

“影子教育”会扩大教育结果不均等吗？ ——基于 PISA 2012 数据的中国、日本、韩国比较研究

胡咏梅，范文凤，丁维莉

[摘要] 私人课外补习的“影子教育”系统在东亚国家相当普遍。本文基于国际学生能力评价项目 PISA 2012 的数据，使用多层线性模型，分析了中国上海、中国香港、日本、韩国四个国家和地区“影子教育”对学生数学素养的影响，并使用再加权倾向得分匹配法(RPSM)估算“影子教育”的净效应。研究发现：课外数学补习对学生数学成绩具有显著的正向效应，且日本、韩国学生的数学补习效应要高于中国上海、香港学生；参加数学课外补习可以缩减家庭经济社会地位带来的成绩差异，从而起到促进教育结果均等的作用。

[关键词] “影子教育”；教育结果不均等；PISA；再加权倾向得分匹配法

一、引言

私人课外补习(private supplementary tutoring)^①因效仿正规教育体系，补习内容随着正规教育体系课程内容的改变而改变，规模随着正规教育体系

[收稿日期] 2017-06-21

[基金项目] 国家自然科学基金面上项目“直面影子教育系统：我国义务教育课外补习研究”(71373165)，北京师范大学“2011 计划”项目“影响学生发展的相关因素测评工具研发”(370331321)。

[作者简介] 胡咏梅，北京师范大学教育学部/首都教育经济研究院，电子邮箱地址：huym0718@bnu.edu.cn；范文凤，北京教育科学研究院教育督导与教育质量评价研究中心，电子邮箱地址：fanaimiao@sina.com；丁维莉，加拿大女王大学公共政策研究院，电子邮箱地址：dingw@queensu.ca。

① 国内许多学者如薛海平、丁小浩、雷万鹏等将“Private Supplementary Tutoring”译成“教育补习”，与本文课外补习涵义相同，本文两种译法混用(薛海平，丁小浩，2009；雷万鹏，2005)。

的生源扩大而扩张,因而被喻为影子教育(Stevenson & Baker, 1992)。^①在世界各地,受全球化和各项经济、社会发展指标排名以及国际学生成就评估的影响,教育的国际竞争日趋激烈,私人课外补习的“影子教育”系统(shadow education system)也正在不断扩张^②(Bray, 2005)。这一现象在东亚地区一直较为普遍,尤其是在中国、日本、韩国等国家更为盛行(Stevenson & Baker, 1992; Ngai and Cheung, 2010; Caritas, Community and Higher Education, 2010; Ministry of Education and Training, 2008; Kim, 2010),这与东亚地区深受儒家文化影响以及这些国家升学通常采用选拔式考试有关。自20世纪90年代中期“影子教育”在中国大陆兴起,目前在中国城镇地区已经相当普遍(薛海平、丁小浩, 2009; 李佳丽等, 2016)。

我们利用2009年参与PISA项目测试的中国上海、中国香港、日本和韩国学生的数据统计发现(参见表1),在四类补习(即测试语言补习、数学补习、科学补习、其他补习)中,参与数学补习的比例相对最高(香港除外),分别是71.3%、48.5%、76.6%、77.5%;参与科学补习的比例相对最低,分别是29%、27.9%、61.6%、57%。日本和韩国学生参与测试语言、数学、科学补习的比例均较为相近。上海和香港学生参与科学补习的比例相差很小,而测试语言补习和数学补习的比例均相差较大,上海学生参与测试语言补习和数学补习的比例均超过香港学生比例约23个百分点。根据表1中参与数学补习的比例可以看出,中国上海、中国香港、日本、韩国学生的补习比例与前述五项调查研究的结果相差不大。

表1 中国上海、中国香港、日本、韩国2009年参与PISA项目测试学生的补习情况统计

国家/地区	测试语言补习比例 ^a	数学补习比例	科学补习比例 ^b	其他补习比例 ^c
中国上海	54.3%	71.3%	29.0%	62.1%
中国香港	30.6%	48.5%	27.9%	55.3%
日本	64.4%	76.6%	61.6%	76.0%
韩国	68.0%	77.5%	57.0%	67.8%

注: a. 测试语言补习是指该国(或地区)学生母语补习,比如上海学生的测试语言补习是指中文补习。

b. 科学补习是指初中科学类课程补习,比如对上海学生来说,包括物理、化学、生

^① 马克·贝磊曾指出,“影子教育”并不包括其他非学科性的课外活动(如绘画、弹琴、体育活动等)的学习。

^② 不仅亚洲国家的教育补习现象较为普遍,而且欧洲、北美、非洲、澳大利亚等国家和地区的教育补习也日渐增长。

物等课程的补习。

c. 其他补习是指前三列未涉及的学科补习，比如对于上海学生来说，英语补习归为其他补习类别。

诸多研究发现学生的家庭社会经济背景影响着学生参与“影子教育”的机会(薛海平、丁小浩, 2009; Southgate, 2009), 而“影子教育”对社会分层也产生了一定影响(薛海平, 2015, 2016, 2017), 因而, 逐渐受到政府、政策制定者以及学界的广泛关注(Bray et al., 2014)。研究教育补习对学生成绩的影响对于家长更好地理解补习作用, 理性制定补习决策, 提高投资收益是相当重要的。同时, 对于政府更有效地配置基础教育资源, 改善基础教育不公平状况也将提供有益参考。昂贵的补习费用很可能使得来自经济收入较低的家庭学生参与补习的比例低于经济收入较高的家庭学生, 从而造成课外教育机会的不均等, 这种不均等是否会扩大教育结果的不均等? 由此, 本研究将利用 PISA 2012 中国上海、中国香港、日本、韩国的数据, 试图探究“影子教育”对于中国、日本、韩国中学生学业成绩的因果效应, 并回答“影子教育”是否会扩大教育结果的不均等问题。

下文的具体安排如下: 第二节是关于课外补习对学生成绩的影响的文献梳理, 以及“影子教育”可能产生的社会不平等问题文献分析, 提出本研究需要回答的主要问题以及待验证的研究假设。第三节简述样本选择的依据, 介绍计量模型方法和变量定义。第四节呈现实证分析结果: 首先基于描述性统计分析不同群体学生参与数学课外补习的比例差异以及参与数学课外补习与未参与数学课外补习学生的数学成绩差异; 其次, 采用再加权倾向得分匹配法(RPSM)估计课外补习的净效应, 利用两水平线性模型分析各类补习时间对数学成绩的影响效应, 并检验“影子教育”是否会扩大教育结果不均等的研究假设。第五节进行总结、讨论并提出政策建议。

二、文献综述

最早研究教育补习现象的学者是 Stevensen 和 Baker(1992), 他们不仅提出了“影子教育”的概念, 还利用对日本高中生接受教育补习的调查数据分析了“影子教育”对高中生升入大学机会的影响。贝磊(2007)的另一项研究表明, 教育补习不仅可以提高学生的数学学业成绩, 而且为学生在校外提供了更有挑战性的学习活动, 从而增强了学生对数学学习的兴趣。而 Smyth(2008)对爱尔兰的学生补习现象研究发现, 接受过教育补习的学生与没有接受过的学生的期末考试成绩不存在显著性差异。张羽采用多水平线性模型和条件分位

数回归模型对中国济南高中生参与课外补习对其高考成绩的因果效应进行了较为全面、深入的探究(Zhang, 2013)。研究发现, 课外补习对城市学生或是农村学生的高考总成绩均没有产生显著的效应, 但是对于来自薄弱学校的城市学生或是成绩较低的城市学生产生了显著的正向效应。薛海平(2016)采用PSM方法发现初中生参加学术类课外补习有助于提高语文、数学、英语各单科成绩, 但对各单科成绩提升幅度有较大差异; 未参加学术类课外补习初中生语数外各科潜在补习收益均高于参加课外补习组的补习收益。

许多社会学者和教育学者对“影子教育”的研究更多关注在“影子教育”可能产生的社会不平等问题。有学者认为, “影子教育”的存在会对社会公平产生威胁, 它可能成为维持和扩大社会的不平等的机制(Verdis, 2002; Murawska & Putkiewicz, 2006; Smyth, 2009)。薛海平(2015)的研究得出如下结论: 我国义务教育阶段的课外补习活动扩大了城乡和不同阶层义务教育学生获得教育资源和教育结果的差距, 影子教育因此成为城乡和阶层差距在代际间维持和传递的一个重要通道。薛海平(2017)的另一项研究中同样提到在家庭资本代际继承过程中, 学校教育和影子教育均发挥了重要的通道作用, 从而建立起了学校教育和影子教育双重再生产机制, 社会阶层将因此而不断被复制, 社会流动将被阻断, 最终导致阶层固化。然而, 李佳丽等(2016)基于Wisconsin模型, 引入父母和子女自我的教育期望作为影响子女学业成绩和学生影子教育参与的中介变量, 研究发现影子教育部分调节了父母的受教育程度、家庭经济水平等家庭背景变量对学生学业成绩的影响。因此对于“影子教育”会维持、甚至扩大社会的不平等尚无统一的定论。

