

教育错配对工资的惩罚效应

——来自中国微观面板数据的证据

颜 敏¹, 王维国²

(1. 东北财经大学 数学学院, 辽宁 大连 116025; 2. 东北财经大学 经济学院, 辽宁 大连 116025)

摘要:教育错配如何影响工资?文章基于2010—2014年中国家庭追踪调查面板数据,基于固定效应模型和工具变量估计,测度了教育错配产生的工资惩罚效应及其产生的理论机制。研究表明:过度教育员工每过度一年将承受1.4%的工资惩罚,教育不足员工将获得2.4%的工资红利。工作所需教育水平是40岁以上员工工资的主要决定因素;随着任职生涯的推进,教育错配会得到改善;教育错配与人力资本之间存在补偿机制。文章的政策涵义是:教育主管部门应从市场需求端改革教育体系,提供市场需要的相应技能毕业生。高校应主动顺应社会主义市场经济需求,及时有效地调整学科专业结构和人才培养结构。企业应建立人才识别和多层奖惩机制,充分释放过度教育者的劳动生产率水平。政府应加强就业指导工作,帮助求职者精准定位与其教育水平相匹配的工作,以减少搜寻成本。

关键词:教育错配;工资;惩罚效应;面板数据

中图分类号:F062.6 **文献标识码:**A **文章编号:**1001-9952(2018)03-0084-13

DOI: 10.16538/j.cnki.jfe.2018.03.007

一、引言

上世纪中期,伴随经济增长向技术和创新驱动方式转变,劳动力需求结构的调整引致许多发达国家高等教育扩张,随后这些国家劳动力市场出现了高等教育溢价降低的现象,Freeman(1976)将这一现象定义为过度教育(Over-Education):即大学生拥有的教育水平超过了其工作所需。过度教育本质上与失业类似,意味着人力资本的浪费与教育资源配置低效。中国自1999年开始高校扩招,高等教育规模迅速扩张,到2014年,包括网络、成人本专科招生数以及研究生招生数在内,我国高等学校在校生总计高达3 891.8769万人,高等教育的毛入学率高达39.39%,^①中国已经进入世界公认的高等教育大众化时期。但据国际劳动组织《就业报》以及麦可思《大学生就业报告》显示:伴随高等教育扩张,我国大学毕业生起薪持续走低,McGuinness(2006)、Leuven和Oosterbeek(2011)等基于实证分析框架证实了过度教育会对工资产生惩罚效应。即“过度教育会降低劳动者的工作积极性和工作效率,阻碍人力资本的有效利用,进而使教育回报降低”。

总体上看,目前针对过度教育工资惩罚效应的研究,总体上得到了相对统一的结论:恰好匹配工人挣得比相同教育水平但过度教育工人要多,但过度教育工人挣得比相同工作又恰好匹配

收稿日期:2017-04-27

基金项目:国家社会科学基金项目(13CJY010);国家自然科学基金项目(71773012)

作者简介:颜 敏(1975—),女,辽宁大连人,东北财经大学数学学院副教授;

王维国(1963—),男,吉林白城人,东北财经大学副校长,教授,博士生导师。

^①相关数据来自《中国统计年鉴》以及《联合国教科文数据库》。

的工人多。对于教育不足则呈现相反的结论：恰好匹配工人挣得比相同教育水平但教育不足工人要少，但教育不足工人挣得比相同工作且教育职位匹配工人要少。换言之，教育不匹配都有成本，因为其暗含着资源的低效率配置。

现有关于教育错配工资效应的研究大多是基于截面数据的分析，因而将面临以下两个重要的计量问题：

首先，不可观测能力因素是评估过度教育工资惩罚效应不可回避的关键问题。Chevalier (2003)研究表明：如果教育错配与不可观测能力之间相关，则最小二乘法估计结果将产生偏误。事实上，假设工资和教育由下列方程决定： $w = \beta S + \alpha A + \varepsilon$ ， $S = \gamma A + \xi$ ， w 是工资的对数， S 是个体获得学校教育年限， A 代表能力， ε 与 A 、 S 都不相关， ξ 与 ε 不相关， ξ 与 w 相关仅因其共同依赖能力因子 A ，但由于 A 无法观测，故基于 OLS 估计 $w = \beta S + \varepsilon$ 将产生 β 有偏估计量，因为 $p\lim(\beta_{ols}) = \beta + \alpha(\sigma_{AS}/\sigma_S^2)$ ，如果 $\gamma > 0$ ， $\sigma_{AS} > 0$ ， $\alpha > 0$ ，则 $\beta_{ols} > \beta$ 。正如Chavelier(2003)、Chevalier和Lindley(2011)的研究所指出的，不可观测能力异质性将大大缩小过度教育的工资惩罚效应。

在教育错配研究中不可忽视的另一个关键问题是教育错配的测量误差。事实上，本文在分别基于平均值法和众数法测算教育错配指标时，发现二者高度相关但却不完全相关，这暗含着存在测度误差。若设 $S_o = S_o^* + e$ ， S_o^* 为过度教育水平真值， S_o 为测算值，测量误差是 e 。若测算值与测量误差相关，但真实值与测量误差不相关，即 $cov(S_o^*, e) = 0$ ，在含误差变量经典假定下(下文简称 CEV)， $p\lim(\hat{\beta}_0) = \beta(\frac{\sigma^2_{S_o^*}}{\sigma^2_{S_o^*} + \sigma_e^2})$ ，由于 $\beta > 0$ ，因而真实的过度教育工资惩罚效应将由于存在测量误差而被低估(Wooldridge, 2002)。

正是基于上述考虑，本文采用微观面板数据固定效应模型，试图控制遗漏能力变量偏差，并进一步实施工具变量法以控制可能存在的测量误差。以期回答如下问题：中国教育错配对工资惩罚效应如何？人力资本补偿假设以及职业流动性假设在中国是否成立？二者之间是否存在相协调的年龄阈值？如何协调教育错配中的工作竞争理论和人力资本理论？进而为国家优化社会资源配置，提高生产率提供全新的政策视角。

二、模型方法、研究数据及变量确认

(一)模型方法。Duncan和Hoffman(1981)将接受的教育水平分解为工作所需教育年限、过度教育年限和教育不足年限，用式子表示如下：

$$S_{ait} = S_{rit} + S_{oit} - S_{uit}$$

其中，当 $S_{ait} > S_{rit}$ 时， $S_{oit} = S_{ait} - S_{rit}$ ，其他情况下 $S_{oit} = 0$ ；当 $S_{ait} < S_{rit}$ 时， $S_{uit} = S_{rit} - S_{ait}$ ，其他情况时， $S_{uit} = 0$ 。 S_{ait} 、 S_{rit} 、 S_{oit} 、 S_{uit} 分别表示接受的教育年限，工作所需教育年限，过度教育年限和教育不足年限。将其纳入经典Mincer工资方程得到如下模型(下文简称其为DH模型)：

$$w_{it} = \mu + \beta_r S_{rit} + \beta_o S_{oit} + \beta_u S_{uit} + \gamma X_{it} + \varepsilon_{it} \quad (1)$$

