

班级规模、同伴特征与初中生学业能力

——来自中国经验数据的证据

曹淑江，谢立艳

[摘要]缩小班级规模是国内外广泛关注的一种提高教学质量并促进教育公平的方式。本文基于中国教育追踪调查(CEPS)数据,采用多层次模型研究班级规模对初中生学业能力的影响。在优质教育资源匮乏的情况下,不同类型的班级生源质量存在明显的差异,大班往往拥有更好的生源,这是导致出现班级规模正向效应估计偏差的主要原因。本研究尝试讨论不同类型班级的班级规模影响学业能力的机制。研究发现,班级规模与学业能力之间存在“倒U型”的非线性关系。中介效应分析结果表明不同类型班级的班级规模影响学业能力的机制不同,其中非大班的班级规模产生正向效应的主要作用路径是同伴效应,但是当班级规模增加到一定程度后,拥挤效应成为班级规模负向效应的主要作用机制。另外,班级规模对不同类型学生的影响存在异质性,适度增加班级人数对所有分位点的学生均产生显著的正向影响,且影响程度基本无差异。然而,当班级规模极度扩张时,学业能力较强的学生对拥挤带来的负向效应更敏感。

[关键词]班级规模;同伴效应;学业能力;多层次模型;CEPS

一、问题的提出

随着我国经济社会的快速发展以及城镇化进程的加快,城乡居民对优质教育资源的需求日益增长,对发展公平且高质量的教育提出了巨大的挑战。如何通过政策调控实现从“有学上”到“上好学”的转变成为现阶段我国基础教育改革的难题。不同于激励性的教育投入方式,缩小班级规模是一种增加资源的教育投入方式。作为衡量学校教育资源配置的重要指标,班级规模成为国内外学者以及政策制定者广泛关注和讨论的话题。在以班级授课为主要教

[收稿日期] 2023-01-20

[作者简介] 曹淑江,中国人民大学教育学院,电子邮箱地址:caosj@ruc.edu.cn;
谢立艳(通讯作者),中国人民大学教育学院,电子邮箱地址:xieliyan@ruc.edu.cn。

学组织形式的学校教育中,班级是学生接受学校教育的基本组织形式,提供了以班级为单位开展教育活动的物理空间,同时限定了班级管理方式、课堂教学模式、师生互动频率以及同伴关系等,并通过错综复杂的作用机制对学生的认知能力和非认知能力产生影响。

为了进一步满足人民对高质量教育的需求,缩减班级规模成为我国提升基础教育质量的重要政策工具,近几年,我国先后出台了诸多消除“大班额”的政策文件。《国家教育事业发展“十三五”规划》指出,“到2020年基本消除56人以上大班额现象”。《国务院关于统筹推进县域内城乡义务教育一体化改革发展的若干意见》以及教育部的相关部署,要求到2018年基本消除66人以上“超大班额”,2019年控制“大班额”要取得突破性、决定性的进展,到2020年基本消除56人以上的大班额。2018年,李克强总理在政府工作报告中指出,“要发展公平而有质量的教育”,并明确提出要“抓紧消除城镇‘大班额’”。2019年出台的《中共中央 国务院关于深化教育教学改革全面提高义务教育质量的意见》要求“加快消除城镇大班额,逐步降低班额标准”。2021年12月教育部等九部门印发的《“十四五”学前教育发展提升行动计划》和《“十四五”县域普通高中发展提升行动计划》指出,“逐步化解和消除学前教育‘大班额’现象”,并明确提出“各地要加大消除普通高中大班额专项规划实施力度,全面消除56人及以上大班额”。通过以上政策文件的梳理不难看出,消除大班额成为我国班级规模调整的共同指向。

国内外学者对班级规模与学业能力的关系进行了大量的理论和实证研究。但是班级规模与学业能力究竟呈现怎样的关系目前尚无定论。实验研究发现较大的班级规模对学业成绩存在负向影响。美国田纳西州的生师比成绩项目(Student-Teacher Achievement Ratio,简称STAR)显示,小班学生的学业成绩明显高于大班学生,所以班级规模对学业成绩存在显著的负效应。英国班级规模与生师比项目(Class Size & Pupil Adult Ratio Project, CSPAR)对儿童的追踪实验结果也表明班级规模越小对学业成绩的提高越有利。Angrist和Lavy(1999)运用准实验的方法对以色列的班级规模效应进行研究,发现班级规模效应对四年级和五年级的学生显著为负。另一类研究的发现与上述结论恰好相反,认为班级人数增加对学业成绩有利,这样的观点与人们的主观预期相悖。例如,研究发现现阶段我国班额与学生学业成绩呈显著的正相关关系(方征,2015)。除以上两种结论外,也有研究表明,班级规模与学业成绩之间没有显著的关系(Leuven et al., 2008; Hanushek, 1999; 约翰·哈蒂,2015)。除了对班级规模与学业能力的线性关系的探究外,也有研究关注班级规模与学业能力的非线性关系(Borland et al., 2005; Ding and Lehrer,

