

# “忧郁的孩子们”：课外补习会 带来负向情绪吗？

社会  
2022·2  
CJS  
第42卷

## 基于中国教育追踪调查(CEPS)两期 数据的因果推断

李适源 刘爱玉

**摘要：**本文关注中国大陆中学生课外补习参与及其负向情绪生成的因果关系。本文基于中国教育追踪调查两期数据，借助工具变量等方法来化解内生性偏误，在量化层面上回答了课外补习参与“是否致郁”以及“使谁抑郁”等问题。本研究主要发现如下：其一，课外补习参与对中学生的情绪健康状况具有显著的“致郁效应”，但这种效应主要存在于家庭背景占劣势的较低阶层的子女当中，而家庭背景具有优势的较高阶层的子女对此类负向情绪有更强的抵御能力，表现出了“心理免疫”的现象；其二，本文从“情绪调节”的角度入手，对“致郁效应”的阶层异质性做出了进一步解释。研究结果有助于我们深入理解课外补习参与可能具有的健康不平等意涵。

**关键词：**社会分层 课外补习 负向情绪 因果推断 工具变量

DOI:10.15992/j.cnki.31-1123/c.2022.02.002

**“Melancholic Children”: Does Cram School Participation Lead to Negative Emotions? Causal Inference Based on Chinese Education Panel Survey (CEPS) Data**

*LI Shiyuan LIU Aiyu*

\* 作者1:李适源 北京大学光华管理学院社会研究中心(Author 1:LI Shiyuan, Center for Social Research, Guanghua School of Management, Peking University);作者2:刘爱玉 北京大学社会学系(Author 2:LIU Aiyu, Department of Sociology, Peking University) E-mail:layby@pku.edu.cn

\*\* 本文系国家社会科学基金项目“当代中国社会转型期的不稳定工作与性别观念的重塑”(21BSH080)的阶段性研究成果。[This research was supported by National Social Science Fund of China(21BSH080).]

特别感谢《社会》两位匿名评审专家的宝贵意见。本研究曾在谢宇教授组织的中国研究工作坊(China Workshop)上做过报告,特别感谢为本文提出宝贵建议的老师与同学。文责自负。

**Abstract:** This study focuses on the causal relationship between middle school students' participation in cram school and their emotional well-being. Our analysis is based on Chinese Education Panel Survey data (CEPS 2013–2015). We construct instrumental variables and introduce lagged variables to mitigate possible endogeneity problems. The major findings are as follows: Firstly, the participation of socio-economically disadvantaged students in cram school has significantly exacerbated their negative emotions. Students with advantaged family socioeconomic status (SES) have shown better ability in emotional regulation, and they are more likely to be psychologically immune to specific negative emotions. Secondly, we explore the class heterogeneity in the aspect of emotional regulation ability. Specifically, we propose a preliminary theoretical framework based on typology of parenting styles and related empirical findings. We speculate that parents from different social background may respond differently to their children's negative emotions, while there may also be intergroup differences in children's own emotional regulation ability. Empirically, we examine the class differences in the performance of emotional regulation of the students in cram school, and apply the counterfactual simulation approach to further illustrate the association between the group difference in emotional regulation and the group heterogeneity of the depressive effects from cram school participation. The finding provides supportive evidences for our theoretical framework. In sum, our study is helpful to better understand the implication of emotional health inequality that may derive from the shadow education system.

**Keywords:** social stratification, cram school, negative emotion, causal inference, instrumental variable

---

## 一、问题提出

伴随着中国经济改革与教育市场化的进程，围绕中小学升学考试科目而展开的课外补习班(cram school)在中国社会蓬勃兴起。近年来，初中和高中阶段的课外补习参与率以及市场规模呈攀升之势，成为教育竞争的典型文化意象(薛海平、丁小浩，2009；中国教育报，2017；孙伦轩、唐晶晶，2019)。

与此同时，青少年抑郁等负向情绪也呈现快速增长态势。在中国教育追踪调查(China Education Panel Survey, CEPS)2013—2014学年的全

国抽样数据中,超过30%的初三学生报告“过去的一周有时或经常感到抑郁”(已经过抽样权重调整)。新近的医学与心理学研究发现,抑郁症在中国青少年群体中的发生率处在高位(苗瑞凯,2020)。在部分发达地区,表现出轻度抑郁症状的中学生比例高达30%左右(张文武等,2014;牛娟等,2019)。值得警惕的是,青少年抑郁是导致自杀最常见的诱因之一(Bridge, *et al.*, 2006; Waldvogel, *et al.*, 2008),近年来不断出现的“中学生跳楼事件”也得到了社会各界的持续关注。此外,还有研究发现,班级其他学生的负向情绪可以通过班级社交网络实现“负向情绪的传染”,从而放大个体负向情绪所造成的社会心理后果(李长洪、林文炼,2019)。可见,考察青少年负向情绪的生成机制以及情绪调节境况具有紧迫的现实意义。

青少年抑郁等负向情绪的增长可能与近年来激烈的教育竞争存在关联(Lin, *et al.*, 2008; McCarty, *et al.*, 2008; Yi, 2013)。已有研究发现,参加课外补习可能会导致青少年抑郁等负向情绪的增加(Kuan, 2018; Zheng, *et al.*, 2020; Chen and Kuan, 2021)。课外补习参与对青少年情绪健康(emotional well-being)的这类副作用,可简约地称之为补习参与的“致郁效应”(depressive effect)。

课外补习与青少年的情绪健康问题,与2021年7月以来国务院颁布的“双减”(减轻义务教育阶段学生作业负担和校外培训负担)政策息息相关。<sup>1</sup>“双减”政策明确强调,“坚持从严治理,全面规范校外培训行为”“不再审批新的面向义务教育阶段的学科类校外培训机构”,政策还提出要“着眼学生身心健康,保障学生休息权利”。可以说,青少年的心理健康问题得到了国家层面的关怀和重视,决策者也试图通过对课外补习的约束和规制以保障学生的身心健康。

保障青少年的心理健康对于经济社会的可持续发展具有关键作用。从经济增长的角度来看,课外补习和青少年的情绪健康问题关涉到“教育人力资本”和“健康人力资本”的积累(Becker, 2007; Hartog and van den Brink, 2007; Azarnert, 2020)。从社会分层的视角看,由于家庭教育资源和教养方式的阶层差异,青少年的身心健康问题有可能会扩大

1. 中共中央办公厅、国务院办公厅印发《关于进一步减轻义务教育阶段学生作业负担和校外培训负担的意见》。参见教育部官网 [http://www.moe.gov.cn/jyb\\_xxgk/moe\\_1777/moe\\_1778/202107/t20210724\\_546576.html](http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/202107/t20210724_546576.html), 网页浏览于2021年9月25日。

既有的教育不平等。由此观之,保障青少年的心理健康与顺利成长对于促进社会公平和可持续发展具有深远的现实意义。

目前,决策者已经注意到补习机构的无节制增长(以及课外补习的过度参与),可能是导致青少年学业负担过重以及引发情绪健康问题的原因。但从学理上看,课外补习与青少年情绪健康的因果关系仍需进一步的澄清和检验——课外补习在多大程度上会威胁到青少年的情绪健康?课外补习究竟威胁到哪些青少年的情绪健康?这些疑问都有待严谨细密的研究。

考虑到不同社会阶层的成员可能在主观认知、情绪体验、行为选择等维度上存在异质性(Duncan, *et al.*, 1988; Xie, 2007; 谢宇, 2006),课外补习给青少年情绪健康带来的后果可能也存在阶层差异。例如,即使参加了类似的课外补习,某些阶层的子女可能在补习过程中持续积累负向情绪,另一些子女却对补习过程中派生的负向情绪具有更好的调节能力,甚至对负向情绪表现出“心理免疫”的现象。我们在评估课外补习对青少年情绪健康的因果影响时,需要更细致地考察不同家庭阶层背景(或称家庭社会经济地位)的学生在补习参与中的情绪健康是否具有相异甚至相反的表现。

从实践层面看,因果效应的异质性对于政策评估和政策制定具有重要意义。课外补习的情绪健康后果若的确存在阶层分异,那么“双减”政策在推行的过程中可以针对不同社会群体的背景与需求状况,做到有的放矢、精准施策,从而增进政策效力。

从学理层面看,课外补习是否会带来抑郁等负向情绪的积累,在多大程度上会带来负向情绪,这些问题仍然存在分歧;其次,课外补习参与的“致郁效应”是否存在阶层异质性,究竟存在何种模式的阶层分异,目前也少有经验研究给出明确答案,有待进一步的实证检验。

本文的研究问题和结构安排如下:首先,基于全国性的调查数据和工具变量方法,考察参加课外补习是否显著增加中学生的负向情绪。在回答补习“是否致郁”的基础上,进一步探究课外补习究竟“使谁抑郁”?“致郁效应”是否在不同家庭阶层背景的学生群体间存在差异?最后,本文尝试解释“致郁效应”为何存在阶层分异。

我们对研究问题做出两点说明:其一,本文关注的课外补习是指“围绕中小学的升学考试科目(如语文、数学、英语),以提升考试成绩

为导向的学业辅导班”(Bray, 2013; Kuan, 2018; Zheng, *et al.*, 2020), 不包括“艺术、体育等非学业导向的课外兴趣班或特长班”;其二, 本文关注的负向情绪主要指“抑郁情绪”(depressive mood), 可将其理解为“沮丧低落的、暂时性的负向情绪状态”。有必要将抑郁情绪与“抑郁症”(major depressive disorder)加以区分, 抑郁症是指“长期持续地处于低落消沉的情绪、对日常生活失去兴趣、且自身难以调节的心境障碍”(De Zwart, *et al.*, 2019)。短暂的抑郁情绪未能及时调节而逐渐累积, 可能会进一步转化为抑郁症。本文对情绪健康的分析主要围绕抑郁情绪展开。

## 二、文献回顾与研究框架

本文的核心问题是课外补习参与“是否致郁”以及“使谁抑郁”这两大问题。针对课外补习“是否致郁”的问题, 现有经验研究存在两种竞争性的观点: 其一, 补习参与有利于缓解同辈压力及其引发的负向情绪, 发挥了“安慰剂”的功效(Zheng, *et al.*, 2020; 孙伦轩、唐晶晶, 2019); 其二, 参加课外补习可能会增加青少年的负向情绪, 具有致郁效应(Bray, 1999, 2013; Kuan, 2018; Zheng, *et al.*, 2020; Chen and Kuan, 2021)。针对“使谁抑郁”的问题, 少有本土研究考察“课外补习的情绪健康后果”是否存在总体异质性, 也未能深入讨论情绪健康后果的组间分异模式。本文的文献回顾正是基于上述两大经验困惑(课外补习“是否致郁”以及“使谁抑郁”)而展开。

### (一) 是否致郁: “致郁效应”与“安慰剂效应”的分歧

课外补习参与对情绪健康状况的影响存在两种竞争性的观点。第一种观点认为, 课外补习参与对于情绪健康状况存在致郁效应, 即补习参与可能增加青少年的抑郁情绪。致郁效应的可能机制是, 长时间、高强度的课外补习对学生可自由支配的闲暇时间以及亲子之间的互动具有“挤出效应”(crowding out effects), 导致学生陷入高度紧张的刚性时间约束当中, 这将对情绪健康产生威胁。其中最直观的影响是参加课外补习对青少年的睡眠剥夺(Kuan, 2018), 而睡眠时间缺乏及各类睡眠问题是导致负向情绪积累的重要诱因(Yoo, *et al.*, 2007; Pasch, *et al.*, 2010; Gujar, *et al.*, 2011; Gruber, *et al.*, 2012)。另一种观点认为, 课外补习参与对情绪健康状况具有“安慰剂效应”。也就是说, 课外补习并不必然提升学业成绩, 却能有效缓解同伴压力, 进而缓解抑郁等负向情绪

