

为贫困生教育赋能

——一项随机对照田野实验研究

何 青 王军辉 甘 犁

【摘 要】教育扶贫重点从提高入学率转向提高学生的在校学习质量。本文通过设计一个针对小学生的随机对照实验,研究了经济激励对贫困地区教育质量的影响效果及其影响机制。激励方案重点在于对学习进步和学习过程进行激励。分析发现实验第一期效果不显著,第二学期和第三学期效果明显且稳定,效果在各种稳健性检验后稳定。激励效应在女生、少数民族学生、父母受教育程度低、中低收入家庭学生、期初成绩较差的学生以及较差学校的学生等相对弱势群体更大。机制分析发现,经济激励提高了家长对孩子教育的关注程度、时间和资金投入,也促进了学生提升学习积极性和增加学习时间。激励效应在实验结束后得到一定程度持续。本研究对提高贫困地区义务教育阶段教育质量的相关政策制定有一定参考价值。

【关 键 词】现金激励;教育质量;学习过程;贫困;田野实验

【作者简介】何青,西南财经大学中国家庭金融调查与研究中心;王军辉(通讯作者),西南财经大学中国家庭金融调查与研究中心;甘犁,西南财经大学中国家庭金融调查与研究中心。

【原文出处】《管理世界》(京),2022.9.63~82,112

【基金项目】本研究得到国家自然科学基金面上项目(基金号:72073108)、国家社科基金青年项目(基金号:20CJL018)、高等学校学科创新引智计划(项目编号:B16040)、中国人口福利基金会和招商局慈善基金会青年学者资助计划之“何为而治——中国当代社会治理研究”二期的资助。

一、引言

经过长期努力,无论是国际还是国内,中小学入学率都已经达到一个比较高的水平。贫困群体教育面临的主要问题已经从如何提高入学率转向如何提高教育质量。根据世界银行的“全球教育质量”数据,虽然国家之间的入学率差异已经很小,但学习质量差异很大,低收入国家普遍存在学习危机,低收入国家的数学和阅读及格率分别为13.9%和4.5%,远远低于高收入国家93.2%和95.3%(World Bank, 2018)。我国贫困地区的学习质量也极其堪忧。有研究发现中西部四年级学生标准化测试的数学成绩比东部学生低30多分(周达等,2017),贫困县的小学四年级学生数学成绩比非贫困县要低30分,并且女生的差距更大(张彩等,2015),即使进入大学后,贫困地区的大学生的学业表现仍比其它学生要差(牛新

春,2017)。在校期间较差的学习质量会降低入学的人力资本积累效率,进而严重影响到后期的收入和人生发展(Hanushek and Wößmann, 2007)。

如何提高贫困地区和贫困家庭学生入学后的学习质量,成为当前农村和贫困地区基础教育阶段迫切需要关心的一个问题。大量文献从家庭背景和收入(Dahl and Lochner, 2012; Lacour and Tissington, 2011)、同伴效应(Peer Effects)(Ding and Lehrer, 2007)、学校绩效责任(Hanushek and Raymond, 2005)、学校管理水平(Bloom et al., 2015)、作业布置(Eren and Henderson, 2008)、教师教学方法(Bonesrønning, 2004)、学校资源投入(Neilson and Zimmerman, 2014)、教师质量(Jackson et al., 2014)等角度出发研究如何提高学生学业表现。中国提高贫困地区教育质量的路径主要是改善教学条件和教师质量,如增加校舍、增加基础

教育阶段信息化技术和设备的投入、实施教师特殊岗位计划和交流轮岗计划等。但仅加大对学校教学资源的投入可能不足以有效改善学习质量。著名的科尔曼报告发现学校资源对教育质量影响非常微弱(Coleman et al., 1966)。如果缺乏激励性措施,仅仅增加对学校的资源投入、增加计算机和网络辅助教学等均不足以改善学生的学业表现(Bulman and Fairlie, 2016; Hanushek, 2003; Martorell et al., 2016)。

对于提高学习质量,除需提升教育供给质量外,提升学生、家长在学习上的投入和学生的学习动力也至关重要。一方面,贫困家庭学生和家長在人力资本投资往往过分关注当前,不会关注长远利益(Lavecchia et al., 2016)。近年来“读书无用论”在农村地区广为流传,家长和学生往往难以意识到教育对人生发展的重要作用,在深度贫困地区更是如此。教育学、心理学和经济学的一些研究也表明,儿童和青少年更容易只关注短期利益,缺乏耐心(Bettinger, 2012)。面对巨大的教育回报率,很多家庭和小孩的教育投入往往没有达到最优水平,在贫困家庭更为明显,这成为教育经济学里一个巨大谜团(Levitt et al., 2016b)。另一方面,缺少学习动力,成绩较差的小孩过早自我放弃也是导致学习危机的重要原因(World Bank, 2018)。有研究发现,贫困家庭针对入学和学习方面的消极态度是贫困家庭学生学习成绩差的一个重要影响因素(Banerjee, 2016)。《中国教育扶贫报告(2016)》发现,越来越多的辍学并非由经济困难导致,而是由学生学习动力不足、知识改变命运的信心不足和自我期望值不够高等非智力因素造成的(司树杰等, 2016)。

为学生提供根据学习目标而设定的短期经济激励是解决贫困学生学习动力不足的一种方式(Fryer, 2011)。学习是一个生产函数(Bowles, 1970; Glewwe et al., 2013),学习成绩是学习产出,学生做作业和学习时间投入、跟老师和同学交流、查阅课外书、认真听讲和按时出勤、学校教师 and 教学设施、家庭时间和资金投入等都是学习生产函数的投入要素。Fryer (2011)认为,很多学生并不清楚学习的生产函数,只对结果进行激励并不一定有效果。据此,学习激励可分为学习结果目标激励和学习投入目标激励。学

习结果目标激励主要针对学生缺乏学习意愿、对学习预期收益估计不足等问题。学习投入目标激励主要针对学生有学习意愿但不清楚学习的生产函数,不知道如何正确学习和分配学习投入,导致学生的学习效率很低的情况。通过给某些关键性学习环节设立奖励,引导学生增加学习投入,提高学习效率,从而达到最终提高学习质量的目标。

本文通过设计一个针对贫困地区小学生的现金激励的随机对照实验,研究经济激励对贫困地区学生学业表现的影响效果及其影响机制。激励方案同时考虑了学习结果和学习过程,针对学习成绩优秀、取得较快进步和平时作业完成情况较好学生分别给予现金奖励。学习优秀奖主要针对学习结果,作业完成奖主要针对学习过程,学习进步奖则同时涉及学习结果和学习过程。

研究发现,实验第一期效果并不明显,到第二期和第三期效果显著。激励效应在女生、少数民族学生、父母受教育程度低、中低收入家庭学生、期初成绩较差的学生以及较差学校的学生等相对弱势群体更大。机制分析发现,现金激励提高了家长对孩子教育的关注程度、时间和资金投入,也促进了学生提升学习积极性和增加学习时间。现金激励效应在实验结束后得到一定程度持续,主要体现在语文成绩进步的保持和获奖学生成绩进步的保持上。作业奖获得者语文成绩保持了与优秀奖获得者相近的进步程度,说明奖励平时学习表现与奖励优秀同等重要。本文的研究结果表明,通过对学习过程和学习结果施加适当的经济激励可有效激发贫困地区学生的学习动力,进而提高贫困地区的教育质量。

本研究采用田野实验的方法研究贫困地区教育问题,是对国内该领域相关研究的重要补充。具体而言本研究的贡献主要体现在以下几方面。

首先,本研究考察了激励效果的动态变化和实验结束后的持续性,有效回应了以往类似实验难以回答的激励效果能否持续性这一问题。现金激励能否产生持续影响是很多学者和政策制订者关心的问题,但多数文献囿于多方面的原因限制,难以跟踪到激励本身的动态效应,也不知道现金激励政策取消之后激励效应是否有持续。本研究通过持续3个学

期的实验干预和实验后持续跟踪学生学业表现,发现在实验期内和实验结束后,激励效果都具有一定可持续性。

其次,与已有研究相比,本研究设计的激励方案创新性地综合考虑了对学习结果和学习过程的激励。以往相关研究多数是对学习成绩等学习结果进行激励,有少部分是对阅读、出勤等学习过程进行激励,较少同时考虑把学习结果激励和学习过程激励结合起来。本研究将对学习结果和学习过程的激励结合,激励方案既考虑了对绩优生的激励,也照顾到以往难以获得奖励的中下游学生。

再次,本文还从家长行为和学生行为两方面研究了现金激励的影响机制,有效弥补了已有文献的不足。现有的文献多数是发现现金激励与学业表现之间的因果关系及异质性影响,对传导机制关注度不够,一方面无法推进对现金激励效果的认识,另一方面也难以以为政策制订提供更有效的指导。

最后,在实践方面,本文对民族贫困地区的教育政策制订有重要的参考价值。实验开展地区为国家级贫困县,近一半人口为少数民族,样本包含大量少数民族学生,本研究为在民族地区开展激励性教育干预提供了经验证据。本研究发现奖励进步和奖励作业完成度高对于鼓励成绩落后的学生好好学习有巨大作用,这种影响对于家境相对较差、少数民族学生和女生更显著,这为民族贫困地区制定教育支持政策提供经验证据支持。

本文后面的安排如下:第二部分对相关文献进行了梳理和评述,第三部分描述了我们的实验设计,第四部分报告了数据描述和描述性统计,第五部分是实证结果分析和机制分析,第六部分分析了实验结束以后激励效应是否依然持续,第七部分对全文进行了简要总结。

二、相关文献

大量文献分别从老师、学生、家长3个角度研究经济激励对学生学习成绩的影响。本文主要关注的是给学生提供经济激励的效果,但目前的相关研究尚无一致结论。

部分研究发现经济激励对学生的成绩只有很小影响。Angrist等(2009)对加拿大的一个大学生

学习帮助项目进行评估,该项目将高中阶段GPA在最高1/4分位数以下的大学生随机分成4组。第1组提供学习帮助,第2组提供奖学金,成绩保持B以上获5000美元奖学金,C+以上获1000美元奖学金;第3组既提供奖学金,也提供学习帮助,第4组不做任何处理,发现提供奖学金对提高学习成绩只有很小的效果。Leuven等(2010)对荷兰阿姆斯特丹大学生学习奖励项目进行评估,该项目将阿姆斯特丹大学一年级新生随机分成3组,进行随机对照实验。第1组学生成绩达到一定标准,奖励681欧元;第2组学习达到一定标准,奖励227欧元;第3组不做任何处理,也发现现金奖励对学习成绩提高只有很小的效果。Sjoquist和Winters(2015)评估了全美25个州的成绩资助项目,这些项目根据高中的GPA、ACT或者SAT成绩、大学学分和GPA来给予奖励,发现24~30岁年龄段拥有大学学历的人口下降了0.2%(但不显著)。Richburg-Hayes等(2011)对纽约社区大学22~35岁之间需要社区矫正的学生,第二和第三学期里如果注册、期中报到且最后6学分里能拿到2.0的GPA,则每学期给予1300美元的奖金,发现第一学期学生多拿了0.6个学分,报到率上升7.4%,但GPA没有显著差异。Bettinger(2012)研究了美国Coshocton县对小学生考试成绩达到一定水平给予现金奖励的随机实验,发现只有数学成绩显著提高了0.15个标准差,但对阅读、社会科学、科学等科目成绩都没有影响。Fryer(2011)通过对美国4个州260所公立小学的4个实验进行评估,发现针对学习成绩等结果指标进行激励无效。

部分研究则发现经济激励对学习成绩有一定的作用。Kremer等(2009)对肯尼亚一个6年级学生奖励项目进行评估,该项目对成绩最好15%的女生给予奖学金,发现女生成绩提高0.2个标准差,并且对于成绩中上的学生影响最大,对男生的成绩也有正的外溢效应。Angrist和Lavy(2009)对以色列的高中生学习激励项目进行评估,该项目将40所学校随机分成两组,对其中一组取得进步的学生提供经济激励,发现进入项目的学校,高中毕业考试通过率提高6~8个百分点。Jackson(2010)研究美国德克萨斯州1994~2005年间的APIP项目(Advanced Placement In-

centive Program), 该项目给通过考试的学生发 100 ~ 500 美元不等的奖励, 同时给老师也发奖励, 发现 11 ~ 12 年级学生考试通过率提高了 2.4%, SAT 分数超过 1100 分以上的人数增长了 13.5%。Oswald 和 Backs-Gellner(2014) 对一个学徒项目进行评估, 这个项目根据学生成绩发放奖励, 发现 GPA 上升了 0.36 个标准差。Riccio 等(2013) 研究了纽约的“机会”项目, 该项目给学生发以下奖励: 出勤率达到 95% 每月发 25 美元; 标准化考试及格或者优秀发 300 或者 600 美元; 参加家长会发 25 美元。发现 4 年级学生考试优秀率在第一年上升 2.1%, 第二年上升 1.7%, 7 年级学生考试优秀率第一年上升 0.8%, 第二年下降 1.6%。Castleman 等(2014)、Cha 和 Patel(2010)、Dynarski(2008)、MacDonald 等(2009) 和 Scott-Clayton(2011) 等研究了美国、墨西哥基于高中成绩和大学成绩给予大学生补贴学费的奖励政策, 发现普遍能提高大学毕业率、减少毕业延迟率, 学生能拿到更多学分。Angrist 等(2014)、Leuven 等(2011) 以大学生为研究对象, 发现根据大学生成绩给予现金奖励能显著提高大学生成绩, 且金额越大效果越大。

