入更高的事实,并将其称为户籍来源地的“反向歧视之谜”。

为什么会出 现户籍来源地的反向歧视?截至目前共有五种观点:(1)被迫创业说。宁光杰(2012)认为,收入水平排序上长期工优于自雇创业,自雇创业优于短期工,当外地人在本地劳动力市场无法获得较好的长期工就业机会时,只能被迫选择次优的自雇创业,创业成功并留下来的外地人抬高了外地人群体的平均收入水平。(2)自选择效应说。徐凤辉和赵忠(2014)以及Kuhn & Shen(2015)的研究均认为,外地人的正向自选择动机使得迁移到本地的外地就业者往往表现出更高的生产力优势,故而能够获得优于本地人的薪资收入。(3)需求效应说。章元和王昊(2011)推测,能够获得来沪工作机会的非农就业人员多为本地短缺的劳动力和人才,故而他们能够享受到工资溢价。(4)收入补偿说。陈昊等(2017)在Kuhn & Shen(2015)的推理基础上认为,深层次的内籍障碍造成外地人很难获得与本地人平等的社会保障待遇,故而雇主倾向于通过工资溢价来补偿社会保障的缺失。(5)社会资本说。亲友圈子组建的社会网络能够让就业信息更加充分,也让就业机会更加突出,国内外部分研究发现社会资本水平与个体收入水平呈现正相关关系(Simon & Warner, 1992; 叶静怡、武玲蔚, 2014)。考虑到户籍来源地的“反向歧视”针对的是基于雇佣的劳动力市场,如果将“反向歧视”的研究视角限定于劳动性收入,那么自营收入偏高并不能被归于其中,故而“被迫创业说”无法从根本上解释当前外地就业者工资偏高的现象。根据已有研究,本文首先将“反向歧视”界定为:在保持劳动生产率一致的情况下仅仅因为户籍来源地不同所导致的收入差异(章元、王昊, 2011)。这意味着,只有当控制所有可能影响劳动生产率的个体特征后,外地人的劳动收入依然显著高于本地人,才能说明劳动力市场存在“反向歧视”。以上除“被迫创业说”之外的四种观点都无一例外地强调了可能影响劳动生产率及其定价过程的潜在因素。在有效界定“反向歧视”内涵的前提下,标准研究范式应该致力于依次解决如下问题:外地人和本地人之间是否存在不能被个体特征解释的收入差异?如果不存在,那么表象上的外地人收入偏高事实因何而造成?现有的实证研究并没有完整地解答以上研究范式提出的问题。其他四种观点虽然一致否认了市场上存在对本地人的“反向歧视”,但其实证结果仍有不能被个体特征所解释的收入差异部分,只是基于推测的方式提出了可能的原因,故而在实证层面既没有完整求证第一个问题,也没有精确求解第二个问题。除了研究范式以外,已有的实证研究在变量设置上也存在不尽完善之处。

1. 对收入的界定不够准确。正如上文所述,“反向歧视”研究的是雇主对雇员户籍来源地的歧视问题,故而收入只能限定于劳动收入,自营收入、财产性收入等其他类型的收入不在检验范围之内,应予剔除。已有的研究如陈昊等(2017)并未明确这一处理细节,使得研究的合理性存疑。同时,反向歧视检验的不应该是总收入口径,使用小时收入口径作为被解释变量才能准确判定是否存在歧视(宁光杰, 2012)。已有的研究除了章元和王昊(2011)以外,均使用了绝对收入口径。实际上,工作时长是劳动时间,对应于供需曲线上的Q,小时收入是单位劳动时间上获得的对价,对应于供需曲线上的P,劳动供需决策应统筹考虑P和Q的关系,而非 $P \cdot Q$ 与Q的关系。基于此,本文以小时收入口径作为研究对象。

2. 控制自选择效应的变量不够充分。倾向得分匹配(PSM)等方法虽然有助于控制自选择偏误,但前提必须是加入与自选择相关的控制变量,故而变量完备是检验结果可信的基本前提。例如,既有研究均未控制社会资本问题,根据已有的研究,社会资本不仅是劳动者做出跨地区流动决策的重要考量因素(孔高文等, 2017),也是影响收入水平的重要因素(陈钊等, 2009; 叶静怡、武玲蔚,

2014)。当然,社会资本也并不必然带来工资溢价,Fang et al(2016)等文献的研究表明,通过熟人推荐获得工作机会的流动人口反而存在显著的工资折价问题,此时社会资本释放出一种“能力不够、关系来凑”的信号,同时阻断了劳动者议价的灵活性,从而给收入水平带来了负面影响。尤其值得一提的是,Kuhn & Shen(2015)使用的是厦门人才网招聘数据,恰恰厦门就是边燕杰和王文彬(2012)调研中社会资本总量最高的城市。社会资本不仅影响迁移决策、工作机会、收入水平,还能影响个体的行业去向,通过多种渠道影响个体劳动报酬率,故而有必要把它作为重要的个体特征纳入控制范围之内。

3. 没有考虑不同的流动类型。无论是 Borjas(1987)等研究者基于 Roy 模型的跨国迁移理论,还是当前国内人口流动的微观实证研究,大多数人口流动的方向都是从经济较差地区到经济较好的地区,以寻求更好的就业机会和更高的收入水平(孙三百等,2012;孟美侠等,2019),本文将将其称为“向上流动”。但是,也有少部分劳动者是从经济较好地区到经济较差地区工作,或者是在经济发展水平相当的地区间流动,我们将其统称为“非向上流动”,其动机很难等同于传统研究中的“向上流动”。如果不加区分统一回归的话,等同于强行认定两类流动人群的行为特征一致。将存在异质性的两类样本作同质性看待,会给研究结果带来偏误。本文按照户籍地和工作地的经济发展水平差异将流动人口精确划分为非向上流动和向上流动两种类型,从划分流动类型视角下对比“反向歧视”的检验结果是否会有所不同。

相比已有的研究,本文主要有三点研究创新:一是界定了不同流动类型,分非向上流动和向上流动两种流动人口样本分别检验“反向歧视之谜”,如果谜题不存在,则应该在每一种流动类型下均不能通过检验;二是以每小时获得的劳动性收入(下文简称“小时收入”)作为被解释变量,加入社会资本、高危职业等各种控制变量以充分控制自选择效应,提高了本文“反向歧视之谜”检验结果的可信性;三是本文研究不仅仅停留在利用处理效应模型以及倾向得分匹配(PSM)方法否定“反向歧视之谜”的存在性上,还进一步利用 Oaxaca-Blinder 分解和 Brown 分解等方法详细探究背后造成“反向歧视”表象的真实原因是什么。