(Zheng, *et al.*, 2020; 孙伦轩、唐晶晶, 2019)。更具体地说, 青少年参加课外补习受到同伴间社会比较的驱动。个体了解到同伴参加了课外补习, 担心同伴成绩超过自己, 于是选择参加补习。在接受课外补习之后, 个体得到一定程度的“心理安慰”(Zheng, *et al.*, 2020)。

目前, “安慰剂效应”与“致郁效应”都得到了不同程度的实证检验。关秉寅(Kuan, 2018)基于中国台湾地区21世纪初的教育追踪调查数据, 利用倾向得分逆概率加权(inverse probability treatment weighting, IPTW)处理内生性问题之后, 发现参加课外补习会显著增加九年级学生的抑郁情绪。此后, 陈怡蓓和关秉寅(Chen and Kuan, 2021)使用时效性更强的“台湾后期中等教育长期追踪调查”(TUSD)2014—2015学年数据, 基于类似的实证策略, 发现对于“初中三年始终参加课外补习的学生”和“在中途退出课外补习的学生”而言, 参加课外补习会显著增加这两类学生的抑郁情绪, 一定程度上印证了关秉寅(Kuan, 2018)之前的研究发现。

与之不同的是, 有学者基于大陆地区的中国教育追踪调查(CEPS)数据发现, 课外补习显著增加了学生的自信心和自我效能感, 而课外补习的“致郁效应”只在长时间过度补习的情形下才显著存在(Zheng, *et al.*, 2020)。孙伦轩和唐晶晶(2019)也印证了补习参与的“安慰剂效应”, 他们使用倾向值匹配基础上的双重差分法(PSM-DID)发现, 课外补习的参与显著改善了学生的情绪健康水平, 能够减少抑郁等负向情绪水平。

## (二) 经验分歧的可能原因

关于“安慰”与“致郁”的分歧, 笔者认为除了研究资料存在时间、地域和结果变量测量上的差异以外, 不同研究对内生性问题的处理方式可能是造成分歧的成因之一。在课外补习对情绪健康的因果效应识别过程中, 主要的内生性问题可归结为两类: 一是遗漏变量与自选择偏误, 二是反向因果问题。

首先讨论遗漏变量与自选择偏误。这一问题的成因是核心自变量取值并非“随机分配”, 即个体是否参与课外补习是由学生自身(或者家长、老师)决定的, 而不是以随机实验的方式来指派。因此, 可能存在未观测到的混淆因素, 该因素既与课外补习的参与决策相关, 又影响到学生的情绪健康, 那么该混淆因素可能导致遗漏变量偏误。例如, “学生的

性情特质”一般难以观测。一方面,性情偏悲观的学生更可能秉持“补课无用论”,参加课外补习的主观意愿相对更低;同时,性情偏悲观的学生或更易产生消极的自我暗示,抑郁等负向情绪也相对更多。因此,性情倾向便是一个典型的遗漏变量,它可能使补习参与对情绪健康状况的因果效应估计出现偏误。

另一个问题是反向因果偏误。本研究的核心自变量是课外补习的参与,因变量是学生的负向情绪水平。心理学和教育学的研究指出,情绪健康状况也会对学生课内外学习投入的积极性产生影响(Chen and Li, 2000; Needham, *et al.*, 2004; Crosnoe, *et al.*, 2007; McLeod and Fettes, 2007; McCarty, *et al.*, 2008)。具体来说,情绪健康状况较差(负向情绪倾向较重)的学生,学习动力较弱,参加课外补习的意愿相对较低。负向情绪(结果变量)对于课外补习(原因变量)具有“负反馈效应”,这可能导致基于普通最小二乘法(ordinary least squares, OLS)得出的补习致郁效应受到系统性的低估。原因在于,课外补习带来的负向情绪积累使学生在未来继续参加补习的可能性下降,进而稀释了未来补习带给学生的致郁效应。

涉及课外补习与负向情绪的本土研究主要使用的分析方法是倾向值匹配(propensity score matching, PSM)以及双重差分法(difference in difference, DID)。但倾向值匹配只能化解由可观测的混淆因素带来的因果推断偏误,较难克服无法观测的遗漏变量的干扰,也难以彻底解决反向因果问题。此外,由于补习参与是微观主体的选择(而非外生的政策冲击),双重差分法的适用性也值得商榷。

为了进一步化解遗漏变量偏误与反向因果带来的内生性问题,本文使用工具变量法(instrumental variable, IV)和滞后变量法(lagged variables)进行因果推断。有效的工具变量能够更好地化解遗漏变量和反向因果问题,较好地捕捉到补习参与对中学生负向情绪的真实因果效应(Angrist, *et al.*, 1996; Angrist and Pischke, 2009; Morgan and Winship, 2015)。

### (三)使谁抑郁:补习参与的异质性后果

关于补习参与“是否致郁”这一问题,只能在总体平均的意义上回答课外补习对青少年情绪健康的后果。无论数据结果表明补习参与是“致郁”还是“安慰”,两者都可能忽略了平均因果影响所掩盖的总体异质性。

总体异质性可以理解为总体内“不同个体之间的差异”,正是因为异质性的存在,“相同的原因会导致不同的结果”(谢宇,2006:48)。在本研究关注的情境中,即使受访学生参与了相同数量和质量的课外补习,不同学生的学业收获、对课外补习的态度评价和情绪体验也可能是不同的。也就是说,补习参与的情绪健康后果很可能是“因人而异”。关注并捕捉总体异质性,不仅有助于在学理上进一步澄清被总体均值掩盖的群间差异,更有助于为社会政策提供更加明确、细致和精准的目标指引。

近年来,在有关青少年发展的经验研究中,针对青少年的“发展型结果”(developmental outcomes)(Kim,2011)的异质性分析日益得到重视。有学者将美国人口调查数据和“佩里学前教育干预”(Perry preschool intervention program)的实验数据相匹配,发现“佩里学前教育干预”更有助于改善家庭社会经济地位处于劣势的子女的非认知能力。教育干预的良性影响在社会阶层较低的孩子当中也更持久和深远(Xie,*et al.*,2020)。再如,有关婚姻家庭和子女养育的西方研究发现,父母离异对子女发展的影响并不都是负面的,父母离异对子女教育成就的不利影响主要存在于社会阶层较高的子女群体中(Brand,*et al.*,2019)。

情绪健康状况作为身心健康的重要指标,也可视为青少年的一项“发展型结果”。在围绕补习参与的经验研究中,值得一提的是关秉寅(Kuan,2018)对中国台湾地区中学生情绪健康的异质性分析。他按“补习参与概率”的高低将样本分为五组,分别估计每组学生参加补习的情绪健康后果。结果发现,“致郁效应”仅在“补习参与概率较高”的第四组学生中统计显著,其他各组学生的抑郁情绪并未因补习参与而显著增加。此外,关秉寅的新近研究也发现补习参与的致郁效应的严重程度可能依赖于“补习参与的模式”(Chen and Kuan,2021)。不过,上述研究的异质性分析主要着眼于补习过程本身,并未直接指向情绪健康后果的阶层异质性。

#### (四)理论框架:教养方式与情绪健康后果的分异模式

捕捉和检验情绪健康后果的阶层异质性及其具体的分异模式,有助于研究者厘清和评估影子教育的教育不平等与健康不平等意涵,并提供更为细致精准的政策指引。

本研究从家庭教养方式(parenting style)的阶层分异出发,在理论

脉络上对情绪健康后果的阶层异质性做出推断与解释。家庭教养方式的阶层分异自 20 世纪以来就得到了持续关注 (Kohn, 1963; Dornbusch, *et al.*, 1987; Lareau, 1987, 2011; Steinberg, *et al.*, 1992)。黄超 (2018) 结合本土情形与经典文献, 按照“沟通频繁度”与“要求严格程度”划分出四种教养方式, 分别是“权威型”(亲子沟通频繁, 要求严格)、“宽容型”(沟通频繁, 要求不严格)、“专制型”(沟通较少, 要求严格)、“忽视型”(沟通少, 要求也少)。黄超 (2018) 发现, 中国家长对其子女普遍具有高要求和高期待, 教养方式分异的关键在于沟通维度——较高阶层的家长多用权威型教养方式, 较低阶层的家长多用专制型教养方式。更重要的发现是, 权威型教养方式更有助于培养子女的非认知能力 (non-cognitive skills), 子女在自我效能、社交能力和集体融入等多方面都有更好的表现 (黄超, 2018)。

家庭教养方式的阶层类型学理论及本土经验为本研究探索情绪健康后果的阶层异质性提供了思路。在补习参与过程中, 家长对子女负向情绪的体察与引导以及学生自身的情绪调节能力, 可能是情绪健康后果因人而异的关键因素。具体来说, 较高社会阶层的家长更可能采取“积极回应, 注重沟通”的教养方式, 更及时地体察和回应子女在补习中派生的负向情绪, 引导子女进行情绪调节。同时, 若将情绪调节能力视作非认知能力, 那么较高阶层的家长的教育方式可能更有利于培养子女的情绪调节能力; 相比之下, 较低社会阶层的家长由于将相对稀缺的经济资源用于课外补习, 可能会要求子女在课外补习中取得适当的学业回报。当子女受到负向情绪困扰时, 较低阶层的家长可能无法及时察觉到子女情绪变化 (并展开有效的沟通、引导与回应)。如果从非认知能力的角度来看, 较低阶层的家长的教养方式也不利于培育子女自身的情绪调节能力, 导致影子教育引发的负向情绪在较低阶层的子女群体中更明显。

上述理论框架提供了一种经验预测——补习参与的情绪健康后果可能存在“劣势—特异”与“优势—免疫”的阶层分异模式。换言之, 课外补习引发的负向情绪可能是较低阶层子女的“特异性后果”, 而较高社会阶层的子女对此类负向情绪具有较好的“心理免疫能力”。进一步看, 由于负向情绪的长期累积会对学生未来的学业表现、身心发展带来持续的不利影响, 课外补习可能借助健康不平等的生产, 隐秘地促成了

教育分层格局的维持与延续。因此,围绕情绪健康后果展开的异质性分析,对于进一步揭示教育与健康不平等的再生产具有启发意义。

### (五) 异质性分析的具体策略

从方法策略的角度看,围绕青少年“发展型结果”而展开的异质性分析仍有待商榷。通常做法是按照接受处理的倾向得分(如参加课外补习的概率)将全样本分组,然后考察组内因果效应在不同组间的差异。

由于倾向得分将家庭背景的各项测量指标(父母收入、教育、职业等)“合众为一”,在某些情形下,倾向得分所表达的社会阶层信息可能是含混的。例如,学生家庭背景的各指标与学生是否参加补习的相关性并不一定是统一且同向的,可能有一些家庭背景指标与补习参与正相关,而另一些家庭背景指标与补习参与负相关。另外,由于分组使用的倾向得分是基于样本数据估计得出的派生变量(*generated variables*),而非总体的真实变量,基于倾向得分而获得的组间分异模式可能受到抽样误差和样本选择偏差的干扰(*Breen, et al., 2015*)。因此,若以倾向得分作为家庭背景的分组变量进行异质性分析,其结果可能具有误导性。

