

教育的信号效应及其市场价值

——基于统计性歧视和雇主学习的视角

王 骏 赵加良*

内容提要 文章从统计性歧视和雇主学习的理论视角,借助 Altonji & Pierret (2001) 和 Lange (2007) 提出的方法考察了教育的信号效应在中国劳动力市场的发挥情况,并测算了教育信号的市场价值。研究发现,随着工作经验的增长,教育对收入的影响显著减弱,而能力对收入的影响却有所增强。雇主学习的速度虽然很快,但雇主的初始预期生产率误差具有相对持久性。当贴现率为 7.18% 时,教育信号的市场价值即多接受一年教育所带来的终生收入现值的增量中信号效应的贡献的上界约为 16%。该结果对贴现率、雇主学习的速度和潜在工作经验长度的变化非常敏感,但对教育成本的变化并不敏感。

关键词 教育信号 市场价值 统计性歧视 雇主学习

一 引言

受教育扩张政策的影响,随着教育机会均等程度的不断提高,教育规模在 20 世纪 90 年代以来实现了迅速扩张,中国正朝着高中阶段教育普及化和高等教育大众化的方向迅速迈进。教育事业的迅速发展改变了劳动者的受教育结构,高学历劳动者占比迅速上升。根据人力资本理论,劳动者受教育水平的提升增加了其人力资本,提高

* 王骏,北京师范大学经济与工商管理学院、教育部高校辅导员培训和研修基地(北京师范大学),电子邮箱:wj@bnu.edu.cn;赵加良,北京师范大学经济与工商管理学院,电子邮箱:932679606@qq.com。作者感谢北京师范大学经济与工商管理学院王善迈教授、袁连生教授、刘泽云教授和杨娟教授提出的宝贵建议,文责自负。

了其劳动生产率进而改善劳动者的就业状况，最终提高其收入水平（Schultz，1975）。但事实并非如此。随着劳动者平均受教育年限的提高，就业率呈现出明显的下降趋势，而失业率却始终在高位徘徊。高学历劳动者特别是高校毕业生的就业状况尤其堪忧。“史上最难毕业季”、“文凭膨胀（credential inflation）”、“知识失业（educated unemployment）”、“学历贬值（diploma depreciation）”、“上学无用论”等成为全社会的热议话题^①。对学历文凭的热衷追求导致私人教育投资过度，过度教育发生率保持在较高水平。基础教育阶段课外补习如火如荼，教育竞争日益加剧。

为什么劳动者受教育水平的提高没有改善其就业状况？为什么虽然知识失业和学历贬值现象严重但人们对学历文凭依然如此热衷？为什么教育机会的可获得性越来越强但教育竞争却越来越激烈？这些现象和问题都统一地指向一个经久不衰的话题：教育在劳动力市场中的作用。人力资本理论认为，教育能够提升劳动者的人力资本进而使其获得更高的工资收入。但教育信号理论（亦称为甄别理论或筛选理论）认为（Spence，1973；Stiglitz，1975；Wolpin，1977），高学历劳动者的工资收入之所以高于低学历劳动者，只是因为受教育水平发送了关于劳动者的能力信号，而并非完全是因为高学历劳动者拥有较高的人力资本。虽然上述提及的文凭膨胀、知识失业、学历贬值和过度教育等现象并非一定由教育发送能力信号的作用导致，还可能由人口、教育、经济和社会等多重因素造成。但不可否认的是，上述现象的出现与教育在劳动力市场中的作用密切相关。教育在劳动力市场中到底扮演着怎样的角色？在提升劳动者人力资本的同时，教育是否还显著地发挥着发送劳动者能力信号的作用？工资在多大程度上取决于劳动者的受教育水平而非能力？教育信号的市场价值到底有多大？本研究旨在分析和解决上述问题。

本研究从统计性歧视和雇主学习的理论视角探讨教育的信号效应及其市场价值。由于雇主在招聘员工时无法充分了解其真实能力，只能借助某些易于观测的特征对其生产率水平做出理性预期。但随着工作经验的增长，雇主通过逐渐“学习”员工

^① 文凭膨胀是指在一定范围内的劳动力市场中，劳动者的学历普遍得到提升，同时高学历劳动者的实际供给超过了劳动力市场对高学历劳动者的实际需求（范皑皑，2013）。知识失业是指接受过良好教育并具备一定专业知识水平与专业技能的高学历劳动者处于不得其用的状态（王效仿，2003）。学历贬值是指劳动者在社会中无法获得相应的工作岗位和发展需求，只能接受低于学历资格的工作，获取的报酬和发展机遇远远低于学历自身的预期值，来自 <https://baike.baidu.com/item/%E6%96%87%E5%87%AD%E8%B4%AC%E5%80%BC/1458737?fr=aladdin>。

的实际工作绩效了解其真实能力。这种现象被称为统计性歧视和雇主学习 (Employer Learning with Statistical Discrimination, 简称为 EL-SD)。本研究借助 Altonji & Pierret (2001) 和 Lange (2007) 提出的方法, 对教育的信号效应在中国劳动力市场的发挥情况进行了实证研究, 并估计了教育信号的市场价值。研究发现, 对刚进入劳动力市场的员工而言, 工资收入主要取决于受教育水平, 而未显著受到能力的影响。但随着工作经验的增长, 教育对收入的影响显著减弱, 而能力对收入的影响却有所增强。雇主初始预期生产率误差在第 1 期下降约 15%, 到第 6 期下降至初始值的一半以下, 但直到第 40 期仍超过 10%, 表明雇主学习的速度虽然很快, 但初始预期生产率误差具有相对持久性。当贴现率为 7.18% 时, 教育信号的市场价值即多接受一年教育所带来的终生收入现值的增量中信号效应的贡献的上界约为 16%。该结果对贴现率、雇主学习的速度和潜在工作经验长度的变化非常敏感, 但对教育成本的变化并不敏感。

二 理论基础与文献综述

(一) 理论基础

本研究本质上从员工在整个职业生涯期间工资与实际能力的偏离程度, 来考察基于受教育水平的统计性歧视和雇主学习在劳动力市场的表现, 进而研究教育的信号效应及其市场价值。该问题尽管从表面上看只涉及劳动力市场本身, 但由于教育信号形成于教育系统, 故它与教育制度关系同样密切。整体来看, 基于受教育水平的统计性歧视和雇主学习主要包括四个阶段。

第一阶段是个体最优教育决策过程, 即个体依据自身的初始能力、预期教育成本和预期教育收益决定自身是否接受更高阶段的教育。预期教育成本不仅包括学费等显性教育成本, 还包括精神、心理等隐性教育成本。在教育成本与个体能力负相关的假定下, 给定合理的教育成本, 高能力者就会选择接受更高阶段的教育, 而低能力者则不会选择继续接受教育。第二阶段是个体人力资本投资过程, 即个体在学校接受正规教育获得学历文凭的阶段。由于教育具有提升人力资本的作用, 故学校正规教育通过知识传授, 技能培养, 情感、态度和价值观的塑造提升个体能力 (包括认知能力和非认知能力)。借助于合理的评价制度和严格的认证程序, 高能力者 (或达到能力要求者) 获得文凭学历, 而低能力者 (或未达到能力要求者) 不被准许毕业。这两个阶段均发生于教育系统内部, 并对应 Arrow (1973) 提出的“双层过滤装置 (dual filter)”。

其中，第一阶段对应“第一层过滤装置”，即在入学阶段将低能力（受教育前的能力）者过滤出去；第二阶段对应“第二层过滤装置”，即在毕业阶段将低能力（受教育后的能力）者过滤出去。这两个阶段是教育信号形成的重要阶段，教育信号的可靠性直接关系到教育的信号效应在劳动力市场的表现。

第三阶段是统计性歧视（statistical discrimination）（Phelps, 1972; Aigner & Cain, 1977; Lundberg & Startz, 1983），即由于不完全信息下劳动力市场的非对称性，雇主难以观测到求职者的真实能力，只能借助其受教育水平对求职者的生产能力和生产率进行评估和理性预测，从而支付其相应的工资。统计性歧视在雇主雇佣和工资支付决策中发挥作用受到两方面因素的影响。一是教育信号的不可靠性，即以学历文凭体现的受教育水平无法切实反映求职者的能力，这与第二阶段的过滤质量较低密切相关。二是市场信息不完全和劳动力市场不对称的程度越高，求职者拥有的有关其能力的私人信息越多，雇主获得的有关求职者能力的信息越少，统计性歧视程度就会越高。第四阶段是雇主学习（employer learning）^①，即雇主随着员工工作经验的增长，通过对员工实际工作绩效的不断观察和学习逐渐了解其真实的能力，从而调整工资支付决策，以使得员工工资与真实能力和生产率相适应。该阶段与第三阶段均发生于劳动力市场内部，并与第二阶段联系紧密，三者共同构成基于受教育水平的统计性歧视和对称雇主学习的理论框架。从理论上讲，如果基于受教育水平的统计性歧视和雇主学习显著存在，则说明教育具有信号效应。本研究借此考察教育的信号效应并测算教育信号的市场价值。