## 二、数据说明、方法介绍与变量构造

### (一)数据来源及说明

本文所使用的数据来源于中国综合社会调查(CGSS)数据库,考虑到本文有必要控制产业虚拟变量和高危职业虚拟变量,2015年以及2012年以前年度的CGSS数据缺少详细的行业统计指标,故本文使用的是2012年、2013年连续两年的CGSS数据库样本,删除雇主、自营劳动者、家庭帮工和无职业者样本。本文按照户籍地和工作地的经济发展水平差异将流动人口划分为非向上流动和向上流动两种类型。虽然CGSS数据库提供了户籍地所在城市,但工作地只具体到所在省份,故而两种流动类型的划分无法具体到城市,只能按照户籍所在省份和工作所在省份的经济发展水平差异粗略划分。此处为每个流动人口样本匹配上户籍所在省份和工作所在省份过去10年的人均生产总值数据,2012年样本使用2002~2011年数据,2013年样本使用2003~2012年数据,数据整理自《中国统计年鉴》。由于过去10年有不少省份的人均生产总值互有高低,如果简单地以均值比较并不能清晰反映过去一段时期内两地经济水平的相对高低,此处以户籍所在省份和工作所在省份过去10年人均生产总值进行成对样本t检验,通过t检验结果来判定工作所在省份人均生产总值是否显著高于户籍所在省份。考虑到流动人口样本较少,为了尽可能保留样本,我们将户籍地城市为本省最发达城市的流动人口(即由最发达城市向省内其他城市流动的人口)作为省内非向上流动人口,与上述跨省非向上流动人口合并记为非向上流动类型。2012年和2013年CGSS流动人口中,非向上流动共计202人,向上流动共计417人。考虑到每一年非向上流动样本较少,我们将2012年和2013年截面样本合并在一起进行实证分析。为了让相邻年份的小时收入具有可比性,对2013年样本的小时收入按照当年居民消费价格指数进行平减以消除年份间价格因素的影响,得到最终混合截面数据的

小时收入。表 1 是本地人和外地人的收入与小时收入的计算对比结果,可以看到,两种流动类型下的外地人绝对收入平均值和小时收入平均值均高于本地人。此处只是纯粹的劳动性收入,剔除了自营收入,仍然在均值上显著占优,表明宁光杰(2012)提出的“被迫创业说”不是造成“反向歧视之谜”的主要原因。

表 1 小时收入的初步对比

年份	2012 年			2013 年		
	样本数	收入(元)	小时收入(元)	样本数	收入(元)	小时收入(元)
本地人	3293	30318.63	13.52	2836	35900.10	15.75
外地人	302	42658.87	20.26	317	53183.28	21.13
非向上流动外地人	114	36210.53	17.11	88	49023.86	19.53
向上流动外地人	188	46569.04	22.17	229	54781.66	21.75

## (二)处理效应模型

我们利用回归方法分析是否存在“反向歧视”现象,构建工资方程的基本形式如下:

$$Wage_i = \alpha + \beta \cdot Migration_i + \lambda \cdot \Theta_i + \mu_i + \varepsilon_i \quad (1)$$

其中, $Wage$  为关键的被解释变量“小时收入”, $Migration$  为“是否流动人口”,方程的控制变量见表 2, $\mu_i$  为个体工作所在地的省份固定效应, $\varepsilon_i$  为随机扰动项。考虑到  $Migration$  存在很强的自选择现象,并非随机抽样的结果,而是一个典型的虚拟内生变量。为了控制自选择偏误,我们参考 Chun & Lee(2001)的研究选择使用处理效应模型(treatment effect model),该模型对“是否结婚”的虚拟内生变量建模,以检验是否存在“结婚者收入更高”的婚姻溢价问题。我们的问题比较类似,虚拟内生变量为“是否流动人口”。处理效应模型的本质接近于 Heckit 两步法回归。第一步是基于 probit 回归的以“是否流动人口”(Migration)为被解释变量的选择方程,解释变量与工资方程一致,额外加入了“兄弟姐妹数量”作为工具变量。兄弟姐妹是就近照顾父母的代替者,兄弟姐妹数量与流动决策相关,但不会直接影响个体收入水平。第二步是工资方程回归,在加入式(1)中所有变量的同时,也加入选择方程 probit 回归基础上获得的逆米尔斯比率以矫正自选择偏误,此时看“是否流动人口”(Migration)的系数估计值是否显著,以判断是否存在“反向歧视”现象。

## (三)变量选取

除了被解释变量和核心解释变量之外,参考章元和王昊(2011)等文献,共选取以下 14 个方面的控制变量,见表 2。

表 2 变量含义

变量名称	变量含义	变量名称	变量含义
$Exp$	工作年限	$Capital$	社会资本
$Exp^2$	工作年限平方	$Time$	每周工作时长,以小时数计
$Gender$	性别,男性为 1	$Insurance$	是否参与社保,参与为 1
$Edu$	受教育年限	$Eddummy$	产业虚拟变量
$Statu$	政治面貌,中共党员为 1	$Danger$	是否高危职业,是为 1
$Marriage$	婚姻状况,有配偶为 1	$Ownership$	单位所有制性质,国有单位为 1
$Farmer$	户口类型,农业户口为 1	$Data12$	是否 2012 年数据,是为 1

考虑到社会资本能够通过多种渠道影响个体的劳动报酬率,本文将社会资本界定为嵌入社会网络中可以被利用的社会资源,参考徐超等(2017)的研究,从社会网络规模角度以亲友交际频次作为社会资本(Capital)的代理指标,具体来源于 CGSS 数据库的问题“过去一年,您是否常在空闲时间

与朋友聚会”,按频率由小到大赋值。每周工作时长(*Time*)存在极端值,对样本进行双边1%缩尾处理。产业虚拟变量(*Eddummy*)是根据国民经济行业分类,分别划分为第一产业、第二产业、第三产业,将党政机关和社会团体归为第四类,针对前三类产业分别构建虚拟变量。由于具有较大风险的高危职业可能需要更多的风险补偿,我们控制了“是否高危职业”(*Danger*),该指标以“是否高危行业”与“是否低级别工人”交乘得到。“是否高危行业”依据《国民经济行业分类》较高工伤风险的行业筛选得到,“是否低级别工人”取自CGSS访谈中“只受别人管理,不管理别人”的受雇者类型。考虑到以上变量不能控制的年份固定效应,我们也设置了样本年份虚拟变量(*Data12*),若数据来源于2012年则为1,反之则为0。