2010; Angrist et al., 2019; Kedagni et al., 2021)。

另外,大量的研究探索了班级规模对学业能力的影响机制,主要包括同伴效应、教师特征以及课堂互动等。中国义务教育阶段的同伴效应对学生存在一定的负面影响,这主要体现在成绩较差的学生当中,而成绩较好的学生能与同层次群体的互动效应得到自我提升(赵颖,2019)。但国外也有研究表明对学业成绩较差的群体,同伴的正向效应更显著(Levin,2001)。有研究发现教师学历与班级规模对学生成绩的影响具有交互作用,缩小班级规模对高学历教师教学水平的发挥更有利(Barrett and Toma,2013)。但也有研究表明教师学历与班级规模可以独立作用于学生成绩,不存在交互效应(Brühwiler and Blatchford,2011;姚昊等,2021)。课堂互动也是班级规模起作用的重要机制,缩小班级规模能够提高学生的课堂参与程度(Finn et al.,2003)。

还有一批研究关注班级规模在城乡、性别、学段和学生类型等维度的异质性影响。例如,班级规模对初中生学业能力的异质性影响与不同行政区域和学校的教育发展、学校布局及家庭选择高度相关(崔盛和吴秋翔,2019)。最佳班级规模与班内学生的自身表现密切相关,当学生表现较差时,最佳班级规模变小(Lazear,2001)。也有研究探讨了班级规模效应在不同教育阶段的异质性,高中阶段的班级规模越大越好(郑琦和杨朴,2018;Addonizio and Phelps,2000)。STAR项目研究发现缩小班级规模对低年级和弱势群体更有利(Jepsen and Rivkin,2009;Angrist and Lavy,1999)。

通过文献梳理发现,尽管学界对班级规模这一主题开展了大量研究,但无论是关于班级规模与学业能力的关系,还是班级规模效应的影响机制,均存在诸多分歧。尤其是不同国家关于班级规模效应的经验证据大相径庭。基于PISA数据的跨国比较结果显示,学生学业能力排名靠前的国家和地区反而拥有较大的班级规模。Blatchford等人(2016)比较了东西方班级规模缩减的政策背景和政策目标,分析发现,东西方政策制定的目标存在明显差异。西方国家缩减班级规模的主要目标是提高学生的学业成绩,而对于包括中国在内的东亚地区,班级规模的缩减是在素质教育改革的大背景下提出的,是深化课程教学改革的配套措施,主要目标是克服应试教育的弊端,促进教学模式的转变。但是,通过控制班级规模促进教育机会公平和过程公平是东西方国家的共同价值指向。

研究班级规模对学生学业能力的影响,一个必要前提是学生的分班是随机的,但在现实中,学校往往是家庭选择的结果,因此这一前提条件并不成立。已有研究发现学生的分班并不是随机的,不同班级规模中的学生特征不同(Hoxby,2000;Angrist and Lavy,1999)。按照中国义务教育政策和相关

法律法规,小学升初中是不需要升学考试的,大多数为免试就近入学,但是民办初中和部分公办重点初中依然举办小升初的选拔性考试,除此之外,还存在非制度性的转学情况。目前义务教育处于优质教育资源匮乏的阶段,口碑好的学校供不应求,好学校能够吸引更多的优质生源,而且招生名额十分紧缺。在没有政府对班级规模上限约束的情况下,资源的紧缺性以及学校的逐利行为必然会导致大班额的出现,而且大班更多出现在好学校,这就意味着大班的生源往往优于非大班的生源。因此,如果在分析班级规模效应时不区分班级类型,得出正向影响的结论可能是一种有偏误的结果。

初中阶段不仅是义务教育的最后一个阶段,也是学业分流的关键阶段。由于不同学段的班级规模效应往往存在差异,本文聚焦于初中阶段的班级规模效应。经过近10年人口结构的变化以及班级规模的调整,初中阶段班级规模的控制取得了阶段性的成果,平均班级规模从2011年的51.83人下降到2020年的45.78人,无论是城市、县镇还是农村都已经基本消除了56人以上的大班额,班级规模基本维持在36—55人。^①但这只是平均水平意义下的班级规模的统计结果,现实中的班级规模依然表现出显著的差异,而且与西方国家相比,我国目前的班级规模仍较大。

通过前文对文献的梳理发现,基于不同数据的班级规模效应的研究,研究结果往往不同。我们分析主要可能存在两方面的原因:一方面,不同国家,或者同一个国家不同年份的班级规模基数不同,可能导致班级规模效应产生差异。比如Angrist等运用相同的研究方法,分别用以色列1991—1992年、2002—2011年五年级的学生数据对班级规模效应进行研究,研究结论却完全不同:1991—1992年的数据结果表明班级规模效应显著为负,而2002—2011年的数据结果显示并不存在显著的班级规模效应。数据对比发现,1991年班级规模的中位数为31,而后面一组数据的中位数为28,这很有可能是造成结果差异的原因(Angrist and Lavy, 1999; Angrist, et al., 2019)。因此,结合我国现阶段较大班级规模的事实,本文假设班级规模与学业能力之间存在非线性关系。

另一方面,基于不同国家经验数据的研究结果不同,也有可能反映了教育生产函数的差异。由于不同国家,尤其是东西方国家的教育体制、社会文化背景、课堂教学模式等存在较大差异,班级规模对学业成绩的影响机制可能完全不同。我国班级规模基数较大,课堂教学主要以讲授为主,每个学生都面临中考升学压力。在课堂教学模式、师资质量以及培养目标相对稳定的

① 资料来源:《中国教育统计年鉴》2011—2020年数据。

情况下,本文假设同伴效应是班级规模影响学业成绩的主要作用机制。此外,本文探索了班级规模对不同类型学生的异质性影响,这对通过调整班级规模和分班决策促进教育公平有着重要意义。

