

幼儿园班级规模与儿童发展： 来自县域追踪调查的证据

杨 钊 刘永烨 梁 净

(北京大学 教育学院/教育经济研究所, 北京 100871)

(北京大学 中国教育财政科学研究所, 北京 100871)

摘 要 降低班级规模是政府近年来提升学前教育结构性质量的政策工具之一。本文基于2017~2018年我国华北某县学前教育质量追踪调查数据,采用多水平增值模型系统考察了在当前学前教育结构性质量普遍提升的背景下县域内幼儿园班级规模效应。研究发现,班级规模对小班儿童升入中班后的发展水平影响不显著,但对中班就读“小额班”的儿童进入大班后的表现有显著负向影响;班级规模对不同儿童群体成绩增值的影响具有较高异质性,并非所有儿童都能在班级规模缩减中获益。经过一系列稳健性检验后,上述结果依然稳健。本文结论表明,降低班级规模虽然可以改善幼儿园结构性质量,但尚未对儿童发展产生显著的积极影响。

关键词 学前教育;班级规模;儿童发展;公共政策

中图分类号: G40-054 **文献标识码:** A **文章编号:** 1671-9468(2020)03-0060-26

DOI: 10.12088/pku1671-9468.202003003

一、问题的提出

近年来,公平与质量成为世界学前教育发展的核心议题。各国政府尝试在扩张学前教育规模的同时提升教育质量,促进学前教育内涵式发展,我国也不例外。自2010年《国家中长期教育改革和发展规划纲要(2010—2020年)》和《国务院关于当前发展学前教育的若干意见》(国发〔2010〕41号,以下简称“国十条”)颁布和实施以来,我国学前教育事业总体规模急速扩张,儿童入园机会显著提升。2019年我国学前三年毛入园率达到83.4%^[1],2017年经济与合作组

作者简介: 杨钊,女,北京大学教育学院/教育经济研究所副教授,博士。

刘永烨,男,北京大学中国教育财政科学研究所研究助理,本文通讯作者。

梁净,女,北京大学中国教育财政科学研究所科研助理/科研项目负责人。

致谢: 本文得到澳门同济慈善会、北京乐平公益基金会及中国发展研究基金会项目的资助。作者感谢北京大学中国教育财政科学研究所学前教育研究课题组提供的数据,同时感谢北京大学宋映泉副研究员和王默博士、中国农业大学经济管理学院讲师龙文进及匿名审稿人的许多建议。文责自负。

织(OECD)国家3~5岁儿童的学前教育参与率平均为87%^[2],中国2019年的学前三年毛入园率已接近2017年OECD国家的水平。“国十条”出台后,国务院作出以县为单位制定学前教育三年行动计划的重大的战略部署,要求各地把内涵发展和质量提高作为编制行动计划的重要内容,并重点支持中西部地区发展农村学前教育。与此同时,中央和地方各级政府加大对学前教育的财政投入,相继出台一系列配套政策或措施,促进农村学前教育的普及和质量提升^①。

义务教育及以上阶段的质量可以通过学生考试成绩来体现,不同的是,学前教育质量概念较为复杂,测量较困难^[3-4]。学界通常从结构性、过程性和结果性三个维度予以定义。结构性质量指学前教育机构的师资水平、幼师比、班级规模和安全标准等可具体规范和控制的要素^[5];过程性质量指师生互动水平、幼儿之间的互动以及相关环境质量等与儿童发展密切相关的动态因素;结果性质量指儿童早期体格、认知、情感、社会适应性及语言等各个方面的综合发展水平^[6]。研究表明,提升学前教育质量有助于促进儿童的早期发展。高质量的学前教育不仅能够促进儿童认知和非认知技能的发展^[7],还是弱势儿童群体成长的重要保护性因素^[8-10]。

近年来,我国学前教育的结构性质量得到明显改善。其中,以平均班级规模、生师比、教师和园长学历结构等三个方面为代表的结构性质量提升最为显著^[11-12]。在众多结构性质量指标中,缩减班级规模被认为是提升教育质量的核心政策工具之一,也是各国政府评估和问责学前教育机构评价的重要指标之一^②。在“国十条”的指导下,中央明确提出地方政府是发展学前教育的责任主体,不仅要逐步建立起对幼儿园结构性质量的规范制度,还要将结构性质量纳入学前教育评价督导体系^③。教育部2013年颁布的《幼儿园教职工配备标准(暂行)》(教师〔2013〕1号)和2016年颁布的《幼儿园工作规程》(中华人民共和国教育部令第39号),进一步明确了幼儿园班级规模和专任教师与保育员配备标准。2017年颁布的《幼儿园办园行为督导评估办法》(教督〔2017〕7号)和2020年教育部督导办颁布的《县域学前教育普及普惠督导评估办法》(教督〔2020〕1号),均强调要将班级规模作为幼儿园保教质量评估指标。同时,《县域

① 例如自2010年起,中央政府主要通过中央财政学前教育项目和农村学前教育推进工程两个专项支持农村学前教育事业发展。在质量提升方面,2011年,财政部和教育部印发《关于加大财政投入支持学前教育发展的通知》(财教〔2011〕405号),明确提出财政支持学前教育发展的基本原则是“广覆盖、保基本、有质量”。

② 例如:美国儿科学会、美国公共健康学会、美国国家早期教育研究所、全美幼教协会等都制定了包括班级规模和幼师比在内的高质量学前教育的评价指标体系。又如,北美国家将班级规模和幼师比作为“质量评价和提升系统”的关键评价指标。

③ 2012年,为贯彻落实《国家中长期教育改革和发展规划纲要(2010—2020年)》和《国务院关于当前发展学前教育的若干意见》精神,进一步推动各地学前教育三年行动计划的实施,教育部研究制定了《学前教育督导评估暂行办法》(教督〔2012〕5号),该《办法》提出在对地方人民政府学前教育督导过程中,要从政府责任、经费投入、园所建设、队伍建设、规范管理、发展水平六个方面进行督导,其配套的“学前教育督导评估指标体系”采用“国十条”规定,强调“合理确定幼儿教师生师比,核定公办幼儿园教职工编制,配足配齐教职工”。

学前教育普及普惠督导评估办法》进一步提出了幼儿园保教质量保障督导评估指标和标准,要求班额普遍达标,县域内 85% 以上的班额符合《幼儿园工作规程》有关规定。

我国现行幼儿园班级规模配备标准,遵循 2013 年教育部印发的《幼儿园教职工配备标准(暂行)》。该《标准》对我国全日制和半日制幼儿园不同年龄班的班级规模做出了明确规定^①,总体上要求平均班级规模控制在 30 人以下。从教育部公布的历年教育统计数据来看(见图 1),2010 年以来城区幼儿园和乡村幼儿园平均班级规模已达到国家标准,但镇区幼儿园平均班级规模从 2001 年至 2017 年一直高于国家标准。从时间趋势上看,镇区幼儿园和乡村幼儿园平均班级规模仍保持较大差距,未呈现出明显的缩减趋势,这体现了学前教育阶段“城镇大班额”和“乡村小班额”并存的现象,这种现象是制约中西部农村地区县域学前教育均衡发展和质量提升的瓶颈之一。

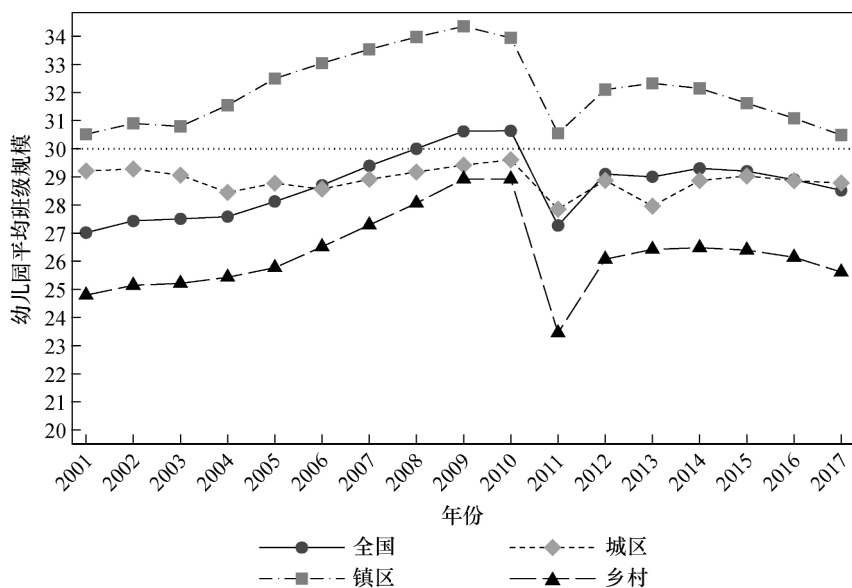


图 1 全国各类幼儿园平均班级规模变化趋势(2001—2017 年)

数据来源:根据历年《中国教育统计年鉴》相关数据计算。

注:2010 年及以前采用城市、县镇和农村的分类标准。由于统计指标的调整,2011 年开始采用城区、镇区、乡村的分类,故在此仅采用 2011 年以后的城乡分类标准。

因此,消除城镇大班额成为各级政府提升学前教育质量的着力点之一。从政策文本的变化中不难发现,以班级规模为代表的结构性质量指标,逐步成为与毛入园率、普惠性幼儿园覆盖率等事业发展指标同样重要的政府督导内容。通过控制班级规模和幼师比以改善学前教育结构在质量的做法,逐渐发展成为