### 三、“反向歧视”的存在性检验

#### (一)基准回归结果分析

处理效应模型基准回归结果如表3所示,全样本、非向上流动和向上流动下的*Migration*的系数均不显著。这意味着,在控制内生性的情况下,无论是否将流动人口进行分类,外地人样本的工资水平均不高于本地人样本,不存在特别针对外地人的工资优待,“反向歧视之谜”被证伪。

表3 基准回归结果

	全样本	非向上流动	向上流动
<i>Migration</i>	-3.154 (-1.33)	-7.168 (-1.44)	-2.818 (-1.52)
<i>Exp</i>	0.250*** (4.30)	0.220*** (3.43)	0.299*** (4.09)
<i>Exp</i> <sup>2</sup>	-0.005*** (-4.68)	-0.004*** (-3.95)	-0.006*** (-4.25)
<i>Gender</i>	4.469*** (10.78)	4.838*** (10.31)	4.736*** (9.45)
<i>Edu</i>	1.164*** (14.03)	0.907*** (10.07)	1.406*** (13.52)
<i>Statu</i>	1.218** (2.04)	0.802 (1.20)	1.780** (2.42)
<i>Marriage</i>	2.065*** (3.55)	2.006*** (2.89)	2.637*** (3.79)
<i>Farmer</i>	0.975 (1.64)	0.050 (0.09)	1.915** (2.52)
<i>Capital</i>	-1.343*** (-5.81)	-1.205*** (-4.72)	-1.188*** (-4.07)
<i>Time</i>	-0.314*** (-25.67)	-0.309*** (-24.24)	-0.348*** (-20.88)
<i>Insurance</i>	0.737 (1.11)	0.856 (1.20)	0.560 (0.72)
<i>Eddummy1</i>	1.246 (0.79)	0.544 (0.33)	2.055 (0.92)
<i>Eddummy2</i>	5.589*** (6.67)	5.221*** (5.52)	6.184*** (5.75)

续表 3

	全样本	非向上流动	向上流动
<i>Eddummy3</i>	4.009*** (5.10)	4.008*** (4.49)	4.912*** (4.82)
<i>Danger</i>	-3.542*** (-5.98)	-3.777*** (-5.73)	-3.555*** (-5.00)
<i>Ownership</i>	-2.382*** (-4.94)	-2.249*** (-3.96)	-2.278*** (-3.95)
<i>Data12</i>	-1.822*** (-4.78)	-1.910*** (-4.48)	-2.102*** (-4.48)
逆米尔斯比率	3.306** (2.19)	6.545*** (2.74)	3.286*** (2.94)
省份固定效应	是	是	是
样本数	6229	4946	3936
adj R <sup>2</sup>	0.303	0.252	0.342

注:(1)\*、\*\*和\*\*\*分别表示在10%、5%和1%水平上显著,下同;(2)括号内为t值,下同。

从其他控制变量来看,工作年限(*Exp*)的符号为正,工作年限的平方( $Exp^2$ )符号为负,符合预期。性别(*Gender*)、受教育年限(*Edu*)、婚姻状况(*Marriage*)、政治面貌(*Statu*)、第二产业就业类型(*Eddummy2*)和第三产业就业类型(*Eddummy3*)都对小时收入有显著的正效应,社会资本(*Capital*)、工作时长(*Time*)、高危职业(*Danger*)、单位所有制性质(*Ownership*)都对小时收入有显著的负效应。其中,社会资本系数显著为负,表明社会资本并不利于劳动者提高收入水平,支持了Fang et al(2013)等文献的研究结论。高危职业系数显著为负,表明工资定价上并没有显著的风险补偿效应,反而越是高危职业的收入水平越低,这可能与该类职业的进入门槛偏低、劳动供给偏多的特征有关。单位所有制性质的系数显著为负,意味着越是国有单位,收入水平越低,这可能与国有单位特殊的薪酬体系有关。国有单位更接近于一种“共享式薪酬体系”,相当于同时设置了收入分布的上限和下限,“旱涝保收”的薪酬特征对低收入能力者非常有利,但对于高收入能力者施加了一个天花板,反而不利(Lang & Palacios, 2018; 刘冲、张红, 2020)。故而,国有单位的薪酬体系胜在方差上,而在均值上反而受损,这是国有单位性质显著降低小时收入的重要原因。另外,社保(*Insurance*)系数不显著,也表明陈昊等(2017)提出的收入补偿效应在这里并没有得到支持。逆米尔斯比率显著为正,意味着确实存在自选择偏误,且自选择偏误偏向于收入的提高。

## (二) 稳健性检验

本文已在基准估计部分加入了变量“是否农业户口”以控制可能的户籍性质歧视,但为了避免存在户籍性质歧视和户籍来源地歧视之间的混淆效应,本文进一步去除农业户口样本,仅对非农业户口样本进行分析。表4中稳健性检验1的结果显示,两种流动类型下,*Migration*的系数均不显著。由此可见,即使完全控制住户籍性质歧视,仍然不存在针对户籍来源地的“反向歧视”。

为了深入论证陈昊等(2017)提出的收入补偿说,本文进一步删除外地人中缴纳社保的劳动者,仅留下本地人全样本和外地人不缴纳社保的样本进行回归,实证结果见稳健性检验2。结果显示,两种流动类型下的*Migration*系数均不显著。这意味着,即使是完全没有参加社保、最该得到收入补偿的外地劳动者,其收入也并未显著高于本地人,这不仅证明了“反向歧视之谜”不存在,也证明了陈昊等(2017)的观点很难成立。实际上,严格的劳动保护并不能完全弥补市场机制带来的待遇差异,而只是推动待遇差异从合同保障程度差异向收入水平差异的形式转变(陈东、刘金东, 2014)。2008年新《劳动合同法》实施后,流动人口的社会保障待遇得到提高,而收入水平随之下降,故而依照“收

入补偿说”的逻辑,“反向歧视”在 2008 年之前要比 2008 年之后更加显著,这与陈昊等(2017)的实证结果恰好相反,也从侧面表明“收入补偿说”是站不住脚的。

