

哪些女性进入了城市高收入行业

——基于CHIP数据的经验分析

鄯 姣 孟大虎

【摘要】利用中国居民收入分配课题组2002年、2013年和2018年城镇住户调查数据,研究了影响女性进入高收入行业的因素。统计分析显示,高收入行业中的女性占比在2002-2018年间有所下降。采用排序Probit估计进行计量分析发现,接受过高等教育、拥有更丰富的社会资本显著增加了女性进入高收入行业的概率,而年龄对女性进入高收入行业的影响从2002年的“年轻就是资本”,转变为2018年的倒U形曲线关系。农业户口对女性进入高收入行业的影响在2002年和2013年不显著,到2018年变得显著正相关。其原因可能有二:一是农业户口劳动力进入高收入行业的户籍壁垒在降低;二是这些女性是经过极度选拔和筛选后进入高收入行业的,这种极度选拔和筛选会带来收入补偿效应。政策启示是:如果要缩小中国劳动力市场上的行业性别差距,消除女性进入高收入行业的准入障碍是至关重要的。

【关键词】女性;高收入行业;城市

【作者简介】鄯姣(1990-),女,新疆五家渠人,首都经济贸易大学经济学院讲师,硕士研究生导师,博士,从事劳动经济学研究(北京 100070);孟大虎,北京师范大学《北京师范大学学报(社会科学版)》编辑部(北京 100875)。

【原文出处】《学术交流》(哈尔滨),2023.10.107~119

【基金项目】研究阐释党的十九届六中全会精神国家社会科学基金重大项目“实现更加充分、更高质量就业研究”(22ZDA094)。

劳动力市场上的性别工资差距一直是学术界研究的焦点问题,而对于劳动力市场上的女性而言,高收入行业的准入障碍便是导致行业性别差距的一个重要原因。早有研究表明,高收入行业中的男性比例远高于女性^[1],而且高收入行业中往往拥有更多高质量的劳动力。^[2]那么,那些已经进入高收入行业的女性具有怎样的特征?为什么她们能够跨进高收入行业的门槛?本文的研究目标是针对女性进入高收入行业的影响因素进行经验分析。

一、文献综述

行业性别差距是一个全世界广泛存在的现象,很早就受到了学者们的关注。国内外有关行

业性别差距的文献主要分为三支。一是关于行业性别差距的诠释与行业的分类。有研究认为行业性别差距可以分为水平差距和垂直差距两种形式,水平差距指男女比例在某一个行业中表现出来的特征与全行业的平均比例分配不一致,说明某一行业存在对男性或者女性的进入门槛;垂直差距指男性与女性在同一行业中的职位与薪资有差异,表明行业内部存在性别不平等。^[3]而在行业分类方面,学者们通常根据研究的需要,按照行业资本来源^[4]、收入水平^{[1][5]}、行业女性占比^[6]、中国城镇住户调查数据中1992-1996年标准等依据进行划分^[7]。

二是关于行业性别差距及其变化趋势的研

究。有研究发现,因为美国就业结构的转变,促使很多女性从事服务业工作^[8],而瑞士的很多女性却被限制在教育、医疗和餐饮等少数行业就业^[9]。国内研究发现,农林牧渔业中女性比例高于男性,建筑业、国家机关和社会团体中男性比例高于女性。^[10]还有研究发现,女性主要集中在收入水平和社会保障程度较低的女性聚集行业和偏女性行业,男性大多集中在中高收入行业。^[11]此外,农民工之间也存在行业性别差距,女性农民工较多集中在劳动密集型低薪行业。^[12]有关行业性别差距的研究表明,近年来行业间的性别差距在逐年增大。^[13]有研究对比了2010年与2000年的行业性别差距,发现女性所在的行业分布呈两极分化,部分传统行业的男性占比提高,垄断行业中女性占比明显下降。^[14]通过分析改革开放四十年来行业性别差距发现,行业性别差距程度总体有所增加,此外,研究还发现接受过高等教育的女性更易在高收入行业就业。^[14]

三是关于行业性别差距的影响因素研究。有研究指出,不但行业间存在性别差异,大多数行业内部也存在着比较严重的性别歧视。^{[15][16]}由于女性在进入高收入行业时面临着更高的进入壁垒,因此更易进入中低收入行业。^[5]但是还有研究发现,女性在进入最高和最低收入行业时均具有优势,尤其是女性的性别特征更利于其进入高收入服务行业,此外研究还发现,对比同样就业于高收入行业的女性和男性劳动者,女性劳动者的教育水平普遍高于男性。^[15]尽管如此,蔡昉和张车伟分析发现,职业天花板现象在各个行业中的女性就业者中表现得更为严峻。^[17]

