

文章编号: 2095-1663(2022)03-0051-06

DOI: 10.19834/j.cnki.yjsjy2011.2022.03.08

博士生科研资助的粘性效应: 比较优势还是能力分层

——以人文社会科学为例

张双志

(成都大学师范学院, 成都 610106)

摘要: 探究博士生科研资助的粘性效应, 即科研资助对科研产出的影响是源于博士生的比较优势还是能力分层, 有着重要的政策启示价值。基于网络爬虫和手工清理获得 1341 份已毕业的博士数据, 应用反事实分析模型 (PSM) 对其进行实证检验, 研究发现: 获得科研资助的博士生平均能力高于未获科研资助者, 其科研产出是能力分层所致的分流结果, 而非遵循比较优势后的分类收益。换言之, 未获得科研资助的博士生, 假如获得科研资助的话, 其边际产出将会比获得科研资助博士生的边际产出高。对此提出的政策建议是, 高校科研资助工作应回归育人的初衷, 大幅提高科研资助的覆盖面, 以便吸纳那些处于申请科研资助边缘的博士生, 激励其产出更多的科研成果从而达到整体提高人才培养质量的目的。

关键词: 博士生; 科研资助; 科研产出; 粘性效应; 反事实分析

中图分类号: G643

文献标识码: A

一、问题提出

习近平总书记精辟指出创新是引领发展的第一动力, “人才是创新的根基, 是创新的核心要素”。创新人才有助于增强国家自主创新能力, 促进经济增长从投资驱动转向创新驱动^[1]。创新来源于实践, 只有在真实情境中的实践才能发现已有知识解决新问题的不足, 从而在解决新问题的过程中创造新的知识。“干中学”是科研资助的一大特色, 创新思维与实践能力的培养关键在于为博士生提供体验式学习场域。那么, 在创新驱动发展战略背景下, 科研资助是否有助于博士生产出更多的科研成果就显得尤为重要。然而, 以往研究大都只关注到受科研资助者的科研产出结果^[2-3], 却忽视了未受科研资助者假如也得到资助的话, 其能够生产的潜在科研成果如何? 此问题的回答涉及到如何对科研资助的绩效进行评估, 核心在于探究博士生科研资助的粘性效应,

即那些边际产出最大的博士生是不是最有可能获得科研资助的群体。如果是的话, 那么科研资助的开展无疑是最有效率的, 实现了其预设的资助目的, 这对于提升研究生人才培养质量大有裨益。

政策项目的评估面临着棘手问题是数据缺失^[4], 即如果博士生获得了科研资助, 我们就不能观测到其未获得资助后的科研产出, 这就缺少了与实际情况相反的另外一种结果数据。也就是说, 科研资助绩效评估的理想状态是比较同一个博士生获得资助和未获得资助所引起的结果变化, 以测算其真实的科研产出。鉴于此, 本文采用反事实分析模型对此进行实证研究。在数据缺失的情况下, 利用降维匹配方法在相反状态的样本中寻找和自身最“接近”的样本来近似替代其反事实情况下的结果^[5]。即为获得科研资助的某个博士生寻找到特征吻合度高的未获得科研资助的另一个博士生, 这样就可以较好地目前获得的调查数据“转变”为实验数据, 以达到类似于用实验数据进行因果推断的目的。换

收稿日期: 2021-10-15

作者简介: 张双志(1990—), 男, 贵州毕节人, 成都大学师范学院特聘副研究员, 教育学博士。

言之,反事实分析可以计算事实结果(科研资助的边际产出)与反事实结果(非科研资助的边际产出)的净差异,得到博士生真实的科研产出。据此可为降低博士生科研资助的粘性效应提供相关对策建议。