表 4 稳健性检验(一)

	稳健性检验 1:仅考虑非农户籍样本		稳健性检验 2:仅考虑无社保外地人	
	非向上流动	向上流动	非向上流动	向上流动
<i>Migration</i>	-9.536 (-1.02)	7.942 (1.46)	2.154 (0.27)	-3.792 (-0.71)
逆米尔斯比率	9.032** (2.10)	-1.671 (-0.58)	2.322 (0.54)	4.598 (1.43)
控制变量	是	是	是	是
省份固定效应	是	是	是	是
样本数	2625	2092	462	383
adj R <sup>2</sup>	0.257	0.348	0.273	0.317

为防止控制变量“空闲时间是否与朋友聚会”按频次主观赋值不能完全代理社会资本,导致结果具有一定偏倚,本文进一步根据 CGSS 数据库问题“请问您与邻居进行社交娱乐活动的频繁程度”与“请问您与其他朋友进行社交娱乐活动的频繁程度”两者加总作为社会资本控制变量进行回归,实证结果见稳健性检验 3。其中,不同流动类型下 *Migration* 的系数仍然不显著。

由于部分本地劳动者可能是刚刚将户口迁入本地,因此,这部分劳动者的小时收入可能会受到其过去是外地人的影响,从而产生了户籍转换的内生性(宁光杰和段乐乐,2017)。为控制这种入籍内生性对估计结果的影响,我们分别删除过去 10 年内加入本地户籍的样本,重新做回归分析。稳健性检验 4 的结果进一步证明了本文的推论,两种流动类型下, *Migration* 的系数均不显著,意味着并不存在“反向歧视”。

表 5 稳健性检验(二)

	稳健性检验 3:替换社会资本指标		稳健性检验 4:考虑入籍内生性	
	非向上流动	向上流动	非向上流动	向上流动
<i>Migration</i>	-3.762 (-0.79)	-2.559 (-1.38)	-7.042 (-1.41)	-2.876 (-1.55)
逆米尔斯比率	4.858** (2.11)	3.025*** (2.71)	6.481*** (2.71)	3.329*** (2.98)
控制变量	是	是	是	是
省份固定效应	是	是	是	是
样本数	4946	3936	4932	3926
adj R <sup>2</sup>	0.249	0.339	0.251	0.342

考虑到 CGSS 数据库中外地人样本偏少可能影响结论的可信性,此处以对应年份的中国流动人口动态监测调查(CMDS)数据与 CGSS 本地人样本嵌套,进行稳健性检验。由于嵌套数据无法找到“是否流动人口”合适的工具变量,此处使用 PSM 方法来检验合并后的样本是否存在“反向歧视”现象。CMDS 数据调查指标与 CGSS 数据库不能完全对应,故而在协变量设置上做几处弱化处理。第一,考虑到 CMDS 数据库中不涉及政治面貌,故而 PSM 分析中舍弃该协变量。第二,CMDS 数据库仅有问题“您业余时间在本地和谁来往最多”,将回答为很少与人来往设为 0,其他设为 1;同理将 CGSS 数据库中“是否社交”问题的回答分为两类,回答“很少”或“从未”的样本设为 0,其他回答设为 1,统一以此指标作为社会资本协变量。第三,由于 CMDS 数据库中没有精确的员工岗位指标,我们

以“是否高危行业”来表征高危职业。第四,由于PSM方法不能像处理效应模型那样加入分省份虚拟变量以控制地区固定效应,此处为了控制不同地区经济文化等多方面差异,我们加入了工作地所在省份的人口迁移率作为协变量,该指标来自2010年人口普查数据。

按照本文的研究需要,将以本地人为对照组,以外地人为处理组,研究两组小时收入的平均处理效应ATT是否存在显著差异,ATT代表了纯粹因为外地人身份而显著有别于本地人的收入差异部分。由于不同的匹配方法有不同的优势,因此本文的基准回归部分选取近邻匹配、核匹配和局部线性回归匹配三种匹配方式。稳健性检验5是基于CMDS和CGSS嵌套数据的全样本进行的PSM分析结果,稳健性检验6是基于CMDS和CGSS嵌套数据的非农业户籍样本进行的PSM分析结果。两个稳健性检验的ATT均不能在5%的水平上显著为正,进一步证明了“反向歧视”确实不存在。

表6 稳健性检验(三)

匹配方式	稳健性检验5:CMDS嵌套数据全样本		稳健性检验6:CMDS嵌套数据非农样本	
	非向上流动	向上流动	非向上流动	向上流动
近邻匹配	-0.45 (-0.15)	0.19 (0.05)	-0.65 (-0.42)	1.34 (0.80)
核匹配	-0.94 (-0.65)	-2.28** (-2.24)	-0.73 (-0.63)	0.03 (0.03)
局部线性回归匹配	-3.82 (-1.29)	-3.11 (-0.79)	-1.89 (-1.23)	0.53 (0.37)

注:(1)由于合并CMDS外地人数据以后样本量达到3万多,故而设置近邻匹配 $k=1$ ;(2)鉴于篇幅所限,稳健性检验部分仅呈现最为关键的ATT估计结果。

#### 四、收入分解结果与分析

“反向歧视”的本质是雇主偏爱外地人而歧视本地人,上文已经否定了“反向歧视”的存在,即外地人在工资收入上并没有获得雇主优待。以上回归分析或者PSM方法都是“裁齐”外地人和本地人基本特征以对比二者之间是否存在因身份带来的显著收入差异,但现实中外地人和本地人的基本特征并不是完全相同的,这也是引发“反向歧视之谜”争论的重要原因。那么,接下来我们关心的问题是:到底是哪些特征变量的不同造成了“外地人收入高于本地人”的假象?换言之,外地人是因为具备哪些基本特征优势才让他们的收入高出本地人?从表3的回归结果中可以看到,有很多特征变量对收入有正效应,也有很多特征变量对收入有负效应。外地人工资收入高于本地人的假象可能是因为他们某些正效应特征变量上有异于本地人的优势,也可能是因为他们某些负效应特征变量上规避了本地人的劣势。这些特征变量中,既有就业者可以自选择的因素,也有诸如单位所有制性质、是否参与社保等被选择的因素。弄清楚以上问题就有了更深一层的意义:如果是自选择因素发挥作用,那么就徐凤辉和赵忠(2014)及Kuhn & Shen(2015)等文献提出的观点一致;但如果也有被选择的因素掺杂其中,那就意味着外地人不仅未“被偏爱”,还可能“被歧视”,而收入偏高的假象只不过是歧视下的被动结果而已。要弄清楚这些问题,我们还需要进一步的实证检验。