其中， w_{it} 是个体 i 在 t 年的实际工资对数， X_{it} 是影响工资的其他解释变量向量，包括所有制、行业、性别、健康、婚姻、经验以及地区特征等， ε_{it} 是残差项。 γ 是各控制变量对应的系数向量。 β_r 是需求教育年限的回报率， β_o 是过度教育年限(与同一职业中工作需要的教育年限相比)回报率($\beta_o > 0$)。 β_u 是教育不足年限回报率($\beta_u < 0$)。

Verdugo和Verdugo(1989)将接受的教育年限(S_e)(而非工作所需教育年限 S_r)过度教育变量和教育不足变量纳入工资方程，但教育错配变量以虚拟变量形式出现，为更好地识别教育错配对工资的惩罚效应，本文借鉴Hartog(2000)的做法，将教育错配虚拟变量转换为教育错配年限变量。因而，本文同时考虑如下模型(下文简称VV模型)：

$$w_{it} = \mu + \beta_a S_{ait} + \beta_o S_{oit} + \beta_u S_{uit} + \gamma X_{it} + \varepsilon_{it} \quad (2)$$

与DH模型意义不同,在VV模型中, β_a 代表的是接受教育水平的回报率, β_o 测度过度教育年限的工资惩罚($\beta_o < 0$),即与具有相同教育水平但教育职位恰好匹配的工人相比,过度教育年限每增加一年所要承受的工资下降百分比。反之, $\beta_u > 0$ 测度的是与具有相同教育水平但恰好匹配的工人相比,教育不足年限获得的工资嘉赏。

进一步,为控制遗漏能力变量偏误,分别基于DH模型和VV模型考虑如下面板模型:

$$w_{it} = \mu + \beta_r S_{rit} + \beta_o S_{oit} + \beta_u S_{uit} + \gamma X_{it} + (\varepsilon_{it} + \alpha_i) \quad (3)$$

$$w_{it} = \mu + \beta_a S_{ait} + \beta_o S_{oit} + \beta_u S_{uit} + \gamma X_{it} + (\varepsilon_{it} + \alpha_i) \quad (4)$$

$$\varepsilon_{it} \sim id(0, \sigma_\varepsilon^2)$$

由于不可观测能力 α_i 包含在复合残差项中,且其与解释变量 x_{it} 相关,OLS估计将得到有偏且非一致估计量, x_{it} 与 α_i 之间的相关程度越高产生的偏误越大。事实上,

$$\hat{\beta}_{ols} = \beta + \left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})(x_{it} - \bar{x})' \right]^{-1} \left\{ T \sum_{i=1}^N (x_{it} - \bar{x})(a_i - \bar{a}) \right\} \quad (5)$$

做一阶差分或者固定效应组内估计可以消除这种偏差。基于式(3)和式(4)的个体平均模型为:

$$\bar{w}_i = \mu + \beta_r \bar{S}_{ri} + \beta_o \bar{S}_{oi} + \beta_u \bar{S}_{ui} + \gamma \bar{X}_i + (\alpha_i + \bar{\varepsilon}_i) \quad (6)$$

$$\bar{w}_i = \mu + \beta_a \bar{S}_{ai} + \beta_o \bar{S}_{oi} + \beta_u \bar{S}_{ui} + \gamma \bar{X}_i + (\alpha_i + \bar{\varepsilon}_i) \quad (7)$$

进而(3)式减(6)式,(4)式减(7)式,恒定变量 μ 和能力变量 α_i 消失,得到:

$$w_{it} - \bar{w}_i = \beta_r (S_{rit} - \bar{S}_{ri}) + \beta_o (S_{oit} - \bar{S}_{oi}) + \beta_u (S_{uit} - \bar{S}_{ui}) + \gamma (X_{it} - \bar{X}_i) + (\varepsilon_{it} - \bar{\varepsilon}_i) \quad (8)$$

$$w_{it} - \bar{w}_i = \beta_a (S_{ait} - \bar{S}_{ai}) + \beta_o (S_{oit} - \bar{S}_{oi}) + \beta_u (S_{uit} - \bar{S}_{ui}) + \gamma (X_{it} - \bar{X}_i) + (\varepsilon_{it} - \bar{\varepsilon}_i) \quad (9)$$

正如前文所言,教育错配的测度误差会引起工资惩罚效应的估计偏误。现实中由于数据的限制,个体真实的过度教育年限 S_o^* 很难观测到,测量误差在所难免。对于规模未知、方向未知的测量误差偏误通常用工具变量法来校正(Green, 1997)。若设 S_o 是过度教育 S_o^* 的一种度量, $S_o = S_o^* + e_1$ (e_1 是测量误差),设 $\ln y = \beta_0 + \beta_1 S_o^* + \beta_2 X + \varepsilon$, $\ln y$ 是对数工资, X 是其他可观测控制变量,将 S_o^* 换成 S_o 得: $\ln y = \beta_0 + \beta_1 S_o + \beta_2 X + (\varepsilon - \beta_1 e_1)$,由于 S_o 与 e_1 相关而导致过度教育工资效应估计有偏和不一致。假设测量误差与观测不到的变量无关, ε 与 S_o^* , S_o 与 X 不相关,在CEV假定下, e_1 与 S_o^* , X 不相关,即 X 是外生的,但 S_o 与 e_1 相关。需要寻找 S_o 的工具变量,其必须与 S_o 相关,与 ε 不相关,并且与测量误差 e_1 不相关。一种解决办法是获取 S_o^* 的第二种度量,即 Z_1 ,既然影响 $\ln y$ 的是 S_o^* , Z_1 显然与 ε 不相关,若 $Z_1 = S_o^* + a_1$ (a_1 是 Z_1 的测量误差),假定 a_1 与 e_1 不相关,换言之, S_o 和 Z_1 都错误地测量了 S_o^* ,但他们的测量误差不相关。 S_o 和 Z_1 通过对 S_o^* 的相依而相关,因而可以用 Z_1 作为 S_o 的工具变量。^①

现有关于教育错配的测量方法,包括平均值法、众数法、工作分析法及自我评估法等各有千秋。任何一种测度都存在测量误差(颜敏等,2017)。本文根据数据的可得性和可靠性,选择平均值法、众数法以及外部评估法测度教育错配指标,发现各个测度指标有很强的相关性但却不完全相关,相关系数介于0.67—0.76之间,说明教育错配度量指标方差中的24%—33%来自于测量误差。因此本文主要基于平均值法测算教育错配指标,另外用众数法、工作分析法测算相应指标做工具变量来矫正测量误差。在假定不同方法测量误差不相关的前提下,额外控制了性别、婚姻、健康等个体特征,地域特征以及行业、职业、所有制等工作特征。此外,个体固定效应又控制

①[美]J·M·伍德里奇. 计量经济学导论现代观点. 北京:中国人民大学出版社,2003:466-467。

了个体不可观测能力因素,所以可以认为在控制了可观测和不可观测因素后工具变量与随机误差项不相关了。经 F 检验和 $Sargan$ 检验,工具变量满足相关性及外生性条件,测量误差可以通过本文的工具变量估计控制。