已有文献为本研究的开展奠定了基础,存在的不足则留下了研究空间。首先,已有文献大都根据自己的研究需要对行业类别进行划分,划分标准并不精确;其次,现有研究均发现行业因素是导致性别工资差距的重要原因,但是并未对什么样的女性可以进入高收入行业获得高收入进行研

究。基于已有研究,本文探讨了女性进入高收入行业的影响因素,主要的贡献有二:其一,利用中国家庭收入调查2002年、2013年和2018年三期数据,借助普通最小二乘法(OLS)分别准确划分三期样本中的高收入行业、中等收入行业和低收入行业包括哪些具体细分行业。其二,运用排序Probit估计方法,分析影响女性进入高收入行业的具体因素,同时观察这些影响因素的变化特征,最后按照女性劳动者所在城市的人口规模不同进行划分,并作了异质性分析。

统计分析结果显示,第一,对比2002年、2013年和2018年中高收入行业、中等收入行业和低收入行业中女性的占比发现,高收入行业中的女性比例呈下降趋势,而中等收入行业中的女性占比在增加,低收入行业中的女性比例并没有发生明显变化;第二,高收入行业中受过高等教育的女性占比呈上升趋势,通过社会资本进入高收入行业的女性占比呈先增加后下降的趋势;第三,高收入行业中女性的工资水平表现出了持续增长的特征。

实证分析结果表明:第一,受教育程度越高的女性,进入高收入行业的概率也就越高;第二,2013年和2018年,拥有社会资本的女性具有更高的概率进入高收入行业;第三,年龄对女性进入高收入行业的影响发生了有趣的变化。2002年的估计结果表明,年轻的女性进入高收入行业的概率更高,而到了2018年,年龄与女性进入高收入行业的概率呈现出一个倒U形曲线的关系,即在倒U形曲线拐点的左侧,年龄与女性进入高收入行业的概率呈正相关关系,而在拐点右侧,则出现了相反的变化。可能的解释是高收入行业对女性准入的门槛提高了,尤其是对教育水平和工作经验的要求均有所增加,但是如果女性的年龄超过某一个临界值,也会降低其进入高收入行业的概率,最终表现出女性年龄与进入高收入行业概率间的倒U形曲线关系;第四,党员身份对女性进入高收

入行业呈显著的负相关关系,可能的原因是拥有党员身份的女性更偏好于选择相对稳定但收入一般的体制内工作;第五,农业户口对女性进入高收入行业的影响仅在2018年显著,且呈正相关关系,一方面,这表明农业户口劳动力进入高收入行业的户籍壁垒可能在降低。另一方面,这些女性是经过极度选拔和筛选后进入高收入行业的,这种极度选拔和筛选会带来收入补偿效应。^[18]

二、数据来源与描述性统计

本文使用的数据来源于中国家庭收入调查2002年、2013年和2018年的城镇住户和外来务工住户样本。CHIP数据主要用于追踪中国居民收入分配的动态情况,是目前关于全国性家庭收入情况的最好数据之一。其中2002年和2013年的样本分为城镇住户、外来务工住户和农村住户三类,而2018年的样本将之前的城镇住户和外来务工住户合并成统一的城镇住户调查,因此,我们将2002年和2013年中的城镇住户与外来务工住户样本合并,加上2018年的城镇住户样本,共形成三组反映城镇住户就业信息的样本数据。研究对象选取16—60周岁、就业身份是雇员且年度工资性收入大于100元的样本。劳动者分为三类群体:高收入行业从业者、中等收入行业从业者和低收入行业从业者。本文关注不同行业的收入水平,因此选取劳动者的年工资收入作为核心变量,包括基本工资加上社保福利折算,具体折算方法参考李实等的处理。^[19]为了保证跨期数据的可比性,本文以2002年为基期,利用省级物价指数进行平减。

为了将行业分为高收入、中等收入和低收入三类,我们首先需要确定具有超额收入的行业。借鉴陈钊等的方法^[1],在明瑟收入方程中加入反映个体特征和就业特征的控制变量,再控制行业虚拟变量,根据行业虚拟变量的符号和显著性判断某一行业属于高收入行业还是低收入行业,其余行业则为中等收入行业。需要说明的是,在明瑟

收入方程中放入行业虚拟变量时,我们参考了Haisken-DeNew和Schmidt^[20]、田柳等^[21]的测算方法,将所有行业的虚拟变量均放入明瑟收入方程中,同时对行业虚拟变量的估计系数施加一个线性限制,即各个行业的劳动者份额乘上相应的估计系数并求和,其结果应为零。作此处理后,行业虚拟变量的估计系数可以解释为在控制了其他决定收入的因素后,该行业劳动者与所有行业劳动者平均收入的差距。控制变量包括性别、年龄、年龄的平方、教育水平、婚姻状况、民族、政治面貌、户口类型、单位所有制类型、职业类型、劳动合同性质、就业地、城市规模、地域等。在此基础上,我们测算出了2002年、2013年和2018年的高收入行业(表1),由于产业结构的变化,不同年份的高收入行业有所差异,其中交通运输、仓储及邮电通信业、金融业、房地产业在三个样本年份中均为高收入行业,金融业的工资比所有行业的平均工资高出20%以上。