部分研究关注到了学习过程的激励。Guryan 等(2016) 研究美国东北部给小学四五年级学生每读一本书给予 5 美元的奖励项目, 发现实验组学生多读了 1.3 本书, 对推送书籍与阅读兴趣高度相符的学生多读了 4.7 本书。Springer 等(2015) 研究了一个针对中小学生的接受教育辅导服务的激励项目, 如果学生参加了分给他们的辅导课时间的 25% 或者 75%, 给他们发一个证书, 或者发 25 美元, 如果参加时间达到 100%, 另外再发 50 美元, 发现发证书组的学生参加辅导课的时间多了 43%, 发奖金组的学生多了 6%。MacDonald 等(2009) 研究了一个针对安大略社区大学生的激励方案, 获得奖励的条件包括完成至少 12 小时辅导课、案例或者讲座课, 发现第一学期 GPA 上升了 0.07, 第二学期上升了 0.12, 第三学期上升了 0.01。女生和大龄学生的效果更为明显。Miller 等(2012) 给新墨西哥大学低收入学生, 4 个学期中对 GPA 不低于 2.0、全日制且每学期多参加两次导师见面会的学生提供 1000 美元的奖金, 发现第一学期学分没有差异, 第二学期多拿了 0.6 个学分。

部分研究比较了不同激励对象、激励方式的效果差异。Fryer(2011) 通过对美国纽约、达拉斯、芝加哥、华盛顿 260 所公立小学在 2007-2009 年间的 4 个实验进行评估, 比较了对学习结果和学习投入分别进行激励的效果, 发现针对学习成绩等结果指标进行激励无效, 对投入激励有效。Barrow 和 Rouse(2018) 对纽约和加州的高三学生和大学生奖学金项目进行评估, 该项目对学生根据绩点进行奖励, 发现学习时间、学习质量显著上升。Blimpo(2014) 对贝宁的一个奖励项目进行评估, 该项目设置了个人学习目标奖励、团队奖励和锦标赛制奖励等 3 类奖励方式, 3 种奖励方式效果都很明显且接近。Behrman 等(2015) 对墨西哥的高中生奖励项目进行评估, 该项目把学生分成 4 组, 第一组只对学生的学习成绩进行奖励, 第二组对老师进行奖励, 第三组对老师和学生都进行奖励, 第四组不做任何处理。发现对学校管理层、老师和学生同时进行奖励的组效果最大; 只对学生进行奖励的组次之; 只对老师进行奖励的没有效果。Levitt 等(2016a) 比较了针对家长和学生给与经济激励的作用差异, 没有发现给家长和学生之间的经济激励在效果上存在显著差异。该项研究是对美国芝加哥在 2008 年的一个九年级学生奖励项目进行评估, 基于出勤、行为、学习成绩等综合表现, 按月发放奖励, 支付方式包括给家长和给学生, 发现整体学习成绩上升 15% ~ 22%。Levitt 等(2016b) 比较了现金奖励额度和现金与非现金奖励的激励效果差异, 发现大额现金奖励(20 美元) 能使得成绩提升 0.068 ~ 0.153 个标准差, 10 美元的奖励额度则没效果, 价值 3 美元的非现金奖励的效果居于前面两种现金奖励额度之间, 另外延迟奖励也没有激励效果, 低龄学生对非现金激励更敏感, 大龄学生对现金激励更敏感, 数学成绩的激励效应比阅读成绩更强, 男生的激励效应要大于女生, 但激励没有持续效应。

我国目前关于现金激励对学生成绩影响的研究较为少见。Li 等(2014) 研究了群体激励对学习成绩的影响, 发现能提高低成绩学生的成绩 0.265 个标准差, 但针对个体的激励没发现有显著效果。这一研究关注到了学生之间的互助行为。

整体来说, 前期研究中的激励方案有以下特点:

(1)对学生的激励多数是针对大学生和高中生,对初中生和小学生的激励项目较少;(2)激励对象普遍是对表现优秀的学生,对表现不够优秀但有进步的学生很少给予奖励,导致大部分成绩不太好的学生都很难得到奖励,激励范围受限,有些“锦上添花”的意味,但反贫困政策更需要的是“雪中送炭”;(3)多数是在发达国家(如美国、加拿大、荷兰等)进行的研究项目,专门针对贫困地区和贫困学生的奖励项目较少,而目前的学习质量危机主要是发生在贫困地区和贫困家庭;(4)对学习过程的关注度不够,多是基于对成绩的奖励,学习投入要素中的阅读、辅导、讲座课等受到关注,自习时间、作业完成等则目前缺少关注。

本文的研究是基于在贫困地区小学开展现金激励的随机对照实验,激励方案的设计既包括了针对学习结果的学习优秀奖,也包括了针对学习过程的作业完成奖,以及同时涉及学习过程和学习结果的学习进步奖,进一步丰富了现有的文献。

三、实验设计和实施

(一)实验地点背景

实验在四川省某典型少数民族自治县开展。该县为国家级贫困县,少数民族居民所占比例达46%以上,2016年该县贫困发生率达13.74%,远高于同期全国贫困发生率4.5%。该县基础教育薄弱,高考成绩不容乐观。如不考虑政策加分,2014-2016年该县高考一本硬上线分别仅1人、14人、15人,当年报名参加普通高考人数分别为508人、790人、499人。前期调研了解到,由于少数民族的一些传统文化习俗和宗教信仰,使得家长对子女教育的总体重视程度不高,尤其是女孩的教育很容易被少数民族家庭

漠视。

(二)实验组和参照组的选取

实验对象为小学四五年级学生。实验从2016年9月秋季学期正式开始,至2017年秋季学期期末考试后结束,共计3个学期。

项目所在县总共有29所小学,实验组和参照组选取的基本原则是在学校层面进行随机抽取。但由于其中有两所学校情况较为特殊,其中一所小学位于县城,是项目县规模最大和平均成绩最好的小学,和其余学校存在显著差异;另有一所小学同时存在中心校和覆盖整个小学阶段的村完小^①。在选取实验组和参照组时,我们对这两个学校进行了特殊处理。

我们首先将规模最大和平均成绩最好的这一所学校单独作为一组,在组内选取实验班班级和参照班级。具体操作方式为,将该校四五年级的班级编号为奇数的班级选为实验班,班级编号为偶数的班级选为参照班。在对县内规模最大和教学质量最高的学校单独选取实验班和参照班后,我们将同时存在中心校和村完小的小学视同两个学校和剩余的其他27所学校混在一起,随机抽取剩余的实验组学校和参照组学校各4所。实验组和参照组选取结果如表1所示。

(三)干预方案

对实验班班级,我们设计的激励方案包括3个现金奖励项目(详细奖励方案请见表2):学习优秀奖、学习进步奖、作业完成奖。其中,学习优秀奖是对班级内期末成绩排名前12%的学生进行奖励,奖励标准为150~300元。学习进步奖针对期末总成绩进步幅度排名前10%的学生,奖励金额为150元。作业优秀

表1实验组和参照组情况

实验组			参照组		
学校	班级个数(四五年级)	人数	学校	班级个数(四五年级)	人数
AB1	7	421	AB1	8	477
A2	10	405	B2	13	444
A3	6	243	B3	6	316
A4	6	201	B4	5	234
AB2	4	113	AB2	3	193
合计	33	1383	合计	35	1664

注:学校AB1的奇数班为实验组,偶数班为参照组,AB1的四年级1班为该校网络教育实验班,没有进入本实验。学校AB2下属的中心校为参照组,AB2下属的村小为实验组。

表2 激励方案设计

激励项目	奖励条件	奖励标准
学习优秀奖	本学期,学习成绩排名班级前4%(含)	300元/人/学期
	本学期,学习成绩排名班级前4%(不含)~8%(含)	200元/人/学期
	本学期,学习成绩排名班级前8%(不含)~12%(含)	150元/人/学期
学习进步奖	本学期,学习进步最大的前10%的同学	150元/人/学期
作业优秀奖	本学期,作业完成最好的前20%	50元/人/学期

注:上述各奖不重复领取。如果某一学生有资格获得一项以上,则给其奖励金额较高的一项。

奖针对平时作业完成得优秀比例排名前20%的学生,奖励金额为50元。

根据官方统计,实验县2016年农村居民人均可支配收入为9433元。奖励金比例相对于当地人均可支配收入比例并不高,50元、150元、200元和300元的奖励标准,相当于2016年当地农村居民人均可支配收入的0.53%、1.59%、2.12%和3.18%。但奖励标准对于低收入家庭收入的比例则相对较高,50元、150元、200元和300元的奖励标准,相当于收入最低1/4组家庭人均年收入的1.6%、4.9%、6.5%和9.7%。

这3个奖项不重复计算,在确定奖励名单时,依次按学习优秀奖、学习进步奖和作业优秀奖进行计算。这种激励方案下,成绩中下游水平学生即使获得成绩优秀奖的机会较小,也有较大机会获得学习进步奖和作业优秀奖。因此,理论上对成绩中下的学生有更强的激励作用,有助于鼓励他们平时认真完成作业,养成良好的学习习惯。

(四)实验实施过程

本实验于2016年秋季学期开始,2017年秋季学期后结束,实验期共计3个学期。具体进展如下。

基准调研。本项目于2016年9月正式实施。本文主要分析结果基于实验设计方案覆盖的完整样本,完整样本实验组和参照组覆盖班级数分别为33个和35个,学生人数分别为1383人和1664人。由于完整样本可获得的信息仅包括性别、民族和年龄,缺少学生父母受教育水平、家庭收入等家庭背景信息,为更准确的估计实验效果和进行异质性分析,我们对部分样本进行了问卷调查。最终问卷调查覆盖的实验组和参照组班级数分别为22个和17个,学生人数分别为609人和387人。在实验正式开始前,我们对实验组和参照组学校的校领导、班主任老师、学生

和学生家长进行抽样调查,获取部分学校、班级和学生基本信息以及学生家庭背景信息。并在实验开始前,采用斯坦福比奈(Stanford-Binet)国际标准智商60题,对部分实验组和参照组学生智商进行测试。

随机对照实验通常可通过计算在给定实验设计方案下的实验最低检测效果(Minimum detectable effect, MDE)来评估实验设计的合理性(Dufo et al., 2008; Bloom, 2006)。在给定显著性水平(α)、把握度(power, κ)、样本量(N)以及实验组样本分配比例(P)时,MDE衡量最低可检测到的实验效果。具体计算公式如下:

$$MDE=(t_{1-\kappa}+t_{\alpha})\times\sqrt{\frac{1}{p(1-p)}}\sqrt{\frac{\sigma^2}{N}}$$

由上式可知,最小化MDE的实验组和参照组最优分配比例为50%(Dufo et al., 2008; Bloom, 2006)。但当实验组比例在20%~80%之间时,实验组和参照组样本非平衡分配对MDE的影响非常小(Bloom, 2006)。本文完整样本实验组比例为45.3%;问卷调查样本实验组比例为61.1%,均在文献建议范围内。

假设对实验效果估计的把握度(κ)为0.8,估计系数的显著性水平为0.05时,经测算采用完整样本进行分析的最低检测效果为0.056个标准差,后文采用完整实验样本估计的系数0.077~0.099个标准差之间(表5),均大于最低检测效果;采用问卷调查样本进行分析的最低检测效果为0.11个标准差,后文采用问卷调查样本控制更多个人特征和家庭特征后的估计系数范围在0.121~0.141之间(表9),均大于最低检测效果。上述分析说明,本文实验设计的样本量和样本分配比例是合理的。

奖励方案告知。在学期初,以集中开会的方式,向实验组学生、教师及家长告知奖励规则。对参照组

学校的学生及老师则不告知任何奖励相关的信息。

奖励金发放。在学期中,老师对学生平时每次作业完成情况进行登记评价,每学期统计得“优”比例;在学期期末,对实验组和参照组学校组织统一的语文和数学测试。期末测试后,根据测试成绩和作业完成情况统计,按照奖励规则计算获奖名单,并在下一学期开学后,以表彰大会的形式对达到相应学习目标的学生公开发放奖励。第一次奖励金发放采取全县表彰大会的形式,后期奖励金发放均在学校层面自行组织,但须将奖励现场照片存档并发给项目组。实验的3个学期中,分别于2017年3月、2017年9月和2018年3月发放3次奖励金共167800元,其中497名学生获学习优秀奖,220名学生获学习进步奖(第一个学期没发学习进步奖),514学生获作业优秀奖。

