

流动人口收入性别差异的实证研究

——以苏浙沪三省(市)数据为依据

徐 慷 田林楠

(南京大学,江苏 南京 210023)

摘 要:通过嵌套模型、奥萨卡模型方法研究苏浙沪三省(市)流动人口的性别收入差异及各因素的贡献率,可以发现,流动人口中存在着显著的性别收入差异,其中,人力资本和职业的性别隔离是流动人口性别收入差异最重要的两个决定因素,人力资本中又以教育年限这一变量的贡献最大,教育上的不平等显著影响到男性流动人口和女性流动人口的收入差异。此外,性别收入差异中不可解释的部分占 75%,表明流动人口收入中可能存在着性别歧视。

关键词:流动人口;收入性别差异;职业性别隔离

中图分类号:C924.4 **文献标识码:**A **文章编号:**1002-6924(2015)5-034-039

DOI:10.13713/j.cnki.cssci.2015.05.006

改革开放以来,在城市化进程加快和户籍制度“社会屏蔽”功能不断弱化的背景下,出现了城乡之间、地区之间的大规模人口流动。《中国流动人口发展报告 2013》显示,及至 2012 年,我国流动人口数量已经达到 2.36 亿,相当于每六个人中就有一个是流动人口。^[1]这种大规模、跨地区的人口流动已经成为我国社会经济转型以及城市化进程中一种必然的发展趋势,同时流动人口也成为城镇非农职业职工乃至整个中国人口的重要组成部分,可以说流动人口的生存境遇是“中国经验”(Chinese Experience)和“中国体验”(Chinese Feeling)的关键内容。并且,改革开放不仅导致了大规模的人口迁移,而且带来了性别关系、性别角色观念和家庭关系的改变,流动使女性离开土地从事非农职业,性别关系和家庭关系的改变则在一定程度上使女性劳动者从家务中相对解放出来,与男性一起进入劳动力市场。

他们在劳动力市场中是什么样的关系?而性别收入差距正是劳动力市场中性别关系的最直接表现形式。^[2]因此,对流动人口性别收入差距的研究是对市场转型过程中性别关系状况的一种透视。

作为城镇非农职业职工重要组成部分的流动人口中是否存在着收入的性别差异?如果存在,性别收入差异程度如何?哪些因素能够影响流动人口的收入性别差异?本文试利用“2012 年全国流动人口动态监测调查”中苏浙沪三省(市)数据,采用嵌套模型、奥萨卡模型等方法对上述问题进行实证研究。

一、相关研究文献回顾

已有的对中国的性别收入差异的研究与倪志伟的市场转型理论密切相关,后者认为市场转型过程中效率优先的市场机制将逐步取代计划

基金项目:江苏高校哲学社会科学研究重大项目“促进江苏人口现代化的关键问题研究”(2012ZDAXM012)。

作者简介:徐慷,南京大学社会学院副教授,主要研究方向:人口社会学、人口经济学;田林楠,南京大学社会学院博士研究生,主要研究方向:人口社会学、情感社会学。

经济中由行政权力主导的再分配机制,^[3]因此人力资本将取代政治资本等成为决定收入的主要因素,在这一理论的影响下,关于性别收入差异的研究多集中在“市场转型与性别平等”这个领域。具体来说,存在着两种截然相反的实证解释,一种认为在改革之前,由于政府的干预和政策的保护和鼓励,在劳动就业领域具有较高的性别平等程度,改革之后市场机制(或人力资本)在劳动就业中发挥更大作用,传统文化中的性别歧视也在挣脱国家行政力量的干预后重新发挥作用,导致收入分配机制开始发生变化,收入的性别差异也愈益明显。另一种观点认为,实证证据表明改革以来性别收入差距并无变化,并因此否认市场化程度对性别收入差距的影响,因而也就无所谓改革前和改革后,也就不存在国家保护力量、传统文化性别歧视以及市场机制之间的交互作用对性别收入差距的影响。

为什么会出现这种截然相反的经验与发现?有学者认为是因为以往的研究忽视了性别—职业间关系这个影响性别收入分层的重要作用机制,因为在现代社会中,人们大多是通过某种职业来获得收入,而男女两性所从事的职业有明显的不同,也即存在着职业性别隔离。他们的研究表明,职业的性别隔离对国有部门收入分配有显著的负面效应,但对体制外的劳动力市场的收入分配以及性别的收入差异没有影响,相反,在体制外的非国有部门,决定收入性别差异的主要因素是男女个体在人力资本方面的差异。^[2]