在消除大班额的政策背景下,本文尝试回答以下问题:消除大班额的举措是否科学有效?在大班额得到控制的情况下,继续缩小班级规模是否对学生学业能力的发展有利?班级规模效应产生的主要作用机制是什么?班级规模对不同学生群体的影响是否存在差异?这些问题的研究对班级规模的政策调整以及基础教育阶段的综合改革具有启示作用。

二、研究设计

(一)模型设定

根据 Hanushek 建立的教育生产函数(Hanushek, 1986),本文构建如下理论模型:

$$cog_{ij} = f(cln_{ij}, I_{ij}, S_j, \epsilon_{ij}) \quad (1)$$

其中 cog_{ij} 表示学校 j 中学生 i 的学业能力, cln_{ij} 表示班级规模, I_{ij} 表示个体层面的变量,包括个人特征变量和家庭特征变量, S_j 表示除班级规模外的学校的其他特征变量, cln_{ij} 和 S_j 属于学校层面的变量, ϵ_{ij} 是扰动项。

鉴于观测数据的结构是分层的,学生个体特征嵌套于班级,班级又嵌套于学校,同属一所学校的学生个体之间的相关性大于来自不同学校的学生个体之间的相关性,学生个体层面的变量受学校层面变量的影响,观测样本不再满足独立同分布和方差恒定的假定,如果使用经典的线性回归模型,就会得到有偏的参数估计和错误的统计推断结果。因此本文采用多层次模型,该模型是由不同层次的自变量解释同一变量的一体化模型,允许放松独立性假设及相关误差结构的存在。由于样本数据仅在每所学校中抽取一个或两个班级,所以不适合构建学生、班级、学校的三层模型。本文构建学生和学校的两层变量,将学校特征变量置于第二层,将学生特征变量置于第一层,并把班级规模作为学校层面的变量。假设班级规模对学业能力的影响在不同的学校是相同的,故采用随机截距模型。分层计量模型如下:

$$\text{第一层: } cog_{ij} = \beta_{0j} + \beta_{1j} I_{ij} + \epsilon_{ij} \quad (2-1)$$

$$\epsilon_{ij} \sim N(0, \sigma^2)$$

$$\text{第二层: } \beta_{0j} = \gamma_{00} + \gamma_{01} cln_{ij} + \gamma_{02} S_j + \mu_{0j} \quad (2-2)$$

$$\beta_{1j} = \gamma_{10} \quad (2-3)$$

$$\mu_{0j} \sim N(0, \tau_{00})$$

将(2-2)与(2-3)代入(2-1)得到混合模型方程:

$$cog_{ij} = \gamma_{00} + \gamma_{10}I_{ij} + \gamma_{01}clsn_{ij} + \gamma_{02}S_{ij} + \mu_{0j} + \varepsilon_{ij} \quad (2)$$

学生学业能力的方差为组内方差和组间方差之和: $\text{Var}(cog_{ij}) = \tau_{00} + \sigma^2$ 。为了评估多层次模型的必要性,利用组内相关系数(ICC)来估计第二层的方差贡献,其定义为: $\text{ICC} = \frac{\tau_{00}}{\tau_{00} + \sigma^2}$ 。计算可得,学校层面的组间相关系数为0.298,这表明因变量中29.8%的方差在第二层,70.2%的方差在第一层。根据Cohen(1988)建议的判断准则,当 $\text{ICC} > 0.138$ 时,因变量与组间是高度关联强度。因此,较高的组间相关系数表明了使用多层次模型的必要性。

在单层次模型中, R^2 表示因变量的变异中有多少是由整个模型解释的。而在多层次模型中,则可以通过估计零模型和全模型来构建一个类似的度量,分别计算第一层和第二层的方差中有多少是由其包含的自变量解释的。本文主要研究学校层面的班级规模变量对学生学业能力的影响,所以重点关注第二层。第二层的计算公式为: $R^2 = \frac{\tau_{00}^{null} - \tau_{00}}{\tau_{00}^{null}}$,其中 τ_{00}^{null} 为零模型第二层次的方差, τ_{00} 为加入解释变量后的第二层次的方差。

(二)数据来源与变量选择

本文基于中国教育追踪调查(China Education Panel Survey, CEPS)初中一年级2013—2014学年的基线调查和2014—2015学年的追踪调查数据进行实证检验。在数据处理过程中,异常值是影响实证分析结果的重要因素,本文将被解释变量和核心解释变量中的缺失值进行了剔除,有效样本为9050人,涉及221个班级。

学校是学生和家庭选择的结果。一方面,个体特征和家庭特征会影响学生择校,同时这些特征也是影响学生学业能力的重要因素;另一方面,学生既是学校教育的消费者,同时也是生产要素。因此在模型设定时,为了减少内生性问题带来的估计偏误,本文加入了一系列控制变量。

模型中的被解释变量为学生上初二时的学业能力,核心解释变量为学生初一时所在班级的班级规模。综合我国各地对班级规模的限定标准以及已有文献对班级类型的划分,本文将40人以下的班级定义为“小班”,56人以上的班级定义为“大班”,其他班级定义为“中班”。数据中共包含31个小班、117个中班、73个大班。由于学生上初中之前,已经接受了六年的小学义务教育,初中入学时学生个体的学业能力存在差异,因此本文把基准期的学业能力作为核心控制变量。CEPS数据中的认知能力测试反映了学生的认知水平,该测试具有国际可比性、全国标准化的特点,测试得分经过3PL模型标

