

“分”至沓来：大学生选课行为中的 分数偏好

张优杨，韩沛言

[摘要]在新时代教育评价改革背景下，学生评价改革日益受到重视。本文构建了一个学生选课行为的理论模型，以探讨大学生选课中分数偏好的形成机制与影响。在目标不一致和信息不对称的框架下，借鉴委托代理问题的分析逻辑，将教师设定为教学活动的委托人，学生作为追求目标成绩的代理人。理论模型显示，教师给分越宽松，学生的最优努力水平越低，从而导致学生选择那些历史给分最好的课程。且在学生能力越强、课程主观性越强的情况下，这一效应越显著。利用某研究型大学3000余门本科课程的选课数据与历史成绩信息，验证了理论模型的核心命题。研究发现，学生选课行为存在显著道德风险特征，课程历史给分情况对选课率具有显著影响，学生更倾向于选择考核宽松、给分较高的课程；这种分数偏好在高保研率群体、高能力学生中表现更为突出，且在主观性更强的课程以及通识课中更为明显。从学生决策机制层面揭示了当前教学评价体系的激励扭曲问题，为“水课”治理、教育评价制度优化及高等教育质量提升提供了理论依据与实践参考。

[关键词]选课；分数偏好；委托代理问题；教育评价改革

DOI:10.19512/j.cnki.issn2096-2088.2026.03.004

一、引言

党的二十大报告指出，要“完善学校管理和教育评价体系”。在新时代高等教育高质量发展的背景下，高等教育评价改革已成为我国教育改革的重要任务之一（刘振天，2024）。2020年10月，中共中央、国务院印发《深化新时代教育评价改革总体方案》，明确提出要破除“唯分数、唯升学、唯文凭、唯论文、唯帽子”的不良倾向，构建科学、合理、多元的教育评价体系，以实现

[收稿日期]2025-11-18

[基金项目]清华大学“计算社会科学研究奖励计划”。

[作者简介]张优杨（通讯作者），武汉大学经济与管理学院，电子邮箱地址：zhangyouyang2003@outlook.com；韩沛言，北京师范大学国际与比较教育研究院，电子邮箱地址：202521010067@mail.bnu.edu.cn。

立德树人根本任务的全面落实。^① 该方案标志着教育评价改革进入系统设计与制度落实的新阶段,强调通过改进评价“指挥棒”引导教育实践的优化与调整。其中,学生评价是教育评价体系的重要组成部分之一。教育部在 2021 年出台的《普通高等学校本科教育教学审核评估实施方案(2021—2025 年)》^②进一步指出,应在教学评价中综合考量学生学习体验、教学过程质量与教育投入产出比。^③ 这一系列改革举措从顶层设计出发,推动我国学生评价理念由单纯关注知识掌握的结果评价,逐步转向更加重视学习过程、能力发展与综合素养的多元评价。在此背景下,深入探究教学评价体系如何影响高校学生的微观决策与行为,对于深化高等教育评价改革、提升高等教育质量具有至关重要的理论与实践意义。

另一方面,“唯分数”评价在高校中长期占据主导地位,学生的学习成绩与保研机会、奖学金等直接绑定。由此催生出的大学生群体在课程选择上所表现出的强烈“分数偏好”——即系统性倾向于选择考核标准宽松、历史给分较高的课程,而规避学术要求严格、评分尺度严谨的课程——已成为一种普遍存在的行为模式(黄亚苹等,2024)。这种集体性的策略选择,不仅催生了所谓的“水课”市场,加剧了“分数膨胀”(Grade Inflation)(哈巍和赵颖,2019;赵颖和哈巍,2020),更在深层次上扭曲了教与学的核心激励结构,使得高等教育的知识传承与创新能力培养目标面临被悬置的风险。因此,精准地揭示大学生选课行为中分数偏好的内在形成机理,并系统性地评估其对学生学习投入的衍生后果,是推动教育评价改革从宏观政策宣导迈向微观实践落实不可或缺的重要前提。