综合考虑,本研究决定采用多个相对客观的家庭背景指标(而非估计出的倾向得分)来划分全样本,通过分组回归的方式考察情绪健康后果的阶层异质性。具体来说,家庭背景的分组标准是多分类的家庭经济条件、父母的最高受教育程度以及父亲的职业类别。

## 三、数据、变量与方法

### (一) 数据来源

本文使用中国教育追踪调查(CEPS)的两期数据,包括基期数据(2013—2014 学年)和追访数据(2014—2015 学年)。该调查采取 PPS 抽样,在基线调查中抽取了 112 所学校、438 个班级的约 2 万名初中生作为样本;第二期调查成功追访到了近 1 万名初一(升初二)的学生。

本研究的样本主要选取的是基线调查(2013—2014 学年)的七年级学生,这批学生在 2014—2015 学年再次受访。后续分析主要使用追访数据,同时也将基期数据中的某些变量作为滞后变量进行统计控制。在本研究使用的所有变量上均不存在缺失的受访者进入研究样本,最终保留的样本共 6 535 人。

### (二) 变量与测量

本文参考了代表性研究的建模策略,对控制变量进行了全面且审慎的选取。

### 1. 原因变量

本文的核心自变量(原因变量)是课外补习参与,操作化为“最近一年来,学生是否参加奥数/普通数学/语文/英语补习班(参加设为1,否则为0)”。研究样本中,2014—2015学年参加课外补习的学生占全样本的38.68%(2 528人),本学年未参加课外补习的学生占全样本的61.32%(4 007人)。

我们比较了不同家庭背景(家庭经济条件、父母文化程度、父亲职业类别)的学生在本学年的补习参与率,发现补习参与率与学生的家庭社会经济地位呈正相关关系。家庭背景处于较高阶层的子女补习参与率相对更高。<sup>2</sup>

### 2. 结果变量

本文的因变量(结果变量)是中学生的情绪健康状况,操作化为“负向情绪得分”。本文借鉴了代表性研究对抑郁情绪的处理方法(Kuan, 2018; Chen and Kuan, 2021),使用项目反应理论(item response theory, IRT)的定序计分模型(graded response model),假定受访者负向情绪水平与受访者对情绪量表题目的应答概率服从逻辑斯蒂函数关系,然后使用最大似然估计法(maximum likelihood estimation, MLE),将CEPS问卷中负向情绪量表的应答结果转化为每位受访者的负向情绪IRT得分。<sup>3</sup>

研究样本2014—2015学年的负向情绪IRT得分在-1.53到2.75之间,均值和中位值都接近0,标准差为0.93。负向情绪得分越高表明受访者的负向情绪积累越多,负向情绪倾向程度越严重。我们比较了不同家庭背景(家庭经济条件、父母文化程度、父亲职业类别)的子女负向情绪得分的组间差异,发现负向情绪的严重程度与学生家庭的社会经济地

---

2. 受篇幅限制,相应的描述性统计结果未在文中展示,有需要的读者可以联系作者索取。

3. 在CEPS(2014—2015)问卷中,负向情绪量表包含以下几个与抑郁情绪相关的问题:在过去的七天内,你是否有以下感觉:(1)沮丧.(2)消沉得不能集中精力做事.(3)不快乐.(4)生活没有意思.(5)提不起劲来做事.(6)悲伤难过。上述每一个问题,应答项为五分定序选项:(1)从不.(2)很少.(3)有时.(4)经常.(5)总是。出于稳健考虑,笔者使用了主成分分析的方法得到负向情绪的标准化因子得分。两种测量结果的样本相关系数达到0.95以上,分析结果也基本一致。简洁起见,下文报告的因变量均为负向情绪的IRT得分。

位呈负相关,家庭阶层较高的子女负向情绪得分更低。<sup>4</sup>

### 3. 控制变量

控制变量分为五类:学生特征、家庭特征、班级特征、时间固定效应、地区固定效应。

(1)学生特征。其一,人口学特征,包括学生性别、年龄、民族、兄弟姐妹数量、户籍类型、迁移流动状况(分为本地学生、省内流动、跨省流动三类)、五分类自评身体健康状况。其二,前期的学业能力,包括七年级时的认知能力标准得分、七年级的语数外标准化期中成绩<sup>5</sup>、(自评的)六年级时学习语数外的吃力程度。其三,居住安排,包括是否与双亲同住、是否寄宿、家庭所在社区的类型。其四,亲子关系,包括与父母亲的亲近程度、学生的日常生活与照料由谁负责、学生在家的学业辅导由谁负责。

(2)家庭特征。其一,家庭收入,包括家庭经济条件<sup>6</sup>、家庭收入的相对水平<sup>7</sup>。其二,家庭文化,包括父亲、母亲的受教育程度,物质性文化资本(家中藏书量、学生是否有独立的书桌、家中是否有电脑与网络)。其三,双亲职业,职业类型根据 CEPS 问卷的职业分类,转化为多个虚拟变量。其四,监督管理,操作为对子女学习生活的要求与监督强度。<sup>8</sup>其五,亲子沟通,操作为与子女讨论学习及生活问题的频率。<sup>9</sup>

在后续的阶层异质性分析中,本研究将家庭经济条件、父母亲的最高受教育程度以及父亲的职业类型作为家庭背景的分组变量。选取理由基于两点:从社会分层的经典理论来看,收入、教育与职业通常被视

---

4. 受篇幅限制,相应的描述性统计结果未在文中展示,有需要的读者可以联系作者索取。

5. 七年级的语数外三科期中考试成绩已经由 CEPS 项目官方进行了标准化处理,在不同的个体之间具有较好的可比性。

6. 这是由受访学生或家长报告的五分类家庭经济条件(非常困难、比较困难、中等、比较富裕、很富裕);考虑到某些类别回答人数较少,CEPS 项目将“非常困难”与“比较困难”合并为一类,将“比较富裕”和“很富裕”合并为一类,生成了三分类家庭经济条件。分析发现,使用学生报告或家长报告的家庭经济条件(无论是五分类或者三分类),主效应系数的效应幅度和标准误并未发生明显变化。因此,若无特别说明,下文统一使用“家长报告的五分类家庭经济条件”作为控制变量。

7. 这是由受访学生的家长报告的相对社区周围人的家庭收入水平(由低到高,共五个序次)。

8. 测量方式是家长对孩子“作业考试、学校表现、上学时间、回家时间、穿着打扮、上网时间、看电视时间”等方面的管理严格程度(不管、管但不严、管得很严)的频次加总。

9. 测量方式是家长主动与孩子讨论“学校发生的事情、孩子与朋友的关系、孩子与老师的关系、孩子心情、心事与烦恼”等事项的频率(从不、偶尔、经常)加总。

作社会经济地位(SES)的核心构成(Blau and Duncan, 1967);从分析策略看,相较于连续型倾向得分,父母教育和职业类型等变量均为多分类变量,天然具备“社会分组”的性质,能更清晰明确地表达出学生所属的家庭阶层信息。当然,这样的分组标准也有不足之处,在分析步骤和结果解读上相对更加繁琐和庞杂。

(3)班级特征。<sup>10</sup>其一,学生是否得到语数外三科老师或班主任老师的提问、表扬、批评。其二,学生对班级班风、班级同学友好程度的评价以及参加班级活动的频率。其三,学生对所在学校的喜爱程度、对同校学生的亲近程度、转校意愿。

(4)时间固定效应。即在回归分析中纳入学期虚拟变量(秋季1,春季0),目的是控制那些只随时间变动而不随个体变化的异质性。

(5)学校固定效应。使用普通最小二乘法(OLS)估计时为每个学校分别设置虚拟变量,将它们纳入回归方程,目的是控制学校层面难以观测的异质性。

(6)区县固定效应。在使用工具变量两阶段最小二乘法(TSLS)估计时,为每个区县分别设置虚拟变量,目的是控制区县层面难以观测的异质性。

### (三)模型设定与实证策略

#### 1. 基准模型设定:OLS及其内生性问题

首先估计基准回归模型,只纳入当期(第二轮CEPS追访数据)的控制变量进行OLS多元回归,进而得到课外补习参与对负向情绪得分的斜率系数估计值。

回归方程如(1)式所示。<sup>11</sup>其一,结果变量 $Y_{iscit}$ 是处在时期 $t$ 、区县 $c$ 、学校 $s$ 的受访学生的负向情绪得分;其二,原因变量 $cramsch_{iscit}$ 为学生 $i$ 在最近一年是否参加课外补习班(学业辅导班)的虚拟变量,参加取值为1,未参加取值为0;其三, $X_{iscit}$ 为学生(及其家庭、班级)层面的控制变量;其四, $\mu_s$ 为学校固定效应(学校虚拟变量);其五, $\lambda_t$ 为学期虚拟变量;其六, $\varepsilon_{iscit}$ 为回归方程的误差项(error term),它包含了所有未观测到的、且对学生 $i$ 负向情绪IRT得分具有影响的因素。

10. 班级特征已通过IRT模型生成潜变量标准化得分,以缓解多重共线性。

11. 出于表达简洁考虑,以下所有回归方程的截距项都包含在了控制变量向量所对应的系数当中。

$\beta_{1,obs}$  为本文所关心的课外补习对负向情绪的因果效应参数。如果在给定控制变量的条件下,误差项  $\varepsilon_{isct}$  与原因变量  $cramsch_{isct}$  不相关,则基于 OLS 得到的样本回归系数  $\hat{\beta}_{1,obs}$  是对因果效应参数  $\beta_{1,obs}$  的一致估计量 (consistent estimator)。

$$Y_{isct} = \beta_{1,obs} \cdot cramsch_{isct} + X'_{isct} \delta + \mu_s + \lambda_t + \varepsilon_{isct} \quad (1)$$

### 2. 引入滞后变量:对基准模型的初步改进

为了初步缓解内生性问题,我们在基准回归模型中加入了受访学生的“前期学业能力”以及结果变量的滞后项(前期负向情绪),作为控制变量。具体而言,纳入了受访学生七年级的语数外三科期中考试成绩和六年级时学习语数外的吃力程度,将其视作对前期学业水平的度量;同时,也加入了学生在七年级时的认知能力测试标准化得分,作为学业能力的代理变量。在下面的回归方程(2)式当中,列向量  $X_{isc,t-1}$  就代表了上述的“前期学业能力”。

另外,本文还纳入了受访学生在上一学年的负向情绪 IRT 得分。假定结果变量(负向情绪)对原因变量(课外补习参与)的反馈效应存在“时滞”,即只有上一期的结果变量才会影响到当期的原因变量。那么给定上一期结果变量后,当期的结果变量就不再影响当期原因变量的取值。在这种情况下,纳入结果变量的滞后项能较好地化解反向因果问题 (Manski, 1993; 陈云松、范晓光, 2010)。

$$Y_{isct} = \beta_{1,obs} \cdot cramsch_{isct} + X'_{isct} \delta + \mu_s + \lambda_t + \rho Y_{isc,t-1} + X'_{isc,t-1} \gamma + u_{isct} \quad (2)$$

