

基于分位数回归分解的农民工性别工资差异研究

王震*

摘要 本文使用基于分位数回归的分解方法研究了农民工的性别工资差异。结果发现,农民工的性别工资差异表现为“天花板效应”;但在工资分布的最高端,性别工资差异又开始缩小。对此,本文给出了基于统计歧视和社会保护政策的解释。而城镇职工的性别工资差异则表现为“粘地板效应”。在农民工工资分布的高端,性别差异表现出与城镇职工工资分布低端相似的趋势。这一点表明在工资分布高端的女性农民工已经开始突破二元分割体制,获得了一定程度的社会保护。

关键词 分位数回归分解 农民工 性别工资差异

一 引 言

近年来,中国农民工群体中的女性比例越来越高。根据第二次全国农业普查数据,2006年我国外出从业女性农民工已达4747万,占全部外出从业农民工的36%。调查显示,相比于男性农民工,女性农民工的就业环境差,工资水平低,在农民工群体中存在明显的性别工资差异(国务院研究室课题组,2006)。那么,在农民工的性别工资差异中,性别歧视起到了多大作用?与城镇职工相比,农民工的性别工资差异有哪些特征?本文基于中国社科院经济所课题组2006年的农民工调查数据,对此进行研究。

已有的对中国性别工资差异的研究,主要集中在城镇职工群体,对农民工群体性别工资差异的研究并不多见。钟甫宁等(2001)使用1998年江苏苏州和无锡的企业调查数据,通过估计Mincer收入方程,研究了不同性别外来劳动力的职业获得和工资差异,发现男性进入的行业平均工资较高,而女性进入的行业平均工资较低;性别系数解释了大约40%的收入差异。Fan(2003)利用两个宏观数据集和三个微观数据集对中国的乡城移民进入城市劳动市场后的性别分工、职业隔离以及工资收入进行了研究,发现不论在哪个数据集中,女性农民工与男性农民工相比在劳动市场上都处于劣势地位。Meng and Miller(1995)研究了中国农村乡镇企业工人的工资性别差异,通过对Brown全要素分解模型的估计,发现在性别工资的总差异中,职业内差异占76.7%,职业间差异占23.3%;职业内性别歧视差异占性别工资差异的98.9%,职业间性别歧视占性别工资差

* 王震,中国社会科学院经济研究所发展经济学研究室,Email: wangzhen@cass.org.cn 通讯地址:北京市月坛北小街2号,邮编:100836 本文为朱玲研究员主持的中国社科院重大课题“和谐社会与社会公平政策选择”以及魏众研究员主持的中国社科院国情调研重大课题“中国居民收入分配体制改革调研”的研究成果之一。感谢朱玲研究员、魏众研究员、邓曲恒博士、金成武博士以及何伟在研究室内部讨论中提出的建设性意见;感谢匿名审稿人的宝贵修改意见。文责自负。

异的 13.2%; 性别歧视所导致的性别工资差异占性别工资差异的 112.1%。Meng (1998)使用 1995年中国社科院人口所在济南的调查数据,研究了农民工的性别工资差异。在农民工性别工资总差异中,职业内差异占 79.3%,其中由于性别歧视导致的工资差异占 43.4%; 职业间差异占 20.7%,其中由于性别歧视导致的工资差异占 36.1%。

这些研究为我们观察农民工群体的性别工资差异提供了参考。但是,已有研究还存在如下问题:第一,所使用的数据都集中在 2000年之前;而 2000年之后,中国关于农民工的政策发生了一系列积极的变化(宋洪远等, 2006),这些变化不能不对农民工群体的劳动市场表现产生影响。第二,这些研究使用的方法主要关注农民工性别工资的条件均值差异,忽略了整个分布上的性别工资差异。第三,这些研究缺乏与城镇职工性别工资差异相比较的视角。基于此,本文将采用基于分位数回归(Quantile Regression)的分解方法,并通过与城镇职工性别工资差异的比较,研究农民工的性别工资差异。

二、方 法

性别工资差异研究的经典方法是 Oaxaca-Blinder分解方法(Oaxaca, 1973; Blinder, 1973)。Oaxaca-Blinder分解建立在 Mincer工资方程基础上。这一分解方法首先分别估计男性和女性的 Mincer工资方程(Mincer, 1974):

$$\ln w_i = X_i \beta_i + u_i, \quad i = m, f \quad (1)$$

m 为男性, f 为女性。 $\ln w_i$ 为对数工资, β_i 为所要估计的系数, X_i 为解释变量。估计出(1)式后,男性和女性的条件均值工资差异可以分解为:

$$\ln w_m - \ln w_f = (\bar{X}_m - \bar{X}_f) \hat{\beta}_f + (\hat{\beta}_m - \hat{\beta}_f) \bar{X}_m \quad (2)$$