现有文献主要沿着两条路径展开:一是从教育学与社会学视角出发,剖析学生选课动机的异化;二是聚焦于“水课”现象的成因与治理策略研究。然而,多数研究集中于宏观政策与现象分析,缺乏对个体微观决策激励机制的严谨形式化分析,因而很难提出具有针对性的政策建议。具体地,如果将学生选课、上课视为一个在特定规则下的决策过程,学生与教师之间的互动可以借鉴委托—代理关系来刻画。虽然师生关系并非严格的契约关系,但其互动本质上符合该理论的两个核心前提:一是目标不一致性,教师(委托人)的

① 《中共中央 国务院印发〈深化新时代教育评价改革总体方案〉》,2020 年 10 月 13 日, http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/202010/t20201013_494381.html, 2025 年 11 月 11 日。

② 《教育部关于印发〈普通高等学校本科教育教学审核评估实施方案(2021—2025 年)〉的通知》,2021 年 2 月 3 日, http://www.moe.gov.cn/srcsite/A11/s7057/202102/t20210205_512709.html, 2025 年 11 月 11 日。

目标包含评教分数、完成教学任务等因素，而学生(代理人)则主要以成绩为驱动；二是信息不对称，教师难以完全观测并量化学生的真实努力程度和学习成效。

本文在此基础上构建了一个理论模型，假定学生成绩同时受自身努力与外生的教师给分宽松度影响，进而推导学生的最小化努力决策。模型证明，给分越宽松将导致学生实际努力水平下降，从而促使学生选择那些历史给分更好的课。在一定条件下，这一效应在能力更高、目标更高的学生中更明显；在主观性更强的课程中更为明显。在理论模型的基础上，利用我国某研究型大学超过3000门本科课程的选课数据，验证理论模型的对应命题。该结论为理解“水课”需求的成因及现有教学评价体系可能面临的激励失效问题提供了理论解释，也为优化学生评价制度提供了新的依据。本文虽聚焦于高等教育中的选课行为，但反映的是一个普遍的组织治理命题：在目标不一致和信息不对称的情况下，过于简单化的评价标准如何诱发个体的策略性行为。这种评价的异化现象在现代组织管理、公共政策实施等领域具有广泛的代表性。

二、文献综述

(一) 学生选课动机相关研究

学生的选课动机受课程质量、教师声誉、同学推荐、评分宽松度、课程工作量等多种因素影响(Wilhelm, 2004; Ma et al., 2021; 郑佳然和张煜, 2022)。然而，评分宽松度已日益成为学生选课的最主要因素。马莉萍和郑翔睿(2024)的研究发现，尽管兴趣与求知仍是多数学生的选课动机，但约三成学生在选课时更关注课程分数与课业负担，其中大二阶段是学生由“学以致用型”转向“学以求分型”的关键时期。郑伦楚等(2017)指出，考试通过率与教师态度是学生最为关注的因素之一。学生的选课行为也引起了经济学者的关注，主要是因为成绩作为学生学业表现的信号，与劳动经济学和教育经济学关注的教育回报等议题密切相关(Ostrovsky and Schwarz, 2010)。Bar等(2009)基于康奈尔大学的数据发现，在线成绩信息的提供会导致学生更倾向于选择给分宽松的课程。然而，高能力学生更少选择给分宽松的课程。Bar等(2012)则进一步构建了一个理论模型，指出公布分数信息看似透明，却会通过改变学生选课行为加剧分数膨胀、弱化成绩的能力信号价值、降低社会福利。