二、计量模型设定

为识别博士生在科研资助中所取得的科研产出,传统方法是采用普通最小二乘法(OLS)模型进行估计,构建如下方程:

$$Y_{ic} = \gamma D_{ic} + \sum \beta Z_{ic} + City_code_i + U_{ic} \quad (1)$$

式(1)中,下标 i, c 分别代表博士生个体和城市; Y_{ic} 表示科研产出,为博士生在读期间以第一作者身份在 CSSCI 来源期刊上发表的论文总篇数; D_{ic} 是虚拟变量,表示博士生若获得科研资助则 $D_{ic} = 1$,反之 $D_{ic} = 0$; Z_{ic} 和 $City_code_i$ 分别代表控制变量与城市效应; U_{ic} 表示期望为零的随机误差项,代表博士生其他无法观测的异质性变量,例如个人能力等。参数 γ 是核心关注系数,可以解释为科研资助对博士生科研产出的贡献率。

然而,OLS 模型假定博士生是否获得科研资助(D_{ic})是随机行为,且科研产出率 γ 对于所有个体都是无偏的。但事实上,博士生能否获得科研资助是非随机行为,其个人能力、数据误差等无法观测的异质性变量会造成 OLS 模型的估计结果有偏。为了处理可能存在的样本选择偏差问题,本文采用 Rosenbaum & Rubin^[6] 提出的反事实分析模型对博士生科研资助的粘性效应进行探讨。

具体来说,通过 $D_{ic} = \{1, 0\}$ 将博士生区分为获得科研资助的处理组和未获得科研资助的控制组,即 Y_{1ic} 为处理组($D_{ic} = 1$)的科研产出, Y_{0ic} 为控制组($D_{ic} = 0$)的科研产出,这样 $Y_{1ic} - Y_{0ic}$ 就可以表示为博士生个体 i 的真实科研产出,也称为处理效应(treatment effect)。然而,现实数据只能获取博士生 Y_{1ic} 或 Y_{0ic} 的其中一个观测值。为解决客观存在的数据缺失问题,需要假定个体 i 同时存在两种不同的结果(Y_{1ic}, Y_{0ic})分别对应着处理组与控制组的科研产出。因此,反事实分析模型可以表达为:

$$Y_{ic} = D_{ic} Y_{1ic} + (1 - D_{ic}) Y_{0ic} \quad (2)$$

式(2)中, Y_{1ic} 与 Y_{0ic} 是可观测解释变量 Z_{ic} 、城市效应 $City_code_i$ 和代表不可观测异质性因素的随机扰动项 (U_{1ic}, U_{0ic}) 的函数。其中,根据 MaCullagh & Nelder^[7] 的建议,采用 Logit 模型来估计倾向得分 D_{ic}^* , 这样可以将选择模型表达为一般化线性模型,即博士生 i 是否获得科研资助可以

用潜变量选择模型来表示:

$$P(D_{ic}^*) = \gamma_{D_{ic}} Z_{ic} - U_{D_{ic}} = 1 / (1 + e^{-x_{ic} \beta_{ic}}) \quad (3)$$

式(3)中, β_{ic} 表示向量的回归参数; 当 $D_{ic}^* \geq 0$ 时,博士生 i 获得科研资助($D_{ic} = 1$); 反之,当 $D_{ic}^* < 0$ 时,博士生 i 未获得科研资助($D_{ic} = 0$); Z_{ic} 代表可观测到影响个体 i 是否获得科研资助的特征变量; $U_{D_{ic}}$ 表示博士生是否获得科研资助不可观测异质性因素的随机误差项,可以理解为个体 i 完成科研资助任务的机会成本。为了完成科研资助所规定的任务,受资助者需要牺牲一些娱乐活动时间来努力完成科研产出。因此,博士生的科研产出与 $\gamma_{D_{ic}} Z_{ic}$ 成正比,与 $U_{D_{ic}}$ 成反比, D_{ic}^* 可以表示为个体获取科研资助所得的净收益。