### 3. 工具变量法:内生性问题的进一步处理

本研究使用工具变量法以进一步化解内生性偏误。本文的工具变量是个体所对应的“位似群体”的补课参与率。“位似群体”是指与学生本人处在同一区(县)但不同学校的其他学生群体。本文工具变量的选取思路借鉴了既有实证研究 (Zhang, 2013), 并对工具变量的构造方法做出了改进。

本研究中工具变量的含义是,“与学生个体  $i$  处在同一区(县)而不同学校的‘其他学生’当中,参加课外补习的人数(占‘其他学生’总人数)的比例”。本文工具变量的数学表达式如(3)式所示。在第二行展示的分式中,分子一项度量了“学生个体所在区(县)其他学生是否参与补习的样本频次加总(在计算中排除了与本人同校的学生)”,分母一项度量了“学生个体所在区(县)其他学生的总人数”(计算中排除了与本人

同校的学生)。

$$IV_{set} = \overline{Cramschi_{c-\alpha}} = \frac{\sum_{c=1}^{N_c} \sum_{i=1}^{N_s} cramschi_{ic} - \sum_{i=1}^{N_s} cramschi_{ic} (s=s_j, c=c_k)}{\sum_{c=1}^{N_c} \sum_{i=1}^{N_s} student_{ic} - \sum_{i=1}^{N_s} student_{ic}} \quad (3)$$

基于工具变量进行因果推断,关键前提是工具变量的有效性,这依赖于工具变量的相关性与外生性。相关性是指工具变量(“位似群体”补课参与率)与原因变量(学生个体是否参与课外补习)的总体(偏)相关系数不为零。本工具变量实质上度量了区(县)层面的中学生补课参与水平。一定程度上,它反映了该地区课外补习市场“局部均衡”的结果。一方面,工具变量可反映学生所在区(县)的其余八年级学生群体对课外补习的需求;另一方面,工具变量也能反映该区(县)补课机构的丰富程度。从需求侧看,地区性补课需求很可能通过地区内人际网络发生信息传递;从供给侧看,地方补课机构的丰富程度也可能影响个体参与课外补习的难度与机会分布(如开班数量、补习费用)。这说明工具变量与原因变量很有可能是相关的。从实证结果角度看,一阶段回归的 F 统计量远大于 10,从统计学意义上来看不存在“弱工具变量”问题。

其二是工具变量的外生性,即要求工具变量与误差项  $\varepsilon_{ic}$  不相关。也就是说,在给定控制变量的条件下,本文的工具变量(“位似群体”的补课参与率)不能与其他未观测到的、影响个体情绪健康的混淆因素产生关联。以下提供三点论证。

首先,本文的工具变量反映了地区层面的补课参与率,可能受到地区间经济文化发展差异的影响,导致工具变量与未观测到的地区异质性发生关联。因此,本文设置了区(县)固定效应,仅在每个区(县)内部进行因果识别,在很大程度上降低了上述问题发生的可能。其次,本研究工具变量度量了地区层面的宏观特征,它不太可能对微观个体的情绪健康状况产生直接影响。第三,最为重要的是,本研究在构造工具变量的过程中有意地排除了“与个体  $i$  处在同一学校的学生群体”。有关同伴效应的代表性研究指出,与个体同班的同学或同校校友可界定为个体的“同伴群体”(杜育红、袁玉芝,2016;杜育红、郭艳斌,2019),而同伴群体的社交网络具有“情绪传染”的可能(李长洪、林文炼,2019)。换言之,课外补习行为可能直接影响到个体的情绪状况。例如,同伴群体竞相参加补课,可能使尚未参与补课的学生感到同辈压力增大,导致负面情绪增加。因此,本文的工具变量中有意排除了基于“强关系”的同

伴效应的干扰。本文认为,对个体情绪健康状况有较大程度和直接影响的人群很可能是距个体较近的同伴群体,即那些与个体同班或同校的学生群体。而那些校外的、距离更远的“位似群体”的补课行为很有可能不会直接影响到个体的情绪健康,需借助个体所在学校或班级的同伴群体行为特征对个体情绪健康带来间接影响,所谓“远亲不如近邻”。本研究在工具变量的构造中已阻断了上述这条间接作用路径;同时,本文并不否认个体有重要或亲密的好友与个体处在不同学校,但是在随机抽样的情形下,身处外校的受访者很难对本校的学生产生系统性的同伴效应。

就具体估计方法而言,本文使用两阶段最小二乘法(two stage least squares, TSLS)得到 $\hat{\beta}_{1, TSLS}$ (即课外补习参与对负面情绪得分的因果效应的估计结果)。基于工具变量的 TSLS 回归方程如下式(4)和(5)所示。

$$Cramschi_{isct} = \pi_1 \cdot IV_{sct} + X'_{isct} \alpha + \mu_c + \lambda_t + \kappa Y_{isct,t-1} + v_{isct} \quad (4)$$

$$Y_{isct} = \beta_{1, TSLS} \cdot \widehat{cramschi}_{isct} + X'_{isct} \delta + \mu_c + \lambda_t + \rho Y_{isct,t-1} + \varepsilon_{isct} \quad (5)$$

为了进一步保障工具变量的外生性,可以在上述两个方程(4)和(5)中加入“学生的前期学业水平和认知能力”这一组滞后变量 $X_{isct,t-1}$ 。理论上而言,如果本文选取的工具变量 $IV_{sct}$ 满足相关性假定和外生性假定,即 $\pi_1 \neq 0$ 且 $Cov(IV_{sct}, \varepsilon_{isct}) = 0$ ,则 $\hat{\beta}_{1, TSLS}$ 能较好地近似衡量总体因果效应参数。

#### 四、研究发现

本部分使用上文介绍的分析方法,基于 CEPS 两期数据主要回答以下三个问题:其一,课外补习“是否致郁”,即考察课外补习参与对负向情绪水平的总体因果效应;其二,课外补习“使谁抑郁”,即考察课外补习致郁效应的阶层异质性;其三,解释致郁效应为何存在阶层分异。

##### (一)描述性统计结果

首先,本文的研究样本可划分为“本学年参加补习”和“本学年未参加补习”两个子样本(简称补习样本、未补习样本),然后分别进行描述性统计。表1报告了结果变量(负向情绪)及部分控制变量的均值和标准差,<sup>12</sup>发现如下:

12. 由于篇幅有限,本文在描述性统计部分将某些分类变量转化为连续变量(如父母教育年限),并省略了某些控制变量的统计结果。在后文的回归分析中,我们仍(转下页)

其一,补习样本的负向情绪得分显著低于未补习样本。平均意义上看,未参加补习的学生似乎表现出了更多的负向情绪。不过,由于暂未考虑到学生的家庭背景、上一学年的负向情绪以及学业表现,而这些因素既可能影响到学生当期的补习参与决策,又可能与当期的负向情绪密切相关,因此简单的负向情绪组间均值比较结果有待进一步的澄清。

其二,从控制变量的分布来看,补习学生更多来自较高的社会阶层,拥有较少的兄弟姐妹、更多来自城市、有更高的认知能力得分和语数外期中考试成绩,拥有更好的家庭经济条件和物质性文化资本。比较而言,未参加补习的学生更多来自较低阶层的家庭,有较多的兄弟姐妹,更多来自农村,认知能力测试和期中考试的表现也相对欠佳,家庭经济条件和物质性文化资本相对较差。

其三,迁移流动状况的组间比较值得注意。在参加补习的学生中,“本地入学比例”反而略低于未参加补习的学生样本,且这两组学生的“跨省流动比例”并无显著差异。这一结果可能与筛选样本的标准有关。本文的研究样本只包含那些在所有控制变量上都拥有观测值的受访者。我们发现,变量(特别是学生户口类型)存在缺失的受访者更多地来自流动人口较多的地区。<sup>13</sup>因此,那些处于迁移流动状态的受访者可能因缺失值未能进入研究样本。考虑到这一点,在后文的回归分析中我们控制了“学生的迁移流动状态”“学生家庭所在的地区类型”“学生所在区县”等变量,以免发生样本选择偏误。<sup>14</sup>

## (二)总效应的因果识别:课外补习是否“致郁”?

为了进一步考察原因变量(课外补习参与)对结果变量(情绪健康)的因果效应以及因果效应的组间(阶层)异质性,研究从描述性统计转向回归分析。

表2展示了在不同的分析方法和模型设定条件下,总效应的系数

---

(接上页)严格按照“变量介绍”中对控制变量的处理方式建立统计模型。完整的描述性统计结果可联系作者索取。

13. CEPS项目文档提到,全国补充样本是“从全国拥有大量流动人口的120个县(区)中抽取10个县(区)”而得到的。详见<http://ceps.ruc.edu.cn/xmwd/cysj.htm>。

14. 我们还对此进行了稳健性检验。例如,纳入不同的样本来源(全国核心样本、上海样本、全国补充样本)作为额外的控制变量,而不再控制学生户口类型(回归分析的样本量上升为8153人,占全样本的75.8%)。原文的分析结果依旧保持稳健,且由于样本量上升,估计精度也随之上升。

表 1：相关变量的描述性统计结果

	本学年参加课外补习 (N=2 528)		本学年未参加课外补习 (N=4 007)	
	均值	标准差	均值	标准差
<b>学生特征</b>				
负向情绪 IRT 得分(本学年)	-0.10	0.873	0.06	0.899
负向情绪 IRT 得分(上学年)	-0.19	0.833	-0.03	0.822
性别(男性=1)	0.46	0.498	0.51	0.500
年龄	14.43	0.634	14.72	0.802
兄弟姐妹数量	0.58	0.709	0.92	0.822
民族(汉族=1)	0.93	0.260	0.84	0.363
户口(农业=1)	0.49	0.500	0.71	0.456
本地入学=1	0.88	0.322	0.90	0.297
跨省入学=1	0.04	0.203	0.04	0.197
认知能力测试标准化得分(本学年)	0.47	0.727	0.14	0.813
认知能力测试标准化得分(上学年)	0.20	0.806	-0.14	0.844
语文期中考试标准化得分(上学年)	72.65	7.885	71.03	9.344
数学期中考试标准化得分(上学年)	72.62	8.363	70.78	9.919
外语期中考试标准化得分(上学年)	72.91	7.960	70.90	9.712
<b>家庭特征</b>				
家庭经济状况(家长自评)				
非常困难	0.03	0.172	0.06	0.239
比较困难	0.13	0.333	0.27	0.445
中等	0.77	0.423	0.63	0.484
比较富裕	0.07	0.252	0.04	0.190
很富裕	0.01	0.084	0.00	0.050
父母双方的最高受教育年限	11.71	3.248	16.09	3.459
父母期望子女接受教育的年数	17.71	3.101	16.61	3.595
物质性文化资本(标准化因子得分)	0.22	0.882	-0.53	1.069
居住安排(子女与双亲同住=1)	0.72	0.449	0.62	0.485

注：1. 描述性统计结果已经过抽样权重调整。2. “本学年”指 2014—2015 学年，“上一学年”指 2013—2014 学年，下同。

估计结果。表 2 的第(1)列是基准 OLS 模型得到的系数估计值。在控制了个体层面协变量、学校固定效应、学期固定效应的基础上，参加课外补习对负向情绪得分的效应系数并不显著区别于 0 值。接下来，笔者在基准模型的基础上纳入滞后变量，以缓解反向因果问题和滞后效应带来的遗漏变量偏误。表 2 的第(2)列系数显示，将滞后变量进行控制之