跟踪调研。2017年6月进行回访,了解实验开展一学年后家庭对子女教育投入变化、关注程度变化和学生学习行为变化。

项目结束。本项目在2018年1月(即2017年的秋季学期)结束。实验结束后我们继续收集了3个学期的成绩情况。

四、数据及研究方法

我们使用的成绩数据为该县每学期全县统考成绩。我们共搜集了实验组和参照组学生从2016年春季学期~2019年春季学期的成绩,共7期的学习成绩情况。分别为:基期(2016年春季学期)、实验期3期(2016年秋季学期、2017年春季学期、2017年秋季学期)和实验结束后3期(2018年春季学期、2018年秋季学期、2019年春季学期)。为保证跨期的可比性,我们将参与项目的实验组和参照组成绩进行标准

化。性别、年龄信息从学籍信息中提取。学生姓名存在同名现象,我们将每个年级在全县范围内出现同名的学生予以剔除,数量为359人,其中4年级148人,5年级211人,占同年级全县学生的4.7%。同时,5年级学生在第5~6期已经进入初中学习,其语文、数学的原始总分均为150分,为保证与小学成绩可比,我们均将其原始分除以1.5,以调整为100分制,在后面做分析时,进一步将成绩得分标准化。

表3比较了实验组和参照组基本人口特征,在性别和年龄方面两组学生没有显著差异,只是实验组少数民族学生占比比参照组高7.7%。在学习成绩方面,实验组和参照组总成绩只相差0.018分且统计意义上不显著。分学科看,两组学生的语文和数学成绩的差距在统计意义上显著,但分差本身并不大。参照组语文平均成绩比实验组显著高2.9分,参照组的数学平均成绩比实验组显著低2.9分,均相当于平均水平的5%左右。

图1展示了实验组学生和参照组学生的成绩分布。两组学生的语文成绩分布存在一定差异,数学成绩分布则非常接近。表4的Kolmogorov-Smirnov检验结果也表明,两组学生成绩差异主要来自语文成绩。

图1中值得注意的一个现象是样本的成绩均呈双峰分布,而不是常见的正态分布(或者至少为单峰分布)。这一特殊现象的出现与我们的实验地点为民族自治县有关,汉语学生成绩偏好,少数民族学生成绩偏差,两者成绩差异较大,导致成绩出现双峰分布。少数民族学生的母语为彝语,但学校课本和课堂教学语言为汉语(部分老师为双语教学),汉语水平成为其学习最重要的壁垒之一。融合程度较高的县城

表3实验组和参照组人口特征和基期成绩比较

变量	参照组		实验组		参照组-实验组
	观测值	均值	观测值	均值	
女性	1624	0.485	1379	0.461	0.023
少数民族	1618	0.630	1363	0.707	-0.077***
年龄	1612	10.614	1375	10.614	0
总成绩	1659	107.917	1326	107.935	-0.018
语文	1659	55.284	1326	52.382	2.902***
数学	1659	52.633	1326	55.553	-2.920***

注:***p<0.01,**p<0.05,*p<0.1。

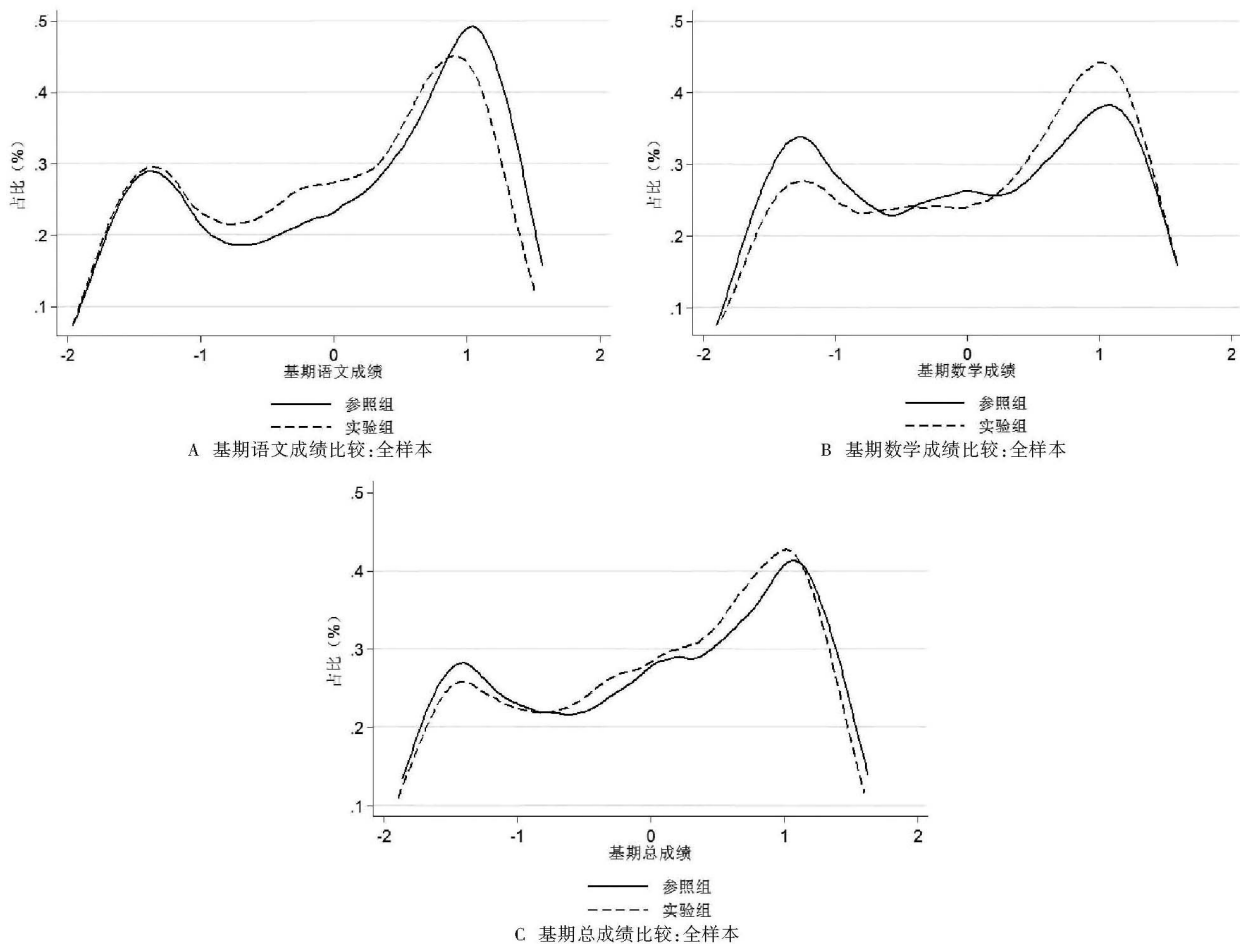


图1 基期成绩分布比较

注:图中的分数均已标准化处理,后面的成绩分布图均相同。

和乡镇少数民族家庭的学生汉语水平会好一些,其学习成绩也会相对较好,但融合程度低的山区少数民族家庭的学生汉语水平也更差,其学习成绩很容易从一年级就开始跟不上,从而导致后期成绩较差。

图2和图3展示了民族与成绩分布的关系。图2中,汉语学生的语文成绩单峰分布(图2A),且向右偏,语文成绩在整个样本中较好。少数民族学生的成绩出现一定程度的双峰分布(图2B),但更高的峰值在左边,成绩整体偏差。如果把少数民族和汉族学生(图2A和图2B)合并到一起,就出现了图1中的双峰分布。图2B中出现双峰分布也主要是因为少数民族学生内部因为融合程度差异较大导致学习成绩差异较大。由于AB1学校是县城学校,县城里民族融合程度较高,所以AB1学校的少数民族学生成绩也呈向右偏的单峰分布(图2C),普遍较好。但如

表4 全样本基期成绩差异分布检验:
Kolmogorov-Smirnov test

Group	D	P-value
语文		
参照组:	0.0015	0.996
实验组:	-0.1032	0
K-S:	0.1032	0
数学		
参照组:	0.0254	0.365
实验组:	-0.0188	0.575
K-S:	0.0254	0.695
总成绩		
参照组:	0.0003	1
实验组:	-0.0498	0.02
K-S:	0.0498	0.041

果把AB1学校排除(图2D),少数民族学生的成绩也呈左偏的单峰分布,普遍较差。如果把图2C和图2D

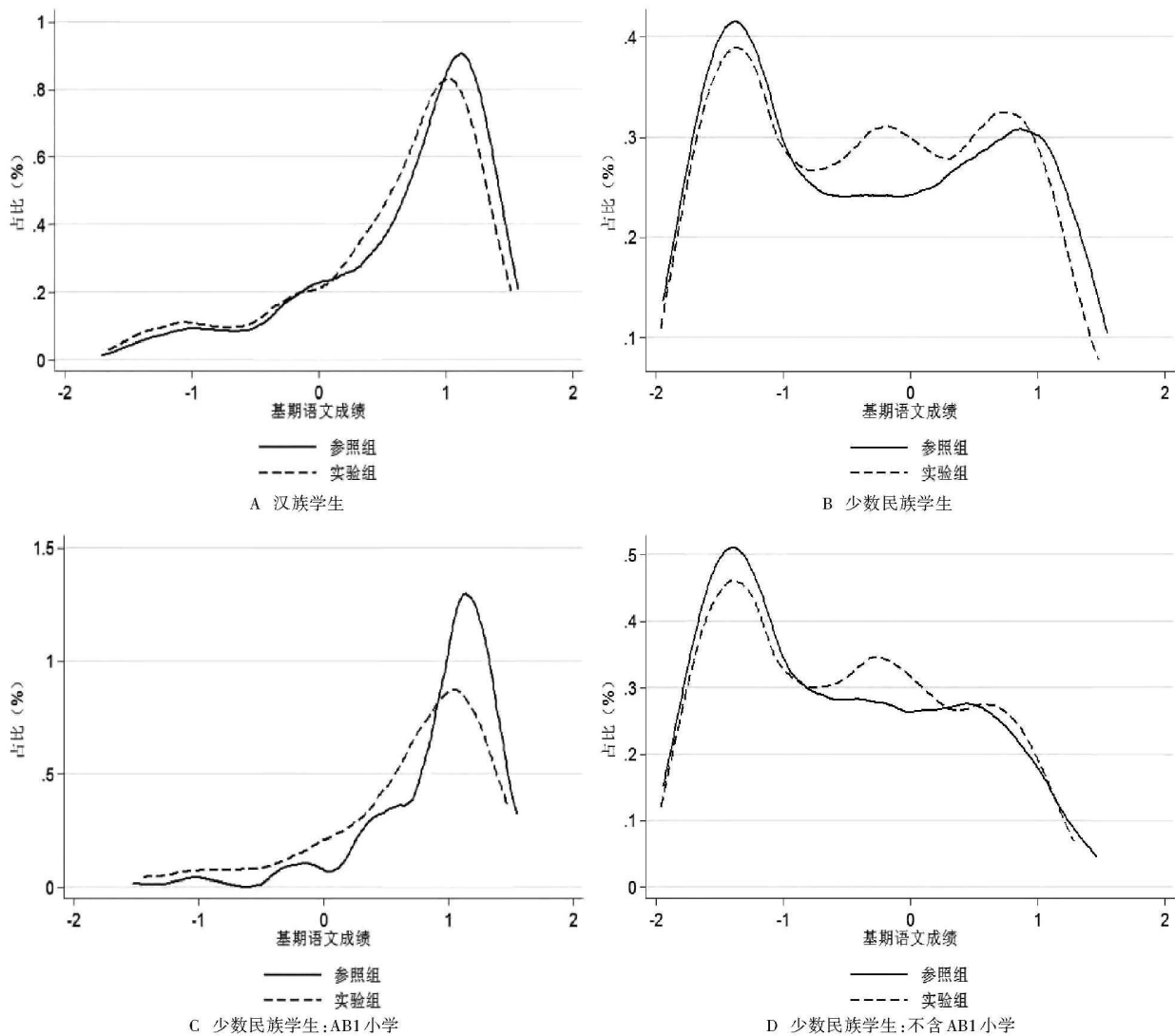


图2 对基期语文成绩分布的进一步说明

合并,则会出现图2B中的结果。图3中的数学成绩的分布特征与语文成绩类似。

图4报告了在实验第一期两组学生成绩变动情况(实验第一期成绩减基期成绩)。实验组学生在数学和语文成绩变动的分布方面,都比参照组学生要更偏右,尤其是语文成绩更加明显,表明实验组学生的进步更大。

图5进一步报告了获奖学生的基准期成绩分布和班内排名分布。获得优秀奖学生期初成绩分布右偏且较为集中,基准期成绩班内排名分布明显左偏且大部分分布在10名以内,表明优秀奖主要获得者主要是原本成绩就比较优秀的学生。作业奖与进步

奖获得学生的基准期成绩分布虽然也呈现右偏特征,但相对于优秀奖而言更加平坦,表明作业奖和进步奖能够覆盖期初成绩较差的群体。另外,作业奖与进步奖基准期班内成绩排名分布虽然也略有左偏,但明显右侧尾部较厚,表明作业奖和进步奖获得者有相当比例属于中下游学生。

