

初中生校外辅导：补差还是培优？

——基于中国教育追踪调查数据的实证研究

徐章星^{1,2}

(1.南京农业大学 金融学院, 江苏 南京 210095; 2.德国哥廷根大学 农业经济与农村发展系, 德国 哥廷根 37073)

摘要：利用中国教育追踪调查(CEPS2013)数据,运用双变量Probit模型,基于参与时间异质性视角,分析初中生学业成绩与校外辅导班参与行为之间的关系,针对不同学生群体进行了分样本讨论,并进行了稳健性检验。研究表明:在“周一到周五校外辅导班参与”的群体中,学生学业成绩与校外辅导班参与负相关,即学业成绩越好,参与辅导班概率越小;在“周末校外辅导班参与”的群体中,随着学业成绩的提高,学生校外辅导班的参与概率经历了一个先上升后下降的“倒U型”趋势,平均而言,当语数外文化课总分为236分时,学生参加校外辅导班的可能性最高。“周一到周五校外辅导班参与”的群体中,学业成绩与校外辅导班参与之间的负相关关系在城市学生和七年级学生中更明显;在“周末校外辅导班参与”的群体中,学业成绩与校外辅导班参与概率之间的“倒U型”关系在农村学生和毕业班学生中更显著,城市学生学业成绩越好,参与校外辅导的可能性越高,存在“成绩越好,补课越疯狂”的现象。

关键词：学业成绩;校外辅导;行为选择;异质性;双变量Probit模型

中图分类号:G632.4

文献标志码:A

文章编号:1009-2013(2020)03-0074-10

Whether junior high school students' tutoring makes up the difference or improves the quality: Empirical evidence from CEPS

XU Zhangxing^{1,2}

(1.College of Finance, Nanjing Agricultural University, Nanjing 210095, China; 2.Department of Agricultural Economics and Rural Development, University of Goettingen, Goettingen 37073, Germany)

Abstract: Based on the perspective of participation time heterogeneity, this paper uses the China Education Panel Survey (CEPS2013) data to analyze the relationship between academic achievement and the choice of students' tutoring behavior via the Bivariate Probit model. With the endogenous treatment, this paper conducts a sample discussion on different students' groups and conducts a robustness test. The results show that: From Monday to Friday, students' academic achievement is negatively correlated with the participation of tutoring, that is, the higher the academic performance, the lower the probability of participating in tutoring classes is; At the weekend, with the improvement of students' academic achievement, the participation probability of tutoring experiences an “inverted U-shaped” trend of rising first and then decreasing. On average, when the total grades of classes is 236, students' possibility of tutoring reaches the maximum. From Monday to Friday, the negative correlation between academic performance and participation in out-of-school tutoring classes is more pronounced among urban students and seventh graders. On weekends, the “inverted U” relationship between academic performance and the probability of participation in out-of-school tutoring classes is more pronounced among rural students and graduating students; the better the academic performance of urban students, the higher the likelihood of participating in out-of-school tutoring is and there is a phenomenon of “the higher the grades, the more crazy the participation of tutoring is”.

收稿日期:2020-05-25

基金项目:国家自然科学基金青年项目(71703070);

江苏省普通高校研究生科研创新计划(KYCX18_0752);

国家建设高水平大学公派研究生项目(CSC201906850095)

作者简介:徐章星(1991—),男,江苏苏州人,博士研究生,主要研究方向为教育经济、农业经济。

Keywords: academic achievement; tutoring; behavior choice; heterogeneity; Bivariate Probit model

一、问题的提出

近年来，出于“减负”目的，中小学课堂教学内容逐步趋于简单，但竞争性的升学机制未发生变化，强烈的社会需求为校外辅导机构带来了巨大的市场契机，“下课不放学”“不是从补习班下课，就是在去补习班的路上”成为很多中小学生学习生活的真实写照。选择参加补习班的不仅有“成绩一般”的同学，也有成绩好的同学，“越是成绩好的学生越补课，好学生补课也疯狂”正成为一种新的趋势。校外辅导班的小班化、分层化教学，是弥补公办学校应试教育不足的一个有效手段。那么，到底是成绩较差的学生还是成绩较好的学生更有可能参加校外辅导班？通常情况下，不论成绩较差还是成绩较好的学生，都有可能参加校外辅导班。一方面，成绩较差的学生参加校外辅导班可以有针对性地查漏补缺，巩固课堂所学知识，并养成良好的学习习惯；另一方面，成绩较好的学生通常在学习上更加自觉和具有目的性，参加校外辅导班能够提升优势科目，在学习上取得更多的优势。因此，对这一问题的回答充满了不确定性。

校外辅导，是指在市场主导下，在学校教育外进行的具有经济行为性质的有偿教育活动，其内容和形式伴随着课堂教育而存在，随主流教育体系课程的变化而改变，因此被形象地称为“影子教育”（shadow education）。2005 年之后，政府一系列减负政策的出台，限制了学校和老师的校外辅导行为，为了取得更好的学习成绩，家长和学生选择了专门的校外辅导机构和老师，与基础教育相关联的规模化、市场化的校外辅导行业开始发展。根据《中国教育新业态发展报告——基础教育》（CIEFR-HS 2017）显示，中小学阶段学生校外培训总体参与率为 48.3%，参与校外培训的学生平均支出为 5 616 元，平摊的生均支出为 2 697 元。根据我国中小学阶段在校生规模 1.871 7 亿进行估计，全国校外培训规模超过 4 900 亿元。从不同学段的差异来看，全国范围内小学生学科补习参与率为 33.4%，初中生为 43.7%，高中生为 48.2%，相对应的生均支出分别为 1 475、2 443 和 2 691 元/年。