在我国，大学生选课动机的功利化倾向与保研制度紧密相关。研究生推免制度在拔尖创新人才选拔中存在“拔尖悖论”：学生在理性追求资源优势的

同时,其学习动机逐渐功利化,表现为由内在求知驱动转向外在绩点导向(黄亚婷和潘仁同,2025)。张洋磊和黄亚苹(2022)进一步揭示了保研竞争所引发的“学习投入内卷化”现象,指出绩点竞争在高校内部形成了以“保研核心圈”为中心的群体性行动结构,学生在这一“内圈”中自我锁定于“唯分数取向”的目标替代逻辑,以最大化绩点和资源收益为主要导向。这一“拔尖—内卷”双重逻辑导致学生在课程选择时更加倾向于低风险、高绩点的性价比课程,从而在制度激励下形成了功利化的选课偏好。

(二)“水课”、分数膨胀的形成与治理相关研究

学生选课动机的异化进一步导致了教学生态的变化,催生了“水课”市场。大学生作为有限理性行为人,在信息不对称的环境中追求预期效用的最大化,从而在选课时高度依赖课程的性价比和口碑等简化指标(罗志敏和马浚锋,2022)。学生在选课与评教过程中,往往会考虑作业简单、得分高等因素,这种理性化短权利衡可能损害长期学习收益,进一步助长“水课”存在(汪雅霜等,2019;李琳琳,2020)。以绩点为核心的“选课秘籍”和策略在学生群体中的广泛传播,强化了功利化策略,使学生过度关注表层指标,从而削弱了知识体系的系统性建构、损害了学生的学业能力发展(林小英和杨芊芊,2023;黄海燕等,2025)。

学生功利性选课的另一个后果是“分数膨胀”。分数膨胀指的是在一段时间内,学生成绩普遍上升,但学生的实际学业能力并未相应提高的现象(Kostal et al., 2016)。分数膨胀的产生受学校成绩制度、学生评教、学生选课策略等诸多因素影响(Jewell et al., 2013; 哈巍和赵颖, 2019; Stroebe, 2016; Gregor, 2021)。基于信号理论的模型说明,当企业无法区分成绩是学生能力还是学校给分宽松所致,均衡的结果是学校会主动选择分数膨胀(Chan et al., 2007)。而且,分数膨胀的发展并不均衡:评分方式的主观性越强,越容易造成分数膨胀;人文社科院系的分数膨胀程度远大于理工科院系,这显著扭曲了学生选课行为(Velazquez et al., 2025; Sabot and Wakeman Linn, 1991)。分数膨胀对高等教育的多个方面造成影响。首先,它削弱了成绩作为学生学业能力指标的可靠性。当所有学生的成绩都偏高时,雇主难以区分优秀学生和普通学生的能力差异(Pattison et al., 2013)。其次,分数膨胀可能会影响学生的学习动机。如果学生认为无论付出多少努力都能获得高分,他们可能会减少学习投入,从而影响其知识的掌握和技能的培养(Babcock, 2010)。然而,宽松评分也可能通过强化学校教育投入激励、提升整体毕业生能力,带来更高的社会福利(Boleslavsky and Cotton, 2015)。为加强“水课”治理,提升学生学习效果,现有研究主要从概念界定、教学转

向及制度改革等层面探讨，认为应通过理念更新与评价优化解决师生博弈及分数膨胀问题，最终实现向高质量教学的回归(刘进和林松月，2019；陆国栋，2018；田华文和李艳萍，2020；解德渤，2020)。

三、理论模型

无论是学校的评分策略还是学生的选课决策，其背后都涉及激励机制与信息不对称问题。为了进一步分析当下教学评价体系中学生选课的内在动机，本文借鉴委托代理理论作为分析框架。委托代理理论用于分析一方将任务委托给另一方，且代理人行为或类型无法被完全观测的各类关系(Ross, 1973; Jensen and Meckling, 1976)。该框架的核心推论为：当代理人在自身努力程度、自身特质方面拥有私人信息，且其利益与委托人存在分歧时，就可能产生道德风险与逆向选择问题。