在解决个体异质性特征导致的估计结果有偏问题后,根据 Bjorklund & Moffitt^[8] 的研究结论,将边际处理效应(MTE)定义为博士生获取科研资助后的边际产出: $MTE = E(Y_{1ic} - Y_{0ic} | D_{ic} = P(Z_{ic}))$, 这里 $P(Z_{ic})$ 代表的是利用 Logit 模型估计的博士生获得科研资助的概率倾向分数。进一步, Heckman^[9]、McCaffrey et al.^[10] 对边际处理效应进行积分处理后得到以下三种不同表达方式的平均处理效应:

①将博士生个体 i 获取资助后科研产出与其未获取资助后科研产出的平均差距定义为“平均处理效应”(ATE): $ATE = E(Y_{1ic} - Y_{0ic} | D_{ic})$; ②将博士生个体 i 获取资助后科研产出定义为“处理组的平均处理效应”(ATT): $ATT = E(Y_{1ic} - Y_{0ic} | D_{ic} = 1)$; ③将博士生个体 i 未获取资助后科研产出定义为“控制组的平均处理效应”(ATC): $ATC = E(Y_{1ic} - Y_{0ic} | D_{ic} = 0)$ 。这样可以直观描述博士生科研资助的粘性效应,方便后文实证分析的开展和结果解读。

三、数据说明及统计特征

(一)数据说明

本文所用数据来源于“双一流”建设高校官网、中国知网数据库(CNKI)和百度搜索数据库。考虑到自然科学与工程学科博士生的学术论文发表以国外期刊为主,与人文社会科学的博士生存有不同偏好。因此,本文的数据搜集样本仅限于人文社会科学,研究结论也只能谨慎地解读该学科范畴。数据搜集的具体步骤为:首先,通过 Pandas 库函数 contains 对“双一流”建设高校官网中马克思主义学院、公共管理学院和教育学院的“师资简介”部分进行关键词模糊搜索,按照姓名、性别、年龄、教育经历

(本硕博就读学校)、研究成果等栏目依次进行整理;其次,接着在 CNKI 数据库中按照姓名、博士毕业学校对第一次整理的数据库再进行关键词精准匹配,以统计高校专任教师在博士就读期间以第一作者身份发表在 CSSCI 来源期刊上的学术论文总篇数;最后,基于百度搜索数据库对网络爬虫所得数据进行手工清洗核实,完善样本中尚在缺失的数据,对于无法补充和确认的数据予以删除,以保证各变量测算的可信度。考虑到数据的可得性和实效性,只统计从 2009 年到 2018 年毕业于国内高校的博士。简言之,样本对象为 2019 年 6 月 30 日前现就职于“双一流”建设高校马克思主义学院、公共管理学院和教育学院的中青年专任教师,经过上述三个步骤的筛选、清洗和编码后获得 1341 份博士数据。

(二)统计特征

1. 变量测度

Y(outcome)是被解释变量,参照鲍威^[11]、梁文艳^[12]等学者的已有处理方法,本文采用在读博士期间以第一作者身份发表在 CSSCI 来源期刊上的学术论文总篇数来衡量博士生的科研产出。虽然“导师一作博士生二作”在学界也被视为衡量博士生学术能力的一个重要指标,但囿于将导师加入到样本

数据后会明显提升手工清洗与核对的困难度,故只整理博士生以第一作者身份发表的学术论文总篇数。

D(subsidy)是核心解释变量,若样本对象在博士就读期间获得过科研资助则赋值为 1,反之赋值为 0。这里的科研资助特指政府、高校等为了提升研究生培养质量而专项设立的研究生科研基金项目,例如天津市研究生科研创新项目、华东师范大学优秀博士学位论文培育资助项目、陕西师范大学博士研究生自由探索项目、对外经济贸易大学研究生科研创新基金等。鉴于研究生科研创新项目一般都规定要在发表论文中明确标注受资助项目的全称作为结项依据,这也成为判断样本对象在博士就读期间是否获得科研资助的区分标准。

Z 则为控制变量,考虑到数据的可得性,主要采用性别、入学方式、是否延期毕业、博士就读学校、本科就读学校等变量用以缓解由于遗漏重要解释变量可能引起的内生性问题。同时,为规避高校所属城市的不同经济发展状况可能会对回归系数产生偏误影响,加入城市固定效应(city_code)以保证结果稳健。变量设计及定义的具体情况如表 1 所示。