（二）文献综述

Farber & Gibbons（1996）第一次建立标准计量模型检验基于受教育水平的统计性歧视和对称雇主学习。他们使用美国 1979 - 1992 年 NLSY79（National Longitudinal Survey of Youth 79）数据中的男性样本，将 AFQT（Armed Forces Qualification Test）得分对受教育水平和初次就业起薪进行回归所得的残差代理能力，采用普通最小二乘估计方法证明了美国劳动力市场存在 EL-SD。基于近似相同的样本和数据，Altonji &

^① 雇主学习可分为对称雇主学习（symmetric employer learning）和非对称雇主学习（asymmetric employer learning）。对称雇主学习又被称为公共学习（public learning），指所有的雇主能够拥有相同的关于难以观测的能力信息，即员工的生产率信息是公共信息（Freeman, 1977; Harris & Holmstrom, 1982）；非对称雇主学习又被称为私人学习（private learning），是指当前雇主比外部雇主拥有更多的关于难以观测的能力信息，即员工的生产率信息是私人信息（Waldman, 1984; Greenwald, 1986）。如无特殊说明，本研究所指的“雇主学习”均为对称雇主学习。

Pierret (2001) 直接使用 AFQT 得分、父亲的受教育年限和同胞兄弟姐妹的工资代理能力, 将潜在工作经验作为实际工作经验的工具变量, 采用两阶段最小二乘估计方法得到了类似的结论。作为该研究领域的奠基性文献, Farber & Gibbons (1996) 提出的模型和 Altonji & Pierret (2001) 提出的模型被广泛应用于经验研究中。

Galindo-Rueda (2003) 利用英国 1981 - 2000 年的 NCDS (National Child Development Study) 和 1991 - 2000 年的 BCS (British Cohort Study) 数据, Pan (2005) 使用加拿大 1994 年第二轮 SLID (Survey of Labour and Income Dynamics) 数据中 16 ~ 69 岁的男性样本, Strobl (2004) 使用 1998 年加纳第五轮 RPED (Regional Program for Enterprise Development) 数据均证实了雇主学习假说。但 Cheung (2010) 使用 1995 - 2004 年 LSAY (Longitudinal Surveys of Australian Youth) 数据中 19 ~ 26 岁的男性样本, Bauer & Haisken-DeNew (2001) 使用 1984 - 1996 年德国 GSOEP (German Socio-Economic Panel) 数据中 60 岁以下的男性样本, Zhangaliyeva & Nakabayashi (2013) 使用 1998 - 2006 年俄罗斯 RLMS (Russia Longitudinal Monitoring Survey) 数据均拒绝了雇主学习假说。

一些研究者关注对称雇主学习的国别差异。比如 Hanushek & Zhang (2006) 使用 1994 - 1998 年 IALS (International Adult Literacy Survey) 追踪调查数据, 发现对称雇主学习存在于美国和智利, 可能存在于荷兰、德国、瑞典和丹麦, 而其余国家则不存在。而 Broecke (2015) 使用 2012 年 PIAAC (Programme for the International Assessment of Adult Competencies) 项目调查数据, 则发现对称雇主学习存在于美国、德国、荷兰和瑞典等国家, 可能存在于澳大利亚、丹麦和挪威, 其余国家则不存在。这两项研究都针对 OECD (Organization for Economic Co-operation and Development) 国家, Strobl (2004) 则提供了来自发展中国家的证据。这项研究发现对称雇主学习只发生在通过正式渠道且未经历过在职培训的个体。由此可见, 由于教育制度和劳动力市场制度在各国间差异巨大, 不同国家的对称雇主学习也表现出不同的特征。即便是针对同一个国家的研究, 由于依托的样本和数据不同, 采用的计量模型和估计方法不同, 使用的变量和衡量指标不同, 研究结果也差异很大。

Altonji & Pierret (1998) 认为雇主学习的速度对理解统计性歧视和雇主学习的经济意义十分重要。Lange (2007) 在 Farber & Gibbons (1996) 和 Altonji & Pierret (2001) 的基础上通过建立理论模型和计量模型估计了雇主学习的速度。他将雇主学习的速度定义为初始预期生产率误差方差占当期员工生产率方差的比重, 然后通过估计明瑟方程得到教育、能力与工作经验虚拟变量的估计系数, 再使用非线性最小二乘 (Non-Linear Least Squares, 简称为 NLS) 方法得到雇主学习速度的估计值。Lange (2007) 使

用1979 – 1998 年 NLSY 数据，发现雇主的初始预期生产率误差在第1 期下降26%，3 期过后就下降51%，但在26 期后才降低至初始值的10% 以下，这表明虽然雇主在员工的职业生涯早期就能够在很大程度上了解其真实的生产率，但预期生产率误差具有相对持久性。

Lange (2007) 在估算雇主学习速度的基础上，测算了教育信号的市场价值。他将多接受一年教育所带来的终生收入现值的增量分解为教育的信号效应和生产力效应，并将教育信号效应所占比重定义为教育信号的市场价值。他通过在模型中引入教育成本，通过求解包含预期的终生收入函数和教育成本函数的教育决策模型，得到教育的人力资本价值的下界进而得到教育信号的市场价值的上界。研究发现，教育信号的市场价值低于15%，且对贴现率和雇主学习速度的变化非常敏感。贴现率越高，雇主学习的速度越快，教育信号的市场价值越低。

三 样本和变量

(一) 样本

本研究使用的数据来源于2013 年“中国居民收入项目”(Chinese Household Income Project, 简称为CHIP) 数据和2012 年、2014 年中国家庭追踪调查(Chinese Family Panel Studies, 简称为CFPS) 数据。CHIP 是一项旨在追踪中国收入分配动态情况的全国性调查项目。该项目由北京师范大学中国收入分配研究院实施，由国家统计局(National Bureau of Statistics of China, 简称为NBS) 完成，相继在1989 年、1996 年、2003 年、2008 年和2014 年进行了五次调查。2013 年CHIP 样本来自国家统计局城乡一体化常规住户调查大样本库，后者覆盖全部31 个省份的16 万户居民。CHIP 按照东、中、西部地区分层，采用系统抽样方法抽取得到CHIP 样本，覆盖了来自15 个省份126 个城市234 个县区的18948 个住户样本和64777 个个体样本。CFPS 是一项旨在反映中国人口与社会经济状况变动趋势的社会追踪调查项目。该调查项目由北京大学中国社会科学调查中心实施。CFPS 正式调查从2010 年开始，样本包括5 个大样本省和20 个小样本省，前者包括辽宁、上海、河南、广东和甘肃，后者包括除5 个大样本省外的20 个省份。2012 年和2014 年针对全部样本进行了两次追踪调查，两年的追访率均超过70%。

本研究首先将样本年龄限定在16 ~ 40 岁。这样处理出于两方面的考虑：一是统计性歧视和雇主学习常发生在员工职业生涯前期，样本包括更为年长的员工可能模糊统

计性歧视和雇主学习的证据,从而弱化教育的信号效应。二是 16~40 岁的员工大多在 20 世纪 90 年代以后进入劳动力市场,随着劳动力市场化程度的不断提高和工资制度改革的持续推进,员工的生产率日益成为工资的决定性因素。因此,不同工作经验员工的工资能够互为反事实。而更为年长的员工大多在 20 世纪 90 年代以前就进入劳动力市场,当时的工资并不取决于员工的生产率,那么不同工作经验员工的工资难以互为反事实。

为满足实证研究需要,本研究剔除不在劳动力市场的个体,包括在校学生、离退休人员、丧失劳动能力(残疾)、家务劳动者等,仅保留就业身份为雇员(或受雇,包括有固定雇主的雇员、劳务工或劳务派遣人员、零工或散工)的个体,剔除就业身份为雇主(包括私营业主、个体工商户)、自我经营(或自雇)和家庭帮工的个体,剔除土地承包者(主要指农民),剔除月工资收入小于 10 元或大于 50000 元、小时工资收入小于 0.5 元或大于 200 元、年工资收入小于 120 元或大于 600000 元的个体,剔除性别、年龄、户口、民族、受教育水平等人口学信息缺失,父母受教育水平信息,工资收入信息缺失的个体。经过上述处理,本研究的样本量为 15218 人。