(二)研究数据。本文数据来自中国家庭追踪调查(*China Family Panel Studies, CFPS*)。CFPS 课题组于 2010 年正式启动,并将 2010 年基线调查的所有个体作为永久追踪对象,于 2012 年、2014 年跟踪调查,这使本文用面板估计技术控制遗漏能力偏误成为可能。数据集中包含详实的个体教育水平、收入、职业、行业、健康以及地区、城乡等特征,为教育错配的工资惩罚效应研究提供了可靠的数据资料。

为保证过度教育年限、教育不足年限两个关键指标测算的准确性,本文首先筛选出历年教育水平以及职业分类不缺失的样本,进一步将样本数太少及无法同其他职业合并的职业样本剔除,得 2010 年 15 997 个,2012 年 17 764 个,2014 年 22 814 个样本。基于上述样本分别测算历年职业需要教育水平和个体教育错配年限。再进一步除去健康、所有制、行业、收入缺失,年龄 21 周岁以下、60 周岁以上,年收入小于 5 000、大于 20 万的样本,2010 年剩余 2 968 个,2012 年剩余 3 611 个,2014 年剩余 6 255 个。然后按个体编码匹配,保留三年全部跟踪的个体,最终形成每年 1 865 个,3 年共计 5 955 个样本的面板数据。

(三)变量的确认。

1. 工资收入。本研究中的收入指把所有工资、奖金、现金福利和实物补贴都算在内,并扣除税和五险一金后过去一年的总收入。将其取自然对数,用 $\ln wage$ 代表。

2. 接受的教育水平(Sa)。CFPS 数据库提供了个体完成的最高教育程度数据,根据我国各层次教育特点,定义文盲半文盲教育年限为 1,小学教育年限为 6,初中教育年限为 9,高中、中专、技校、大专的教育年限为 12,大专教育年限为 15,大学本科教育年限为 16,硕士教育年限为 19,博士教育年限为 22。

3. 教育错配变量。本文旨在测度教育错配对工资的惩罚效应,因此模型中用到的是教育错配年限变量。为设法规避教育错配变量的测量误差,本文借鉴 Verdugo 和 Verdugo(1989)、Kiker 等(1997)分别提出的平均值法(*mean*)和众数法(*mode*)以及基于《国际职业分类标准》、《国际教育分类标准》的工作分析法(*JA*)测算各教育错配指标。具体地,平均值法计算每一职业教育年限平均值和标准差,设定均值的一个标准差范围为适度教育(Sq),超出部分为过度教育年限(So),低于部分为教育不足年限(Su)。并且认为每一职业的平均教育年限为该职业需要的教育年限(Re)。众数法认为每一职业员工教育年限的众数即表示适度教育水平或该职业需要的教育水平(Sqm),将实际教育年限与职业所需教育年限的差定义为过度教育年限(Som),职业需要教育年限与实际教育年限的差定义为教育不足年限(Sum)。工作分析法的思想来自于劳动力市场专家的分析,由劳动力市场一些专家对各种职业所需教育水平(Sqj)进行评价,超出专家设定的教育年限为过度教育年限(Soj),低于专家评定的教育水平为教育不足年限(Suj)。

CFPS 数据库依据《中华人民共和国标准职业分类与代码》(*GB/T6565-2009*)将职业划分为 8 个大类,83 个中类,595 个小类。本研究首先将教育水平指标和职业类别指标都没有缺失的样本筛选出来,进而将每年样本较少又无法与其他类别合并的职业样本剔除,最终基于 46 个中类职业测算职业所需教育年限,进而估算个体教育错配年限。^①

^①本研究中绝大多数职业年均样本数超过 100,其中种植业生产人员、工程施工、商业服务业、运输设备操作等职业超过 1 000,印刷工人、运输服务、邮政电信、文化艺术工作人员等职业样本数较少,但历年超过 40,这为准确测量职业所需教育年限提供了保证。限于篇幅并未列出这些样本数据,有需求读者可向作者索取。

3. 经验(*Ex*)。CFPS数据中并没有提供工作经验的度量指标,本文用年龄-6-教育年限代理,对于教育年限低于12年的,用年龄-18代理。同时将经验平方(Ex^2)纳入工资模型中。

4. 健康(*Health*)。健康人力资本是影响工资效率的重要因素之一。CFPS数据库详细询问了个体的健康状况,包括非常不健康、比较不健康、不健康、一般、健康、比较健康、很健康、非常健康。将其分别赋值1-8,数值越大代表其越健康。

5. 其他控制变量。借鉴相关文献,本研究工资模型中还包括婚姻虚拟变量(*Married*):设定在婚为1,其他未婚、丧偶、离婚、同居为0。性别虚拟变量(*Male*):男性为1,女性为0。设置东部、中部地区虚拟变量(*Dong*、*Zhong*),基础组为西部。设定三个单位所有制虚拟变量:港澳台投资企业、外商投资企业设为1,其他为0(*Gat*),国有、集体事业单位、政府机关为1,其他为0(*Gj*),国有企业为1,其他为0(*Gq*),基础组为个体、私营、民办以及股份合作、有限责任公司。设置5个行业虚拟变量:电力、燃气、采矿、交通仓储为1,其他为0(*Cd*);建筑、房地产为1,其他为0(*Fj*);教育、金融、信息、科学、文化传播为1,其他为0(下文中用*Jj*表示);制造业为1,其他为0(*Zj*);公共管理、卫生、社保、租赁、商务服务为1,其他为0(*Gj*);基础组是批发零售、住宿餐饮行业。设置户籍虚拟变量(*City*):城市户口为1,农村户口为0。表1给出了关键变量的描述统计。

表1 关键变量的描述性统计

	平均值	标准差(整体)	标准差(组间)	标准差(组内)	最小值	最大值	观测值
<i>lnwage</i>	10.037	0.797	0.633	0.485	6.194	14.000	5 595
<i>Sa</i>	11.147	3.765	3.766	0.000	1	22	5 595
平均值法测算的教育错配指标							
<i>Sq</i>	11.024	2.334	2.137	0.940	6.364	16.2	5 595
<i>So</i>	0.166	0.596	0.498	0.328	0	6.7874	5 595
<i>Su</i>	0.285	1.034	0.985	0.317	0	10.1569	5 595
众数法测算的教育错配指标							
<i>Sqm</i>	11.550	2.940	2.633	1.309	9	16	5 595
<i>Som</i>	0.856	1.715	1.484	0.861	0	10	5 595
<i>Sum</i>	1.250	2.269	2.059	0.954	0	14	5 595
工作分析法测算的教育错配指标							
<i>Sqj</i>	11.64307	2.5059	2.123	1.353	6	16	5 595
<i>Soj</i>	0.96103	1.5953	1.046	0.978	0	10	5 595
<i>Suj</i>	1.45612	2.0971	1.978	1.025	0	13	5 595