表 1 变量设计、定义与特征

变量名称	变量符号	变量定义	均值	标准差	最小值	最大值
科研产出	outcome	博士就读期间以第一作者身份在 CSSCI 来源期刊上发表的学术论文总篇数	2.10	3.76	0	8
科研资助	subsidy	若博士就读期间获得过科研资助则赋值为 1,反之赋值为 0	0.27	0.44	0	1
性别	gender	若性别为男性则赋值为 1,反之赋值为 0	0.64	0.32	0	1
入学方式	exam	若博士入学方式是通过公开招考(含申请审核)获得则赋值为 1,反之赋值为 0	0.48	0.16	0	1
是否延期毕业	delay	若博士毕业时就读年限超过学制规定则赋值为 1,反之赋值为 0	0.26	0.34	0	1
博士就读学校	doc985	若博士就读学校为“985”高校则赋值为 1,反之赋值为 0	0.46	0.41	0	1
	doc211	若博士就读学校为“211”高校(不包括“985”高校)则赋值为 1,反之赋值为 0	0.39	0.29	0	1
本科就读学校	bac985	若本科就读学校为“985”高校则赋值为 1,反之赋值为 0	0.38	0.52	0	1
	bac211	若本科就读学校为“211”高校(不包括“985”高校)则赋值为 1,反之赋值为 0	0.42	0.26	0	1
高校所属城市	city_code	博士就读高校所属城市依次划分为东、中、西三个不同类别	—	—	—	—

注:city_code 是城市固定效应,对其赋值不具有实际统计意义,故不汇报其描述性指标。

2. 描述性说明

表 1 汇报了均值、标准差、最小值和最大值等变

量的描述性指标。具体而言,样本对象的男性占比为 64%;通过公开招考(含申请审核)方式取得博士

入学资格的比例为 48%;延期毕业率为 26%;在博士就读期间获得过科研资助的比例为 27%;以第一作者身份在 CSSCI 来源期刊上发表的论文人均篇数为 2 篇,其中最大值是 8 篇,最小值为 0 篇;博士就读于“985”高校的占比为 46%,而就学于“211”高校(不包含“985”高校)的比例为 39%;本科就读于“985”高校的占比为 38%,而就学于“211”高校(不包含“985”高校)的比例为 42%。样本数据说明“211”及“985”高校的博士毕业生已成为“双一流”建设高校马克思主义学院、公共管理学院和教育学院中青年专任教师队伍的最主要来源。

四、实证结果

(一)平衡性检验

从理论上说,完美的倾向值匹配可以严格忽略处理分配机制的影响,直接对被解释变量进行处理效应分析。然而,在现实调查之中我们无法完全获得所有的控制变量使倾向值匹配符合严格可忽略分配的假定,这就使得为了消除样本选择偏差的倾向值匹配方法自身也存在选择偏差问题。为了较好地验证倾向值匹配之后的样本是否仍然存在选择偏差问题,需要对控制变量的分布平衡性进行检验。根据方程(3)运用 Logit 模型对博士生个体 i 的倾向值 D_{ic}^* 进行估算,使用卡尺内最近邻匹配法在共同支持域上实现处理组与控制组的配对。卡尺内最近邻匹配法具有独特的优势,那就是匹配后的处理组与控制组在可观测的控制变量上是平衡的。新样本可以像随机实验获得的数据一样允许研究者使用几乎所有类型的多元回归方法来分析因果效应,这无疑对因果效应的研究很有帮助。Rosenbaum & Rubin^[13]指出当控制变量标准化偏差的绝对值小于 20% 时才能通过平衡性检验。如图 1 所示,控制变量在倾向值匹配前后的标准化差异比较大,匹配后所有控制变量标准化偏差的绝对值均小于 5%,且明显集中在零点附近。说明匹配后控制变量分布较匹配前更为平衡,其标准化平均值之差更趋近于零,由此判断经过倾向值匹配后的样本数据适合进行反事实分析。

