

在线数据揭示预期薪金的影响因素

王 军^{1*}, 高 见^{1,2}, 杨 泉^{1,3}, 刘金虎¹, 周 涛^{1,2}

(1. 电子科技大学大数据研究中心 成都 611731; 2. 成都新经济发展研究院 成都 610094;
3. 电子科技大学数学科学学院 成都 611731)

【摘要】数据资源的丰富和分析方法的创新,促使社会经济学逐渐转变为数据驱动的定量化学科。作为量化人力资源的组成部分,薪金研究对社会经济发展有重要意义。然而,以往研究大多基于规模有限的普查数据,对不同经济和文化背景的考虑也不足。本文基于中国人力资源网站获取的大规模简历数据,分析了多种因素对求职者预期薪金的影响,结果发现身高、工作经验和教育程度等因素影响预期薪金,并且男女存在显著的差异。其中,女性平均预期低于男性,相比男性差大约五年工作经验或一个教育学位。最后,多变量回归方法验证了分析结果的鲁棒性。

关键词 大数据; 计算社会经济学; 数据驱动; 多变量回归; 预期薪金

中图分类号 TP391 **文献标志码** A **doi**:10.3969/j.issn.1001-0548.2019.02.023

Online Data Reveal Key Factors on Salary Expectation

WANG Jun^{1*}, GAO Jian^{1,2}, YANG Xiao^{1,3}, LIU Jin-hu¹, and ZHOU Tao^{1,2}

(1. Big Data Research Center, University of Electronic Science and Technology of China Chengdu 611731;
2. Institution of New Economic Development Chengdu 610094;
3. School of Mathematical Sciences, University of Electronic Science and Technology of China Chengdu 611731)

Abstract The enrichment of data resources and the innovation of analytic methods are gradually facilitating the transformation of socioeconomics into a data-driven and quantitative discipline. As a part of quantitative human resources, the investigation of salary has a significant role on social and economic development. However, previous studies are mainly based on census data with limited sizes and lack of considerations in a different economic and cultural background. Based on large-scale resume data that were crawled from websites of Chinese human resource service providers, this paper analyzes key factors on job seekers' salary expectation. Results suggest that height, working experiences, and educational degree affect salary expectation, and there are significant gender differences. In particular, females have lower salary expectation on average and lag behind males for five years' working experience or one educational degree. Finally, the robustness of the analytical results is checked using the multivariate regression method.

Key words big data; computational socioeconomics; data-driven; multivariate regression; salary expectation

数据资源和分析方法的创新为社会经济和人力资源管理等相关领域的量化分析提供了新思路^[1]。一方面,在线信息平台广泛收集和存储了大量的社会经济数据^[2]。另一方面,跨学科交叉研究催生了很多新的分析工具^[3]。利用这些新工具定量分析在线数据,能够揭示社会经济状态^[4]、绘制全球贫穷地图^[5]、刻画产业结构^[6-7]、感知城市发展不平衡^[8]、制定经济发展策略^[9-10]、推断个体财富状况^[11]和预测个体失业率^[12-13]等。在量化人力资源方面,不仅升职和离职对员工的发展很重要^[14],薪金水平更加直接影响到员工的生活水平和工作热情,也进一

步影响组织的工作效率^[15]。

薪金对社会经济发展有非常重要的作用,能够配置劳动力资源。反过来,薪金也受很多社会经济因素影响。在身高因素上,文献[16]发现薪金与身高呈非线性关系,高于平均身高的员工有明显的薪金优势;文献[17]发现增加10 cm身高能分别增加男性8%和女性13%的薪金;文献[18]发现身高对女性预期薪金的影响大于男性。在教育因素上,文献[19]发现受教育水平对收入有显著正向的影响;文献[20]发现我国教育水平和薪金收入总体上呈正相关关系;文献[21]发现高等教育率和教育公平对于改善收

收稿日期: 2018-05-10; 修回日期: 2018-08-30

基金项目: 国家自然科学基金(61703074, 61673086)

作者简介: 王军(1988-), 男, 博士生, 主要从事大数据分析等方面的研究. E-mail: junwang31@gmail.com

入平等性至关重要。在工作经验方面,文献[22]发现技术工人的企业任职和工作经验对薪金收入有正面作用;文献[23]发现有15~20年企业工作经验的、自然科学专业毕业的女博士拥有最高的薪金收入。

虽然已经有一些针对薪金影响因素的分析,但这些研究一般基于规模有限的问卷和普查数据^[1],且大多关注单一变量的影响,针对中国背景下的研究也不足。中国虽然近些年经济发展快、经济复杂性高^[24],但发展不平衡,直接影响不同地区的薪金水平。更为重要的是,以往的研究一般使用实际薪金数据,无法剔除职场中普遍存在的身高和性别歧视对薪金的影响^[18]。如今,得益于互联网和信息技术的发展,已经能够获得非干预的大规模数据,可以全面分析预期薪金的影响因素。

本文基于人力资源网站收集的14余万求职者简历数据,分析了影响中国求职者预期薪金的核心因素。结果发现,预期薪金存在明显的性别差异,男性的平均预期薪金高于女性;求职者的身高、以往工作经验、受教育程度和期望工作地的经济发展水平对预期薪金有显著影响;最后,本文使用多变量回归验证了分析结果的鲁棒性。