自2013年底PISA项目2012测试结果公布以来, 亚洲国家和地区学生的优异成绩举世瞩目, 也引来许多国外媒体的关注。比如, 《华盛顿邮报》2013年12月5日曾以“勤奋的亚洲学生称霸全球性考试”为标题总结PISA的测试结果。2012年共有65个国家及地区的大约51万名15岁学生参加了PISA测试, 中国上海学生在数学、语言、科学素养三项测试中均排名首位。中国香港、新加坡、中国台湾、日本、韩国等地的学生也都名列前茅^①。美国学生则在各科排名中都未进入20强。如何正确理解和解释亚洲部分国家和地区学生的PISA成绩排名是本文试图回答的问题之一。我们希望通过对中国上海、中国香港、日本、韩国学生参与教育补习的对比分析, 揭示参与“影子教育”的学生群体的特征, 探究“影子教育”是否真正能够带来学生成绩的提

^① 2009、2012年的PISA排名靠前的国家或地区多数课外补习比例较高, 这意味着其较高的成绩不仅是学校教育的成果, 也与校外的教育补习相关。

升以及可能存在的异质性，并且估算补习时间对成绩的影响效应。此外，着重探讨“影子教育”是否会扩大教育结果的不均等问题，以期为国家 and 政府制定规范和控制“影子教育”的相关政策提供实证依据。由此，我们试图借助 PISA2012 中国上海、中国香港、日本以及韩国的数据检验以下四个研究假设：

研究假设 1：来自家庭经济社会地位高的学生参与“影子教育”的概率显著高于来自家庭经济社会地位低的学生。

研究假设 2：参与数学课外补习可以显著提升学生的数学成绩。

研究假设 3：增加数学课外补习时间可以显著提升数学成绩。

研究假设 4：“影子教育”会扩大教育结果的不均等。

三、样本选择与计量模型

(一)数据来源

本研究使用的数据来源于 PISA 官方网站(<http://www.pisa.oecd.org>)。PISA 作为国际权威的学生学业成就测试项目，其学科素养测试以及背景问卷的设计均由国际学科专家和测量专家合作完成，学生学科素养的估算采用了现代教育测量理论(如项目反应项目等)并利用专业的统计软件(如 Conquest)获得的。PISA 2012 的主要评估领域是数学，测试的数学素养既体现学生对各种数学知识(空间与图形、变化与关系、数量、不确定性)的掌握，又反映了其数学认知能力(数学表述、数学运用和数学阐释能力)的表现，学生的数学素养得分具有较好的综合性和代表性。

(二)样本选择

本研究选取 PISA 2012 数据库中中国上海、中国香港、日本、韩国四个地区和国家的数据。之所以选择这四个地区和国家，首先是由于前文所提到的，“影子教育”在这些国家和地区规模都较大，都是深受儒家文化影响的国家和地区；其次，中国上海、中国香港、日本、韩国在 2012PISA 三科素养排名中名列前茅，而本文试图解释“影子教育”在其中的作用。

2012 年中国上海、中国香港、日本、韩国参加 PISA 项目测试的共有 21231 名 15 岁 3 个月—16 岁 2 个月的在校学生，中国上海、中国香港、日本、韩国参与 PISA 项目的学校数分别为 155、148、191、156，学生数分别为 5177、4670、6351、5033。PISA 抽样是采用两阶段按规模大小成比例的概率抽样(PPS)以及处理特殊规模学校(如特小规模学校和特大规模学校)的

方法,因而各地区和国家抽取的学校数和学生数并不完全相同,但都要求在150所学校和4500名学生以上,以保证估计校间方差和校内方差的精度。由上述数据知,中国上海、日本、韩国的学校、学生被试量均符合要求,中国香港基本符合要求。

(三) 变量选取与计量模型

本研究从PISA 2012学生问卷和学校问卷数据构成的嵌套关系中,考察教育补习对学生数学素养得分(以下简称数学成绩)的影响。尽管在线性回归模型中加入控制变量可使回归模拟实验结果,但很难保证找到合适的控制变量,使得条件独立假设(Conditional Independence Assumption, CIA)^①成立。在本研究中,是指即使加入了学生水平和学校水平的部分控制变量,仍很难检验条件独立假设成立。近年来,通过匹配策略保障条件独立假设成立的PSM方法得以广泛使用,本研究将采用Rosenbaum和Rubin(1983)提出的再加权倾向得分匹配法(Reweighting on Propensity Score Matching,以下简称RPSM)来估算补习带来的学生成绩的因果效应。

除了用再加权倾向得分匹配法揭示补习带来的学生成绩的因果效应外,我们还将稳健性检验中给出加入相同控制变量后的多水平线性回归模型的补习效应估计,并将之与RPSM估计结果相比较。

本研究的因变量是学生的数学成绩。之所以没有选择阅读成绩和科学成绩,是由于2012年PISA的主要评估领域是数学,阅读和科学素养的测评题量相对较小,覆盖的范围不足以全面评估学生的阅读素养和科学素养,因而本研究没有将这两科素养成绩作为计量模型的因变量。此外需要说明的是,在PISA2012原始数据中,为阅读、数学、科学各提供了5个PV值(似真值,plausible value)^②。一般来说,由于似真值具有随机性,因而不适合作为个体分数向学生报告,但它们在估计总体参数时具有不可替代的优势,比如利用似真值可以得到总体参数的无偏估计、在复杂的抽样设计中,可以得到标准误的精确估计等等。

本研究参照PISA数据分析的做法,在多水平模型中同时使用这5个PV值。自变量分为两层:第一层是学生水平变量,包括学生性别(gender)、学

① 条件独立假设是能对回归赋予因果解释的核心假设。这个假设有时又被称为选择性偏误来自可观察变量。即加入这些可观察变量后,选择性偏误会消失。

② 所谓PV值,是根据项目反应理论中的边际估计技术,基于学生的作答反应和背景变量获得学生能力的一个后验分布,从这个后验分布中随机抽取5个值即为PV值。其直观解释是指一个学生可能拥有的能力范围的代表值。

生家庭经济社会文化地位(*escs*)、数学课外补习(*mathoutschool*)；第二层是学校水平变量，包括学校类型(*public*)和所在地区(是否城镇，*location*)。

$$MS = \alpha_0 + \alpha_1 gender + \alpha_2 escs + \alpha_3 mathoutschool + \varepsilon$$

$$\alpha_0 = \gamma_{00} + \gamma_{01} public + \gamma_{02} location + \mu_0$$

其中 MS 代表学生的数学成绩，其他变量的含义详见表 2。在分析学生成绩变异的教育生产函数模型设定时，通常将性别、家庭经济社会地位以及学校类型等作为控制变量，以使得在缺少随机分配的情况下来模拟实验结果，从而使得关键解释变量的回归系数可以视作因果效应的近似估计(安格里斯特、皮施克，2012)。选择这些变量是参考了已有相关研究的结果。根据雷万鹏(2005)和薛海平(2009)的研究发现，家庭收入水平、父母受教育程度、学业成绩、城乡差异、地区差异等是学生参加教育补习的重要影响因素，研究发现家庭社会经济背景水平越好(即家庭收入水平越高、父母受教育程度越高)、大城市学生(相对于农村学生)、公办学校(相对于民办学校)参加教育补习的可能性越大。而在学业成绩上两者的结论并不相同。薛海平(2009)的研究发现，学习基础好的学生参与补习的可能性更大，而雷万鹏(2005)的研究表明，学习成绩差的学生参与补习的可能性更大。Bray 等(2014)对香港中学生的调查发现，女生在课外补习上的花费显著高于男生，家庭收入水平对课外补习支出有显著的正向效应。据韩国近期的一项调查也发现，父母文化程度、家庭经济状况与孩子参加课外辅导与否则有直接关系。拥有大学学历的家长有 90% 让孩子参加课外辅导班，而高中学历的家长让子女接受课外辅导的比例为 50%。月收入 700 万韩元的家庭为子女支付的课外辅导费用，比月收入 100 万韩元的低收入家庭高出近 10 倍(中国教育报，2011-08-16)。因而，本研究也将这些变量作为可能影响学生参与课外补习机会的自变量纳入计量模型中。因为倾向得分匹配模型中自变量的选择需要既考虑影响结果变量的因素，同时也要考虑影响接受处理的因素，因此，本研究也将学生性别、学生家庭经济社会地位、学校类型、所在地区变量作为 RPSM 模型的协变量，使得补习效应的估计更接近因果效应。^①