在教育管理中，委托代理理论被用于分析国家、学校、教师，乃至学生之间多层级的委托关系(包海芹，2004)。目前，委托代理理论在高校治理中的相关研究主要聚焦政府、高校行政管理层与教师三者间的关系，几乎只被应用于宏观层面的教育政策或院校管理，还较少延伸至课程层面的互动机制、学生决策行为及师生关系研究(Yallem et al., 2018; 韩晓敏和孙元涛，2021; 李鹏虎和王传毅，2024)。事实上，在高等教育体系中存在三重委托代理关系：国家与高校之间、学校管理者与教师之间，以及教师与学生之间的关系(隋学礼等，2009)。其中，师生之间，学生是购买服务的委托人，而教师则是提供教学服务的代理人。然而，从激励结构的角度出发，师生关系也可以从反向进行理解：教师作为教学活动的设计与主导者，希望实现教育价值、完成教学任务，乃至追求评教分数；而学生作为教学任务的执行者，则在既定评价规则下追求自身效用(如绩点、学分、荣誉)的最大化(Zubrickas, 2015)。同时，教师无法完全观测学生课下的真实努力程度，而学生对自身的努力策略拥有更充分的信息。这意味着师生之间存在目标不一致与信息不对称的问题。需要明确的是，与经典委托代理理论不同，教师并不是雇佣学生来完成教学任务；且学生并非为教师的目标提供努力，而是在共同制度约束下追求自利目标；加之教师对学生是否选课没有强约束，因此师生之间并未形成严格意义上的契约委托关系。然而，由于目标不一致和信息不对称的存在，本文试图借鉴委托代理理论的基本思想和逻辑，用于分析既定制度约束下师生行为。

在此分析框架下，学生在选课时可能采取隐藏行动的“道德风险”行为。

具体表现为：在选课阶段，利用信息优势系统性偏好考核宽松、课业负担轻的“水课”；在学习过程中，则根据评价规则策略性地分配努力，以最小化投入达成目标分数，而非追求知识获取的最大化。这种基于理性计算的决策行为，正是学生选课动机异化与水课市场形成的微观机理。

(一)基本设定

考虑一个包含学生和教师的选课场景。学生的能力为 θ ，服从某个分布 $F(\theta, \bar{\theta})$ 。学生对于课程的成绩设有不同的目标 $G(\theta)$ 。假设学生能力越强，目标也越高，即 $G'(\theta) > 0$ 。

令 I 为教师集合， J 为课程集合。定义学生的选课空间为集合 $\Omega \subset I \times J$ ，其中每一个元素 $ij \in \Omega$ 代表由教师 i 授课的课程 j 。目前，中国大学的评分制度通常为两部分结构，例如，某校《本科生学籍管理实施细则(修订)》规定，“课程考核成绩根据平时成绩(包括平时测验、期中考试、作业、课堂讨论、论文、考勤等)和期末考试成绩综合评定，平时成绩占 30%—50%，期末考试成绩占 50%—70%”。但是，这种评分模式的根本问题在于很难准确观测学生的努力程度(Tippin et al., 2012)。虽然平时成绩名义上与学生表现挂钩，但是也有很强的主观性(李忠伟和赵富学，2024)。

在这种评分方式下，设单门课程 ij 的最终成绩 g_{ij} 的形式基本一致，由两部分构成：

$$g_{ij} = \alpha_j k_{ij} + (1 - \alpha_j) p_j(e, \theta) \quad (1)$$

其中，“课程主观性” α_j 是由课程本身的性质决定的，而教师 i 可以选择自己在课程 j 的“给分宽松度” k_{ij} 。^① 这两项共同决定了课程的主观成绩 $\alpha_j k_{ij}$ 。相对地， $(1 - \alpha_j) p_j(e, \theta)$ 被称为课程的客观成绩，因为它由学生的努力投入 $e \geq 0$ 决定。 $p_j(e, \theta)$ 是客观成绩函数，假设满足单调递增与边际递减，即学生越努力，成绩越高；但努力带来成绩的边际提升递减；同时，学生能力越强，达到某个分数就越容易，即 $p_e > 0$ 、 $p_{ee} < 0$ 和 $p_\theta > 0$ 、 $p_{e\theta} > 0$ 。