1 数据与方法

1.1 数据描述

本文的简历数据来源于中国的两大人力资源服务商,即前程无忧网站(<http://www.51job.com>)和中华英才网(<http://www.chinahr.com>)。求职者在使用人力资源服务平台寻找工作时,需要在线提交求职简历,公开个人基本信息和预期薪金,HR根据这些公开信息筛选合适的求职者。本文于2014年从这两大招聘网站上爬取公开的简历信息,涵盖来自中国大陆31个省份的142 190位求职者。

简历数据主要是求职者主动公开的个人信息,包括性别(gender)、出生日期、身高(height)、工作经历、毕业院校(school)、获得学位(degree)、求职省份和预期薪金(salary)。根据出生日期,计算得到年龄(age);根据工作经历,计算得到工龄(seniority)。毕业院校根据教育部文件分为4类进行数值化:4代表“985工程”院校(985 PRG),3代表“211工程”院校(211 PRG),2代表普通高等院校(college),1代表其他院校(others)。获得学位分为4类,同样进行数值化:4代表博士学位(PhD),3代表硕士学位(master),2代表学士学位(bachelor),1代表其他学位(others)。求

职省份涵盖31个中国大陆的省份、直辖市和自治区,省份名称和缩写见文献[18]。

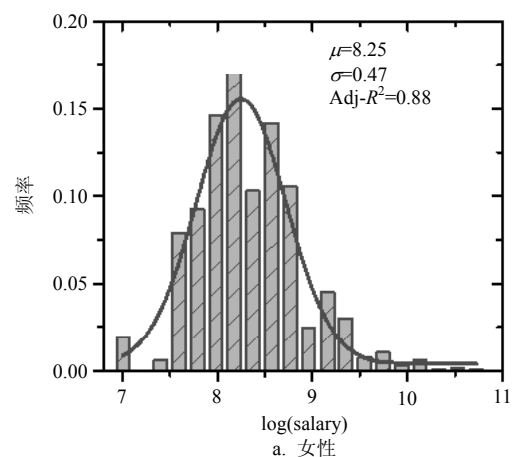
为了保障样本数量中有足够的合理预期薪金的普通求职者,本文通过限定求职者的身高和预期薪金来筛选数据,剔除不符合限定的数据。其中,男性的身高范围限定为[160, 185] cm,女性的身高范围限定为[150, 175] cm;预期薪金的范围限定为[1 000, 50 000]元/月。如表1所示,使用数据涵盖141 064人,包括78 413位男性和62 651位女性;男性平均身高高于女性;男性的平均工龄和年龄都大于女性;男性的平均教育水平高于女性,男性毕业于更好的学校、取得更高的学位;男性的平均预期薪金为8 039元/月,明显高于女性的5 017元/月。

表1 简历数据的统计信息

性别	样本数量	身高/cm	年龄/岁	工龄/年	学校	学位	预期薪金/元
男	78 413	173.39	29.67	5.48	2.31	1.81	8 039
		[160, 185]	[18, 66]	[0, 36]	[1, 4]	[1, 4]	[1 000, 50 000]
女	62 651	162.04	27.64	3.80	2.21	1.77	5 017
		[150, 175]	[18, 57]	[0, 31]	[1, 4]	[1, 4]	[1 000, 50 000]

注:学校和学位为向量化后的数值。

在对预期薪金进行自然对数运算之后,男性和女性的预期薪金呈现正态分布。也就是说,自然对数的预期薪金 $\log(\text{salary})$ 服从以 μ 为平均值、 σ 为标准差的正态分布 $\log(\text{salary}) \sim N(\mu, \sigma^2)$ 。如图1所示,自然对数运算之后的预期薪金,可以很好地被正态分布拟合 ($\text{Adj. } R^2 \approx 0.85$);男性的平均预期薪金 ($\mu = 8.65$) 高于女性的平均值 ($\mu = 8.25$);男性的预期薪金分布更宽广 ($\sigma = 0.58$),女性的预期薪金分布更集中 ($\sigma = 0.47$)。



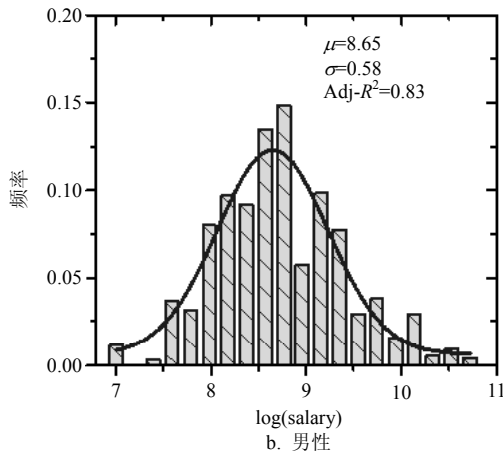


图1 预期薪金的分布

(条形图表示频率分布, 实线为正态拟合曲线)

除了简历数据, 后续分析还使用了省份的经济发展数据, 来源于《中国统计年鉴-2015》, 由国家统计局(<http://www.stats.gov.cn>)编纂, 发布全国31个省份在2014年的人均生产总值(GDPpc)数据。