(二)反事实分析

参照 Abadie et al.^[14]的研究成果,采用核匹配(kernel matching)和样条匹配(spline matching)模型对整体平均处理效应(ATE)、资助者平均处理效

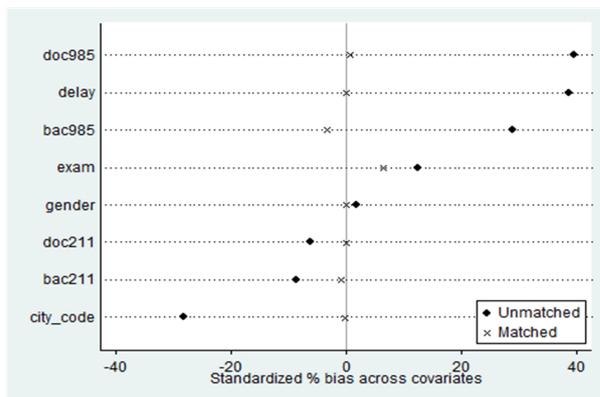


图 1 控制变量的标准化偏差对比图

应(ATT)与未资助者平均处理效应(ATC)进行估算,并与 OLS 模型估算结果进行科研资助选择偏差、科研产出分类收益等方面的比较分析,具体结果如表 2 所示。

1. 基本回归结果

OLS 模型对博士生科研产出的估算值为 0.22,通过了 1% 的显著性水平检验,说明在不考虑个体不可观测异质性变量的情况下,科研资助对博士生科研产出的贡献率为 22%。第(1)列报告了整体平均处理效应(ATE)的估算值范围为 0.1500~0.1520,说明随机给予一名博士生以科研资助要比其未获得资助情况下提高 15% 的科研产出。第(2)列汇报了资助者的平均处理效应(ATT)估算范围为 0.0922~0.0993,意味着 OLS 模型对科研资助对博士生科研产出的贡献率明显高估了 12%,也验证了对其进行倾向值匹配分析的必要性。第(3)列报告了非资助者的平均处理效应(ATC)估算范围在 0.1717~0.1719 之间,揭示了未获得科研资助博士生的边际产出要比获得科研资助者高出 7 个百分点。需要解释的是,这是基于反事实分析模型估算的结果,比较的是博士生是否获得科研资助后的不同潜在结果,并不是资助前与资助后的实际科研产出结果。

2. 科研资助的选择偏差

第(4)列汇报了 ATE 的偏差范围为 0.0719~0.0739,表示不考虑样本选择偏差问题的 OLS 模型结果要比整体平均处理效应高 7%,说明忽视博士生的不可观测异质性变量将导致对科研产出的过高估计。显然,不可观测的异质性变量对博士生的科研产出结果产生了重要影响。进一步,第(5)列汇报了 ATT 的选择偏差范围在 0.1246~0.1317 之间,第(7)列报告了 ATC 的选择偏差范围为 -0.0520

~-0.0522。也就是说,科研资助者的选择偏差在不同的半参数估算方法中都是正数,非科研资助者的选择偏差皆为负数。根据 Heckman et al.^[15] 的研究结论,如果资助者与非资助者的选择偏差都是负数的话,那么博士生是否获得科研资助是遵循了比较优势原则,即非资助者即使获得科研资助,其边际产出结果也比不上实际的受资助者,反之亦然。如果资助者的选择偏差为正,而非资助者的选择偏差为负,说明博士生是否获得科研资助是由于能力分层导致的选择结果,即受资助者的平均能力高于非受资助者。据此由第(5)列与第(7)列的汇报结果可知,博士生是否获得科研资助并不遵循个体的比较优势,而是由于资助者与非资助者之间明显的的能力差距所导致的分流结果。这揭示了资助者与非资助者的科研产出不能简单以是否获得科研资助作为评判依据,还需要考虑博士生之间客观存在的能力差距,否则会明显高估科研资助的成果产出效应,在