^① 其中,小班(3~4 岁)、中班(4~5 岁)、大班(5~6 岁)和混龄班的班级规模分别为 20~25 人、25~30 人、30~35 人和 30 人以下。

当前督导中评估政策的达标性和高利害问责标准的政策工具^①。

班级规模和幼师比下降会对教育机会和教育质量产生重大影响。宋映泉指出,当前我国学前教育事业发展的主要矛盾之一是结构性学前质量改善与城乡儿童发展差距不断扩大之间的矛盾,即随着幼儿园班级规模和幼师比的下降、教师和园长学历水平的明显提升,以《教育部3~6岁儿童发展指南》儿童发展工具、《早期儿童发展能力指数》(中国版)和《亚太早期儿童发展量表》(East Asia Pacific Early Child Development Scales,简称EAP-ECDs)测量的城乡儿童发展水平之间的差距不但未随之下降,反而有明显扩大的趋势^[13]。

幼儿园班级规模与儿童发展之间的复杂关系,需要进一步的研究予以澄清。遗憾的是,我国学术界尚未系统分析幼儿园班级规模对儿童发展的影响。一方面,在“以县为主”的学前教育管理体制下,国家学前教育督导多以县级政府为对象。因此,需要以县域大样本数据为基础来考察学前教育结构性质量的改善对儿童发展的影响,而这方面的实证研究较少。另一方面,由于测量方面的困难,鲜有大规模、系统性的追踪调查来测量农村地区学前教育质量,并考察其结构性质量提升对儿童发展的影响。有鉴于此,本文拟使用北京大学中国教育财政科学研究所2017—2018年在华北某县收集的学前教育阶段儿童发展追踪调查数据,系统考察贫困地区县域幼儿园班级规模与儿童发展之间的关系,并分析在不同儿童发展领域、不同年段、城乡之间、不同所有制、不同家庭背景等方面班级规模效应的差异。在此基础上,本文对我国学前教育阶段以消除“大班额”为导向的结构性质量提升策略进行了反思。

二、幼儿园班级规模效应

作为结构性教育质量的指标,缩减班级规模和降低幼师比历来是学术界和政策制定者长期关注的议题。幼儿园班级规模指一个幼儿园班级中幼儿的数量;幼师比指在一定班级规模条件下教师所负责照顾的幼儿人数与教师人数的比值,因此,二者相互联系且同方向变动。二者之中,幼师比是与幼师互动质量等更相关的概念,但班级规模是低幼师比发挥作用的条件^[14]。班级规模和幼师比是容易定义、测量和通过立法规制的学前教育结构性质量指标,这些指标可以通过营造安全和积极的保育环境来间接影响儿童发展^[15]。理论分析指出,缩减班级规模可能通过提升师生互动^[16-18]、提升教学质量^[19]、改进儿童和教师行为^[20-24]等多种机制影响学前儿童的教育成就。

^① 以《幼儿园办园行为督导评估办法》为例,该《办法》适用于教育督导机构对面向3~6岁儿童提供保育服务的幼儿园(班、点)实施的督导评估(第一章第三条)。《办法》要求幼儿园办园行为督导评估报告除了向社会发布,还要报送本级人民政府,作为制定学前教育政策、加强幼儿园管理的依据(第四章第十三、十四条),督导评估结果将作为幼儿园年检、确定级类和园长评优评先的重要依据(第四章第十六条)。

虽然理论分析表明学前教育阶段较小的班级规模和较低的幼师比有助于儿童发展,但经验分析却未能达成一致结论^[25-26]。20 世纪 70 年代的实验研究发现,被随机分配到小规模班级中的儿童(幼师比低于 7:1,儿童数量小于 14 人)比被分配到大规模班级的儿童(幼师比高于 9:1,儿童数量小于 18 人)在一般性知识、语言和社会行为测量上的得分更高^[27]。80 年代至 90 年代,美国田纳西州实施了一项学前教育阶段的大规模随机干预实验,即生师比改进计划(Student-Teacher Achievement Ratio,简称 STAR)。基于这项全州层面的大规模随机干预实验,大量研究发现,小规模班儿童的标准化成绩比常规班级的儿童高 0.22 个标准差^[28],班级规模的显著影响至少持续到 7 年级^[29]。为了更好地评估幼儿园班级规模的长期影响,拉吉·切蒂(Raj Chetty)等学者将 STAR 项目数据与儿童成年后的税收数据连接起来,对 STAR 项目的长期影响进行了评估。他们发现尽管班级规模对儿童 8 年级以后考试成绩的影响逐渐消失,但对其非认知技能方面的积极影响仍然存在^[30]。格罗斯(J. Gross)等学者也发现缩减班级规模会提高儿童意志力,这种能力对儿童的学业成绩有重要的预测作用,还对个人在劳动力市场的表现有显著影响^[31]。

元分析发现班级规模和幼师比与儿童认知和非认知发展存在复杂的关系。一项针对美国 1960—2007 年间 38 项学前教育评估项目的元分析发现,班级规模和幼师比与儿童发展存在非线性关系^①。当采用非线性的班级规模测量时(大于或者小于 15 人),元分析发现班级规模降低至 15 人以下时,每减少一名儿童,儿童认知和社会性情感发展的效应量就提高 0.1 个标准差^[32],这表明只有在较低的班级规模水平上,才会出现显著相关关系。

国内学术界也关注班级规模对儿童发展的影响,但这些研究往往更关注中小学或更高教育阶段,较少涉及学前教育阶段,很难将基础教育阶段班级规模效应的研究结论推广到学前教育阶段^[33]。在少数将学前阶段班级规模作为控制变量的研究中,胡碧颖等学者发现,在控制其他因素的情况下,班级规模和儿童字词认知水平存在显著的负相关关系^[34]。近期一项实验研究发现,降低在场幼师比可以提高幼儿学习结果和参与度,提高教师对个体幼儿的关注度和评价的准确度。当在场幼师比从 20:1 降到 10:1,教师教学行为从全班集体教学转变为不同幼师比和分组形式的教学以后,幼儿的学习结果显著优于全班集体教学的结果。其原因在于此时幼儿参与的专注性、积极性和主动性显著提升,且教师关注的幼儿比例从 60%~65%提高到 85%~100%,教师能够准确评价的幼儿比例从 40%~60%提升到 80%~85%^[35]。

① 该研究以 3~5 岁儿童参与的学前教育项目为研究对象,所选择项目要满足儿童每周至少在园 10 小时、至少在园 4 个月的筛选要求。基于上述标准,研究者筛选出 38 个独立研究,这些独立研究提供了 53 个对照组和 328 个效应量。以效应量为因变量,班级规模和幼师比为自变量,采用两水平随机截距模型进行估计。

现有对幼儿园班级规模与儿童发展的研究尚存在一些问题。第一,以往研究多集中在西方社会,且尚未得出一致结论。第二,多数非实验研究基于单一时点的横截面数据,鲜有研究采用追踪调查数据来考察班级规模对儿童发展增值的影响。第三,多数研究未考虑班级规模的内生性问题^①,这些研究结论难以排除因自选性和遗漏变量带来的潜在混淆偏差^[36],从中获得的政策建议也需要谨慎对待。第四,在结果变量的选取上,已有研究关注班级规模对儿童认知技能的影响,较少关注其对非认知技能影响,忽视班级规模对儿童非认知技能的影响将会造成有偏估计。第五,近年来我国政府致力于通过缩减班级规模提高教育质量,但这种做法是否同时提升了县域内城乡以及不同家庭背景儿童群体的发展水平,尚未得到证实。

三、研究设计

本文使用的数据来源于北京大学中国教育财政科学研究所2017—2019年在华北某县开展的为期三年的“县域内农村学前教育发展质量评估指标体系”研究项目。该研究采用前瞻性队列研究设计(Prospective Cohort Study),从结构性质量、过程性质量和儿童发展指标三个维度,全面了解县域内学前教育发展现状及其变化趋势。

(一) 样本县基本情况

样本县位于华北某省东南部,辖10镇10乡,是国家扶贫开发工作重点县之一。根据国家统计局2017年统计用区划和城乡划分代码对该县城乡类型分类可知,该县的城乡类型覆盖镇中心区、镇乡结合区、乡中心区和村庄四类,前两类属于城镇,后两类属于农村。经统计,该县城镇占比31.34%,农村占比68.66%,是典型的以县镇和乡村为主的农村地区。近年来,该县累计投入1.2亿元发展学前教育。2018年,该县共有公办园189所,民办园41所,在园儿童4.2万余人,学前三年毛入园率达96%。该县幼儿园数量和入园幼儿规模居全市之首,学前三年毛入园率远超全省平均水平,是我国贫困农村地区实现“基本普及学前教育”和“增加普惠性学前教育资源占比”两个发展目标的典型代表县之一。

^① 班级规模的内生性问题主要体现在校际和校内的筛选两个方面:校际的筛选主要指家长自我选择把子女送入私立学校或者班级规模较小的公立学校;校内的筛选主要指学校可能根据学生的需要来分班,例如将低水平的学生集中在小班授课。

(二) 抽样设计

该项目抽样的总体是该县 217 所幼儿园中的园长、教师、儿童、家长。在抽样方法上,采用多阶段概率抽样和非概率抽样相结合的方式,抽样分三个阶段进行。第一阶段是对幼儿园的分层随机抽样。研究团队根据该县教育局提供的 20 个乡镇 217 所幼儿园的名称、规模和类型等基础信息建构抽样框,通过简单随机抽样方法抽取了 60 所幼儿园,同时按照幼儿园质量等级补充抽取 2 所质量等级最高的幼儿园。上述抽样过程保证了每个乡镇至少有 1 所幼儿园进入调查,幼儿园样本的分布在所有制(公办或民办)、质量等级和城乡属性等维度上与该县幼儿园总体情况吻合,可以视为该县学前教育发展基本状况的代表性样本。