1.2 分析方法

在挖掘预期薪金的核心影响因素方面, 首先采用统计分析方法, 分析单个或同一类变量对预期薪金 $\log(\text{salary})$ 的影响, 包括身高 (height)、工作经验 (age 和 seniority)、教育程度 (school 和 degree) 和地理经济 (GDPpc) 等相关变量。然后, 使用多变量回归分析方法验证结果的鲁棒性, 并比较各个变量对预期薪金影响的性别差异。

多变量分析方法采用普通最小二乘 (ordinary least square, OLS) 回归模型, 被解释变量为预期薪金 $\log(\text{salary})$, 解释变量为影响预期薪金的各个变量。由于教育水平中的 school 和 degree 都是分类变量, 所以将他们作为虚拟变量 (dummy variable) 引入回归模型。为了比较解释变量对于男性和女性预期薪金的影响, 分析男性和女性回归结果的截距和斜率差异, 本文在混合样本 (pooled sample) 上进行回归, 控制男性和女性在回归分析中的残差变化一致。

在回归方程中, 将女性 (female) 作为虚拟变量, 其他解释变量与 female 形成交叉项 (interaction term), 所估计的回归方程如下:

$$\begin{aligned} \log(\text{salary}) = & \varepsilon + \beta_0 + \delta_0 \text{female} + \\ & \beta_1 \text{height} + \delta_1 \text{female} \cdot \text{height} + \\ & \beta_2 \text{age} + \delta_2 \text{female} \cdot \text{age} + \\ & \beta_3 \text{seniority} + \delta_3 \text{female} \cdot \text{seniority} + \\ & \beta_4 D^{\text{school}} + \delta_4 \text{female} \cdot D^{\text{school}} + \\ & \beta_5 D^{\text{degree}} + \delta_5 \text{female} \cdot D^{\text{degree}} + \\ & \beta_6 \log(\text{GDPpc}) + \delta_6 \text{female} \cdot \log(\text{GDPpc}) \end{aligned}$$

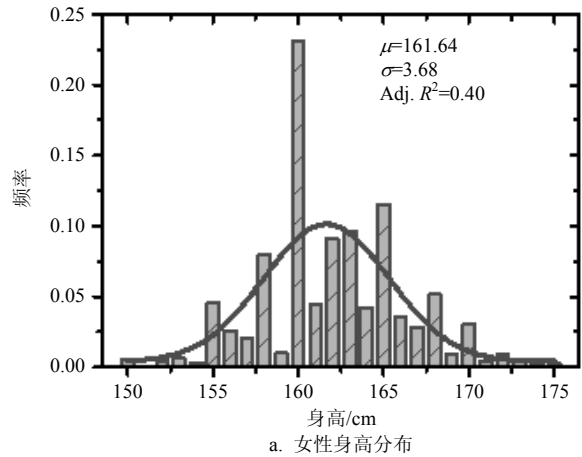
式中, female = 0 对应于基准的男性组; β 为解释变量对于男性预期薪金的回归系数; δ 为相应解释变量在男女回归系数上的差异; D^{school} 为学校的虚拟变量; D^{degree} 为学位的虚拟变量; ε 为误差项。

2 实验结果

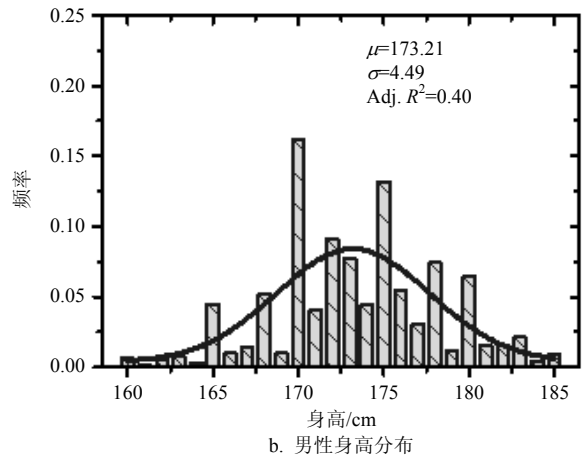
实验中首先依次分析身高、工作经验、教育水平和经济地理对男女预期薪金的影响, 考虑单个或同一类变量的作用。然后, 利用多变量回归模型, 验证分析结果的鲁棒性, 定量刻画和比较男女预期薪金影响因素的不同, 侧重于性别差异。

2.1 身高的影响

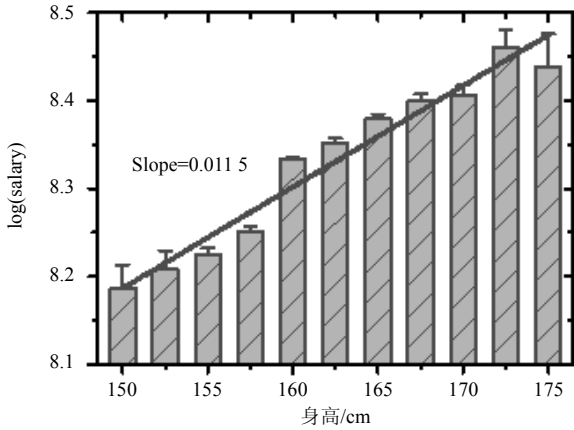
首先, 考虑男性和女性求职者的身高分布。如图2a和图2b所示, 男女的身高频率分布都可以很好地通过正态分布拟合 ($\text{Adj. } R^2 \approx 0.40$), 但男性和女性的身高分布存在明显差异。具体而言, 男性的平均身高 ($\mu = 173.21$) 显著高于女性的平均身高 ($\mu = 161.64$); 男性身高的分布更宽广 ($\sigma = 4.49$), 而女性身高的分布更狭窄 ($\sigma = 3.68$)。