以往的研究对于改革以来的性别收入差距已经形成了比较成熟的研究策略,得出了一系列经验性的结论,并通过人力资本理论、市场转型理论、传统文化的性别歧视、职业的性别隔离理论等对这些经验性结论做出了解释,但这些研究结论和理论解释都是针对城镇的非农职业职工而言。我们知道,改革开放以来,由于户籍制度的松动,劳动力开始在城乡和地区之间流动融合,因此在城镇非农职业劳动力市场中,如果以流动和户籍作为分类标准的话,有四类劳动力群体:①本地常住人口,城镇户口;②本地常住人口,农业户口;③流动人口,城镇户口;④流动人口,农业户口。而以往的研究并未详细区分这四类劳动力群体,只将其笼统地作为城镇非农职业职工进行研究,忽视了作为资源分配机制和途径的户籍制度对流动人口的各种限制和负面效应,

使得流动人口就业、收入、所嵌入的社会制度背景等方面与流入地的户籍职工有着相当大的差异,因此有必要把流动人口作为专门的研究对象进行研究。

二、数据、变量与方法

(一)数据

本研究所采用的数据来自原国家计生委2012年进行的全国流动人口动态监测调查,该调查每年开展一次,按照随机原则在全国31个省(区、市)和新疆生产建设兵团抽取样本点,抽样总体为在流入地居住一个月以上,非本区(县、市)户口的15—59周岁(即1952年6月至1997年5月间出生)流动人口,需要注意的是抽样总体中不包括调查时在车站、码头、机场、旅馆、医院等地点的流入人口。抽样方法为分层、多阶段、与规模成比例的PPS抽样的混合。2012年调查全国总样本量为15.9万人,其中上海1.5万人,浙江1万人,江苏8000人,最终上海有效问卷14056份,浙江9371份,江苏7491份,有效样本量合计30918,问卷有效回收率约为93.69%,符合要求。本研究即以江苏、浙江、上海三地数据的合并数据作为研究数据。之所以选择苏浙沪的样本作为研究对象是因为这三地构成了长三角经济圈,集中了全国45.2%的跨省流动人口,具有很强的典型性和广泛的代表性。

(二)变量

1. 因变量。本研究关心的是劳动力市场中流动人口收入的性别差异及影响因素,因此我们的因变量应该是流动人口通过在劳动力市场中就业而获得的收入。由于“2012年全国流动人口动态监测调查”所提供的是流动人口的月收入信息,因此我们以此作为本研究的因变量。按照学界的惯常做法,我们对收入变量取对数值,以便使其接近正态分布。

根据以往研究的经验以及流动人口的特征,影响流动人口收入性别差异的变量可以归为以下五类:人力资本要素、制度要素、结构要素、与流动本身有关的要素以及人际关系网络要素。

2. 人力资本变量:a. 性别。在模型中,我们将性别处理成虚拟变量,其中女性=1,男性=0;b. 年龄;c. 教育程度。根据学界通常的做法,我们将教育程度转换成相应的教育年限,使之变成定距变量;d. 总工龄与本地工龄。总工龄从流动

人口第一次离开县级户籍地开始算起,而本地工龄则用 2012 减去本次流入本地年份表示;e. 每周工作时间;f. 就业身份。分为雇员、雇主、自营劳动者、家庭帮工四类。

3. 制度变量:a. 户口性质,我们将其处理为虚拟变量,其中非农户口=1,农业户口=0;b. 是否持有有效暂住证,我们同样将其处理为虚拟变量,其中持有有效暂住证=1,没有持有有效暂住证=0。

4. 结构变量:a. 就业单位性质,其中国有部门=1,非国有部门=0;b. 职业的性别隔离。根据学界通用的做法,我们将职业的性别隔离操作化为职业的女性比例。这个变量直接从 2012 全国流动人口动态监测调查的数据中计算得到。

5. 流动变量:a. 流动范围,分为跨省流动、省内跨市、市内跨县三类;b. 流动模式,分为独自流动、仅夫妻两人流动、仅夫妻两人与孩子一起流动、其他模式四类;c. 流动人口的来源地,包括华北、东北、华东、中南、西南和西北。