不难看出平均值法和众数法以及工作分析法测算的各教育错配指标整体模式一致,但是基于众数法以及工作分析法测算的各种教育错配指标具有较大的标准差。这与众数法、工作分析法基于点估计的特征有关。事实上,当基于2010年^①数据比较平均值法和众数法测度的各类结果时,我们发现67.37%的面板观测点融入同一种类(表2)。这为后文工具变量的引入奠定了基础。

表2 基于平均值法和众数法测算的教育错配指标的融合度(%)

众数法平均值	过度教育	适度教育	教育不足	总计
过度教育	0.1822	0.0004	0	0.1826
适度教育	0.1378	0.3794	0.1584	0.6756

^①本研究也测算2012年以及2014年两种方法测算的教育错配的融合度,其中2012年为61.56%,2014年为61.33%。2012年以及2014年教育错配发生率与2012年大致相近。基于工作分析法与众数法测算的各个指标也具有相似的融合度,限于篇幅本文未列示这些测算结果。

续表 2 基于平均值法和众数法测算的教育错配指标的融合度 (%)

众数法平均值	过度教育	适度教育	教育不足	总计
教育不足	0.0016	0.0281	0.1121	0.1418
总计	0.3216	0.4079	0.2705	1

注：行表示基于平均值法测算的过度教育、适度教育、教育不足的比例；列表示基于众数法测算的过度教育、适度教育、教育不足的比例；对角线数值的和(67.37%)即为二者相融比例。

三、实证结果

(一)基于明瑟方程教育回报率估计。本文首先测算了各个截面时点个体教育回报率演变,进而基于面板数据估计了混合模型、随机效应模型以及固定效应模型以展示并控制可能存在的个体异质性和测量误差。具体估计结果呈现在表 3、4 中。

由表 3 可见,2010 年样本期教育回报率最高,达 7.23%,意味着个体教育年限每增加一年,将使工资上涨 7.23%,但 2012 年下降到 4.83%,2014 年又回升到 5.44%,教育回报呈现波动性下降的态势。颜敏和王维国(2014)的研究表明:中国近些年已经呈现出技能偏向特征,说明我国对高技能劳动力需求是增加的,那么在大学劳动力迅猛增加的同时教育回报率的下降意味着劳动力市场配置效率的下降。表 3 后两列是基于面板数据混合模型以及随机效应广义最小二乘估计的结果。*pool-ols* 估计结果以及随机效应估计与前述各年教育回报率相差不大,其中 *pool-ols* 估计三年平均教育回报率为 6.74%,但是 *F* 检验结果表明存在明显的个体异质性,因而基于同质性假设的混合最小二乘估计是不一致的。随机效应估计结果显示,教育回报率轻微下降到了 6.1%。

表 3 基于明瑟方程教育回报率估计^①

	2010 年	2012 年	2014 年	<i>Pool-ols</i>	<i>Random-effect</i>
<i>Sa</i>	0.0723 ^{***} (0.005)	0.0483 ^{***} (0.0059)	0.0544 ^{***} (0.0055)	0.06540 ^{***} (0.0032)	0.061 ^{***} (0.0044)
<i>ex</i>	0.0294 ^{***} (0.007)	0.0113(0.0086)	0.0102(0.0084)	0.0258 ^{***} (0.0045)	0.0380 ^{***} (0.0058)
<i>Ex²</i>	-0.0006 ^{***} (0.0002)	-0.0003(0.0002)	-0.0003(0.0002)	-0.0005 ^{***} (0.0001)	-0.0007 ^{***} (0.0001)
<i>Health</i>	0.0251(0.0216)	0.0183(0.0132)	0.0224 [*] (0.0123)	0.0877 ^{***} (0.0071)	0.1105 ^{***} (0.007)
<i>Married</i>	0.2230 ^{***} (0.0527)	-0.0023(0.0654)	-0.0028(0.0632)	0.1087 ^{***} (0.0354)	0.0434(0.0408)
<i>Male</i>	0.3328 ^{***} (0.032)	0.3091 ^{***} (0.0380)	0.2979 ^{***} (0.0355)	0.2959 ^{***} (0.0208)	0.2908 ^{***} (0.0277)
<i>R²</i>	0.2592	0.1399	0.1722	0.1915	0.1867

注:1.()里是估计的标准差,其中面板模型是基于个体层面的聚类稳健标准差(下同)。2.由于本研究中关键变量个体教育年限 *Sa* 没有包含在职培训、干中学等非学校教育,样本期内不变,因此 *IV* 模型并没有考虑固定效应估计。3.随机效应估计的 *R²* 是整体 *R²*。混合模型中个体同质性假设 *F* 统计量值是 54.76, *P*=0.000。4.*, **, ***分别代表在 10%, 5%, 1% 水平下显著。下表同。

(二)教育错配对工资的惩罚效应估计。表 4 呈现的是基于 *ORU* 模型教育错配对工资的惩罚效应估计。其中第 2、3、4 列是基于 *DH* 模型的估计结果,第 5、6 列是基于 *IV* 模型估计的教育错配对工资的惩罚效应。

表 4 基于 *ORU* 模型教育错配对工资惩罚效应估计

	<i>DH</i> 模型			<i>IV</i> 模型		
	混合回归	随机效应	固定效应	<i>IV</i> 固定效应	混合回归	随机效应
<i>Sa/Sq</i>	0.0895 ^{***} (0.0045)	0.0777 ^{***} (0.0052)	0.014131(0.02627)	0.1061 ^{**} (0.0239)	0.0858 ^{***} (0.0053)	0.0776 ^{***} (0.0051)
<i>So</i>	0.0682 ^{***} (0.0064)	0.0734 ^{***} (0.0073)	0.011965(0.026729)	0.0922(0.096)	-0.0224 ^{***} (0.0081)	-0.0070 [*] (0.0024)