第二阶段是班级抽样。结合前瞻性队列研究设计、幼儿园入园和编班^①规定,研究团队制定了班级抽样的原则:如果某幼儿园覆盖了小、中、大三个年级,且每个年级只有一个班,则在小班年级和中班年级中各抽取 1 个班作为调查对象;如果该幼儿园每个年级的班级数大于 1,则使用随机数表在小班和中班两个年级中随机抽取 1 个班作为调查对象。另外,考虑到某些幼儿园没有小班或中班这两个年级,则随机抽取 1 个大班或 1 个混龄班作为替代。抽取小班和中班两个年级为基线调查起始群,能够更为完整地了解儿童在幼儿园阶段的成长轨迹并方便后续追踪调查。

第三阶段是儿童抽样。完成班级抽样后,研究团队获取了被抽样班级的儿童花名册,并按照班级内男童和女童的比例,采用简单随机抽样方法抽取 10 名儿童进行儿童发展水平测试。需要说明的是,在一般情况下,每个幼儿园将有 2 个班共计 20 名儿童进入调查样本。但部分幼儿园因规模较小,抽取儿童数量可能不足 20 名。在基线调查样本中,有 5 所幼儿园抽样不足 20 名的情况。完成所有阶段的抽样后,入样儿童及家长、所在班级的教师和所在幼儿园园长被邀请参与本研究项目。

(三) 数据收集

截至 2020 年,研究团队已完成了对样本中的儿童的三轮追踪调查。2017 年 12 月,研究团队前往抽样幼儿园开展第一轮基线调查,基线调查的最终样本包括 62 个幼儿园的 119 个班级中 1216 名儿童、950 名儿童的家长 and 164 名教师。其中,小班、中班、大班和混龄班分别有 564 名、606 名、11 名和 35 名儿童,专任教师、保育员和其他教师分别有 143 名、17 名和 4 名。2018 年 11 月和 2019 年 6 月,研究团队分别对基线调查的样本儿童、家长和教师进行了两次追

^① 《幼儿园工作规程》第二章第十一条规定:幼儿园可以按年龄分别编班,也可以混合编班。年龄班即小班、中班和大班。

踪调查。第一次追踪调查成功追踪儿童 887 名,追访率为 72.9%。需要说明的是,在第一轮和第二轮追踪调查中,部分基线受访儿童因升学、请假、转园和退学而未追踪成功。为此,研究团队在第一次追踪调查时补充抽取了 311 名儿童作为新增样本。此外,部分儿童家长或教师因不配合、拒绝回答等导致部分儿童的家庭背景和班级教师信息数据缺失。其中,2017 年家长问卷的基线应答率为 78.13%,家长问卷可以完全匹配 867 名儿童信息;1 个班级没有教师数据。

2018 年第一次追踪调查时,因儿童升学或追踪时无应答等因素导致部分样本损耗(sample attrition),样本损耗主要集中在基线调查的大班和混龄班。其中,基线调查时 3 个大班的儿童已全部升入小学;4 个混龄班儿童进行了重新分班。重新分班可能对儿童发展产生影响,进而对班级规模效应产生有偏估计。因此,本文将分析样本限定在 2017 年调查的小班和中班两个年级及其第一次追踪成功的 887 名儿童,被成功追踪的样本儿童所在样本班级与基线所在样本班级保持一致。换言之,这个分析样本可以被视为是小班和中班两个年级的代表性样本。剔除家庭背景信息大量缺失的儿童后,最终剩下 58 所幼儿园的 111 个班级中 867 名儿童作为分析样本,其中包括 2017 年基线调查时 57 个小班的 443 名儿童和 58 个中班的 434 名儿童。

(四) 测量工具

儿童发展能力通过《亚太早期儿童发展量表》测量。家长在签署知情同意后,入样儿童接受 45~60 分钟的一对一个别化测试。EAP-ECDs 量表由香港大学研究团队编制,主要考察学前儿童在认知思维、社会情绪、运动技能、语言阅读、健康安全、文化参与及学习品质七大领域的发展状况^[37]。调研当日,儿童根据调查员要求进行口头应答、指认、认读或者排列图片、手部操作、直线行走、接球以及书写或者绘画等。调查员根据儿童在测试中的表现对各题项进行 0、1、2 计分,儿童在各领域得分越高则说明其发展水平越高^[38]。该量表曾在中国 5 个省市以及亚太地区共计 6 个国家中进行了前期预测试和两次正式测试,测试结果表明该量表在各国之间均表现出良好的信效度^[39],可用于国家层面大规模追踪和评估儿童的发展状况,并为不同国家和地区制定学前教育政策提供实证数据^[40]。

该量表完整版共计 85 个项目,每个领域包含的题目完整信息及各国测试的测量学属性信息(包括难度、区分度及项目功能差异),参见刘丽薇(Nirmala Rao)等学者的研究^[41]。2017 年 EAP-ECDs 各分领域和总量表的信度系数(Cronbach's α)介于 0.53~0.95 之间,2018 年介于 0.56~0.95 之间。除了运动发展领域的信度较低,其他维度的信度系数均高于 0.7,说明该量表具有较好的内部一致性^[42]。为了比较儿童在不同年份的发展水平,本文基于项目反应理论(IRT)中的广义分部计分模型(Generalized Partial Credit Model,简称

GPCM)^[43]对儿童发展能力值进行估计^①,随后将能力估计值进行标准化。经 GPCM 模型调整后的儿童发展水平得分不仅可应用于增值性评价,还能确保儿童的基线和追踪测试得分在同一尺度下可比。

家长问卷主要包括两个模块。第一模块收集了儿童个人和家庭背景信息,包括儿童的性别、出生日期、兄弟姐妹数、父母受教育程度、婚姻状况等信息。第二模块主要询问儿童的家庭社会经济地位(Socio Economics Status, 简称 SES)信息,包括父母职业和月收入、家庭年收入情况、农村家庭常见耐用品清单^②。根据菲尔默(Deon Filmer)和普里切特(Lant H. Pritchett)提出的基于农村家庭常见耐用品清单构建的家庭资产指数的方法^[44],本文构建了家庭 SES 代理变量。

在豪斯(Carollee Howes)等学者提出的结构性质量要素框架下^[45],研究团队结合中国幼儿园的实际情况设计了《幼儿园发展情况调查表》和《幼儿园教师问卷》两类收集结构性质量信息的问卷。其中,《幼儿园发展情况调查表》包括幼儿园园舍、设施设备、教职工数、经费以及幼儿园近五年来发展状况;《教师问卷》主要包括教师基本人口学信息、教育背景、职业流动、培训经历以及培训需求等。

(五) 变量定义和描述统计

表 1 概括了本文变量的来源、定义及描述统计。连续型变量的描述统计结果以均值(标准差)的形式呈现,分类变量的描述统计频数和百分比的形式呈现。

① GPCM 模型是分部计分模型(PCM)的一般化形式,适用于每题有多个分值、难度随分值的增加而增加的情况,同时增加了每个题项的区分度。国际学生能力测试(PSIA)自 2015 年起也开始使用双参数 logistic 模型和 GPCM 模型对学生能力分进行计算。具体方法参见:OECD. (2017). *PISA 2015 technical report*, 143—201. http://www.oecd.org/pisa/data/2015-technical-report/09_Chapter_09_PISA2015.pdf.

② 虽然工资性收入是 SES 的重要测量指标之一,但本文样本中有 90% 以上的儿童来自农村家庭。这些农村家庭缺乏工资性收入,因此用收入作为农村家庭 SES 的代理变量可能导致估计结果有偏。采用农村家庭常见耐用品清单来构建农村家庭社会经济背景的方法被众多学者以及包括世界银行在内的研究机构采纳,是一种综合衡量资产或财富的测量方式,比询问家庭年收入更恰当。参见:Gwatkin, D. R., Rutstein, S., Johnson, K., Suliman, E., Wagstaff, A., & Amouzou, A. (2007). *Socio-economic difference in health, nutrition and population within developing countries*. Washington, DC: World Bank; Hong, R., & Hong, R. (2007). Economic inequality and undernutrition in women: Multilevel analysis of individual, household, and community levels in Cambodia. *Food and Nutrition Bulletin*, 28 (1), 59—66. 和 Zere, E., Mandlhate, C., Mbeeli, T., Shangula, K., Mutirua, K., & Kapenambili, W. (2007). Equity in health care in Namibia: Developing a needs-based resource allocation formula using principal components analysis. *International Journal for Equity in Health*, 6(1), 3.