一定程度上过分夸大科研资助的作用。

3. 科研产出的分类收益

第(6)列与第(8)列分别报告了资助者和非资助者科研产出结果分类收益。具体来说,资助者边际产出的分类收益范围为-0.0527~-0.0578,非资助者边际产出的分类收益范围在-0.0199~-0.0217之间,说明资助者与非资助者的分类收益都是负数。依然根据 Heckman et al. 的研究结论,正的分类收益与个体的自我选择有关,而负的分类收益与不可观测的异质性变量有关。换言之,博士生是否获得科研资助受到一些不可控因素的影响,可能与高校科研资助的管理机制有关,即存在组织层面的选择偏差。高校为了保证科研资助工作能在短时间内取得较好的成绩,自然会优先资助能力较高的博士生,这本身无可厚非。然而,挑选赢家的管理方式可能会背离科研资助的初衷,造成“强者愈强、弱者愈弱”的马太效应。

表 2 选择偏差与分类收益

	ATE	ATT	ATC	偏差	资助者		非资助者	
					选择偏差	分类收益	选择偏差	分类收益
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
核匹配	0.1520*** (0.0430)	0.0993** (0.0479)	0.1719*** (0.0444)	0.0719	0.1246	-0.0527	-0.0520	-0.0199
样条匹配	0.1500*** (0.0416)	0.0922** (0.0472)	0.1717*** (0.0430)	0.0739	0.1317	-0.0578	-0.0522	-0.0217

注:***、**、* 分别表示通过 1%、5% 和 10% 的显著性水平检验;() 内为估计系数自抽 500 次后获得的 Bootstrap 标准差。

五、结语

在对样本选择偏差问题进行处理后,ATT 的估计值为 9% 明显低于 OLS 模型的估计值,说明不可观测异质性变量对于博士生是否获得科研资助有着重要的影响,忽视其将对博士生的科研产出结果产生误判。进一步,通过核匹配和样条匹配模型探究博士生科研资助的粘性效应。发现资助者的平均处理效应(ATT)低于整体平均处理效应(ATE),而后者又低于未受资助者的平均处理效应(ATC)。说明那些最有可能被高校挑选出来的高能力者,从科研资助中获得的边际产出结果是较低的;反而,那些不太可能被挑选出来的博士生,假如获得科研资助的话会取得较高的边际产出结果。简言之,科研资助对博士生科研产出所带来的影响并不遵循个人的

“比较优势”,而是其“能力分层”导致的分流结果。

为了更好地达成科研资助的目的,结合本文实证研究结论,提出如下对策建议:

第一,发挥基金资助的指挥棒作用,引导博士生的科研服务于国家战略发展需要。创新是引领发展的第一动力,具有创新思维与实践技能的高素质人才在产业升级换代与新旧动能转换过程中发挥着日益重要的人力资本作用。这就要求基金资助要对接创新驱动发展战略的实际需要,引导广大博士生将论文写在祖国大地之上,在教育链、人才链与产业链、创新链的深度融合中实现青年学者的人生价值。

第二,扩大基金资助的覆盖面,增强博士生的科研获得感。博士生是否获得科研资助是高校挑选赢家的客观结果,高校为了取得较好的科研资助绩效,自然会优先资助能力较高的博士生,而忽略了为其他博士生提供适宜的科研支持环境。因此,高校应

该让科研资助回归育人的初衷,尽量规避挑选赢家造成的育人功能扭曲。鼓励那些处于科研申请边缘的博士生积极参与科研活动,通过“干中学”引导其找到正确的研究方法。

第三,完善“院→校→省”三级科研资助机制,发挥“以赛促学”的资助目的。充分发挥省级科研资助的示范引领作用,支持校级科研资助和院级科研资助结合自身实际条件,创新性推动科研资助的特色化与规范化。省级和校级科研资助可以挑选赢家,为能力较高的博士生提供施展才华的平台。同时,考虑到博士生群体总量不大的实际情况,院级科研资助则应强调全员参与性,提升科研资助的普惠度在财政资金可承受方面也具有一定的可行性。