①限于篇幅本文并未列出地区、所有制以及行业等控制变量的估计结果,有需求的读者可向作者索取。

续表4 基于ORU模型教育错配对工资惩罚效应估计

	DH模型				IV模型	
	混合回归	随机效应	固定效应	IV固定效应	混合回归	随机效应
<i>Su</i>	-0.0396***(0.0046)	-0.0467***(0.0054)	-0.0035(0.0103)	-0.0826(0.0809)	0.0445***(0.0066)	0.0256***(0.0061)
<i>Ex</i>	0.0295***(0.0045)	0.0388***(0.0055)	0.152767***(0.0118)	0.1547***(0.0119)	0.0288***(0.0061)	0.0390***(0.0058)
<i>Ex</i> ²	-0.0006***(0.0001)	-0.0007***(0.0001)	-0.00145***(0.00024)	-0.0024***(0.0003)	-0.0006***(0.0001)	-0.0007***(0.0001)
<i>health</i>	0.0883***(0.0071)	0.1108***(0.0066)	0.054393***(0.0087)	0.0538***(0.0086)	0.0880***(0.0079)	0.1104***(0.0070)
<i>Married</i>	0.0984***(0.0352)	0.0443(0.0388)	-0.11766***(0.0535)	-0.11856****(0.0535)	0.0992***(0.0414)	0.0425(0.0407)
<i>male</i>	0.3175****(0.0208)	0.3038****(0.0275)	—	—	0.3159****(0.0275)	0.3018****(0.0276)
<i>R</i> ²	0.2028	0.1989	0.1660	0.1551	0.2015	0.1946
人力资本理论检验	<i>F</i> =40 <i>P</i> > <i>F</i> =0.00	$\chi^2(2)$ =25.7 <i>P</i> > χ^2 =0.00	<i>F</i> =0.38 <i>P</i> > <i>F</i> =0.684	$\chi^2(2)$ =0.039 <i>P</i> > χ^2 =0.96	<i>F</i> =34.12 <i>P</i> > <i>F</i> =0.00	$\chi^2(2)$ =19.71 <i>P</i> > χ^2 =0.00
工作竞争理论检验	<i>F</i> =105 <i>P</i> > <i>F</i> =0.00	$\chi^2(2)$ =173.2 <i>P</i> > χ^2 =0.0	<i>F</i> =0.34 <i>P</i> > <i>F</i> =0.715	χ^2 =0.628 <i>P</i> > χ^2 =0.71	<i>F</i> =101.6 <i>P</i> > <i>F</i> =0.00	$\chi^2(2)$ =180.3 <i>P</i> > χ^2 =0.00

注: 1. 豪斯曼检验的卡方统计量 $\chi^2(19) = 1034.9, p > \chi^2(19) = 0.0000$ 。2. 人力资本理论检验在DH模型等价于 $\beta_r = \beta_o - \beta_u$, 在IV模型中等价于 $\beta_o = \beta_u = 0$; 工作竞争理论检验在DH模型中等于检验 $\beta_o = \beta_u = 0$, 在IV模型中等价于检验 $\beta_a = -\beta_o = \beta_u$ 。

首先由IV模型的估计结果可见, 加入过度教育年限和教育不足年限后, 混合模型中对接受教育水平的回归从6.54%升到8.58%, 即当员工的教育水平与其职位相匹配时的教育年限每提高一年, 工资将增长8.58%。由IV模型可知, 过度教育每增加一年, 员工要承受2.24%的工资惩罚, 而教育不足员工将得到的4.45%的工资嘉赏。与混合估计结果相比, 本研究关键变量的广义最小二乘估计结果都存在一定幅度下降。但由于不可观测个体异质性的存在, 无论pool-ols估计还是随机效应估计都将有偏且不一致。

对于DH模型, 本文结论与Korpi和Tählin(2007)相同。 $\beta_s > 0, \beta_o > 0, \beta_u < 0, |\beta_s| < \beta_o, \beta_o < \beta_r$ 。相同教育水平下, 恰好匹配工人比过度教育工人挣得多($\beta_s > \beta_o$), 若从事相同工作, 过度教育工人挣得比恰好匹配工人多($\beta_o > 0$)。反之, 相同教育水平下, 恰好匹配工人比教育不足工人挣得少($\beta_s + \beta_u > 0$), 若从事相同工作, 教育不足工人挣得比恰好匹配工人少($\beta_u < 0$)。具体在混合模型中, 工作所需的教育水平回报率为8.95%, 与IV模型的估计结果具有很强的一致性。过度教育年限回报率为6.82%。过度教育者将面临每年2.13%的工资惩罚。教育不足年限回报率为-3.96%, 说明教育不足工人将比相同工作适度教育工人少挣3.96%的工资, 但与相同教育水平且与工作匹配的工人相比, 工资将多赚近5%(8.95%-3.96%)。同样如果存在不可观测能力偏误, 无论基于pool-ols估计还是随机效应估计都将有偏且不一致, 豪斯曼检验结果显示 $\chi^2(19) = 1034.9, p > \chi^2(19) = 0.0000$, 说明模型中确实存在与解释变量相关的不可观测异质性, 因而固定效应模型结果更可靠。

值得注意的是, 表4固定效应估计结果中关键变量系数大幅减小, 其中工作所需教育年限回报下降到1.41%, 过度教育年限回报下降到1.19%, 教育不足年限的工资效应下降到-0.35%。换言之, 能力偏误使教育错配的工资效应产生向上偏误, 其中过度教育工资惩罚效应由2.1%(0.0895-0.0682)下降到0.22%(0.0141-0.00119), 教育不足的工资红利从5%(8.95%-3.96%)下降到1.06%(0.01413-0.0035)。Tsai(2010)及Iriando和Pérez-Amaral(2015)分别基于美国、欧洲劳动力数据的分析也得到类似的结论。可能的解释是: 首先, 我们的样本期是高等教育持续扩张十年后, 劳动力教育年限普遍增长, 但是教育获得水平和能力获得水平之间的关系变得模糊不清, 高层次教育不再是高能力可靠的信号, 这意味着过度教育形成的工资惩罚效应可能更多反映的是不可观测的能力异质性, 过度教育工人工资比完全匹配工人工资低的真正原因是他们能力偏低, 并被分配到与其边际产量相匹配的职位。因而当控制了不可观测异质性后, 过度教育的工资

惩罚效应将大大减少。另外,本研究中面板的时期较短,不足以捕捉到诸如能力等个体异质性引致的所有教育错配效应。由于时期短,很多控制变量在样本期内变化不大,比如地区变量2010年隶属东部地区的个体在后续两年有变化的只有15个,因而在固定效应中并未考虑地区、城市变量,固定效应模型无法成功地测度例如性别不随时间改变的因素对工资的影响。最后,更重要的是当用不同方法测度教育错配时,测量误差在所难免,测度误差会低估教育错配的工资惩罚效应。

为进一步控制测量误差,我们对固定效应面板模型实施了工具变量(IV)估计,并严格检验工具变量有效性和外生性。^①当使用众数法和工作分析法测度指标作为平均值法相应测度指标的工具变量时,系数的绝对值明显增大。工作所需教育年限回报率上升到10.61%,过度教育年限回报率由0.019上升到0.092,教育不足年限回报率从-0.0035进一步下降到-0.0826,因此不可观测异质性引致的向上工资偏误大部分被测量误差引致的负偏误补偿了。但过度教育年限回报并不显著,说明过多的教育年限并未呈现显著的工资效应,过度教育工人承受至少1.4%的工资惩罚效应(0.1061-0.922),而教育不足工人的工资红利上升到2.4%。这与Verhaest等(2012)基于弗兰德年轻工人的估计结果相当。表4第4、5列估计结果表明,教育错配工资效应估计中,能力遗漏和测量误差都不容忽视。这与Dolton和Silles(2008)基于英国大学生面板数据的研究结论相似,其研究表明,工人异质性导致工资惩罚的向上偏误几乎被测量误差引致的向下偏误完全抵消。但与Verhaes等(2012)的研究结论不同,其估计结果表明能力偏差几乎可以被忽略。当然由于样本不同,国家间教育体制、就业环境以及大学生能力的差别,估计结果间不具有完全的可比性。

由此可见,在测度教育错配的工资效应时,能力偏差以及测量误差都是至关重要的因素。当控制了测量误差和遗漏变量偏误后,工作需要的教育回报上升到10.61%,但教育错配年限并未产生显著的工资效应,说明教育错配都有成本,将面临一定程度的工资惩罚。因而工资主要依赖于工作需要的教育水平,表4最后两行的卡方检验结果也验证了这一论断。那么在中国,教育错配工资效应产生的内在机制是什么呢?表面上观测到的教育错配效应(混合模型和随机效应估计)仅仅反应的是个体间不可观测的异质性和测量误差吗?是员工职业生涯早期的短暂现象还是会随着职业过程和工作流动性而改善呢?看似过度的工人除了不可观测能力偏低外,那些诸如表达能力、理解能力、健康资本又如何?是否存在人力资本补偿机制呢?