上式右边第一项为特征差异,第二项为系数差异,即由不可观测因素导致的工资差异^①。但是, Oaxaca-Blinder分解存在所谓的“指数问题”(Index Problem),即使用男性还是女性的回归系数作为没有“歧视”的工资结构。在(2)式中,女性面临的工资结构作为没有“歧视”的工资结构。实际上,若以男性回归系数作为没有“歧视”的工资结构,那么(2)也可以分解为:

$$\ln w_m - \ln w_f = (\bar{X}_m - \bar{X}_f) \hat{\beta}_m + (\hat{\beta}_m - \hat{\beta}_f) \bar{X}_f \quad (3)$$

为解决这一指数问题,Neumark(1988)提出使用男性和女性全部样本的回归系数作为没有“歧视”的工资结构^②,并将性别工资差异分解为如下三项:

$$\ln w_m - \ln w_f = (\bar{X}_m - \bar{X}_f) \hat{\beta}^* + (\hat{\beta}_m - \hat{\beta}^*) \bar{X}_m + (\hat{\beta}^* - \hat{\beta}_f) \bar{X}_f \quad (4)$$

其中, $\hat{\beta}^*$ 为男性和女性全部样本回归的系数。(4)式右边第一项为特征差异,第二项和第三项为系数差异,又可以分为男性因为性别歧视获得的“好处”(第二项)和女性因为性别歧视得到的“坏处”(第三项)。

① 在一些文献中,这个差异被称作“性别歧视”导致的差异。实际上,这里的“歧视”是一种残差歧视,即不能被可观测因素解释的差异,有别于理论上的“歧视”。在经验研究中,对这一“歧视”效应的估计会因为研究者对工资方程自变量的选择及数据的可获得性而不同,可参见 Jones(1983)。

② Cotol(1988)提出使用男性和女性回归系数的加权平均作为没有“歧视”的工资结构。

上述分解方法所关注的是性别工资的条件均值差异。但是, 均值差异掩盖了整个工资分布上的性别差异。在现实劳动市场中, 性别工资差异在分布的不同位置是不同的。如果工资分布的高端存在更大的性别差异, 那么就存在“天花板效应”(Glass Ceiling Effects); 如果工资分布的低端存在更大的性别差异, 那么就存在“粘地板效应”(Sticky Floor Effects) (Albrecht et al, 2003)。不论是 Oaxaca-Blinder 分解还是 Neumark 分解, 所分析的都是条件均值上的性别工资差异, 不能揭示整个分布上的性别工资差异特征。而建立在分位数回归 (Quantile regression) 上的分解方法则可以用来分析整个分布上的性别工资差异。分位数回归模型可设定如下 (Koenker and Bassett, 1978; Koenker and Hallock, 2001):

$$\ln w_i = x_i \beta_{\theta_i} + u_i; \quad i = m, f \tag{5}$$

m 和 f 分别为男性和女性; $\theta \in (0, 1)$ 为分位数; X_i 为解释变量, β_{θ} 为第 θ 分位数回归的系数。第 θ 分位点工资的条件分位数为: $Q_{\theta}(\ln w_i | X_i) = X_i \beta_{\theta}; \quad i = m, f$ β_{θ_i} 可以通过下式求得:

$$\hat{\beta}_{\theta_i} = \text{Min}_i \left[\sum_j \rho_{\theta}(\ln w_{ij} - X_{ij} \beta_{\theta_{ij}}) \right]; \quad i = m, f; \quad j = 1, \dots, n \tag{6}$$

其中, ρ_{θ} 为检查函数 (Check function), $\rho_{\theta} = \begin{cases} \theta u, & u \geq 0 \\ (\theta - 1) u, & u < 0 \end{cases}$

建立在分位数回归基础上的分解通过构造反事实的工资分布 (Counterfactual Wage Distribution), 并将之与实际的工资分布进行比较, 将全部性别工资差异分解为特征差异和系数差异。在性别工资差异研究中, 使用比较广泛的是 Machado and Mata (2005) 提出的构造反事实工资分布的方法 (M-M方法)。MM方法首先从均匀分布 $U[0, 1]$ 中生成一个大小为 n 的随机样本, 然后分别估计男性和女性的 n 个分位数回归系数, $\hat{\beta}_{\theta_{ij}} (i = m, f; j = 1, \dots, n)$ 。通过对解释变量矩阵 X_i 进行有放回地重新抽样, 对男性和女性分别生成一个大小为 n 的随机样本。最后, 在男性回归系数的基础上模拟女性的工资分布, 并重复进行 n 次, 生成反事实的女性工资分布 $\{X_i^f \hat{\beta}^m\} (i = 1, \dots, n)$ 。这一方法的基本思想是通过概率积分转换定理的方法来构造反事实的工资分布。一般而言, 通过条件分位数方程转换来估计条件分布是可行的, 但是我们不能保证估计出来的条件分位数方程是单调的, 从而也不能保证是否能够转换 (Nicodemò, 2009)。