准化，基准期与追踪调查的测试得分具有可比性，因此本文选择两期数据的认知能力测试的3PL标准化成绩作为学生学业能力的代理变量。

由于影响学生学业能力的因素具有明显的分层特点，因此变量中应包含学生个体层面、家庭层面以及学校层面的信息。个体特征主要包括性别、是否独生子女、是否为农业户口、是否住校等。家庭特征主要指家庭文化资本和家庭经济资本，本文分别选取父母的受教育程度及家庭经济状况作为两类家庭资本的代理变量。学校特征主要包括学校所在地区、学校质量、是否按成绩分班、同伴能力等，同伴能力主要包括同伴学业能力的均值和方差。表1是对主要变量的描述性统计。

表1 主要变量的描述性统计

变量	小班	中班	大班	总体
学业能力	0.196 (0.899)	0.275 (0.845)	0.442 (0.730)	0.303 (0.832)
班级规模	33.07 (5.111)	47.85 (4.251)	65.31 (6.799)	49.69 (11.73)
基准学业能力	0.011 (0.875)	-0.001 (0.874)	0.133 (0.845)	0.034 (0.869)
性别	0.524 (0.500)	0.520 (0.500)	0.523 (0.500)	0.522 (0.500)
是否独生子女	1.465 (0.499)	1.584 (0.493)	1.545 (0.498)	1.554 (0.497)
是否为农业户口	0.471 (0.499)	0.548 (0.498)	0.506 (0.500)	0.524 (0.499)
是否住校	0.183 (0.387)	0.330 (0.470)	0.325 (0.469)	0.304 (0.460)
家庭文化资本	4.594 (2.042)	4.530 (2.006)	4.781 (2.066)	4.604 (2.030)
家庭经济资本	1.911 (0.482)	1.834 (0.499)	1.840 (0.494)	1.848 (0.496)
学校区位	2.942 (1.615)	2.736 (1.596)	1.993 (1.205)	2.585 (1.551)
学校质量	1.847 (0.571)	1.998 (0.618)	2.257 (0.666)	2.037 (0.638)

续表				
变量	小班	中班	大班	总体
分班方式	0.910 (0.287)	0.788 (0.409)	0.925 (0.263)	0.845 (0.362)
同伴能力均值	0.011 (0.488)	-0.001 (0.490)	0.133 (0.462)	0.034 (0.486)
同伴能力方差	0.729 (0.121)	0.724 (0.106)	0.706 (0.097)	0.721 (0.107)
观测数	1610	5094	2346	9050
百分比(%)	17.79	56.29	25.92	100

注：表中每个变量对应的第一行数值为均值，括号内为标准差。

三、研究发现

(一) 班级规模对学业能力的影响

图1显示了不同类型班级学生的基准学业能力的分布情况。小班和中班学生的能力分布情况相似，但大班学生的能力分布则表现出明显的差异，大班当中认知水平较低的学生所占比例明显低于小班和中班，而认知水平较高的学生所占比例明显高于小班和中班。由于大班的生源特征存在明显的差异，而小班和中班的生源特征类似，下文将不再区分小班和中班，将两类班级统称为“非大班”。大量研究表明，同伴效应是班级规模影响学业能力的重要作用机制，因此分别对加入同伴特征和不加入同伴特征的两种情况进行检验。

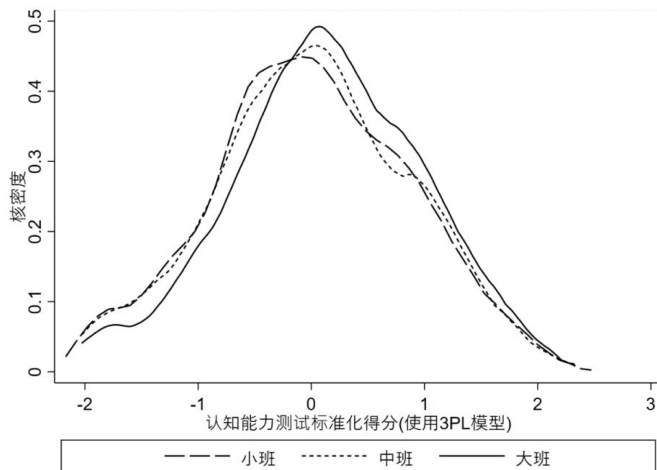


图1 不同类型班级学生的基准学业能力分布情况

根据模型的设定, 首先使用全样本数据检验班级规模对初中生学业能力的影响。随机截距模型的估计结果如表 2 第(1)、(2)列显示, 班级规模会影响初中生的学业能力。班级规模对学生的学业能力的影响在 1%的水平上显著为正, 班级规模每增加一个标准差, 学生的学业能力提高约 0.005 个标准差。上述回归结果表明, 对于不区分班级类型的全样本而言, 班级规模的增大对学生学业能力的发展有利, 正如前文所述, 这是大班拥有好生源所致。第二层次的 $R^2=0.705$, 所以学校层面上 70.5%的方差可以由设定的模型来解释, 这意味着该模型的整体解释度比较高。但值得注意的是, 班级规模能解释的学校层面的方差仅占 2.43%, 说明班级规模在学校间的方差解释程度不高。

表 2 班级规模对学业能力的影响(KLM 模型)