四、教育错配对工资惩罚效应的理论检验

对过度教育产生的工资惩罚效应的理解有四种主流观点:人力资本理论(Becker, 1964)认为,工资是按照边际产量计算的,增加人力资本投入会提高劳动生产率,工资因而会增加。在教育上投资多少直接影响着劳动生产率。但如果某职位员工的供给超过需求,结果就会导致工资暂时性下降,因而过度教育会对工资产生惩罚效应。但需求和供给间的错配并不会导致技能的不充分利用,因而相对工资下降只是暂时性的(Green等, 1999)。工作竞争理论(Thurow, 1975)认为,边际产量和工资是由工作性质决定的,而非员工特质决定的。工人间为获得工作机会而竞争,并不是为工资而竞争。因而很多工人会通过在职培训途径获得更多技能以增加自己的工作机会。雇主们会按照学历、经验、年龄等因素将工人进行潜在排名,为了减小培训成本,雇主会更倾向于优先选择排名靠前的工人,因此工人得到的工作机会,与其在排序中的相对位置一致。排

^①正如Verhaest和Omey(2006)所言,理想的工具变量应该是从事这份工作需要的教育水平,但针对中国样本没有这样的调查数据,本文中,当严格实施工具变量有效性和可靠性检验时,一阶段弱工具变量 F 检验值都大于10,且 P 值都为0,即不存在弱工具变量问题,进一步实施过度识别约束的Sargan检验, χ^2 统计量值为0.23, P 值为0.899,因而没有理由拒绝工具变量外生性的原假设。

名最前的工人得到最好的工作。但如果高收入职位相对不足,高学历劳动力的供给量相对过多,高学历劳动力中的一部分将被迫接受仅需要较低教育水平就能从事的职位,并且他们将得到相应于这些职位的报酬,因而,与相同教育水平但从事恰好匹配职位的工人相比,其工资偏低,过度教育对工资产生了惩罚效应。人力资本补偿理论强调正规学校教育不是个体生产能力的完美度量,个体不可观测异质性解释了工资差异的大部分,过度教育工人可能用他们剩余的教育补偿了其他方面人力资本的缺陷,因此,过度教育者与那些拥有相同教育水平但与工作恰好匹配的工人相比更缺乏效率,因而会比恰好匹配的工人获得较少的报酬(Korpi 和 Tählin, 2007)。职业流动性理论认为,过度教育是真实存在的,但只是短暂的,过度教育水平的上升仅限于在工人职业生涯的早期阶段,这与整体劳动力市场中的技能工资差距相协调,人力资本投资战略和劳动力市场搜寻一样可能在初期产生不匹配,并且会产生对工资的惩罚效应,然而随着职业过程和工作流动性,未来工资高速增长会补偿初始的工资惩罚。

表4的最后两行分别基于两种模型检验人力资本理论和工作竞争理论。由IV模型估计结果可知:当控制了接受的教育水平,人力资本理论认为过度教育年限和教育不足年限对工资没有重要的影响($\beta_o = \beta_u = 0$),无论在混合面板估计还是随机效应估计中均被拒绝。工作竞争理论认为,工资大小依赖于工作需要的教育水平而非工人获得的教育水平。即($\beta_a = -\beta_o = \beta_u$)。同样在混合估计和随机效应估计中被拒绝。从DH模型估计结果可得相似的结论:无论在混合模型还是随机效应模型设定中,人力资本理论($\beta_r = \beta_o = -\beta_u$)和工作竞争理论($\beta_o = \beta_u = 0$)均被拒绝。但固定效应估计及基于固定效应模型的工具变量估计显示,无论人力资本理论还是工作竞争理论都不能被拒绝。即控制了能力偏误及测量误差时,人力资本理论和工作竞争理论在中国成立。工作所需的教育水平和接受的教育水平都是影响工资的重要因素。

Iriondo 和 Pérez-Amaral(2015)根据欧盟国家的数据研究表明:员工在一定年龄之前,其工资取决于个人的受教育水平,达到一定年龄后,其工资取决于职业所需教育水平。刚进入劳动力市场时,工资水平依赖于教育和工作经验,老板付工资多少主要看他是谁,随着职业生涯推进,老板付工资会看他做了什么,即在长期职业生涯中,工作所需教育水平会逐渐对工资产生主导决定作用。只有在进入劳动力市场初期,教育水平才是员工生产率的代理指标。那么,中国劳动力市场是否存在这样的年龄阈值?在阈值前后教育对工资作用机制有何不同?表5给出了分年龄组基于DH模型的教育错配工资效应估计。如表5所示,当我们把样本分成40岁以上和40岁以下年龄组时,教育对工资的作用机制完全不同。40岁以下群组,教育对工资作用变得模糊不清,不仅职业所需平均教育年限的提高未对工资产生显著的向上拉动,过度教育回报率竟然显著为负。可能的原因是这部分群体大多数在学校期间经历了中国高等教育扩张,过度教育员工可能更多意味着能力偏低,人力资本会因为技能缺失而贬值。与此相反,40岁以上员工群体,工资不

表5 分年龄教育错配对工资惩罚效应工具变量估计

	Age ≤ 40	Age > 40
S_i	0.071(0.6256)	0.205***(0.060)
S_o	-0.582*** (0.0494)	0.122*** (0.029)
S_u	0.183(0.2281)	-0.109(0.099)
constant	9.975*** (0.5908)	7.564*** (0.524)
人力资本理论检验 $\beta_r = \beta_o = -\beta_u$	$\chi^2(2) = 21.53 P > \chi^2(2) = 0.000$	$\chi^2(2) = 23.53 P > \chi^2(2) = 0.000$
工作竞争理论检验 $\beta_o = -\beta_u = 0$	$\chi^2(2) = 24.78 P > \chi^2(2) = 0.000$	$\chi^2(2) = 1.28 P > \chi^2(2) = 0.542$
R^2	0.0985	0.1489