##### (一)Oaxaca-Blinder分解

Oaxaca-Blinder分解方法能够帮助我们从群体的组间差异之中分解出不同变量的绝对贡献值。该分解方法可以在回归方程的基础上将两个群组的组间收入差异分解为特征变量可解释部分和特征变量不可解释部分。假设A、B分别表示外地人和本地人,基本分解步骤如下:

$$\bar{W}_A - \bar{W}_B = \beta_A \bar{X}_A - \beta_B \bar{X}_B = (\bar{X}_A - \bar{X}_B) \beta_A + \bar{X}_B (\beta_A - \beta_B) \quad (2)$$

其中,第一部分 $(\bar{X}_A - \bar{X}_B) \beta_A$ 表示特征变量不同所导致的收入差异,属于组间收入差异的可解释部分,第二部分 $\bar{X}_B (\beta_A - \beta_B)$ 表示特征变量不可解释部分。我们分两种流动类型的外地人样本与

本地人样本互为对照组计算分解组间收入差异,考虑到我们仅是为了检验各个变量上的差异对组间收入差异的影响,故而表7中只罗列了特征变量可解释部分。结合本文研究需要,显著为正的即是对外地人工资溢价有正面贡献的特征变量,反之,则是对本地人工资溢价有正面贡献的特征变量。收入依然采用小时收入口径。

表7 Oaxaca-Blinder 分解结果

变量	非向上流动		向上流动	
	分解值	贡献率	分解值	贡献率
<i>Exp</i>	-1.257***	-34.79%	-2.059***	-34.40%
<i>Exp</i> <sup>2</sup>	1.010***	27.96%	1.646***	27.51%
<i>Gender</i>	0.065	1.81%	-0.010	-0.17%
<i>Edu</i>	-0.494	-13.68%	0.106	1.78%
<i>Status</i>	-0.066	-1.84%	-0.103	-1.72%
<i>Marriage</i>	-0.231***	-6.40%	-0.254***	-4.25%
<i>Farmer</i>	0.053	1.46%	0.009	0.15%
<i>Capital</i>	0.166*	4.60%	0.078	1.31%
<i>Time</i>	-1.483***	-41.06%	-0.494**	-8.25%
<i>Insurance</i>	-0.123*	-3.40%	-0.205**	-3.43%
<i>Eddummy1</i>	-0.015	-0.43%	-0.019	-0.33%
<i>Eddummy2</i>	-0.232	-6.42%	-0.310**	-5.18%
<i>Eddummy3</i>	0.447***	12.37%	0.576***	9.62%
<i>Danger</i>	0.084	2.33%	0.087	1.46%
<i>Ownership</i>	0.387***	10.71%	0.462***	7.72%
<i>Data12</i>	-0.051	-1.41%	0.113**	1.88%

注:已控制省份固定效应。

从表7中分解结果对比来看,工作年限平方(*Exp*<sup>2</sup>)、单位所有制性质(*Ownership*)、第三产业虚拟变量(*Eddummy3*)始终显著为正,提高了外地人的工资溢价,从而推动了“反向歧视”假象的形成。在非向上流动和向上流动下,三者合计能解释外地人平均工资溢价的51.04%和44.85%。社会资本(*Capital*)在非向上流动中也显著提高了外地人工资溢价。

1. 产业就业分布对外地人工资溢价的影响。虽然第三产业虚拟变量起到正向影响,但在其他两类产业虚拟变量的抵消之下,产业分布的总影响并不突出。从表3回归结果可以看到,第二产业和第三产业虚拟变量都对收入有显著的正效应。根据CGSS统计显示,有61.71%的外地人从事第三产业,显著高于本地人46.86%的占比;有35.86%的外地人从事第二产业,低于本地人42.09%的占比。外地人虽然在第三产业中占比较多,抬高了他们相比本地人的工资优势,但在第二产业中占比偏少,又抵消掉了一部分工资优势。

2. 工作年限结构对外地人工资溢价的影响。虽然工作年限平方(*Exp*<sup>2</sup>)的影响为正且贡献率居首,但工作年限(*Exp*)本身影响为负,且大小完全抵消了工作年限平方的影响。根据劳动经济学理论,劳动者的收入水平存在职限效应,随着工作年限呈现先上升后下降的“倒U”型,每个人都存在达到收入顶峰的工作年限拐点。图1是外地人和本地人工作年限核密度分布对比,工作年限本身影响为负而工作年限平方影响为正,是因为选择跨地区流动的外地劳动者更为年轻,工作年限较短,故而使得工作年限较长的本地人在工作年限拐点前的爬坡区间占优,而在工作年限拐点后的下降区间居于劣势。优劣相抵后,工作年限因素对反向歧视问题没有产生实质性的影响。