表1 相关变量数据来源以及描述统计

变量	数据来源及说明	观测值	均值/百分比	标准差
因变量:儿童发展水平				
认知发展领域得分 2017	2017 年得分来自基线调查,2018 年得分来自第一次追踪调查。儿童发展水平得分使用 GPCM 模型估计出分领域发展水平得分和总分,再转换为 Z 分数	867	0.00	0.95
社会性情绪发展领域得分 2017		867	0.00	0.90
总体发展水平得分 2017		867	0.00	0.98
认知发展领域得分 2018		867	-0.01	0.94
社会性情绪发展领域得分 2018		867	0.00	0.91
总体发展水平得分 2018		867	0.00	0.97
核心自变量:班级规模				
连续变量测量	所在班级注册儿童数	111	26.84	8.97
分类变量—“小额班”	根据政策规定将班级规模虚拟化编码:小于政策规定规模的班级称为“小额班”,政策规定之内规模的班级称为“常规班”,大于政策规定规模的班级称为“大额班”	33	29.73%	/
分类变量—“常规班”		25	22.52%	/
分类变量—“大额班”		53	47.75%	/
控制变量:儿童个体及家庭特征				
儿童性别	男孩=1,女孩=0	867	49.94%	/
儿童年龄	利用 2017 年调查基线调查年月日减去儿童的出生年月日	867	4.30	0.64
父母最高受教育年限	选取受教育年限更高的一方	867	10.08	2.52
父母是否为主要监护人	是=1,否=0	867	80.39%	/
兄弟姐妹数	被调查儿童亲兄弟姐妹数量	846	1.65	0.87
农村户口	农村户口=1;城镇户口=0	856	92.87%	/
家庭资产指数	基于问卷中列出的 26 类农村家庭常用物品,利用主成分分析计算所得	867	0.03	2.33
控制变量:班级特征				
班级教师平均月收入	2017 年班级内教师平均月收入(元);回归时取对数	108	1436.54	395.82
班级教师平均年龄	先利用基线调查年份(2017)减去每个教师(含保育员)的出生年份,再计算班级内所有教师的平均年龄	110	27.90	6.89
班级教师平均教龄	先利用基线调查年份(2017)减去每名教师(含保育员)开始从教的年份,再计算班级内所有教师的平均教龄	109	3.62	4.15
班级教师平均受教育年限	先将每名教师(含保育员)的受教育程度转化成受教育年限,再计算班级内所有教师的平均受教育年限	108	12.44	1.92
教师人数	班级内专任教师和保育员的数量总和	111	2.06	0.72
幼师比	班级儿童总数除以教师数所得	111	13.90	5.07
控制变量:幼儿园特征				
所有制类型—公办园	公办园=1;民办园=0	46	79.30%	/
所有制类型—民办园		12	20.70%	/
所在地—乡村	基于 2017 年国家统计局的城乡分类;城镇=1,乡村=0	36	62.10%	/
所在地—城镇		22	37.90%	/

需要说明的是,有 12 个样本班级有 2 名专任教师,而多数样本班级仅有一名专任教师,并未按照相关政策规定配备相应数量的保育员或助教。考虑到专任教师和保育员对儿童发展的影响不同,在回归分析中应对不同类型教师赋予不同的权重^[46]以考查其对儿童发展的影响。但受数据限制,本文借鉴菲利普森(Leeslie C. Phillipsen)等学者的做法对班级层面教师个体特征相关变量进行处理并纳入模型分析,即如果一个班超过一名专任教师,则计算均值代表该班专任教师的平均受教育年限、培训经历、教龄和工资^[47]。

(六) 模型与分析策略

本文采用教育生产函数模型对班级规模与儿童发展的关系进行估计。在传统教育生产函数模型基础上,本文引入增值模型中的协变量校正模型(Covariate Adjustment Model),即以儿童的追踪测试成绩作为结果变量,基期发展水平得分作为协变量进行估计。该模型的引入可以控制不随时间变化且无法观测到的变量,进而缓解遗漏变量偏误所造成的内生性问题。另外,考虑到分析数据呈现个体、班级和幼儿园等多重嵌套结构特征^①,参照杨钊^[48]、郑琦和杨钊^[49]等研究,本文在单一水平增值模型基础上进一步将模型拓展为多水平模型。本文使用组内相关系数(ICC)来衡量儿童发展水平在班级和幼儿园层面的差异。幼儿园层面基线调查和第一次追踪调查的 ICC 分别为 0.13 和 0.16,这说明大约 13%~16% 儿童个体发展水平差异来自园际差异。班级层面基线调查和第一次追踪调查的 ICC 分别为 0.57 和 0.496,这说明有一半左右的儿童个体发展水平差异可以由班级层面差异来解释,可见班级间差异大于园际差异,应采用“个体—班级”两水平随机截距模型^②。具体模型设定如下:

第 1 水平:儿童个体层面

$$Y_{ijs}^t = \mu_0 + \mu_1 Y_{ijs}^{t-1} + \mu_2 I_{ijs}^{t-1} + \varepsilon_{ijs} \quad (1)$$

第 2 水平:班级层面

$$\begin{aligned} \mu_{0j} &= \lambda_{00} + \lambda_{01} CS_{js}^{t-1} + \lambda_{03} T_{js}^{t-1} + \theta_{0j} \\ \mu_{kj} &= \lambda_{kj} \quad (k = 1, 2, 3 \dots) \end{aligned} \quad (2)$$

其中,被解释变量 Y_{ijs}^t 表示在第 s 个幼儿园中第 j 个班内第 i 个儿童在 t 期的发展水平的标准化得分。第 1 水平变量 I_{ijs}^{t-1} 涵盖了基线调查时儿童个体及其家庭背景特征变量,包括儿童的性别、年龄、家庭背景等协变量。 Y_{ijs}^{t-1} 代表了基线调查时儿童发展水平的标准化得分。第 2 水平变量 CS_{js}^{t-1} 是基线调查时儿童所在班级的规模; T_{js}^{t-1} 为基线调查时儿童所在班级和幼儿园层面的协变量,包括教

① 如果将跨层变量加入同一层模型,可能违背了线性回归中独立同分布和同方差假定,故使用单一层次的最小二乘法分析班级规模效应可能因不同层级数据误差项相关而导致估计偏倚。

② 在多水平模型的分析框架下,首先使用受限极大似然估计法对嵌套模型进行模型检验。在嵌套模型检验中,原假设是斜率方差为 0,即随机截距模型拟合较好,若检验结果拒绝原假设($p < 0.05$),则选择随机斜率模型。模型检验结果没有拒绝原假设,所以选择随机截距模型用于后续分析。

师特征、公民办、幼儿园所在地等变量。

在儿童发展早期阶段,孩子年龄即便只相差1岁或几个月,他们之间在认知能力、身体发育、心理成熟度上也存在明显差异^[50]。换言之,班级规模对不同年龄儿童的影响可能存在差异。南鲁普(Anne Brink Nandrup)利用丹麦公立学校所有在校生的行政数据发现,班级规模对小学阶段不同年级之间学生数学成绩的影响不同^[51]。因此,本文在估计班级规模的效应时,将样本分为小班组和中班组分别进行估计。小班组考察了儿童在小班阶段(K1, 3~4岁)进入不同规模班级对其中班阶段(K2, 4~5岁)儿童发展指标的影响,中班组考察了儿童在中班阶段(K2, 4~5岁)进入不同规模班级对其大班阶段(K3, 5~6岁)儿童发展的影响。

四、研究发现

(一) 县域内班级规模特征描述

该县教育局为研究团队提供了全县217所幼儿园班数和在园人数信息。基于上述信息,本文计算了全县幼儿园班级规模整体分布情况^①。基于政策规定,本文首先对该县幼儿园平均班级规模进行描述(见表2),再对抽样园班级规模进行描述(见表3)。2017年全县幼儿园平均班级规模的均值为26.66人,标准差为10.5人,中位数为28.26人,最小值5人,最大值69.33人。除了民办园,全县幼儿园、公办园、城镇园和农村园的平均班级规模均在《幼儿园工作规程》所规定的30人以内,但全县仍有三成及以上的幼儿园平均班级规模超过国家标准,其中以民办园班级规模超标占比最高。

从表2的描述统计可以发现,不同所有制和城乡属性的幼儿园平均班级规模存在较大的异质性。公办园平均班级规模为25.89人,民办园平均班级规模为30.08人,二者的平均班级规模在统计上存在显著差异。公办园平均班级规模的整体分布更为离散,极差较大;民办园平均班级规模分布更为集中。城镇园平均班级规模为28.05人,略高于农村园(26.21人),但二者不存在显著差异。城镇园平均班级规模的整体分布也呈现出更为离散的趋势,标准差和极差都高于农村园;农村园平均班级规模更为集中,极差较小。由此可见,在该县域内部,不同所有制和城乡属性的幼儿园在以班级规模为代表的结构性质量方面存在较大差异。

^① 因教育局没有按年龄班分别统计民办园的班数和在园儿童数,所以只能计算幼儿园层面的平均班级规模。

表 2 华北某县全县幼儿园平均班级规模的描述统计

	均值	标准差	中位数	最小值	最大值	超过平均班级规模占比(%)
全县平均	26.66	10.50	28.26	5.00	69.33	38.71
公办园	25.89	11.41	25.83	5.00	69.33	37.50
民办园	30.08	2.96	30.00	23.33	37.50	43.90
城镇园	28.05	11.32	29.50	5.00	69.33	36.99
乡村园	26.21	10.14	28.00	5.50	53.33	39.39

在 119 个样本班级中,“小额班”“常规班”和“大额班”分别占比 31.09%、22.69%和 46.22%。其中,小班阶段三类班级的占比分别为 29.82%、21.05%和 49.12%;中班阶段三类班级的占比分别为 31.03%、24.14%和 44.83%。值得注意的是,班级规模大小往往与办园规模之间呈现显著的正相关关系(相关系数为 0.32, $p < 0.001$)。研究团队的实地调查和访谈发现,班级规模大的幼儿园意味着有更多的家长选择让孩子进入该园接受教育,这体现了家长对该园教育质量的认可^[52]。

表 3 华北某县抽样园班级规模的描述统计

	平均班级规模	不同规模班级占比(%)		
		小额班	常规班	大额班
抽样班级	26.43	31.09	22.69	46.22
小班	25.00	29.82	21.05	49.12
中班	28.55	31.03	24.14	44.83

(二) 基准回归

基准回归结果如表 4 所示。模型(1)和(3)是未控制基期儿童发展水平得分的一般多水平模型,模型(2)和(4)是控制了儿童基期发展水平得分的多水平增值模型。回归结果显示,不论儿童在小班阶段就读何种规模的班级,班级规模对其升入中班后的发展水平都无显著影响;中班阶段就读“小额班”的儿童,其在大班阶段的发展水平比中班阶段就读“大额班”的同伴显著低 0.296 个标准差,中班阶段就读“常规班”的儿童与中班阶段就读“大额班”的儿童在大班阶段的发展水平上无显著差异。换言之,当儿童在中班阶段时处于“小额班”时,他们在大班阶段的表现会显著低于在中班阶段就读“大额班”的儿童。