模型中使用的自变量以及后文描述性分析使用的变量定义和计分方式如表 2 所示。

^① 本研究将多水平模型的分析结果作为倾向得分匹配模型结果的稳健性检验的一部分，因而在模型选择上主要依据倾向得分匹配模型的自变量选择方式，采用与该模型相同的变量作为多水平模型中的自变量，从而使得两种模型方法的结果有可比性。

表2 变量名称及其编码方式

变量	定义	计分方式
学生水平变量		
性别 (gender)	学生性别	离散变量, 2点计分, 1—女, 0—男
家庭经济社会地位 (escs)	根据父母最高职业地位、最高受教育水平及家庭拥有物三项指标合成	连续变量
数学课外补习 (mathoutschool)	每周是否参加数学课外补习	离散变量, 2点计分, 1—是, 0—否
语言课外补习 (langoutschool)	每周是否参加语言课外补习	离散变量, 2点计分, 1—是, 0—否
科学课外补习 (scioutschool)	每周是否参加科学课外补习	离散变量, 2点计分, 1—是, 0—否
数学课外补习时间 (mathtime)	每周参加数学课外补习时间	没有参加补习记为0小时, 每周参加少于2小时记为1小时, 每周参加2—4小时记为3小时, 每周参加4—6小时记为5小时, 每周参加6及6小时以上记为6小时。
语言课外补习时间 (langtime)	每周参加语言课外补习时间	
科学课外补习时间 (scitime)	每周参加科学课外补习时间	
其他课外补习时间 (othertime)	每周参加其他课外补习时间	
学校水平变量		
班级规模 (classsize)	所在班级的学生数	连续变量
学校类型 (public)	所在学校是公立还是私立学校	离散变量 2点计分, 1—公立学校, 0—私立学校
学校所在地 ^a (location)	学校所在地是城镇还是农村	离散变量 2点计分, 1—城镇, 0—农村

注 a: PISA 样本中将学校所在地分为 5 类, 分别是: 农村、小城镇、城镇、城市和大城市, 此处将农村和小城镇归为农村, 将城镇、城市、大城市归为城市。

四、实证分析结果

(一) 学生和学校变量描述性统计

从中国上海、中国香港、日本、韩国学生水平和学校水平变量的统计描述(见表3)可以看出:(1)学生水平的变量中,中国上海的学生数学成绩最高,其次是中国香港、韩国,日本相对最低,而且四个国家/地区学生数学成绩均高于OECD国家的平均水平(494)。中国上海参与PISA项目的女生比例略多于男生,而其他三个国家/地区则是男生略多于女生。韩国、日本学生的家庭经济地位变量均值高于中国上海和中国香港学生。中国上海和日本学生中参与数学补习的比例均达到或超过了70%,而语言补习、科学补习、其他补习的比例也均在50%—60%之间,韩国和中国香港学生参与各类补习的比例相对低一些,不过数学补习比例均比其他课程补习要高;四个地区和国家学生每周参与数学补习的时间均多于其他科目的补习时间,且韩国、中国上海学生每周平均数学补习时间(分别是2.56小时、2.01小时)要多于日本学生(1.74小时)和中国香港学生(1.07小时)。(2)学校水平变量中,中国上海、日本、韩国抽样学校均是以公立学校居多,中国上海公立学校比例甚至达到91%,日本为70%,韩国为53%,中国香港公立学校比例仅有7%,这与香港教育体制有关^①。中国上海、日本的平均班级规模分别为39.29、36.29,略高于中国香港(34.08)、韩国(33.6)。中国上海和香港抽样学校所在地均属城镇,日本、韩国分别有98.4%、93.7%的抽样学校在城镇,农村学校所占比例相对较低。

表3 变量的描述性统计(均值/标准误)

	中国上海	中国香港	日本	韩国
学生水平变量				
数学成绩	612.68(100.979)	561.24(96.31)	536.41(93.524)	553.77(99.077)

^① 根据现在香港教育局的统计分类,香港的中小学校分为:官立学校、津贴/资助学校、直资学校(“直接资助计划”学校)、私立学校、国际学校。所谓官立学校,是指政府创办和经营的学校,政府承担这些学校的所有开支,并且学校教职工属于公务员。2011—2012学年,在香港1092所(同时亦有中学和小学单独计算)中小学校中,仅有6%属于官立学校,即公立学校。[http://www.edu.cn/xsc_12533/20130614/t20130614_963980.shtml\[OB/EL\]-2014-10-07](http://www.edu.cn/xsc_12533/20130614/t20130614_963980.shtml[OB/EL]-2014-10-07)

	中国上海	中国香港	日本	韩国
女生比例	0.51(0.500)	0.46(0.499)	0.47(0.499)	0.47(0.499)
家庭经济社会地位	-0.36(0.964)	-0.79(0.973)	-0.07(0.713)	0.012(0.743)
数学补习比例	0.71(0.455)	0.47(0.499)	0.70(0.459)	0.66(0.474)
语言补习比例	0.51(0.500)	0.24(0.430)	0.58(0.494)	0.53(0.499)
科学补习比例	0.55(0.497)	0.29(0.454)	0.54(0.498)	0.39(0.488)
其他补习比例	0.57(0.495)	0.42(0.493)	0.69(0.461)	0.65(0.478)
数学补习时间	2.01(1.905)	1.07(1.527)	1.74(1.814)	2.56(1.370)
语言补习时间	1.33(1.786)	0.44(1.019)	1.04(1.323)	1.99(1.108)
科学补习时间	1.49(1.820)	0.68(1.378)	0.95(1.279)	1.70(1.007)
其他补习时间	1.41(1.736)	0.90(1.438)	1.62(1.758)	2.37(1.281)
学校水平变量				
公立学校比例	0.91(0.290)	0.07(0.251)	0.70(0.458)	0.53(0.499)
班级规模	39.29(7.558)	34.08(5.018)	36.29(5.662)	33.60(6.041)
城市学校比例 ^a	1.00(0)	1.00(0)	0.98(0.127)	0.94(0.244)

注 a: PISA 样本中将学校所在地分为 5 类, 分别是: 农村、小城镇、城镇、城市和大城市, 此处将农村和小城镇归为农村, 将城镇、城市、大城市归为城市。

(二)不同群体学生参与数学补习的比例分析

由表 4 可以看出, 无论哪个国家或地区, 家庭经济社会地位高于平均值的 学生参加数学课外补习的比例要显著高于家庭经济社会地位低的学生。这一结果验证了研究假设 1。中国上海、韩国两者分别相差 9.7 和 12.8 个百分点, 而中国香港、日本甚至达到 17 个百分点。尽管 PISA 项目对学生和家长的调查问卷中没有涉及补习费用, 但根据前文已述的几项在中国、日本、韩国的调查结果均表明, 课外补习费用已占据家庭教育支出的重要地位, 高昂的课外补习费用对于家庭经济收入较低的家庭来说, 是难以负担的。因此, 上述四个国家或地区来自较高家庭经济社会地位的参与课外补习的学生比例高于来自较低家庭经济社会地位的学生参与补习比例也是情理之中。这种由于家庭经济背景不同造成的补习机会差异是否会导致这两类学生群体学业成绩的更大差异呢? 换言之, 补习是否会扩大教育结果的不均等程度呢? 这也是下文要着重探讨的问题。

除了香港以外, 其他三个国家或地区私立学校的学生参与课外数学补习的比例显著高于公立学校, 这与来自私立学校的学生家庭经济条件相对较好, 有能力支付高昂的补习费用有关。香港公立学校的学生参与课外补习的比例高于私立学校, 我们推测与公立学校的教育质量不如私立学校有关。香港人

均收入水平很高^①，无论是公立学校的学生家庭，还是私立学校学生的家庭，都有能力支付补习费用^②。由于私立学校已经提供了很好的教育服务，教育质量也较高，因而该类学校学生无需再进入课外补习市场。日本、韩国来自城镇的学生参与数学补习的比例均显著高于农村学生，不过韩国来自农村的学生参与数学补习的比例也较高，达到 58.5%。这与韩国为推进“教育平准化”政策，构筑“新教育体制”的改革方案有关。2005 年起，韩国正式推出由公立学校实施的“放学后教育计划”。按照该计划，公立学校利用放学后的时间，向学生提供保育、托管、学术课程补习、艺术课程辅导和娱乐活动等服务。中央和地方政府提供专项经费支持，各大学和研究机构开发具体的课外辅导项目。在政府经费支持下，不仅城市学生能以低于市场的价格满足其个性化学习需求，而且农村和低收入家庭子女也有条件接受免费的课外辅导，缩小与高收入家庭子女因收入差距带来的教育差距。近几年来，韩国政府不断加大“放学后教育计划”的实施力度和经费投入。据韩国教育科学技术部统计，目前已有 99.8% 的中小学开办了“放学后教育”，近半数中小學生参与其中(中国教育报，2011—08—16)。