Melly (2006) 提出了通过在解释变量集中进行条件分布积分的方法来获得非条件分布的估计, 并给出了估计结果的渐进性质。Melly (2006) 首先估计工资的条件分布, 并给出 X_i 和 θ 下工资条件分布的估计量公式。然后, 将这一工资条件分布方程在 X_i 上进行积分得到工资的非条件分布方程, 并以此非条件分布方程为基础构造反事实的工资分布。Melly (2006) 将第 θ 分位数的工资差异做如下分解:

$$Q_{\theta}(\ln w_m) - Q_{\theta}(\ln w_f) = Q_{\theta}(X_m \hat{\beta}_m) - Q_{\theta}(X_m \hat{\beta}_f) + Q_{\theta}(X_m \hat{\beta}_f) - Q_{\theta}(X_f \hat{\beta}_f) \tag{7}$$

特征差异
系数差异

在本文中, 我们首先给出 Neumark 分解的结果, 然后使用 Melly (2006) 的方法对农

民工的性别工资差异进行分位数回归分解^①。

三、数据说明与描述

本文使用的数据来自于中国社科院经济所课题组 2006年在大连、上海、武汉、深圳、重庆五城市的农民工调查数据,数据清理后共有样本 2 398 个。这次调查首先按照所有制性质、规模及所在行业抽取不同类型的企业,然后在企业中按照性别平衡的原则,随机抽取农民工填写调查问卷。因此,这一数据没有包括自我雇佣类型的农民工样本。

表 1 相关变量描述统计

	全 部	男 性	女 性
教育年限(年)	9 40	9 55	9.26
非农工作年限(年)	7 14	7 93	6.42
老家接受培训(%)	19 97	25 11	15.13
进城接受培训(%)	64 10	65 67	62.40
健康测度得分	2 34	2 04	2.62
年龄(岁)	29 12	30 25	28.08
有配偶(%)	58 47	58 38	58.59
职业(%)			
职业 1	8 23	9 16	7.36
职业 2	5 64	4 36	6.86
职业 3	37 21	46 58	28.43
职业 4	48 92	39 91	57.36
企业所有制性质(%)			
所有制 1	10 11	11 38	8.94
所有制 2	6 93	7 99	5.96
所有制 3	46 53	50 56	42.83
所有制 4	25 27	16 16	33.66
所有制 5	11 15	13 90	8.62

表 1给出了相关变量的描述统计。本次调查的农民工的平均年龄为 29.12岁;女性的平均年龄低于男性。58.47%的农民工为已婚。平均教育年限为 9.40年,女性的平均受教育年限低于男性。考虑到农业工作经历对农民工在城市劳动市场的表现影响不大,因此我们使用非农务工经商年限表示农民工的工作经验。平均而言,全部样本的非农务工年限为 7.14年,女性低于男性。不论是在老家接受培训还是在进入城市后接

^① 实际上, Melly(2006)证明,当 $n \rightarrow \infty$, 其结果与 MM方法一致。本文在实际计算过程中,使用 Melly根据其文章编写的 Stata程序。

受培训的比例,女性都低于男性。根据樊明(2002),我们在问卷中还设计了自评健康指标。根据这些自评健康指标得到的健康测度得分越高则自评健康状况越差。从平均健康测度得分来看,女性的健康状况也弱于男性。教育、工作经验、培训和健康是经典的人力资本的四个变量。从上述几个方面看,女性农民工的人力资本积累低于男性。

我们将农民工的职业分位四类:第一类为进入管理层的农民工,包含专业技术人员;第二类为办事人员;第三类为技术工人;第四类为非技术工人。如果将第一类和第二类作为“白领职业”,后两类作为“蓝领职业”,那么农民工主要集中在“蓝领职业”中。在“白领职业”中,女性则主要集中在办事人员职业;在“蓝领职业”中,女性主要集中在非技术工人中。这表明在农民工群体中存在性别职业隔离^①。

我们将农民工所在企业的所有制性质分为五类:第一类为国有企事业单位;第二类为城镇集体所有制;第三类为私营企业;第四类为外资与合资企业;第五类为其他股份制企业。

被调查农民工的平均月工资为1002.05元,男性为1100.24元,女性为910.78元;剔除工作时间影响后,平均小时工资为4.51元,男性为4.87元,女性为4.17元。男性和女性农民工的月均工资差异(对数;男性/女性)为1.03,小时工资(对数)的性别差异(男性/女性)上升到1.11。

从性别工资差异的经验分布上看(图1),农民工的性别工资差异表现为“天花板效应”,随着工资的上升,性别工资差异不断扩大^②。这种性别工资差异特征与城镇职工的性别工资差异特征正好相反。我们使用国家统计局2006年城镇住户抽样调查数据(子样本)^③测算了城镇职工2006年的性别工资差异(年工资)。图1显示城镇职工的性别工资差异呈现“粘地板效应”:在收入分布的高端,性别工资差异较小。