随着校外辅导行业的发展与扩张，国内外学者从不同角度对课外补习的影响因素进行了研究。许多学者发现校外辅导参与和家庭经济条件有关，与

中等和低收入家庭相比，高收入家庭其子女参与校外辅导的概率较高^[1-3]。来自爱尔兰和土耳其的经验证据表明，校外辅导班参与具有社会阶层化的特征，社会地位较高的阶层收入和受教育程度较高，更加注重子女教育，也更能够支付得起高质量的校外辅导费用^[4,5]。关于学业成绩与学生校外辅导参与的关系，Bray & Kwok^[6]从教育需求的角度出发，通过对香港学生校外辅导参与进行分析与讨论，认为出于同伴间的竞争压力，成绩较好的学生参与课外补习的概率高于成绩较差的学生。方晨晨和薛海平^[7]利用 2012 年北京大学中国社会科学调查中心数据，探讨了课外补习与学生成绩之间的关系，发现重点学校的学生补习概率显著高于非重点校，存在“成绩越好，补课越疯狂”的现象。雷万鹏^[8]通过实地调研发现，中国高中生教育补习主要是“补差”，即教育补习主要帮助成绩落后学生提高成绩。贺建清^[9]通过对江西、湖南两省在校城镇初中生进行问卷调查发现，就读于重点班和学业成绩较好的学生参加校外补习的概率较低，认为当前的校外补习是“补差”教育，而非“培优”教育。乐志强、杜育红^[10]根据中国家庭动态跟踪调查（CFPS2012）数据，通过 Logistic 回归方法研究发现，学科类校外辅导具有刚性需求的特征，不论学生成绩如何，只要有可能，父母都会选择学校课程辅导以增加子女的竞争优势。此外，城乡差异^[11]、性别差异^[12]、学校教育质量^[13]等因素也将影响学生校外辅导的行为选择。

根据“补课热”的现状，基于参与时间异质性视角，笔者利用 2013 年中国教育追踪调查（CEPS）数据库，研究学业成绩对学生校外辅导行为选择的影响，以期为学生校外辅导行为选择以及当前教育改革提供依据。

二、研究设计

（一）研究方法

1. 双变量 Probit 模型

为了测度学业成绩对学生校外辅导行为选择的影响，按照学生参与校外辅导班时间差异，将学生划分为“周一到周五校外辅导班参与”和“周末校外辅导班参与”两个群体。因为学生参加辅导班的选择可能不相互独立，在周一到周五参加校外辅

导班的选择可能会影响学生周末校外辅导班的参与选择；同时，学生周末校外辅导班参与选择也可能影响到周一到周五校外辅导班的参与选择。据此，本文采用双变量 Probit 模型 (Bivariate Probit model) 估计学生校外辅导行为选择。双变量 Probit 模型有两个方程，第一个方程表示学生周一到周五校外辅导班的参与行为，第二个方程表示学生周末校外辅导班的参与行为。在双变量 Probit 模型中，两个方程具有相同的控制变量，学生周一到周五校外辅导班的参与行为和周末校外辅导班的参与行为相关，误差项也相关，同时误差项的协方差等于一个固定的常数。模型的具体设定如下：

以 J_m 表示学生周一到周五校外辅导班的参与情况， $J_m=1$ 表示参加， $J_m=0$ 表示未参加；以 J_n 表示学生周末校外辅导班的参与情况， $J_n=1$ 表示参加， $J_n=0$ 表示未参加； J_m^* 和 J_n^* 为学生校外辅导班参与行为的潜变量 (latent variable)，满足：

$$\begin{cases} J_m^* = \alpha_1 A + \alpha_2 X + \eta_1 \\ J_n^* = \beta_1 A + \beta_2 X + \eta_2 \end{cases} \quad (1)$$

在方程 (1) 中， A 表示学生学业成绩，以语数外三门文化课总分衡量； X 为控制变量，表示影响学生校外辅导行为选择的其他可能因素，包括个体情况、家庭特征以及所在学校特征等； α 和 β 为待估计参数，假定随机扰动项 η_1 和 η_2 满足均值为 0，标准差为 1 的联合正态分布。如果 η_1 和 η_2 相关系数等于 0，则该模型等价于两个独立的单变量 Probit 模型，只需分别对其进行估计；如果 η_1 和 η_2 相关系数不等于 0，则说明两个方程的误差项相关，学生周一到周五参加校外辅导班的选择和周末参加校外辅导班的选择受到一些不可观测的因素共同影响，其中 η_1 和 η_2 相关系数大于 0，说明增加学生周一到周五参加校外辅导班概率的不可观测因素同时也会增加其周末参加校外辅导班的概率， η_1 和 η_2 相关系数小于 0 则说明增加学生周一到周五参加校外辅导班概率的不可观测因素会同时降低学生在周末参加校外辅导班的可能性。

双变量 Probit 模型定义为：

$$J_m = \begin{cases} 1 & \text{如果 } J_m^* > 0 \\ 0 & \text{如果 } J_m^* \leq 0 \end{cases} \quad J_n = \begin{cases} 1 & \text{如果 } J_n^* > 0 \\ 0 & \text{如果 } J_n^* \leq 0 \end{cases} \quad (2)$$

通过以上模型的设定可以发现，对于校外辅导参与，学生会有以下四种选择：既在周一到周五参

加校外辅导班又在周末参加校外辅导班、只在周一到周五参加校外辅导班、只在周末参加校外辅导班以及不参加校外辅导班，即 (1, 1)、(1, 0)、(0, 1) 和 (0, 0) 四种可能的结果。为了提高模型的估计结果及效率，此处采用最大似然法对学生校外辅导行为选择进行联合估计。

当 $J_m=1, J_n=1$ 时，

$$\Pr(J_m=1, J_n=1) = F(\alpha_1 A + \alpha_2 X, \beta_1 A + \beta_2 X; \rho) \quad (3)$$