教师的决策是通过确定自己的给分宽松度 k_{ij} 促使学生达到某些基本的努力程度，以完成教学任务。同时，教师还会考虑评教分数、完成教学任务等因素。由于本文关心的核心问题是学生的行为，故把教师的行为视为外生给定的。

假定学生遵循以下核心决策准则：

^① 例如，“高等数学”的主观性通常会小于“英语写作”之类的课程，这意味着 α 的不同。而对于同样教高等数学的甲、乙两位教师来说，可能甲给分比较宽松，乙给分比较严格，这意味着 k 的不同。

(1)学生的决策过程包括选课和努力两个阶段。第一阶段，学生在课程空间中选择一门课 $ij \in \Omega$ ；第二阶段，学生确定自己的努力水平 e 。

(2)学生无法直接观测到教师的给分好坏。只能通过教师的历史给分信息 s_{ij} 形成一个对教师给分宽松度的推断： $\mathbb{E}(k_{ij} | s_{ij}) = f(s_{ij})$ 。如果这名教师所授课程的历史成绩较高，那么学生认为他的给分更宽松，即 $f' > 0$ 。因此，学生的预期成绩是：

$$\mathbb{E}(g_{ij} | s_{ij}) = \alpha_j f(s_{ij}) + (1 - \alpha_j) p_j(e, \theta) \quad (2)$$

(3)在任何选定的课程中，学生的首要约束是确保预期成绩不低于其目标分数 $G(\theta)$ ，即学生是目标导向型而非单纯的成绩最大化型。

(4)在满足目标分数的条件下，学生通过最小化客观努力 e 来最大化自身的间接效用 V_{ij} 。这意味着学生将努力视为一种带来负效用的成本。

(二)决策过程

用逆向归纳法求解。在第二阶段(努力阶段)，学生的目标是选择努力程度 e ，以最小化达成目标分数 $G(\theta)$ 的成本：

$$\min e \quad \text{s.t.} \quad \alpha_j f(s_{ij}) + (1 - \alpha_j) p_j(e, \theta) \geq G(\theta) \quad (3)$$

这个问题的约束条件等号应该成立，设 e^* 为满足目标成绩 $G(\theta)$ 的最小努力水平，可以解出 $e^*(s_{ij}, \theta)$ 应满足下式：

$$\alpha_j f(s_{ij}) + (1 - \alpha_j) p_j(e^*, \theta) = G(\theta)$$

对这个式子关于 s_{ij} 求全微分可得：

$$\frac{\partial e^*}{\partial s_{ij}} = - \frac{\alpha f'(s_{ij})}{(1 - \alpha) p_{j,e}} < 0 \quad (4)$$

由此可得引理1：

引理1(历史给分对学生努力的影响)：对于能力和目标分数给定的任意学生，如果学生能从这门课的历史给分中形成对主观给分 k_{ij} 合理的推断，那么历史给分越高，其最优努力程度越低。

由于在选课阶段，学生会最大化间接效用来选择课程。效用由最小努力成本决定：

$$\max_{ij \in I \times J} V_{ij} = -e^*(\mathbb{E}(k_{ij} | s_{ij}), \theta) \quad (5)$$

结合引理1，得到命题1：

命题1(学生的分数偏好)：在其他条件相同的情况下，学生将选择历史给分 s_{ij} 最高的课程。

对(4)关于 θ 进行求导，可得到：

$$\frac{\partial^2 e^*}{\partial s_{ij} \partial \alpha_j} = \frac{f'}{(1 - \alpha_j)^2 p_{j,e}} \left[1 - \frac{\alpha_j p_{j,ee} (p_j - f)}{p_{j,e}^2} \right] \quad (6)$$