	全样本		非大班		大班	
	(1)	(2)	(3)	(4)	(5)	(6)
班级规模	0.005*** (0.002)	0.006** (0.003)	0.002 (0.002)	-0.008** (0.004)	-0.009** (0.004)	
学校区位	-0.043** (0.018)	-0.039** (0.018)	-0.070*** (0.023)	-0.043** (0.020)	0.004 (0.032)	-0.012 (0.033)
学校质量	-0.019 (0.043)	-0.037 (0.044)	-0.018 (0.061)	-0.094* (0.051)	0.157*** (0.059)	0.143** (0.057)
分班方式	0.027 (0.070)	0.033 (0.071)	0.050 (0.095)	0.028 (0.079)	0.056 (0.123)	0.085 (0.125)
同伴能力均值	0.242*** (0.030)	0.241*** (0.030)		0.348*** (0.038)		-0.029 (0.062)
同伴能力方差	0.504*** (0.099)	0.485*** (0.010)		0.410*** (0.118)		0.268 (0.225)
基准学业能力	0.405*** (0.009)	0.405*** (0.009)	0.440*** (0.010)	0.414*** (0.011)	0.373*** (0.016)	0.380*** (0.017)
性别	-0.065*** (0.013)	-0.064*** (0.013)	-0.075*** (0.016)	-0.069*** (0.016)	-0.048* (0.025)	-0.048* (0.025)
是否独生子女	-0.023 (0.016)	-0.024 (0.016)	-0.021 (0.019)	-0.017 (0.019)	-0.044 (0.032)	-0.043 (0.032)

续表

	全样本		非大班		大班	
	(1)	(2)	(3)	(4)	(5)	(6)
是否农业户口	0.004 (0.024)	-0.002 (0.024)	0.009 (0.029)	0.008 (0.029)	-0.033 (0.045)	-0.032 (0.045)
是否住校	0.026 (0.016)	0.026 (0.016)	0.020 (0.019)	0.021 (0.019)	0.053* (0.031)	0.053* (0.031)
家庭经济资本	-0.006 (0.015)	-0.006 (0.015)	-0.004 (0.017)	-0.005 (0.017)	-0.013 (0.027)	-0.013 (0.027)
家庭文化资本	0.023*** (0.004)	0.022*** (0.004)	0.022*** (0.005)	0.020*** (0.005)	0.027*** (0.008)	0.027*** (0.008)
常数项	-0.005 (0.152)	-0.205 (0.166)	0.146 (0.213)	0.109 (0.195)	0.473 (0.327)	0.420 (0.417)
第二层剩余方差	0.061	0.063	0.092	0.064	0.021	0.020
第一层剩余方差	0.394	0.394	0.410	0.406	0.346	0.346
观测数	9050	9050	6704	6704	2346	2346

注：* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ ；括号内为标准误。

进一步加入班级规模的二次项来考察班级规模与学业能力之间是否存在非线性关系。如图2所示，班级规模与学业能力呈现“倒U型”关系，在峰值之前，班级规模的增大有助于学业能力的提升，但超过峰值时，班级规模的增大对学生的学业能力产生负向影响。另外，在峰值附近区域内曲线比较平缓，这意味着班级规模对学业能力的边际影响相对较小。因此，通过改变峰值附近区域的班级规模提高学业成绩是一种低效的投资方式。这一结论与利用希腊数据对班级规模与学业能力的关系进行实证检验的结果大体一致，但是峰值存在差异(Kedagni et al., 2021)。由此可见，控制大班额是提高教育质量的必要举措。

分样本检验结果如表2所示，班级规模对学业能力的影响在不同类型的班级中存在差异。首先，对于超过56人的大班，班级规模对学业能力的影响在5%的水平上显著为负，班级规模每增加一个标准差，学业能力降低约0.009个标准差，说明对于大班，班级规模的增大对学业能力的发展产生负面影响。同时，对于大班，无论控制同伴特征与否，班级规模对学业能力都有显著的负效应。而且在控制同伴特征的模型中，同伴特征对学业能力的影

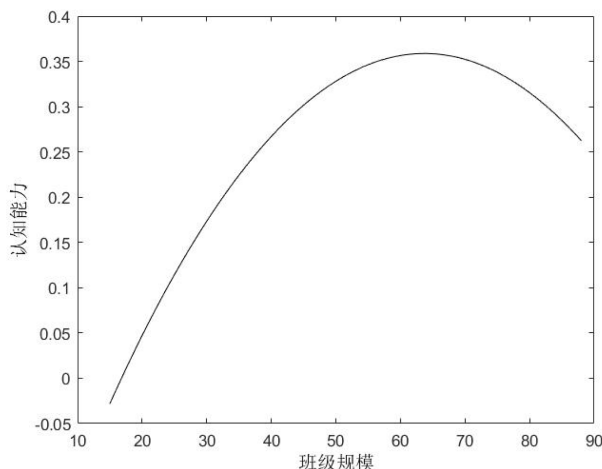


图2 班级规模与学业能力的关系

响并不显著，这意味着如果班级中的人数非常多，同伴效应将不再是班级规模负效应的作用机制，此时拥挤效应占主导。其次，对于非大班，在不控制同伴特征时，班级规模对学业能力的影响在5%的水平上显著为正，控制同伴特征后，班级规模对学业能力的影响不再显著，但是两个同伴特征变量对学业能力的影响都在1%的水平上显著为正，这说明在非大班中，同伴效应很有可能是班级规模对学业能力影响的重要机制。