当 $J_m=1, J_n=0$ 时，

$$\Pr(J_m=1, J_n=0) = \Phi(\alpha_1 A + \alpha_2 X) - F(\alpha_1 A + \alpha_2 X, \beta_1 A + \beta_2 X; \rho) \quad (4)$$

当 $J_m=0, J_n=1$ 时，

$$\Pr(J_m=0, J_n=1) = \Phi(\beta_1 A + \beta_2 X) - F(\alpha_1 A + \alpha_2 X, \beta_1 A + \beta_2 X; \rho) \quad (5)$$

当 $J_m=0, J_n=0$ 时，

$$\Pr(J_m=0, J_n=0) = 1 - \Phi(\alpha_1 A + \alpha_2 X) - \Phi(\beta_1 A + \beta_2 X) + F(\alpha_1 A + \alpha_2 X, \beta_1 A + \beta_2 X; \rho) \quad (6)$$

对数似然函数为：

$$\text{LnL} = \sum_i \left\{ \begin{aligned} & J_m J_n \text{Ln} F(\alpha_1 A + \alpha_2 X, \beta_1 A + \beta_2 X; \rho) \\ & + J_m (1 - J_n) \text{Ln} [\Phi(\alpha_1 A + \alpha_2 X) - F(\alpha_1 A + \alpha_2 X, \beta_1 A + \beta_2 X; \rho)] \\ & + (1 - J_m) J_n \text{Ln} [\Phi(\beta_1 A + \beta_2 X) - F(\alpha_1 A + \alpha_2 X, \beta_1 A + \beta_2 X; \rho)] \\ & + (1 - J_m) (1 - J_n) \text{Ln} [1 - \Phi(\alpha_1 A + \alpha_2 X) - \Phi(\beta_1 A + \beta_2 X) \\ & + F(\alpha_1 A + \alpha_2 X, \beta_1 A + \beta_2 X; \rho)] \end{aligned} \right\} \quad (7)$$

(3)~(7) 式中， (\cdot) 为一元标准正态分布累积函数， $F(\cdot)$ 为二元正态累积分布函数。

2. 模型的内生性问题

模型 (1) 中可能存在内生性问题。一方面，在回归模型中，可观测变量以核心变量和控制变量的形式进入模型，但是有一些与学生校外辅导行为选择相关的不可观测因素的缺失导致估计结果有偏。如果影响学生校外辅导行为选择的不可观测变量不对其学业成绩产生影响，即学业成绩是外生的，那么可以得到学业成绩对学生校外辅导行为选择的一致估计。另一方面，可能存在反向因果，学生参加校外辅导班的行为也会影响其学业成绩。为了解决模型的内生性问题，本文通过在双变量 Probit 模型中引入工具变量来处理模型的内生性。以学生六年级时学习态度作为工具变量：第一，对于初中生，学业成绩与学习态度具有较强的相关性。相关研究表明，学习态度越认真，学生在课堂上的反应能力越快，有助于培养独立意识和分析问题的能力，进而提高学业成绩^[14-16]。第二，学生的学习态度与学生校外辅导行为选择之间具有较强

的外生性，学生初中时的校外辅导行为选择不会影响到学生六年级时的学习态度，而且没有证据和理由认为学生六年级时候的学习态度影响到了初中时期校外辅导行为选择。

由于双变量 Probit 模型缺乏对模型内生性的直接处理方法，借鉴 Roodman David^[17]的研究思路，采取两步法对具有内生性的双变量 Probit 模型进行估计。第一阶段，用内生变量学业成绩作为因变量对工具变量和其他外生变量做 Probit 回归，得到学业成绩的估计拟合值。第二阶段，将学业成绩的估计拟合值作为核心变量代入估计模型，采用条件混合过程估计方法对双变量 Probit 模型进行二阶段回归。

（二）数据来源

本文数据来源于中国教育追踪调查（CEPS）。CEPS 调查是中国人民大学调查与数据中心开展的大型追踪项目，目的在于揭示学校、家庭以及社会等因素对于个人教育产出的影响，并进一步探讨教育产出在个人生命历程中发挥的作用。中国教育追踪调查(CEPS 2013—2014)以七年级（初一）和九年级（初三）两个同期群为调查起点，以人口平均受教育水平和流动人口比例为分层变量，从全国随机抽取 28 个县级单位（县、区、市）作为调查点。本调查以学校为基础，在入选的县级单位随机抽取 112 所学校、438 个班级，被抽中班级的学生全体入样，基线调查共调查了 1 万多名学生。

在实证分析中，首先将学生问卷、家庭问卷和学校问卷进行匹配，然后将学生每天学习、睡眠以及校外辅导参与等时间加总，剔除时间超过 24 小时的样本。最后剔除变量数据缺失或记录为“不知道”等的样本，最终得到 8 527 个有效样本。

（三）变量说明

1. 因变量

在衡量学生校外辅导行为选择时，根据 CEPS 问卷“上周一到上周五（周末），你平均每天上校外辅导班（与课业有关）时间为多少？”的回答情况来确定，若上校外辅导班时间为 0，则其未参加辅导班；如上校外辅导班时间不为 0，则认为其参加校外辅导班。在统计过程中，对样本学生周一到周五和周末校外辅导班的参与情况分别进行统计，将其作为学生校外辅导参与行为方程中的被解释

变量。考虑到不同时间段校外辅导过程中可能存在的补差和培优行为，该指标基于参与时间的异质性，分离了学生校外辅导班的参与行为，比学生是否参与校外辅导更能客观地反映学业成绩对于学生校外辅导班参与决策的真实影响。