(二) 变量

本研究的核心变量包括教育、能力、潜在工作经验和工资收入。教育由受教育年限表征,由最高学历转化得到。文盲/半文盲、小学、初中、高中、大专、大学本科、硕士、博士学历对应的受教育年数分别为 0、6、9、12、15、16、19 和 22 年。能力由父母的受教育水平代理。这是基于两个方面的考虑:一是丰厚的家庭资本或良好的家庭背景能够在很大程度上促进个体能力的发展,故父母的受教育水平能够在很大程度上代理个体的综合能力;二是数据的可获得性,CFPS 和 CHIP 均提供了受访者父母的受教育水平。为剔除父母受教育水平的代际差异(即时间趋势),本研究具体用父母受教育年限残差的均值表征,该残差分别来源于针对父亲(或母亲)受教育年限和出生年份不缺失的全部样本、对父亲(或母亲)受教育年限关于出生年虚拟变量进行回归的计量模型(Farber & Gibbons, 1996)。潜在工作经验由年龄减去受教育年限再减 6 得到(Altonji & Pierret, 2001; Pinkston, 2006; Mansour, 2012)。这种定义方式将入职后的正规受教育年限排除在外,并将个体未就业和从事非正式工作的时间涵盖其中,但忽视了上学年龄和学制因素对计算潜在工作经验的影响。为剔除工作时间对工资收入的影响,工资收入用小时工资性收入来衡量。月工资性收入除以每月工作的天数再除以每天工作的小时数即可得到小时工资性收入。对于 CHIP 数据,月工资性收入由报告的年工资性收入与全年工作月数相除得到。对于 CFPS 数据,月工资性收入包括工

资、奖金、现金福利和实物福利折合现金。

控制变量包括个体特征和工作特征变量。个体特征变量均为虚拟变量，包括性别、户口性质、民族和居住地类型。对于性别，男性赋值为 1，反之为 0。对于民族，少数民族赋值为 1，反之为 0。对于户口性质，农业类户口赋值为 1，反之为 0。其中，CFPS 中的农业类户口包括农业户口和没有户口两类，非农类户口包括非农户口和外籍两类；CHIP 中的农业类户口包括农业户口和户籍改革前是农业户口的居民户口两类，非农类户口包括非农户口、户籍改革前是非农户口的居民户口和外籍三类。对于居住地类型，城镇赋值为 1，农村赋值为 0。CFPS 使用国家统计局分类标准划分城乡，CHIP 根据问卷类型划分城乡，将城镇住户和外来务工住户视为城镇，将农村住户视为农村。

工作特征变量均为虚拟变量，包括职业培训、行业、职业、工作类型、省份和年份。检验对称雇主学习的重要假定是人力资本投资中性（human capital investment neutrality），即并不存在额外的人力资本投资过程影响教育回报和能力回报随工作经验增长的变动趋势。由于职业培训可能使得员工更具生产性，那么具有特定能力或教育特征的员工往往接受更多的职业培训，从而干扰对统计性歧视和对称雇主学习的检验，故有必要控制个体参与职业培训的情况。CFPS 提供了个体参与职业培训的相关信息，故可以直接生成个体是否参与职业培训虚拟变量。CHIP 并未提供受访者参加职业培训的信息，但 NBS 过录数据中提供了家庭教育支出数据。成人教育支出包含在家庭教育支出中，并主要体现为职业培训。一般来说，如果家庭成人教育支出为零，那么该家庭的所有成员应均未参加过需要个人承担培训成本的培训项目。但如果家庭的成人教育支出不为零，那么该家庭中至少有一个成员参加过培训项目。本研究据此定义个体是否参与职业培训虚拟变量。针对受雇的工资性收入者，行业变量涉及农业、工业和建筑业、传统服务业、现代服务业、公共服务业和其他等六大类。职业变量涉及国家机关和事业单位、专业技术人员、办事人员和有关人员、商业服务业人员、农林牧副渔和水利生产人员、生产运输人员、其他等七大类。单位或工作类型变量涉及党政机关团体、事业单位与社会组织、国有类企业、集体类企业、私营类企业、外资类企业、其他等七大类。

为更准确地报告回归和模拟结果，本研究进行加权处理。CFPS 数据提供了基于全国总抽样样本和全国再抽样样本的横截面权数和面板权重。由于本研究将 CFPS 数据视为混合横截面数据，故使用基于全国总样本的横截面权数。CHIP 数据没有权重，本研究按照 2010 年全国人口普查数据分地区（大都市、东部、中部和西部）、户籍和省份三层计算权重。针对由两个数据库构成的混合样本，本研究赋予三个数据库相同的权

重，并按照数据库和调查年分组对权重进行标准化处理，故混合样本权重范围在 0 ~ 1 之间。

四 雇主学习的速度

本研究使用 Lange (2007) 提出的理论模型测算教育信号的市场价值。在测算教育信号的市场价值前，需要首先定义雇主学习的速度，然后以 Farber & Gibbons (1996) 和 Altonji & Pierret (2001) 模型为基础估计雇主学习的速度。

(一) 理论模型

假定员工在 x 期的生产率的对数 χ 由两部分构成：其一是线性函数 $\tilde{\chi}(s, q, \eta, z)$ ， s 表示雇主和研究者均可观测的变量，如受教育水平，用 q 表示雇主能够观测但研究者无法观测的变量，如员工应聘简历上的信息，用 z 表示雇主无法观测但研究者可以观测的变量，如能力，用 η 表示雇主和研究均无法观测的变量，并假定这四个因素均不随 x 的变化而变化；其二是工作经验的多项式 $\tilde{H}(x)$ 。

$$\chi = \tilde{\chi}(s, q, \eta, z) + \tilde{H}(x) = rs + \alpha_1 q + \lambda z + \eta + \tilde{H}(x) \quad (1)$$

假定 (s, q, η, z) 服从联合正态分布。由于雇主无法观测到 z 和 η ，故雇主根据 s 和 q 形成条件期望 $E(z|s, q)$ 和 $E(\eta|s, q)$ 。在线性假定的条件下：

$$z = E(z|s, q) + v = \gamma_1 q + \gamma_2 s + v \quad (2)$$

$$\eta = E(\eta|s, q) + e = \alpha_2 s + e \quad (3)$$

将式 (2) 和式 (3) 带入式 (1) 中，则：

$$\begin{aligned} \chi &= (r + \lambda\gamma_2 + \alpha_2)s + (\alpha_1 + \lambda\gamma_1)q + (\lambda v + e) + \tilde{H}(x) \\ &= E(\tilde{\chi}|s, q) + (\lambda v + e) + \tilde{H}(x) \end{aligned} \quad (4)$$

根据雇主学习模型，雇主在劳动力市场每一期 τ 均会获得关于员工的生产率信息 y_τ ，则：

$$y_\tau = \tilde{\chi} + \varepsilon_\tau \quad (5)$$

ε_τ 表示雇主对员工的预期生产率误差，它与模型中的其他变量无关，且服从独立同分布 (independently, identically and normally)，方差为 σ_ε^2 。对于每一期 x ，雇主可利用的员工 x 维生产率信息矩阵为 $y^x = (y_0, y_1, \dots, y_{x-1})$ 。因此，雇主对员工预期生产率的后验分布 (posterior distribution) 满足均值为 μ_x 、方差为 $\rho_x = \frac{1}{\sigma_x^2}$ 的正态分布，其中 (μ_x, σ_x) 为：

$$\mu_x = (1 - \theta_x)E(\tilde{\chi} | s, q) + \theta_x \left(\frac{1}{x} \sum_{\tau=0}^{x-1} y_\tau \right) \quad (6)$$

$$\frac{1}{\sigma_x^2} = \frac{1}{\sigma_0^2} + \frac{1}{\sigma_\varepsilon^2} \quad (7)$$

μ_x 为雇主基于员工可观测特征的初始预期生产率和当期之前所有期预期生产率均值的加权和，权重分别为 $1 - \theta_x$ 和 θ_x 。 σ_0^2 为 χ 基于 (s, q) 的条件方差，即初始预期生产率误差的方差，故 ρ_x 为初始预期生产率误差方差与当期预期生产率误差方差的和。回归系数 θ_x 可以被表示成：

$$\theta_x = \frac{xK_1}{1 + (x - 1)K_1} \quad (8)$$

其中：

$$K_1 = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_\varepsilon^2} \quad (9)$$

K_1 即雇主学习的速度，它等于初始预期生产率误差方差占当期员工生产率误差方差的比重，反映了员工可观测特征所包含的初始生产率信息与员工随后 τ 期生产率信息的关系。回归系数 θ_x 和参数 K_1 均位于区间 $[0, 1]$ 。随着 x 的增加， θ_x 趋近于 1，且 θ_x 与 K_1 成正相关。 K_1 越小，说明相对于员工可观测特征给雇主带来的初始信息，随后 τ 期的生产率信息 y_τ 就相对越少，雇主就会为员工当期生产率信息施加越小的权重，即 θ_x 越小，故雇主赋予新信息的权重随着 K_1 的增加而增加。

在对称雇主学习和竞争性劳动力市场的假定下，员工工资 $W(s, q, y^x)$ 等于预期生产率，即：

$$W(s, q, y^x) = E(e^x | s, q, y^x) = e^{E(\tilde{\chi} | s, q, y^x) + \tilde{H}(x) + \frac{1}{2}\sigma_x^2} \quad (10)$$