(二) 班级规模的影响机制分析

前文分析发现，对于非大班而言，在加入同伴学业能力的均值和标准差两个控制变量时，班级规模的正向效应不再显著。为了进一步检验班级规模对学业能力的影响机制，下面对中小班和大班分别进行同伴特征的中介效应检验。

中介效应分析(Mediating Effects)常用来分析自变量对因变量的作用机制。本文以班级规模为自变量，以学业能力为因变量，以同伴平均能力和同伴能力方差分别作为中介变量，构建多层次固定中介效应模型，使用Bootstrap方法进行中介效应检验。与其他中介效应检验方法相比，Bootstrap检验法具有较高的统计效力，是公认的可以取代Sobel方法而直接检验系数乘积的方法(温忠麟和叶宝娟，2014)。本文将抽样次数设为500，表3呈现了中介效应的检验结果，图3反映了班级规模对学业能力影响的路径。

表 3 同伴特征的中介效应分析结果

路径	非大班(样本量: 6704)			大班(样本量: 2346)		
	总效应	直接效应	间接效应	总效应	直接效应	间接效应
同伴平均能力	0.006**	0.003	0.006***	-0.009**	-0.011**	0.020**
同伴能力方差	0.006**	0.005*	0.002***	-0.009**	-0.008**	0.001***

注: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; 括号内为标准误。

中介效应检验结果显示, 对于非大班, 班级规模对学生学业能力的总效应在 5% 的水平上显著为正。如图 3 所示, 班级规模通过同伴平均能力影响学业能力的间接效应在 1% 水平上显著为正, 与此同时, 班级规模影响学业能力的直接效应并不显著, 因此同伴均值能力在班级规模与学业能力之间发挥完全中介的作用。Preacher 和 Hayes(2008)呼吁放弃完全中介的概念, 将所有中介都看作部分中介。因此我们将这一结论解释为: 班级规模主要通过同伴均值能力影响学业能力。班级规模通过同伴能力的方差影响学业能力的间接效应在 1% 的水平上显著为正, 且直接效应在 10% 的水平上显著为正, 因此, 同伴能力方差在班级规模与学业能力之间发挥部分中介作用, 此时中介效应占总效应的 33.63%。这表明班级规模会通过同伴能力方差影响学业能力。Sacerdote(2011)称这种同伴的差异性带来的正向效应为“彩虹效应”。综上所述, 对于非大班, 同伴学业能力特征是班级规模影响学业能力的主要路径。

对于超过 56 人的大班, 班级规模对学生学业能力的总效应在 5% 的水平上显著为负。如图 4 所示, 通过同伴能力特征影响学业能力的间接效应显著为正, 而班级规模对学业能力的直接效应显著为负。由此可知, 在大班中, 同伴能力特征并不是班级规模影响学业能力的重要路径, 而是在班级规模与学业能力之间发挥了遮掩效应, 此时拥挤效应在班级规模负效应中占主导。

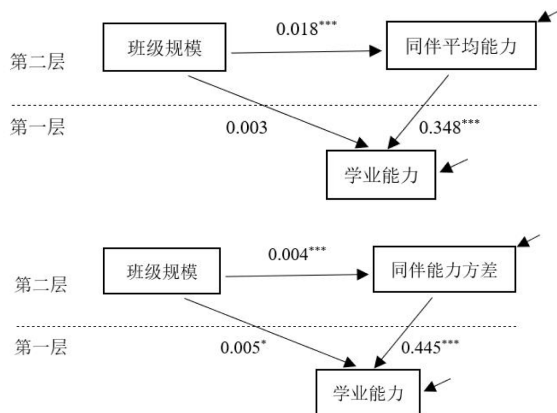


图 3 非大班的多层次中介模型

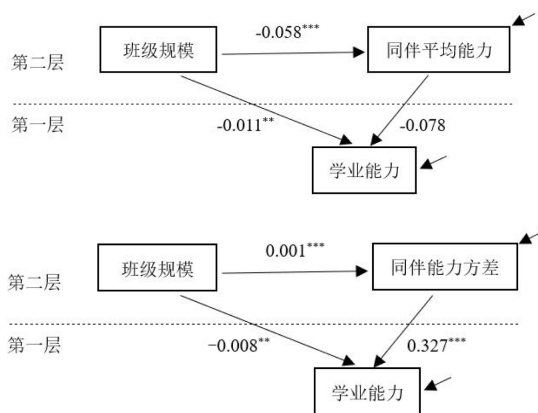


图4 大班的多层次中介模型

(三) 班级规模对不同能力分位学生的异质性影响

为了探究班级规模对不同能力分位学生的影响，我们使用分位数回归进行估计。表4呈现了班级规模效应的分位数回归结果。前三列的结果表明，在非大班中，班级规模对不同能力分位学生的影响均在1%的水平上显著为正，班级规模提高1个标准差，学生的学业能力提高约0.005个标准差。后三列结果显示，对班级人数超过56人的大班，班级规模的增大对所有类型的学生都有负向影响，但对学困生和中等生的负向影响在统计上不显著，对学优生的影响在5%的水平上显著为负，班级规模每增加1个标准差，学优生的学业能力降低0.006个标准差。