表 2: 课外补习对负向情绪的总效应 (基于 OLS 和 TSLS 的估计结果) (N=6 535)

	(1)全样本 OLS	(2)全样本 OLS	(3)全样本 TSLS	(4)全样本 TSLS
参加补习班=1	0.039	0.052**	0.355**	0.357**
(语/数/外)	(0.026)	(0.023)	(0.139)	(0.138)
滞后变量组	NO	YES	YES	YES
个体层面协变量	YES	YES	YES	YES
学校固定效应	YES	YES	NO	NO
区县固定效应	NO	NO	YES	YES
学期固定效应	YES	YES	YES	YES
一阶段 F-stats			181.29	102.33
Hansen's J-stats				0.01

注:1. 括号内是以班级聚类的稳健标准误;\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ 。2. 滞后变量包括前期学业能力的度量以及上一学年的负向情绪 IRT 得分; 个体层面协变量包括学生特征、家庭特征、学生所在的班级特征。3. 第(1)列和第(2)列是基于普通最小二乘法(OLS)得到的系数估计结果。4. 第(3)列是基于当期工具变量和两阶段最小二乘法(TSLS)得到的影响系数估计值。5. 第(4)列是基于当期工具变量及滞后期工具变量,使用两阶段最小二乘法(TSLS)得到的影响系数估计结果。

后,课外补习参与对负向情绪得分的影响系数估计值增大,系数估计值在  $p < 0.05$  的水平上显著为正。这在一定程度上印证了“负向情绪对课外补习参与”的负反馈效应可能导致因果效应系数被低估。

在多元 OLS 回归的基础上,本研究使用工具变量法来进一步处理内生性偏误。首先,使用当期数据(2014—2015 学年 CEPS 截面数据)构造了学生个体对应的“位似群体”课外补习参与率作为工具变量,使用两阶段最小二乘法(TSLS)来估计课外补习参与对负向情绪得分的因果效应,结果如表 2 的第(3)列所示。第(3)列末尾报告了第一阶段回归的 F 值,数值上远大于临界值 10,说明此处使用的工具变量满足相关性假定,暂不存在弱工具变量问题。<sup>15</sup>

基于工具变量的 TSLS 估计结果显示,课外补习参与对负向情绪得分的因果效应系数显著为正。数值上看,给定控制变量,参加课外补习(相较于未参加)使学生的负向情绪得分增加约 0.38 个标准差(这是因为致郁效应的点估计值约为 0.35,而负向情绪得分的样本标准差约为 0.93),即课外补习对情绪健康表现出显著的致郁效应。

15. 简洁起见,若不存在弱工具变量问题,下文将不再专门解读一阶段 F 值。

考虑到基于当期数据构造的工具变量(即2014—2015学年“位似群体”的课外补习参与率)可能与结果变量(即2014—2015学年学生的负向情绪得分)存在互为因果的问题,本文用2013—2014学年的CEPS基期数据,构造了滞后期的“位似群体补课参与率”作为第二个工具变量;再将两个工具变量同时放入一阶段回归方程,进行两阶段最小二乘估计。结果如表2第(4)列所示,与第(3)列系数非常接近。此外,本文在表2的第(4)列下方报告了过度识别检验的J统计量。J统计量数值约为0.01,p值超过0.9,过度识别检验不显著。这在一定程度上印证了本研究工具变量的外生性(陈强,2014)。

需要注意的是,与加入了滞后变量的OLS回归相比,工具变量法对致郁效应的估计值增大了7倍左右。存在两种可能的解释:其一,由于遗漏变量和反向因果等问题,OLS回归可能低估了真实的致郁效应;其二,当因果效应具有个体异质性时,工具变量法所估计的其实是“局部平均因果效应”(local average treatment effect, LATE)(详细介绍参见 Angrist, *et al.*, 1996; Angrist and Pischke, 2009)。也就是说,本研究用工具变量估计出的致郁效应可能反映的是“(总体当中)课外补习的参与决策真正受到了‘位似群体’影响的那些学生”的平均因果效应(Stock and Watson, 2015: 506—508)。这些学生的补习参与决策更易受到外界的影响,由此推测他们对补习效果的认知评价以及补习参与中的情绪体验也相对更敏感,更易受社会环境因素的影响。因此,补习参与引发的负向情绪可能在这部分学生当中更强烈和鲜明。这也说明本研究的结果可能不宜直接推广到所有学生,而是更适用于那些补习参与易受到他人或周遭环境影响的学生群体。不过,工具变量和(加入滞后变量的)OLS回归虽在量上有较大区别,但是从致郁效应的方向和显著性上看并没有质的差别。

为了检验课外补习参与的“致郁效应”是否受到课外补习强度的调节(孙伦轩、唐晶晶, 2019; Zheng, *et al.*, 2020),本文将学生的“周末日均补习时长”转化为4个虚拟变量(不补习、0~2小时、2~4小时、4小时以上)。放入后三个虚拟变量(基准组是未参加补习的学生)作为回归方程的核心自变量,生成与之对应的工具变量,进行两阶段最小二乘回归。然后,本研究使用瓦尔德检验(Wald test),原假设为“上述三个虚拟变量对应的系数之间没有显著差异”,卡方统计量显示,这三个系数在  $p <$

0.1 的水平上并不存在显著差异。我们使用类似的方法将“补习科目数量”转化为 5 个虚拟变量(未参加补习、只补习 1 门、补习 2 门、补习 3 门、补习 4 门),并放入后四个虚拟变量(基准组是未参加补习的学生)作为回归方程的核心自变量,其余设定不变。统计结果显示,这四个虚拟变量对应的系数也不存在显著差异。上述结果可在一定程度上说明,补习参与的致郁效应并不明显依补习时长和补习科目的数量发生变化。<sup>16</sup>

总的来看,课外补习参与对中学生的情绪健康具有显著的致郁效应,即参加课外补习会显著增加学生的负向情绪,其效应大致为 0.38 个负向情绪得分的标准差。此外,补习参与的致郁效应并不明显依补习时长和科目数量而发生变化。

### (三) 阶层异质性分析:课外补习使谁抑郁?

本节考察学生的家庭阶层背景对课外补习致郁效应的调节作用。核心问题是,对于不同家庭阶层背景子女而言,课外补习参与的致郁效应是否存在组间异质性。

本研究按照家庭背景变量的取值将全样本分组,以分组回归的方式考察各组的组内因果效应的显著性差异,而未设置交互项来直接比较各组系数的大小。因果推断研究者范德韦勒(VanderWeele, 2015)将基于分组回归系数的显著性比较方法称为“定性交互分析”(qualitative interaction)。定性交互分析的目的并非考察不同群体在致郁效应具体数值上的组间差异是否显著,而是在定性层面上考察致郁效应在哪些子群体中显著存在。

本文使用父母双方的最高受教育程度、父亲现从事的职业类型来划分不同的子群体,然后分组进行回归分析。<sup>17</sup>表 3 展示了按照父母最高受教育程度划分的分组回归结果,估计方法和变量选择与总效应分析保持一致,详见表 2 中第(4)列的模型设定。将表 3 后三列系数进行横向比较可以发现,只有在父母文化程度为初中及以下的子女群体中,

16. 简洁起见,文中并未展示对应的回归结果。有需要的读者可以联系作者索取。

17. 需要说明的是,在 CEPS 数据中没有对家庭收入的直接测量,只有受访者报告的自评家庭经济条件(五分类或三分类)。该变量分布存在严重的“趋中现象”,回答家庭经济条件中等的人数比例超过 60%,而回答家庭条件较富裕或很富裕的人数不足 5%。如果按家庭经济条件进行分组回归,将会使某些组别的样本量较少(不到 500 人),工具变量的相关性受到较大威胁,可能引发弱工具变量问题,因此在分组回归中笔者没有使用自评家庭经济条件来划分子群体。

表 3: 课外补习因果效应的异质性(按父母受教育程度分组)

	(1)TSLs 全样本	(2)TSLs 初中及以上	(3)TSLs 中专或高中	(4)TSLs 大专及以上
参加补习班=1 (语/数/外)	0.357*** (0.138)	0.561*** (0.160)	0.231 (0.252)	0.313 (0.332)
滞后变量组	YES	YES	YES	YES
个体层面协变量	YES	YES	YES	YES
区县固定效应	YES	YES	YES	YES
学期固定效应	YES	YES	YES	YES
一阶段 F 值	102.33	48.35	29.21	17.29
样本量	6 535	3 130	1 807	1 598

注:1.括号内是以班级聚类的稳健标准误;\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ 。2.滞后变量包括前期学业能力的度量以及上一学年的负向情绪 IRT 得分;个体层面协变量包括学生特征、家庭特征、学生所属的班级特征。3.第(1)列是全样本回归结果,第(2)(3)(4)列是以父母最高受教育程度进行分组的子样本回归结果。

补习参与对负向情绪的因果效应系数显著为正。相比之下,父母文化程度为中专或高中、大专及以上的两组子女,参加课外补习对负向情绪的回归系数在  $p < 0.1$  的水平上并不显著区别于 0 值。父母文化程度相对较高的子女对课外补习引发的负向情绪表现出心理免疫现象。

类似地,表 4 中展示了按照父亲当前从事的职业类型进行分组的回归结果,估计方法和变量选择与表 3 保持一致。将表 4 中的后五列系数进行比较,只有在较低职业阶层(工人、农业劳动者和城市初级劳动者)中,其子女参与课外补习的致郁效应显著为正。相比之下,在较高职业阶层中(体制内或体制外的管理人员、专业技术人员等),致郁效应的估计值在  $p < 0.1$  的水平上不显著区别于 0 值。也就是说,父亲职业阶层较高的子女对课外补习诱发的负向情绪表现出心理免疫现象。

#### (四)“致郁效应”为何存在阶层分异?

依据理论框架,笔者推测,家长教养方式的阶层分异模式以及与之关联的子女非认知能力(情绪调节能力)的差异可能有助于解释致郁效应为何存在阶层异质性。但是,受数据和变量可及性的限制,本文暂且无法完整验证理论框架的解释力。本研究主要考察了“情绪调节境况”这个具体维度上的阶层分异,进而为前文的理论框架提供支持性证据,仍有待更完善的质性解释和量化检验。

具体来说,我们将研究样本进一步限定为“最近一年参加了课外补

表 4: 课外补习因果效应的异质性(按父亲当前职业类型分组)

	(0)TSLS 全样本	(1)TSLS 较高阶层	(2)TSLS 较高阶层	(3)TSLS 较低阶层	(4)TSLS 较低阶层	(5)TSLS 较低阶层
参加补习班=1 (语/数/外)	0.357*** (0.138)	0.081 (0.247)	0.346 (0.225)	0.500*** (0.180)	0.536** (0.214)	0.818** (0.347)
滞后变量组	YES	YES	YES	YES	YES	YES
个体层面协变量	YES	YES	YES	YES	YES	YES
区县固定效应	YES	YES	YES	YES	YES	YES
学期固定效应	YES	YES	YES	YES	YES	YES
一阶段 F 值	100.43	27.23	33.16	44.02	42.38	18.92
样本量	6 535	1 052	1 282	3 166	1 841	1 049