表2分别展示了2002年、2013年和2018年高收入行业、中等收入行业和低收入行业中的女性占比。可以看出,高收入行业中女性的占比总体呈下降趋势,从2002年的40.31%下降至2013年的33.30%,直至2018年的32.59%;中等收入行业的女性占比波动较大,从2002年的41.23%上升至2013年的60.03%,继而在2018年又下降至55.65%;低收入行业中女性的占比相对稳定,2002年至2018年均保持在48%左右。

表3展示了高收入行业中女性的主要特征。可以看出,2002年至2018年处于高收入行业中的女性平均年龄并没有太大的变化,均在37—38岁之间。而2002年至2018年行业中受过高等教育的女性占比呈上升趋势,从2002年的36.51%上升至2013年的37.28%,继而上升至2018年的40.80%,表明高收入行业对女性教育水平的要求更高了。从社会资本发挥的重要性来看,女性通过亲朋好友的介绍才进入高收入行业的占比在

表1 高收入行业估计结果及其工资溢价

年份	行业	系数	标准误	工资溢价/%
2002年	电力、燃气及水的生产和供应业	0.212***	0.032	23.6
	地质勘察业、水利管理业	0.121*	0.070	12.9
	交通运输、仓储及邮电通信业	0.081***	0.021	8.4
	金融保险业	0.201***	0.038	22.3
	房地产业	0.145***	0.054	15.6
	教育、文化艺术和广播电影电视业	0.064**	0.029	6.6
	科学研究和综合技术服务业	0.120**	0.055	12.7
2013年	采矿业	0.256***	0.042	29.2
	制造业	0.051**	0.020	5.2
	电力、燃气及水的生产和供应业	0.179***	0.047	19.6
	建筑业	0.110***	0.032	11.6
	交通运输、仓储和邮政业	0.095***	0.026	10.0
	信息传输、软件和信息技术服务业	0.082**	0.040	8.5
	金融业	0.271***	0.041	31.1
	房地产业	0.236***	0.064	26.6
2018年	制造业	0.046***	0.014	4.7
	建筑业	0.111***	0.019	11.7
	交通运输、仓储和邮政业	0.060***	0.021	6.2
	信息传输、软件和信息技术服务业	0.075***	0.028	7.8
	金融业	0.231***	0.031	26.0
	房地产业	0.152***	0.045	16.4
	科学研究和技术服务业	0.117**	0.054	12.4

注:1.*、**、***分别表示在10%、5%和1%水平上显著;2.工资溢价表示工资方程中行业虚拟变量的估计系数。

表2 不同类别收入行业中女性的占比(%)

行业分类	2002年	2013年	2018年
高收入行业	40.31	33.30	32.59
中等收入行业	41.23	60.03	55.65
低收入行业	48.21	46.67	49.18

表3 高收入行业中女性特征

特征变量	2002年	2013年	2018年
年龄/岁	37.71	37.46	37.73
受过高等教育的占比/%	36.51	37.28	40.80
依赖社会资本找工作的占比/%	8.85	49.68	39.49
已婚占比/%	83.58	84.42	83.65
年工资/元	13297.35	45908.60	62318.25
年工作时间/小时	2054	2266	2264
小时工资/元	7.12	22.25	29.74

注:受过高等教育指受过大专及以上的教育水平。

2002年最低,仅为8.85%,但是2013年却上升至49.68%,而2018年又回落至39.49%,表明2013年以后,女性进入高收入行业的途径更加公平了。此外,高收入行业中已婚女性的占比相对平稳,样本年份中均在84%左右。

高收入行业中女性的年工资水平在这一期间表现出了持续增长的特征,2002年至2018年期间,高收入行业中女性年工资均值上升了368.7%,2002年至2013年间上升了245.3%,2013年至2018年间增长速度要低一些,也达到了35.8%。高收入行业中女性的年工作时间在2002年至2013年有所上升,从2054小时上升至2266小时,2018年略微下降至2264小时。总体来看,高收入行业中女性的小时工资呈上升趋势,从2002年每小时7.12元上升至2018年每小时29.74元,增长了317.7%,表明女性的人力资本有较大提高。