2. 核心自变量

为考察学业成绩对学生校外辅导班参与选择的影响，本文的核心自变量用语数外三门文化课的总分来衡量，同时为了克服总分不同带来的差异性，将学业成绩进行了标准化处理。周一到周五，学业成绩较好的学生在完成课堂内的学业任务后，其对于休闲时间的渴望大于对成绩提升的需求，相较于学业成绩较差的学生，其参与校外辅导的概率相对较低。在周末，学生的闲暇时间相对较多，能够挤出部分娱乐时间参与校外辅导，巩固课堂知识、查漏补缺，促进学业成绩的提升。值得注意的是，对于学业成绩很差的学生，其学习积极性不高导致其校外辅导参与可能性较低；对于学业成绩很好的学生，其已达学业成绩的“高原地带”，如若继续增加校外辅导参与时间，学业负担不断加重，校外辅导未能实现预期效果。相对而言，中等学业成绩的学生“可塑性”相对较高，参与校外辅导的可能性最高。因此，在周末学生学业成绩与校外辅导参与之间存在“倒 U 型”关系。

3. 其他控制变量

其他控制变量主要考虑了学生校外辅导行为选择的其他影响因素，包括学生性别、户口、家庭经济情况、学校教学质量等。学生性别对校外辅导班的参与可能存在负向影响。根据近年来初中阶段出现的“女优于男”特征，可以看出家长对女孩的期望逐步提升，这在一定程度上会促进女生参与校外辅导。探讨农村学生、留守学生和流动学生的受教育机会是教育公平问题的关注重点：由于农村教育资源相对稀缺，学生参与校外辅导的可能性相对较低；对于留守学生，父母长期在外地工作，缺乏关爱和学业上的监督，他们参与校外辅导可能性较低；流动学生从农村迁至城市/县城上学，在短期内可能跟不上学校的教学进度，在课后需要补习以弥补与同学之间的差距，因此其参与校外辅导的可能性相对较高。父母受教育年限越长，重视子女教育的概率越大，越有可能敦促子女参与校外辅导以提

升学业成绩。家庭经济条件衡量的是与学生校外辅导参与相关的经济成本,家庭经济条件越好,父母更愿意将子女送至高质量的校外辅导机构或参与一对一的辅导方式以促进子女学业成绩的提升。独生子女除了能从家庭中获得较多的物质保障外,也能通过参与校外辅导得到父母在学业上的关心和监督。学校教学质量对于学生校外辅导参与行为的影响是不确定的。一方面,教学质量越高的学校学生学业成绩较好的可能性越高,参与校外辅导的需求相对较低;另一方面,教学质量越高的学校学生和家长对于学业成绩的要求也可能越高,更有可能

参与校外辅导,即存在“成绩越好,补课越疯狂”的现象。关于学校属性,一方面公立学校教师队伍结构良好,教学经验丰富,有助于学生认知能力的开发,减少校外辅导参与;另一方面公立学校虽然进行免费义务教育,但大班式教育不能有效提升学生学业成绩,因此学生参与校外辅导概率相对较高。此外,私立初中老师对学生的督促较强,可能降低学生校外辅导参与。因此,学校属性对于学生校外辅导的参与的影响仍需检验。各变量的定义及描述性统计如表1所示。

表1 变量的定义及描述性统计结果

变量名称	变量定义及描述	均值	标准差
校外辅导班参与选择	学生在周一到周五参加校外辅导班(是=1,否=0)	0.090 9	0.287 4
	学生在周末参加校外辅导班(是=1,否=0)	0.255 8	0.436 3
性别	学生性别(男=1,女=0)	0.498 6	0.500 0
学业成绩	学生2013年期中考试标准化成绩(语文、数学和英语成绩之和)	213.741 5	24.478 4
户口	学生户口(农业户口=1,非农业户口=0)	0.535 3	0.498 7
流动学生	学生从农村移至此市/县上学(是=1,否=0)	0.165 0	0.371 2
留守学生	学生父母长期在外地工作(是=1,否=0)	0.096 9	0.295 9
父亲受教育年限	按照我国教育体制将其转化为相应的受教育年限(年)	11.614 9	2.801 4
母亲受教育年限	按照我国教育体制将其转化为相应的受教育年限(年)	9.578 4	3.360 2
家庭经济	学生家庭经济水平(困难=1,中等=2,富裕=3)	1.855 7	0.482 9
独生子女	学生是独生子女(是=1,否=0)	0.456 6	0.498 1
教学质量	学生所在学校学历为本科及以上的教师所占比例	0.732 1	0.282 4
学校属性	学校属性(公立=1,私立=0)	0.928 1	0.258 3

三、结果与讨论

(一) 描述性统计分析

根据表1可知,周一到周五参加校外辅导班和周末参加校外辅导班的受访者分别占比为9.09%和25.58%,说明在周一到周五参与校外辅导班的学生相对较少。核心自变量学生学业成绩的平均值为213.74分,标准差为24.28。其他控制变量方面,男学生占比为49.86%,农村学生占比为53.53%,可见样本男女比例和城乡比例较为均衡。在全样本中,流动学生和留守学生分别为1407和827人,占比分别为16.50%和9.69%,该部分群体是教育的公平性所需要重点关注的对象。父亲受教育年限均值为11.61年,母亲受教育年限均值为9.58年,同时在全样本中,父母一方拥有大专以上学历占比为20.00%。家庭经济条件均值为1.86,并且在全样本中,家庭经济条件在中等以上的有6829人,占比

为80.09%,说明家庭经济支持是学生参与校外辅导的一个必要条件。样本独生子女占比为45.66%,该指标能够从另一个角度反映家庭经济能力对于子女校外辅导参与的影响。学校特征方面,学生所在学校本科及以上学历的教师所占比例均值为73.21%,学校本科及以上学历的教师所占比例越高,教学质量越好。公立学校占比为92.81%,说明在当前初中教育体系下,公立学校仍占绝大部分。