以上分析表明，班级规模效应对于不同能力分位的学生具有异质性。首先，对于非大班的学生而言，班级规模的增加对所有能力分位的学生都有显著的正向影响，而且对不同分位的学生影响差别不大。这说明适度增加班级人数有利于不同类型学生的学业能力的发展。但是对于大班而言，班级规模的极度增加会对学优生产生显著的负向效应，而对于中等生和学困生，班级拥挤带来的负效应并不显著，因此学优生对超大规模班级的负向效应比较敏感。

表4 班级规模对不同能力分位学生的影响

	非大班			大班		
	0.25RQ	0.5RQ	0.75RQ	0.25RQ	0.5RQ	0.75RQ
班级规模	0.005***	0.005***	0.004***	-0.004	-0.003	-0.006**
	(0.002)	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)

续表

	非大班			大班		
	0.25RQ	0.5RQ	0.75RQ	0.25RQ	0.5RQ	0.75RQ
控制变量	包含	包含	包含	包含	包含	包含
观测数	6704	6704	6704	2346	2346	2346

注：* $p < 0.1$ ，** $p < 0.05$ ，*** $p < 0.01$ ；括号内为标准误。

四、结论与启示

本文基于中国教育追踪调查数据中七年级学生的数据，采用多层次模型、多层次中介效应分析以及分位数回归等方法，对班级规模与初中生学业能力的关系进行研究，主要讨论了班级规模与学业能力的非线性关系，探索了不同类型班级的班级规模影响学业能力的机制，并检验了班级规模对不同层次学生的异质性影响。

研究发现，班级规模与学生学业能力之间并不是简单的正向或负向的线性关系，而是“倒U型”的非线性关系，即班级规模对学业能力的影响呈现先正向后负向的变化趋势。换言之，在不改变大班教学模式的情况下，班级规模过大或过小都不利于学生学业能力的发展。此外，由于峰值附近的曲线相对平缓，所以通过改变峰值附近的班级规模提高学生的学业能力是一种低效的投资方式。不同类型班级的班级规模的影响路径不同，具体而言，对于小班和中班，适当增加班级规模带来的正向效应主要来自于同伴效应，同伴学业能力均值和方差在班级规模与学业能力之间均发挥中介作用，同伴平均学业能力越高对学生学业能力发展越有利，而且，同伴的异质性也有助于学业能力的提高，这体现了同伴之间的“彩虹效应”。但是对于班级规模已经很大的大班，继续增加班级规模会对学生的学业能力产生负向影响，此时拥挤效应占主导。另外，班级规模对不同能力分位学生的影响具有异质性。对于大班，班级规模对学优生的学业能力具有显著的负向影响，而对学困生和中等生则没有显著影响，这说明学业能力较强的学生对班级规模的拥挤效应更加敏感。

以上结论的政策含义有两点。首先，我国现阶段义务教育阶段的主要矛盾是城乡居民对优质教育的渴望与优质教育资源的不足，而班级规模的缩减不仅是对班级人数的控制，更是一项以提高教育质量为目标系统工程。考虑到班级规模与学业能力发展“倒U型”的非线性关系，在小班化教学模式探索阶段，可以放宽班级规模调整的范围，允许学校适度增加班级人数，避免

因过度追求班级规模的缩减造成的资源浪费。但是,超大班级规模带来的拥挤效应会对学生产生显著的消极影响,因此政府强制消除“大班额”是兼顾效率与教育公平的重要举措。其次,目前我国讲授制大班教学模式为主导的情况下,班级规模主要通过同伴效应对学业能力产生正向影响,因此学校应该科学地做好分班决策,重新审视分层分班的方式,重视班级学生构成的差异性和多样性,充分发挥同伴之间的“彩虹效应”,从而保证班级规模的正向效应得以充分发挥。

[参考文献]

- 崔盛、吴秋翔, 2019:《班级规模对初中生学业能力的异质性影响——基于中国教育追踪调查数据的实证研究》,《中国教育学刊》第3期。
- 杜育红、袁玉芝, 2016:《教育中的同伴效应研究述评:概念、模型与方法》,《教育经济评论》第3期。
- 方征, 2015:《班额调整须与教育发展阶段相适应:班额效应“异常”的思考》,《教育发展研究》第4期。
- 唐一鹏, 2014:《从上海 PISA2012 透视学生能力分位的影响因素》,《上海教育科研》第6期。
- 王卫东, 2016:《中国教育追踪调查(CEPS)基线报告》,北京: 中国社会科学出版社。
- 温忠麟、叶宝娟, 2014:《中介效应分析:方法和模型发展》,《心理科学进展》第5期。
- 姚昊、胡耀宗、马立超, 2021:《班级规模、教师学历如何影响学生学业成绩?——基于 PISA 2018 的国际比较研究》,《清华大学教育研究》第5期。
- 约翰·哈蒂, 2015:《可见的学习——对 800 多项关于学业成就的元分析的综合报告》,彭正梅,邓莉,高原,北京:教育科学出版社。
- 赵颖, 2019:《同群效应如何影响学生的学业能力》,《财贸经济》第8期。
- 郑琦、杨钊, 2018:《班级规模与学生学业成绩——基于 2015 年 PISA 数据的研究》,《北京大学教育评论》第6期。
- Addonizio, M. F. and J. L. Phelps, 2000, “Class Size and Student Performance: A Framework for Policy Analysis”, *Journal of Education Finance*, 26(2): 135–156.
- Angrist, J. D. and V. Lavy, 1999, “Using Maimonides’ rule to Estimate the Effect of Class Size on Scholastic Achievement”, *Quarterly Journal of Economics*, 114(2): 533–575.
- Angrist, J. D. et al., 2019, “Maimonides’ Rule Redux”, *American Economic Review*, 1(3): 1–16.
- Barrett, N. and E. F. Toma, 2013, “Reward or Punishment? Class Size and Teacher Quality”, *Economics of Education Review*, 35: 41–52.
- Blatchford, P., K. W. Chan, M. Galton, K. C. Lai and J. C. K. Lee, 2016, *Class Size: Eastern and Western Perspectives*, Routledge Press.