对式 (10) 进行对数化处理，并令 $\hat{H}(x) = \tilde{H}(x) + \frac{1}{2}\sigma_x^2$ ，得到工资收入对数 $w(s, q, y^x)$ 的表达式：

$$w(s, q, y^x) = (1 - \theta_x)E(\tilde{\chi} | s, q) + \theta_x \left(\frac{1}{x} \sum_{\tau=0}^{x-1} y_\tau \right) + \hat{H}(x) \quad (11)$$

式 (11) 描述了工资收入对数与雇主获得的初始生产率信息 (s, q) ，以及在员工职业生涯各阶段反映的生产率信息 y^x 的关系。但实际调查数据只能提供 (s, z, y^x) ，而非 (s, q, y^x) ，故必须建立工资收入对数与 (s, z, y^x) 关系，而非与 (s, q, y^x) 的关系。假定 (q, η) 和 (s, z) 具有如下形式的线性关系：

$$q = \gamma_3 s + \gamma_4 z + u_1 \quad (12)$$

$$\eta = \gamma_5 s + \gamma_6 z + u_2 \quad (13)$$

则基于可观测数据 (s, z, x) 的工资收入对数可被表示成:

$$E[w(s, q, y^x) | s, z, x] = (1 - \theta_x) E[E(\tilde{\chi} | s, q) | s, z] + \theta_x E(\tilde{\chi} | s, z) + \hat{H}(x) \quad (14)$$

式 (14) 的重要意义在于它将工资收入对数分解为两部分, 其一是依赖 (s, z) 的两类生产率信息的加权平均, 其二是依赖 (s, z) 的工作经验多项式。 $E[E(\tilde{\chi} | s, q) | s, z]$ 和 $E(\tilde{\chi} | s, z)$ 均是受教育水平 s 和能力 z 的线性函数。权重 θ_x 依赖于工作经验 x 和雇主学习的速度 K_1 。当 $x=0$ 时, $\theta_x \rightarrow \infty$, 而当 $x \rightarrow \infty$ 时, $\theta_x \rightarrow 1$ 。故当 $x=0$ 时, (s, z) 的估计系数完全取决于 $E[E(\tilde{\chi} | s, q) | s, z]$; 而当 $x \rightarrow \infty$ 时, (s, z) 的估计系数完全取决于 $E(\tilde{\chi} | s, z)$ 。因此, 随着工作经验的增长, (s, z) 的估计系数将从由 $E[E(\tilde{\chi} | s, q) | s, z]$ 决定的某个初始值收敛到由 $E(\tilde{\chi} | s, z)$ 决定的某个极限值。这个收敛过程的速度就是本研究要估计的雇主学习的速度。利用 $x=0$ 和 $x=\infty$ 两种情况下的估计系数, 可以将式 (14) 改写成如下形式:

$$E[w(s, q, y^x) | s, z, x] = [(1 - \theta_x) b_{s,0} + \theta_x b_{s,\infty}] + [(1 - \theta_x) b_{z,0} + \theta_x b_{z,\infty}] + \hat{H}(x) \quad (15)$$

其中, $b_{s,0}$ 和 $b_{z,0}$ 分别表示 $\theta_x=0$ 时教育和能力对工资收入对数的影响, $b_{s,\infty}$ 和 $b_{z,\infty}$ 分别为在 $\theta_x=1$ 时教育和能力对工资收入对数的影响。由于权重 θ_x 只是 K_1 和 x 的函数, 故式 (15) 只包括 $\{K_1, b_{s,0}, b_{s,\infty}, b_{z,0}, b_{z,\infty}\}$ 五个参数。如果能够获得 (s, z) 和员工在职业生涯各阶段 (从零到工作经验最大值 T) 的工资收入水平, 我们就能够通过教育和能力对工资收入对数回归估计这五个参数。

首先, 通过估计式 (16) 得到教育和能力在每个工作经验取值处对工资收入对数的影响的 $2 \times (T+1)$ 维估计系数矩阵 $\{\hat{\beta}_{s,x}, \hat{\beta}_{z,x}\}_{x=0}^T$:

$$w_x = \sum_x \beta_{s,x} (sD_x) + \sum_x \beta_{z,x} (zD_x) + M\beta_M + \varepsilon_x \quad (16)$$

其中, D_x 表示每个工作经验取值对应的虚拟变量, sD_x 和 zD_x 分别表示这些虚拟变量与 s, z 的交互项, M 表示其他控制变量, 包括性别、民族、户口、居住地类别、行业、职业、工作类型、省份和年份虚拟变量。

然后, 将 $\{\hat{\beta}_{s,x}, \hat{\beta}_{z,x}\}_{x=0}^T$ 表示成结构化参数 $\{K_1, b_{s,0}, b_{s,\infty}, b_{z,0}, b_{z,\infty}\}$ 的函数, 即:

$$\{\hat{\beta}_{s,x}, \hat{\beta}_{z,x}\}_{x=0}^T = \{(1 - \theta_x) b_{s,0} + \theta_x b_{s,\infty}, (1 - \theta_x) b_{z,0} + \theta_x b_{z,\infty}\}_{x=0}^T \quad (17)$$

为估计参数 $\{K_1, b_{s,0}, b_{s,\infty}, b_{z,0}, b_{z,\infty}\}$, 我们将每一个估计系数 $\{\hat{\beta}_{s,x}, \hat{\beta}_{z,x}\}_{x=0}^T$ 作为一个观测值。对 $\{\hat{\beta}_{s,x}\}_{x=0}^T$ 而言, 非线性函数 $b_s(x; K_1, b_{s,0}, b_{s,\infty}) = (1 - \theta_x) b_{s,0} + \theta_x b_{s,\infty}$ 是包含三个参数 $(K_1, b_{s,0}, b_{s,\infty})$ 函数, 估计这三个参数最自然而直接的方式是, 最小化 $\{\hat{\beta}_{s,x}\}_{x=0}^T$ 与 $b_s(x; K_1, b_{s,0}, b_{s,\infty})$ 之间的距离来得到最优的 $(K_1, b_{s,0},$

$b_{s,\infty}$)。 $\{\hat{\beta}_{z,x}\}_{x=0}^T$ 情况类似。非线性最小二乘方法能够帮助我们实现上述过程从而得到最优的 $\{K_1, b_{s,0}, b_{s,\infty}, b_{z,0}, b_{z,\infty}\}$ 估计量。

使用上述方法会得到 K_1 的两个估计量，分别来自于教育的估计系数和能力的估计系数，故 K_1 是被过度识别的。可行的解决策略有两个：一是检验这两个 K_1 估计值是否存在显著差异，如果不存在显著差异，说明教育和能力的估计系数在职业生涯期间的变化反映的是统一学习过程，故可以同时接受这两个估计值。二是如果存在显著差异，则需要同时使用教育和能力的估计系数来计算 K_1 。本研究同时最小化 $\{\hat{\beta}_{s,x}, \hat{\beta}_{z,x}\}_{x=0}^T$ 在预测值 $b_s(x; K_1, b_{s,0}, b_{s,\infty})$ 和 $b_z(x; K_1, b_{z,0}, b_{z,\infty})$ 附近的误差方差，即最小化目标函数：

$$\sum_{x=0}^{x=T} w_x \{ [\hat{\beta}_{s,x} - b_s(x; K_1, b_{s,0}, b_{s,\infty})]^2 + [\hat{\beta}_{z,x} - b_z(x; K_1, b_{z,0}, b_{z,\infty})]^2 \} \quad (18)$$

其中，权重 w_x 为每个工作经验取值下观测值所占比重的倒数。

(二) 估计结果

1. 各潜在工作经验取值下的教育回报和能力回报

各潜在工作经验取值下的教育回报和能力回报如图 1 所示^①。图中的实点表示各潜在工作经验取值下的教育（或能力）回报，曲线表示教育（或能力）回报关于潜在工作经验的非线性回归拟合线。可以看出，教育回报随着潜在工作经验的增长呈现出明显下降趋势，且教育回报下降的速度随潜在工作经验的增长而减小；能力回报随着潜在工作经验的增长呈现出明显上升趋势，且能力回报上升的速度随潜在工作经验的增长而减小。由此可见，各潜在工作经验取值下教育回报和能力回报的变动趋势与雇主学习假说一致。

2. 雇主学习的速度

本研究采用 NLS 方法对式 (17) 进行估计，以得到 $\{K_1, b_{s,0}, b_{s,\infty}, b_{z,0}, b_{z,\infty}\}$ 的一致估计量。估计结果如表 1 所示。第 (1) 列和第 (2) 列分别报告了仅根据各潜在工作经验取值下的教育回报和能力回报估计的 K_1 值。运用 Clogg et al. (1995) 提出的方法对这两个 K_1 估计值是否存在显著性差异的检验结果显示， z 值为 14.08，在 1% 的水平上拒绝两个 K_1 估计值不存在显著差异的原假设，说明教育回报和能力回报在职业生涯期间的变动趋势反映的雇主学习过程存在显著差异，需要同时使用各潜在工作经验取值下的教育回报和能力回报联合估计 K_1 ，估计结果见第 (3) 列。可以看出， $b_{s,0}$ 、 $b_{s,\infty}$ 、 $b_{z,0}$ 、 $b_{z,\infty}$ 和 K_1 的