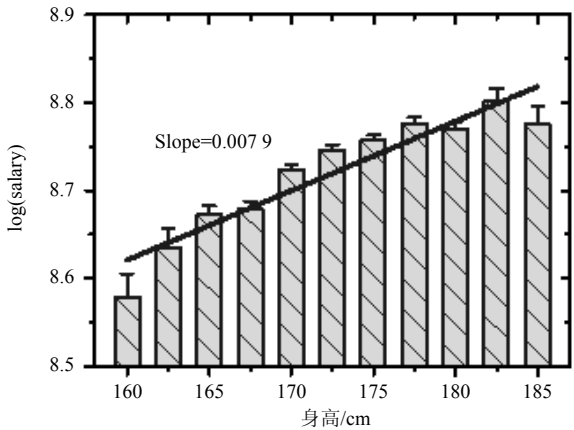
a. 女性身高分布



b. 男性身高分布



c. 身高对女性预期薪金的影响



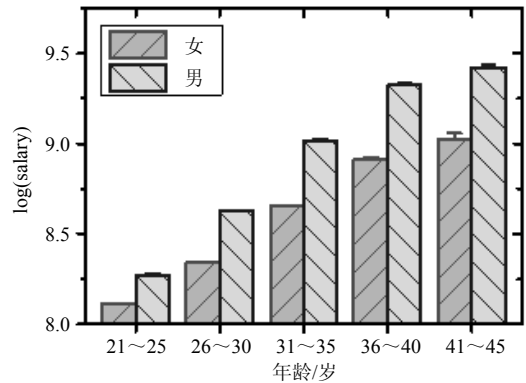
d. 身高对男性预期薪金的影响

图2 身高对预期薪金的影响(a、b中条形图为概率分布, 实线为正态拟合曲线; c、d中条形图为平均值, 误差线为标准误差, 实线为拟合直线)

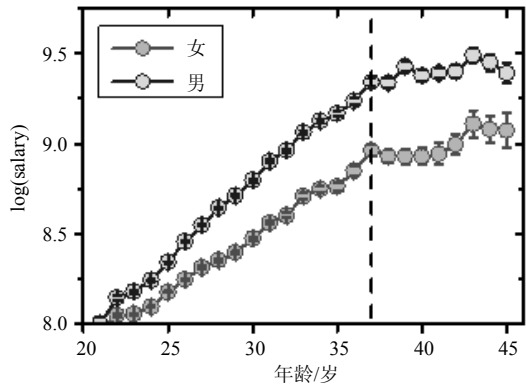
然后, 分析身高对预期薪金的影响。如图2c和图2d所示, 平均预期薪金与平均身高呈现线性关系(男性: $Adj. R^2 \approx 0.87$, 女性: $Adj. R^2 \approx 0.95$)。不论对于男女, 求职者的平均身高越高, 平均预期薪金越高, 存在Height Premium现象^[25-26]。女性拟合直线的斜率(Slope=0.011 5)显著高于男性拟合直线的斜率(Slope=0.007 9), 说明身高对女性预期薪金的影响大于身高对于男性预期薪金的影响。

2.2 工作经验的影响

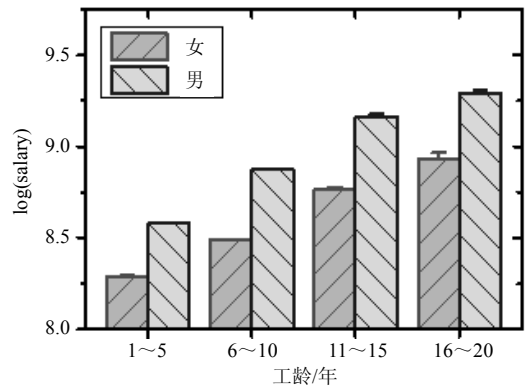
通过年龄和工龄(age 和 seniority)来近似估计求职者的工作经验, 两者数值越大, 则代表工作经验越丰富。如图3a所示, 随着平均年龄的增加, 平均预期薪金增加, 男女预期薪金的差异也增大。其中, 21~25岁男性平均预期薪金与26~30岁女性相当, 31~35岁男性平均预期薪金与41~45岁女性相当。如图3b所示, 以37岁左右为界, 预期薪金首先随着年龄的增加而线性增长, 然后稳中有增。