- Blatchford, P., P. Bassett, H. Goldstein and C. Martin, 2003, "Are Class Size Differences Related to Pupils' Educational Progress and Classroom Processes? Findings from The Institute of Education Class Size Study of Children Aged 5-7 Years", *British Educational Research Journal*, 29(5): 709-730.
- Borland, M. V., R. M. Howsen and M. W. Trawick, 2005, "An Investigation of the Effect of Class Size on Student Academic Achievement", *Education Economics*, 13(1): 73-83.
- Brühwiler, C. and P. Blatchford, 2011, "Effects of Class Size and Adaptive Teaching Competency on Classroom Processes and Academic Outcome", *Learning and Instruction*, 21(1): 95-108.
- Cohen, J., 1988, *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edition. Academic Press.
- Ding, W. and S. Lehrer, 2010, "Estimating Treatment Effects from Contaminated Multiperiod Education Experiments: The Dynamic Impacts of Class Size Reductions", *Review of Economics and Statistics*, 92(1): 31-42.
- Finn, J. D., G. M. Pannozzo and C. M. Achilles, 2003, "The 'Why's' of Class Size: Student Behavior in Small Classes", *Review of Educational Research*, 3: 321-368.
- Hanushek, E. A., 1986, "The Economics of Schooling: Production and Efficiency in Public School", *Journal of Economic Literature*, 24(3): 1141-1177.
- Hanushek, E. A., 1999, "Some Findings from an Independent Investigation of the Tennessee STAR Experiment and from Other Investigations of Class Size Effects", *Educational Evaluation and Policy Analysis*, 21(2): 143-163.
- Hoxby, C. M., 2000, "The Effects of Class Size on Student Achievement: New Evidence from Population Variation", *Quarterly Journal of Economics*, 115(4): 1239-1285.
- Jepsen, C. and S. Rivkin, 2009, "Class Size Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size", *Journal of Human Resources*, 44(1): 223-250.
- Kedagni, D., K. Krishna, R. Megalokonomou and Y. Zhao, 2021, "Dose Class Size Matter? How, and at What Cost?", *European Economic Review*, 133: 103664.
- Krueger, A. B., 1999, "Experimental Estimates of Education Production Functions", *Quarterly Journal of Economics*, 114(2): 497-532.
- Lazear, E. P., 2001, "Educational Production", *Quarterly Journal of Economics*, 116(3): 777-803.
- Leuven, E., H. Oosterbeek and M. Ronning, 2008, "Quasi-experimental Estimates of the Effect of Class Size on Achievement in Norway", *Scandinavian Journal of Economics*, 110(4): 663-693.
- Levin, J., 2001, "For Whom the Reductions Count: A quantile Regression Analysis of

- Class Size and Peer Effects on Scholastic Achievement”, *Empirical Economics*, 26(1): 221–246.
- Preacher, K. J. and A. F. Hayes, 2008, “Assessing Mediation in Communication Research”. In Hayes, A. F., M. D. Slater and L. B. Snyder (Eds.), *The Sage Sourcebook of Advanced Data Analysis Methods for Communication Research*: 13–54. Sage Publications, Inc.
- Sacerdote, B., 2011, “Peer Effects in Education: How Might They Work, How Big are They and How Much do We Know Thus Far”, In E. A. Hanushek, S. Machin and L. Woessmann(Eds.), *Handbook of the Economics of Education*, Vol. 3: 250–273.
- Urquiola, M., 2006, “Identifying Class Size Effects in Developing Countries: Evidence from Rural Bolivia”, *Review of Economics and Statistics*, 88(1): 171–177.

Class Size, Peer Characteristics and Junior Middle School Students’ Academic Ability: Evidence from Empirical Data in China

CAO Shu-jiang, XIE Li-yan

(School of Education, Renmin University of China)

Abstract: Reducing class size is an educational investment method that is widely concerned to improve quality and promote education equity. Based on the data of China Education Panel Survey (CEPS), we use Hierarchical Linear Model to study the effects of class size on junior middle school students’ academic ability. In the absence of high-quality education resources, there are obvious differences in the quality of students in different types of classes. Large classes often have better students, which is the main reason for the estimation bias of the positive effect of class size. We found that there is an approximately inverted-U relationship between class size and academic ability. Mediation effect test shows that there are differences in the impact mechanism of class size effect in different type of classes, performance may increase as class size rises initially because of peer effects, and fall when congestion effects start to dominate. In addition, research shows that the influence of class size on different types of students is heterogeneous, moderate increase in class size has a significant positive impact on all quantile students, and the degree of the impact is similar. However, when the class size is extremely expanded, students with strong academic ability are more sensitive to congestion effects.

Key words: class size; peer effect; academic ability; Hierarchical Linear Model; CEPS

(责任编辑: 郑磊 责任校对: 郑磊 刘泽云)