三、基本估计策略与稳健性检验

(一)排序Probit估计的结果

根据以上分析可以将行业分为低收入行业、中等收入行业和高收入行业三类,因此本文通过构建排序Probit模型(ordered probit model)研究教育、社会资本、婚姻状况等个体特征对女性进入高收入行业的影响。具体模型如式(1)所示:

$$\text{Ind}_i^* = \alpha_0 + \beta_0 H_i + \beta_1 X_i + \varepsilon_i, \text{ 其中 } \text{Ind}_i = f(x) = \begin{cases} 1 & \text{if } \text{Ind}_i^* \leq \mu_1 \\ 2 & \text{if } \mu_1 < \text{Ind}_i^* \leq \mu_2 \\ 3 & \text{if } \mu_2 < \text{Ind}_i^* \leq \mu_3 \end{cases} \quad (1)$$

被解释变量 Ind_i^* 表示个体 i 的潜在进入行业, ε_i 为随机干扰项。重点关注的解释变量 H_i 为个体 i 的教育水平、社会资本、婚姻状况、年龄、民族、政治面貌、户籍类型,控制变量 X_i 为个体 i 的工作特征变量,为单位所有制类型、职业、劳动合同性质,此外,模型还控制了就业地城市规模和地域类型。

表4报告了整体样本排序Probit模型的估计结果。模型(1)和模型(2)报告了2002年的估计结果,模型(3)和模型(4)报告了2013年的估计结果,模型

(5)和模型(6)报告了2018年的估计结果,模型(1)、(3)、(5)仅加入了女性的个体特征变量,而模型(2)、(4)、(6)在加入女性个体特征的基础上还控制了女性的工作特征变量和所在城市规模与地域特征变量。

从整体回归来看,各主要变量系数显著,模型整体回归效果较好。2002年、2013年和2018年均显示出了教育对女性进入高收入行业的显著性,并随着教育程度的增加,女性进一步进入高收入行业的概率也得到提高。比较2002年、2013年和2018年中通过社会资本找工作这一变量发现,2002年,社会资本对女性进入高收入行业的影响并不显著,但是到了2013年和2018年社会资本开始显著影响女性找到高收入行业的工作了,而且这一现象在2013年表现得更加明显。

通过三个样本的比较发现,年龄对女性进入高收入行业的影响发生了有趣的变化。2002年的估计结果表明,年轻的女性进入高收入行业的概率更高,2013年年龄变量却不显著了,到了2018年,会发现年龄与女性进入高收入行业的概率呈现出一个倒U形曲线的关系,即在倒U形曲线拐点的左侧,年龄越大女性进入高收入行业的概率越高,但是在拐点的右侧,则出现了相反的变化,也就是女性年龄越大进入高收入行业的概率越低。从2002年到2018年年龄对女性进入高收入行业的变化特征来看,在早期,年轻对女性进入高收入行业是一个优势,但是随着时间推移,由于高收入行业对工作经验和高等教育水平的要求不断增加,最终表现出进入高收入行业的女性年龄呈增长趋势,但是也并非年龄越大越好。因此,当女性到达一个年龄峰值后,进入高收入行业的概率开始下降。样本年份中,婚姻状况对女性进入高收入行业的影响并不显著。

汉族仅在模型(1)和模型(3)中显著,且加入工作特征变量、城市规模和地域特征后,便不显著了,表明综合其他因素来看,民族状况对女性进入高收入行业的影响非常有限。此外,各年份均显示出党

表4

整体样本排序 Probit 估计结果^①

	2002年		2013年		2018年	
	模型(1)	模型(2)	模型(3)	模型(4)	模型(5)	模型(6)
初中	0.275** (0.155)	0.279 (0.172)	0.295*** (0.089)	0.268** (0.107)	0.090 (0.065)	0.060 (0.069)
高中	0.579*** (0.153)	0.358** (0.173)	0.149 (0.094)	0.185 (0.113)	0.315*** (0.072)	0.246*** (0.077)
中专技校	1.022*** (0.157)	0.531*** (0.183)	0.264*** (0.101)	0.404*** (0.122)	0.266*** (0.080)	0.207** (0.086)
大专及以上学历	1.320*** (0.155)	0.622*** (0.182)	0.021** (0.095)	0.206* (0.119)	0.217*** (0.071)	0.100** (0.080)
社会资本	0.070 (0.068)	0.112 (0.087)	0.221*** (0.040)	0.089* (0.049)	0.058** (0.033)	0.012* (0.035)
年龄	-0.036* (0.020)	-0.062*** (0.025)	0.003 (0.017)	-0.008 (0.020)	0.034** (0.014)	0.027** (0.015)
年龄平方	0.001** (0.000)	0.001** (0.000)	-0.017 (0.022)	0.002 (0.026)	-0.055*** (0.018)	-0.042** (0.019)
已婚	-0.006 (0.077)	-0.123 (0.102)	-0.063 (0.061)	-0.089 (0.069)	-0.047* (0.047)	-0.030 (0.050)
汉族	0.234** (0.097)	-0.150 (0.129)	0.230* (0.089)	0.144 (0.103)	0.092 (0.073)	0.005 (0.079)
党员	-0.054 (0.050)	-0.261*** (.067)	-0.490*** (0.057)	-0.189*** (0.068)	-0.448*** (0.047)	-0.295*** (0.051)
农业户口	-0.159 (0.353)	0.621 (0.428)	0.014 (0.054)	0.016 (0.064)	0.209*** (0.034)	0.124*** (0.036)
工作特征变量	否	是	否	是	否	是
城市规模	否	是	否	是	否	是
地域	否	是	否	是	否	是
R ²	0.0593	0.1233	0.0324	0.2679	0.0247	0.1195
样本	3717	2642	3661	3261	5754	5507