表 4 班级规模和儿童总体发展水平关系的基准回归结果

固定效应部分	小班样本		中班样本	
	(1)	(2)	(3)	(4)
班级特征				
“小额班”(以“大额班”为参照)	-0.177 (0.216)	-0.175 (0.208)	-0.334** (0.149)	-0.296* (0.153)

(续表)

固定效应部分	小班样本		中班样本	
	(1)	(2)	(3)	(4)
“常规班”(以“大额班”为参照)	-0.184 (0.209)	-0.086 (0.203)	0.077 (0.140)	0.114 (0.145)
公办园(民办园=0)	-0.014 (0.196)	-0.083 (0.189)	-0.215 (0.149)	-0.054 (0.154)
城镇(乡村=0)	-0.003 (0.199)	-0.191 (0.192)	-0.166 (0.123)	-0.176 (0.126)
教师平均最高受教育年限	-0.038 (0.050)	-0.075 (0.048)	0.025 (0.031)	0.020 (0.032)
教师平均教龄	-0.033 (0.026)	-0.012 (0.025)	0.029 (0.019)	0.008 (0.019)
教师平均月工资(取对数)	0.288 (0.406)	0.441 (0.392)	-0.322 (0.246)	-0.339 (0.253)
教师平均年龄	-0.006 (0.013)	-0.003 (0.012)	-0.038*** (0.013)	-0.027** (0.014)
教师人数	-0.112 (0.157)	-0.097 (0.151)	-0.001 (0.098)	0.013 (0.102)
儿童个体及家庭特征				
男孩(女孩=0)	-0.027 (0.074)	-0.045 (0.064)	-0.094 (0.069)	-0.071 (0.060)
基线年龄	0.608*** (0.094)	0.220** (0.088)	0.532*** (0.085)	0.174** (0.083)
父母最高受教育年限	0.041** (0.019)	0.025 (0.016)	0.046** (0.018)	0.028* (0.016)
父母是否为主要监护人(0=否)	0.048 (0.096)	0.040 (0.083)	-0.024 (0.091)	0.009 (0.080)
农村户口(城镇户口=0)	-0.001 (0.194)	-0.061 (0.168)	-0.088 (0.152)	-0.208 (0.133)
兄弟姐妹数	-0.012 (0.045)	0.025 (0.039)	-0.059 (0.043)	-0.081** (0.038)
家庭资产指数	0.007 (0.019)	0.004 (0.016)	0.005 (0.020)	0.009 (0.017)
基线测试得分		0.569*** (0.051)		0.519*** (0.048)
截距项	-4.252 (2.729)	-2.989 (2.611)	1.036 (1.789)	2.503 (1.839)
随机效应部分				
随机截距方差	0.214*** (0.062)	0.213*** (0.057)	0.082*** (0.031)	0.107*** (0.035)
误差项的方差	0.490*** (0.037)	0.361*** (0.028)	0.429*** (0.033)	0.320*** (0.024)

(续表)

固定效应部分	小班样本		中班样本	
	(1)	(2)	(3)	(4)
观测值	403	403	405	405
对数似然值	-486.508	-434.058	-450.542	-402.715
Wald chi2	57.788	199.281	84.467	211.687

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

(三) 异质性分析

本文尝试考察班级规模对来自不同背景(城乡、所有制、家庭社会经济背景)的儿童发展影响的异质性。首先,幼儿园所处的城乡位置不同,可能在教育资源获得和质量上存在差异。其次,不同所有制的幼儿园在服务对象、服务内容和收费方面存在较大的差异^[53],来自优势家庭的孩子更有可能入读“低价优质”的公办幼儿园^[54]。再次,来自不同家庭社会经济背景的儿童对班级环境差异的反应可能存在差异。上述三种差异可能造成班级规模差异,进而对儿童发展水平产生异质性影响。

采用基准模型,本文首先分析了班级规模对城乡幼儿园和不同所有制幼儿园儿童影响的差异。表5回归结果显示,在小班样本中,小班阶段就读城镇园“小额班”的儿童到中班阶段的发展水平比其他儿童显著低0.8个标准差;在中班样本中,中班阶段就读农村幼儿园“常规班”的儿童升入大班阶段后的发展水平比中班阶段就读“大额班”的儿童升入大班阶段后的发展水平显著高0.442个标准差。换言之,若农村幼儿园的儿童在中班阶段就读于符合政策规定的“常规班”,那么其升入大班阶段后表现会显著优于中班阶段就读于“大额班”的同伴,这说明中班阶段农村幼儿园儿童从班级规模缩减中可能受益更多。

表5 班级规模对城乡幼儿园儿童发展影响的异质性检验

	小班样本		中班样本	
	农村幼儿园	城镇幼儿园	农村幼儿园	城镇幼儿园
“小额班”	-0.089 (0.263)	-0.800* (0.470)	-0.166 (0.172)	-0.554 (0.340)
“常规班”	0.049 (0.312)	-0.424 (0.349)	0.442** (0.179)	-0.153 (0.264)
控制变量	是	是	是	是
观测值	258	145	239	166

注: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; 括号内为标准误; 除未控制城乡幼儿园类型变量外,其他控制变量与基准回归增值模型部分相同。

表6回归结果显示,不同所有制幼儿园的班级规模效应的异质性体现在公

办园的中班样本中。在小班样本中,不论儿童在小班阶段进入公办园还是民办园的“小额班”或“常规班”,其发展水平在升入中班阶段后与其“大额班”的同伴相比不存在显著差异;在中班样本中,中班阶段进入公办园“小额班”的儿童,其升入大班阶段后的发展水平比中班阶段就读“大额班”的同伴显著低 0.259 个标准差;中班阶段进入“常规班”儿童的发展水平比中班阶段就读“大额班”的同伴显著高 0.333 个标准差,这意味着中班阶段公办园儿童能从“大额班”的班级规模缩减中受益更多。

表 6 班级规模对不同所有制幼儿园儿童发展影响的异质性检验

	小班样本		中班样本	
	公办园	民办园	公办园	民办园
小额班	-0.103 (0.246)	-0.725 (0.670)	-0.259* (0.146)	/ /
常规班	-0.123 (0.258)	-0.131 (0.447)	0.333** (0.163)	-0.207 (0.628)
控制变量	是	是	是	是
观测值	307	96	312	93

注：* $p<0.1$ ，** $p<0.05$ ，*** $p<0.01$ ；括号内为标准误；除未控制城乡幼儿园类型变量外，其他控制变量与基准回归增值模型部分相同。

其次,本文还考察了班级规模对来自不同家庭社会经济背景儿童的影响。按家庭资产指数的中位数,划分高家庭社会经济背景和低家庭社会经济背景两组家庭,再考察班级规模对不同家庭社会经济背景儿童的发展水平的影响。表 7 的回归结果显示,在小班样本中,班级规模对儿童发展无显著影响;在中班样本中,来自高家庭社会经济背景的儿童在中班阶段进入“小额班”后,其大班时的发展水平比“大额班”儿童显著低 0.362 个标准差,“常规班”和“大额班”儿童发展水平之间不存在显著。换言之,中班阶段来自高家庭社会经济背景的儿童进入“常规班”或“大额班”将对其进入大班阶段的增值更为明显。

表 7 班级规模对不同家庭社会经济背景儿童的影响

	小班样本		中班样本	
	低家庭社会经济背景组	高家庭社会经济背景组	低家庭社会经济背景组	高家庭社会经济背景组
小额班	-0.113 (0.239)	-0.264 (0.284)	-0.196 (0.154)	-0.362* (0.204)
常规班	0.046 (0.262)	-0.086 (0.241)	0.275 (0.168)	0.176 (0.185)
控制变量	是	是	是	是
观测值	193	210	202	203

注：* $p<0.1$ ，** $p<0.05$ ，*** $p<0.01$ ；括号内为标准误；除未控制家庭资产指数变量外，其他控制变量与基准回归增值模型部分相同。

(四) 稳健性检验

在基准回归之后,本文采用三种方法检验回归结果的稳健性。第一,使用班级规模连续变量的一次项和二次项代替班级规模虚拟变量进行稳健性检验,从而判断班级规模与儿童发展之间是否存在线性关系和阈值(threshold)。第二,用幼师比代替班级规模连续变量重新进行估计。第三,班级规模对儿童发展的影响可能因发展领域不同而存在差异。认知和非认知技能作为新人力资本理论的两大内容^[55],在儿童早期发展过程中发挥着重要作用。认知技能指个体掌握的知识以及获得新知识的能力,非认知技能则指和其他人社会交往中表现出的思想、情感和行为^[56-58]。为此,下文将使用 EAP-ECDs 量表中的认知和以社会性情感发展领域为代表的非认知技能得分替代儿童总体发展水平得分作为因变量进行稳健性检验。

1. 班级规模、幼师比与儿童总体发展水平

既有研究发现班级规模、生师比与认知和成就之间存在非线性关系^[59],因此本文在基准模型设定中以国家规定为参照将幼儿园班级分为“小额班”“常规班”和“大额班”进行检验,并发现了非线性关系。已有研究也指出幼儿园班级规模或幼师比可能存在阈值,只有在阈值范围内班级规模或幼师比才有可能对儿童发展产生积极影响。对我国幼儿园的研究发现,班级规模的阈值约为 35 人,幼师比的阈值约为 13 人^[60]。