进一步研究发现,在全样本中,周一到周五参加校外辅导班的学生人数为775人,占比为9.09%;周末参加辅导班的人数为2181人,占比为25.58%(表2)。具体来看,有663名受访学生在周一到周五和周末同时参加了校外辅导班,占比7.78%;周一到周五参加辅导班但周末未参加校外辅导班人数为112人,占比为1.31%,可能的原因是该部分学生喜欢在工作日保持高强度的学习,在周末选择休息;有1518名受访学生选择了在周末

参加辅导班但在周一到周五未参与，一个合理的解释是该部分学生为了防止“本末倒置”，在周一到周五选择了以学校老师布置的课内学习为主，在周末选择参加校外辅导以进一步巩固课堂知识；在周一到周五和周末都没有参加辅导班的学生人数为 6 234 人，占比 73.11%。

表 2 学生校外辅导行为选择及占比

	周一到周五参加		周一到周五未参加		合计	
	人数	占比/%	人数	占比/%	人数	占比/%
周末参加	663	7.78	1 518	17.80	2 181	25.58
周末未参加	112	1.31	6 234	73.11	6 346	74.42
合计	775	9.09	7 752	90.91	8 527	100.00

(二) 学业成绩对校外辅导行为选择的影响

图 1 显示了学业成绩与校外辅导班参与选择之间的相关性。从全样本来看，在“周一到周五校外辅导班参与”群体中，随着学业成绩的上升，学生参与校外辅导班的概率会降低；在“周末校外辅导班参与”群体中，学业成绩与学生校外辅导班参与概率之间呈现“倒 U 型”关系。在“周一到周五学生校外辅导班参与”模型中引入学业成绩的一次项；在“周末学生校外辅导班参与”模型中引入学业成绩的一次项和二次项，回归结果如表 3 所示。

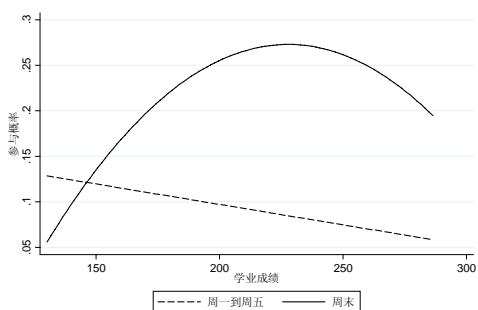


图 1 学业成绩与校外辅导班参与的相关性

双变量 Probit 模型的 ρ 值为 0.76906 (表 3)，同时通过了似然比显著性检验，说明学生周一到周五辅导班参与方程和周末辅导班参与方程之间存在联立关系。在“周一到周五校外辅导班参与”群体中，学生校外辅导班参与概率与学业成绩负相关，回归系数为 -0.004 06，在 1% 的水平上显著；在“周末校外辅导班参与”群体中，学生学业成绩与校外辅导班参与概率之间呈现一个先升后降的“倒 U 型”，学业成绩一次项和二次项系数分别为 0.018 88 和 -0.000 04，均在 5% 的显著水平上显著。由此可见，随着学业成绩的上升，学生参与校外辅导班的概率

上升，平均而言，当学业成绩到达 236 分时，学生参与校外辅导班的概率最高，随着学业成绩的进一步上升，学生在周末参与校外辅导班的概率将降低。对于以上现象的合理解释是在周一到周五，成绩较差的学生更迫切地需要提分，参与校外辅导班的概率更大，成绩较好的学生在完成学校老师布置的课堂任务后，更愿意将余下的时间用来休息，防止过度疲惫。因此，随着学生学业成绩的提升，将降低学生在周一到周五参与校外辅导的概率。在周末，随着学业成绩的增加，学生对于学业成绩提升需求增加，参与校外辅导班的概率开始上升，学生在保证休息的前提下挤出部分娱乐时间，当学业成绩达到并超过拐点后，学生参与校外辅导班的概率降低。相比较，学业成绩中等的学生参与校外辅导班概率较高，可能的原因是成绩较差和较好的学生学业成绩已处于“饱和状态”，“可塑性”不高，难以继续提升，成绩中等学生还有潜力，因此参与校外辅导班的概率较高。

从控制变量来看，“周一到周五校外辅导班参与”群体中，性别的回归系数为 -0.146 99，在 1% 的水平上显著，说明女生参与校外辅导班的概率较高，这与我国近些年初中阶段“女优于男”的现象一致^[18]，可能的原因是现阶段家长和社会对女孩的社会期望较高，加强了对女孩的教育投资，外在表现为女生参与校外辅导班的可能性增大。户口系数为 -0.286 03，在 1% 水平上显著，说明农村家庭经济水平相对较低，而参与校外辅导班花费较高，因此学生参与校外辅导班的概率相对较低，家庭经济条件系数及显著水平验证了此观点(系数为 0.086 41，5% 的显著水平上显著)。独生子女系数为 0.090 38，在 5% 水平上显著，说明独生子女在家庭中获得了足够的教育资源，提升了其参与校外辅导班的概率。学校属性系数为 0.153 49，在 10% 水平上显著，一个合理的解释是公立学校相对较多的空余时间提升了学生参与校外辅导班的可能性。在“周末校外辅导班参与”群体中，流动学生系数为 0.173 20，在 1% 的水平上显著，说明流动学生从农村到城市来学习，随着学习条件的改善，对于校外辅导的需求也提升；留守学生系数为 -0.160 47，在 1% 的水平上显著，可能的原因是留守学生家庭经济条件较差，参与校外辅导的概率较低；学校教育质量的系数