注:1.模型汇报了回归系数值和稳健标准误,在解释模型时使用边际效应值。2.*、**、***分别表示在10%、5%和1%水平上显著;3.括号内为标准误。

员对女性进入高收入行业呈显著负向影响,可能的原因是拥有党员身份的女性更愿意选择相对稳定但收入一般的体制内工作。有意思的是,农业户口对女性进入高收入行业的影响在2002年和2013年并不显著,但是到了2018年转而变得显著正相关了。其原因可能有二:一是农业户口劳动力进入高收入行业的户籍壁垒在降低,二是这些女性是经过极度选拔和筛选后进入高收入行业的,这种极度选

拔和筛选会带来收入补偿效应。^[18]

(二)稳健性检验

为了检验结果的稳健性,本文采用了两种方法:一是替换估计方法,将被解释变量替换为是否在高收入行业工作(是=1,否=0),并采用普通最小二乘法进行估计;二是替换解释变量,用教育年限替换教育水平离散组,用工作经验及其平方替换年龄及其平方。从估计结果可以看出,模型各变

量系数显著,主要变量估计结果与表4基本一致,这说明排序Probit估计结果是稳健的。

由于排序Probit估计的参数含义不直观,本文进一步分别以表4的模型(2)、(4)、(6)回归结果为基础,计算各个解释变量影响女性进入高收入行业的边际效应,计算当所有解释变量处于其均值水平,某一解释变量变动一个单位,女性进入各类行业的概率如何变化^②,如表5所示。总体来看,教育水平越高的女性进入高收入行业的概率越高,但是其余解释变量在不同年份表现出了不同的变化特征。

具体而言,在2002年和2013年,即便是教育

水平较低的女性,也有很大可能进入高收入行业,这说明教育并不是决定能否进入高收入行业的关键变量;而到了2018年,仅有受过较高教育水平的女性有更高的概率进入高收入行业,如接受过大专及以上教育的女性,进入高收入行业的概率在5%的显著性水平上提高3.1%,表明高收入行业对女性受教育水平的要求在逐步提高。

通过比较2002年、2013年和2018年中社会资本对女性进入高收入行业的边际效应来看,在2002年,社会资本变量对女性进入高收入行业的影响并不显著,但是2013年和2018年的结果显示,拥有社会资本的女性进入高收入行业的概率