① 所有估计系数均来源于对式 (15) 的估计结果。由于潜在工作经验超过 26 年的样本仅占全部样本的 2.69%，故将这些样本的潜在工作经验一律赋值为 26。

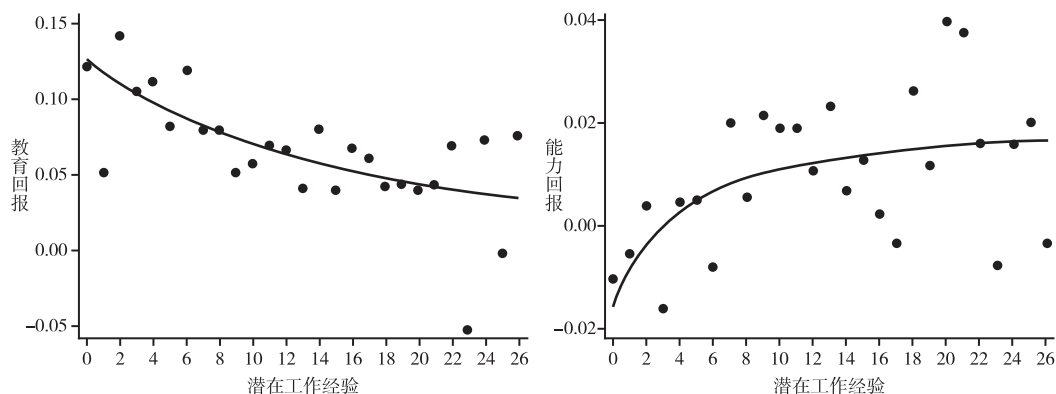


图1 各潜在工作经验取值下的教育回报和能力回报的变动趋势

注：回归的被解释变量为小时工资对数，核心解释变量为教育、能力、教育与潜在工作经验的交互项、能力与潜在工作经验的交互项，控制变量包括性别、户口性质、民族、居住地类型、参加过职业培训、行业、职业、工作类型、省份和年份虚拟变量。回归使用标准化权重进行加权。

资料来源：根据2013年CHIP数据、2012年和2014年CFPS数据计算得到。

估计值分别为0.11、-0.03、0.01、0.11和0.15，表明员工在职业生涯期间内，教育回报从11.46%下降至2.51%，而能力回报从0.85%上升至10.86%，雇主学习速度 K_1 为0.15。

表1 非线性最小二乘回归估计结果

	(1)	(2)	(3)
K_1	0.05 *** (0.00)	0.20 *** (0.01)	0.15
$b_{s,0}$	0.13 *** (0.00)		0.11
$b_{s,\infty}$	-0.03 *** (0.00)		-0.03
$b_{z,0}$		-0.02 *** (0.00)	0.01
$b_{z,\infty}$		0.02 *** (0.00)	0.11
调整的 R ²	0.89	0.54	
F 值	38793.68	5368.46	
差异性检验	z = 14.08 (p < 0.01)		

注：第(1)列的被解释变量为各潜在工作经验取值下的教育回报，第(2)列的被解释变量为不同潜在工作经验取值下的能力回报，第(3)列的被解释变量同时包括各潜在工作经验取值下的教育回报和能力回报；括号内为标准误，通过Bootstrap方法重复操作5000次获得；* p < 0.1, ** p < 0.05, *** p < 0.01；回归使用各潜在工作经验取值下样本量进行加权。

资料来源：根据2013年CHIP数据、2012年和2014年CFPS数据计算得到。

根据式(8)提供的 K_1 和 θ_x 的关系,本研究计算了在不同 K_1 值下 θ_x 随潜在工作经验的变动趋势,如图2所示。可以看出,随着潜在工作经验的增长,雇主在 x 期预期员工生产率时对当期之前所有期预期生产率均值施加的权重 θ_x 逐渐收敛到1。这是因为随着雇主对员工生产率的学习,雇主会给予初始预期生产率越来越小的权重,而给予员工当前生产率信息越来越大的权重。 θ_x 与雇主学习的速度 K_1 呈现出明显的正相关。这是因为,雇主学习的速度越小,说明相对于员工可观测特征给雇主带来的初始信息,随后各期的生产率信息就会相对越少,雇主就会为员工当期生产率信息施加越小的权重。特别地,当 $K_1 = 0.15$ 时,初始预期生产率误差在第1期下降约15%,到第3期下降约1/3,到第6期下降超过50%,到第24期下降幅度超过80%,但直到第40期,初始预期生产率误差占初始值的比重依然超过10%。

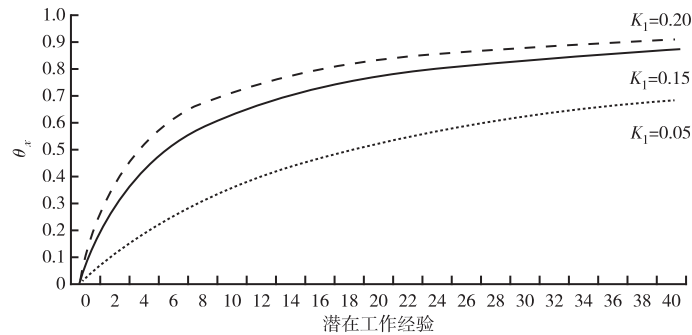


图2 随潜在工作经验增长的变动趋势

资料来源:根据2013年CHIP数据、2012年和2014年CFPS数据计算得到。

五 教育信号的市场价值

(一) 理论模型

本研究将教育信号的市场价值定义为:多接受一年教育所带来的终生收入现值的增量中教育信号的贡献。这里的终生收入现值是指个体从首次进入劳动力市场年份到其退休年份 T 期间的工资收入现值。设 i 为劳动力市场中劳动收入的贴现率,则具有特征 (s, q, η, z) 的个体的预期终生收入现值为:

$$\sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) | s, q, \eta, z] = \sum_{\tau=0}^T e^{-i\tau} E[e^{E(\tilde{\chi} | s, q, y^\tau) + H(\tau)} | s, q, \eta, z] \quad (19)$$

对式 (19) 关于 s 求一阶导数, 得到多接受一年教育所带来的预期终生收入现值的增量, 即:

$$\sum_{\tau=0}^T e^{-i\tau} \frac{\partial E[W(s, q, y^\tau) | s, q, \eta, z]}{\partial s} = \sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) | s, q, \eta, z] [(1 - \theta_\tau)(\lambda\gamma_2 + \alpha_2) + r] \quad (20)$$

由式 (20) 可以看出, 教育对预期工资收入对数的影响主要体现在两个方面: 一是直接的生产率效应, 由 r 表示; 二是信号效应, 由 $(\lambda\gamma_2 + \alpha_2)$ 表示。由于 $\lambda\gamma_2$ 反映的是由于 s 与 z 相关而导致的预期工资收入对数随 s 的增加而增长, α_2 反映的是由于 s 与 η 相关而导致的预期工资收入对数随 s 的增加而增长, 故 $(\lambda\gamma_2 + \alpha_2)$ 反映的是由统计性歧视造成的 s 与 (η, z) 的相关关系。权重 $(1 - \theta_x)$ 表明, 信号效应对预期工资收入增长的贡献随着工作经验的增长而减小。对受教育水平相同的个体进行平均化处理后, 本研究将教育信号的市场价值 Γ 定义为:

$$\Gamma = \frac{\sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) | s] [(1 - \theta_\tau)(\lambda\gamma_2 + \alpha_2)]}{\sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) | s] [(1 - \theta_\tau)(\lambda\gamma_2 + \alpha_2) + r]} \quad (21)$$

它度量了多接受一年教育所带来的终生收入现值的增量中教育信号的贡献。由于我们能够估计出不同教育分组的收入生命周期曲线, 故 $E[W(s, q, y^\tau) | s]$ 能够被观测到。因此, 在给定贴现率 i 的情况下, 求解只需要识别 $(\lambda\gamma_2 + \alpha_2, r)$ 。

Lange (2007) 指出, 由于信号模型的核心就是教育与其他不可观测的反映员工生产率的因素呈正相关, 故如果要估计教育信号的市场价值, 就不能限制不可观测因素与教育的相关性。因此, 无法通过对模型中的参数施加某些限制来识别 $(\lambda\gamma_2 + \alpha_2, r)$ 。虽然基于雇主学习模型无法识别参数 $(\lambda\gamma_2 + \alpha_2, r)$, 但由于最优教育决策要求教育成本等于教育的生产率效应和信号效应之和, 通过在模型中引入教育成本, 就能够有效识别参数, 从而估算教育信号的市场价值。