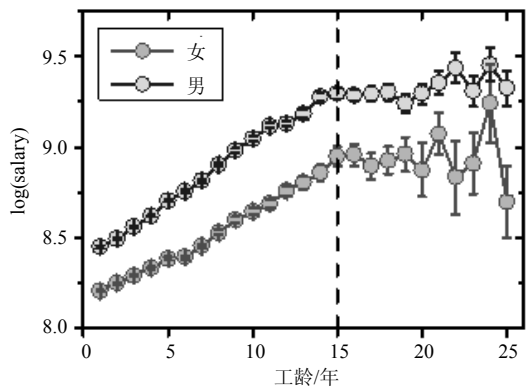
a. 不同年龄段的平均预期薪金



b. 预期薪金随着年龄的变化



c. 不同工龄段的平均预期薪金



d. 预期薪金随着工龄的变化

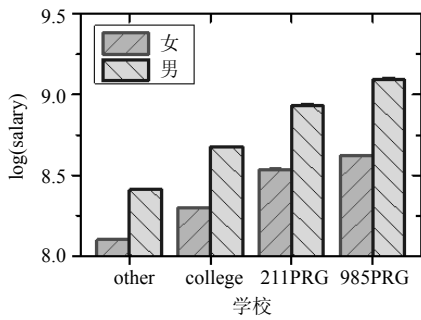
图3 工作经验对预期薪金的影响(误差线表示标准误差, 竖直虚线对应曲线拐点处的年龄和工龄)

工龄同样显著影响预期薪金。如图3c所示, 随着平均工龄的增加, 平均预期薪金增加, 男女的差异基本保持一致。如果男女要有相近的预期薪金, 那么女性比男性要增加5年以上的工龄。其中, 1~5年工龄男性平均预期薪金与6~10年工龄女性相当, 6~10年工龄男性平均预期薪金与11~15年工龄女性相当。如图3d所示, 以15年工龄为界, 预期薪金首先随着工龄增加而线性增长, 然后基本保持平稳。

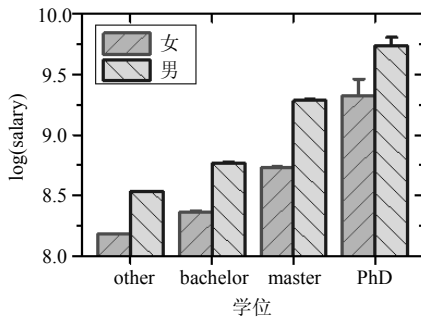
2.3 教育程度的影响

通过毕业院校 (school) 和获得学位 (degree) 来估计求职者的教育程度, 分析预期薪金受教育程度的影响。如图4a所示, 求职者的毕业院校越好, 平均预期薪金越高。毕业院校对男女预期薪金的影响存在显著差异, 普通高校(college)毕业的男性与“985工程”院校(985PRG)毕业的女性有相近的预期薪金, 其他院校(others)毕业的男性与“211工程”院校(211PRG)毕业的女性有相近的预期薪金。

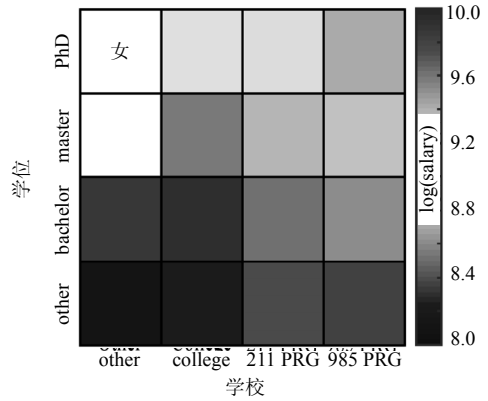
预期薪金显著受到获得学位的影响。如图4b所示, 随着学位的增加, 平均预期薪金增加。女性的预期薪金比男性相差一个学位, 如: 女性硕士(master)与男性本科(bachelor)的预期薪金相当, 女性博士(PhD)与男性硕士(master)的预期薪金相当。事实上, 毕业院校和获得学位共同影响预期薪金。如图4c和图4d所示, 工作经验对男女的影响有显著差异, 男性预期薪金普遍高于女性。其中, “985”工程院校毕业的女性博士与“211工程”院校毕业的男性硕士有相近的预期薪金。



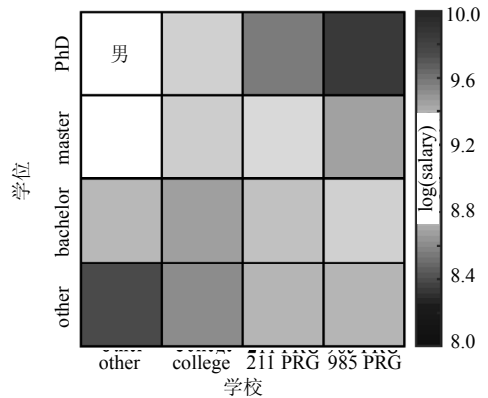
a. 预期薪金随着毕业院校的变化



b. 预期薪金随着获得学位的变化



c. 预期薪金随着工龄的变化

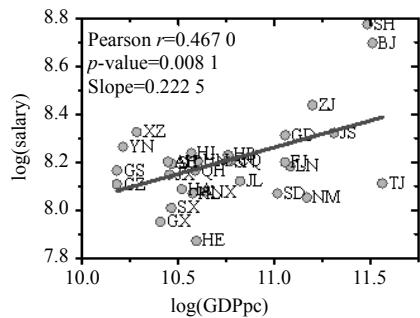


d. 毕业院校和获得学位共同对男性预期薪金的影响

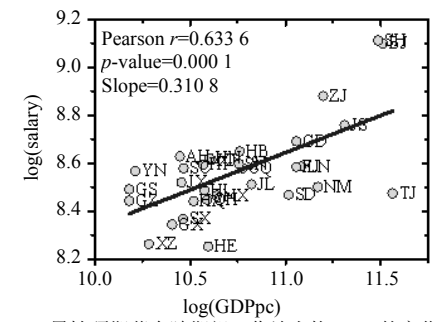
图4 教育水平对预期薪金的影响(a、b中条形图表示平均值, 误差线表示标准误)