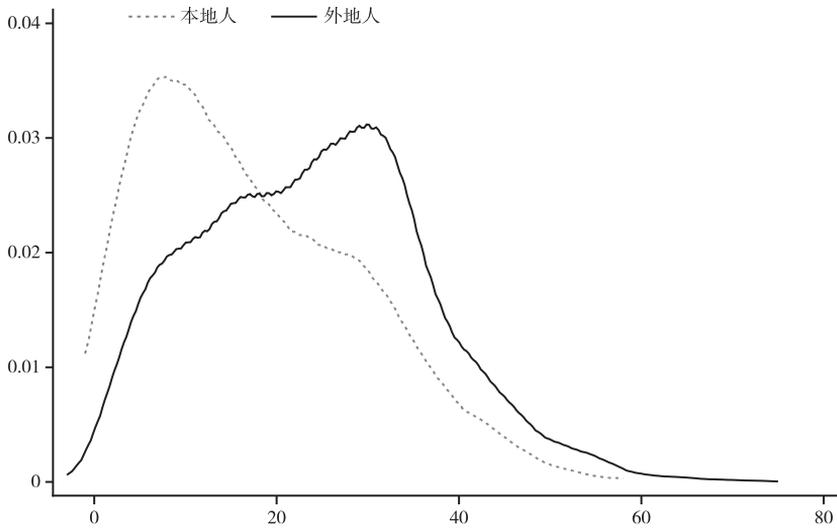


图1 工作年限核密度分布对比

3. 单位所有制性质对外地人工工资溢价的影响。单位所有制性质让外地人的小时收入相比本地人偏高,这一点与直观经验不符。国内文献如陈钊等(2009)的研究认为,存在针对外地人的所有制歧视,外地人更难进入本地的国有单位工作。2012年和2013年的CGSS统计结果也确实如此,有35.57%的本地人在国有单位工作,只有17.29%的外地人在国有单位工作。t统计检验显示,同在国有单位工作的外地人和本地人小时收入没有显著差别,国有单位本地人和非国有单位外地人的小时收入也没有显著差别,而同在非国有单位的外地人小时收入则要显著高过本地人,前者达到了19.92元,后者则为13.01元。这可能是劳动力市场自发筛选的结果,由于优秀的本地人更多涌入国有单位工作(王伟同等,2019),使得进入非国有单位的本地人更多的是生产力偏低的劳动者,而做出迁移决策的外地人往往是更为年轻的高生产力劳动者。我们计算对比了不同群体的平均受教育年限,发现国有单位本地人的平均受教育年限达到13.14年,而非国有单位本地人仅为10.18年,外地人平均受教育年限为10.98年,虽然低于国有单位的本地人,但相比非国有单位的本地人依然占有优势。这种“田忌赛马式”的竞争结果就是外地人更有比较优势,获得的竞争性薪资水平也更高。以“共享式薪酬分配”为特征的国有单位本身收入差距就不大,无论是外地人还是本地人,同在国有单位工作的报酬水平很难拉开差距,而以“贡献式薪酬分配”为特征的非国有单位则不同,其收入方差(离散程度)要显著高于国有单位(邢春冰,2005;张义博,2012)。在完全市场化的竞争体系下,生产力偏高的外地人获得的收入回报能够迅速脱颖而出,拉开与本地人之间的差距。此处为了有一个更加直观的认识,我们利用Akita(2003)提出的Theil指数嵌套分解方法计算了国有单位和非国有单位的小时收入差异分布情况,其分解原理如下:

$$Theil = \sum_i \frac{Y_i}{Y} \log \frac{y_i}{y} + \sum_i \sum_j \frac{Y_{ij}}{Y} \log \frac{y_{ij}}{y_i} + \sum_i \sum_j \sum_k \frac{Y_{ijk}}{Y} \log \frac{y_{ijk}}{y_{ij}} \quad (3)$$

其中, $i, j, k$ 分别为省份、行业、个人的标号, $y$ 表示对应层次的小时收入均值, $Y$ 表示对应层次的小时收入总和。以上分解得到的三个部分依次为省份间差异、行业间差异和个体间差异,通过嵌套分解能够将区域(省份)间和行业间的固有差异剔除,只保留了纯劳动者个体间的小时收入差异。表8的嵌套分解结果显示,无论是2012年还是2013年,非国有单位的小时收入差异均要显著高过国有单位,特别是在剔除区域间和行业间的固有差异后,非国有单位劳动者个体间的小时收入差异达到了国有单位的两倍左右,反映了国有单位的收入分配更为均匀,而非国有单位收入分配则更为离散。在这种收入分配格局下,所有制进入歧视驱使生产力更高的外地人进入非国有单位与生产力偏低的本地人展开竞争,从而间接造成了“外地人收入高于本地人”的假象。换言之,不仅不存在倾

向于外地人的“反向歧视”，反而存在针对外地人身份的正向歧视，外地人收入偏高也只不过是所有制歧视带来的被动性占优假象而已。

表8 Theil 指数嵌套分解结果

年份	所有制性质	省份间差异	省份内行业间差异	行业内个体间差异	总泰尔指数
2012	国有	0.0612	0.1070	0.1483	0.3165
	非国有	0.0917	0.1343	0.2948	0.5207
2013	国有	0.0655	0.1107	0.1270	0.3031
	非国有	0.0812	0.1394	0.2144	0.4350

4. 社会资本对外地人工资溢价的影响。分流动类型来看，社会资本的绝对贡献值在非向上流动中显著为正，能使外地人的小时收入溢价 0.166 元。此处，社会资本之所以推动了外地人工资溢价假象，是因为其对收入整体水平呈显著的负效应（见表 3），而外地人虽然可能是基于偏高于他人的社会资本做出迁移决策，但其社会资本总量相比本地人而言仍然处于偏低水平。CGSS 样本统计显示，本地人社会资本指标均值为 3.369，外地人均值为 3.307，其中非向上流动外地人均值更低，约为 3.228。偏低的社会资本水平使得外地人规避了社会资本带来的减收效应，故而相比本地人表现出一定的收入优势。社会资本虽然对外地人工资溢价表现出一定的正影响，但其作用机制与“社会资本说”仍有显著不同，其背后是一种“负负得正”的逻辑，而不是外地人社会资本偏高引致收入偏高的逻辑。这一作用逻辑与单位所有制性质比较相近，两种因素都对收入有负效应，外地人限于身份认同和社会资本不足，无法凭借亲友圈子和本地身份获得就业机会，反而规避了两种因素带来的减收效应。无论是社会资本因素还是国有单位招聘的本地人身份偏好，都是非市场因素在发生作用，外地人在两个因素均不占优势的情况下更多依靠于自身能力参与市场竞争，反而获得了更高的收入水平，这其中既有自选择效应，也有“被选择”的影响。此前研究中，宁光杰（2012）提出的“被迫创业说”、章元和王昊（2011）提出的“需求效应说”都不同程度强调了“被选择”因素。章元和王昊（2011）认为，非农就业的外地人多为本地短缺的劳动力和人才，这种劳动力岗位的稀缺性和互补性特征让他们能够享受到工资溢价，这一观点在本文中没有得到有效的支持，虽然外地人更年轻，产业分布结构也有所不同，但上文显示两者并没有带来实质性的工资差距。本文所论证的“被选择”发生逻辑更加类似于宁光杰（2012）提出的“被迫创业说”，都是收入上的被动性占优。