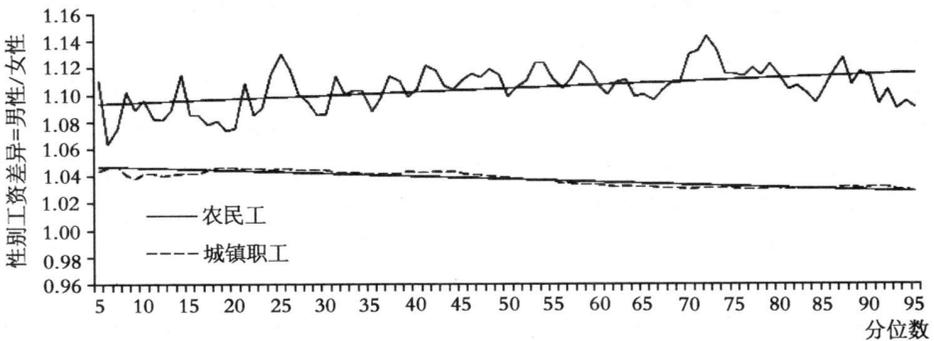


图1 农民工与城镇职工的性别工资差异

注释:1 性别工资差异=男性/女性;2 农民工为小时工资(对数),城镇职工为年工资(对数);3 图中直线为趋势线。

① 测算的性别职业隔离的杜肯指数(Duncan Index)为19.95%。

② 为剔除分布两端的异常值,本文中使用了第5到第95个分位点。

③ 数据来源于2006年国家统计局城镇住户抽样调查的子样本,样本涵盖11个省(直辖市、自治区),样本量为7320户,10379口人。由于没有得到工作时间的数据,因此这里使用城镇职工的年工资。年工资中实际包含了工作时数的影响,因此与农民工小时工资的比较只能是大致的比较,意在说明一种趋势。

四、估计结果及解释

我们首先使用 Neumark分解方法对农民工性别工资的均值差异进行分解。在 Mincer工资方程的回归中,我们使用的变量包括两组,一组是供给方的个人特征变量,包括教育年限、非农工作年限、非农工作年限的平方、在老家接受培训、在城市接受培训、年龄和是否有配偶;另一组包括农民工的职业和所在单位的所有制性质。

从 Mincer方程的 OLS回归结果中(表 2),可以发现个人特征变量的回报率存在明显的性别差异:男性教育年限、非农工作年限,以及健康测度的回报率都高于女性,而女性在城市接受培训的回报率则高于男性。如果不考虑职业和所在单位所有制性质,那么以供给方变量回归为基础的 Neumark分解结果显示(表 3),系数差异占到了总差异的 62.76%。如前所述,在农民工群体中还存在性别职业隔离,这也影响了女性的工资和性别工资差异。此外,所在单位不同的所有制性质也会影响农民工的性别工资差异。因此,我们将职业类别和所在单位所有制性质加入回归方程,并以此为基础进行 Neumark分解(表 3),结果显示在所有变量回归下,系数差异占总差异的比例有所下降,但是仍然占到了总差异的一半以上(53.02%)。这表明在农民工的工资性别差异中,还存在较为严重的性别歧视。

为了发现在整个分布上的性别工资差异特征,我们使用上节给出的基于分位数回归分解的方法对农民工的性别工资差异进行分解^①。在分位数回归中,我们使用与 OLS回归中相同的解释变量。表 2给出了分位数回归的第 25和第 75个分位点的回归结果。从分位数回归结果中可以发现人力资本回报率的性别差异在工资分布的高端出现了缩小的趋势。除了老家培训不显著外,教育年限、非农工作经验、健康测度和在城市培训的回报率差异都出现了缩小的趋势。图 2和图 3给出了教育年限和非农工作经验在不同分位点上的回报率。女性教育年限的回报率在工资分布的低端呈下降趋势,但是在第 45个分位点左右开始上升;从回报率的性别差异看,则在整个分布上呈现出缩小的趋势。女性非农工作经验的回报率从最低端开始上升,在第 85个分位点左右又出现了下降的趋势;而在整个分布上回报率的性别差异呈现出缩小趋势。

从全部解释变量分位数回归分解的结果分析(图 4),农民工的性别工资差异表现出“天花板效应”。但是,在最高端又出现了性别差异的缩小。在总差异中,系数差异所占比例高于特征差异所占比例,除了在第 5到第 10个分位点上系数差异所占比例低于 50%外,其他分位点上的系数差异都超过了 50%,最高达到了 70%(第 54个分位点)。而在最高端(第 85到第 95个分位点)系数差异又表现出了缩小的趋势。这表明在农民工的性别工资差异中,包括性别歧视在内的不可观察因素是主要原因,特别是在工资分布的高端;而在最高端,虽然系数差异开始缩小,但是占总差异的比例还是在 60%以上。