我们借鉴 Todd 和 Wolpin(2003)、Mbiti 等(2019)等研究方法,基于学习生产函数的基本理论,主要采用以下形式的方程来估计激励方案对学习成绩的影响:

$$y_{i,t>0}=\alpha_0+\alpha_1T+\beta_y y_{i,t=0}+\alpha_2X_i+S_{i,t=0}+T_i+\varepsilon_{i,t>0} \quad (1)$$

其中, y 为标准化成绩, T 为 0~1 变量,1 表示实

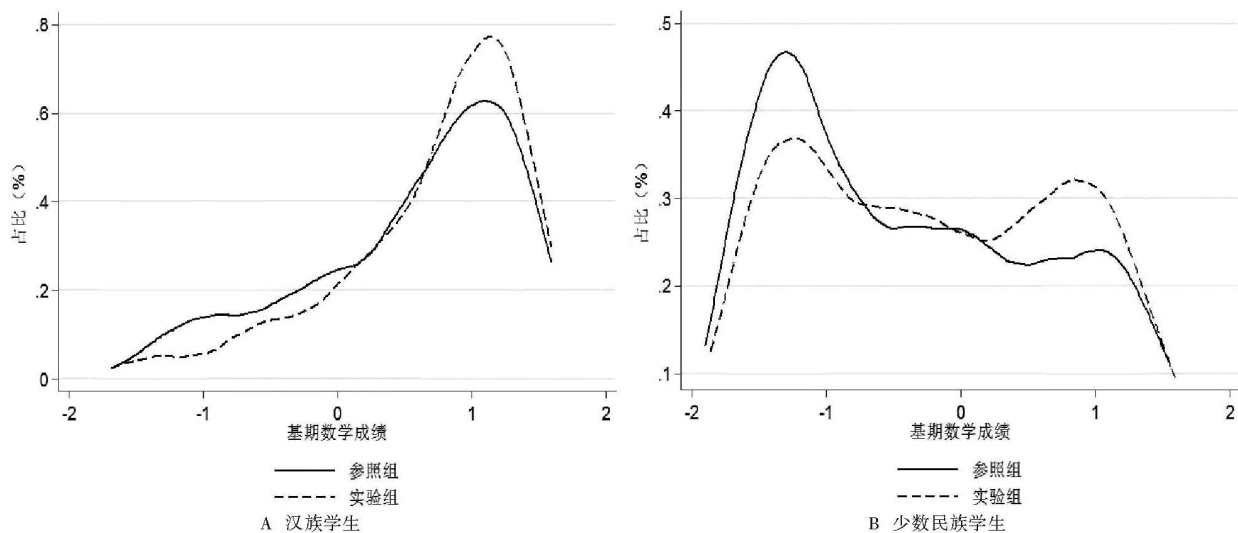
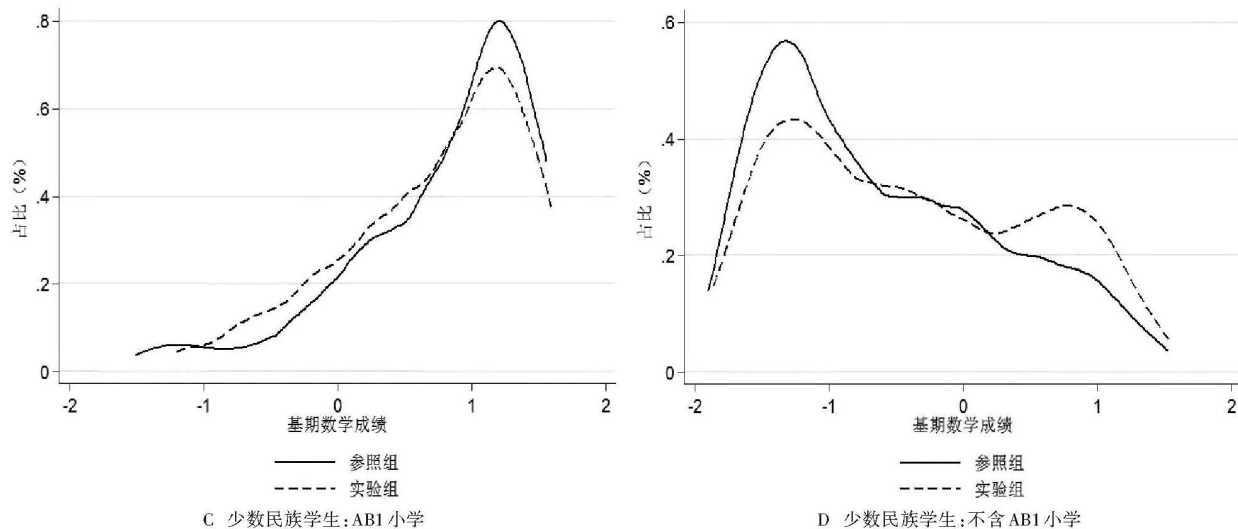


图3 基期数学成绩分布的进一步说明



续图3 基期数学成绩分布的进一步说明

验组,有现金激励,0表示参照组,没有现金激励。 $y_{i,t=0}$ 是基期标准化成绩。 X_i 为个人特征,包括性别、年龄、民族等基本个人信息, $S_{i,t=0}$ 为基期学生所在学校的分年级平均成绩,以控制学校质量, T 为学期固定效应。 ε 为残差项。下标中, i 是学生个人, t 为学期。估计系数的标准误都在“班级+学期”层面进行聚类调整。

这一估计方程的主要特点是控制了基期的学习成绩这一最重要的学习投入,基期学习成绩代表了该学生现有的知识水平,是后期学习最重要的学习基础。

五、估计结果:现金激励对学生学业表现的影响

(一)基准估计结果

实验进行了3个学期。表5报告了前三期的平均处理效应(ATE),表6报告了按基期成绩分组和动态时间上的激励效果的差异。

表5第(1)列中,实验组系数为0.099,且在1%的显著性水平上显著,表明相对于参照组而言,实验组在给予奖励后,总成绩增加了0.099个标准差。分科目来看,相对于参照组,实验组的语文成绩和数学成绩分别上升了0.077个标准差和0.093个标准差,现金激励对数学成绩的正面影响更加明显。

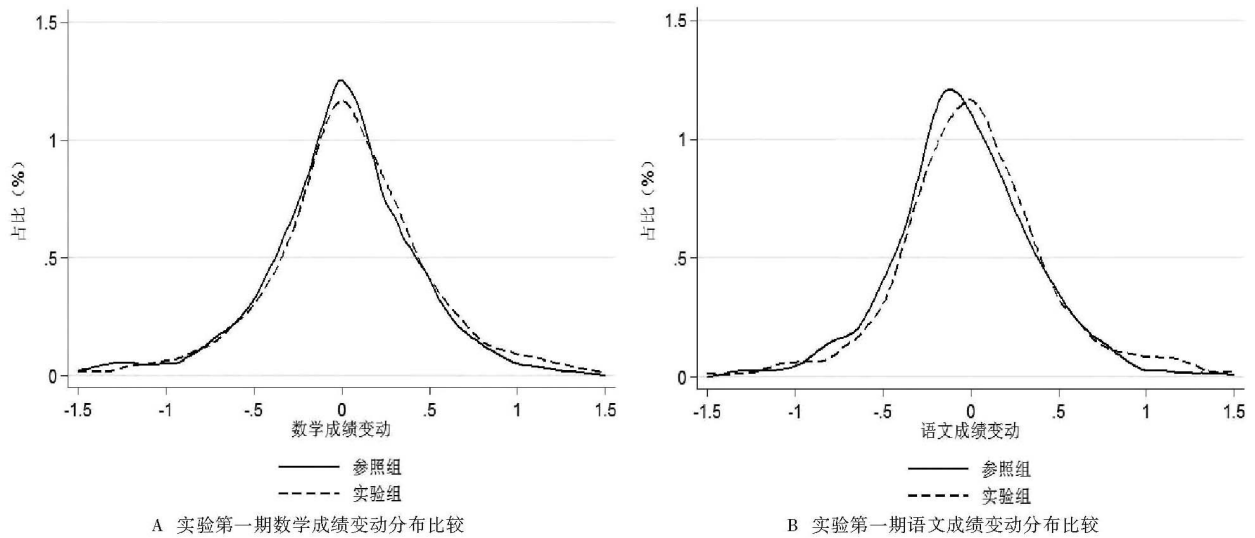


图4 实验第一期成绩变动分布

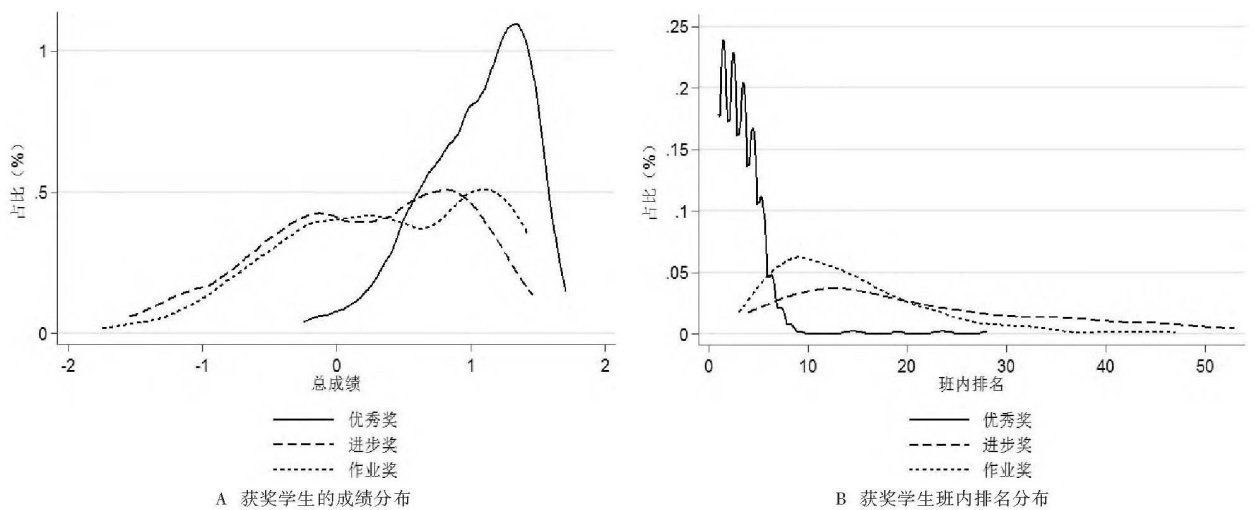


图5 获奖学生成绩及排名分布

在控制变量方面,基期成绩与实验期成绩相关性非常明显,语文成绩的相关性要强于数学成绩,即语文成绩更依赖前期的学习基础,这也很可能使得数学成绩受学生当期努力的影响相对会大些。年龄的系数为负,可能主要是当地部分偏远地区学生入学年龄较大,这部分学生因家庭背景、社会文化等诸多因素导致学习条件和学习基础均有较大欠缺。女生的语文比男生要高0.140个标准差,这也与目前的文献整体相符,但数学成绩在性别之间没有显著差异。少数民族学生的语文成绩系数显著为正,这一点颇为令人惊讶,表明如果给定期初成绩、性别和学

校的整体质量,少数民族学生的学习表现并不一定会更差。

该实验奖励方案的最大特点是设立了学习进步奖和作业优秀奖,把奖励范围扩大到初始成绩处于中下水平的学生,让原先那些很难获得老师关注的成績处于中下水平的学生也有机会获得奖励。这种奖励方式是否能有效激励中差生成绩增长是我们关注的重点。我们把学生按基期在班内成绩从高到低排名分为5组:前20%,中上20%,中等20%,中下20%和后20%。表6的第(1)~(3)列报告了按成绩分组的回归结果。

变量	实验的平均处理效应		
	总成绩	语文	数学
	(1)	(2)	(3)
实验组	0.099*** (0.026)	0.077*** (0.026)	0.093*** (0.033)
基期总成绩	0.880*** (0.009)		
基期语文		0.853*** (0.011)	
基期数学			0.801*** (0.012)
年龄	-0.018*** (0.006)	-0.012** (0.006)	-0.021*** (0.008)
女性	0.058*** (0.010)	0.140*** (0.012)	0.015 (0.011)
少数民族	0.033** (0.016)	0.050*** (0.017)	-0.010 (0.020)
学校平均成绩	0.126*** (0.034)	0.151*** (0.035)	0.242*** (0.047)
常数项	-0.483*** (0.169)	-0.698*** (0.180)	-0.920*** (0.225)
学期	Y	Y	Y
N	8339	8339	8339
adj.R ²	0.847	0.813	0.774