2.4 地理经济的影响

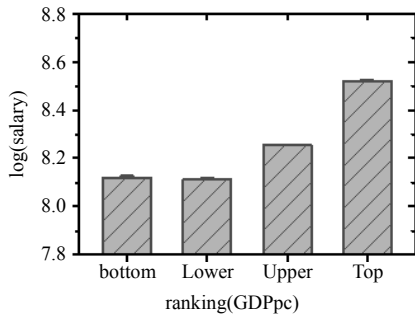
预期薪金与地理经济状况紧密相关。一般而言, 期望工作地的经济发展水平越高, 求职者的预期薪金越高。如图5a和图5b所示, 不论对于男性还是女性, 他们的平均预期薪金都随着期望工作地人均GDP的增加而增加, 即两者存在显著的正相关。其中, 男性的皮尔森相关系数($r = 0.6336$)大于女性的皮尔森相关系数($r = 0.4670$), 说明男性预期薪金与期望工作地经济发展水平更加相关。线性拟合的结果显示, 男性拟合直线的斜率(Slope = 0.3108)大于女性拟合直线的斜率(Slope = 0.2225), 说明期望工作地的经济水平对男性预期薪金的影响大于女性。



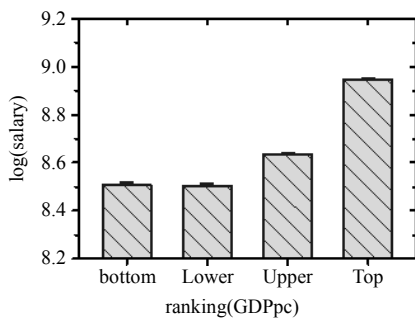
a. 女性预期薪金随期望工作地人均GDP的变化



b. 男性预期薪金随期望工作地人均 GDP 的变化



c. 期望工作地人均GDP排序对女性期望薪金的影响



d. 期望工作地人均GDP排序对男性期望薪金的影响

图5 地理经济对预期薪金的影响(a、b中实线为拟合直线,数据标签为省份缩写; c、d中期望工作地人均GDP从高到低等分为4段: Top、Upper、Lower和Bottom; 条形图表示平均值, 误差线为标准误)

为了更加细致地分析期望工作地的经济发展水平对求职者预期薪金的影响, 将预期工作地按照人均GDP从高到低等分为4段, 计算相应期望工作地的平均预期薪金。如图4c和图4d所示, 男性和女性的结果有同样的规律。具体而言, 求职者在经济发展水平后50%的期望工作地(即Lower和Bottom)有相同的预期薪金, 随着期望工作地经济发展水平提高到前50%(即Upper和Top), 预期薪金显著增加。

2.5 结果鲁棒性检验

为了检验分析结果的鲁棒性, 定量对比解释变量对男性和女性预期薪金影响上的差异, 在男女混合样本上采用多变量回归模型。回归分析中的被解释变量为预期薪金 $\log(\text{salary})$, 解释变量包括身高 (height)、工作经验 (age 和 seniority)、教育程度 (school 和 degree) 和地理经济 (GDPpc)。其中, school 和 degree 作为虚拟变量引入回归方程。

表格1总结了回归分析结果。在没有控制任何其他变量时, 从表格第1列看到, 身高对于男性预期薪金有显著正向的影响 ($\beta_1 = 0.0072$), 身高对女性的影响显著高于男性 ($\delta_1 = 0.0057$), 女性的平均预期薪金显著低于男性 ($\delta_0 = -1.2431$)。当只考虑工作经验的影响时, 从表格第2列看到, 年龄对男性有显著正向影响 ($\beta_2 = 0.0706$) 的同时, 工龄有显著负向影响 ($\beta_3 = 0.0030$); 年龄对女性的影响显著小于男性 ($\delta_2 = -0.0099$)。当同时考虑身高和工作经验的影响时, 如表格第3列所示, 男女的平均预期薪金没有显著差异。当增加考虑教育程度时, 从表格第4列看到, 身高的影响降低, 男女差异不显著; 年龄的影响降低, 工龄的影响增加但男女差异不显著。