^① Melly给出的 Stata程序同时给出了估计结果的标准误,限于篇幅本文没有给出,有需要者可与作者联系索取。

表2 农民工工资方程的OLS回归和分位数回归结果

	OLS回归				分位数回归			
	男性		女性		男性		女性	
	0.25	0.75	0.25	0.75	0.25	0.75	0.25	0.75
教育年限	0.0297***	0.0224***	0.0232***	0.0120*	0.0222**	0.0272***	0.0099	0.0087
非农工作年限	0.0312***	0.0166**	0.0314***	0.014	0.0212**	0.0200*	0.0103	0.0102
非农工作年限平方	-0.0010***	-0.0004	-0.0011***	-0.0003	-0.0007	-0.0006	0.0004	0.0005
健康测度	-0.0108***	-0.0153***	-0.0116***	-0.0155***	-0.0092*	-0.0142***	0.0055	0.004
老家培训	0.0034	0.0028	0.0038	0.003	0.0741	0.0032	0.0454	0.0472
城市培训	0.0462	0.0009	0.0418	0.0186	0.0732*	0.0839**	0.042	0.038
年龄	0.0298	0.0345	0.034	0.0381	0.042	0.038	-0.0054	0.0056*
有配偶	0.0516*	0.0944***	0.0938***	0.1110***	0.0029	0.0031	0.1834***	-0.0218
职业(以第一类职业为基准)	0.0277	0.0281	0.0308	0.032	0.0602	0.0512	0.0033	0.0031
职业2	-0.0026	-0.0046*	-0.001	0.0039	-0.0054	-0.0056*	0.0033	0.0031
	0.1869***	-0.0057	0.2024***	-0.005	0.1834***	-0.0218	0.0033	0.0031
	0.0385	0.0397	0.0427	0.0429	0.0602	0.0512	0.0033	0.0031
	-0.1763**	0.0226	-0.1322	0.0937	-0.1526	-0.0299	0.0033	0.0031
	0.0726	0.0718	0.0814	0.08	0.1082	0.0969	0.0033	0.0031

续表

	OLS 回归				分位数回归			
	0.25		0.75		0.25		0.75	
	男性	女性	男性	女性	男性	女性	男性	女性
职业 3	-0.2565 ***	-0.2440 ***	-0.1234 **	-0.1684 ***	-0.3147 ***	-0.3210 ***		
	0.0464	0.0562	0.0526	0.0638	0.0683	0.0751		
职业 4	-0.4552 ***	-0.4024 ***	-0.2928 ***	-0.2523 ***	-0.5496 ***	-0.4974 ***		
	0.0489	0.0558	0.0555	0.0645	0.0713	0.0732		
单位所有制性质(以国有企业单位为基准)								
所有制 2	-0.0225	0.1293 **	-0.0833	-0.0432	0.1184	0.1102		
	0.0603	0.0625	0.0674	0.0687	0.0912	0.0842		
所有制 3	-0.0511	0.0949 **	-0.0203	0.1209 **	-0.0565	-0.051		
	0.0423	0.0425	0.0479	0.0472	0.0641	0.0573		
所有制 4	0.0276	0.2705 ***	0.0443	0.2851 ***	0.0224	0.1275 **		
	0.0513	0.0483	0.0585	0.0534	0.0769	0.0647		
所有制 5	0.0738	0.1511 ***	0.0741	0.1584 ***	0.1114	0.0703		
	0.05	0.055	0.0565	0.061	0.0762	0.0739		
截距	1.3074 ***	1.0302 ***	0.9031 ***	0.8011 ***	1.7851 ***	1.3969 ***		
	0.114	0.12	0.1339	0.1374	0.1707	0.1583		
N	1 010	923	1 010	923	1 010	923		
AdjR2/Pseudo R2	0.245	0.2384	0.1352	0.0999	0.155	0.1615		

注:系数下方为标准误;*、**、***分别代表在10%、5%、1%的程度上显著;OLS估计为 Adjusted R2,分位数回归为 Pseudo R2。

表3 Neumark分解结果及部分分位数回归分解结果

	Neumark分解结果		分位数回归分解结果				
	供给方变量回归	全部变量回归	0.1	0.25	0.5	0.75	0.9
总差异	0.193	0.1941	0.1342	0.1611	0.2024	0.2448	0.2423
特征差异	0.0719	0.0912	0.0679	0.0543	0.0587	0.0734	0.0875
(%)	37.24	46.98	50.61	33.67	28.99	29.97	36.11
系数差异	0.1211	0.1029	0.0663	0.1069	0.1437	0.1714	0.1548
(%)	62.76	53.02	49.39	66.33	71.01	70.03	63.89
男性高估	0.0582	0.0491					
(%)	30.17	25.32					
女性低估	0.0629	0.0538					
(%)	32.59	27.71					