为0.263 22,在1%水平上显著,说明学校教育质量较高也加强了学生间的竞争,进而提升了学生参与校外辅导班的概率。值得注意的是,在“周一到周五校外辅导班参与”群体中,父亲受教育年限对子女校外辅导参与影响为正,但不显著,母亲受教育年限在1%水平上显著为正,在“周末校外辅导班参与”群体中,父母受教育年限同时正向影响子女校外辅导参与,可能的原因是,在工作日内母亲照顾子女的可能性较大,对子女学业监督和校外辅导的参与影响较大。其他控制变量影响方向和显著程度与“周一到周五学生校外辅导班参与”方程的估计结果基本一致。

表3 双变量 Probit 模型回归结果

变量	周一到周五	周末
学业成绩一次项	-0.004 06*** (0.000 8)	0.018 88** (0.007 8)
学业成绩二次项		-0.000 04** (0.000 0)
性别	-0.146 99*** (0.039 4)	-0.163 26*** (0.032 6)
户口	-0.286 03*** (0.045 1)	-0.443 34*** (0.036 1)
流动学生	0.084 52 (0.053 3)	0.173 20*** (0.043 7)
留守学生	0.002 54 (0.071 3)	-0.160 47*** (0.062 2)
父亲受教育年限	0.009 72 (0.008 8)	0.024 52*** (0.007 5)
母亲受教育年限	0.036 25*** (0.007 7)	0.065 02*** (0.006 7)
家庭经济	0.086 41** (0.042 9)	0.239 16*** (0.036 6)
独生子女	0.090 38** (0.045 3)	0.306 24*** (0.036 1)
教学质量	0.088 01 (0.071 1)	0.263 22*** (0.062 1)
学校属性	0.153 49* (0.092 3)	0.262 74*** (0.081 7)
常数项	-1.185 58*** (0.216 6)	-4.355 74*** (0.810 1)
观测值	8 527	
对数似然值	-6 023.783 6	
Wald卡方值	1 176.18	
ρ	0.769 06(0.014 8)	
P值	Prob> $\chi^2=0.000 0$	

注: *、**和***分别表示10%、5%和1%的显著性水平; (·)表示标准误。下同。

(三) 内生性问题的处理

为了证明双变量 Probit 模型估计结果的可靠性,同时考虑到模型可能存在的内生性问题,根据上文分析,采用学生学习态度作为工具变量,根据学生对于 CEPS 问卷的问题“六年级时,就算是我喜欢的功课,我也会尽全力去做?”的答案(“完全不同意、不太同意、比较同意、完全同意”),将“完全不同意”“不太同意”“比较同意”和“完全同意”四个回答分为赋值为1、2、3、4,并采用条件混合过程估计方法(CMP)进行估计。估计结果如表4所示。

表4 考虑变量内生性的双变量 Probit 模型一阶段回归结果

变量	学业成绩
学习态度	3.364 1*** (0.318 5)
其他控制变量	已控制
观测值	8,527
R^2	0.081 4
F值	66.75
P值	Prob > F=0.000 0

根据一阶段的回归结果,学习态度对学生学业成绩影响系数为3.3641,在1%水平上显著。该结果表明:学习态度具有连续性,学生六年级时形成的优良学习态度依旧能够促进其初中时候学业成绩的提升。不同于博士阶段的科研型学习,中学阶段的学习仍是知识传播型,端正学习态度并形成良好的学习习惯,学生学业成绩提高相对更容易。

采用条件混合估计法(CMP)对双变量 Probit 模型进行第二阶段回归,结果如表5所示。结果表明,在处理了学业成绩的内生性问题后,选择周一到周五参与校外辅导的,学业成绩与校外辅导班参与选择负相关,影响程度有所提高;选择周末参与校外辅导的,学业成绩一次项和二次项分别在1%和5%的水平上显著为正或负,学业成绩与校外辅导班参与选择之间呈现“倒U型”,其余解释变量方向和显著程度与前文基本保持一致,说明前文的估计结论是稳健的。

(四) 异质性考察

上文实证检验了学生学业成绩与校外辅导班参与选择之间的关系,并且在处理了内生性后结果依旧稳健。那么,对于不同学生群体而言,学业成绩对校外辅导班参与选择的影响是否具有差异

性？据此，本研究按学生户籍和年级两个方面对学生进行分类，分别采用工具变量进行两阶段估计，

以考察学生校外辅导行为选择的异质性，结果见表 6 和表 7。

表 5 考虑变量内生性的二阶段双变量 Probit 模型回归结果

变量	周一到周五	周末	变量	周一到周五	周末
学业成绩一次项	-0.006 77** (0.002 9)	0.021 73*** (0.007 5)	家庭经济	0.086 33** (0.043 3)	0.237 65*** (0.036 0)
学业成绩二次项		-0.000 04** (0.000 0)	独生子女	0.091 73** (0.045 0)	0.303 38*** (0.036 2)
性别	-0.178 25*** (0.050 9)	-0.113 46** (0.048 4)	教学质量	0.086 13 (0.070 7)	0.265 43*** (0.060 2)
户口	-0.283 66*** (0.045 9)	-0.444 14*** (0.036 4)	学校属性	0.147 94 (0.094 4)	0.268 58*** (0.077 1)
流动学生	0.090 04* (0.053 4)	0.163 20*** (0.043 5)	常数项	-0.611 57 (0.632 8)	-5.101 38*** (0.924 6)
留守学生	-0.001 16 (0.071 3)	-0.153 25** (0.061 0)	观测值		8,527
父亲受教育年限	0.011 58 (0.008 7)	0.021 37*** (0.007 5)	对数似然值		-43897.841
母亲受教育年限	0.037 07*** (0.007 7)	0.062 97*** (0.006 6)	Wald卡方值		2108.60
			P值		Prob> $\chi^2=0.0000$