注:括号内为在“班级+学期”层面聚类调整的标准误。
***p<0.01,**p<0.05,*p<0.1。

变量	平均处理效应:按基期班内成绩分组和动态效应					
	总成绩	语文	数学	总成绩	语文	数学
	(1)	(2)	(3)	(4)	(5)	(6)
实验组	0.089*** (0.031)	0.032 (0.035)	0.103** (0.042)			
实验组×最高20%	-0.005 (0.037)	0.004 (0.044)	0.001 (0.050)			
实验组×中上20%	-0.038 (0.032)	-0.001 (0.039)	-0.053 (0.046)			
实验组×中等20%	0.044 (0.028)	0.067* (0.034)	0.025 (0.040)			
实验组×中下20%	0.043* (0.024)	0.102*** (0.029)	-0.009 (0.031)			
实验组×第1期				0.058 (0.041)	0.046 (0.038)	0.056 (0.054)
实验组×第2期				0.122** (0.050)	0.097** (0.049)	0.116* (0.059)
实验组×第3期				0.119*** (0.045)	0.091** (0.045)	0.107* (0.061)
基期成绩	Y	Y	Y	Y	Y	Y
控制变量	Y	Y	Y	Y	Y	Y
学期	Y	Y	Y	Y	Y	Y
N	8339	8339	8339	8339	8339	8339
adj.R ²	0.847	0.817	0.776	0.847	0.813	0.774

注:括号内为在“班级+学期”层面聚类调整的标准误。***p<0.01,**p<0.05,*p<0.1。常数项略去。控制变量包括年龄、性别、民族、学校平均成绩。

表6前3列的结果整体而言,在有奖励后,期初不同层次的学生成绩都有所提高,但初期成绩处于中下水平的学生进步会更明显。表6中,由于把基期成绩前四组分别与实验组做了交叉项,“实验组”系数反映的是基期成绩处于最低的20%那一组的处理效应。第(1)列中,“实验组”系数为0.089,且在1%显著性水平上显著。“实验组×中下20%”系数为0.043,在10%显著性水平上显著,“实验组×前20%”和“实验组×中上20%”系数符号为负,但均不显著,期初成绩处于班内中间20%与实验组的交叉项系数不显著。表明,基期成绩处于班内最低20%的学生在有奖励后,成绩上升了0.089个标准差,且基期成绩处于不同水平的学生成绩都有不同程度的提升,但基期成绩处于班内中下20%的学生进步最快,成绩上升了0.132个标准差(=0.089+0.043)。语文成绩方面,期初成绩排中等20%和中下20%学生进步最快,数学成绩则各组进步情况没有显著差异。

本实验进行了3个学期,使得我们有机会观察现金激励的动态效应。表6中第(4)~(6)列报告了实验的动态效果。在(4)~(6)列中,“实验组×第1期”系数

均不显著,这很有可能与实验刚开始,一些家长和学生没有关注到项目的开展,或者对实验项目执行的真实性仍然存在疑虑,激励效应没有发生。第(4)列中,“实验组×第2期”和“实验组×第3期”的系数均在1%或5%的显著性水平上显著为正,系数分别为0.122和0.119,这表明在第一轮奖励发放后,激励效应非常明显。数学成绩和语文成绩的激励效应也均类似,在第一期没有显现,在第二期和第三期变得非常明显。

由于不同的现金激励实验的激励方案差别较大,不同研究之间很难直接比较。但从激励效果本身来看,本实验的激励效果与Kremer等(2009)的0.2个标准差、Bettinger(2012)的0.15个标准差、Levitt等(2016b)的0.068~0.153个标准差等比较接近,并且对差生的影响为0.141~0.226个标准差,也与Li等(2014)的0.265个标准差比较接近。本实验发现数学成绩的激励效果比语文成绩要更大,这与Bettinger(2012)和Levitt等(2016b)的发现也一致。

另外,本研究的动态分析也表明,考虑到学生和家长对于干预项目的理解和关注可能需要一个过程,只做一期干预很可能难以产生预期效果。

(二)稳健性检验
1.控制更多个人特征和家庭特征

表5~表6中的估计方程没有控制更多的个人特征和家庭特征,可能会存在遗漏变量问题。实验开展时,随机抽取了8所学校996个学生进行问卷调

查,了解学生的IQ、学习行为、学习态度、家庭人口和经济特征、父母对小孩教育的时间和资金投入情况等信息。图6展示了调查样本与非调查样本基期的成绩分布。调查样本相对来说成绩好的学生占比更高。Kolmogorov-Smirnov检验结果也表明调查样本与非调查样本之间存在显著的差异(表7)。

表8比较了调查样本中实验组和参照组的人口特征和家庭特征,整体而言,调查样本中实验组学生的个人能力和家庭条件要比参照组好一些。主要是因为调查样本中,只有实验组中有AB1小学的学生样本,但参照组中没有。具体可见图7。图7A中调查样本中汉族学生成绩较好,且呈单峰分布,但图7B中少数民族学生成绩呈双峰分布,两图合并就会使得全调查样本出现双峰分布。图7C中可以看到,只

表7 调查样本与非调查样本基期成绩差异分布检验:
Kolmogorov-Smirnov test变量

Group	D	P-value
语文		
参照组:	0.1038	0
实验组:	-0.0057	0.957
K-S:	0.1038	0
数学		
参照组:	0.1523	0
实验组:	-0.0021	0.994
K-S:	0.1523	0
总成绩		
参照组:	0.1647	0
实验组:	-0.0072	0.93
K-S:	0.1647	0

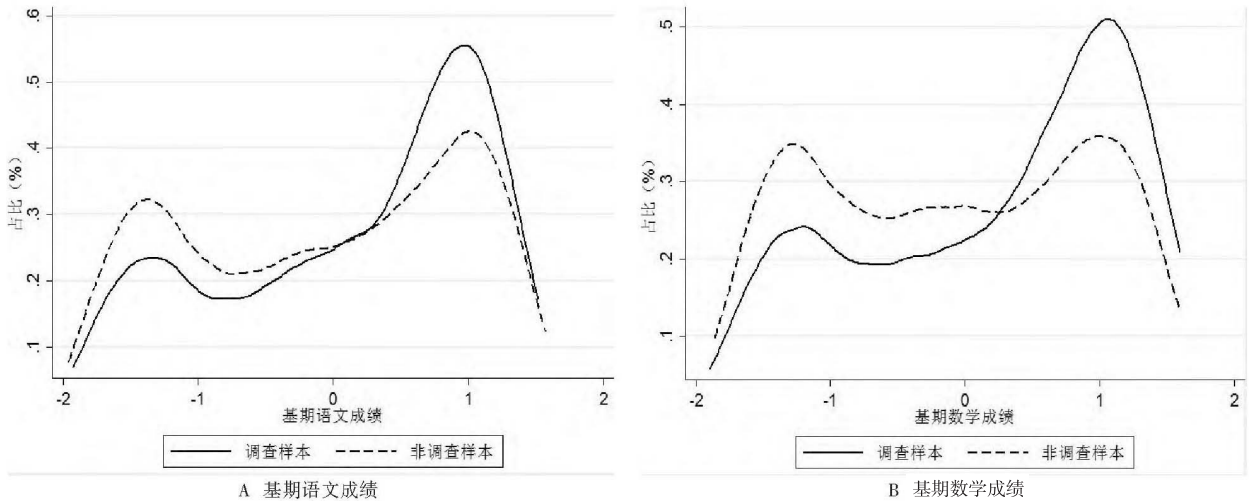


图6 调查样本与非调查样本基准期成绩分布比较

部分报告了基本回归结果,第2部分报告了按基期成绩分组的回归结果,第3部分报告的是动态效应估计结果。

在控制更多学生个人特征和家庭特征后,第1部分的实验组系数仍然显著,且比表5的系数更大。第2部分中,“实验组”系数显著为正,4个成绩分组与实验组交叉项的系数随着基期成绩升高而降低且逐渐显著为负,即基期成绩越差的学生成绩上升越快。第3部分中,第1期系数均不显著,但在第2期和第3期,语文成绩分别显著上升了0.210个标准差和0.155个标准差,数学成绩在第3期显著增长了0.202个标准差,使得激励效应随实验时间推延越来越大。整体而言,表9的回归结果与表5、表6结果比较一致。

2. 不含成绩最好的学校

该县内部各学校之间成绩差异较大,该县最大的小学(学校AB1)整体学生成绩也最好,实验前基期

成绩总分平均分达157分(语文+数学),远超过其它学校学生的平均分90分(语文+数学)。同时,该校由于在城区,该校学生的家庭经济状况和父母受教育程度均较高,对孩子的学习也更关心,教育质量更高,这可能会削弱该校学生与其它学校的学生的可比性,现金激励效应也可能存在差异。我们把这个学校的样本剔除,进一步观察了现金激励的效果(请见表10),整体与表5~表6的结果比较接近。

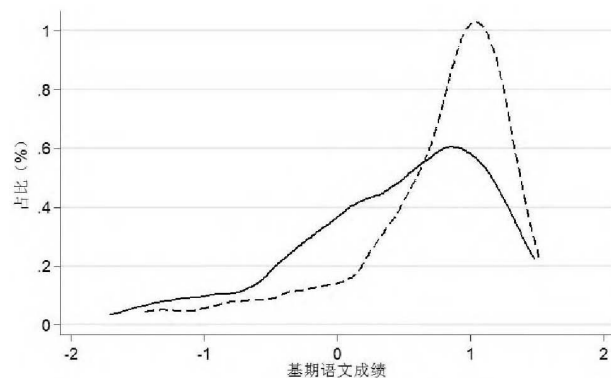
3. 不含学校内部同时存在实验组和参照组的学校

在抽样时,AB1和AB2同时有实验组和参照组,这有可能导致学校内部学生和老师之间存在外溢效应,从而影响激励效果。我们把这个学校排除以考察激励效果,估计结果见表11第一部分。估计结果显示,估计系数仍显著且高于表5和表6估计结果,当不存在奖励信息外溢效应时,实际激励效果可能更大。

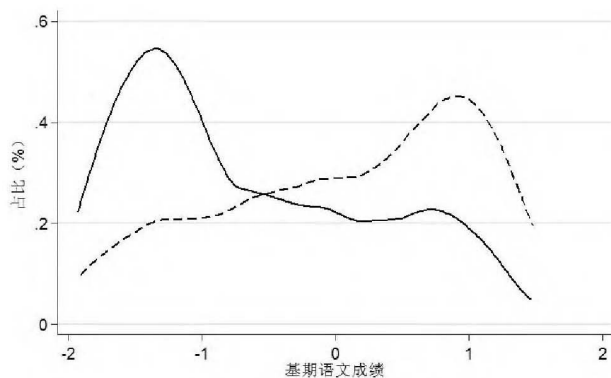
表 10 稳健性检验二:不含成绩最好的学校

10组别和变量	全样本		
	总成绩	语文	数学
	(1)	(2)	(3)
不分组			
实验组	0.144*** (0.031)	0.116*** (0.031)	0.141*** (0.038)
按基期班内成绩分组			
实验组	0.114*** (0.032)	0.072* (0.040)	0.116*** (0.035)
实验组×前 20%	0.015 (0.043)	-0.018 (0.053)	0.061 (0.054)
实验组×中上 20%	-0.022 (0.040)	-0.019 (0.049)	-0.003 (0.056)
实验组×中等 20%	0.078** (0.036)	0.055 (0.040)	0.100** (0.048)
实验组×中下 20%	0.075*** (0.029)	0.125*** (0.035)	0.033 (0.033)
动态效应			
实验组×第 1 期	0.093* (0.048)	0.070 (0.045)	0.099 (0.061)
实验组×第 2 期	0.185*** (0.056)	0.154*** (0.055)	0.179*** (0.065)
实验组×第 3 期	0.157*** (0.053)	0.125** (0.057)	0.146** (0.066)
控制变量	Y	Y	Y
学期	Y	Y	Y

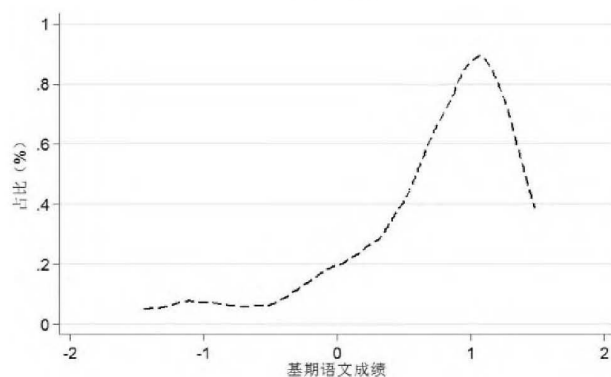
注:括号内为在“班级+学期”层面聚类调整的标准误。***p<0.01,**p<0.05,*p<0.1。常数项略去。控制变量包括年龄、性别、民族、个人IQ、父母受教育程度、家庭收入、学校平均成绩等。