识别教育信号回报的上界要求 $\frac{\partial E[w(s, q, 0) | s]}{\partial s} \geq \frac{\partial w(s, q, 0)}{\partial s}$ 。在该假定下, 利用调查数据估计的教育回报应至少不小于实际的教育回报。由于雇主在为员工制定工资时能够利用研究者不可观测的但能够反映员工生产率的因素 q , 故相比于研究者, 雇主对员工受教育水平的依赖度更高。因此, $\frac{\partial E[w(s, q, 0) | s]}{\partial s} \geq \frac{\partial w(s, q, 0)}{\partial s}$

的一个充分条件是，在平均水平上，教育与这些研究者不可观测的但能够反映员工生产率的因素（或生产率本身）呈正相关，即 $\frac{\partial E [q|s]}{\partial s} \geq 0$ 且 $\frac{\partial E [z|q, s]}{\partial s} = \gamma_1 \geq 0$ 。其经济直觉是，信号模型反映的就是教育与雇主不可观测的反映员工生产率的因素 η 的正相关性，而实际数据揭示的是教育与研究者观测到的反映员工生产率的因素 z 的正相关性，故与之类似，该假定暗含的是教育与雇主可观测的反映员工生产率的因素 q 的正相关性。因此，雇主在预测员工生产率时就会较少依赖员工的受教育水平。

最优教育决策要求额外接受一年教育的收益必须等于额外接受一年教育的成本，故利用式（20）得：

$$\sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) | s, q, \eta, z] [(1 - \theta_\tau)(\lambda\gamma_2 + \alpha_2) + r] = \Lambda(s) \quad (22)$$

由式（20）可得， $\frac{\partial w(s, q, 0)}{\partial s} = (\lambda\gamma_2 + \alpha_2) + r$ 。对受教育水平相同的个体进行标准化处理后， $\frac{\partial E [w(s, q, 0) | s]}{\partial s} \geq \frac{\partial w(s, q, 0)}{\partial s}$ 意味着：

$$\sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) | s] \{ (1 - \theta_\tau) \left[\frac{\partial E [w(s, q, 0) | s]}{\partial s} - r \right] + r \} \geq \Lambda(s) \quad (23)$$

式（23）的左边唯一不可观测的参数是教育的生产率效应参数 r 。由于式（23）的左边是关于 r 的严格增函数，故一定存在唯一的 r 满足式（23）的取等条件。由于 r 表示的是教育的生产率效应的下界，故 $\frac{\partial E [w(s, q, 0) | s]}{\partial s} - r$ 表示的是当 $x=0$ 时教育的信号效应的上界。

假设个体最大化终生工资收入的现值，则目标函数为：

$$e^{-is} \sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) | s, q, \eta, z] - \Pi(s) \quad (24)$$

$\Pi(s)$ 表示教育的学费成本。对受教育水平相同的个体而言，教育的平均边际成本为：

$$\Lambda(s; i) = i \sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) | s] + \psi'(s) \quad (25)$$

$\Lambda(s; i)$ 的估计量可以由可观测的工资收入生命周期曲线和学费成本得到。将式（23）取等并将式（25）代入得到如下等式：

$$\sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) | s] \left\{ (1 - \theta_\tau) \left[\frac{\partial E [w(s, q, 0) | s]}{\partial s} - r \right] + r \right\} = \Lambda(s; i) \quad (26)$$

求 r (26) 即可得到教育的生产率效应的下界:

$$r = \frac{\psi'(s) + \sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) | s] \left\{ i - (1 - \theta_\tau) \frac{\partial E[w(s, q, 0) | s]}{\partial s} \right\}}{\sum_{\tau=0}^T e^{-i\tau} \theta_\tau E[W(s, q, y^\tau) | s]} \quad (27)$$

将 r 代入式 (21) 即可得到 Γ 。

(二) 参数校准

T 表示潜在工作经验, 即从首次进入劳动力市场年份到其退休年份的时间跨度。本研究将 T 的基准值设定为 40。

序列 $\{\theta_\tau\}_{\tau=0}^T$ 取值与雇主学习的速度 K_1 有关。本研究将 $\{\theta_\tau\}_{\tau=0}^T$ 的基准值设定为 $K_1 = 0.15$ 时的值。

$E[W(s, q, y^\tau) | s]$ 表示工作经验—工资收入曲线, 即个体在职业生涯期间的年平均工资收入。对于潜在工作经验为 0~26 年的个体, 本研究使用 16~40 岁的样本计算各潜在工作经验取值下个体的年平均工资收入, 而对于潜在工作经验超过 26 年的个体, 本研究将这些样本的年平均工资收入赋值为常数, 即潜在工作经验为 26 年的个体的年平均工资收入。由混合横截面数据得到的各工作经验取值下的个体的平均工资收入未考虑工资收入的增长趋势。根据国家统计局公布的 2013—2016 年居民人均可支配收入工资性收入, 本研究将工资水平的年均增长速度设定为 8%, 得到各工作经验取值下的个体的平均名义工资收入。

$\frac{\partial E[w(s, q, 0) | s]}{\partial s}$ 表示员工刚进入劳动力市场时的教育回报, 即 $\frac{\partial w_s(s, 0)}{\partial s}$ 的估计值, 同时也是教育的生产率效应和教育的信号效应之和。本研究通过估计以工资收入对数为因变量, 以受教育年限、工作经验、潜在工作经验平方项、受教育年限与潜在工作经验的交互项、受教育年限与潜在工作经验平方项的交互项、性别、户口性质、民族、居住地类型、省份、年份虚拟变量为自变量的回归模型, 得到受教育年限的估计系数为 9.67%。

$\psi'(s)$ 表示接受教育的年均学费成本。本研究使用个体就读最高学历阶段 (小学及其以上学历) 的最后一年的学费成本数据, 并使用调查年份与潜在工作经验的差为每个个体确定学费成本数据的年份。年均学费成本用生均学费衡量。其中, 学费数据来源于《中国教育经费统计年鉴》, 在校生数数据来源于《中国统计年鉴》。本研究以 2000 年为基期对学费数据进行调整以剔除价格水平的影响, 并采用标准化权重计算得

到平均学费成本为 1702.48 元。

i 表示贴现率。由于 r 是随 i 的变化而变化的，故需要为估计 r 确定合适的贴现率水平。为保证估计结果的稳健性，本研究为参数设定的取值范围至少包括如下贴现率水平：

第一，Lange (2007) 将市场利率作为贴现率的下界。市场利率通常用无风险债券利率衡量。无风险债券利率常用银行一年期定期存款利率代理。中国人民银行一年期定期存款利率（零存整取或整存零取）在 1990 - 2000 年间的变化范围为 1.98% ~ 9.00%，在 2001 - 2015 年间的变化范围为 1.71% ~ 3.33%。

第二，代际内贴现率。张颖等 (2008) 认为，个体跨期选择行为的时间背景既可以在生命周期内，也可以超过生命周期。代际内贴现率反映的是个人在其生命周期内进行跨期选择时所表现出来的时间偏好，偏重于个体的选择问题。代际间贴现率是代际间时间偏好的直观表示，反映当代对未来代的关系程度，它反映的是公众对社会的责任感和前瞻性，偏重于影响期长达数代的政府政策或社会决策评估。该研究采用实验调查法，对 400 位 18 ~ 76 岁来自制造业、金融业、IT 业和事业单位的学生、在职人员和离退休人员进行了问卷实验调查，发现代际内贴现率呈现出指数贴现形式。时间跨度为 3 个月、6 个月、1 年、10 年和 20 年的代际内贴现率中值分别为 30.00%、18.50%、30.00%、22.00% 和 13.00%。本研究将时间跨度设定为 40 年，并将贴现率与时间跨度的关系设定为如下形式：

$$i = \beta_0 e^{\beta_1 T + \varepsilon} \quad (28)$$

为估计出 β_1 ，本研究将式 (28) 两边取对数得：

$$\ln i = \ln \beta_0 + \beta_1 T + \varepsilon \quad (29)$$

估计结果表明， β_1 的估计值为 -0.03，标准误为 0.01，该估计系数在 10% 的水平上统计显著； $\ln \beta_0$ 的估计值为 -1.32，标准误为 0.16，该估计系数在 1% 的水平上统计显著。因此，当 $T=40$ 时的预测值为 7.18%。

第三，Lange (2007) 将潜在工作经验为 0 年即刚进入劳动力市场时的教育回报率作为贴现率的下界。在本研究中，贴现率的上界应为 9.67%。

综合以上分析，本研究将报告不同 $i=2.00\%$ 、 3.00% 、 4.00% 、 5.00% 、 6.00% 、 7.00% 、 7.18% 、 8.00% 、 9.00% 和 9.67% 时 r 的估计结果。