仅取决于接受的教育水平,工作所需的教育水平对工资作用更强更显著。职业所需教育年限每提高1年,工资将提升20%。工资形成机制的人力资本理论对40岁以上群体不成立,但工作竞争理论成立。说明工作所需教育对我国40岁以上员工工资起了主导作用。换言之,随着员工工作年限的增长,员工的工资逐渐取决于他们现有的生产率和技能,而不是取决于他们在离开教育系统时所拥有的教育水平。这可以解释为只有那些在工作中不断学习的员工,他们的人力资本才会继续保持或增加,而那些技能未被充分利用、没有经历干中学或者经历事业中断的员工,他们的人力资本可能贬值。换言之,具有相同教育水平的个体,随着工作年限增长,他们之间的技能和生产率异质性也会越来越大。

事业中断可能是人力资本贬值的一个影响因素。接下来本研究基于40岁及40岁以上样本,估计一个多元Logit模型阐述职业流动性对教育错配的影响。模型中除了接受的教育年限(*Sa*)、性别(*gender*)、健康状况(*health*)以外,更关注职业变动对教育过度和教育不足(基础组是适度教育)发生率的影响,因而引入以下两个变量:跳槽换工作(*Jump*),跳过槽*Jump*取值为1,否则取值为0;本文中跳槽是指主动跳槽,主动离开在任的职位和机构,到其他机构工作是自愿换工作,因而几乎没有空档期。鉴于数据的可得性,本文考虑的另一个雇佣状态改变指标是辞职(*Demission*),辞过职*Demission*取值为1,否则为0。本研究中辞职指非自愿辞职,工作所在机构通过解聘等方式让其离开在任的职位和机构,因而暗含着一定的人力资本损失。

表6中,我们计算了在两种最易受教育错配影响的较高教育水平(教育年限为16年)和较低教育水平(教育年限为9年)上各个因素对教育错配发生的边际影响。如预期一样,教育水平越高越容易过度教育,工作经验越长越不容易发生过度教育。教育不足则呈现相反的结论。男性过度教育的可能性更大而教育不足可能性偏小。身体越不健康越可能过度教育。被动辞职的员工过度教育可能性更大,教育不足的可能性更小。相对而言,主动跳槽的员工不容易发生过度教育。因而随职业生涯推进,员工通过不断寻找与自己教育技能相匹配的工作,教育错配会得到改善,这与Burdett(1978)提出的职业流动性假设相一致。

表6 职业流动性理论检验

$dy/dx(Sa=16)$		$dy/dx(Sa=9)$	
解释变量	教育过度	解释变量	教育不足
<i>Sa</i>	0.106*** (0.002)	<i>Sa</i>	-0.026*** (0.002)
<i>gender</i>	0.118*** (0.026)	<i>gender</i>	-0.014*** (0.005)
<i>exp</i>	-0.011*** (0.001)	<i>exp</i>	0.001*** (0.0001)
<i>health</i>	-0.043*** (0.02)	<i>health</i>	0.002 (0.003)
<i>Jump</i>	-0.038 [†] (0.034)	<i>Jump</i>	0.012 [†] (0.008)
<i>Demission</i>	0.106 [†] (0.066)	<i>Demission</i>	-0.039** (0.016)

最后,本研究将进一步检验人力资本补偿理论在中国劳动力市场的存在性。将教育错配各指标与健康、语言表达能力和理解能力等人力资本指标建立联系。其中理解能力和语言表达能力都赋值1-7,数值越大表示能力越强。^①基于统计数据的可得性,在此基于2014年CFPS的截面数据进行检验。表7呈现了基于DH模型人力资本补偿假设的检验结果。不难发现工作所需

^①Korpi 和 Tählin(2009)基于瑞典数据选择经验代理一般技能,检验教育水平和经验之间的互补关系。本文舍弃这一指标的原因是CFPS并没有统计个体参加工作时间等信息,因而前文工资方程中我们用年龄-教育年限-6代理,其本身就与教育年限正相关,因而用其检验过度教育与人力资本补偿之间的关系是不可靠的。

教育水平和口头表达能力、理解能力均呈现显著正相关关系,过度教育年限对两种能力的提升作用表现得更强更显著,高于具有相同教育水平但教育职位匹配的工人(0.095>0.047, 0.132>0.098)。反之,相同职位恰好匹配工人比教育不足工人能力更强。表7最后一列呈现了健康和错配之间的关系。与前文估计结果类似:工作所需教育水平越高,健康状况越差,过度教育者比具有相同工作但教育职位匹配的工人健康状况差,比相同教育水平但匹配的工人更健康,教育不足员工要比相同工作但教育职位匹配的员工更健康。即过度教育工人用其更多的教育年限补偿了较差身体素质。

表7 人力资本补偿假设检验

解释变量	口头表达能力	理解能力	健康
工作所需年限	0.047 ^{***} (0.007)	0.098 ^{***} (0.007)	-0.045 ^{***} (0.003)
过度教育年限	0.095 ^{***} (0.024)	0.132 ^{***} (0.022)	-0.043 ^{***} (0.008)
教育不足年限	-0.226 ^{***} (0.023)	-0.265 ^{***} (0.020)	0.056 ^{***} (0.007)
R^2	0.19	0.18	0.13

根据以上研究结论,本文发现过度教育工人相比从事相同工作恰好匹配工人具有更好的口头表达能力和理解能力,教育不足工人则呈现相反的结论。过度教育工人可能实际上用他们剩余的教育补偿了其他人力资本方面的缺陷,即教育错配与人力资本之间存在补偿机制。

五、结论及启示

本文基于CFPS微观面板数据研究了教育错配对工资的影响效应及其产生的理论机制。基于固定效应模型以及工具变量法解决遗漏变量偏差与测量误差,进一步检验教育错配的相关经济理论在中国的适用性。本文结论有:(1)2010—2014年间,教育回报呈曲折下降态势:2010年样本期最高,达7.23%,但2012年下降到4.83%,2014年又回升到5.44%。教育回报率的下降暗含着劳动力市场配置效率的下降。(2)控制了测量误差和遗漏变量偏误后,工作需要教育回报率上升到10.61%,过度教育工人将面临1.4%的工资惩罚,教育不足员工将获得2.4%的工资红利。工资不仅依赖于接受的教育水平,更依赖工作需要的教育水平。(3)随着任职生涯的推进,员工通过不断寻找与自己教育技能相匹配的工作,教育错配会得到改善。(4)教育错配与人力资本之间存在补偿机制。在工作中,表达能力、理解能力和健康等各方面的缺失可以通过过度教育而得到补偿。

根据上述研究结论,本文得如下启示:第一,教育主管部门应改革教育体系。过度教育年限并未呈现明显的工资效应,其原因之一是过度教育者相对于相同教育水平恰好匹配者能力偏低,过度教育并非由于高学历毕业生供给大于需求,更主要的原因是高学历毕业生在教育体系中所获得技能的缺失。当前的教育体系导致他们不能获得与其知识结构相匹配的工作。因此,应该改革教育体系,高校应及时有效地调整学科专业结构和人才培养结构,主动顺应社会主义市场经济的需求,根据经济结构、产业结构的变动以及劳动力市场需求的变化,对学科专业结构和人才培养模式进行及时调整,从而提供市场需要的拥有技能的毕业生。第二,企业应该建立人才识别和多层级奖惩机制。过度教育存在一定程度的工资惩罚效应。抛开过度教育者自身能力偏低外,企业可能缺乏有效的人才识别策略。本质上是过度教育者自身的个人劳动生产率水平高于相同职业但教育职位匹配者的平均劳动生产率水平,事实上,从本文的研究结论(过度教育者的工资惩罚低于教育不足者的工资红利)可推知,过度教育者提高了行业的平均劳动生产率水平。因此,企业应该通过建立人才识别机制,识别出过度教育这一部分人才,再通过多层级的奖惩激励制度,充分释放他们的劳动生产率水平,进而提高企业整体的生产率水平。第三,政府