表5 主要解释变量的边际效应

	2002年			2013年			2018年		
	低收入行业	中等收入行业	高收入行业	低收入行业	中等收入行业	高收入行业	低收入行业	中等收入行业	高收入行业
初中	-0.084* (0.049)	0.030 (0.054)	0.054* (0.030)	-0.068** (0.028)	0.006 (0.004)	0.062** (0.024)	-0.017 (0.019)	0.001 (0.002)	0.019 (0.021)
高中	-0.110** (0.049)	0.038** (0.020)	0.072*** (0.030)	-0.047 (0.029)	0.005 (0.004)	0.042* (0.025)	-0.073*** (0.022)	0.001 (0.002)	0.073*** (0.023)
中专技校	-0.167*** (0.053)	0.053*** (0.020)	0.114*** (0.034)	-0.100*** (0.031)	0.005 (0.004)	0.095*** (0.027)	-0.061** (0.025)	0.001 (0.002)	-0.062** (0.026)
大专及以上	-0.199*** (0.053)	0.059*** (0.020)	0.139*** (0.033)	-0.052 (0.030)	0.005 (0.004)	0.047* (0.026)	-0.028 (0.023)	0.002 (0.002)	0.031** (0.025)
社会资本	-0.036 (0.028)	0.009 (0.007)	0.027 (0.021)	-0.022* (0.012)	0.001 (0.001)	0.021* (0.012)	-0.003* (0.010)	-0.001 (0.000)	0.004* (0.011)
年龄	0.020** (0.008)	-0.005** (0.002)	-0.015** (0.006)	0.002 (0.005)	-0.000 (0.000)	-0.002 (0.005)	-0.008* (0.005)	-0.001 (0.000)	0.008* (0.005)
年龄平方	-0.002** (0.000)	0.006** (0.000)	0.001** (0.000)	-0.001 (0.006)	0.001 (0.000)	0.001 (0.006)	0.013** (0.006)	0.001 (0.000)	-0.013** (0.006)
已婚	0.039 (0.033)	-0.010 (0.008)	-0.029 (0.024)	0.022 (0.017)	-0.001 (0.001)	-0.021 (0.016)	0.009 (0.015)	0.001 (0.001)	0.009 (0.015)
汉族	0.048 (0.041)	-0.012 (0.011)	-0.036 (0.031)	-0.036 (0.025)	0.001 (0.001)	0.034 (0.024)	0.002 (0.023)	0.001 (0.000)	-0.002 (0.024)
党员	0.084*** (0.021)	-0.021*** (0.006)	-0.062 (0.016)	0.047*** (0.017)	-0.002* (.001)	-0.045*** (0.016)	0.088*** (0.015)	0.001 (0.001)	-0.088*** (0.015)
农业户口	-0.200 (0.138)	0.051 (0.035)	0.149 (0.103)	-0.004 (0.016)	0.001 (0.001)	0.004 (0.015)	-0.037*** (0.011)	-0.001 (0.001)	0.037*** (0.011)
样本	2642	2642	2642	3261	3261	3261	5507	5507	5507

注:*、**、***分别表示在10%、5%和1%水平上显著。

在10%的显著性水平上分别增加2.1%和0.4%。

观察女性的年龄变量发现,2002年,女性的年龄越小进入高收入行业的概率越高,如女性年龄长一岁在5%的显著性水平上进入高收入行业的概率下降1.5%,2013年,年龄对女性进入高收入行业的影响并不显著,但是到了2018年,女性的年龄与进入高收入行业的概率在10%的显著性水平上形成了倒U形曲线的关系,也就是说在倒U形曲线拐点的左侧,女性年龄越大,进入高收入行业的概率越高,在拐点右侧,则恰恰相反。

婚姻和民族状况对女性进入高收入行业并没有显著影响。女性如果是党员进入高收入行业的概率更低,这种现象在2013年和2018年表现得更加明显,例如,在2013年和2018年,相较于非党员女性,党员女性进入高收入行业的概率在1%显著性水平上分别下降4.5%和8.8%。

从户籍类型来看,2002年和2013年户口性质并不影响女性进入高收入行业,但是到了2018年却出现了农业户口的女性进入城市高收入行业的概率更高的情况,具体而言,拥有农业户口的女性进入高收入行业的概率在1%的显著性水平上高出3.7%。如上文分析所言,我们认为,在很大程度上,这并不能说明在农业户口和非农业户口的竞争中出现了“逆转”,这反而说明,我国的城市劳动力市场仍然在基于户籍身份进行选拔和筛选,只有那些能力更强的拥有农业户口的女性在经过极度选拔和筛选后,才能克服户籍身份的不利影响进入高收入行业。

四、内生性检验

女性进入高收入行业的影响因素很复杂,受到个体因素、家庭因素等多维度的影响。由于数据和方法上的限制,我们无法控制所有可能影响女性进入高收入行业的变量,因此模型中可能存在遗漏变量问题,从而使基准回归结果有偏,影响了模型估计的准确性和稳健性。本文采用工具变量法解决由于遗漏变量带来的内生性问题。

根据工具变量外生性和相关性的选择要求,在讨论教育水平对女性进入高收入行业的影响时,本文采用地区整体平均受教育水平作为女性受教育程度的工具变量。家庭对子女的教育投资往往具有同群效应,会受到周围整体的教育投入和教育水平的影响,这种地区层面的变量往往反映了当地的基本社会风气和开放程度,体现了人们对教育的重视和投入程度,而地区平均受教育水平又很难直接影响女性个体进入高收入行业的概率。考虑到工具变量模型中的识别问题以及教育分组过程存在一定的主观性,仅在受教育水平作为连续变量时的情况下使用工具变量,不再讨论教育细分为离散组时的情况。

考虑到行业分类赋值的有序性,因此本文采用Roodman^[22]提出的工具变量两阶段混合过程方法(two-stage mixed-process estimator)进行估计。已有文献表明,面对被解释变量是离散变量时,这种估计策略比普通两阶段回归方法更有效率。^[23-24]表6报告了基于工具变量方法的估计结果。^③