(三) 模拟结果

表 2 报告了在不同贴现率水平下终生收入现值、教育的生产率效应的下界和教育信号的市场价值的上界的估计结果。可以看出，当贴现率取基准值 7.18% 时，教育的

生产率效应的下界为 6.28%，教育信号的市场价值的上界为 15.91%。但模拟结果对贴现率变化的反映非常敏感。从数理看，贴现率对教育的生产率效应的下界的偏效应为：

$$\frac{\partial r}{\partial i} = \frac{\tau\psi'(s) + \sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) | s]}{\sum_{\tau=0}^T e^{-i\tau} \theta_\tau E[W(s, q, y^\tau) | s]} \quad (30)$$

故 $\partial r / \partial i > 0$ 。从经济直觉看，贴现率越高，表明个体越关注当期收入而越不关注未来收入，个体就越不可能花费时间和精力投资教育以获得教育信号的收益。因此，随着贴现率的上升，教育的生产率效应呈现出迅速上升的趋势，而教育信号的市场价值呈现出迅速下降的趋势。

表 2 教育信号的市场价值

贴现率 i (%)	终生收入现值 $\sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) s]$ (万元)	教育的生产率效应的下界 \underline{r} (%)	教育信号的市场价值的上界 \bar{T} (%)
2.00	356.77	<0.00	>100.00
3.00	278.73	0.94	73.55
4.00	220.44	2.14	52.26
5.00	176.55	3.38	37.52
6.00	143.22	4.67	26.38
7.00	117.66	6.03	17.37
7.18	113.72	6.28	15.91
8.00	97.88	7.46	9.69
9.00	82.42	8.97	2.86
9.67	73.93	10.03	<0.00

注：假定 $T=40$ ， $K_1=0.15$ ， $\frac{\partial E[w(s, q, 0) | s]}{\partial s} = 9.67\%$ ， $\psi'(s) = 1702.48$ 。

资料来源：根据 2013 年 CHIP 数据、2012 年和 2014 年 CFPS 数据计算得到。

具体而言，当贴现率低于 3.00% 时，教育的生产率效应的下界为负，教育信号的市场价值的上界超过 100%，故教育的功能几乎完全体现为信号效应。当贴现率为

3.00%时，教育的生产率效应的下界虽然由负转正，但依然不足为1.00%，教育信号的市场价值的上界却超过70%。由此可见，当贴现率处于较低水平时，教育的生产率效应很低，但教育信号的市场价值很高。当贴现率为5.00%时，教育的生产率效应的下界为3.38%，教育信号的市场价值的上界下降至50%以下。当贴现率为9.67%（即潜在工作经验为0时的教育回报率）时，教育的生产率效应的下界已经超过10%，教育信号的市场价值的上界已经下降至0以下，故教育的功能几乎完全体现为生产率效应。由此可见，当贴现率处于较高水平时，教育的生产率效应很高，但教育信号的市场价值很低。

（四）进一步讨论

如上述分析，除贴现率*i*外，教育的生产率效应的下界和教育信号的市场价值的上界依赖于雇主学习的速度 K_1 、学费成本 $\psi'(s)$ 和潜在工作经验长度*T*的取值。因此，在上述模拟结果的基础上，本研究对模拟结果进行敏感性分析。

1. 雇主学习速度的变化

由于 θ_τ 是雇主学习的速度 K_1 的单调增函数，故探讨雇主学习速度的变化对教育的生产率效应的下界的影响只需要考察 θ_τ 的变化对模拟结果的影响。从数理看，雇主学习的速度对教育的生产率效应的下界的偏效应为：

$$\frac{\partial \underline{r}}{\partial \theta_\tau} = \frac{\left\{ \frac{\partial E[W(s, q, 0) | s]}{\partial s} - i \right\} \left\{ \sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) | s] \right\}^2}{\left\{ \sum_{\tau=0}^T e^{-i\tau} \theta_\tau E[W(s, q, y^\tau) | s] \right\}^2} \quad (31)$$

故 $\partial \underline{r} / \partial \theta_\tau > 0$ 。从经济直觉看，雇主学习的速度越快，雇主对员工真实生产率的了解程度就越高，教育的生产率效应就会越强，教育信号的市场价值就会越低。因此， \underline{r} 随着 K_1 的变大而变大，但 \bar{T} 随着 K_1 的变大而变小^①。

表3报告了当贴现率为7.18%时，在不同雇主学习速度取值下教育的生产率效应的下界和教育信号的市场价值的上界的估计结果。可以看出，模拟结果对雇主学习速度变化的反应是非常敏感的。当单独使用各潜在工作经验取值下的教育回报估计雇主学习的速度即 $K_1 = 0.05$ 时，教育的生产率效应的下界仅为4.59%，教育信号的市场价值的上界却高达39.43%。当单独使用各潜在工作经验取值下的能力回报估计雇主学习

① Habermalz (2011) 指出，之所以雇主学习的速度与教育信号的市场价值呈负相关，是因为两个基本假定：一是雇主学习是对称的，即当前雇主和外部雇主拥有相同的信息；二是雇主学习是完全的，即雇主能够获得关于员工生产率的全部信息。

的速度即 $K_1 = 0.20$ 时，教育的生产率效应的下界仅为 6.58%，教育信号的市场价值的上界却高达 11.57%。结合 $K_1 = 0.15$ 的模拟结果，教育的生产率效应的下界的取值范围为 [4.59%，6.58%]，教育信号的市场价值的上界的取值范围为 [11.57%，39.43%]。

表 3 教育信号的市场价值：雇主学习速度的变化

雇主学习的速度 K_1	教育的生产率效应的下界 $\underline{r}(\%)$	教育信号的市场价值的上界 $\bar{T}(\%)$
0.05	4.59	39.43
0.15	6.28	15.91
0.20	6.58	11.57
0.25	6.74	9.22

注：假定 $T=40$ ， $i=7.18\%$ ， $\frac{\partial E[w(s, q, 0) | s]}{\partial s} = 9.67\%$ ， $\psi'(s) = 1702.48$ 。

资料来源：根据 2013 年 CHIP 数据、2012 年和 2014 年 CFPS 数据计算得到。

2. 学费成本的变化

从数理看，学费成本对教育的生产率效应的下界的偏效应为：

$$\frac{\partial \underline{r}}{\partial \psi'(s)} = \frac{1}{\sum_{\tau=0}^T e^{-i\tau} \theta_{\tau} E[W(s, q, y^{\tau}) | s]} \quad (32)$$

故 $\partial \underline{r} / \partial \psi'(s) > 0$ 。从经济直觉看，随着信号成本的增长，教育信号带来的收益就越不足以抵消所支付的信号成本，个体也就越不可能投资教育信号。因此，随着学费成本的提高，教育的生产率效应就会越强，教育信号的市场价值就会越低。因此， \underline{r} 随着 $\psi'(s)$ 的增加而变大，但 \bar{T} 随着 $\psi'(s)$ 的增加而变小。

表 4 报告了当贴现率为 7.18% 时，在学费成本取值下教育的生产率效应的下界和教育信号的市场价值的上界的估计结果。可以看出，模拟结果对学费成本变化的反应并不十分敏感。当学费成本从 500 元增长至 2500 元时，教育的生产率效应的下界从 6.13% 增长至 6.38%，教育信号的市场价值的上界从 16.85% 下降至 15.29%。表 4 还特别报告了学费成本分别为 834.15 元（以 1990 年为基期的平均学费成本）和 2066.34 元（以 2010 年为基期的平均学费成本）的模拟结果。结合 $\psi'(s) = 1702.48$ 的模拟结果，教育的生产率效应的下界的取值范围为 [6.17%，6.33%]，教育信号的市场价值的上界的取值范围为 [15.63%，16.58%]。

表 4 教育信号的市场价值：学费成本的变化

学费成本 $\psi'(s)$ (元)	教育的生产率效应的下界 r (%)	教育信号的市场价值的上界 \bar{T} (%)
500.00	6.13	16.85
834.15	6.17	16.58
1000.00	6.19	16.45
1500.00	6.26	16.06
1702.48	6.28	15.91
2000.00	6.32	15.68
2066.34	6.33	15.63
2500.00	6.38	15.29

注：假定 $T=40$, $i=7.18\%$, $K_1=0.15$, $\frac{\partial E[w(s, q, 0) | s]}{\partial s} = 9.67\%$ 。

资料来源：根据 2013 年 CHIP 数据、2012 年和 2014 年 CFPS 数据计算得到。

3. 潜在工作经验长度的变化

考虑到潜在工作经验长度对模拟结果的影响，本研究将原本设定的潜在工作经验长度的基准值从 40 年分别缩短至 35 年和延长至 45 年，与之对应的贴现率分别变为 8.51% 和 6.13%。模拟结果如表 5 所示。可以看出，潜在工作经验长度从 35 年增长至 45 年，终生收入现值从 82.67 万元增长至 157.83 万元。教育的生产率效应的下界从 8.20% 下降至 4.96%，教育信号的市场价值的上界从 5.89% 上升至 20.52%。由此可见，模拟结果对潜在工作经验长度的变化也比较敏感。