参考文献:

- [1] 闵维方. 新时代教育发展的战略重点是提高国家创新能力[N]. 中国教育报,2018-12-13.
- [2] 管佳,李奇涛. 我国研究生科研资助的现状、问题与对策分析[J]. 学位与研究生教育,2012(7):48-52.
- [3] 马荣康,金鹤. 高校技术转移对科研产出的影响效应研究——科研资助的中介作用与调节作用[J]. 科研管理,2020(5):279-288.
- [4] 郭申阳,Mark W. Fraser. 倾向值分析:统计方法与应用[M]. 重庆:重庆大学出版社,2016:1-3.
- [5] 温兴祥,杜在超. 匹配法综述:方法与应用[J]. 统计研究,2015(4):104-112.
- [6] Rosenbaum P R, Rubin D B. The Central Role of The Propensity Score in Observational Studies for Causal Effects [J]. *Biometrika*,1983(70):41-45.
- [7] McCullagh P, Nelder J. *Generalized Linear Models* (2nd ed.) [M]. London:Chapman & Hall,1989:22-28.
- [8] Bjorklund A, Moffitt R. The Estimation of Wage Gains and Welfare Gains in Self-selection Models [J]. *The Review of Economics and Statistics*,1987,69(1):42-49.
- [9] Heckman J J. Comment on “Identification of Causal Effects Using Instrumental Variables” by Angrist, Imbens, & Rubin [J]. *Journal of the American Statistical Association*,1996,91:459-462.
- [10] McCaffrey D F, Ridgeway G, Morral A R. Propensity Score Estimation With Boosted Regression for Evaluating Causal Effects in Observational Studies [J]. *Psychological Methods*,2004(9):403-425.
- [11] 鲍威,金红昊,田明周. 我国研究型大学教师队伍年龄结构与科研产出的关系 [J]. *高等教育研究*,2020(5):54-62.
- [12] 梁文艳,周晔馨,于洪霞. 社会资本与大学教师学术创新能力研究 [J]. *经济研究*,2019(11):133-148.
- [13] Rosenbaum P R. Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score [J]. *The American Statistician*,1985,39(1):33-38.
- [14] Abadie A, Drukker D, Herr J L, et al. Implementing Matching Estimators for Average Treatment Effects in Stata [J]. *Stata Journal*,2004(4):290-311.
- [15] Heckman J J, Vytlačil E. Structural Equations, Treatment Effects, and Economic Policy Evaluation [J]. *Econometrica*,2005,73(3):669-738.

The Stickiness Effect of Doctoral Research Fund: Comparative Advantage or Ability Stratification? —A Case Study based on Humanities and Social Sciences

ZHANG Shuangzhi

(College of Teachers, Chengdu University, Chengdu 610106, China)

Abstract: Exploring the stickiness effect of the scientific research fund granted to doctoral students, i. e., exploring whether the influence of the fund on scientific research output is due to comparative advantage or ability stratification, has important policy enlightenment value. Based on the data of 1,341 doctoral graduates obtained through web crawler and manual mining, the author uses the counterfactual analysis model (PSM) to carry out an empirical test. The study finds: the average ability of the doctoral students who have obtained research fund is higher than those who have not and the scientific research output is the result of regrouping caused by ability stratification, rather than the categorized benefits from the comparative advantage. In other words, if the unfunded doctoral students can obtain the fund, the marginal output they produce will be larger than that by the doctoral students who have been funded. This means a suggestion for policy adjustment: universities return to the original mission for cultivating students when using scientific research fund, increase the coverage of the fund greatly so as attract more doctoral students who wish to apply for the fund, encourage them to yield more scientific research achievements, and improve the overall quality of talent cultivation.

Keywords: doctoral students; research grants; scientific research output; stickiness effect; counterfactual analysis