则有:

$$\begin{cases} \frac{\partial^2 e^*}{\partial s_{ij} \partial \theta} > 0, \text{ if } p_{j, e\theta} + p_{j, ee} \left[\frac{G'(\theta) - (1 - \alpha_j) p_{j, \theta}}{(1 - \alpha_j) p_{j, e}} \right] > 0; \\ \frac{\partial^2 e^*}{\partial s_{ij} \partial \theta} < 0, \text{ if } p_{j, e\theta} + p_{j, ee} \left[\frac{G'(\theta) - (1 - \alpha_j) p_{j, \theta}}{(1 - \alpha_j) p_{j, e}} \right] < 0 \end{cases} \quad (7)$$

由此得到命题 2:

命题 2(学生能力对努力下降的调节效应): 当历史给分更高时, 可能出现两种情形: (1) 能力越强的学生, 最优努力受历史给分的影响越小; (2) 能力越强的学生, 最优努力受历史给分的影响越大。

对(4)关于 α 进行求导, 可得到:

$$\frac{\partial^2 e^*}{\partial s_{ij} \partial \alpha_j} = - \frac{f'}{(1 - \alpha_j)^2 p_{j, e}} \left[1 - \frac{\alpha_j p_{j, ee} (p_j - f)}{p_{j, e}^2} \right] \quad (8)$$

则有:

$$\begin{cases} \frac{\partial^2 e^*}{\partial k_i \partial \alpha_j} > 0, \text{ if } \frac{\alpha_j p_{j, ee} (p_j - f)}{p_{j, e}^2} > 1; \\ \frac{\partial^2 e^*}{\partial k_i \partial \alpha_j} < 0, \text{ if } \frac{\alpha_j p_{j, ee} (p_j - f)}{p_{j, e}^2} < 1. \end{cases} \quad (9)$$

由此得到命题 3:

命题 3(课程主观性对努力下降的调节效应): 当历史给分更高时, 可能出现两种情形: (1) 课程主观性 α 越高, 最优努力的下降越小; (2) 课程主观性 α_j 越高, 最优努力的下降越大。

只有当努力的边际回报递减到几乎为零, 且课程几乎全是主观分时, 情形 1 才会成立。这时再增加主观权重, 对努力的进一步挤出空间反而变小。我们忽略这种情况。可以说明, 命题 2 的情形 2 和命题 3 的情形 2 可以同时成立, 这通常意味着 G' 足够大; 比起能力, 努力是获取高分更重要的因素。在这种条件下, 高能力学生目标很高, 原本需要非常努力才能达到目标分数。但由于努力回报递减显著, 使得他们努力的边际成本很高, 通常高于低能力学生。因此当主观分数增加时, 他们可以减少大量高成本的边际努力, 努力下降幅度比低能力学生更大。

四、数据来源与研究设计

(一) 数据来源

本文通过实证检验上述命题。使用的数据来自某研究型大学。通过该校教务系统等平台, 获取了该校 2021—2022 学年共三个学期 3000 余门本科课程的开课情况, 以及这些课程 2013—2019 年的历史成绩信息。

作为示意，图1展示了该校2013—2019年以六门经济与管理学院学生为授课对象的大类公共课程的成绩数据，共开课413次。选择经济与管理学院的理由是，该院是全校规模最大的学院，同一课程开设教师数量多，有利于比较不同教师的给分差异。选择这些课程的理由是，学生均为经济与管理学院经济学大类大一大二学生，学生能力得到了较好的控制。