表 6 城乡异质性考察

变量	农村		城市	
	周一到周五	周末	周一到周五	周末
学业成绩一次项	-0.005 97 (0.004 2)	0.018 02* (0.010 8)	-0.007 46* (0.004 0)	0.026 20** (0.010 8)
学业成绩二次项		-0.000 04* (0.000 0)		-0.000 04* (0.000 0)
其他控制变量	已控制	已控制	已控制	已控制
观测值		4,565		3,962
对数似然值		-22 697.382		-21 173.741
Wald卡方值		650.19		774.81
P值		Prob> $\chi^2=0.000 0$		Prob> $\chi^2=0.000 0$

表 7 年级异质性考察

变量	七年级		九年级	
	周一到周五	周末	周一到周五	周末
学业成绩一次项	-0.007 01* (0.004 1)	0.003 52 (0.010 4)	-0.005 65 (0.004 1)	0.044 97*** (0.010 7)
学业成绩二次项		0.000 00 (0.000 0)		-0.000 09*** (0.000 0)
其他控制变量	已控制	已控制	已控制	已控制
观测值		4284		4243
对数似然值		-21 834.239		-22 021.914
Wald卡方值		880.85		1255.55
P值		Prob> $\chi^2=0.000 0$		Prob> $\chi^2=0.000 0$

从城乡差异的角度来看（表 6），农村学生中，选择周一到周五参与校外辅导班的，学业成绩与校外辅导班参与选择没有显著关系，选择周末参加校外辅导班的，学业成绩与校外辅导班参与之间呈现“倒 U 型”，根据学业成绩一次项系数和二次项系数计算出拐点为 225.25 分，农村样本的 38.03% 处于该分数。由此可以推断，随着学业成绩的增加，

学生对于成绩的提升需求增加，参与校外辅导班的概率开始上升。同时，由于农村学生家庭经济基础较差，学业成绩拐点出现较早，当学业成绩达到并超过拐点后，其参加校外辅导班的概率降低。城市学生周一到周五学业成绩与校外辅导班参与选择负相关，与前文一致；在周末，城市学生学业成绩一次项系数为 0.026 20，二次项系数为-0.000 04，

分别在5%和10%水平上显著(拐点为327.50分),因此在学业成绩取值区间内,城市学生校外辅导参与和学业成绩显著正相关,其中可能的原因是城市学生竞争观念相比农村学生更强,提分的需求更高,加上城市学生家庭经济条件较好,能够支付补课费,因此学业成绩越高,参加校外辅导班的概率越高,存在“成绩越好,补课越疯狂”的现象。

从年级差异来看(表7),七年级学生“周一到周五校外辅导班参与”与学业成绩显著负相关,学业成绩越高,参加校外辅导班的概率越低,这与前文的研究结论相一致,学业成绩系数为-0.007 01,在10%水平上显著,其中可能的原因是七年级学生相对毕业班学生而言,学业压力较低,对成绩提分的要求相对较低,表现为“学业成绩越高,周一到周五参与校外辅导班的概率下降越多”;七年级学生“周末校外辅导班参与”与学业成绩没有显著关系。对于九年级学生而言,周一到周五是否参加校外辅导班与其学业成绩无显著关系;周末参加校外辅导班与学业成绩呈“倒U型”关系,可能原因是成绩中等的毕业班学生成绩提升的可能性较高,其参加校外辅导的可能性较高,而成绩较高和较差的学生学业成绩提分较难,参加校外辅导的概率较低。

(五) 稳健性检验

采用“周一到周五校外辅导班参与”和“周末校外辅导班参与”的时间作为因变量,同时保持控制变量不变,采用双变量Tobit模型进行估计,估计结果如表8所示。结果显示,双变量Tobit模型通过了似然比检验,其他控制变量方向和显著程度与前文一致,未出现显著性变化,说明前文的实证结果是稳健的。

表8 双变量Tobit的稳健性检验

变量	周一到周五	周末
学业成绩一次项	-0.010 87*** (0.002 0)	0.070 17** (0.030 1)
学业成绩二次项		-0.000 17** (0.000 1)
其他控制变量	已控制	已控制
观测值	8,527	8,527
对数似然值		-11 286.289
Wald卡方值		1 396.51
ρ		0.581 67(0.014 0)
P值		Prob> $\chi^2=0.000 0$

四、研究结论与启示

本文利用中国教育追踪调查(CEPS2013)数据,

运用双变量Probit模型,基于参与时间异质性视角,分析了学业成绩与校外辅导班行为选择之间的关系,在处理模型内生性基础上,针对不同学生群体进行了分样本讨论,并进行了稳健性检验。主要结论如下:

第一,在“周一到周五校外辅导班参与”的群体中,学生学业成绩与校外辅导班参与负相关,即学业成绩较高参与辅导班概率越小。同时,女生、城市学生、家庭经济较好、独生子女和公立学校的学生在周一到周五参加校外辅导班的概率较高。

第二,在“周末校外辅导班参与”的群体中,随着学生学业成绩的提高,学生参加校外辅导班的概率经历了一个“先上升后下降”的“倒U型”趋势,平均而言,当语数外文化课总分为236分时,学生参加校外辅导班的可能性最高。同时,女生、独生子女、流动学生、非留守学生和来自教学质量较高学校的学生在周末参加校外辅导班的概率较高。

第三,分样本来看:在“周一到周五校外辅导班参与”的农村学生中,其学业成绩与校外辅导班参与无显著关系,在“周末校外辅导班参与”的农村学生群体中,学业成绩与校外辅导班参与呈现“倒U型”关系;在“周一到周五校外辅导班参与”的城市学生中,学业成绩与校外辅导班参与呈负相关,在“周末校外辅导班参与”的城市学生中,学业成绩与校外辅导班参与正相关,存在“成绩越好,补课越疯狂”的现象。七年级学生群体中,周一到周五参与校外辅导班的学生的学业成绩与校外辅导班参与负相关,而在周末参与校外辅导班的学生,其学业成绩与是否参与校外辅导之间无明显关系;九年级学生在周一到周五参加校外辅导班,其学业成绩与校外辅导班参与无显著相关性,在周末参加校外辅导班,其学业成绩与参加校外辅导班之间呈现“倒U型”关系。