注：男性高估部分为上文所说的男性从性别歧视中获得的“好处”，女性低估部分为女性从性别歧视中获得的“坏处”；分位数分解基于全部变量的分位数回归。

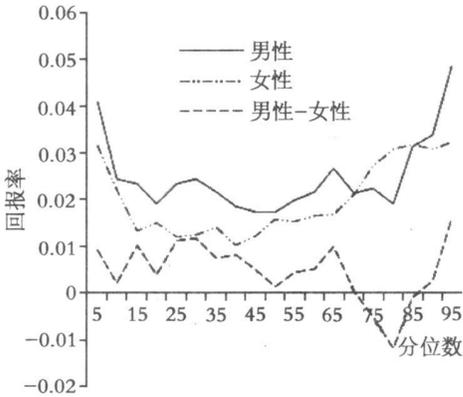


图2 教育年限回报率差异

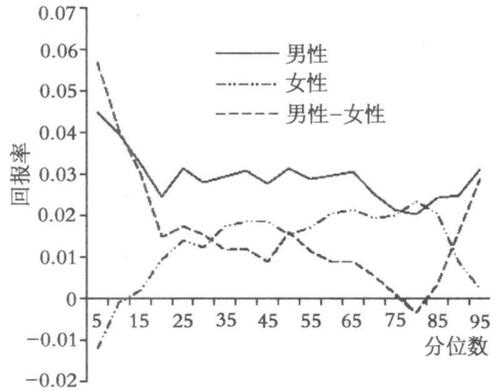


图3 非农工作经验回报率差异

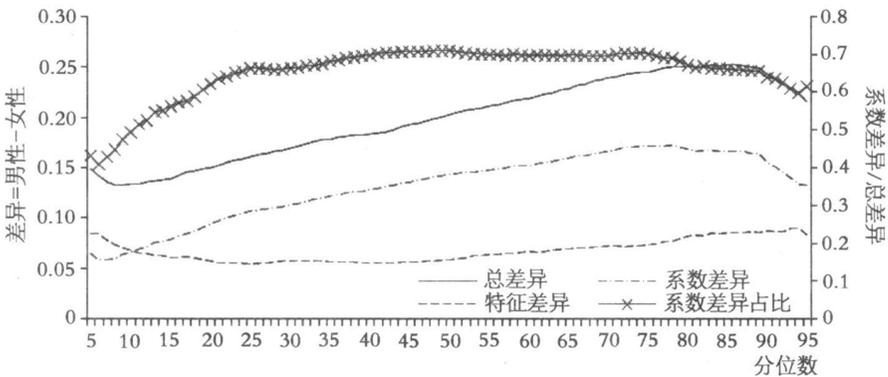


图4 农民工性别工资差异的分位数回归分解结果(全部变量回归)

为了分析性别职业隔离和所在单位所有制性质对农民工性别工资差异分布的影响,我们仅使用供给方变量进行了分位数回归分解,并与全部变量回归结果进行了比较(图5)。在加入职业和单位所有制性质后,性别工资差异在工资分布的高端出现了明显的下降,但在低端则变化不明显。这表明性别职业隔离和单位所有制性质对农民工性别工资的影响主要是在工资分布的高端。

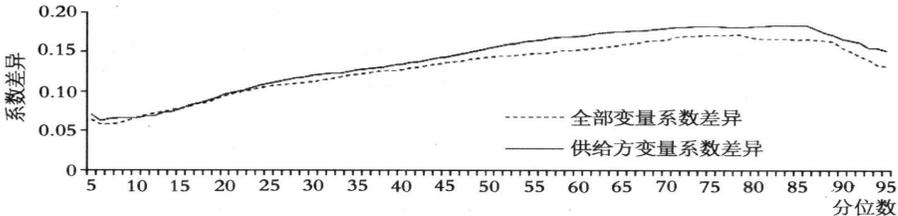


图5 全部变量和供给方变量的分位数回归分解结果

农民工性别工资差异表现出来的“天花板效应”表明性别歧视因素在工资分布的高端影响更大,而在最高端因性别歧视导致的工资差异又出现了下降趋势。那么,如何解释这种性别工资差异特征呢?在劳动经济学中,对性别歧视的解释主要有偏好歧视理论(Becker, 1957)和统计歧视理论(Aigner and Cain, 1977; Autor, 2003)。偏好歧视是说对女性的歧视来源于雇主或顾客对女性的厌恶偏好。统计歧视理论则认为之所以产生对女性的歧视,是因为在雇主和工人之间存在信息不对称,雇主依靠性别作为甄别工人生产能力的信号。在信息不对称的条件下,雇主不能有效鉴别男性和女性的真实生产能力和努力程度,雇主会根据全社会女性的平均生产能力来给自己雇佣的女性职工发放工资,这样就构成了对生产能力超过平均水平的女性的歧视。

我们认为统计歧视理论更适合用来解释农民工性别工资差异的分布特征。在工资分布最低端的农民工主要是那些从事最不需要技术和经验的低端工作,他们往往与雇主不签订劳动合同或签订短期劳动合同,大多数是计件工资或能够很容易转化为计件工资的计时工资。在这种条件下,性别作为鉴别生产能力的信号作用不大,因此性别工资差异以及性别“歧视”都不大。随着工资的上升,农民工所从事工作的技术含量也逐渐提高,信息不对称的程度增大,对工人监督的成本越来越高,从而性别作为生产能力信号的作用也开始增大。在信息不对称的条件下,这导致雇主根据平均工资水平给女性工人发放工资,从而损害了生产能力高于平均水平的女性工人,造成他们的低工资。如果没有社会政策的干预,那么可以预计随着工资水平的提高,性别工资差异还会继续增大。

但是,在农民工性别工资差异的分解中,在工资分布的最高端性别工资差异出现了下降。我们认为这主要是在性别工资分布的最高端,社会保护政策开始覆盖这部分农民工,从而矫正了性别工资差异继续拉大的趋势。也就是说,在工资分布的低端和中高端,农民工群体缺乏有效地社会保护,性别歧视没有得到矫正;而到了工资分布的高端,一部分农民工开始获得社会性别保护。图6给出了按工资五等分组下的女性农民工与所在企业签订劳动合同的比例,以及工伤保险、医疗保险、养老保险、失业保险的覆盖率情况。可以发现,随着工资的提高,女性农民工与所在企业签订劳动合同的比例以及四项社会保险的覆