注:1. 括号内是以班级聚类的稳健标准误;\*\*\* p<0.01,\*\* p<0.05,\* p<0.1。2.滞后变量包括前期学业能力的度量以及上一学年的负向情绪 IRT 得分;个体层面协变量包括学生特征、家庭特征、学生所属的班级特征。3.模型 1 是父亲职业为管理人员和专业技术人员的子样本回归;模型 2 是父亲职业为管理人员、专业技术人员、一般办事人员的子样本回归;模型 3 是父亲职业为技术工人、普通工人、农业劳动者以及城市初级劳动者的子样本回归;模型 4 是父亲职业为普通工人、农业劳动者以及城市初级劳动者的子样本回归;模型 5 是父亲职业为普通工人、农业劳动者的子样本回归。

习的学生”,发现不同社会阶层的子女在负向情绪调节方面存在与致郁效应类似的组间差异。具体而言,CEPS 问卷中有多项问题涉及学生产生负向情绪时,学生及家长的应对状态。主要有以下三方面问题:其一,由受访学生家长作答的“家长主动与子女讨论心事或烦恼的频率”,以及由受访学生作答的“父亲、母亲与自己讨论心事或烦恼的频率”,这三个变量都是三分类的定序变量(从不、偶尔、经常)。上述三个变量在一定程度上反映了家长主动体察子女负向情绪的情况。其二是受访学生对于“出现负向情绪时,自己通常可以很快调整过来”的同意程度,该问题采取的是四分定序测量(从“完全不同意”到“完全同意”),在一定意义上衡量了学生自身的情绪调节能力。其三是受访学生对“出现负向情绪时,可以很快在他人帮助下调整过来”的同意程度,该问题同样采取了四分定序测量方式。这个变量在一定意义上体现了学生自身的情绪调节能力,以及通过他人帮助来调节情绪的难易程度。

本研究将上述五个变量的取值进行加总平均,然后将其转化为均值为 0、方差为 1 的“情绪调节境况的标准化得分”,它综合了家长对子

女情绪的主动体察状况以及子女依靠自身或通过他人帮助以实现情绪调节的状况。该变量得分越高,说明情绪调节境况越佳。

本研究用组间均值比较的方式来考察情绪调节境况的阶层差异。表5按照父母受教育程度分组,报告了最近一年参加了课外补习的学生情绪调节境况的分异模式。第(1)列系数是“家长主动与子女讨论心事、烦恼的频率”的组间比较。第(2)列和第(3)列系数分别展示了各组子女对“自己可以很快调整负向情绪”以及“能在他人帮助下很快调整负向情绪”的同意程度。从趋势上看,父母文化程度越高,子女对这两项陈述的同意程度越高。第(4)列系数是对前三列系数的综合,展示了“情绪调节境况的标准化得分”的组间均值差异。可以看出,子女情绪调节的境况随父母文化程度的上升而不断向好发展。如果父母学历为研究生及以上,子女的情绪调节境况得分比父母文化程度为小学及以下的子女高出近0.6个标准差,比父母文化程度为初中的子女高出约0.4个标准差。表5最后一行报告了组间均值检验结果。以班级聚类且异方差稳健的F统计量显示,在“家长主动与子女谈心”“在他人帮助下很快能调整负向情绪”及“情绪调节境况的标准化得分”这三个变量上,都存在显著的组间差异( $p < 0.01$ )。而按父母文化程度划分的各组子女虽

表5:情绪调节境况的阶层差异(按父母最高受教育程度划分的组间均值比较)

	(1)家长主动与子女讨论心事或烦恼的频率	(2)自己可以很快调整负向情绪(同意程度)	(3)在他人帮助下可以很快地调整负向情绪(同意程度)	(4)情绪调节境况(标准化得分)
小学及以下	2.082	2.994	2.677	-0.061
初中	2.182	3.115	2.798	0.109
中专	2.229	3.010	2.727	0.129
普通高中	2.261	3.120	2.765	0.188
职业高中	2.322	3.113	2.782	0.125
大专	2.452	3.073	2.860	0.219
本科	2.488	3.232	2.969	0.398
研究生及以上	2.530	3.402	3.175	0.500
组间均值F检验	18.74***	1.63	4.38***	6.16***

注:1. F统计量由班级聚类的稳健方差-协方差矩阵得到;\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ 。2. 各组均值的描述性统计结果已经过抽样权重调整。3. 分析样本限定在2014—2015学年参加了课外补习的学生群体( $N=3\ 421$ )。4. 为更接近总体分布,此处所用样本是原始数据中在最近一年参加补习的学生,这包括了在其他控制变量上存在缺失的受访者。若排除存在变量值缺失的受访者,主要结论仍然保持稳健(下表6的情形与此相同,不再赘述)。

在“自己能很快调整负向情绪”的同意程度上表现出阶层差异,但并不显著( $p=0.13$ )。

接下来,我们以学生父亲现从事的职业类别作为分组标准,发现在最近一年参加了课外补习的学生当中,各组子女的情绪调节境况也存在类似的分异模式。组间均值比较的结果如表 6 所示。总的来说,相较于父亲位于较低职业阶层(例如普通工人、农业劳动者、城市初级劳动者)的子女,父亲处在较高职业阶层(如体制内外的管理人员、专业技术人员)的子女拥有更好的情绪调节境况。例如,父亲为政府机关领导干部的学生“情绪调节境况标准化得分”比父亲为普通工人或农业劳动者的学生高出了约 0.45 个标准差。此外,表 6 的最后一行报告了组间均值检验结果。除了“自己可以很快调整负向情绪”这一变量仅在  $p<0.1$  的水平上边缘显著,其余三个变量均在  $p<0.05$  的水平上存在显著的组间差异。

为进一步验证情绪调节境况的阶层差异是否稳健,本研究进行了

表 6: 情绪调节境况的阶层差异(按父亲职业类别划分的组间均值比较)

	(1)家长主动与子女讨论心事或烦恼的频率	(2)自己可以很快调整负向情绪(同意程度)	(3)在他人帮助下可以很快地调整负向情绪(同意程度)	(4)情绪调节境况(标准化得分)
政府机关领导/干部	2.501	3.242	2.945	0.520
企事业单位领导/干部	2.414	3.151	2.933	0.321
科学家、大学教师等	2.453	3.274	2.914	0.528
医生、律师、中小学教师	2.385	3.130	2.910	0.246
一般职工、办事人员	2.298	3.091	2.888	0.212
个体工商户	2.290	3.166	2.856	0.198
技术工人	2.320	3.200	2.748	0.157
普通工人	2.161	3.077	2.778	0.082
农业劳动者	2.178	2.986	2.768	0.059
城市初级劳动者	2.039	2.743	2.628	-0.267
退休、无业、失业、下岗	2.025	2.955	2.681	-0.294
组间均值 F 检验结果	4.62***	1.65*	2.05**	3.54***

注:1. F 统计量由班级聚类的稳健方差-协方差矩阵得到;\*\*\*  $p<0.01$ ,\*\*  $p<0.05$ ,\*  $p<0.1$ 。2. 各组均值的描述性统计结果已经过抽样权重调整。3. 样本限定在 2014—2015 学年参加了课外补习的学生( $N=3\ 421$ ),部分职业类别未列出。

以下两个 OLS 回归。在回归方程中,“情绪调节境况的标准化得分”作为因变量,学生的人口学特征、学期和学校固定效应作为控制变量,父母最高受教育程度和父亲职业类别(处理为多个二值变量)分别作为两个模型的核心自变量。回归结果显示,在控制了学生特征、学期和学校固定效应后,情绪调节境况在不同阶层的子女间仍存在显著差异(联合检验  $p < 0.05$ )。

上述结果表明,不仅由补课引发的致郁效应存在组间异质性,不同家庭背景的学生在情绪调节境况方面也存在类似的阶层分异模式。接下来要回答的问题是,情绪调节境况的阶层差异是否能够解释(或影响)致郁效应的阶层差异。这是一个“高维度”的因果推断问题,原因变量是情绪调节境况的阶层差异(而非个体层面的情绪调节境况),结果变量是由补习引发的致郁效应的阶层差异(而非个体层面的负向情绪水平)。针对上述原因变量和结果变量,均只有一个相应的观测值,难以用结果变量对原因变量进行回归来估计出因果效应,因此,我们选择“反事实模拟”的分析框架来回答这一因果问题。

反事实模拟的思路是:对于参加补习的较低阶层的子女而言,假定他们的情绪调节境况变得和较高阶层家庭的子女同样好,其他条件不变,那么(原有的)致郁效应的阶层差异是否也会相应地缩小(或者说,致郁效应的组间异质性是否会相应减弱)?在这个反事实情境中,参加补习的较低阶层家庭的子女与那些较高阶层家庭的子女在情绪调节境况上存在的“事实性”阶层差异不复存在(平均而言)。这时,如果致郁效应在较高与较低阶层子女间(事实上)存在的差异也因情绪调节境况的阶层差异被“反事实地抹平”而相应地缩小甚至消失,那么就可以印证情绪调节的阶层差异在一定程度上能解释致郁效应的阶层差异。

在实证策略上,本研究借鉴了代表性研究使用的反事实分析框架(Liu 2019; VanderWeele 2015),并基于研究情景的特殊性(高维因果推断)进行了相应调整,具体步骤如下:<sup>18</sup>

第一步,使用全样本(6 535 人)数据和主效应分析中所用的工具变量回归模型设定,在此基础上引入“情绪调节境况”的一次项、“情绪调节境况”与“补习参与”的交互项。<sup>19</sup>模型设定如下所示, $Y_i$ 为负向情

18. 简洁起见,这部分的分析是对两组家庭阶层背景差异较大的子群体进行比较。

19. 需要注意的是,这里设置交互项的目的并不是考察情绪调节或“致郁效”(转下页)

绪得分,  $D_i$  为补习参与变量,  $M_i$  为情绪调节境况,  $X_i$  为控制变量。我们用两阶段最小二乘法估计出“情绪调节境况”与“补习参与”的交互项对应的系数  $\beta_3$ , 其估计值为  $-0.384$  ( $p < 0.01$ ), “情绪调节境况”所对应的系数  $\beta_2$  的估计值为  $-0.202$  ( $p < 0.01$ )。

$$E(Y_i | D_i, M_i, X_i) = \beta_1 D_i + \beta_2 M_i + \beta_3 (D_i \cdot M_i) + X_i' \gamma$$

第二步, 分别在较低阶层子女样本、较高阶层子女样本中进行工具变量回归, 模型设定和第一步相同。将这两组子女的情绪调节境况的取值 ( $M_i$ ) 设定在各自的样本均值处 (表示为  $\bar{m}_{Low}$  和  $\bar{m}_{High}$ ), 然后分别估计出这两组子女的“致郁效应”(基于事实情形)。较低阶层的子女“事实上的致郁效应”表示为  $\delta_{Low, act}$ , 较高阶层子女“事实上的致郁效应”表示为  $\delta_{High, act}$ 。进一步计算出较高、较低阶层这两组子女在致郁效应上的事实性差异, 记为  $\Delta_{actual}$ 。计算方式如下所示:

$$\delta_{Low, act} = E(Y_i | D_i = 1, SES_i = low, M_i = \bar{m}_{Low}) - E(Y_i | D_i = 0, SES_i = low, M_i = \bar{m}_{Low})$$

$$\delta_{High, act} = E(Y_i | D_i = 1, SES_i = high, M_i = \bar{m}_{High}) - E(Y_i | D_i = 0, SES_i = high, M_i = \bar{m}_{High})$$

$$\Delta_{actual} = \delta_{High, act} - \delta_{Low, act}$$

接下来, 将参与了课外补习的较低阶层子女的情绪调节境况“设定”在  $\bar{m}_{High}$ , 即“较高阶层子女的情绪调节境况的均值处”,<sup>20</sup> 其他条件不变, 重新估计出较低阶层子女的致郁效应, 即“反事实的致郁效应”, 记为  $\delta_{Low, ct}$ :

$$\delta_{Low, ct} = E(Y_i | D_i = 1, SES_i = low, M_i = \bar{m}_{High}) - E(Y_i | D_i = 0, SES_i = low, M_i = \bar{m}_{Low})$$