6. 人际关系网络:在本研究中,我们通过业余时间与谁来往最多来测量,按照异质性从低到高分:很少与人来往、同乡(户口在老家)、其他外地人、同乡(户口在本地)、其他本地人。

(三)统计模型

对于性别收入差距的测量和比较,我们选择通过估计收入模型里的性别系数来测算。因此,我们使用嵌套(nested)模型的方式建立 8 个最小二乘法线性回归模型:首先是基准模型,只加入一个自变量性别,单独估计性别的效应,然后在接下来的 7 个模型中累计加入控制变量人力资本(其中剔除性别要素,因为性别已经单独作为一个变量纳入)、制度要素(即户口性质)、结构要素之一(即就业单位的性质)、结构要素之二(即职业的性别隔离)、流动要素之一(即流动范围和来源地)、流动要素之二(即流动模式)和人际关系网络(其中,最后一个模型称作完全模型)以观察性别系数的显著性情况及变化来确定各变量对收入性别差异的影响。为了更为精确地确定各变量对流动人口收入性别差异的贡献率,我们还采用了奥萨卡模型对各要素的影响进行分解。它把性别收入差距分成两个部分,一部分是由性别因素所导致的,也即未被各变量解释的部分,另一部分是由除性别之外的其他因素所导致的,也即被可观察到的因素解释的部分。其中,未被

解释的部分通常被称之为“工资歧视”。^[4]其公式为:

$$\ln W_m - \ln W_f = B_f(X_m - X_f) + \text{未被解释的部分} \quad (\text{公式 1})$$

其中,ln W 是月收入的自然对数,B 是估计收入的回归模型中各变量的系数,X 是回归模型中所有自变量的均值,下标 m 代表男性(male),f 代表女性(female)。因此 ln W_m - ln W_f 表示男性与女性收入的自然对数的差值,(X_m - X_f)表示变量 X 在男性样本中的均值与在女性样本中的均值的差值,而 B_f(B_m)则表示相应的变量 X 在男性样本的回归模型或女性样本回归模型中的系数,因为男性样本模型和女性样本模型中该系数可能不一致,因此学界通常的做法是采用两者的均值来计算。还有学者用所有样本的回归模型中的相应变量的系数来进行计算,^[2]公式为:

$$\ln W_m - \ln W_f = B_a(X_m - X_f) + \text{未被解释的部分} \quad (\text{公式 2})$$

此时,B_a 为所有样本回归模型中的系数,其他部分则与公式 1 完全一致。以下,我们将按照这两种方式同时计算,以公式 2 的计算结果来解释,以公式 1 的结果作为参照。

三、数据分析结果

(一)收入、人力资本、结构要素、制度要素以及流动要素的性别差异

1. 因变量的性别差异。表 1 报告了所有变量的均值在男女两性之间的差异及这种差异的显著性。数据显示,男女两性在月平均收入方面存在着显著差异,女性流动人口的月平均收入比男性少 768 元,是男性月平均收入的 80% 左右(2994.32/3762.58=0.7958)。

2. 人力资本变量的性别差异。男性流动人口的平均年龄要比女性流动人口大 1.39 岁,并且差异显著。在受教育年限方面,男女两性只有 0.37 年的差距。在总工龄和本地工龄方面,男性流动人口分别比女性流动人口显著地多出 1.23 年和 0.21 年。同时,表 1 显示在每周工作时间内,男女两性间不存在显著差异。在就业身份上方面,两性间存在着显著差异。从表 1 中可以看到,绝对多数的流动人口都属于雇员身份,其中属于雇员身份的男性流动人口的比例为 71%,女性为 74%,男性比女性流动人口少 3 个百分点;

在雇主和自营劳动者这两个身份中,男性比女性分别高出 2 个百分点和 3 个百分点;几乎没有男性流动人口属于家庭帮工的就业身份,而女性流动人口在这一就业身份上比男性高出 2 个百分点。总之,在就业身份男性流动人口比女性流动人口更占优势。综上可以看出,在人力资本方面男性流动人口总体而言稍占优势,无论是在工龄所代表的的工作经验、教育年限以及就业身份方面都比女性流动人口更有优势。