表 4 不同学生群体参与数学课外补习的比例及差异检验

	中国上海		中国香港		日本		韩国		
	比例	卡方值	比例	卡方值	比例	卡方值	比例	卡方值	
SES ^③	高 ses	75.4%	648***	55.6%	1332***	78.2%	24210***	72.1%	7151***
	低 ses	65.7%		38.6%		61.5%		59.3%	
公私立	公立	70.1%	89***	53.7%	68***	67.8%	3264***	63.5%	1207***
	私立	76.3%		46.2%		74.5%		68.8%	
性别	男生	67.0%	345***	45.3%	47***	67.2%	2561***	65.4%	75***
	女生	74.1%		48.4%		72.6%		66.7%	
地区	城镇	70.7%		46.7%		70.4%	7878***	66.5%	652***
	农村	—		—		32.0%		58.5%	

注：“*”表示在 0.1 水平上显著，“**”表示在 0.05 水平上显著，“***”表示在 0.01 水平上显著。下同。

① 由世界银行发布的数据可知，2012 年香港人均 GDP 36708 美元，世界排名第 27 位。

② 据马克·贝磊等人的研究表明(Bray, et. al., 2014)，香港中学生课外补习的费用占家庭月收入的平均比重为 8.7%。

③ 由 escs 变量生成，离散变量，2 点计分，1 为高于所在参与国学生 escs 平均值，0 为低于平均值。

以上统计分析发现不同群体学生参与数学补习的比例存在显著性差异,将在后文 Logit 模型中进一步检验,此处的分析结果也是 Logit 模型中加入 SES、性别、学校类型以及所在地区变量作为解释变量的依据。

(三)不同群体补习与不补习学生成绩差异

由表 5 可以看出,在分性别比较中,中国上海补习与不补习的男生数学成绩相差 22 分,而补习与不补习的女生的成绩差异为 17 分,男生补习与不补习成绩差值比女生高 4.5 分,但是男女生补习与不补习成绩差值没有显著性差异。在参与数学补习的上海学生中,女生数学成绩显著低于男生(低 9.7 分),而在没有参与数学补习的上海学生中,女生数学成绩比男生低 5.2 分左右,没有显著性差异。显然,补习与不补习者的成绩差异不完全是补习的效应,我们的基本推测是,数学补习给上海男生带来的成绩收益与女生大致相同,补习也不会扩大男女生成绩差距。日本的补习与不补习的男生成绩相差 60 分,补习与不补习的女生成绩相差 65 分,补习与不补习男生的成绩差值比女生低 5 分,两者也没有显著性差异。在参与数学补习的日本学生中,女生数学成绩显著低于男生(低 19 分),而在没有参与数学补习的日本学生中,女生数学成绩比男生低 24.3 分,存在极其显著性差异。我们推测,与上海情况相反,在日本,数学补习给男生带来的成绩收益也与女生基本相同,但补习可能会缩小男女生成绩差距。

在分家庭经济社会地位的比较中,可以看出上海参与数学补习的来自家庭经济社会地位低于平均值的学生(以下简称低 SES 学生)比未参与补习的此类学生数学成绩高 20.3 分,而参与数学补习的来自家庭经济社会地位高于平均值的学生(以下简称高 SES 学生)比未参与补习的此类学生数学成绩仅仅高出 1 分,两类学生补习与不补习者的成绩差值存在显著性差异。在参与数学补习的上海学生中,高 SES 学生数学成绩高于低 SES 学生 60 分,而在没有参与数学补习的上海学生中,高 SES 学生数学成绩要高于低 SES 学生 79.4 分。因而,我们推测,数学补习给上海低 SES 学生会带来更大的成绩收益,而且可能会缩小高 SES 学生与低 SES 学生成绩差距。后面的计量模型将验证这一假设。

在日本参与数学补习的高 SES 学生比未参与补习的此类学生数学成绩高 58.4 分,参与数学补习的低 SES 学生比未参与补习的此类学生数学成绩高 48.2 分,但是两类学生补习与不补习者的成绩差值不存在显著性差异。在参与数学补习的日本学生中,高 SES 学生数学成绩高于低 SES 学生 44.2 分,而在没有参与数学补习的日本学生中,高 SES 学生数学成绩要高于低 SES 学生 34.1 分。因而,我们推测,数学补习给日本高 SES 学生会带来更大的成

绩收益，而且可能会扩大高 SES 学生与低 SES 学生成绩差距。这一假设也有待后面的计量模型的验证。

限于篇幅，此部分仅给出中国上海与日本各类不同群体补习与补习学生成绩差异的结果及分析，中国香港和韩国学生的成绩差异统计结果参见附表 1。

表 5 中国上海、日本不同群体补习与补习学生数学成绩差异

		中国上海				日本			
		数学补习	未参加数学补习	补习与不补习成绩差别	不同群体补习与不补习成绩差别的比较	数学补习	未参加数学补习	补习与不补习成绩差别	不同群体补习与不补习成绩差别的比较
性别	女	614.446 (93.476)	596.942 (109.792)	17.503*** (t=2.807)	-4.528 (t=-0.51)	547.875 (84.887)	482.687 (79.464)	65.188*** (t=10.617)	5.174 (t=0.68)
	样本量	1312	456			1463	564		
	男	624.164 (97.651)	602.132 (112.588)	22.032*** (t=3.040)		566.981 (89.086)	506.967 (94.698)	60.014*** (t=11.628)	
	样本量	1123	545			1438	694		
	女生-男生	-9.718** (t=2.446)	-5.189 (t=0.681)			-19.107*** (t=-4.091)	-24.280*** (t=-3.778)		
SES	高 SES	646.010 (88.178)	645.040 (104.066)	0.97 (0.142)	-19.348** (t=-2.23)	577.098 (83.885)	518.714 (90.031)	58.383*** (t=9.990)	10.136 (t=1.48)
	样本量	1320	422			1605	452		
	低 SES	585.946 (93.798)	565.628 (104.276)	20.319*** (t=3.319)		532.896 (85.004)	484.648 (85.978)	48.248*** (t=10.155)	
	样本量	1112	579			1266	784		
	高 SES-低 SES	60.064*** (t=10.390)	79.413*** (t=8.591)			44.202*** (t=8.620)	34.067*** (t=4.905)		

(四) 基于 RPSM 的补习效应分析

正如前文所述，匹配策略的优势在于经过某种倾向得分匹配后，实验组与对照组接近随机分配的结果，从而保证条件独立假设成立，以获得回归系数的因果解释。本研究采用再加权倾向得分匹配模型(RPSM)来揭示数学补

习带来的学生成绩因果效应。

表 6 基于再加权倾向得分匹配法的数学补习效应估计值

	处理组 样本量	控制组 样本量	ATET (标准误)	ATENT (标准误)	ATE (标准误)
中国上海	2432	1001	3.105 (23.576)	7.679 (57.017)	4.439 (16.660)
中国香港	1427	1608	-11.281 (34.457)	-9.911 (23.635)	-10.555 (14.224)
日本	2871	1236	50.529*** (5.983)	46.947*** (17.132)	49.451*** (4.194)
韩国	2168	1107	50.222*** (9.524)	49.550*** (17.894)	49.994*** (6.150)

统计结果表明，匹配后处理组和控制组各变量均值均不存在显著性差异，且通过了 LR 检验。中国上海、中国香港、日本、韩国学生样本的 LR 检验统计量的 P 值分别为 0.323、0.395、0.959、0.063，满足条件独立假设，表明匹配质量是比较好的，结果是可信的。

由表 6 知，中国上海学生数学补习的平均处理效应为 4.4 分，补习给真正参加了补习的学生所带来的成绩净效应为 3.1 分，比未参加补习的学生的补习效应要低 4.5 分，说明未参与补习的学生可以从补习中获益更大。而日本、韩国学生数学补习的平均处理效应分别为 49.5 分、50 分，补习给真正参加了补习的学生所带来的成绩净效应分别为 50.5 分、50.2 分，比未参加补习的学生的补习效应分别要高 3.6 分、0.7 分。显然，日本、韩国学生接受数学课外补习的收益要比上海学生高很多。这一结果表明，研究假设 2 对于日本、韩国学生接受课外数学补习是成立的，对于上海、香港接受课外数学补习的学生并不成立。香港学生的补习效应出现了负值，尽管三类效应值均不具有统计显著性。我们推测在不同国家/地区出现补习效应结果有较大差异的原因：一是上海、香港数学补习质量不高，不论“培优”还是“补差”都没有明显效果；二是上海的学生，尤其是高 SES 学生，能够得到的学校数学教育已经足够好，他们的数学学业成绩也已经很高（参见表 5），根据学习中的“高原现象”，上海高 SES 学生很可能已经处于学习高原期，即使通过课外补习也很难再使成绩获得显著提升（即存在“门槛效应”），即补习不会带来额外的好处，这与上海高 SES 学生家长更愿意将学生送入补习提高班的愿望是相悖的，本研究结果看不出上海学生参加数学补习这种投资的明显产出。而日