由表6可以看出,与基准回归估计结果相似,

表6 工具变量回归结果

变量	2002年	2013年	2018年
地区平均受教育水平	0.705**(0.054)	0.561***(0.026)	0.535***(0.019)
加入控制变量	是	是	是
通过弱工具变量检验	是	是	是
通过过度识别检验	是	是	是
观测值	2642	3261	5507

注:1.模型汇报了回归系数值和稳健标准误,在解释模型时使用边际效应值;2.*、**、***分别表示在10%、5%和1%水平上显著;3.地区平均受教育水平根据CHIP2002年、2013年和2018年全国数据计算得到。

工具变量估计结果依然显著且影响方向相同,由于控制了遗漏变量带来的内生性问题,在每个临界值处的边际效应值大小不同于基准回归估计。具体而言,2002年、2013年和2018年,女性的受教育水平每增加1年,在1%的显著性水平上,女性进入高收入行业的概率分别增加3.0%、3.5%和3.5%。这表明受教育水平对女性进入高收入行业的影响力在增加。

五、异质性检验

我们按人口数量将城市分为中小城市、大城市、超大及特大城市三类,分样本进行回归分析。^④

第一,样本年份里受教育水平对女性进入高收入行业的影响显示出了相似的特征,即在中小城市,相比于小学及以下教育水平,高中、中专技校、大专及以上受教育水平均对女性进入高收入行业有显著的正向影响,但是在大城市、超大及特大城市中,仅有高等教育可以显著提高女性进入高收入行业的概率。

第二,社会资本只有在2002年的大城市样本中对女性进入高收入行业有显著正向影响,其余样本年份中均没有显著影响。

第三,年龄在2002年的大城市样本中对女性进入高收入行业呈显著负向影响,即年龄越小的女性进入高收入行业的概率越高。而到了2018年的中小城市、超大及特大城市中,女性的年龄与其进入高收入行业的概率却呈倒U形曲线的关系,即在拐点左侧,女性的年龄越大进入高收入行业的概率越高,在拐点的右侧则恰好相反,表明进入高收入行业对女性受教育水平、工作经验的要求都更高了。

第四,相对而言,婚姻和民族状况对女性进入高收入行业的影响并不显著,但是拥有党员身份的女性进入高收入行业的概率明显降低,原因仍然可能是我们上面所提到的,拥有党员身份的女性更倾向于进入稳定但收入不是很高的体制内工作。

第五,在2018年的中小城市、超大及特大城市中,农业户口的女性进入高收入行业的概率显著升高,但在其余年份并没有表现出这个特征,说明随着人口流动性的增强,更多有能力的外地劳动力进入城市,也正是因为她们的能力更强,因此进入高收入行业的概率更高。

六、结论与政策启示

本文使用2002年、2013年和2018年中国家庭收入调查的城镇数据,首先采用限制最小二乘法准确识别出样本年份高收入行业、中等收入行业和低收入行业具体包括哪些细分行业,继而运用排序Probit估计分析了女性劳动力进入高收入行业的影响因素,最后进行了稳健性检验及异质性检验。结论如下:首先,教育水平与社会资本对女性进入高收入行业具有显著正向影响,且这种现象在2013年和2018年表现得更加明显。其次,年龄对女性进入高收入行业的影响发生了有趣的变化,2002年“年轻就是资本”,到了2018年,年龄与女性进入高收入行业的概率呈倒U形曲线的关系,即在倒U形曲线拐点的左侧,年龄与女性进入高收入行业的概率呈正相关关系,在拐点右侧则恰恰相反。再次,样本年份中,党员身份对女性进入高收入行业的概率均呈显著的负相关关系,而农业户口在2018年中对女性进入高收入行业表现出了显著的正向影响。最后,通过城市规模估计发现,个体特征因素对女性进入高收入行业概率的影响在不同规模的城市中反映出了不同的结果。例如,大城市、超大及特大城市对女性教育水平的要求更高,社会资本仅在大城市中影响女性进入高收入行业的概率,年龄与女性进入高收入行业概率的倒U形曲线关系仅在中小城市、超大及特大城市中显著,与此相似,农业户口变量也仅在中小城市、超大及特大城市的子样本中显著。

本文的主要政策启示是,要致力于实现劳动力市场中的性别平等,就需要打破女性进入高收入行业的障碍,实施有利于行业性别收入均等化

的政策。一方面,高等教育的性别比已经发生了逆转^[25],即女性接受过高等教育的比例超过了男性,但是即便是受过高等教育的女性更多了,我们仍然发现,2002年至2018年,高收入行业中的女性占比在下降,表明应给予高学历女性平等入职的个人发展机会,尽快扭转高收入行业中女性占比下降的态势,促进高收入行业内部性别均衡发展。另一方面,应继续强化劳动力市场的规范化发展,弱化社会资本对女性求职的作用,促进行业和性别均衡发展。