盖率都随之提高。这表明在工资分布的高端,女性农民工所受到的社会政策保护要好于工资分布的低端,而社会保护政策部分冲抵了统计性性别歧视造成的性别工资差异。

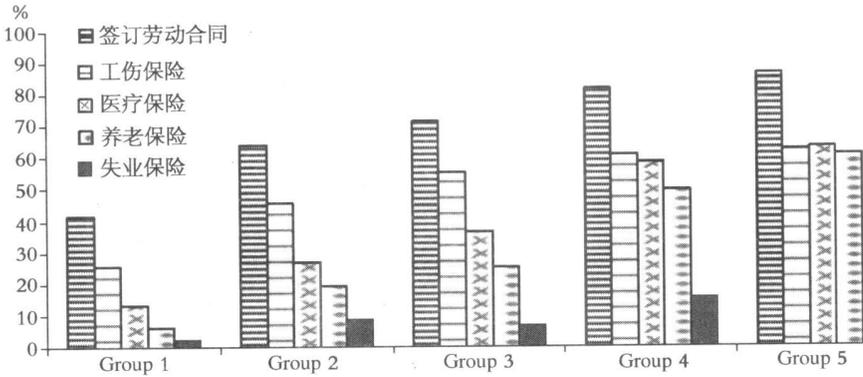


图 6 按工资五等分组的女性农民工签订合同及社会保险覆盖率情况

与城镇职工性别工资差异的比较则表明,在工资分布高端的部分女性农民工已经开始突破城乡二元分割体制,逐步融入城市社会。城镇职工性别工资差异表现为“粘地板效应”,即在工资分布的低端性别工资差异较大,随着工资水平的上升,性别工资差异缩小。图 7 给出了我们使用国家统计局 2006 年城镇住户抽样调查数据进行的城镇职工性别工资差异的分位数分解结果^①,该结果显示城镇职工性别工资差异与系数差异呈现出明显的“粘地板效应”。比较图 4 农民工性别工资差异分布的高端与图 7 城镇职工性别工资差异分布的低端,可以发现二者分布的趋势有相似之处:即都出现了性别工资差异逐渐减小的趋势。这种趋势大致可以表明在工资分布高端的女性农民工已经开始部分突破城乡二元分割体制,获得了与工资分布低端城镇职工相似的社会保护。

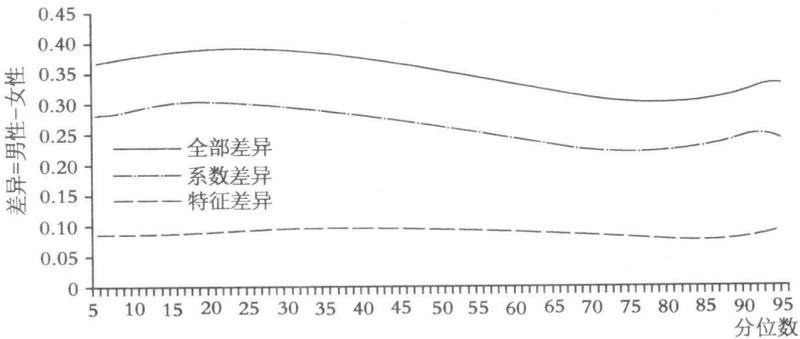


图 7 城镇职工性别工资差异的分位数回归分解结果

① 此处使用的是城镇职工的年工资(对数);自变量包括:年龄、教育程度、婚姻状况、就业性质、职业。教育程度分为小学以下、初中、高中和中专、大学专科、本科及以上;就业性质分为:国有经济单位、城镇集体及其他经济类型单位、非正规就业人员(城镇个体或私营企业被雇佣者、离退休再就业人员及其他就业者);职业为国家统计局城镇住户抽样调查方案中的八类。此处的估计虽然由于使用的变量不同,估计结果的绝对值与农民工的分解结果没有直接的可比性,但是仍可给出城镇职工性别工资差异的分布状况。

五、结 论

本文在中国社科院经济所课题组 2006 年农民工调查数据的基础上,使用 Neumark 分解和分位数回归分解的方法研究了农民工的性别工资差异;我们还对农民工群体的性别工资差异和城镇职工的性别工资差异进行了比较。有如下结论值得关注:

第一,农民工工资方程的 OLS 估计显示,人力资本的回报率存在性别差异,男性教育年限、非农工作年限,以及健康测度的回报率都高于女性;而女性在城市接受培训的回报率则高于男性。Neumark 分解结果显示,系数差异占总差异的比例占到了总差异的一半以上(53.02%)。这表明在农民工的工资性别差异中,还存在较为严重的性别歧视。

第二,分位数回归结果显示,农民工的性别工资差异及系数差异表现为“天花板效应”;但是,在工资分布的最高端,性别工资差异又出现了缩小的趋势。性别职业隔离和农民工所在单位的所有制性质对农民工性别工资差异的影响主要在工资分布的高端。