本、韩国多数学生还未进入学习高原期，借助补习可以获得掌握知识、技能的一些方法和技巧，补薄弱环节和知识漏洞，从而使其成绩明显提升。即补习可以给日本、韩国学生带来显著收益，无论是已经参与补习的学生，还是尚未参与补习的学生，均可从数学补习中获益颇丰，这也反映了日本、韩国课外补习投资的高回报，同时说明日本、韩国学生家长乐意投资子女参与课外补习是理智的。香港学生参与数学补习、科学补习、语言补习的比例均相对最低，可能也是部分香港家长理性决策的结果。另外，值得关注的是，上海学生未参与补习的学生的补习效应高于已经参与补习学生的效应。未参与补习的学生很可能多数属于低收入家庭，前文也提及高 SES 学生比低 SES 学生参与补习的比例高出近 10 个百分点(表 7 也表明，家庭社会经济地位是影响学生参与补习的重要因素)。如果这些学生能够参与补习，则可以提升他们的学业成绩。

由表 7 可以得出，性别、家庭社会经济地位、学校类型以及所在地区均是影响学生参与数学补习的重要因素，这与前文表 4 的统计结果一致，而且在四个地区和国家模型中，家庭经济地位变量的系数显著为正，也再次验证了研究假设 1，即来自家庭经济社会地位高的学生参与“影子教育”的几率显著高于来自家庭经济社会地位低的学生。

表 7 影响学生是否参与数学补习的 logit 模型

变量	中国上海	中国香港	日本	韩国
Gender	0.324*** (1.382)	0.156** (1.169)	0.247*** (1.280)	0.087 (1.091)
ESCS	0.296*** (1.345)	0.410*** (1.507)	0.710*** (2.033)	0.532*** (1.702)
Location			1.223*** (3.397)	0.166 (1.181)
Public				-0.203*** (0.816)
截距	0.852*** (2.345)	0.135** (1.145)	-0.382 (0.682)	0.593*** (1.810)

注：1. 表格中为 logit 模型的系数，括号里为机会比率。

2. 中国上海、中国香港、日本模型没有纳入学校类型变量，是由于如果加入该变量，则无法通过匹配后处理组与控制组的平衡性检验。

(五) 稳健性检验

表 8 给出了在相同控制变量下 OLS 与 RPSM 数学补习效应估计值的比

较,可以看出在四个国家/地区模型中两种方法补习效应的符号与显著性基本相同,仅是数值上有差异。说明RPSM的结果是基本可靠的。OLS估计效应的前提假设是:是否参与补习是完全随机的,而实际情况很可能是低能力的学生选择参加补习的几率更大,高能力学生选择不参加数学补习的几率更大(雷万鹏,2005;Yu Zhang.,2013),高能力学生可以通过自学获得更好的成绩。因而,OLS模型很可能遗漏了与参与补习呈负相关的能力变量而低估了补习收益。至于为何上海、香港学生的补习效应与日本、韩国学生有较大差异,下文将从补习时间和补习其他科目对数学成绩的影响效应的视角来进一步分析、阐释。

表8 在相同控制变量下OLS和RPSM模型中数学补习效应的估计值

	OLS	RPSM(ATET)
中国上海	0.332(3.274) [3433]	3.105(23.576) [2432/1001]
中国香港	-15.838*** (2.826) [3035]	-11.281(34.457) [1427/1608]
日本	13.468*** (2.964) [4107]	50.529*** (5.983) [2871/1236]
韩国	24.682*** (3.366) [3275]	50.222*** (9.524) [2168/1107]

注:1. 四个地区和国家OLS模型中的控制变量与表7中的自变量相同。

2. 括号“()”中的数据为标准误,OLS结果下方的方括号“[]”中的数据为样本量;RPSM结果下方的方括号“[]”中的数据为处理组/控制组的样本量。

(六) 补习时间对成绩的影响效应

为什么中国上海、中国香港、日本、韩国学生的数学补习效应存在较大差异?从前文描述性统计发现,四个地区和国家学生在各类补习的比例上存在较大差异,而且在补习时间上也存在明显差异,比如香港学生的数学、科学平均补习时间均低于其他三个地区和国家学生。不同学科的补习是否存在叠加效应还是挤出效应?我们拟采用两水平线性模型从补习时间的角度考察各国/地区学生补习效应的差异。

表 9 补习时间对数学成绩影响效应的两水平模型结果

变量	模型 1		模型 2	
	中国上海	日本	中国上海	日本
固定效应				
截距	192.184***	330.885***	201.503***	329.177***
学生水平变量				
Gender	-14.148***	-16.556***	-14.799***	-16.569***
Escs	7.716***	5.385**	7.641***	5.166**
Escs ²	-5.805***	0.870	-5.655***	0.714
Mathtime	1.107	8.025***	-0.038	6.344***
Mathtime ²	-0.336	-0.593*	0.003	-0.524
Langtime			-5.301***	0.567
Scitime			3.485***	-0.687
Othertime			1.547	2.261***
学校水平变量				
Public	-33.611***	-3.681***	-33.376***	-3.796***
Classsize	24.466***	7.843***	23.899***	7.958***
	-0.308***	-0.057***	-0.301***	-0.059
随机效应				
学校内	5172.398***	3876.623***	5122.177***	3864.359***
学校间	3431.134***	3280.015***	3237.169***	3272.459***
N	3433	4107	3425	4095

表 9 补习时间对数学成绩影响效应的两水平模型结果(续)

变量	模型 1		模型 2	
	中国香港	韩国	中国香港	韩国
固定效应				
截距	524.043***	741.693***	532.539***	735.406***
学生水平变量				
Gender	-23.166***	-10.015**	-22.267***	-10.737**
Escs	8.352***	17.646***	9.221***	17.402***
Escs ²	0.581	5.348**	1.106	5.364**

变量	模型 1		模型 2	
	中国香港	韩国	中国香港	韩国
Mathtime	-7.697**	13.709***	-10.002***	13.766***
mathtime ²	0.621	-0.562	0.876	-0.705
Langtime			-9.649***	0.359
Scitime			8.647***	-3.096
Othertime			0.792	3.804***
学校水平变量				
Public	31.150***	-17.033***	29.688***	-15.710***
Classsize	-1.175	-16.154***	-1.615	-16.043***
Classsize ²	0.083**	0.286***	0.087*	0.285***
随机效应				
学校内	5142.270***	5713.179***	4962.542***	5686.631***
学校间	3207.368***	2596.605***	3026.882***	2619.061***
N	3035	3275	3016	3238

注:***表示在 0.01 水平上显著,**表示在 0.05 水平上显著,*表示在 0.1 水平上显著。

从表 9 中模型 1 与模型 2 的结果比较看出,性别、家庭经济社会地位、学校类型、班级规模的系数值相差很小,说明模型的稳健性很好,研究结果的信度较高。在这里主要分析补习时间对学生数学成绩的影响效应。模型 1 和模型 2 的结果显示,数学补习时间对日本学生的数学素养成绩有显著的正向影响,但是影响效应呈倒 U 型曲线,即每周数学补习时间在 7 个小时左右,数学补习效应最大(目前日本学生参与数学补习的每周平均时间为 1.74 小时);韩国学生的数学补习时间对数学成绩的影响没有呈现显著的非线性影响,其一次项效应极其显著,单位补习时间的成绩效应高达 13.7 分。因此,关于补习时间对数学成绩产生显著正向效应的研究假设 3 在日本、韩国数据中得到了验证。日本与韩国学生参与数学补习单位时间的效应如此显著(且数值较大),也是这两个国家学生成绩的 RPSM 模型得到相比上海、香港较大的 ATT 结果的原因之一。而对上海学生样本的分析可以发现,数学补习时间对数学成绩没有显著的影响,但是语言补习时间对数学成绩有显著的负向影响,科学补习时间对数学成绩有显著的正向影响,这一结果与表 10 中的模型 2、3 的科学补习的系数符合相同,验证了科学学习与数学学习效果的“叠加效应”(additive effect)存在。香港学生的数学补习时间呈现极其显著的负向效应,我们推测这与参与补习的学生每周补习时间太短有关(平均 1.07 小

时)，只有提高每周补习时间到 6 小时以上才有可能产生实质性的成绩提升。另外，很可能是香港课外补习的数学教师水平较低，不能因材施教，有针对性地采取有效的提分策略。相对而言，香港学生参与科学补习获益更大，科学补习的单位时间产生的数学成绩效应高达 8.6 分。