3. 制度变量的性别差异。在这方面,男女两性间存在着显著差异,并且女性流动人口要比男性流动人口更有优势,无论是在拥有非农户口方面还是持有有效暂住证方面。

4. 结构变量的性别差异。在表 1 中我们可以发现,男性所在职业的平均女性比例显著地低于女性所在职业的平均女性比例,差值约为 9 个百分点,这表明在流动人口中存在着一定程度的职业的性别隔离。在就业单位性质方面,男性比女性更多地国有部门,并且这种差异是显著的。因此,可以认为,在结构变量方面,男性流动人口相对于女性处于优势地位。

5. 流动变量与人力关系网络变量的性别差异。在流动范围上,无论男女都以跨省流动为主,各自有 88%和 89%的男性流动人口和女性流动人口属于跨省流动,在这一类型上男性比女性少 1 个百分点,而在市内跨县这一类型上男性又比女性多 1 个百分点。二者在省内跨市这一类型上没有显著差异。总体而言,女性的流动范围要略微大于男性。在流动模式方面,男女两性只在独立流动和仅夫妻两人流动这两个模式上有显著差异,其中,男性在独自流动方面要比女性高出 5 个百分点,而在仅夫妻两人一起流动方面女性又比男性多出 5 个百分点。在人际关系网络方面,男女两性也存在着显著差异,男性与户口在老家的同乡以及其他外地人来往更多,而女性与其他本地人来往更多,同时,在很少与人来往方面,女性比男性的占比更高,多出一个百分点。

(二)性别收入差异中的性别效应、职业性别隔离效应与流动模式效应

在表 2 中,我们可以发现,模型 1(基准模型)中的性别系数为 - 0.216,并且在 0.001 的水平上是显著的。在加入了教育年限、工龄等人力资本变量之后(模型 2)性别系数上升为 - 0.184,这表明约有 14.8%([(- 0.216+0.184)/-

表 1 所有变量的均值及其性别差异 a

变量	男性	女性	差异 ^a
月平均收入(元)	3762.58	2994.32	768.26***
月平均收入自然对数	8.06	7.85	0.21***
年龄	33.75	32.36	1.39***
教育年限	9.84	9.47	0.37***
总工龄	8.85	7.62	1.23***
本地工龄	5.13	4.92	0.21***
每周工作时间(小时)	58.33	58.07	0.26
就业身份			
雇员	0.71	0.74	- 0.03***
雇主	0.09	0.07	0.02***
自营劳动者	0.20	0.17	0.03***
家庭帮工	0.00	0.02	- 0.02***
户籍性质(非农业=1)	0.15	0.17	- 0.02***
是否持暂住证(是=1)	0.94	0.95	- 0.01***
职业的性别隔离(职业的女性比例)	0.40	0.49	- 0.09***
单位性质(国有部门=1)	0.06	0.05	0.01**
流动范围			
跨省流动	0.88	0.89	- 0.01*
省内跨市	0.09	0.09	0
市内跨县	0.03	0.02	0.01**
流动模式			
独自流动	0.30	0.25	0.05***
夫妻两人流动	0.25	0.30	- 0.05***
夫妻与子女一起流动	0.37	0.37	0
其他	0.08	0.08	0
户籍地			
华北	0.01	0.02	- 0.01
东北	0.01	0.02	- 0.01***
华东	0.62	0.61	0.01*
中南	0.18	0.18	0
西南	0.15	0.15	0
西北	0.02	0.03	- 0.01
业余时间向来往最多			
同乡(户口在本地)	0.09	0.09	0
同乡(户口在老家)	0.54	0.53	0.01**
其他本地人	0.13	0.15	- 0.02**
其他外人	0.13	0.12	0.01***
很少与人来往	0.10	0.11	- 0.01***
样本数量	15379	11995	——

注:a 差异即为表内男性的数值与相应女性的数值的差,对于定序变量而言,数值是指取值所占比例;* $p < 0.05$ (双尾 T-test 检验,下同);** $p < 0.01$;*** $p < 0.001$ 。