为了检验阈值是否存在,本文纳入班级规模的一次项和二次项进行阈值检验。表 8 回归结果显示,在线性模型设定下,班级规模与儿童发展之间存在不显著的正相关关系;在非线性模型设定下,班级规模的一次项系数和二次项系数均不显著,这说明班级规模与儿童发展之间不存在阈值。幼师比是与班级规模高度相关、但又不同的结构性质量指标^[61],本文采用幼师比代替基准模型中的班级规模来检验结果的稳健性。表 9 回归结果显示,幼师比与儿童发展之间存在显著但微弱的正相关关系,但不存在非线性关系和阈值。采用幼师比与利用班级规模的发现一致,结果较为稳健。上述回归结果与现有文献的发现一致,即结构性质量与儿童发展之间只有微弱的关系^[62-63]。

表 8 班级规模的非线性检验结果

	小班样本		中班样本	
	(1)	(2)	(1)	(2)
班级规模	0.004 (0.010)	0.016 (0.050)	0.007 (0.008)	0.050 (0.049)
班级规模的二次项		-0.000 (0.001)		-0.001 (0.001)
控制变量	是	是	是	是
观测值	403		405	

注: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; 括号内为标准误; 控制变量与基准回归增值模型部分相同。

表 9 幼师比的非线性检验结果

	小班样本		中班样本	
	(1)	(2)	(1)	(2)
幼师比	0.079* (0.042)	0.011 (0.138)	0.036 (0.036)	-0.028 (0.138)
幼师比的二次项		0.002 (0.003)		0.002 (0.003)
控制变量	是	是	是	是
观测值	403		405	

注：* $p<0.1$ ，** $p<0.05$ ，*** $p<0.01$ ；括号内为标准误；在基准回归增值模型控制变量的基础上进一步控制了班级规模连续变量。

2. 班级规模对儿童不同发展领域的影响

表 10 呈现了以认知和社会性情绪发展作为因变量的回归结果。结果显示，不论儿童在小班阶段就读“小额班”还是“常规班”，其升入中班阶段后的认知发展和社会性情绪发展水平与小班阶段就读“大额班”的同伴不存在显著差异，但均低于“大额班”同伴。中班阶段处于“小额班”的儿童，其升入大班阶段后的社会性情感发展水平比“大额班”同伴显著低 0.34 个标准差。由此可见，中班同伴阶段就读“大额班”的同伴进入大班阶段后的社会性情感发展水平优于“小额班”同伴。

表 10 班级规模对儿童不同发展领域的影响

	认知发展		社会性情绪发展	
	小班样本	中班样本	小班样本	中班样本
“小额班”	-0.306 (0.194)	-0.195 (0.128)	-0.102 (0.228)	-0.340** (0.148)
“常规班”	-0.120 (0.187)	-0.114 (0.119)	-0.173 (0.220)	0.213 (0.138)
控制变量	是	是	是	是
观测值	403	405	403	405

注：* $p<0.1$ ，** $p<0.05$ ，*** $p<0.01$ ；括号内为标准误；控制变量与基准回归增值模型部分相同。

五、结论与讨论

(一) 研究结论

本文首次利用县域学前教育阶段儿童发展追踪调查数据，采用多水平增值模型分析了我国华北农村地区幼儿园班级规模与儿童发展之间的关系。研究

发现在县域内的幼儿园中,不同类型幼儿园的平均班级规模和幼师比存在较大差异,班级规模对儿童个体发展产生了较为复杂的影响。基准回归发现,小班儿童发展并未受到班级规模影响,但对中班阶段就读“小额班”的儿童升入大班阶段后的表现有显著负向影响,中班阶段就读“小额班”的儿童进入大班阶段后的发展水平得分比其他同伴显著低 0.302 个标准差。

异质性分析显示,对就读“小额班”的城镇幼儿园小班儿童与高家庭社会经济背景的中班儿童,其成绩增值低于其他儿童。同时,相比中班阶段就读“大额班”儿童,就读“常规班”的农村儿童、公办园儿童和低家庭社会经济背景儿童的成绩增值更高。稳健性分析未发现班级规模存在“阈值”,幼师比与儿童发展之间呈现不显著的正相关关系;用认知发展和非认知能力替代整体儿童发展得分后,结果依然稳健。研究结果表明,降低班级规模虽然可以改善幼儿园结构性质量,但是尚未对儿童发展产生显著的积极影响。

基于上述研究发现,本文得到两条初步结论。第一,作为学前教育结构性质量核心的班级规模,对我国幼儿园儿童以《亚太早期儿童发展量表》衡量的儿童发展的贡献有限。与“大额班”相比,小班阶段的“小额班”效果不显著,中班阶段的“小额班”效果为负向显著;以线性形式测量的班级规模与班级规模无显著相关关系。二者意味着降低幼儿园班级规模虽然能提升结构性质量,对儿童发展却未发挥积极作用,甚至在中班阶段存在异常的班级规模效应。

第二,班级规模对我国学前教育阶段儿童发展的影响具有较高的异质性,并非所有群体的儿童都能从班级规模缩减中受益。一方面,中班阶段来自农村幼儿园或低社会经济背景家庭的儿童从班级规模缩减中可能受益多,这意味着在中班阶段开展的班级规模缩减有可能对弱势群体产生补偿效应,但这一积极效果在小班阶段并不存在;另一方面,从城乡、所有制还是家庭背景上看,“小额班”儿童的发展水平在所有模型中均低于“大额班”儿童,这也说明在县域内“小额班”保育质量不佳,存在大量“无效的小班”,这与对我国初高中班级规模效应的发现一致^[64-65]。

(二) 为何存在“无效的小班”

为何幼儿园阶段“小额班”对儿童发展增值未能产生积极效果?为何存在大量“无效的小班”?本文认为有三种可能的解释:一是县域内存在幼儿园规模的家庭选择,自我选择导致了不同类型幼儿园规模及班级规模的差异。城市化进程导致农村的人口流动和教育格局发生深刻变化。由于大规模人口流动、学龄人口减少和为时十年的农村“撤点并校”政策,许多县域呈现出“城挤、乡空、村弱”的教育新格局^[66],本文也发现这种教育格局正逐步向学前教育阶段蔓延。当前学前教育实践中,各地省级、市级或区级示范性幼儿园通常拥有当地最好的教育资源与政策优势,办园规模和班级规模都比较大^[67]。办园规模大的幼儿园,通常班级规模也较大。2017 年我国农村幼儿园平均班级规模为 25.6 人,而

城区园为 28.8 人。本文样本中的农村园平均班级规模为 25.9 人,低于城镇园的 28.4 人。在县域内,城乡幼儿园规模差异可能是优质学前教育资源向城区集中所致,乡村中家庭社会经济背景较好的家长为追求较高质量的学前教育而向城镇迁移,这使得大量乡村儿童脱离农村园,进而造成幼儿园班级规模上“城挤乡空”的现象。有能力选择幼儿园的多为社会经济背景较高的家庭,因此会导致城镇园中高家庭社会经济背景的儿童聚集。高家庭社会经济背景与儿童能力发展呈正相关关系,这就会导致城镇“大额班”儿童发展水平显著高于农村“小额班”儿童的结果。基准回归中父母受教育程度与儿童发展显著正相关,这支持了上述“自我选择”的假说。

二是班级规模或者幼师比超过理论阈值。胡碧颖等学者曾以班级规模和幼师比与幼儿园过程性质量的关系为例,讨论结构性质量的可能阈值。采用广义可加模型(Generalized Additive Modeling),该研究发现在 177 个幼儿园教室中,以“情感支持和教室组织”衡量的过程性质量与班级规模和幼师比之间均存在非线性关系,班级规模的阈值为 35 人,幼师比的阈值为 13 人^[68]。参考上述方法,本文以班级规模和幼师比与儿童发展水平为例来检验是否存在阈值。遗憾的是,稳健性分析中采用班级规模和幼师比的一次项和二次项的非线性模型未发现存在阈值。此外,本文样本幼儿园 2017 年的平均班级规模为 26.84 人,低于胡碧颖等确定的班级规模阈值^[69]。因此,“小额班”积极效应的缺乏不能用样本幼儿园班级规模整体过大、超过理论阈值来解释。

三是班级规模未能通过影响过程性质量来提升儿童发展水平。文献综述显示班级规模降低以提升过程性质量,并通过过程性质量提升儿童发展水平^[70-72]。为了检验这一假设,本文利用罗伯特·皮亚塔(Robert C. Pianta)等学者开发的《课堂评价得分系统(幼儿版)》(CLASS Pre-K,以下简称 CLASS)测量了幼儿园教室的过程性质量^[73]。CLASS 通过情感支持、班级管理和教学支持三个领域十个子维度,来评估过程性质量中最为核心的部分——师幼互动。该量表在中国幼儿园环境使用中也表现出良好的测量学属性^[74]①。

根据温忠麟和叶宝娟提出的中介效应检验流程^[75],本文依次对小班和中班样本儿童进行过程性质量的中介效应检验^②。中介效应分析显示,在小班样本中,班级规模对 CLASS 的情感支持领域和教学支持领域产生了显著的积极影响,但后者对儿童发展水平无显著影响;在中班样本中,虽然班级规模对

① CLASS 可采用现场观察和视频观察两种评分方式,本项目采用的是视频观察评分法。调查员在调查当日对抽样班级进行了至少 2 小时的连续拍摄,拍摄内容包括了集体教学、区角活动、户外活动等不同活动类型。录像完成后交澳门大学教育学院团队由经过 TeachStone® 公司专业培训和认证的评分者进行评分。根据 CLASS 得分情况,可将过程性质量划分为三个等级:1~2 分代表低质量,3~5 代表中等质量,6~7 分代表高质量。在本研究中,情感支持、班级组织和教学支持三个领域的 Cronbach's α 系数分别为 0.7、0.799 和 0.881,各领域表现出较好的内部一致性。