(七)“影子教育”是否会扩大教育结果的不均等？

此部分我们将探讨“影子教育”是否会扩大教育结果的不均等问题，采用加入交互项的两水平线性回归模型来分析课外补习对于来自不同家庭经济社

表 10 课外补习对数学成绩影响效应的两水平模型结果

变量	模型 1		模型 2		模型 3	
	中国上海	日本	中国上海	日本	中国上海	日本
固定效应						
截距	191.859***	326.862***	196.190***	325.535***	200.650***	325.810***
学生水平变量						
Gender	-14.161***	-16.119***	-15.151***	-16.009***	-15.152***	-16.008***
Escs	7.594***	5.539***	7.916***	5.498***	13.713***	6.923*
Escs ²	-5.842***	0.887	-5.798***	0.911	-5.319***	1.207
Mathoutschool	0.026	13.480***	4.224	8.481**	0.876	8.087**
Langoutschool			-11.959***	5.500*	-12.638***	5.475*
Scioutschool			4.438	2.260	4.625	2.271
escs * mathoutschool					-7.774***	-1.985
学校水平变量						
Public	-33.603***	-4.197***	-33.600***	-4.479***	-33.511***	-4.484***
Classsize	24.461***	8.064***	24.276***	8.113***	24.180***	8.110***
Classsize ²	-0.308***	-0.059***	-0.306***	-0.060***	-0.305***	-0.060***
随机效应						
学校内	5175.638***	3890.851***	5157.777***	3885.222***	5146.864***	3884.697***
学校间	3427.266***	3491.201***	3302.115***	3480.021***	3290.641***	3482.441***
N	3433	4107	3428	4104	3428	4104

注：*** 表示在 0.01 水平上显著，** 表示在 0.05 水平上显著，* 表示在 0.1 水平上显著。

表 10 课外补习对数学成绩影响效应的两水平模型结果(续)

变量	模型 1		模型 2		模型 3	
	中国香港	韩国	中国香港	韩国	中国香港	韩国
固定效应						
截距	518.973***	746.548***	524.957***	746.078***	524.028***	746.206***
学生水平变量						
Gender	-22.741***	-10.076**	-22.051***	-10.128**	-22.071***	-10.136**
Escs	8.118**	18.980***	8.780***	18.919***	11.062***	18.333***
Escs2	0.466	5.728***	0.826	5.535**	1.167	5.439**
Mathoutschool	-15.777***	24.511***	-18.484***	23.588***	-21.434***	23.575***
Langoutschool			-17.490***	3.759	-17.749***	3.807
Scioutschool			21.544***	-2.207	21.571***	-2.200
escs * mathoutschool					-3.924	0.882
学校水平变量						
Public	30.847***	-17.300***	29.705***	-16.674***	29.640***	-16.688***
Classsize	-0.761	-15.669***	-1.063	-15.680***	-0.897	-15.692***
Classsize2	0.076*	0.280***	0.079*	0.280***	0.076*	0.280***
随机效应						
学校内	5139.319***	5771.229***	5012.810***	5774.342***	5009.332***	5774.030***
学校间	3233.064***	2664.185***	3079.023***	2686.767***	3081.177***	2686.811***
N	3035	3275	3025	3248	3025	3248

首先考察学生水平的自变量系数,研究发现:四个地区和国家学生的数学成绩均有显著的性别差异,男生比女生成绩表现好^①;家庭经济社会地位对日本、中国香港的学生数学成绩均有显著的正向影响,但是对中国上海学生来说,家庭经济社会地位对数学成绩的影响呈倒U型曲线,即当家庭经济社会地位低于某一值(0.649)时,随着家庭经济社会地位的提升,其对数学素养的影响也增大,但是高于这一值之后,随着家庭经济社会地位的提升,其

^① 在参与PISA2012测试的65个国家(地区)中,37个国家(地区)的男生数学成绩显著高于女生数学成绩,仅有5个国家女生数学成绩显著高于男生。2003年PISA测试结果与此类似,在41个参与国家(地区)中,27个国家(地区)男生数学成绩显著高于女生,11个国家(地区)男生数学成绩高于女生,但不具有统计显著性,仅有1个国家男生数学成绩显著低于女生。

对数学素养的影响反而减小^①，而对韩国学生的数学成绩的影响是正U型。参加数学课外补习对日本、韩国学生的数学素养成绩有显著的正向影响，对上海学生成绩有正向影响，但不显著，对香港学生的数学成绩有显著的负向效应。这一结果与前文的RPSM结果以及补习时间模型结果相吻合。参加语言课外补习对中国上海和香港学生的数学素养成绩有显著的负向影响^②，但是日本和韩国的结果恰恰相反，参加语言课外补习对日本和韩国学生的数学素养成绩正向影响，且对日本学生的影响具有统计显著性。我们推测，这很可能与上海学生每周平均语言补习时间高于日本学生有关(表3可知，高出约半个小时)，上海学生语言补习时间过长挤占了其数学学习时间，而且语言学习与数学学习的方式、方法以及思维特性都存在很大差异，不像数学与科学课程学习可以产生“叠加效应”(additive effect)，因而造成上述结果；参加科学补习对香港学生的数学成绩有显著的正向效应，对上海、日本学生的数学成绩也均有正向的影响，但是不具有统计显著性。这一结果也是这两大学科学习存在“叠加效应”的验证。

其次，考察学校层面的自变量系数，研究发现，学校类型对四个地区和国家学生的数学素养成绩均有显著的影响，中国上海、韩国公立学校学生的平均数学成绩分别比私立学校低33分、17左右，日本的差值则相对较小，公立比私立低4分左右，而中国香港正好相反，公立学校学生的平均数学成绩比私立学校高30分；班级规模对上海和日本学生的数学成绩均有显著的影响，并且均呈倒U型效应，上海的班级规模达到40人时，对学生数学成绩的正向影响最大，而日本的班级规模达到67人时，影响效应最大。上海平均班级规模约为39人，已经很接近最优规模状态，而日本学校的班级规模还有较大的提升空间。班级规模对香港与韩国学生数学成绩的影响与上海、日本不同，香港的班级规模还有提升空间(班级规模的二次项显著为正，一次项不显著)，韩国学校的班级规模达到56人时其效应值才为正。因而，韩国学校的班级规模也有较大提升空间。

最后，从模型3中家庭经济社会地位与参加数学课外补习的交互项的系数可以看出，参加数学课外补习可以减小家庭经济社会地位带来的成绩差异，

① 这一结果与任春荣和辛涛(2013)的结果基本一致。我们推测可能是在中国，家庭经济社会地位高的子女父母一般工作任务繁重，难以有时间陪伴和指导子女学习，甚至对子女的教育期望也不高，而且“富二代”或“官二代”的学习动力也可能不足，从而造成这类家庭的子女成绩与其父辈经济社会地位不相符的现象。

② 香港学生参与语言补习的平均时间每周仅0.44小时，可能难以起到实质性作用，反而会因占用时间而减少了学生数学学习时间，导致其数学成绩下降。

并且在上海这种缩减效应是极其显著的,大约可以缩小近8分的差距,在中国香港、日本分别可以缩小4分、2分的差距,韩国学生成绩模型的系数为正,也不具有统计显著性。上海学生成绩模型的这一结果推翻了“影子教育会扩大教育结果不均等”的研究假设4。课外数学补习对于上海高SES家庭学生与低SES家庭学生的数学成绩影响效应是存在异质性的,与前文描述性统计结果一致,补习给上海低SES家庭学生可以带来更大的成绩收益。课外补习可以缩小家庭经济地位带来的成绩差异,对于促进教育结果均等具有重要作用。上海低SES学生的数学平均成绩显著低于高SES学生的数学平均成绩(低62.6分),而且低SES未补习的学生平均成绩仅为565分,因而,这部分学生如果参与补习,基本属于“补差”类型的补习,可以获得较高的补习收益;而高SES学生未参与补习的平均成绩为645分,如果他们参与补习,则属于“培优”类型,不仅其补习收益微小,而且会加重他们的学习负担,挤占他们发展其他方面潜能的时间。近两年来,随着我国中小学减负政策的推行,中小学放学时间提前,家庭作业量大幅缩减,刺激了更多家庭和学生寻求课外补习,不同家庭背景的学生围绕“影子教育”系统的竞争日益激烈(薛海平等2014)。在此背景下,低SES家庭可能会选择让孩子参加课外补习活动,从而弥补父母在学业指导上的不足,缩小与其他儿童的学业成绩差距。本研究的结果证明,在中国上海,低SES家庭为孩子的课外补习投资决策是明智的,而高SES家庭为孩子的课外补习投资决策是不够明智的,课外补习的成绩收益甚微。由表10中模型1—3各控制变量的系数值可以看出,模型结果的稳健性很好,即研究结果的信度较高。

五、结论与讨论

综上所述,本研究基于PISA2012年中国上海、中国香港、日本以及韩国学生样本数据,采用RPSM模型和多水平线性回归模型分析了“影子教育”——课外补习对学生数学成绩的影响效应。主要得到以下研究结论:

第一,无论是中国上海学生还是日本、韩国学生,参与课外数学补习均能带来明显的成绩提升,而且,在日本、韩国,课外数学补习对学生数学成绩具有显著的正向效应。这一结论与OECD使用PISA2012年的数据得出的结论类似(OECD,2013),尽管OECD的PISA报告中使用的计量方法与我们的方法不相同。前文已述,中国、日本、韩国都受儒家文化影响,家长都希望子女能够考入理想的大学,以至将来有很好的工作和事业发展。借助补习可以帮助子女在学业上取得进步,因而他们愿意进行此方面的教育投资,

以使子女获得更高的人力资本，从而将来可以获得更高的人力资本投资回报。而且大量实践结果让他们看到，补习的确能够让学生成绩提升，因此，他们愿意支付高昂的补习费用，以换取子女在选拔性考试中的收益。

第二，日本、韩国学生的数学补习效应高于中国上海、中国香港学生。我们推测这是由于中国上海和日本、韩国学生的补习群体差异造成的，上海学生数学素养整体水平高于日本、韩国学生，上海参加补习的学生平均成绩已经高达618.9分，日本、韩国参加补习的学生平均成绩分别是557.3分、575.4分，因而上海学生的补习多数属于“培优”，而日本、韩国学生补习多数属于“补差”。由于学习进步规律中的“门槛效应”，上海参与补习的学生借助补习使得成绩获得较大提升的空间相对日本、韩国学生要小。另外，日本与韩国学生参与数学补习单位时间的效应较大，这既是日本、韩国课外补习教师水平相对较高的反映，也可能是这两个国家学生补习效应相比上海、香港学生更大的原因之一。香港学生课外数学补习效应为负，我们推测这与参与补习的学生每周补习时间太短(平均每周1.07小时)，难以产生实质性的成绩提升，另一方面也可能与香港课外补习市场中的数学教师对于正规学校课程中的考评内容理解不够，补习针对性不强，造成补习无效果。还有一种可能性，香港学生补习效应的RPSM模型中遗漏了关键的解释学生参与补习的特征变量，导致出现不符合实际的结果。由此，我们也推测，中国上海、中国香港学生在PISA2012中表现优异，排名名列前茅，是正规学校教育系统的作用，“影子教育”系统在其中的作用很有限，而日本、韩国学生表现突出，“影子教育”系统的作用相对较大。近年来韩国政府加大投资力度推行的“放学后教育计划”功不可没。

第三，不同学科的补习对数学成绩的影响不同。科学补习与数学补习存在“叠加效应”。无论是对中国上海、中国香港学生还是日本学生，参与科学类课外补习均能够起到促进其数学成绩提升的作用，而且对于中国上海、中国香港的学生来说，增加科学类课外补习时间，其成绩提升的效应是正向显著的。科学补习与数学补习产生“叠加效应”主要基于这两大类学科的相通性，不仅研究范式相近，而且其学习过程都离不开逻辑思维。相反，上海、香港学生参与语文补习对其数学成绩的影响是显著负向的，而日本学生从日语补习中仍能获得显著正效应。我们推测这是中国上海学生平均每周课外补习时间多于日本学生，语文课外补习会挤占其数学学习时间所致，而香港学生是每周语文补习时间过短，难以起到实质性作用，反而会因占用时间而减少了学生数学学习时间，导致其数学成绩下降。

第四，来自不同家庭经济地位的学生参与补习的机会存在较大差异。

四个地区和国家的高 SES 学生参加课外数学补习的比例均显著高于低 SES 学生。中国上海两者相差 10 个百分点左右，中国香港、日本甚至达到 17 个百分点。前文已述，无论是在中国，还是在日本、韩国，课外补习费用对于低收入家庭来说，都是相当大的开支。高昂的课外补习费用对于家庭经济收入较低的家庭来说是难以负担的，因而很可能会影响低收入家庭对子女的课外补习决策。

第五，参加数学课外补习可以缩减家庭经济社会地位带来的成绩差异，从而起到促进教育结果均等的作用。家庭经济社会地位高于平均值的学生的成绩要显著高于家庭经济社会地位低于平均值的学生，中国上海、中国香港、日本、韩国这两类学生数学成绩分别约差 63 分、41 分、49 分、53 分。参照联合国教科文组织统计研究所(UIS)开发的性别不平等指数，我们给出的四个地区和国家的家庭背景不平等指数均超过了平等范围(大于 1.03)，性别不平等指数均在平等范围(参见附表 2)。因此，有必要采取教育干预措施缩小不同家庭背景学生的学业成绩差距。

这种主要源于家庭背景的差异导致的学生学业成绩的巨大差异可以借助课外补习而缩小。在上海这种缩减效应是极其显著的，大约可以缩小近 8 分的差异。这一结论可能与社会以及媒体的共识相左，人们通常会认为课外补习是富裕阶层才能为子女做出的投资选择，普通收入阶层由于难以支付高额的补习学费而少有家庭会为子女做此方面的投资。因而，课外补习会扩大教育结果不均等，甚至会扩大社会不平等。然而，从上海 2009 年和 2012 年的 PISA 测试群体的数据来看，接受数学课外补习的比例很大(70%左右)，即使来自较低家庭经济社会地位的学生选择数学补习的比例仍超过了 60%。“中国义务教育阶段城市家庭教育成本研究”课题组的调查结果显示，76%的家庭为子女选择了课外补习。需要提及的是，根据上海学生补习效应的 RPSM 模型结果发现，上海学生未参与补习的学生的补习效应高于已经参与补习学生的效应。未参与补习的学生很可能多数属于低收入的家庭，前文也提及高 SES 的学生比低 SES 学生参与补习的比例高出近 10 个百分点(表 7 也表明，家庭社会经济地位是影响学生参与补习的重要因素)。如果这些学生能够参与补习，则可以明显提升他们的学业成绩，缩小与高 SES 学生的学业成绩差距，从而使得课外补习成为人力资本积累的“平等器”。从这一角度来说，政府和教育政策制定者应当看到教育补习的合理性成分，赋予课外补习以“合法”地位(胡咏梅等，2015)。

本研究发现，尽管在不同国家或地区，数学课外补习的效应不尽相同，但课外补习确实可以提高学生的学业成绩，而且参与数学课外补习可以缩减

家庭经济社会地位产生的成绩差异，对于促进教育结果均等来说具有一定作用(在中国上海，这一缩减效应是显著的)。一些研究发现，在东亚国家，课外补习已经成为低收入家庭教育支出的必需项目，课外补习支出占低收入家庭收入的比重相当大，甚至会影响他们的日常生活水平(Bray & Kwok, 2013; Lee et. al., 2009; Lin & Chen, 2006)。因此，为了充分发挥“学校教育是社会平等器”^①的作用，我们建议政府和学校为家庭经济收入较低且学业成绩较低的学生提供必要的课外补习机会，免收其补习费用。效仿韩国政府推行的“放学后教育计划”，学校利用放学后的时间，有专职教师为低收入家庭学生提供“补差型”的学术课程补习，中央和地方政府为此提供专项经费支持。此外，政府可以通过减免教育补习集团税收的方式鼓励和倡导这些补习机构提供具有公益性质的教育服务，为来自收入特低的家庭学生减免学费，为进步较大的贫困生提供奖学金，从而达成政府、学校和社会共同促进义务教育阶段的教育结果均等的局面(胡咏梅等，2015)。

[参考文献]

- 安格里斯特、皮施克，2012：《基本无害的计量经济学——实证研究者指南》，上海：格致出版社。
- 贝磊，2007：《教育补习与私人教育成本》，北京：北京师范大学出版社。
- 何瑞珠、卢乃桂，2011：《从国际视域析教育素质与平等》，北京：教育科学出版社。
- 胡咏梅、范文凤、丁维莉，2015：《影子教育是否扩大教育结果的不均等——基于PISA2012上海数据的经验研究》，《北京大学教育评论》第3期。
- 雷万鹏，2005：《高中生教育补习支出：影响因素及政策启示》，《教育与经济》第1期。
- 李佳丽、胡咏梅、范文凤，2016：《家庭背景、影子教育和学生学业成绩成就：基于Wisconsin模型的经验研究》，《教育经济评论》第1期。
- 薛海平，2015：《从学校教育到影子教育：教育竞争与社会再生产》，《北京大学教育评论》第3期。
- 薛海平，2016：《课外补习、学生成绩与社会再生产》，《教育与经济》第2期。
- 薛海平，2017：《家庭资本与教育获得：影子教育的视角》，《教育科学研究》第2期。
- 薛海平、丁小浩，2009：《中国城镇学生教育补习研究》，《教育研究》第1期。
- 薛海平、王东、巫锡炜，2014：《课外补习对义务教育留守儿童学业成绩的影响研究》，《北京大学教育评论》第3期。