表 5 教育信号的市场价值：潜在工作经验长度的变化

潜在工作经验 T (年)	贴现率 i (%)	终生收入现值 $\sum_{\tau=0}^T e^{-i\tau} E[W(s, q, y^\tau) s]$ (万元)	教育的生产率 效应的下界 r (%)	教育信号的 市场上界的上界 \bar{T} (%)
35	8.51	82.67	8.20	5.89
40	7.18	113.72	6.28	15.91
45	6.13	157.83	4.96	20.52

注：假定 $K_1=0.15$, $\frac{\partial E[w(s, q, 0) | s]}{\partial s} = 9.67\%$, $\psi'(s) = 1702.48$ 。

资料来源：根据 2013 年 CHIP 数据、2012 年和 2014 年 CFPS 数据计算得到。

六 结论与政策建议

本研究从统计性歧视和雇主学习的理论视角,借助 Altonji & Pierret (2001) 和 Lange (2007) 提出的方法,考察了教育的信号效应在中国劳动力市场的发挥情况,并测算了教育信号的市场价值。

首先,本研究用初始预期生产率误差方差占当期员工生产率误差方差的比重衡量雇主学习的速度,并采用非线性最小二乘法,利用各潜在工作经验取值下的教育回报和能力回报估计了雇主学习的速度。研究发现,对刚进入劳动力市场的员工而言,工资收入主要取决于受教育水平,而未显著受到能力的影响。但随着工作经验的增长,教育对收入的影响显著减弱,而能力对收入的影响却有所增强。雇主初始预期生产率误差在第1期下降约15%,到第6期下降至初始值的一半以下,但直到第40期仍超过10%,表明雇主学习的速度虽然很快,但初始预期生产率误差具有相对持久性。

其次,本研究将教育信号的市场价值定义为多接受一年教育所带来的终生收入现值的增量中教育信号的贡献。本研究通过引入教育成本,借助个体最优教育决策模型识别教育的生产率效应的下界,进而测算教育信号的市场价值的上界。模拟结果显示,当贴现率为基准值7.18%时,教育的生产率效应的下界为6.28%,教育信号的市场价值的上界为15.91%。模拟结果对贴现率、雇主学习的速度和潜在工作经验长度的变化非常敏感,但对学费成本的变化并不十分敏感。

据此,本研究提出如下政策建议。一方面,完善教育评价制度,增强教育信号的可靠性。从教育评价的标准看,要强调教育评价的全面性。学历文凭的颁发必须建立在科学多样的评价标准基础之上,更加重视学生的综合素质。从教育评价的方式看,要强调教育评价的持续性,重视教育的过程评价,适当增加教育决策阶段和人力资本投资过程的心理和精神成本等非货币性成本。从教育评价的程序看,要强调教育评价的规范性,由“严进宽出”向“宽进严出”转变,确保学历文凭的质量。另一方面,要完善劳动力市场制度,促进劳动力信息的有效流动,削弱劳动力市场信息不对称的程度,降低雇佣和工资支付决策的不确定性。

参考文献：

- 范皑皑 (2013), 《文凭膨胀：高校扩招与毕业生就业的困境与选择》, 《教育学术月刊》第4期, 第16-23页。
- 王效仿 (2003), 《知识失业：一个时代的来临》, 《社会》第4期, 第15-17页。
- 张颖、任若恩、黄薇 (2008), 《中国代际内与代际间贴现率的实验研究》, 《金融研究》第9期, 第164-175页。
- Aigner, Dennis & Glen Cain (1977). Statistical Theories of Discrimination in Labor Markets. *ILR Review*, 30 (2), 175-187.
- Altonji, Joseph & Charles Pierret (1998). Employer Learning and the Signalling Value of Education. In Isao Ohashi & Toshiaki Tachibanaki (eds.), *Internal Labour Markets, Incentives and Employment*. London: Macmillan Press, pp. 159-195.
- Altonji, Joseph & Charles Pierret (2001). Employer Learning and Statistical Discrimination. *The Quarterly Journal of Economics*, 116 (1), 313-350.
- Arrow, Kenneth Joseph (1973). Higher Education as A Filter. *Journal of Public Economics*, 2 (3), 193-217.
- Bauer, Thomas & John Haisken-DeNew (2001). Employer Learning and the Returns to Schooling. *Labour Economics*, 8 (2), 161-180.
- Broecke, Stijn (2015). Experience and the Returns to Education and Skill in OECD Countries. *OECD Journal: Economic Studies*, 2015 (1), 123-147.
- Cheung, Stephen (2010). A Test of Employer Learning in the Labour Market for Young Australians. *Applied Economics Letters*, 17 (1), 93-98.
- Clogg, Clifford, Eva Petkova & Adamantios Haritou (1995). Statistical Methods for Comparing Regression Coefficients Between Models. *American Journal of Sociology*, 100 (5), 1261-1293.
- Farber, Henry & Robert Gibbons (1996). Learning and Wage Dynamics. *The Quarterly Journal of Economics*, 111 (4), 1007-1047.
- Freeman, Smith (1977). Wage Trends as Performance Displays Productive Potential: A Model and Application to Academic Early Retirement. *The Bell Journal of Economics*, 8 (2), 419-443.

- Galindo-Rueda, Fernando (2003). Employer Learning and Schooling-Related Statistical Discrimination in Britain. *IZA Discussion Paper*, No. 778.
- Greenwald, Bruce (1986). Adverse Selection in the Labor Market. *The Review of Economic Studies*, 53 (3), 325 – 347.
- Habermalz, Steffen (2011). The Speed of Employer Learning and Job Market Signalling Revisited. *Applied Economics Letters*, 18 (7), 607 – 610.
- Hanushek, Eric & Lei Zhang (2006). Quality-Consistent Estimates of International Returns to Skill. *NBER Working Paper*, No. 12664.
- Harris, Milton & Bengt Holmstrom (1982). A Theory of Wage Dynamics. *The Review of Economic Studies*, 49 (3), 315 – 333.
- Lange, Fabian (2007). The Speed of Employer Learning. *Journal of Labor Economics*, 25 (1), 1 – 35.
- Lundberg, Shelly & Richard Startz (1983). Private Discrimination and Social Intervention in Competitive Labor Markets. *The American Economic Review*, 73 (3), 340 – 347.
- Mansour, Hani (2012). Does Employer Learning Vary by Occupation? *Journal of Labor Economics*, 30 (2), 415 – 444.
- Pan, Shih-Yi (2005). Employer Learning and Statistical Discrimination in the Canadian Labour Market. MA. diss., University of Saskatchewan.
- Phelps, Edmund (1972). The Statistical Theory of Racism and Sexism. *The American Economic Review*, 62 (4), 659 – 661.
- Pinkston, Joshua (2006). A Test of Screening Discrimination with Employer Learning. *ILR Review*, 59 (2), 267 – 284.
- Schultz, Theodore (1975). The Value of the Ability to Deal with Disequilibria. *Journal of Economic Literature*, 13 (3), 827 – 846.
- Spence, Michael (1973). Job Market Signaling. *The Quarterly Journal of Economics*, 87 (3), 355 – 374.
- Stiglitz, Joseph (1975). The Theory of “Screening,” Education, and the Distribution of Income. *The American Economic Review*, 65 (3), 283 – 300.
- Strobl, Eric (2004). Do Employers Use Education as A Signal for Ability in Developing Countries? Evidence from Ghana. *Applied Economics Letters*, 11 (4), 259 – 261.
- Waldman, Michael (1984). Job Assignment, Signalling, and Efficiency. *The RAND Journal*

of Economics, 15 (2), 255 – 267.

Wolpin, Kenneth (1977). Education and Screening. *The American Economic Review*, 67 (5), 949 – 958.

Zhangaliyeva, Aigerim & Masaki Nakabayashi (2013). Heritage from Czar: The Russian Dual System of Schooling and Signaling. *ISS Discussion Paper Series*, No. F163.

Educational Signaling and Its Market Value: From the Theory of Statistical Discrimination and Employer Learning

Wang Jun^{1,2} & Zhao Jialiang¹

(Business School, Beijing Normal University¹;

Training and Research Base for College Counselors Affiliated
to the Ministry of Education, Beijing Normal University²)

Abstract: From the theoretical perspective of statistical discrimination based on education and employer learning, this paper empirically analyzes educational signaling in China's labor market and measures its market value using the method proposed by Altonji & Pierret (2001) and Lange (2007). It is found that as working experience grows, the effect of education decreases while the effect of ability increases. Although the speed of employer learning is fast, the initial expected productivity error is relatively persistent. When the discount rate remains 7.18 percent, the upper bound on market value of educational signaling, defined by the percentage contribution of signaling to the increase in lifetime earnings due to an additional year of schooling, is about 16 percent. The simulation results are very sensitive to the discount rate, the speed of employer learning, and the length of potential experience, but not sensitive to educational cost.

Keywords: educational signaling, market value, statistical discrimination, employer learning

JEL Classification: I21, J24, J71

(责任编辑：封永刚)