基于上述研究结论得出以下启示:第一,在周一到周五,学业成绩较差的学生首先应以学校课堂内的知识巩固为主,可以在学有余力的基础上参与校外辅导;在周末,学业成绩中等的学生可以适度增加校外辅导班时间,达到激发学习潜力又不增加学业负担的效果。第二,农村学生和留守学生校外辅导资源短缺,应有条件地推动公益性质的校外辅导,拓宽校外辅导资源的广度和深度,促进教育结果的公平化。第三,城市学生和毕业班学生在参与校外辅导的同时,应当关注学业压力可能带来的负向影响,应适度增加体育锻炼,实现校外辅导效果

的最大化。

参考文献：

- [1] STEVENSON D L ,BAKER D P .Shadow education and allocation in formal schooling : Transition to university in Japan[J] . American Journal of Sociology , 1992 , 97(6) : 1639-1657 .
- [2] 李佳丽, 胡咏梅, 范文凤. 家庭背景、影子教育和学生学业成绩——基于 Wisconsin 模型的经验研究[J]. 教育经济评论, 2016, 1(1): 70-89 .
- [3] 薛海平. 课外补习、学习成绩与社会再生产[J]. 教育与经济, 2016(2): 32-43 .
- [4] TANSEL A , BIRCAN F . Demand for education in Turkey : A tobit analysis of private tutoring expenditures[J]. Economics of Education Review , 2006 , 25(3) : 303-313 .
- [5] SMYTH E. Buying your way into college? Private tuition and the transition to higher education in Ireland[J]. Oxford Review of Education , 2009 , 35(1) : 1-22.
- [6] BRAY M , KWOK P . Demand for private supplementary tutoring : Conceptual considerations , and socio-economic patterns in Hong Kong[J] . Economics of Education Review , 2003 , 22(6) : 611-620 .
- [7] 方晨晨, 薛海平. 课外补习对义务教育阶段学生成绩影响的实证研究[J]. 上海教育科研, 2014(12): 5-9 .
- [8] 雷万鹏. 高中生教育补习支出：影响因素及政策启示[J]. 教育与经济, 2005(1): 39-42 .
- [9] 贺建清. 城市初中生课外补习的影响因素研究[J]. 现代基础教育研究, 2018, 30(2): 19-29 .
- [10] 乐志强, 杜育红. 家庭背景对学生校外辅导行为选择的影响研究[J]. 教育发展研究, 2018, 38(10): 7-14 .
- [11] 钱国英, 唐丽静. 城市义务教育阶段学生参加课外补习机率的影响因素分析——基于武汉、深圳的调查[J]. 教育财会研究, 2009, 20(3): 9-12 .
- [12] ZHANG , Yu . Does private tutoring improve students' National College Entrance Exam performance?—A case study from Jinan , China[J] . Economics of Education Review , 2013 , 32 : 1-28 .
- [13] SONG K O , PARK H J , SANG K A . A cross-national analysis of the student- and school-level factors affecting the demand for private tutoring[J] . Asia Pacific Education Review , 2013 , 14(2) : 125-139 .
- [14] 鲁志祥, 董翠敏, 冯秀梅, 等. 物理学习态度与成绩的关联研究[J]. 中学物理教学参考, 2015, 44(18): 2-4 .
- [15] 彭红兵. 高中学生英语学习态度与成绩的相关分析[J]. 江西教育科研, 2006(4): 18-19, 46 .
- [16] 吴灵敏, 林娴珍, 李玲, 等. 大学生学习“植物生理学”课程的学习态度与成绩分析[J]. 高校生物学教学研究(电子版), 2019, 9(4): 28-33 .
- [17] ROODMAN David . Fitting fully observed recursive mixed-process models with cmp[J] . The Stata Journal , 2011 , 11(2) : 159-206 .
- [18] 顾天竹, 顾锡宏. 教育质量提高更有利于谁：男生还是女生?——来自我国初中教育调查的经验证据[J]. 教育与经济, 2018(3): 38-45, 88 .

责任编辑：黄燕妮

我刊 2019 年度优秀论文评选结果揭晓

为激励广大作者撰写学术论文的积极性，吸引更多高水准的论文，进一步提高我刊的学术水平和稿件质量，我刊决定从 2020 年起，对上一年度刊发的论文进行评选。经过栏目编辑初选、九位审稿专家复选和编辑部最终审定，我刊 2019 年度优秀论文评选揭晓，张宗毅、刘玉萍、程中培、龙子泉、金江峰和刘建的论文获评为优秀论文，每位作者将获得编辑部颁发的荣誉证书和物质奖励。

以下是详细名单：

张宗毅, 王许沁, 葛继红. 中国农机化效率：区域差异及购置补贴影响效应, 2019(3) .

刘玉萍, 郭郡郡. 农民工不同层次教育回报的异质性——基于收入分布与不同群体的比较, 2019(2) .

程中培. 城乡居民家庭基本生活需要的测量及保障路径分析, 2019(6) .

龙子泉, 张媛媛, 周玉琴, 等. 产权改革、社会资本与小型农田水利设施管护效果, 2019(4) .

金江峰. 农村征地拆迁中行政包干制的机制与效应, 2019(2) .

刘建. 贫困治理中的国家自主性及其提升路径——基于 L 乡农村精准扶贫的案例分析, 2019(6) .