^① 学校教育是人们突破社会出身等的制约，向上流动、改善经济生活的开放路径。人力资本理论尤其强调，在完全竞争的市场条件下，所有人都可以通过投资教育来提高未来收入，从而促进经济平等。

- Bray, M. and O. Kwok, 2013, "Behind the Facade of Fee-free Education: Shadow Education and its Implications for Social Justice", *Oxford Review of Education*, 39(4): 480—497.
- Bray, M. and P. Kwok, 2003, "Demand for Private Supplementary Tutoring: Conceptual Considerations, and Socio-economic Patterns in Hong Kong", *Economics of Education Review*, 22(6): 611—620.
- Bray, M. , 2005, "Private Supplementary Tutoring: Comparative Perspectives on Patterns and Implications", Sept. 13—15. Oxford: International Conference on Education and Development.
- Bray, M. , S. Zhan, C. Lykins, D. Wang and K. Ora, 2014, "Differentiated Demand for Private Supplementary Tutoring: Patterns and Implications in Hong Kong Secondary Education", *Economics of Education Review*, 38: 24—37.
- Caritas, Community and Higher Education Service, 2010, *Private Supplementary Tutoring of Secondary Students: Investigation Report*, Hong Kong: Caritas.
- Kim, K. , 2010, "Educational Equality," In Lee, C. J. , S. Kim and D. Adams (Eds.), *Sixty Years of Korean Education*, Seoul: Seoul National University Press, pp. 285—325.
- Lee, C. J. , H. J. Park, and H. Lee, 2009, "Shadow Education Systems". In G. Sykes, B. Schneider and D. N. Plank (Eds.), *Handbook of Education Policy Research*, New York: Routledge for the American Educational Research Association. pp. 901—919.
- Li, D. and M. C. Tsang, 2003, "Household Education Decisions and Implications for Gender Inequality in Education in Rural China", *China: An International Journal*, 1(2): 224—248.
- Lin, D. S. and Y. F. Chen, 2006, "Cram School Attendance and College Entrance Exam Scores of Senior High School Students in Taiwan", *Bulletin of Educational Research (Taiwan)*, 52(4): 35—70, (in Chinese).
- Ministry of Education and Training, 2008, *Report on the Situation of Academic Learning Activities of Children*, Tokyo: MonbukagakushoHokokusho, (In Japanese).
- Murawska, B. and E. Putkiewicz, 2006, "Poland", In Silova, I. , V. Būdiene and M. Bray (eds.), *Education in a Hidden Marketplace: Monitoring of Private Tutoring*, New York: Open Society Institute, pp. 257—277.
- Ngai, A. and S. Cheung, 2010, *Students' Participation in Private Tuition*, Youth Poll Series No. 188, Hong Kong: Hong Kong Federation of Youth Groups.
- OECD, 2013, *PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices*, Vol. IV: 112—115.
- Rosenbaum, P. R. , and D. B. Rubin, 1983, The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika*, 70: 41—55.

- Smyth, E. , 2008, ‘The More the Better? Intensity of Involvement in Private Tuition and Examination Performance’, *Educational Research and Evaluation*, 14 (5): 465—476.
- Smyth, E. , 2009, ‘Buying your Way into College? Private Tuition and the Transition to Higher Education in Ireland’, *Oxford Review of Education*, 35 (1): 1—22.
- Southgate, D. E. , 2009, “Determinants of Shadow Education: A Cross-national Analysis”, Unpublished Doctoral Dissertation, The Ohio State University.
- Stevenson, D. L. and D. P. Baker. 1992, “Shadow Education and Allocation in Formal Schooling: Transition to University in Japan”, *American Journal of Sociology*, 97(6): 1639—1657.
- Tansel, A. and F. Bircan, 2006, “Demand for Education in Turkey: A Tobit Analysis of Private Tutoring Expenditures”, *Economics of Education Review*, 25(3): 303—313.
- Tsang, M. C. , 2002, “Comparing the Costs of Public and Private Schools in Developing Countries”, In Levin, H. and P. McEwan (Eds.), *Yearbook of the American Education Finance Association*, New York: Eye on Education.
- Verdis, A. , 2002, “School Effectiveness Research for Educational Evaluation in Greece”, PhD thesis, Institute of Education, University of London.
- Zhang, Y. , 2013, “Does Private Tutoring Improve Students’ National College Entrance Exam Performance?: A Case Study from Jinan, China”, *Economics of Education Review*, 32: 1—28.

Does “Shadow Education” Enlarge the Inequality of Educational Outcomes?

: An Empirical Study Based on PISA 2012 data from China, Japan and Korea

HU Yong-mei¹, FAN Wen-feng², DING Wei-li³

(1. Faculty of Education/Capital Institute for Economics of Education, Beijing Normal University;

2. Educational Supervision and Quality Assessment Research Center,
Beijing Academy of Education Science;

3. School of Policy Studies and Department of Economics, Queen’s University)

Abstract: The “shadow education” system of private supplementary tutoring has become quite common in East Asian countries nowadays. Based on the data of Programme for International Student Assessment 2012 (PISA 2012), the paper analyzes the influence of shadow education on the mathematical literacy of students of Shanghai, Hong Kong, Japan

and Korea by means of a hierarchical linear model, and estimates the net effect of shadow education using the method of reweighting on propensity score matching (RPSM). The following findings are obtained from the study: first, supplementary math tutoring has a significant positive effect on the math score of students, and such an effect is more obvious on Japanese and Korean students than on Hong Kong and Shanghai students; second, attending supplementary math tutoring may narrow the gap between students in learning performance that is caused by the difference in their families' economic, social and cultural status (ESCS), thus promoting the equality of educational outcomes.

Key words: “shadow education”; inequality of educational outcomes; Programme for International Student Assessment (PISA); reweighting on propensity score matching (RPSM)

(责任编辑: 刘泽云 责任校对: 刘泽云 孙志军)

附表 1: 中国香港与韩国不同群体补习与不补习学生成绩差异

		中国香港				韩国			
		数学补习	未参加数学补习	补习与不补习成绩差别	不同群体补习与不补习成绩差别的比较	数学补习	未参加数学补习	补习与不补习成绩差别	不同群体补习与不补习成绩差别的比较
gender	女	553.662 (84.467)	554.047 (95.755)	-0.385 (t=-0.067)	-0.965 (t=-0.127)	560.404 (89.206)	513.631 (94.902)	46.773*** (t=7.100)	-28.664*** (t=-3.251)
	样本量	681	723			1013	503		
	男	570.909 (96.719)	570.329 (102.672)	0.580 (t=0.105)		588.641 (95.652)	513.203 (101.713)	75.438*** (t=10.269)	
	样本量	753	900			1157	606		
	女生与男生成绩差别	-17.247** (t=-2.221)	-16.282** (t=-2.385)			-28.237*** (t=-4.193)	0.427 (t=0.051)		

		中国香港				韩国			
		数学补习	未参加 数学补习	补习与不 补习成绩 差别	不同群体 补习与不 补习成绩 差别的 比较	数学补习	未参加 数学补习	补习与 不补习 成绩差别	不同群体 补习与不 补习成绩 差别的 比较
SES	高 SES	580.102 (89.134)	591.430 (96.475)	-11.328* (t=-1.920)	-6.931 (t=-0.909)	595.341 (92.033)	540.900 (102.153)	54.441*** (t=8.490)	-2.008 (t=-0.239)
	样本量	808	628			1237	476		
	低 SES	540.199 (89.678)	544.596 (97.190)	-4.397 (-0.895)		549.391 (89.333)	492.942 (90.849)	56.450*** (t=7.797)	
	样本量	619	980			931	631		
	高 SES 与低 SES 学生 成绩 差别	39.903*** (t=6.066)	46.834*** (t=6.446)			45.950*** (t=8.368)	47.958*** (t=6.691)		

附表 2：性别不平等指数与家庭背景不平等指数

	女生数学 平均成绩	男生数学 平均成绩	性别不平等 指数(GPI) ¹	高 SES 学生 平均数 学成绩	低 SES 学生 平均数学 成绩	家庭背景 不平等 指数 ²
中国上海	609.888	615.607	0.991	643.246	580.588	1.108
日本	527.011	544.884	0.967	563.015	514.120	1.095
中国香港	552.957	568.378	0.973	584.382	543.451	1.075
韩国	544.193	562.114	0.968	579.537	526.501	1.101

注：1. 性别不平等指数(GPI, Gender Parity Index), 为联合国教科文组织统计研究所(UIS)用于考察某一指标的性别均等状况, 计算方法为女性指标数值与男性指标数值的比值, 数据越偏离 1, 越不平等, 0.97 至 1.03 为性别平等范围。该指数在 2010、2011 年的全球教育概览中均有使用。

2. 我们参照性别不平等指数计算家庭背景不平等指数。