用较高阶层子女的“事实上的致郁效应”减去较低阶层子女的“反事实致郁效应”, 即可得到两组子女在致郁效应上的反事实差异, 记为

$$\Delta_{counter}$$

(接上页) 应”的阶层差异, 而是为了在下一步使用“回归拟合法”模拟出较低阶层家庭的子女在反事实情形下的致郁效应。这是因为, 如果“情绪调节-补习参与”交互项对应的系数为 0, 则意味着无论较低阶层子女拥有了何种水平的情绪调节境况, 其致郁效应都不会发生任何改变; 因此, “情绪调节-补习参与”交互项系数显著区别于 0 值, 是我们印证“情绪调节的阶层差异能够影响(或解释)致郁效应的阶层差异”的必要条件。

20. 此处的较高阶层子女并未区分是否参与课外补习。出于稳健考虑, 也可以将之设定在“参与了补习的、较高阶层子女的情绪调节境况的样本均值处”。我们发现这两种设定对反事实模拟的结果并不会带来较大改变。为公式推导方便及行文简洁, 本文只报告了前一种设定的反事实模拟结果。

“忧郁的孩子们”：课外补习会带来负向情绪吗？

$$\Delta_{counter} = \delta_{High, act} - \delta_{Low, ct}$$

最后，计算致郁效应的“事实差异”与“反事实差异”的差异（简称“差中差”），它反映了假定（参与补习的）较低阶层子女的情绪调节境况“调整到”与较高阶层子女一样好的水平后，致郁效应的阶层差异缩小的幅度。线性模型设定下，致郁效应的“差中差”可以简单地用下式计算得出：

$$\Delta_{counter} - \Delta_{actual} = \delta_{Low, act} - \delta_{Low, ct} = (\beta_2 + \beta_3) (\bar{m}_{Low} - \bar{m}_{High})$$

假定参加了补习的较低阶层子女的情绪调节境况变得与较高阶层子女同样好，保持其他条件不变，将会使较高阶层家庭与较低阶层家庭的子女致郁效应的“事实性差异”缩小 $(\beta_2 + \beta_3) (\bar{m}_{Low} - \bar{m}_{High})$ 个单位。<sup>21</sup>

接下来，将致郁效应的“差中差”除以致郁效应原有的事实差异，就可以得出致郁效应的阶层差异缩小的百分比大小（差异缩小的幅度）。

$$\left| \frac{\Delta_{counter} - \Delta_{actual}}{\Delta_{actual}} \right| = \left| \frac{(\beta_2 + \beta_3) (\bar{m}_{Low} - \bar{m}_{High})}{\delta_{High, act} - \delta_{Low, ct}} \right|$$

本研究考察了父母受教育程度较高和较低两组子女在致郁效应上的阶层差异及其反事实结果，以及父亲职业阶层较高和较低两组子女在致郁效应上的阶层差异及其反事实结果，如表7所示。表7第一行系数表示，假若“父母文化程度为初中及以下且参加了补习的子女”的情绪调节境况变得和“父母文化程度为大专及以上的子女”同样好，那么较高与较低阶层子女的致郁效应差异将会（相较事实情形）缩小约42%。第二行将组间比较限定在“父母文化程度为初中”和“父母文化程度为本科及以上”的两组子女，情绪调节境况的反事实模拟显示，致郁

21. 在这里，我们假定“情绪调节境况”与“补习参与”的交互效应并不依赖学生的家庭阶层背景，因此直接使用全样本估计出交互效应系数。更理想的情形是只在较低阶层子女（或较高阶层子女）的子样本中估计“情绪调节境况”与“补习参与”的交互效应，但实践中子样本量不够大，工具变量回归得到的交互效应系数估计精度出现较大损失，不足以准确模拟出较低阶层子女的反事实致郁效应。另外，我们检验了交互项系数在不同家庭背景的女儿间是否存在阶层差异，发现在 $p < 0.1$ 的水平上并不存在显著差异。基于此，我们使用全样本而不是分样本来估计交互效应，并用全样本（含交互项）的回归结果计算致郁效应的“差中差”。同时，为了与上节的组间异质性分析结果保持一致，我们仍然使用分样本回归的方式来计算致郁效应的“事实差异”。我们也换用了其他的计算方式（例如，只使用全样本回归或者只使用分样本回归）来计算上述三项指标，得到了较类似的反事实模拟结果。简洁起见，其他算法的结果并未展示。

表 7: 致郁效应阶层差异的反事实模拟结果

	(1)致郁效应的 事实性差异	(2)致郁效应的 反事实差异	(3)“差中差” = (2)-(1)	(4)差异缩小幅度 = (3)/(1)
教育分层	-0.270	-0.157	0.113	41.9%
教育分层	-0.151	-0.025	0.126	83.5%
职业分层	-0.386	-0.287	0.099	25.6%
职业分层	-0.192	-0.081	0.111	57.7%

效应的阶层差异将会(相较事实情形)缩小约 84%。<sup>22</sup> 第三行系数表示, 假若“父亲职业为普通工人、农业劳动者或城市初级劳动者”的情绪调节境况变得与“父亲职业为管理人员、专业技术人员”的子女一样好, 那么较高与较低阶层子女的致郁效应差异将会(相较事实情形)缩小约 26%。最后, 将组间比较限定在“父亲职业为普通工人”和“父亲职业为管理人员”的两组子女, 我们发现情绪调节境况的反事实模拟使致郁效应的阶层差异(相较事实情形)缩小约 58%。

上述反事实模拟结果表明, “情绪调节境况”的阶层差异可能在一定程度上影响和解释了致郁效应的阶层差异(组间异质性)。不过, 这是一个高维因果问题, 我们难以保证情绪调节境况及其阶层差异满足因果推断框架要求的“可忽略性假定”(ignorability assumption)。因此, 反事实模拟的结果只能为本文的理论解释提供支持性证据, 无法直接和严谨地证明情绪调节的阶层差异对致郁效应的阶层差异具有因果影响, 该问题仍有待进一步研究论证。

## 五、结论与讨论

### (一) 结论

课外补习在当下已是全球性的普遍文化现象, 伴随着日益激烈的教育竞争, 青少年的负向情绪问题持续得到舆论关注, 也成为社会科学研究的重要议题。在此现实背景下, 本文关注中国大陆中学生课外补习参与和情绪健康状况的因果关系问题。本文基于两期 CEPS 数据, 用工具变量法和滞后变量法化解因果推断中的内生性偏误, 首先分析了课外补习对负向情绪得分的总体因果效应; 然后使用分组回归的方式考察课外补习的致郁效应可能存在的阶层异质性; 最后对致郁效应的阶

22. 这一情形的下降比例较大, 可能是由于这两组子女致郁效应的事实性差异原本就比较小。

层异质性进行初步解释和量化检验。主要结论如下:

其一,本文回答了课外补习“是否致郁”的问题。加入滞后变量的OLS回归和工具变量回归结果均显示,参加课外补习会显著增加学生的负向情绪得分。其中,基于工具变量法得到的致郁效应估计值约为0.38个标准差。

其二,在回答了“是否致郁”的基础上,本文探索了课外补习“使谁抑郁”的问题。本研究按照父母最高受教育程度、父亲当前从事的职业类型划分出不同的子群体,使用分组回归进行异质性分析,发现课外补习的致郁效应表现出“劣势—特异”和“优势—免疫”的阶层分异模式。

其三,本文尝试对课外补习致郁效应的阶层异质性做出解释。本研究基于家庭教养方式的理想型和经验发现提出了一个初步的理论框架:在课外补习参与过程中,不同社会阶层的家长对子女负向情绪的应对方式可能有所不同,子女自身的情绪调节能力可能也存在组间差异。我们具体考察了参加补习的学生“情绪调节境况”的阶层分异模式,并采用反事实模拟的方式展示了“情绪调节境况的阶层差异”与“致郁效应的阶层差异”的内在关联,为本文的理论框架提供了支持性的证据,但仍有待进一步的质性解释和量化检验加以完善。

## (二) 讨论

本文的发现可能有助于弥合既有研究的经验分歧。其一,就课外补习对负向情绪的总体因果效应而言,补习参与的影响存在“安慰”与“致郁”的分歧。本文通过比较OLS多元回归与工具变量法的TSLs回归结果,认为安慰剂效应值得商榷,它可能是未充分考虑到反向因果问题,导致课外补习的致郁效应受到系统性低估(孙伦轩、唐晶晶,2019)<sup>23</sup>。其二,本文使用工具变量处理内生性问题后,课外补习的致郁效应并未明显受到补习强度的调节。相较于未参与补习的学生,不同时长、不同科目数量的补习参与都会显著增加学生的负向情绪。而有学者认为补习的致郁效应只在补习强度较大的情形下才会凸显(Zheng, *et al.*, 2020)。本研究发现的情绪健康后果所对应的补习参与阈值相比既有研究更低,该发现或更具警示性的政策意涵。

本研究对课外补习致郁效应的阶层异质性分析,有助于我们理解课外补习可能具有的教育与健康不平等后果。本文发现,课外补习引发

23. 另一种可能的解释是,本文所用的工具变量法捕捉到的是“局部平均因果效应”。

的负向情绪很可能是较低阶层子女的特异性后果，而较高阶层的子女对此类负向情绪具有较好的心理免疫能力。这可能意味着，即使课外补习参与机会的阶层差异在逐步弥合，课外补习参与结果的不平等仍可能继续。负向情绪的长期累积又将对学生的学业表现、身心发展带来多方面、持续性的不良后果。由此，课外补习可能借助健康不平等的生产，隐秘地维持和延续了现有的教育不平等格局。这为未来的教育分层研究和进一步的“双减”政策落地提供了有益启发。

### （三）局限与进路

本研究仍存在诸多局限。其一，本文并不能完全印证家庭教养方式与课外补习的致郁效应及其阶层分异之间存在因果关联。其二，受数据可及性的限制，本文难以捕捉到课外补习的类型与质量的差别，如培优（enrichment）与补差（remediation）的区分，补习班的班级规模、教师授课水平、补习讲授的内容等。课外补习类型的阶层分异也可能是致郁效应呈现阶层异质性的解释之一。其三，本文在阶层异质性分析中的分组标准还有待改进，主要参考了经典理论的做法，用父母文化程度、职业类型来划分样本以展示课外补习与负向情绪的阶层分异图景。