表2 预期薪金的最小二乘法回归分析结果

变量	OLS Regression Model with Dependent Variable: log(salary)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
female	-1.2431*** (0.1281)	0.0182 (0.0265)	-0.1514 (0.1181)	0.0195 (0.1739)	-3.4696*** (0.0845)	-3.6280*** (0.1423)	-3.1455*** (0.1932)	-1.7180*** (0.1786)
height	0.0072*** (0.0005)		0.0116*** (0.0004)	0.0092*** (0.0004)		0.0054*** (0.0005)	0.0030*** (0.0004)	0.0075*** (0.0004)
female-height	0.0057*** (0.0008)		0.0020*** (0.0007)	0.0003 (0.0007)		0.0022*** (0.0008)	0.0010 (0.0007)	-0.0021*** (0.0006)
age		0.0706*** (0.0006)	0.0711*** (0.0006)	0.0481*** (0.0007)				0.0482*** (0.0007)
female-age		-0.0099*** (0.0010)	-0.0104*** (0.0010)	-0.0083*** (0.0011)				-0.0105*** (0.0011)
seniority		-0.0030*** (0.0007)	-0.0026*** (0.0007)	0.0205*** (0.0007)				0.0209*** (0.0007)
female-seniority		-0.0022* (0.0012)	-0.0023** (0.0012)	-0.0014 (0.0012)				-0.0002 (0.0012)

(续表)

变量	OLS Regression Model with Dependent Variable: log(salary)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDPpc					0.314 1*** (0.004 6)	0.311 9*** (0.004 6)	0.294 7*** (0.004 3)	0.307 7*** (0.003 9)
female-GDPpc					0.277 3*** (0.007 6)	0.265 2*** (0.007 7)	0.227 4*** (0.007 4)	0.193 4*** (0.006 7)
D^{school}	NO	NO	NO	YES	NO	NO	YES	YES
D^{degree}	NO	NO	NO	YES	NO	NO	YES	YES
Constant	7.485 5*** (0.081 7)	6.661 9*** (0.016 5)	4.636 1*** (0.075 7)	5.393 0*** (0.073 4)	5.227 2*** (0.051 0)	4.312 2*** (0.091 8)	4.583 9*** (0.087 4)	2.237 7*** (0.081 5)
Observations	141 064	141 064	141 064	141 064	141 064	141 064	141 064	141 064
F	4 914	10 348	7 658	3 663	8 148	4 960	2 543	4 052
Adj. R^2	0.094 6	0.268 3	0.275 4	0.330 3	0.147 7	0.149 5	0.234 5	0.376 2
RMSE	0.634 2	0.570 1	0.567 4	0.545 4	0.615 3	0.614 6	0.583 1	0.526 4

注: 回归分析在男女混合样本上使用基于最小二乘的多变量回归模型, female=0对应于基准的男性组。 D^{school} 和 D^{degree} 分别表示毕业院校和获得学位的虚拟变量。如果标记为YES, 则表示考虑在回归分析中; 如果标记为NO, 则表示没有考虑在回归分析中。表格中是变量的回归系数, 括号中是标准误。统计显著性水平为: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ 。

当仅考虑期望工作地经济发展水平时, 如表格第5列所示, 人均GDP显著正向影响平均预期薪金($\beta_0 = 0.314 1$), 对于女性的影响程度显著高于对于男性的影响程度($\delta_0 = 0.277 3$)。当增加考虑身高和教育水平的影响时, 这些结果得到保持(见表格第6列和第7列)。当考虑所有解释变量的影响时, 如表格第8列所示, 女性的预期薪金显著低于男性($\delta_0 = -1.718 0$); 身高、年龄、工龄和经济地理都对预期薪金有显著正向影响。其中, 身高和年龄对女性影响小于男性, 工龄对男女的影响没有显著差异, 经济地理对女性影响大于男性。所有的解释变量在一起, 能够最大程度的解释预期薪金, 解释能力可以达到 $\text{Adj. } R^2 \approx 0.376 2$, 即所有变量可以解释37.62%的预期薪金的变化。

3 结束语

本文基于求职者简历数据对预期薪金的影响因素进行了全面的分析。结果发现, 性别、身高、工作经验、教育程度和地理经济等因素显著影响求职者的预期薪金。具体而言, 男性的平均预期薪金高于女性; 身高与平均预期薪金正相关, 身高对女性的影响大于男性; 平均预期薪金随年龄和工龄的增加而增加, 但达到37岁年龄和15年工龄之后保持稳定; 毕业院校越好、获得学位越高, 预期薪金越高, 并且女性预期薪金与男性相比相差一个学位; 期望工作地经济发展水平越高, 求职者预期薪金越高。多变量回归分析结果显示, 性别对预期薪金的影响显著, 身高和年龄对女性影响小于男性, 工龄的影

响没有显著差异, 经济地理对女性影响大于男性。

单变量以及回归分析结果显示, 男性和女性在预期薪金上存在显著差异。如果男性和女性要达到相近的预期薪金, 在工作经验上, 女性需要比男性要增加5年以上的工龄, 如: 有6~10年工龄的女性与有1~5年工龄的男性有相近的预期薪金; 在教育程度上, 女性需要比男性多读一个学位, 如: 女性硕士与男性学士有相当预期薪金。

性别不平等一直是社会关注的热点之一。本文的分析结果也揭露出职场中潜在存在的性别不平等现象^[27]。在呼吁社会共同努力消除男女不平等的同时, 更希望女性同胞增加职场自信。在性别和身高等先天条件不能改变的情况下, 通过增加工作经验和教育经历来提升自身能力, 更加自信和主动地争取职场平等待遇。