(二) 稳健性检验：基于 Brown 分解

Oaxaca-Blinder 分解忽视了不同群体因部门分割导致的进入概率的差别，这种差别无法通过简单加入虚拟变量的方式解决（王湘红等，2016）。从上文可以看到，所有制带来的部门分割显著存在于外地人和本地人之间，虽然加入了单位所有制性质（Ownership）虚拟变量，但 Oaxaca-Blinder 分解没有考虑所有制性质与其他变量之间的交互效应，潜在扩大了收入差异的整体不可解释部分，也使得各个变量绝对贡献值估计结果可能存在偏差。Brown et al(1980)提出的 Brown 分解方法能够有针对性地控制所有制部门分割导致的进入歧视，从而让各个变量绝对贡献值的估计结果更为可靠。Brown 分解采用 logit 模型预测不同群组劳动者的无歧视部门选择机制，在收入回归方程基础上将组间收入差异分解为部门间和部门内两部分。

首先，使用 logit 模型，估算第  $i$  个个体进入部门  $j$  的概率：

$$P_{ij} = e^{\beta_j X_j} / \sum_{j=1}^J e^{\beta_j X_j} \tag{4}$$

其中， $X_j$  为影响个体进入部门  $j$  的变量组。

组间收入差异分解如下：

$$\bar{W}_A - \bar{W}_B = \sum_j P_j^B (\bar{W}_j^A - \bar{W}_j^B) + \sum_j \bar{W}_j^A (P_j^A - P_j^B) \tag{5}$$

$$\sum_j P_j^B (\bar{W}_j^A - \bar{W}_j^B) = \sum_j P_j^B (\bar{x}_j^A - \bar{x}_j^B) \beta_j^A + \sum_j P_j^B (\beta_j^A - \beta_j^B) \bar{x}_j^B \quad (6)$$

$$\sum_j \bar{W}_j^A (P_j^A - P_j^B) = \sum_j \bar{W}_j^A (P_j^A - \tilde{P}_j^B) + \sum_j \bar{W}_j^A (\tilde{P}_j^B - P_j^B) \quad (7)$$

其中,  $\tilde{P}_j^B$  表示以 B 群组平均条件按照 A 群组进入方程估算的部门 j 进入概率。式(6)为部门内收入差异的进一步分解,其中,第一部分表示特征变量可解释的部门内收入差异,第二部分表示特征变量不可解释的部门内收入差异。式(7)为部门间收入差异的进一步分解,其中第一部分表示部门进入歧视可解释的收入差异,第二部分表示部门进入歧视不可解释的收入差异。我们主要计算的是特征变量可解释部分  $\sum_j P_j^B (\bar{x}_j^A - \bar{x}_j^B) \beta_j^A$ , 对应于除所有制性质以外的其他特征变量可解释的组间收入差异,这一收入差异绝对贡献值是完全剔除了所有制歧视之后的结果。需要说明的是,因为外地人进入国有单位的人员较少,无法区分国有单位和非国有单位各自的回归系数  $\beta_j^A$ , 故而我们假定外地人在两类单位中的回归方程系数一致,对外地人全部样本不区分国有单位和非国有单位做回归得到统一的系数  $\beta^A$ 。计算结果见表 9 所示。

表 9 Brown 分解结果

变量	非向上流动	向上流动
<i>Exp</i>	-1.755***	-6.600***
<i>Exp</i> <sup>2</sup>	1.463***	4.884***
<i>Gender</i>	0.191	-0.013
<i>Edu</i>	0.195	1.393**
<i>Status</i>	-0.008	-0.122
<i>Marriage</i>	-0.383	0.008
<i>Farmer</i>	-0.182	-0.800
<i>Capital</i>	0.152***	0.014**
<i>Time</i>	-0.936***	-0.138***
<i>Insurance</i>	-0.062	0.024
<i>Eddummy1</i>	-0.004	-0.042
<i>Eddummy2</i>	0.054	-0.287
<i>Eddummy3</i>	0.797**	0.794*
<i>Danger</i>	-0.262	0.053
<i>Data12</i>	-0.035	0.084

注:(1)表中展示的是各个变量对组间小时收入差异的可解释部分,即绝对贡献值;(2)Brown 分解的过程不涉及统计量,此处只根据所涉系数显著性标注了星号;(3)已控制省份固定效应。

Brown 分解结果显示,在剔除所有制进入歧视因素后,各个变量对组间小时收入差异的绝对贡献值有所变化。工作年限(*Exp*)及其平方(*Exp*<sup>2</sup>)的绝对贡献值依然是相互抵消的态势。社会资本(*Capital*)和第三产业就业类型(*Eddummy3*)对外地人工工资溢价的贡献值依然显著为正。社会资本在向上流动中也变得显著,不再仅仅局限于非向上流动下显著为正。总体来看,Brown 分解的结果与 Oaxaca-Blinder 分解的结果基本一致,也说明了本文结论的稳健性。此处参考 Brown 分解部门歧视的计算原理,计算对比了  $P_j^B$  和  $\tilde{P}_j^B$  的差值,前者代表本地人实际进入国有单位的比例(实际为 35.57%),后者则代表了按照外地人进入国有单位的 logit 方程本地人进入国有单位的比例,发现后者严重偏低,按照非向上流动外地人口的 logit 回归方程系数代入本地人平均水平,得到本地人国有单位占比仅为 9.49%,按照向上流动外地人口的 logit 回归方程系数代入本地人平均水平,得到本地人国有单位占比为 23.71%。这意味着,国有单位对外地人设置的“门槛”更为严苛,同样条件下本地人能够获得优先进入的权利。

## 五、本文结论与政策启示

本文利用2012年和2013年的CGSS数据对近几年广受关注的“反向歧视之谜”进行了再检验。在加入充足的控制变量后,外地人和本地人组间小时收入差异变得不再显著,从而直接证伪了“反向歧视之谜”。此外,本文的基本研究结论还有如下几点:一是支持了以往研究“不存在主观歧视”的既有结论,同时在更完备的研究范式下证伪了“反向歧视之谜”,让该谜题的解决不再仅仅停留在推理猜测层面;二是利用Oaxaca-Blinder分解和Brown分解探究了到底是哪些特征变量的差异造成了外地人工资偏高的虚假表象,否定了陈昊等(2017)提出的“收入补偿说”,部分验证了徐凤辉和赵忠(2014)及Kuhn & Shen(2015)提出的自选效应,同时也发现了被选择因素在其中发挥了显著作用,表明外地人的工资溢价假象反而是缘于被歧视的结果。