此外，本文使用分组回归的方式捕捉阶层异质性的做法也值得商榷。即使面对同一份情绪量表，不同家庭背景子女对量表的主观理解可能存在差异；另外，即使参加质量和数量相同的补习班，不同家庭背景子女面临的情绪健康问题的来源可能不同。一个可能的猜想是，较高阶层的子女在补习参与过程中或更易产生向上的社会比较或平行比较，并与其父母的“中产焦虑”相互叠加，对自己在补习中取得的进步并不满意，在“内卷”浪潮裹挟下倍感竞争压力；而弱势阶层的子女或受制于家境约束，忧心补习无用、浪费稀缺的家庭经济资源。这样看来，不同家庭背景的子女在情绪健康等主观指标上的可比性仍有待改进。

本研究有两个改进方向：其一是获取信息量更为丰富的全国代表性数据，进而对家庭教养方式、课外补习与负向情绪的因果机制进行更扎实的实证分析；其二是走进田野，观察不同阶层的子女在补习参与过程中的真实境遇与心路历程，探索不同类型的家长帮助其子女化解负向情绪的具体策略及阶层分异的图景，以期充实和发展量化发现，迈向更鲜活饱满又接地气的社会研究。

## 参考文献(References)

- 陈强. 2014. 高级计量经济学及 Stata 应用(第二版)[M]. 北京:高等教育出版社.
- 陈云松、范晓光. 2010. 社会学定量分析中的内生性问题——测估社会互动的因果效应研究综述[J]. 社会 30(4):91-117.
- 杜育红、郭艳斌. 2019. 初中生的同伴效应:基于发育成熟度的自然实验[J]. 教育与经济(3):40-48.
- 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[J]. 教育经济评论(1):77-91.
- 黄超. 2018. 家长教养方式的阶层差异及其对子女非认知能力的影响[J]. 社会 38(6):216-240.
- 李长洪、林文炼. 2019. “近墨者黑”:负向情绪会传染吗?——基于“班级”社交网络视角[J]. 经济学(季刊)(2):597-616.
- 苗瑞凯. 2020. 中国中学生负性生活事件与抑郁症状关系的元分析[J]. 上海教育科研(3):47-50.
- 牛娟、阮明成、周亦佳. 2019. 中学生心理素质、主观社会地位与抑郁的关系及其变化发展模式[A]. 第二十二届全国心理学学术会议摘要集.中国心理学会.
- 孙伦轩、唐晶晶. 2019. 课外补习的有效性——基于中国教育追踪调查的估计[J]. 北京大学教育评论(1):123-141.
- 谢宇. 2006. 社会学方法与定量研究(第一版)[M]. 北京:社会科学文献出版社.
- 薛海平、丁小浩. 2009. 中国城镇学生教育补习研究[J]. 教育研究(1):39-46.
- 张文武、周东升、胡珍玉. 2014. 宁波市中学生抑郁情绪调查及相关因素分析[J]. 中国学校卫生(10):1503-1505.
- 中国教育报. 2017. 中国教育学会发布调查报告显示 2016 年我国中小学课外辅导“吸金”超八千亿[J]. 教育发展研究(4):63.
- Angrist, Joshua D. and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. New Jersey: Princeton University Press.
- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. “Identification of Causal Effects Using Instrumental Variables.” *Journal of the American Statistical Association* (91): 444-455.
- Azarnert, Leonid V. 2020. “Health Capital Provision and Human Capital Accumulation.” *Oxford Economic Papers* 72(3): 633-650.
- Becker, Gary S. 2007. “Health as Human Capital: Synthesis and Extensions.” *Oxford Economic Papers* 59(3): 379-410.
- Blau, Peter M. and Otis D. Duncan. 1967. *The American Occupational Structure*. New York: The Free Press.
- Brand, Jennie E., Ravaris Moore, Xi Song, and Yu Xie. 2019. “Parental Divorce is Not Uniformly Disruptive to Children's Educational Attainment.” *Proceedings of the National Academy of Sciences* 116(15): 7266-7271.
- Bray, Mark. 1999. *The Shadow Education System: Private Tutoring and Its Implications for Planners*. Paris: UNESCO.
- Bray, Mark. 2013. “Benefits and Tensions of Shadow Education: Comparative Perspectives on The Roles and Impact of Private Supplementary Tutoring in the Lives of Hong Kong Students.” *Journal of International and Comparative Education* 2(1): 18-30.
- Breen, Richard, Seungsoo Choi, and Anders Holm. 2015. “Heterogeneous Causal Effects and Sample Selection Bias.” *Sociological Science* 2: 351-369.

- Bridge, Jeffrey A., Tina R. Goldstein, and David A. Brent. 2006. "Adolescent Suicide and Suicidal Behavior." *Journal of Child Psychology and Psychiatry* 47(3-4): 372-394.
- Chen, I-Chen and Ping-Yin Kuan. 2021. "The Heterogeneous Effects of Participation in Shadow Education on Mental Health of High School Students in Taiwan." *International Journal of Environmental Research and Public Health*. DOI:10.3390/ijerph18031222.
- Chen, Xinyin and Bo-shu Li. 2000. "Depressed Mood in Chinese Children: Development Significance for Social and School Adjustment." *International Journal of Behavioral Development* 24(4): 472-479.
- Crosnoe, Robert, Riegle-Crumb Catherine, and Muller Chandra. 2007. "Gender, Self-Perception, and Academic Problems in High School." *Social Problems* 54(1): 118-138.
- De Zwart, P. L., B. F. Jeronimus, and P. de Jonge. 2019. "Empirical Evidence for Definitions of Episode, Remission, Recovery, Relapse and Recurrence in Depression: A Systematic Review." *Epidemiology and Psychiatric Sciences* 28(5): 544-562.
- Dornbusch, Sanford M., Philip L. Ritter, P. Herbert Leiderman, Donald F. Roberts, and Michael J. Fraleigh. 1987. "The Relation of Parenting Style to Adolescent School Performance." *Child Development* 58(5): 1244-1257.
- Duncan, Otis D., Magnus Stenbeck, and Charles Brody. 1988. "Discovering Heterogeneity: Continuous versus Discrete Latent Variables." *American Journal of Sociology* 93(6): 1305-1321.
- Gruber, Reut, Jamie Cassoff, Sonia Frenette, Sabrina Wiebe, and Julie Carrier. 2012. "Impact of Sleep Extension and Restriction on Children's Emotional Lability and Impulsivity." *Pediatrics* 130(5): e1155-e1161.
- Gujar, Ninad, Seung-Schik Yoo, Peter Hu, and Matthew P. Walker. 2011. "Sleep Deprivation Amplifies Reactivity of Brain Reward Networks, Biasing the Appraisal of Positive Emotional Experiences." *The Journal of Neuroscience* 31(12): 4466-4474.
- Hartog, Joop, Henriette Maassen, and Van den Brink. 2007. *Human Capital: Advances in Theory and Evidence*. New York: Cambridge University Press.
- Kim, Hyun Sik. 2011. "Consequences of Parental Divorce for Child Development." *American Sociological Review* 76(3): 487-511.
- Kohn, Melvin L. 1963. "Social Class and Parent-Child Relationships: An Interpretation." *American Journal of Sociology* 68(4): 471-480.
- Kuan, Ping-Yin. 2018. "Effects of Cram Schooling on Academic Achievement and Mental Health of Junior High Students in Taiwan." *Chinese Sociological Review* 50(4): 391-422.
- Lareau, Annette. 1987. "Social Class Differences in Family-School Relationships: The Importance of Cultural Capital." *Sociology of Education* 60(2): 73-85.
- Lareau, Annette. 2011. *Unequal Childhoods: Class, Race, and Family Life*. Los Angeles: University of California Press.
- Lin, Huang-Chi, Tze-Chun Tang, Ju-Yu Yen, Chin-Hung Ko, Chi-Fen Huang, Shu-Chun Liu, and Cheng-Fang Yen. 2008. "Depression and Its Association with Self-Esteem, Family, Peer and School Factors in a Population of 9586 Adolescents in Southern Taiwan." *Psychiatry and Clinical Neurosciences* 62(4): 412-420.
- Liu, Airan. 2019. "Can Non-Cognitive Skills Compensate for Background Disadvantage? — The Moderation of Non-Cognitive Skills on Family Socioeconomic Status and Achievement during Early Childhood and Early Adolescence." *Social Science Research* 83: 102306.
- Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* 60(3): 531-542.
- McCarty, Carolyn A., W. Alex Mason, Rick Kosterman, J. David Hawkins, Liliana J. Lengua,

- and Elizabeth McCauley. 2008. “Adolescent School Failure Predicts Later Depression among Girls.” *Journal of Adolescent Health* 43(2): 180–187.
- McLeod, Jane D. and Danielle L. Fettes. 2007. “Trajectories of Failure: The Educational Careers of Children with Mental Health Problems.” *American Journal of Sociology* 113(3): 653–701.
- Morgan, Stephen and Christopher Winship. 2015. *Counterfactuals and Causal Inference: Methods and Principles for Social Research* (Second Edition). New York: Cambridge University Press.
- Needham, Belinda, Robert Crosnoe, and Chandra Muller. 2004. “Academic Failure in Secondary School: The Inter-Related Role of Health Problems and Educational Context.” *Social Problems* 51(4): 569–586.
- Pasch, Keryn E., Melissa N. Laska, Leslie A. Lytle, and Stacey G. Moe. 2010. “Adolescent Sleep, Risk Behaviors, and Depressive Symptoms: Are They Linked?” *American Journal of Health Behavior* 34(2): 237–248.
- Steinberg, Laurence, Sanford M. Dornbusch, and B. Bradford Brown. 1992. “Ethnic Differences in Adolescent Achievement: An Ecological Perspective.” *The American Psychologist* 47(6): 723–729.
- Stock, James H. and Mark W. Watson. 2015. *Introduction to Econometrics* (Third Edition). New York: Pearson.
- VanderWeele, Tyler J. 2015. *Explanation in Causal Inference: Methods for Mediation and Interaction*. New York: Oxford University Press.
- Waldvogel, Jacki L., Martha Rueter, and Charles N. Oberg. 2008. “Adolescent Suicide: Risk Factors and Prevention Strategies.” *Current Problems in Pediatric and Adolescent Health Care* 38(4): 110–125.
- Xie, Yu, Christopher Near, Hongwei Xu, and Xi Song. 2020. “Heterogeneous Treatment Effects on Children’s Cognitive/non-cognitive Skills: A Reevaluation of an Influential Early Childhood Intervention.” *Social Science Research* 86: 102389.
- Xie, Yu. 2007. “Otis Dudley Duncan’s Legacy: The Demographic Approach to Quantitative Reasoning in Social Science.” *Research in Social Stratification and Mobility* 25(2): 141–156.
- Yi, Chin-Chun. 2013. *The Psychological Well-being of East Asian Youth*. Dordrecht: Springer.
- Yoo, Seung-Schik, Ninad Gujar, Peter Hu, Ferenc A. Jolesz, and Matthew P. Walker. 2007. “The Human Emotional Brain without Sleep—A Prefrontal Amygdala Disconnect.” *Current Biology* 17(20): R877–R878.
- Zhang, Yu. 2013. “Does Private Tutoring Improve Students’ National College Entrance Exam Performance?—A Case Study from Jinan, China.” *Economics of Education Review* 32: 1–28.
- Zheng, Xiaodong, Chengcheng Wang, Zheng Shen, and Xiangming Fang. 2020. “Associations of Private Tutoring with Chinese Students’ Academic Achievement, Emotional Well-Being, and Parent-Child Relationship.” *Children and Youth Services Review* 112. DOI:10.1016/j.childyouth.2020.104934.