注释:

- ①为了节约篇幅,该表未展示出其他控制变量的系数以及排序 Probit 模型的临界点。
- ②受篇幅所限,我们仅展示了个体特征变量的边际效应值。
- ③为了节约篇幅,该表未展示出其他控制变量的系数以及排序 Probit 估计的临界点。其他控制变量包括:社会资本、年龄、年龄的平方、婚姻状况、民族、政治面貌、户口类型、工作特征变量、城市规模、地域。
- ④为节约篇幅,本小节未展示异质性检验的估计结果。

参考文献:

- [1]陈钊,陆铭,佐藤宏.谁进入了高收入行业?——关系、户籍与生产率的作用[J].经济研究,2009(10):121-132.
- [2]MARTINS P S. Industry Wage Premia: Evidence from the Wage Distribution[J]. *Econometrics Letters*, 2004(2): 157-163.
- [3]魏真兰.我国行业性别隔离对性别工资差异的影响研究[D].长沙:湖南大学,2015:78-99.
- [4]岳希明,李实,史泰丽.垄断行业高收入问题探讨[J].中国社会科学,2010(3):77-93,221-222.
- [5]王美艳.中国城市劳动力市场上的性别工资差异[J].经济研究,2005(12):35-44.
- [6]蒋永萍,杨慧.妇女的经济地位[M].北京:中国妇女出版社,2013:101-125.
- [7]葛玉好.部门选择对工资性别差距的影响:1988-2001年[J].经济学(季刊),2007(2):607-628.
- [8]RENDALL M. Brain versus Brawn: the Realization of Women's Comparative Advantage[J]. IEW- Working Papers, 2010, 491.
- [9]IRENE K, MARLIS B, STEFAN S. Variation in Job Opportunities for Men and Women in the Swiss Labor Market 1962-1989[J]. *Research in Social Stratification and Mobility*, 2010, (28): 309-323.
- [10]谭琳,卜文波.中国在业人口职业、行业性别隔离状况及其成因[J].妇女研究论丛,1995(1):24-28.
- [11]杨慧.新世纪女性行业结构与性别隔离——基于对第三期中国妇女社会地位调查数据分析[C]//2013年中国社会学年会“性别发展与美丽中国建设”论坛文集,2013:105-125.
- [12]罗俊峰.农民工行业分布对性别工资差异的影响[J].人口与经济,2017(6):105-115.
- [13]CHI W, BO L. Trends in China's Gender Employment and Pay Gaps: Estimating Gender Pay Gaps with Employment Selection[J]. *Journal of Comparative Economics*, 2014(3): 708-725.
- [14]杨慧,张子杨.40年来中国行业性别构成变化趋势——平等还是隔离? [J].人口与经济,2019(4):122-134.
- [15]王湘红,曾耀,孙文凯.行业分割对性别工资差异的影响——基于CGSS数据的实证分析[J].经济动态,2016(1):44-53.
- [16]罗楚亮,滕阳川,李利英.行业结构、性别歧视与性别工资差距[J].管理世界,2019(8):58-68.
- [17]蔡昉,张车伟.中国人口与劳动问题报告 No. 17[M].北京:社会科学文献出版社,2016:237-239.
- [18]鄢姣,许敏波,孟大虎.地域歧视、补偿性溢价与户籍工资差距[J].人口与经济,2021(4):108-125.
- [19]李实,吴珊珊,孟凡强.“五险一金”扩大了城镇职工收入不平等吗? [J].社会科学辑刊,2019(2):73-87.
- [20]HAISKEN-DENEW J P, SCHMIDT C M. Interindustry and Interregion Differentials: Mechanics and Interpretation[J]. *Review of Economics & Statistics*, 1997(3): 516-521.
- [21]田柳,周云波,沈扬扬.不可观测能力、群分效应和行业工资差距:行业分割视角[J].世界经济,2018(12):98-120.
- [22]ROODMAN D. Fitting Fully Observed Recursive Mixed-Process Models with Cmp[J]. *The Stata Journal*, 2011(2): 159-206.
- [23]NORTON E, LINDROOTH R, ETNETT S. Controlling for the Endogeneity of Peer Substance Use on Adolescent Alcohol and Tobacco Use[J]. *Health Economics*, 1998(5): 439-453.
- [24]RIVERS D, VUONG Q. Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models[J]. *Journal of Econometrics*, 1988(3): 347-366.
- [25]鄢姣,孟大虎.高等教育性别差异逆转:驱动因素与典型表征——来自宏观层面和微观层面的经验证据[J].河南师范大学学报(哲学社会科学版),2023(5):95-100.