② 受篇幅限制,中介效应模型未在此处展示。

CLASS 情感支持、班级管理和教学支持三个领域均有显著的积极影响,但只有班级管理对儿童发展水平产生了显著的积极影响。换言之,班级规模未能通过提升过程性质量全面促进儿童发展。“无效小班”与过程性质量积极效应的缺乏密切相关:若“小额班”的教师不能改变传统教学行为,那么衡量师幼互动的 CLASS 量表可能就无法识别出过程质量的改善,因此也难以发现班级规模通过影响过程质量而促进儿童发展水平的提升。

综上所述,现阶段我国县域内的农村地区“小额班”优势的缺乏可能与家庭对幼儿园的自我选择相关,也可能与教师在“小额班”中并未提供更多的情感支持、班级管理或教学支持相关,从而错失了利用小班额改进过程性质量和儿童发展的机会。班级规模与儿童发展关系的逻辑与影响机制有待进一步的研究来确定。

本文克服了部分方法论的缺陷,拓展了已有研究领域。与此前研究将幼儿视为同质性群体的做法不同^[76-77],本文关注结构性质量对不同儿童群体的异质性影响,在一定程度上验证了宋映泉提出的学前教育结构质量改善但是城乡儿童发展差异扩大的悖论^[78]。由于采用了增值评估方法,本文得以在控制儿童基线发展水平的条件下,考察前期班级规模对后期儿童发展的影响。本文采用了班级规模是否符合政府标准来定义班级规模的虚拟变量,即采用非线性的方程形式来考察班级规模效果,是一种比较理想的方程形式^[79]。通过上述方法论方面的调整,本文的结论较为稳健。

诚然,本文仍有一些局限。首先,该县的数据只能说明部分农村县域内学前教育发展状况,将该结论外推时需要谨慎。其次,择“园”行为是完全开放的,基本上由市场调节^[80],择“园”过程往往与班级规模选择的过程密切相关。即便使用多水平增值模型部分控制了遗漏变量带来的内生性问题,也难以解决班级规模自选择的内生性问题。

(三) 学前教育公共政策工具的选择

在学前教育公共政策实践中,我国政府不断提高办园标准,以“消除学前大班额”为政策工具来提升学前教育的结构性质量,促进普及普惠幼儿园的发展。已有研究主要关注以班级规模为核心的结构性质量对过程性质量或儿童发展的影响,但是尚未对我国政府学前教育公共政策导向与政策工具进行反思。

公共政策设计的核心是思考“做什么?怎么做?谁来做?”的问题,其中“做什么”是要进行政策问题的诊断和政策目标的界定,而政策问题的诊断又决定了政策方案的优先排序问题和组织实施。因此,评价学前教育结构性质量提升政策的效果,首先要考虑当前政策是否解决了学前教育发展过程中的主要矛盾问题。换言之,就是要评价班级规模缩减是否应该优先得到解决的问题。那么消除学前“城镇大班额”是不是我国学前教育事业发展的主要矛盾?政府财政性教育经费是否应该优先解决班级规模问题?已有研究指出“政策制定要以解

决主要矛盾为目标,而不是过程性手段为目标”^[81]。当前我国学前教育事业发展的主要矛盾之一是结构性质量改善与城乡儿童发展差距扩大之间的矛盾,而不是简单的优化结构性质量的问题。换言之,改善以班级规模为代表的学前教育质量是解决矛盾的过程性手段,而不是解决主要矛盾本身。因此,以班级规模缩减本身作为政策目标的做法值得商榷。

从政策工具的有效性和异质性效果来看,学前教育阶段的班级规模缩减也未必是提升结构性质量的理想方式。本文的分析指出,当前阶段在县域内农村通过缩减班级规模很难直接促进小班和中班儿童的发展,这表明学前教育阶段的班级规模缩减可能不是有效的政策干预工具。更为重要的是班级规模缩减政策的效果具有较大的异质性,即不稳定性。“小额班”对小班儿童无显著影响,对中班儿童有显著负向影响,对家庭社会经济地位较高的儿童有显著的负面影响,对农村幼儿园和家庭社会经济地位较低的中班儿童有积极影响。具有较大异质性效果的政策干预不是理想的干预措施,它可能带来很多意料之外的后果。

此外,缩减班级规模作为一项政策会带来许多变化,这些变化会随着社会情境的变化而变化,并有可能改变其有效性。有研究表明,在其他条件相同的情况下,15名儿童的班级比20名儿童的班级增加了近三分之一的教师、行政管理和设施费用。一方面,班级规模缩减并不意味着幼师比的降低,幼师比高意味着教师仍然无法与儿童进行更多一对一有效互动,从而无法提高过程性质量。另一方面,缩减班级规模会大量增加对教师的需求,这样可能会有更多低质量或不合格的教师加入,进而无法保障保教质量^[82]。虽然我国政府将缩减班级规模作为一项重要的政策工具,但是这项政策工具需要在其他政策工具的配合下才有可能产生预期的政策效果。

我国学前教育发展经历了从规模扩张到普及普惠的发展,从结构性质量向过程性质量的提升是未来发展方向^[83]。如何通过调整班级规模来促进我国学前教育事业发展是对学术界和政策制定者智慧的考验。本文在此方面做了初步探索,未来研究将关注班级规模对教师和儿童行为的影响,从而探索其影响机制。

参考文献

- [1] 中华人民共和国教育部. 2019年全国教育事业发展统计公报[EB/OL]. 2020-05-20, http://www.moe.gov.cn/jyb_sjzl/sjzl_fztjgb/202005/t20200520_456751.html.
- [2] OECD. (2019). *Education at a glance 2019*. Paris: OECD Publishing.
- [3] 黄晓婷, 宋映泉. 学前教育的质量与表现性评价——以幼儿园过程性质量评价为例[J]. 北京大学教育评论, 2013(01): 2—10.
- [4][6][12][13][54][78][81] 宋映泉. 我国学前教育事业发展主要矛盾与公共财政投入改革方向[J]. 教育经济评论, 2019(03): 19—48.

- [5][21][45] Howes, C., Phillips, D. A., & Whitebook, M. (1992). Thresholds of quality: Implications for the social development of children in center-based child care. *Child Development*, 63(2), 449—460.
- [7] Seifert, K. L. (2013). Cognitive development and the education of young children. In O. N. Saracho & B. Spodek(Eds.), *Handbook of research on the education of young children* (3ed., pp. 480). New York: Routledge.
- [8] Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., Nores, M., & Ypsilanti, M. I. (2005). *Lifetime effects: The High/Scope Perry Preschool Study through age 40*. Ypsilanti, MI: High/Scope Press.
- [9] Vitiello, V. E., Moas, O., Henderson, H. A., Greenfield, D. B., & Munis, P. M. (2012). Goodness of fit between children and classrooms: Effects of child temperament and preschool classroom quality on achievement trajectories. *Early Education and Development*, 23(3), 302—322.
- [10] Williford, A. P., Maier, M. F., Downer, J. T., Pianta, R. C., & Howes, C. (2013). Understanding how children's engagement and teachers' interactions combine to predict school readiness. *Journal of Applied Developmental Psychology*, 34(6), 299—309.
- [11] 宋映泉. 学前教育公共财政投入、事业发展趋势及公平性挑战[A]. 杨东平编. 中国教育发展报告(2014) [M]. 北京:社会科学文献出版社, 2014: 48—72.
- [14] OECD. (2018). *Engaging young children: Lessons from research about quality in early childhood education and care*. Paris: OECD Publishing.
- [15][19][26][32][33][59][61] Bowne, J. B., Magnuson, K. A., Schindler, H. S., Duncan, G. J., & Yoshikawa, H. (2017). A meta-analysis of class sizes and ratios in early childhood education programs: Are thresholds of quality associated with greater impacts on cognitive, achievement, and socioemotional outcomes? *Educational Evaluation and Policy Analysis*, 39(3), 407—428.
- [16][47] Phillipsen, L. C., Burchinal, M. R., Howes, C., & Cryer, D. (1997). The prediction of process quality from structural features of child care. *Early Childhood Research Quarterly*, 12(3), 281—303.
- [17] NICHD. (2002). Child-care structure→process→outcome: Direct and indirect effects of child-care quality on young children's development. *Psychological Science*, 13(3), 199—206.
- [18] Finn, J. D., Pannozzo, G. M., & Achilles, C. M. (2003). The “why's” of class size: Student behavior in small classes. *Review of Educational Research*, 73(3), 321—368.
- [20] Howes, C., & Rubenstein, J. L. (1985). Determinants of toddlers' experience in day care: Age of entry and quality of setting. *Child Care Quarterly*, 14(2), 140—151.
- [22] Frede, E. C. (1995). The role of program quality in producing early childhood program benefits. *Future Child*, 5(3), 115—132.
- [23] Burchinal, M. R., Cryer, D., Clifford, R. M., & Howes, C. (2002). Caregiver