本文的研究结论对当前政策制定有如下三点启示:首先,无论是“反向歧视”还是“正向歧视”,如果是真实存在的主观情感因素所致,都必然被市场机制所惩罚,歧视的存在只能说明市场化程度不高,尚需进一步推进市场化改革进程,让市场在资源配置中发挥基础性作用;其次,无论是“被选择”还是“自选”,都是阻碍劳动力市场化的重要因素,不仅损害了社会公平,也损害了要素配置效率,未来的市场化改革仍要致力于解决信息不对称、生活成本过高、外地人市民化进程缓慢的问题;最后,不同类型的人口流动的迁移动机不同,影响外地人收入水平的因素也有所不同,故而未来有必要分类施策,破除不同流动类型下的求职壁垒和收入增长的阻碍因素。

### 参考文献:

- 边燕杰 王文彬,2012:《跨体制社会资本及其收入回报》,《中国社会科学》第2期。
- 陈东 刘金东,2014:《劳动保护有助于缩小就业弱势群体的相对收入差距吗——以新〈劳动合同法〉的实施为例》,《财贸经济》第12期。
- 陈昊 赵春明 杨立强,2017:《户籍所在地“反向歧视之谜”:基于收入补偿的一个解释》,《世界经济》第5期。
- 陈钊 陆铭 佐藤宏,2009:《谁进入了高收入行业?——关系、户籍与生产率的作用》,《经济研究》第10期。
- 邓曲恒,2007:《城镇居民与流动人口的收入差异——基于Oaxaca-Blinder和Quantile方法的分解》,《中国人口科学》第2期。
- 孔高文 刘莎莎 孔东民,2017:《我们为何离开故乡?家庭社会资本、性别、能力与毕业生就业选择》,《经济学(季刊)》第2期。
- 刘冲 张红,2020:《偏低的公共部门工资与改革前景——来自微观视角的新证据》,《财经研究》第4期。
- 孟美侠 等,2019:《城市工资溢价:群聚、禀赋和集聚经济效应——基于近邻匹配法的估计》,《经济学(季刊)》第2期。
- 宁光杰,2012:《自我雇佣还是成为工资获得者?——中国农村外出劳动力的就业选择和收入差异》,《管理世界》第7期。
- 宁光杰 段乐乐,2017:《流动人口的创业选择与收入——户籍的作用及改革启示》,《经济学(季刊)》第2期。
- 孙三百 黄薇 洪俊杰,2012:《劳动力自由迁移为何如此重要?——基于代际收入流动的视角》,《经济研究》第5期。
- 王金营 李竞博,2016:《人口与经济增长关系的再检验——基于人口活跃度—经济模型的分析》,《中国人口科学》第3期。
- 王美艳,2005:《城市劳动力市场上的就业机会与工资差异——外来劳动力就业与报酬研究》,《中国社会科学》第5期。
- 王伟同 张玲 谢佳松,2019:《国企环境、体制身份与人口迁移》,《财经研究》第9期。
- 王湘红 曾耀 孙文凯,2016:《行业分割对性别工资差异的影响——基于CGSS数据的实证分析》,《经济动态》第1期。
- 邢春冰,2005:《不同所有制企业的工资决定机制考察》,《经济研究》第6期。
- 徐凤辉 赵忠,2014:《户籍制度和特征对企业特征对工资收入差距的影响研究》,《中国人民大学学报》第3期。
- 徐超 吴玲萍 孙文平,2017:《外出务工经历、社会资本与返乡农民工创业——来自CHIPS数据的证据》,《财经研究》第12期。
- 叶静怡 武玲蔚,2014:《社会资本与进城务工人员工资水平——资源测量与因果识别》,《经济学(季刊)》第4期。
- 张义博,2012:《公共部门与非公共部门收入差异的变迁》,《经济研究》第4期。
- 章元 王昊,2011:《城市劳动力市场上的户籍歧视与地域歧视:基于人口普查数据的研究》,《管理世界》第7期。
- Akita, T. (2003), “Decomposing regional income inequality in China and Indonesia using two-stage nested Theil decomposition method”, *Annals of Regional Science* 37(1):55-77.
- Borjas, G. J. (1987), “Self-selection and the earnings of immigrants”, *American Economic Review* 77(4):531-553.

- Chun, H. & I. Lee(2001), “Why do married men earn more: Productivity or marriage selection”, *Economic Inquiry* 39(2):307—319.
- Fang, T. et al(2016), “The use and impact of job search procedures by migrant workers in China”, *China Economic Review* 37:154—165.
- Kuhn, P. & K. Shen(2015), “Do employers prefer migrant workers? Evidence from a Chinese job board”, *IZA Journal of Labor Economics* 4(1):1—31.
- Lang, K. & M. D. Palacios(2018). “The determinants of teachers’ occupational choice”, NBER Working Paper, No. 24883.
- Simon, C. J. & J. T. Warner (1992), “Matchmaker, matchmaker: The effect of old boy networks on job match quality, earnings, and tenure”, *Journal of Labor Economics* 10(3):306—330.

## Do Migrants Enjoy Wage Premium?

— Re-examining the Puzzle of Reverse Discrimination

LIU Jindong<sup>1</sup> QIN Ziyang<sup>1</sup> KONG Peijia<sup>2</sup>

(1. Shandong University of Finance and Economics, Jinan, China;

2. Shandong University, Jinan, China)

**Abstract:** There has been a special phenomenon in the labor market that the income level of migrants is significantly higher than that of local workers. Contrary to the traditional view of discrimination based on household registration, it is characterized as a “reverse discrimination mystery” by researchers. Whether the puzzle really exists in China and why there is a wage premium in the labor market need to be examined with great care. Based on the CGSS data for the reference years of 2012 and 2013, this paper divides population flow into two types, upward flow and non-upward flow. Keeping the locals’ and non-locals’ characteristics at the same level under the treatment effect model (TEM), the results show that there is no significant “reverse discrimination” for these two types of flow. The robustness tests of tackling the problem of endogeneity and using the combined data with CMDS also support this conclusion. Furthermore, the Oaxaca-Blinder decomposition and Theil nested decomposition results show the wage-premium illusion is not only caused by the self-selection, but also by the effects of being selected. To be specific, the ownership discrimination forces the migrants with high productivity to enter the non-state-owned units to compete with the locals with low productivity, resulting in this passive result. The conclusion of this paper suggests that the migrants are still discriminated against rather than preferred.

**Keywords:** Population Flow; Wage Premium; Household Registration; Reverse Discrimination

(责任编辑:陈建青)

(校对:武鹏)