0.216=0.148])的性别效应来源于男女两性在人力资本变量上的差异,这也表明人力资本要素是流动人口收入性别差异的十分重要的决定因

素。而表 1 所显示的男女两性在除了每周工作时间之外的所有人力资本要素上都存在显著差异,并且总体而言男性流动人口比女性流动人口更占优势这一点正好能够说明上述性别系数的变化。但在模型 3 中,加入了制度变量之后,性别系数下降到 - 0.193,即性别不平等程度反而上升了,这是因为在拥有城市户籍和持有有效居住证方面女性流动人口比男性流动人口更占优势,也即制度变量方面的女性占优使得收入差异只能是性别因素所造成的。在加入了就业单位性质这一结构变量之后(模型 4),性别系数没有发生任何变化,这表明就业单位性质几乎不解释收入的性别差异。模型 5 显示,职业的性别隔离是流动人口收入性别差异的最主要来源之一,加入职业性别隔离这一结构变量之后,性别系数由 - 0.913 上升到 - 0.169,收入的性别效应与模型 4 相比下降了约 12.4%([(- 0.193+0.169)/- 0.193=0.124])。而模型 6、模型 7 和模型 8 显示,流动范围、来源地、流动模式以及人际关系网络等变量几乎不解释任何的性别收入差异,因为在加入这些变量之后,性别系数的变化非常微小。总之,在流动人口中的确存在着显著的收

表 2 各变量对性别效应的直接影响

	性别系数	T 值	决定系数(R ²)
模型 1(基准模型):性别(女=1)	-0.216	- 34.39***	0.041
模型 2:性别+人力资本变量(年龄、教育年限、总工龄、本地工龄、工作时间、就业身份)	-0.184	- 32.02***	0.233
模型 3:性别+人力资本+制度变量(户籍制度、是否持有居住证)	-0.193	- 33.61***	0.244
模型 4:性别+人力资本+制度变量+结构变量 1(就业单位性质)	-0.193	- 33.53***	0.245
模型 5:性别+人力资本+制度变量+结构变量 1+结构变量 2(职业的性别隔离)	-0.169	- 28.26***	0.251
模型 6:性别+人力资本+制度变量+结构变量+流动变量 1(流动范围、来源地)	-0.170	- 28.44***	0.254
模型 7:性别+人力资本+制度变量+结构变量+流动变量 1+流动变量 2(流动模式)	-0.168	- 28.17***	0.254
模型 8(完全模型):性别+人力资本+制度变量+结构变量+流动变量+人际关系网络(业余时间与谁来往最多)	-0.168	- 28.20***	0.257

入的性别差异,而很大一部分的这种性别的负面效应可以归因于人力资本要素和职业的性别隔离,而流动模式变量并不能对收入的性别效应做出有效解释。

表 3 的数据表明,人力资本、制度变量、结构变量、流动变量和人际关系网络可以解释大约 25.04%的性别收入差异,其中结构变量的贡献率最大,为 15.77%;人力资本的贡献率次之,为 12.37%;而制度变量、流动变量和人际关系网络对性别收入差异的贡献率为负,分别为 - 1.83%,- 0.94%和 - 0.33%,这表明人力资本和结构变量是解释流动人口性别收入差异的最主

表 3 各因素对收入性别差异的贡献率(%)

	使用所有样本系数计算 ^a	使用男性样本与女性样本模型系数的平均数计算 ^b
人力资本变量	12.37	13.34
年龄+年龄平方	-3.16	-2.65
教育年限	8.32	8.15
总工龄+总工龄平方	1.71	1.80
本地工龄+本地工龄平方	-0.19	-0.22
每周工作时间	0.14	0.15
就业身份	5.55	6.11
制度变量	- 1.83	- 1.84
户口性质(非农户口=1)	-1.59	- 1.61
是否持有居住证(是=1)	-0.24	- 0.23
结构变量	15.77	13.07
就业单位性质(国有部门=1)	-0.13	- 0.11
性别隔离(职业女性比)	15.90	13.18
流动变量	- 0.94	-0.82
流动范围	-0.59	-0.63
流动模式	0.33	0.50
来源地	-0.68	-0.69
人际关系网络(业余时间与谁来往最多)	-0.33	-0.33
被以上变量解释了性别差异	25.04	23.42
未被解释的性别差异	74.96	76.58
收入的性别差异(自然对数差)	0.21	0.21

注:a 百分比是根据公式 $B_m(X_m - X_f) \div 0.21 \times 100$ 计算得到;b 根据 $[B_m(X_m - X_f) + B_f(X_m - X_f)] \div 0.21 \times 100$ 计算得到。