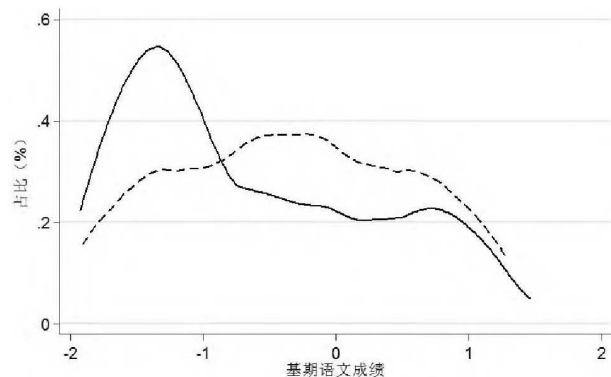
A 调查样本汉族学生



B 调查样本少数民族学生



C 调查样本少数民族学生:AB1小学



D 调查样本少数民族学生:不含AB1小学

图7 调查样本的进一步说明

4. 不含改变组别的学生

有10个学生在实验期间,因为家庭住址变动原因变换了就读学校,导致在实验组和参照组之间进行过调整,我们把这些学生删除,发现估计结果与表5~表6比较接近,请见表11第二部分。

5. 不含大班和小班

各学校班级规模差异较大,最大的班接近70人,最小的班仅为10人,班级规模本身会影响到教学质量(Angrist and Lavy, 2009),并且也可能影响到班内的学习竞争程度。我们把少于15人的班认定为小班(6个),多于60人的班认为定为大班(20个班)。把大班和小班样本剔除后,我们进一步估计了激励效应,发现估计结果与表5~表6都比较接近,请见表11第三部分。

(三)激励效应的异质性分析

前面的分析都只着眼于平均效应,在人力资本

和反贫困的研究中,性别平等、民族、家庭社会经济条件都是非常重要的议题。我们实验的样本点处于国家级贫困县,该地少数民族家庭的收入整体低于汉族家庭,贫困发生率较高,人力资本是切断贫困代际传递最有效的手段。因此,我们进一步从性别、民族、父母受教育程度和家庭收入的角度考察了现金激励效应的差异。父母低教育是指当父母的学历都低于初中则为1,否则为0。

表12分别考察了性别、民族、父母受教育程度和家庭收入对激励效果的影响。表中第一部分中的“实验组×女性”“实验组×少数民族”和第二部分中“实验组×父母低教育”的系数均显著为正,表明女生、少数民族和低受教育程度家庭的学生激励效果更大。

另外,第二部分中,“实验组×收入最低25%”的语文成绩和总成绩系数均不显著,只有数学成绩系

表 11 稳健性检验三、四、五：内部分组、改变组别及大小班

组别和变量	全样本		
	总成绩	语文	数学
	(1)	(2)	(3)
不含学校内部分组的学校			
实验组	0.134*** (0.033)	0.136*** (0.032)	0.106*** (0.039)
不含改变组别的学生			
实验组	0.099*** (0.027)	0.077*** (0.026)	0.093*** (0.034)
不含大班和小班			
实验组	0.119*** (0.029)	0.109*** (0.029)	0.104*** (0.035)
基期成绩	Y	Y	Y
控制变量	Y	Y	Y
学期	Y	Y	Y

注：括号内为在“班级+学期”层面聚类调整的标准误。***p<0.01，**p<0.05，*p<0.1。常数项略去。控制变量包括年龄、性别、民族、个人IQ、父母受教育程度、家庭收入、学校平均成绩等。

表 12 激励效应的差异：性别、民族、父母受教育程度与家庭收入

变量	总成绩	语文	数学	总成绩	语文	数学
	(1)	(2)	(3)	(4)	(5)	(6)
Panel 1：性别与民族						
实验组	0.094*** (0.030)	0.083*** (0.030)	0.078** (0.037)	0.070* (0.037)	0.055 (0.035)	0.060 (0.045)
实验组×女性	0.158*** (0.029)	0.217*** (0.029)	0.109*** (0.036)			
实验组×少数民族				0.128*** (0.034)	0.124*** (0.032)	0.079* (0.041)
Panel 2：父母受教育程度与家庭收入						
实验组	0.061 (0.038)	0.187*** (0.052)	-0.064 (0.042)			
实验组×父母低教育	0.092** (0.040)	-0.054 (0.047)	0.228*** (0.047)			
实验组×收入最低 25%				0.097 (0.061)	0.038 (0.055)	0.150** (0.067)
实验组×收入次低 25%				0.190*** (0.046)	0.164*** (0.047)	0.200*** (0.056)
实验组×收入次高 25%				0.189*** (0.053)	0.207*** (0.051)	0.158** (0.065)
实验组×收入最高 25%				0.074** (0.035)	0.120** (0.048)	0.017 (0.054)
基期成绩	Y	Y	Y	Y	Y	Y
控制变量	Y	Y	Y	Y	Y	Y
学期	Y	Y	Y	Y	Y	Y

注：括号内为在“班级+学期”层面聚类调整的标准误。***p<0.01，**p<0.05，*p<0.1。常数项略去。控制变量包括年龄、性别、民族、学校平均成绩及各交叉项变量。

数显著，但“实验组×收入次低 25%”和“实验组×收入次高 25%”的系数最大且均在 1% 显著性水平上显著，表明中间收入阶层的激励效应最大，但该地区为国家级贫困县，收入水平远低于全国平均水平。这

语文成绩和数学成绩均是基期成绩处于中等 20% 和中下 20% 组的学生激励效果明显,而五所差学校中,基期不同成绩水平的各组学生均有明显激励效果,并且基期成绩最差的两组学生成绩进步最为明显。

表 15 报告了分位数回归结果,结果显示实验组成绩各分位数均有显著提升,且对低分位数提升较大。现金激励使实验组学生总成绩的 25、50 和 75 分位分别显著提升了 0.098、0.085 和 0.077 个标准差。分学科来看,语文成绩的分位数变化与总成绩类似,

表 15 激励效应差异:分位数回归				
变量	0.25	0.5	0.75	1
	(1)	(2)	(3)	(4)
总成绩				
实验组	0.098*** (0.010)	0.085*** (0.009)	0.077*** (0.011)	0.057 (0.067)
语文				
实验组	0.085*** (0.012)	0.071*** (0.010)	0.058*** (0.012)	-0.017 (0.072)
数学				
实验组	0.075*** (0.012)	0.065*** (0.010)	0.090*** (0.014)	0.049 (0.044)
基期成绩	Y	Y	Y	Y
控制变量	Y	Y	Y	Y
学期	Y	Y	Y	Y
N	8339	8339	8339	8339

注:***p<0.01,**p<0.05,*p<0.1。常数项略去。控制变量包括年龄、性别、民族、学校平均成绩。

表 16 家庭家长对教育的关心和投入

变量	生活关心程度		对学习关心程度		陪伴时间		学习指导时间		辅导费支出	书本费支出
	父亲	母亲	父亲	母亲	父亲	母亲	父亲	母亲		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
实验组	-0.128 (0.166)	0.982** (0.154)	-0.125 (0.136)	-0.882** (0.295)	0.447** (0.076)	0.283** (0.085)	0.856** (0.050)	0.123 (0.073)	1.753** (0.126)	0.137* (0.069)
基期	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
控制变量	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
班级	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
观测值	737	738	738	735	747	748	717	714	742	726
adj.R ²					0.240	0.315	0.221	0.271	0.207	0.697
Pseudo R ²	0.0570	0.110	0.0693	0.110						
对数似然率	-780.2	-736.7	-771.5	-755.5						

注:括号内为在班级层面聚类调整的标准误。***p<0.01,**p<0.05,*p<0.1。常数项略去。控制变量包括年龄、性别、民族、家庭收入、是否住校、父母的教育背景。第(1)~(4)列为 Ordered Logit 模型估计结果,且因变量的选项均为 5 级量化:1(非常关心)、2(比较关心)、3(一般)、4(不太关心)、5(非常不关心)。第(5)~(9)列均为 OLS 估计结果,因变量均已对数化。

低分位提升较大。但数学成绩提升略有不同,数学成绩的 75 分位数提升效果最大,提升了 0.09 个标准差。总的来看,本激励性干预方案对整体较差的学校激励效果更为明显,成绩提升主要体现在提升了中下、中等学生成绩。

(五)机制分析:家长行为和学生行为

前面分析表明,现金激励对学生成绩产生了显著的积极影响,并且基期成绩越差的学生成绩上升越快,随着奖励时间延长,激励效应也更加明显。但现金激励是怎么影响到学生成绩的?一种可能是现金激励使得家长对小孩学习的督促和教育投入增加了,另一种可能是现金激励使得学生自己更加努力学习。本节我们从家长行为和学生行为两个角度来探讨现金激励的影响机制。

表 16 报告了现金奖励对家长在孩子教育方面的时间和资金投入的影响。父亲对孩子的生活和学习关心程度均没有受到显著影响,但母亲对孩子的关心出现了有趣的变化。在有现金奖励后,母亲对孩子的生活关心程度显著下降了,但对孩子学习的关心程度出现了显著上升。另外,父母对孩子的陪伴和学习指导时间都有不同程度的上升。家庭的辅导费和书本费等教育投资也都显著增加。那么现金奖励对父母行为的影响是否在少数民族群体更加突出呢?分析发现,现金奖励对少数民族家庭父亲和母亲对小孩的陪伴时间和学习指导时间的增加方面与

汉族家庭并无显著差异。但在少数民族家庭中,母亲对女孩的学习指导时间的增加显著高于男孩(囿于篇幅限制,本部分结果见《管理世界》网络发行版附录附表1)。

表17进一步报告了按基期成绩分组时家长在孩子教育投入方面的差异。第(1)~(2)列中,不同成绩组父亲的系数都不显著,不同成绩组母亲对小孩的生活关心都有下降。在第(3)~(4)列中,只有中下20%和中上20%成绩组学生的父亲对小孩的学习关心程度显著上升,但在母亲方面,所有成绩组的母亲对学习关心程度都显著上升,且中下20%和中上20%成绩组上升最快。在陪伴时间方面,各组有不同程度上升,但在学习指导时间方面,各组学生的父

表 17 家长对教育的关心和投入:按基期班内成绩分组

变量	生活关心程度		对学习关心程度		陪伴时间		学习指导时间		辅导费支出	书本费支出
	父亲	母亲	父亲	母亲	父亲	母亲	父亲	母亲		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
实验组	-0.221 (0.250)	1.148*** (0.261)	0.253 (0.226)	-0.635* (0.352)	0.385** (0.158)	0.161 (0.118)	0.737*** (0.110)	0.059 (0.102)	1.541*** (0.160)	0.182* (0.100)
实验组×前20%	0.186 (0.354)	-0.197 (0.354)	-0.415 (0.287)	-0.506* (0.301)	0.137 (0.208)	0.268* (0.132)	0.242 (0.154)	0.169 (0.151)	0.001 (0.188)	-0.053 (0.116)
实验组×中上20%	-0.045 (0.313)	-0.659* (0.389)	-0.632*** (0.235)	-0.674** (0.301)	0.255 (0.265)	0.413** (0.152)	0.204 (0.138)	0.025 (0.123)	0.551 (0.353)	0.025 (0.093)
实验组×中等20%	0.090 (0.241)	-0.102 (0.316)	-0.526 (0.370)	-0.308 (0.365)	-0.015 (0.197)	0.131 (0.145)	0.118 (0.153)	0.085 (0.108)	0.547* (0.309)	-0.085 (0.116)
实验组×中下20%	0.128 (0.235)	-0.357 (0.243)	-0.670*** (0.207)	-0.432* (0.225)	0.235 (0.155)	0.128 (0.127)	0.185 (0.154)	0.096 (0.121)	0.452* (0.266)	-0.116 (0.126)
基期	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
控制变量	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
班级	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
观测值	737	738	738	735	747	748	717	714	742	726
adj. R ²					0.245	0.322	0.225	0.273	0.216	0.698
Pseudo R ²	0.0575	0.113	0.0731	0.113						
对数似然率	-779.8	-734.1	-768.4	-753.1						

表 18 学生的学习态度和学习投入

变量	认真听讲	积极回答问题	与班主任交流	与任课老师交流	周末与同学交流	工作日作业时间	工作日复习预习时间	工作日辅导班时间	周末作业时间	周末复习预习时间	周末辅导班时间
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
实验组	2.665*** (0.238)	2.268*** (0.161)	1.310*** (0.284)	0.392* (0.211)	1.028*** (0.222)	0.968*** (0.061)	1.656*** (0.083)	2.715*** (0.101)	1.779*** (0.067)	1.732*** (0.075)	2.845*** (0.109)
基期	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
控制变量	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
班级	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

续表 18

变量	认真听讲	积极回答问题	与班主任交流	与任课老师交流	周末与同学交流	工作日作业时间	工作日复习预习时间	工作日辅导班时间	周末作业时间	周末复习预习时间	周末辅导班时间
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
观测值	695	685	692	680	661	682	693	695	700	693	694
adj.R ²						0.255	0.170	0.318	0.229	0.160	0.315
Pse.R ²	0.103	0.134	0.0996	0.0671	0.0670						
对数似然率	-722.8	-767.5	-779.8	-825.8	-832.8						

注:括号内为在班级层面聚类调整的标准误。***p<0.01,**p<0.05,*p<0.1。常数项略去。控制变量包括年龄、性别、民族、家庭收入、是否住校、父母的教育背景。第(1)~(5)列为Ordered Logit模型估计结果,因变量的选项均为五级量化:1(非常不同意)、2(不太同意)、3(无所谓)、4(比较同意)、5(非常同意)。第(6)~(11)列因变量均为对数化后的值。