第三,农民工性别工资差异出现的“天花板效应”及分布最高端性别工资差异缩小的趋势,可以用统计歧视理论和社会保护政策差异来解释。在工资分布的低端,农民工所从事的大多是最不需要技术和经验的低端工作,性别作为鉴别生产能力的信号作用不大,因此性别工资差异以及性别“歧视”都不大。随着工资的上升,农民工所从事工作的技术含量也逐渐提高,对工人监督的成本越来越高,从而性别作为生产能力信号的作用也开始增大,信息不对称导致的性别歧视影响增大。最高端性别工资差异的缩小则主要是因为社会保护政策开始覆盖这部分农民工,从而矫正了性别工资差异继续拉大的趋势。

第四,比较农民工性别工资差异分布的高端与城镇职工性别工资差异分布的低端,可以发现二者分布的趋势有相似之处:即都出现了性别工资差异逐渐减小的趋势。这种趋势大致可以表明在工资分布高端的女性农民工已经开始部分突破城乡二元分割体制,获得了与工资分布低端的城镇职工相似的社会保护。

上述结论的政策含义在于,需要通过社会保护政策的实施来矫正农民工群体的性别工资差异。实际上,当前中国的社会保护政策网络已经形成,但是,问题在于这些政策并没有平等覆盖到所有劳动群体,而农民工显然是没有被覆盖的群体之一。可以预计,如果缺乏有效的社会保护政策,农民工群体中的性别工资差异还将继续扩大。乐观的是,在工资分布高端的女性农民工已经开始获得与工资分布低端城镇职工相似的社会保护。就当前的具体政策而言,实际上并不需要制定新的针对女性农民工群体的社会保护政策,只要将当前覆盖城镇女性职工的社会保护政策扩展到女性农民工群体,就可以大大缩小这个群体的性别工资差异。

参 考 文 献

- tions Review, 30(2), 175-187.
- Albrecht, J. A. Bjorklund, and S. Vroman, 2003, "Is There A Glass Ceiling in Sweden?" *Journal of Labor Economics*, 20(1), 145-177.
- Author, David H., 2003, "The Economics of Discrimination Theory" *Lecture Note of Labor Economics*, MIT 14, 661.
- Becker, G. S., 1957, *The Economics of Discrimination*, University of Chicago Press.
- Blinder, Alan S., 1973, "Wage Discrimination, Reduced Form and Structural Estimates" *Journal of Human Resources*, 8(4), 436-455.
- Cotton, J., 1988, "On the Decomposition of Wage Differentials" *The Review of Economics and Statistics*, 70(2), 236-243.
- Fan, C. C., 2003, "Rural-Urban Migration and Gender Division of Labor in Transitional China" *International Journal of Urban and Regional Research*, 27(1), 24-47.
- Jones, F. L., 1983, "On Decomposing the Wage Gap: A Critical Comment on Blinder's Method" *Journal of Human Resources*, 18(1), 126-130.
- Koenker, R., and G. B. Bassett, 1978, "Regression Quantiles" *Econometrica*, 46(1), 33-50.
- Koenker, R., and K. Hallock, 2001, "Quantile Regression" *Journal of Economic Perspectives*, 15(4), 143-156.
- Machado, J., and J. Mata, 2005, "Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression Models" *Journal of Applied Econometrics*, 20(4), 445-465.
- Melly, B., 2006, "Estimation of Counterfactual Distributions Using Quantile Regression" *University of St. Gallen, Discussion Paper*.
- Meng, Xin and P. Miller, 1995, "Occupational Segregation and Its Impact on Gender Wage Discrimination in China's Rural Industrial Sector" *Oxford Economic Papers*, 47(1), 136-155.
- Meng, Xin, 1998, "Gender Occupational Segregation and Its Impact on the Gender Wage Differential among Rural-Urban Migrants: A Chinese Case Study" *Applied Economics*, 30(6), 741-752.
- Mincer, Jacob A., 1974, *Schooling, Experience, and Earnings*, National Bureau of Economic Research.
- Neumark, D., 1988, "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination" *Journal of Human Resources*, 23(3), 279-295.
- Nicodemo, C., 2009, "Gender Pay Gap and Quantile Regression in European Families" *IZA Discussion Paper No. 3978*.
- Oaxaca, R., 1973, "Male-female Wage Differentials in Urban Labor Markets" *International Economic Review*, 14(3), 693-709.
- 樊明, 2002, 《健康经济学——健康对劳动市场表现的影响》, 社会科学文献出版社。
- 国务院研究室课题组, 2006 《中国农民工调研报告》, 言实出版社。
- 宋洪远、黄华波和刘光明, 2006 关于农村劳动力流动的政策问题分析, 载蔡昉、白南生主编, 《中国转轨时期劳动力流动》, 社会科学文献出版社, 59—76页。
- 钟甫宁、徐志刚和栾敬东, 2001 《经济发达农村地区外来劳动力的性别差异研究》, 《人口与经济》第2期, 31—37页。