要因素。但如果将这两类变量进一步分解,我们可以发现人力资本变量 12.37% 的贡献率主要是由教育年限和就业身份提供,这两个变量的贡献率分别为 8.32% 和 5.55%, 共计 13.87%, 而年龄十年龄平方和本地工龄+本地工龄平方则稀释了这二者的贡献。结构变量的情况与此相似,该变量的贡献率由职业的性别隔离提供,并且被就业单位性质减少了 0.13%。因此,我们可以总结,教育年限、就业身份、职业的性别隔离是解释流动人口性别收入差异的主要因素。对照按照男性样本系数和女性样本系数的均值所计算的贡献率,可以发现,虽然具体数字没有完全一致,但是各组变量的贡献率的结构和关系以及各组变量内部各次级变量之间的贡献率的结构和关系与按照所有样本的系数计算出来的结果完全一致,这也从另一个角度证明了我们的分析的可靠性或信度。并且,这种分析方式得出了与表 2 完全一样的结论。

四、结论与讨论

通过上述分析我们发现:

第一,在由流动人口构成的劳动力市场中存在明显的收入性别差异,在控制人力资本、制度、结构、流动以及人际关系网络等变量之后,女性流动人口的收入是男性的 84%,也就是说约有 16% 的收入差距是完全由性别因素引起的。通过比较我们可以发现,流动人口中的由性别因素导致的收入差异与整个城镇劳动力市场基本一致:李春玲、李实的研究表明,在中国东部地区,约有 14% 的性别收入差距完全是由性别导致的;^[5] 吴愈晓、吴晓刚的研究发现,在非国有部门,控制其他变量之后,约有 15% 的性别收入差异是由性别因素造成的。^[2] 因此,可以认为在由性别因素导致的性别收入差异方面不存在流动人口群体和城镇非农职工群体的差异,也就是说在流动人口群体中也存在着同样的性别歧视。

第二,人力资本和职业的性别隔离是流动人口性别收入差异中最重要的两个决定因素,分别解释了流动人口性别收入差异的 12% 和 16%

(详见表 3),也即流动人口的性别收入差异有很大一部分是由男女两性在人力资本和职业的性别隔离方面的差距所造成的。其中,人力资本中又以教育年限这一变量的贡献最大,男性流动人口和女性流动人口的收入差异约有 8% 归因于两性在教育上的不平等。

一项最新研究表明,2005 年以来中国的收入不平等已经达到了一个非常高的水平,2010 年中国的基尼系数已经达到了 0.53。^[6] 在这种收入不平等中,人们更多地关注城乡差距、地区差距以及阶层差距,强调社会阶层结构的“断裂”,但我们的研究表明性别收入差距也显著存在,并且流动人口的收入性别差异虽然有约 25% 左右可以被人力资本等因素所解释,但仍有 75% 的性别收入差距未被解释,因此在流动人口中可能存在着性别工资歧视。在这样的背景下,我们应该重新关注流动人口劳动力市场中的因性别以及粘附在性别之上的传统观念、结构歧视(如职业的性别隔离)等因素所造成的收入不平等,以期提高收入分配中的性别公平水平以及全社会的收入分配的公平程度。

参考文献:

- [1] 国家卫生和计划生育委员会流动人口司. 中国流动人口发展报告(2013)[Z]. 北京: 中国人口出版社, 2013: 19.
- [2] 吴愈晓, 吴晓刚. 城镇的职业性别隔离与收入分层[J]. 社会学研究, 2009(4): 88—111.
- [3] Nee, Victor. A Theory of Market Transition: From Redistribution to Markets in State Socialism [J]. American Sociological Review, 1989(54): 663—681.
- [4] Oaxaca, R. Male-Female Wage Differentials in Urban Labor Markets [J]. International Economic Review, 1973(14): 693—709.
- [5] 李春玲, 李实. 市场竞争还是性别歧视——收入性别差异夸大趋势及其原因解释[J]. 社会学研究, 2008(2): 94—117.
- [6] Xie, Yu & Xiang, Zhou. Income inequality in today's China [J]. Proceedings of the National Academy of Sciences of the United States of America, 2014(19): 6928—6933.

[责任编辑: 李桃]