表 19 学生的学习态度和学习投入:按基期班级内成绩分组

变量	认真听讲	积极回答问题	与班主任交流	与任课老师交流	周末交流	工作日作业时间	工作日复习预习时间	工作日辅导班时间	周末作业时间	周末复习预习时间	周末辅导班时间
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
实验组	2.554*** (0.293)	1.147*** (0.357)	0.165 (0.292)	2.110*** (0.288)	1.328*** (0.408)	0.783*** (0.077)	1.574*** (0.119)	2.748*** (0.167)	1.589*** (0.117)	1.674*** (0.108)	2.823*** (0.154)
实验组×前20%	0.830** (0.414)	0.818*** (0.294)	0.565* (0.297)	0.541 (0.444)	-0.883* (0.459)	0.286*** (0.083)	0.107 (0.112)	-0.064 (0.202)	0.381** (0.143)	0.138 (0.146)	0.032 (0.180)
实验组×中上20%	0.530* (0.315)	0.550** (0.279)	0.117 (0.254)	0.336 (0.354)	-0.258 (0.276)	0.376*** (0.074)	0.142 (0.105)	0.028 (0.221)	0.458*** (0.137)	0.074 (0.125)	0.207 (0.155)
实验组×中等20%	0.289 (0.321)	0.496 (0.422)	0.283 (0.319)	0.331 (0.284)	-0.305 (0.269)	0.249*** (0.089)	0.068 (0.096)	-0.082 (0.235)	0.406*** (0.139)	0.186 (0.113)	0.051 (0.138)
实验组×中下20%	0.168 (0.297)	0.195 (0.235)	0.386 (0.394)	0.090 (0.348)	-0.237 (0.389)	0.198** (0.089)	0.139 (0.113)	-0.019 (0.212)	0.087 (0.144)	-0.031 (0.107)	0.005 (0.083)
基期	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
控制变量	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
班级	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
观测值	695	692	680	685	661	682	693	695	700	693	694
adj.R ²						0.269	0.172	0.318	0.251	0.164	0.316
Pse.R ²	0.107	0.104	0.0696	0.136	0.0723						
对数似然率	-719	-776.2	-823.6	-765.9	-828.1						

注:括号内为在班级层面聚类调整的标准误。***p<0.01,**p<0.05,*p<0.1。常数项略去。控制变量包括年龄、性别、民族、家庭收入、是否住校、父母的教育背景。第(1)~(5)列为Ordered Logit模型估计结果,因变量的选项均为五级量化:1(非常不同意)、2(不太同意)、3(无所谓)、4(比较同意)、5(非常同意)。第(6)~(11)列因变量均为对数化后的值。

影响分析来看,基期成绩处于中下水平的学生进步最快,很有可能是学习投入对学习成绩的影响也存在边际效应递减,成绩越差的学生的学习投入的边际效应越大,所以虽然基期成绩较差的学生学习投入不如学习成绩好的学生,但学习成绩进步却更明显。激励方案除对学生和家长产生了影响外,还可能对教师行为也产生了一定影响。在调研走访中,我们发现实验组教师对学生作业的批改相比以往更加认真。这主要是由于作业奖要求任课教师对学生

作业完成情况进行统计,这可能导致任课教师在对 学生作业完成情况的监督更认真。对于这一潜在影响机制,由于我们没有收集教师行为相关数据,暂时无法进行检验。

六、激励效应在实验结束后是否持续?

文献中争议非常大的一个问题是,现金激励是 否会有长期持续性影响?这涉及外部激励与内部动力是否存在替代问题以及短期的激励是否有助于改 善学习习惯和态度问题。有文献发现现金激励有持

续效应(Bettinger, 2012; Leuven et al., 2010),但也有文献发现现金激励取消后激励效应消失了(Levitt et al., 2016a; List et al., 2018; Visaria et al., 2016)。本轮实验进行了3个学期,但我们同时跟踪收集了结束后3个学期的数据,以考察现金奖励停止后学生的学业表现,表20报告了这一估计结果。

表20的估计结果显示,实验结束后对实验组语文成绩的激励效应仍然持续,数学成绩的激励效应则没有持续。实验结束后的3个学期,实验组平均总成绩比参照组高0.057个标准差,语文平均成绩比参照组显著高0.104个标准差,数学平均成绩则没有显著差异。按基期成绩分组的估计结果表明,语文成绩的持续效应在基期成绩中下20%组的学生最

表 20 实验结束后三期效应			
组别和变量	全样本		
	总成绩	语文	数学
	(1)	(2)	(3)
不分组			
实验组	0.057** (0.029)	0.104*** (0.026)	-0.007 (0.040)
按基期班内成绩分组			
实验组	0.073* (0.040)	0.086* (0.046)	0.022 (0.052)
实验组×前20%	-0.035 (0.046)	-0.058 (0.051)	0.001 (0.055)
实验组×中上20%	-0.075* (0.040)	-0.051 (0.047)	-0.070 (0.052)
实验组×中等20%	-0.054 (0.040)	-0.025 (0.046)	-0.056 (0.049)
实验组×中下20%	0.046 (0.037)	0.090** (0.044)	0.018 (0.043)
动态效应			
实验组×第4期	0.090* (0.054)	0.101** (0.048)	0.053 (0.075)
实验组×第5期	0.061 (0.050)	0.116** (0.045)	-0.011 (0.068)
实验组×第6期	0.018 (0.044)	0.094** (0.038)	-0.065 (0.064)
基期成绩	Y	Y	Y
学期	Y	Y	Y
控制变量	Y	Y	Y
N	7777	7777	7777

注:括号内为在“班级+学期”层面聚类调整的标准误。
***p<0.01, **p<0.05, *p<0.1。常数项略去。所有回归都控制了基期成绩、年龄、性别、民族、学校平均成绩。

大,与对实验期成绩的估计结果一致。对数学成绩的激励效果则在各成绩组均没有得到延续。动态效应的估计结果表明,语文成绩的效果在实验结束后的三期仍持续且均在0.1个标准差附近保持稳定,甚至略高于表6中实验期第二、第三期的估计系数。而数学成绩的效果则在实验结束后的第一期就消失了。

另外两个值得追问的问题是:实验对未获得奖金的学生是否产生了影响?各类获奖学生中,获得哪一类奖项的学生学习表现得到了更好的延续?表21通过对未获奖学生和获得各类奖项学生成绩在实验结束后的变化进行估计回答了这两个问题。

表21中“实验组”的估计系数反映的是没有获得奖项的学生在实验结束后相对于参照组期初成绩相同学生的成绩变化。估计结果显示,实验结束后,未获奖学生的语文成绩变化不大,但数学学习成绩可能受到一定程度的负面影响。这可能是由于数学相较于语文,要持续保持较好的成绩,需要更多持续的学习投入。但我们并不确定这一结果是由于没有获得奖励的学生因没有获得奖励遭受的负面打击导致,还是由于随着实验停止导致的奖励停止发放导致。

表21中“实验组”和三类奖项交乘项的系数反映的是获得各类奖项学生在实验结束后成绩变化情况。第(1)列估计结果显示,实验结束后,获优秀奖和作业奖学生的总成绩相较于期初显著提升,获进步

表 21 获奖学生在实验结束后的学业表现

变量	总成绩	语文	数学
	(1)	(2)	(3)
实验组	-0.036 (0.032)	-0.002 (0.032)	-0.114*** (0.042)
实验组×优秀奖	0.295*** (0.035)	0.298*** (0.032)	0.325*** (0.048)
实验组×进步奖	0.049 (0.042)	0.114*** (0.035)	-0.033 (0.057)
实验组×作业奖	0.149*** (0.031)	0.232*** (0.031)	0.093** (0.044)
基期成绩	Y	Y	Y
控制变量	Y	Y	Y
学期	Y	Y	Y
N	7777	7777	7777
adj.R ²	0.729	0.693	0.631

注:括号内为在“班级+学期”层面聚类调整的标准误。
***p<0.01, **p<0.05, *p<0.1。常数项略去。所有回归都控制了基期成绩、年龄、性别、民族、学校平均成绩、所获奖项。

奖学生总成绩与期初水平没有显著变化。第(2)列估计结果显示,获得优秀奖、进步奖和作业奖的学生语文成绩均显著高于参照组。这说明实验组语文成绩的持续效应主要体现在获奖学生成绩进步得以继续保持。获得优秀奖的学生在实验结束后语文成绩保持最好,与基准期相比提升0.298个标准差。获得作业奖的学生实验结束后进步0.232个标准差,略低于获得优秀奖学生。获得进步奖的学生实验结束后,成绩提升相对较低,比基准期提升0.114个标准差。第(3)列估计结果显示,获优秀奖和作业奖学生的数学成绩保持了较好水平,但获进步奖的学生均较实验前没有显著变化。

上述结果表明,激励效应在实验结束后虽然持续了,但这种持续没有普遍性,优秀奖学生和作业奖学生的持续性更强,进步奖学生持续性相对较弱。结合前面表14的分析来看,在数学成绩方面激励效应有持续可能主要是整体成绩较差的学校的那部分优秀学生。而语文成绩激励效应持续性较强,则很可能与表14中好学校和差学校的低成绩组语文成绩均表现出了明显的进步有关。

七、结论与启示

如何提升贫困地区和贫困家庭学生的教育水平是我国政府和社会长期关心的问题。在多方面的努力下,教育基础设施已经有了很大改善,义务教育阶段入学率也处于非常高的水平。如何提升贫困学生入学后的学习质量成为下一步贫困地区教育工作面临的主要问题之一。这是一个长期性的难题,受学校条件、教师质量和动力、家庭环境和父母背景、个人的学习动力等多个因素的影响。现金激励是一种执行成本较低、操作性较强的教育扶贫方式,获得了学术界、政府机构和社会的广泛关注。

本文运用随机田野实验的方法,研究了外部激励对贫困地区学生学习成绩的影响效果及其动态变化。研究发现,兼顾学习结果和学习过程的综合性激励方案,显著提升了学生学习成绩,虽然效果在实验的第一个学期不显著,但在实验的第二、第三个学期显著且持续稳定。异质性分析发现,激励效应在女生、少数民族学生、父母受教育程度低、中低收入家庭学生、期初成绩较差的学生以及较差学校的学

生等相对弱势群体更大。机制分析发现,经济激励提高了家长对孩子教育的关注程度、时间和资金投入,也促进了学生提升学习积极性和增加学习时间。对实验结束学生成绩变化的分析发现,激励效应在实验结束后得到一定程度持续。

本文研究表明,设计得当的现金激励方案可有助于提升中下游学生的学习积极性,进而提升其学习质量。在贫困地区设计教育资助政策时,要给予外部经济激励一定的空间。学习本身具有较强的累积性,前期的学习质量对后期的学习有至关重要的影响。外在经济激励本身是对学习成绩有积极影响。在基础教育阶段打下良好基础有助于后期的进一步学习,具有较大的正外部性。在有些深度贫困地区,基础教育质量与其它地区相比差距较大,较差的基础教育会导致后期教育和培训难以弥补的后果,采取激励的手段来保证基础教育质量达到一定的水准是有必要的。

本文的研究,对贫困学生激励方案的设计有以下两点启示。

首先,激励范围应考虑到难以获得绩优奖励的中下游学生,奖励条件要具备可达性。传统的学生激励方案更注重对优秀学生的奖励,这类类似做锦上添花的工作,最需要关注的成绩较中下游学生往往因难以达到获奖条件而难以受到激励。我们设立的学习进步奖和作业奖获奖难度相对较低,成绩中下游的学生只要努力就有较大机会获得奖励,这扩大了激励的对象范围,让成绩不是太好的学生能获得激励,有动力提高自己的学习成绩。

其次,在对学习结果进行激励的同时,还要同时重视对学习过程的激励。Fryer(2011)发现近些年的研究发现,学生往往不太清楚学习的生产函数,对学习过程的引导比对学习结果的引导更为重要。在设计经济激励机制时,有必要对学习的关键环节设置激励,引导学生养成良好的学习习惯。

虽然我们发现现金奖励对后进生产生了巨大影响,但这个影响到底是因为现金带来的还是因为荣誉(或者老师因为有后进生的奖励增加对学生的关心)带来的,这是我们在实验设计之初所忽略的。如果仅仅荣誉和关心就能大大提高后进学生的学习动力,无疑会大大降低教育干预的成本,提高教育政

策的投入产出效率,这是我们在后期的研究中拟进一步深入挖掘的一个方向。另外,本文的研究也表明现金激励也不是万能的,数学成绩的激励效应明显较弱。尤其需要注意的是,家庭收入最低的那部分学生并没有受到现金激励的影响,这部分学生值得我们加强关注,他们的学习过程可能受到更多不同寻常的约束,以致仅仅提高外在的激励无法帮助他们摆脱这些约束。如何有效提高最贫困家庭学生的学习质量,将是我们后期研究的重要主题。

注释:

①中心校主要指我国设立在乡镇一级,对所辖小学或初中具有一定的管理、指导职能的学校。村完小的全称是村级完全小学,即在村一级设有初级和高级两部的小学,一般学制为五年或六年。村完小和村完小在行政上可能同属一所小学管理,但教学相互独立。

参考文献:

- [1]牛新春:《招生倾斜政策下重点大学农村学生的学业准备和初期学业表现——基于X大学的实证案例研究》,《复旦教育论坛》,2017年第4期。
- [2]司树杰、王文静、李兴洲:《教育扶贫蓝皮书:中国教育扶贫报告(2016)》,社科文献出版社,2016年。
- [3]张彩、陈福美、李勉、张平平、王耘:《贫困地区学生数学与科学学业表现状况调查》,《教育学报》,2015年第1期。
- [4]周达、杜霄丰、刘坚:《我国东、中、西部地区四年级学生数学学业表现的现状分析》,《上海教育科研》,2017年第4期。
- [5]Angrist, J. and Lavy, V., 2009, "The Effects of High Stakes High School Achievement Awards: Evidence from a Randomized Trial", *American Economic Review*, 99(4), pp. 1384 ~ 1414.
- [6]Angrist, J., Lang, D. and Oreopoulos, P., 2009, "Incentives and Services for College Achievement: Evidence from a Randomized Trial", *American Economic Journal: Applied Economics*, 1(1), pp. 136 ~ 163.
- [7]Angrist, J., Oreopoulos, P. and Williams, T., 2014, "When Opportunity Knocks, Who Answers? New Evidence on College Achievement Awards", *Journal of Human Resources*, 49(3), pp. 572 ~ 610.
- [8]Banerjee, P. A., 2016, "A Systematic Review of Factors Linked to Poor Academic Performance of Disadvantaged Students in Science and Maths in Schools", *Cogent Education*, 3, 1178441.