应加强就业指导工作。员工在职业生涯的早期存在着更为严重的教育错配问题。可以通过加强对无工作经验或经验较少的大学生的就业指导,通过分析求职者的受教育情况以及市场的需求情况,帮助求职者精准定位与其教育水平相匹配的职位,减少搜寻成本,缓解过度教育的工资溢价不足问题。

最后,需要指出的是,社会关系网络等对职位获得及工资的影响效应已得到了诸多研究的关注(李宏彬等,2012;孔高文等,2017)。相同教育水平的大学生,拥有的社会关系网络越丰厚,越可能在职业生涯初期获得更高收入的职业,但也越可能弱化其求职及追求知识增加技能的努力程度。如何搜集详实可靠的数据准确测算有效社会网络规模、质量、异质性进而挖掘其对教育错配及工资的长期影响方向和程度是本文进一步的研究方向。总之,我国收入分配不均和教育错配现象并存,因此市场因素以及更重要的制度因素、社会因素应同时纳入分析框架才能更好地理解并解决经济不平等问题。

参考文献:

- [1]范皓皓. 大学生人力资本的过度与不足——基于弥补型过度教育视角的实证分析[J]. 北京大学教育评论, 2012, (10): 100-119.
- [2]罗润东,彭明明. 过度教育及其演变趋势分析——基于CGSS受高等教育职员的调查[J]. 经济社会体制比较, 2010, (5): 173-179.
- [3]武向荣. 教育扩展中的过度教育现象及其收入效应[J]. 北京师范大学学报(社会科学版), 2007, (3): 132-136.
- [4]武向荣,赖德胜. 过度教育发生率及其影响因素——基于北京市数据的分析[J]. 教育发展研究, 2010, (19): 36-41.
- [5]颜敏,王维国. 中国技能偏态性技术变迁的实证检验: 兼论大学教育溢价[J]. 统计研究, 2014, (10): 55-62.
- [6]Bauer T K. Educational mismatch and wages: A panel analysis[J]. Economics of Education Review, 2002, 21(3): 221-229.
- [7]Becker G. Human capital: A theoretical and empirical analysis, with special reference to education[M]. New York, NY: National Bureau of Economic Research, 1964.
- [8]Dolton P J, Silles M A. The effects of over-education on earnings in the graduate labour market[J]. Economics of Education Review, 2008, 27(2): 125-139.
- [9]Duncan G J, Hoffman S D. The incidence and wage effects of overeducation[J]. Economics of Education Review, 1981, 1(1): 75-86.
- [10]Hartog J. Over-education and earnings: Where are we, where should we go?[J]. Economics of Education Review, 2000, 19(2): 131-147.
- [11]Iriando I, Pérez-Amaral T. The effect of educational mismatch on wages in Europe[J]. Journal of Policy Modeling, 2015, 38(2): 304-323.
- [12]Mcguinness S. Overeducation in the labour market[J]. Journal of Economic Surveys, 2006, 20(3): 387-418.
- [13]Tomas K, Michael T. Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974-2000[J]. Labour Economics, 2009, 16(2): 183-193.
- [14]Verdugo R R, Verdugo N T. The impact of surplus schooling on earnings: Some additional findings[J]. Journal of Human Resources, 1989, 24(4): 629-643.
- [15]Verhaest D, Omey E. Overeducation, undereducation and earnings: Further evidence on the importance of ability and measurement error bias[J]. Journal of Labor Research, 2012, 33(1): 76-90.
- [16]Wooldridge J M. Econometric analysis of cross-section and panel data[J]. Mit Press Books, 2001, 1(2): 206-209.

The Penalty Effect of Educational Mismatch on Wages: An Empirical Study Based on Micro Panel Data of China

Yan Min¹, Wang Weiguo²

(1. School of Mathematics, Dongbei University of Finance and Economics, Dalian 116025, China;

2. School of Economics, Dongbei University of Finance and Economics, Dalian 116025, China)

Summary: With the expansion of higher education for many years, the difficulty in graduates' employment and the continuous decrease in the starting salaries have become the typical features of China's labor market in recent years. Under pressure from all facets, many university graduates have to engage in jobs that do not match their education. The mismatch between education and jobs not only implies the inefficient allocation of resources invested in education, but also imposes penalty effect on labor wages. How is the effect of educational mismatch on wages? How to avoid measurement error and the heterogeneity from unobservable variable "ability"? What is the theoretical mechanism of the effect of educational mismatch on wages? Thus, this paper analyzes the effect of educational mismatch on wages on the basis of the models proposed by Duncan & Hoffman(1981), and Verdugo, R.R & Verdugo, NT(1989), by using the panel data of China Family Panel Studies(CFPS) from 2010 to 2014. We use a fixed effects model and instrumental variable estimates to address the two econometric problems faced by the empirical literature: omitted variable bias and measurement error. The fixed-effect instrumental variable estimates in this paper confirm that overeducated workers suffer from a wage penalty of 1.4%. Conversely, under-educated employees receive a bonus of 2.4%. Both omitted variable bias and measurement error are inevitable and important factors. The upward bias which is caused by the omission of unobservable variable "ability" is offset by the downward bias caused by measurement error to a large extent. In addition, this paper further tests the theoretical mechanism of the effect of education mismatch on employment. we find educational requirement for jobs is the vital determinant of wages for employees aged 40 or above, while it has no significant effect on wages for those under the age of 40. In other words, as employees' working years are getting longer and longer, their wages gradually depend on their current productivity and skills, rather than the education they have received when they left the education system. Workers who take voluntary job changes are less prone to over-education. Therefore, as their career progresses, the employees constantly seek jobs that match their educational skills, so that the mismatch has been improved. It means that the occupation liquidity hypothesis holds in China. There is a compensation mechanism between education mismatch and human capital. The over-education can compensate for the lack of expression ability, understanding capability, and health at work. The enlightenment contained in the conclusion of this paper is as follows: firstly, the education authorities should reform the education system, so as to provide the skilled graduates needed by the market; secondly, colleges and universities should adjust the subject structure and personnel training structure actively and effectively to meet the needs of the market; what's more, companies should establish the talent identification and multi-level reward and punishment mechanism so as to identify the over-educated workers and release their labor productivity; finally, the governments should strengthen career guidance and help university graduates to find more matched jobs so as to reduce the time of job seeking and weaken the wage premium problem.

Key words: educational mismatch; wage; penalty effect; panel data

(责任编辑 石头)