参 考 文 献

- [1] EINAV L, LEVIN J. Economics in the age of big data[J]. Science, 2014, 346(6210): 1243089.
- [2] PENTLAND A. Social physics: How social networks can make us smarter[M]. London: Penguin, 2015.
- [3] BARABÁSI A L. Network science[M]. New York: Cambridge University Press, 2016.
- [4] 高见, 周涛. 大数据揭示经济发展状况[J]. 电子科技大学学报, 2016, 45(4): 625-633.
GAO Jian, ZHOU Tao. Big data reveal the status of economic development[J]. Journal of University of Electronic Science and Technology of China, 2016, 45(4): 625-633.
- [5] BLUMENSTOCK J, CADAMURO G, ON R. Predicting poverty and wealth from mobile phone metadata[J]. Science,

- 2015, 350(6264): 1073-1076.
- [6] HIDALGO C A, KLINGER B, BARABÁSI A L, et al. The product space conditions the development of nations[J]. *Science*, 2007, 317(5837): 482-487.
- [7] GAO J, JUN B, PENTLAND A, et al. Collective learning in China's regional economic development[EB/OL]. (2017-03-04). <http://arxiv.org/abs/1703.01369>.
- [8] SALESSES P, SCHECHTNER K, HIDALGO C A. The collaborative image of the city: Mapping the inequality of urban perception[J]. *PLoS One*, 2013, 8(7): e68400.
- [9] ALSHAMSI A, PINHEIRO F L, HIDALGO C A. Optimal diversification strategies in the networks of related products and of related research areas[J]. *Nature Communications*, 2018, 9(1): 1328.
- [10] GAO J. Maximizing the collective learning effects in regional economic development[C]//2017 14th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP). Chengdu, China: IEEE, 2017: 337-341.
- [11] LUO S, MORONE F, SARRAUTE C, et al. Inferring personal economic status from social network location[J]. *Nature Communications*, 2017, 8: 15227.
- [12] YUAN J, ZHANG Q M, GAO J, et al. Promotion and resignation in employee networks[J]. *Physica A*, 2016, 444: 442-447.
- [13] LLORENTE A, GARCIA-HERRANZ M, CEBRIAN M, et al. Social media fingerprints of unemployment[J]. *PLoS One*, 2015, 10: e0128692.
- [14] 高见, 张琳艳, 张千明, 等. 大数据人力资源: 基于雇员网络的绩效分析与升离职预测[M]//刘怡君. 社会物理学: 社会治理. 北京: 科学出版社, 2014: 38-56.
GAO Jian, ZHANG Lin-yan, ZHANG Qian-ming, et al. Big data human resources: Performance analysis and promotion resignation in employee networks[M]//LIU Yi-jun. *Social Physics: Social Governance*, Beijing: Science Press, 2014: 38-56.
- [15] PFEFFER J, LANGTON N. The effect of wage dispersion on satisfaction, productivity, and working collaboratively: Evidence from college and university faculty[J]. *Administrative Science Quarterly*, 1993, 38(3): 382-407.
- [16] HÜBLER O. The nonlinear link between height and wages in Germany, 1985-2004[J]. *Economics and Human Biology*, 2009, 7(2): 191-199.
- [17] SOHN K. The height premium in Indonesia[J]. *Economics and Human Biology*, 2015, 16: 1-15.
- [18] YANG X, GAO J, LIU J H, et al. Height conditions salary expectations: Evidence from large-scale data in China[J]. *Physica A*, 2018, 501: 86-97.
- [19] WEI X, TSANG M C, XU W, et al. Education and earnings in rural China[J]. *Education Economics*, 1999, 7(2): 167-187.
- [20] 马范文. 教育程度与工资收入相关性的中外比较[J]. 广州大学学报: 社会科学版, 2002, 1(4): 90-93.
MA Fan-wen. Correlation between education and salary abroad and in China: a comparative study[J]. *Journal of Guangzhou University (Social Science Edition)*, 2002, 1(4): 90-93.
- [21] GREGORIO J D, LEE J W. Education and income inequality: New evidence from cross-country data[J]. *Review of Income and Wealth*, 2002, 48(3): 395-416.
- [22] DUSTMANN C, MEGHIR C. Wages, experience and seniority[J]. *The Review of Economic Studies*, 2005, 72(1): 77-108.
- [23] ALSULAMI H. The effect of education and experience on wages: the case study of Saudi Arabia[J]. *American Journal of Industrial and Business Management*, 2018, 8(1): 129-142.
- [24] GAO J, ZHOU T. Quantifying China's regional economic complexity[J]. *Physica A*, 2018, 492: 1591-1603.
- [25] HEINECK G. Up in the skies? The relationship between body height and earnings in Germany[J]. *Labour*, 2005, 19(3): 469-489.
- [26] KIM T H, HAN E. Height premium for job performance[J]. *Economics and Human Biology*, 2017, 26: 13-20.
- [27] DING W W, MURRAY F, STUART T E. Gender differences in patenting in the academic life sciences[J]. *Science*, 2006, 313(5787): 665-667.

编辑 蒋晓