- [9]Barrow, L. and Rouse, C. E., 2018, "Financial Incentives and Educational Investment: The Impact of Performance-based Scholarships on Student Time Use", *Education Finance and Policy*, 13(4), pp. 419 ~ 448.
- [10]Behrman, J. R., Parker, S. W., Todd, P. E. and Wolpin, K. I., 2015, "Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools", *Journal of Political Economy*, 123(2), pp. 325 ~ 364.
- [11]Bettinger, E. P., 2012, "Paying to Learn: The Effect of Financial Incentives on Elementary School Test Scores", *Review of Economics and Statistics*, 94(3), pp. 686 ~ 698.
- [12]Bloom, H. S., 2006, "Randomizing Groups to Evaluate Place-based Programs", *MDRC Working Papers on Research Methodology*.
- [13]Blimpo, M. P., 2014, "Team Incentives for Education in Developing Countries: A Randomized Field Experiment in Benin", *American Economic Journal: Applied Economics*, 6(4), pp. 90 ~ 109.
- [14]Bloom, N., Lemos, R., Sadun, R. and Van Reenen, J., 2015, "Does Management Matter in Schools?", *The Economic Journal*, 125(584), pp. 647 ~ 674.
- [15]Bonesrønning, H., 2004, "Do the Teachers' Grading Practices Affect Student Achievement?", *Education Economics*, 12(2), pp. 151 ~ 167.
- [16]Bowles, S., 1970, "Towards an Educational Production Function", in *NBER Chapters: Education, Income, and Human Capital*.
- [17]Bulman, G. and Fairlie, R. W., 2016, "Technology and Education: Computers, Software, and the Internet", in *Handbook of the Economics of Education*(Vol. 5), Elsevier.
- [18]Castleman, B. L., Page, L. C. and Schooley, K., 2014, "The Forgotten Summer: Does the Offer of College Counseling After High School Mitigate Summer Melt among College-intending, Low-income High School Graduates?", *Journal of Policy Analysis and Management*, 33(2), pp. 320 ~ 344.
- [19]Cha, P. and Patel, R., 2010, "Rewarding Progress, Reducing Debt: Early Results from the Performance-based Scholarship Demonstration in Ohio", *MDRC*, October.
- [20]Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F. and York, R., 1966, *Equality of Education Opportunity*, Washington, DC: National Center for Educational Studies.
- [21]Dahl, G. B. and Lochner, L., 2012, "The Impact of Fami-

ly Income on Child Achievement: Evidence from the Earned Income Tax Credit", *American Economic Review*, 102(5), pp. 1927 ~ 1956.

[22]Ding, W. and Lehrer, S. F., 2007, "Do Peers Affect Student Achievement in China's Secondary Schools?", *The Review of Economics and Statistics*, 89(2), pp. 300 ~ 312.

[23]Duflo, E., Glennerster, R. and Kremer, M., 2008, "Using Randomization in Development Economics Research: A Toolkit", in *Handbook of Development Economics*(Vol.4), pp. 3895 ~ 3962, Elsevier.

[24]Dynarski, S., 2008, "Building the Stock of College-educated Labor", *Journal of Human Resources*, 43(3), pp. 576 ~ 610.

[25]Eren, O. and Henderson, D. J., 2008, "The Impact of Homework on Student Achievement", *The Econometrics Journal*, 11(2), pp. 326 ~ 348.

[26]Fryer Jr, R. G., 2011, "Financial Incentives and Student Achievement: Evidence from Randomized Trials", *The Quarterly Journal of Economics*, 126(4), pp. 1755 ~ 1798.

[27]Glewwe, P. W., Hanushek, E. A., Humpage, S. D. and Ravina, R., 2013, "School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010", NBER Working Paper, 17554.

[28]Guryan, J., Kim, J. S. and Park, K. H., 2016, "Motivation and Incentives in Education: Evidence from a Summer Reading Experiment", *Economics of Education Review*, 55, pp. 1 ~ 20.

[29]Hanushek, E. A., 2003, "The Failure of Input-based Schooling Policies", *The Economic Journal*, 113(485), pp.F64 ~ F98.

[30]Hanushek, E. A. and Raymond, M. E., 2005, "Does School Accountability Lead to Improved Student Performance?", *Journal of Policy Analysis and Management*, 24(2), pp. 297 ~ 327.

[31]Hanushek, E. A. and Wößmann, L., 2007, "The Role of Education Quality for Economic Growth", World Bank Policy Research Working Paper.

[32]Jackson, C. K., 2010, "The Effects of an Incentive-based High-school Intervention on College Outcomes", Nber Working Paper, No. 15722.

[33]Jackson, C. K., Rockoff, J. E. and Staiger, D. O., 2014, "Teacher Effects and Teacher-related Policies", *Annual Review of Economics*, 6(1), pp. 801 ~ 825.

[34]Kremer, M., Miguel, E. and Thornton, R., 2009, "Incentives to Learn", *The Review of Economics and Statistics*, 91(3), pp. 437 ~ 456.

[35]Lacour, M. and Tissington, L. D., 2011, "The Effects of Poverty on Academic Achievement", *Educational Research and Reviews*, 6(7), pp. 522 ~ 527.

[36]Lavecchia, A. M., Liu, H. and Oreopoulos, P., 2016, "Behavioral Economics of Education: Progress and Possibilities", in *Handbook of the Economics of Education*(Vol. 5), Elsevier.

[37]Leuven, E., Oosterbeek, H. and Van der Klaauw, B., 2010, "The Effect of Financial Rewards on Students' Achievement: Evidence from a Randomized Experiment", *Journal of the European Economic Association*, 8(6), pp. 1243 ~ 1265.

[38]Leuven, E., Oosterbeek, H., Sonnemans, J. and Van Der Klaauw, B., 2011, "Incentives versus Sorting in Tournaments: Evidence from a Field Experiment", *Journal of Labor Economics*, 29(3), pp. 637 ~ 658.

[39]Levitt, S. D., List, J. A. and Sadoff, S., 2016a, "The Effect of Performance-based Incentives on Educational Achievement: Evidence from a Randomized Experiment", NBER Working Paper, No. 22107.

[40]Levitt, S. D., List, J. A., Neckermann, S. and Sadoff, S., 2016b, "The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance", *American Economic Journal: Economic Policy*, 8(4), pp. 183 ~ 219.

[41]Li, T., Han, L., Zhang, L. and Rozelle, S., 2014, "Encouraging Classroom Peer Interactions: Evidence from Chinese Migrant Schools", *Journal of Public Economics*, 111(C), pp. 29 ~ 45.

[42]List, J. A., Livingston, J. A. and Neckermann, S., 2018, "Do Financial Incentives Crowd out Intrinsic Motivation to Perform on Standardized Tests?", *Economics of Education Review*, 66, pp. 125 ~ 136.

[43]MacDonald, H., Bernstein, L. and Price, C., 2009, "Foundations for Success: Short-Term Impacts Report", Canada Millennium Scholarship Foundation.

[44]Martorell, P., Stange, K. and McFarlin Jr, I., 2016, "Investing in Schools: Capital Spending, Facility Conditions, and Student Achievement", *Journal of Public Economics*, 140, pp. 13 ~ 29.

[45]Mbiti, I., Muralidharan, K., Romero, M., 2019, "Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania", *The Quarterly Journal of Economics*, 134, pp. 1627 ~ 1673.

[46]Miller, G., Luo, R., Zhang, L., Sylvia, S., Shi, Y., Foo, P., Zhao, Q., Martorell, R., Medina, A. and Rozelle, S., 2012, "Effectiveness of Provider Incentives for Anaemia Reduction in Rural China: A Cluster Randomised Trial", *BMJ*, 345, doi: <https://doi.org/10.1136/bmj.f7682>.

org/10.1136/bmj.e4809.

[47]Neilson, C. A. and Zimmerman, S. D., 2014, "The Effect of School Construction on Test Scores, School Enrollment, and Home Prices", *Journal of Public Economics*, 120, pp. 18 ~ 31.

[48]Oswald, Y. and Backes-Gellner, U., 2014, "Learning for a Bonus: How Financial Incentives Interact with Preferences", *Journal of Public Economics*, 118, pp. 52 ~ 61.

[49]Richburg-Hayes, L., Sommo, C. and Welbeck, R., 2011, "Promoting Full-Time Attendance among Adults in Community College: Early Impacts from the Performance-Based Scholarship Demonstration in New York", in MDRC.

[50]Riccio, J., Dechausay, N., Miller, C., Nuñez, S., Verma, N. and Yang, E., 2013, "Conditional Cash Transfers in New York City: The Continuing Story of the Opportunity NYC-Family Rewards Demonstration", in MDRC(Issue September).

[51]Sjoquist, D. L. and Winters, J. V., 2015, "State Merit-based Financial Aid Programs and College Attainment", *Journal*

of Regional Science, 55(3), pp .364 ~ 390.

[52]Scott- Clayton, J., 2011, "On Money and Motivation a Quasi-experimental Analysis of Financial Incentives for College Achievement", *Journal of Human Resources*, 46(3), pp. 614 ~ 646.

[53]Springer, M. G., Rosenquist, B. A. and Swain, W. A., 2015, "Monetary and Nonmonetary Student Incentives for Tutoring Services: A Randomized Controlled Trial", *Journal of Research on Educational Effectiveness*, 8(4), pp. 453 ~ 474.

[54]Todd, P. E., Wolpin, K. I., 2003, "On the Specification and Estimation of the Production Function for Cognitive Achievement", *The Economic Journal*, 113(485), pp. 3 ~ 33.

[55]Visaria, S., Dehejia, R., Chao, M. M. and Mukhopadhyay, A., 2016, "Unintended Consequences of Rewards for Student Attendance: Results from a Field Experiment in Indian Classrooms", *Economics of Education Review*, 54, pp. 173 ~ 184.

[56]World Bank, 2018, *Learning to Realize Education's Promise*, Washington: The World Bank Publications.

附表 1 对家长行为影响的民族和性别差异

变量	Ln(陪伴时间+1)		Ln(学习指导时间+1)	
	父亲	母亲	父亲	母亲
	(1)	(2)	(3)	(4)
实验组	0.479 (0.329)	0.249 (0.337)	0.865*** (0.198)	-0.262 (0.280)
少数民族	-0.193 (0.271)	-0.204 (0.280)	0.072 (0.185)	-0.330 (0.250)
实验组×少数民族	-0.031 (0.311)	0.034 (0.313)	-0.009 (0.203)	0.380 (0.269)
基期	Y	Y	Y	Y
控制变量	Y	Y	Y	Y
班级	Y	Y	Y	Y
N	747	748	717	714
adj.R ²	0.240	0.315	0.221	0.274
少数民族样本				
实验组	0.480** (0.183)	0.433*** (0.144)	0.735*** (0.111)	-0.020 (0.111)
女	-0.008 (0.187)	-0.134 (0.161)	-0.148 (0.128)	-0.289* (0.164)
实验组×女	0.005 (0.276)	-0.146 (0.209)	0.281 (0.213)	0.426** (0.204)
基期	Y	Y	Y	Y
控制变量	Y	Y	Y	Y
班级	Y	Y	Y	Y
N	371	373	351	354
adj.R ²	0.242	0.247	0.247	0.268