- training and classroom quality in child care centers. *Applied Developmental Science*, 6(1), 2—11.
- [24] de Schipper, E. J., Riksen-Walraven, J. M., & Geurts, S. A. (2006). Effects of child-caregiver ratio on the interactions between caregivers and children in child-care centers: an experimental study. *Child Development*, 77(4), 861—874.
- [25] Huntsman, L. (2008). *Determinants of quality in child care: A review of the research evidence*. Centre for Parenting and Research, Service System Development Decision, NSW Department of Community Services.
- [27] Ruopp, R., Travers, J., Glantz, F., et al. (1979). *Children at the center: Final Report of the National Day Care Study*. Cambridge, MA.
- [28] Krueger, A. B. (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics*, 114(2), 497—532.
- [29] Nye, B., Boyd-Zaharias, J., & DeWayne, F. (1994). *The lasting benefits study: A Continuing analysis of the effect of small class size in kindergarten through third grade on student achievement test scores in subsequent grade levels-seventh grade (1992—93)*. Technical report. Academic Achievement, 54.
- [30] Chetty, R., Friedman, J. N., Nathaniel, H., Emmanuel, S., Whitmore, S. D., & Danny, Y. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project Star. *The Quarterly Journal of Economics*, 126(4), 1593—1660.
- [31] Gross, J., Backes-Gellner, U., & Balestra, S. (2017). *Does class size affect student 'Grit'? Evidence from a randomised experiment in early grades*. economics of education working paper series 0129; University of Zurich, Department of Business Administration (IBW).
- [34] Hu, B., Zhou, Y., Chen, L., Fan, X., & Winsler, A. (2017). Preschool expenditures and Chinese children's academic performance: The mediating effect of teacher-child interaction quality. *Early Childhood Research Quarterly*, 41, 37—49.
- [35] 李相禹, 刘焱. 师幼比对幼儿园集体教学质量影响的实证分析[J]. 学前教育研究, 2016(5): 3—14.
- [36] Chingos, M. M. (2013). Class size and student outcomes: Research and policy implications. *Journal of Policy Analysis and Management*, 32(2), 411—438.
- [37][39][41] Rao, N., Sun, J., Ng, M., et al. (2014). *Validation, finalization and adoption of the East Asia-Pacific Early Child Development Scales (EAP-ECDS)*. Retrieved from <http://www.arnec.net/wp-content/uploads/2015/07/EAP-ECDS-Final-Report1.pdf>
- [38][40] 周兢, 张莉, Rao N. 我国早期儿童学习与发展的现状分析:一项来自 EAP-ECDS 的研究报告[J]. 全球教育展望, 2018(07): 114—128.
- [42][52] 北京大学中国教育财政科学研究所. “县域内农村学前教育发展质量评估指标体系”项目追踪调研报告[R]. 北京: 北京大学, 2019.
- [43] Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *ETS Research Report Series*, 1, i—30.

- [44] Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure data-or tears: An application to educational enrollments in States of India. *Demography*, 38(1), 115—132.
- [46] Mukerjee, S., & Witte, A. D. (1993). Provision of child care: Cost functions for profit-making and not-for-profit day care centers. *Journal of Productivity Analysis*, 4(1), 145—163.
- [48] 杨钊. 同伴特征与初中学生成绩的多水平分析[J]. 北京大学教育评论, 2009(4): 50—64.
- [49] 郑琦, 杨钊. 班级规模与学生学业成绩——基于 2015 年 PISA 数据的研究[J]. 北京大学教育评论, 2018(4): 105—127.
- [50] 张春泥, 谢宇. 入学年龄限制真的造成了“七八月陷阱”吗?——兼评刘德寰、李雪莲《“七八月”的孩子们》[J]. 社会学研究, 2017(1): 54—77.
- [51] Nandrup, A. B. (2016). Do class size effects differ across grades? *Education Economics*, 24(1), 83—95.
- [53] 曾晓东, 刘莉. 不同所有制幼儿园服务行为分析及其政策启示[J]. 学前教育研究, 2016(08): 14—23.
- [55] 吴要武, 宋映泉. 非认知技能及干预措施研究[J]. 劳动经济研究, 2019, 7(6): 64—79.
- [56] Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Chapter 1—Personality psychology and economics. In E. A. Hanushek, S. Machin & L. Woessmann (Eds.), *Handbook of the Economics of Education* (Vol. 4, pp. 1—181): Elsevier.
- [57] Glewwe, P., Huang, Q., & Park, A. (2017). Cognitive skills, noncognitive skills, and school-to-work transitions in rural China. *Journal of Economic Behavior & Organization*, 134, 141—164.
- [58] Humphries, J. E., & Kosse, F. (2017). On the interpretation of non-cognitive skills—what is being measured and why it matters. *Journal of Economic Behavior & Organization*, 136, 174—185.
- [60][68][69][77] Hu, B., Fan, X., Wu, Y., & Yang, N. (2017). Are structural quality indicators associated with preschool process quality in China? An exploration of threshold effects. *Early Childhood Research Quarterly*, 40, 163—173.
- [62] Early, D. M., Maxwell, K. L., Burchinal, M., Bender, R. H., Ebanks, C., Henry, G. T., Iriondo-Perez, J., Mashburn, A. J., Pianta, R. C., Alva, S., Bryant, D., Cai, K., Clifford, R. M., Griffin, J. A., Howes, C., Jeon, H.-J., Peisner-Feinberg, E., Vandergrift, N., & Zill, N. (2007). Teachers' education, classroom quality, and young children's academic skills: results from seven studies of preschool programs. *Child Development*, 78(2), 558—580.
- [63][79] Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, 25(2), 166—176.

- [65] 郑琦. 班级规模的效应和机制研究——基于本土化和国际比较的证据[D]. 北京: 北京大学, 2019.
- [66] 21 世纪教育研究院. 西部农村基础教育发展报告(2019)[R]. 北京: 21 世纪教育研究院, 2019.
- [67] 赵南. 公办幼儿园的重新界定与区域发展策略——基于学前教育公共服务体系的视角[J]. 湖南师范大学教育科学学报, 2014(4): 108—114.
- [70] OECD. (2001). *Starting strong: Early childhood education and care*. Paris: OECD Publishing.
- [71] OECD. (2017). *Starting strong 2017: Key OECD indicators on early childhood education and care*. Paris: OECD Publishing.
- [73] Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS) Manual, Pre-K*. Baltimore, MD: Paul H. Brookes Publishing.
- [74] Hu, B., Fan, X., Gu, C., & Yang, N. (2016). Applicability of the classroom assessment scoring system in Chinese preschools based on psychometric evidence. *Early Education and Development*, 27(5), 714—734.
- [75] 温忠麟, 叶宝娟. 中介效应分析: 方法和模型发展[J]. 心理科学进展, 2014(5): 731—745.
- [76] Hu, B., Mak, M. C. K., Neitzel, J., Li, K., & Fan, X. (2016). Predictors of Chinese early childhood program quality: Implications for policies. *Children and Youth Services Review*, 70, 152—162.
- [80] 陈晋玲, 丁延庆, 叶晓阳. 普通高中学生择校的因果效应[J]. 教育学术月刊, 2014(5): 36—45.
- [82] Francis, J., & Barnett, W. S. (2019). Relating preschool class size to classroom quality and student achievement. *Early Childhood Research Quarterly*, 49, 49—58.
- [83] 杨宁, 任越境, 罗丽红, 周静娴, 史九倩. 过程性质量转向: 城乡学前教育均衡发展的新着力点[J]. 中国教育学刊, 2020(5): 21—27.

(责任编辑 范皑皑)

development and preschool quality collected in a national-poverty stricken county in North China during 2017—2019, the current paper analyzes the gap of urban-rural child development and disparity in both structural and process quality in kindergartens and relating trends over time. The main findings include the following: Firstly, the overall and major domains of development of children in county are significantly advanced comparing to their counterparts in rural areas. Secondly, no evidence shows that there is an overall widening trend of urban-rural child development gap over time. Additionally, while both structural quality at institutional level and process quality at classroom level have been improved significantly, the two quality indicators in county are found significantly higher than those in village. Meanwhile, with exception in domain of child care quality, there is no evidence that there is a significant increased gap over time. After exploring the institutional causes of the persisting rural-urban child development gap, this paper concludes by offering suggestions on public policies and future research regarding the inequality of early childhood education.

Preschool Class Size and Child Development: Evidence from a County Level Follow-Up Survey

YANG Po, LIU Yongye, LIANG Jing

Page 60

Class size is one of the key indicators for early childhood education's structural quality. Nationwide, the average class size has been declining since 2010; however, prior research has yet to reach a consensus on the effect of class size reduction. Based on a follow-up survey conducted in 2017 to 2018 of preschool children in a Northeastern county in China, this study utilizes a multilevel value-added model to identify the impact of class size on child development measured by the East Asia-Pacific Early Child Development Scales (EAP-ECDs). Once controlling the baseline EAP-ECDs score, class size of K1 classroom has no significant effect on children's performance at K2; K2 children in smaller classes perform significantly worse in K3 than their counterparts in regular or large K2 classes. The subgroup analysis shows that K2 children from rural households, low socio-economic status families, or public kindergartens benefit more in K3 if they attend regular rather than large K2 classes. The robustness check demonstrates that the results are robust under alternative definitions for the structural quality and the child

development. Class size reduction can improve the structural quality of early childhood education; however, it has not produced a homogeneous, positive, and significant effect on child development. There is a need for future research to identify the mechanisms behind class size reduction.

Relationship between Family's Socioeconomic Status and Parental Investments in Early Learning: Evidence from the China Family Panel Studies

SU Yufen, Nirmala Rao

Page 86

The disparities in parental investment might result in inequality in children's early learning experiences, which differentiate children's early developmental outcomes and subsequent academic learning. This study examined the relation between family socioeconomic status (SES) and parental investment in early learning. Analyses based on 6034 three- to five-year-old children from the China Family Panel Studies (2010-2016) showed that parents, on average, spent 3200.99 yuan on every child's preschool education and 93.57 yuan on enrichment activities. These two expenditures accounted for 9% and 0.3% of annual household income, respectively. Hierarchical linear modeling analyses indicated significant but modest associations among household income, parental education, and parental expenditure on children's preschool education and enrichment activities. Poorer parents spent a larger percentage of their household incomes on their children than did wealthier parents. Urban parents invested more in their children's early learning than did rural parents. From 2010 to 2016, the amount of parental expenditure on preschool education and its ratio to household income increased consistently. The findings revealed that preschool education has been an important aspect of parental investment, and it might have added economic burden to lower-SES families. Potential explanations for and implications of the findings are discussed.

On the Holistic Development of Rural Left-behind Children and the Construction of a Comprehensive Support System

ZHU Xudong, BO Yanling

Page 104

This study takes the connotation elements of children's holistic development