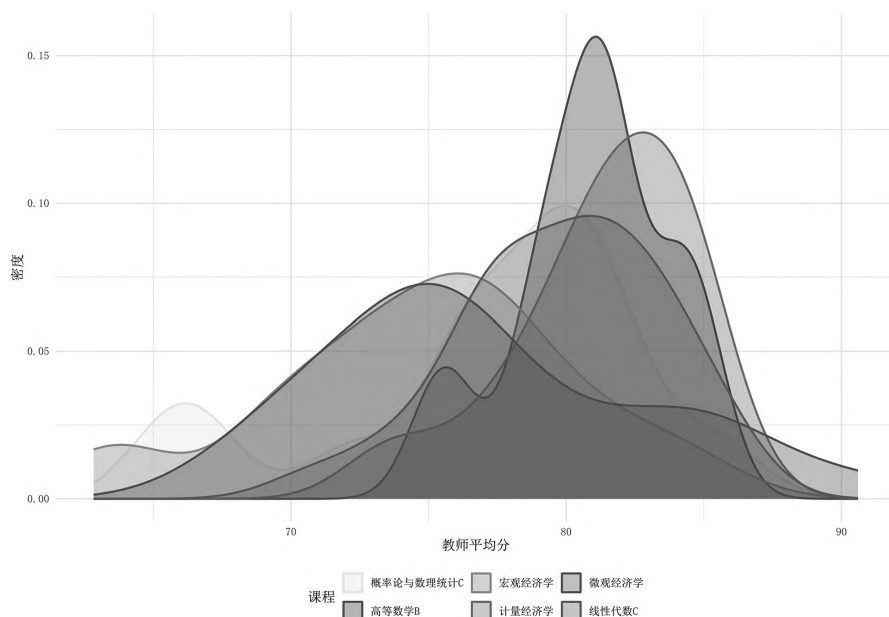


图1 不同教师均分的概率分布

图1绘制了这六门课程不同教师所教班级平均成绩的概率密度图。如果教师的给分能够完整准确反映学生努力，那么可以预测，各课程成绩应该集中于各自的均值处呈现单峰分布。然而不难看出，课程给分的离散程度均较大。其中，评分客观性更强的数学类课程的平均成绩分布相对经济学课程更加集中。对于考试方式、教学内容、学生水平等因素均类似的课程，成绩分布却较为离散，这可能是由于评分机制并不能完全反映学生能力，而是受其他因素干扰。这是由前文所述的两部分评分结构所致：一部分是期末卷面成绩，另一部分是平时作业、课堂表现等。后者往往被教师用于“捞人”；在美国，教师则通常采用“curve”(曲线评分)等相对评分而非绝对评分方法来调整原始成绩(Bowen and Cooper, 2022)。不同教师给分之间存在的这种系统性差异，使学生直接体会到给分的宽、严差异，从而为学生的策略行为提供了基础的制度空间。

(二)变量选择

1. 被解释变量

被解释变量是课程的选课率，根据实际选课人数/课程容量计算。课程容

量通常根据各专业总人数确定,因此选课率这一变量实际上能够较为准确地表示课程的受欢迎程度。有时选课人数会超出课程容量,导致选课率大于1。

2. 解释变量

解释变量是课程的给分情况,主要包含三个指标,课程的平均成绩、满绩率(90分以上的比例)和不及格率(60分以下的比例),均为2013—2019年的平均值。

3. 控制变量

控制变量包括三个层次。第一个层次是学生特征,包括学生所属学科(以信息科学专业为基准组)、^①选课时学生在校时长(学期)、学生所在学院的保研率、学院总人数(取对数),以及2021年在该校所在省份的理科录取分数线。其中,2021年理科录取分数线用于代理学生的能力。首先,高考录取分数能够较好地预测学生的学业表现(Bai et al., 2014)。其次,除了通识选修课以外,其他课程类型的选课群体同质性比较强,即基本上来自相同学院的相同年级。由于该校实行基于学院的大类招生,因此学院层面的高考分数线反映了不同学院学生的相对能力。此外,选取理科录取分数线是因为文理科的分数线不具备可比性,而绝大多数专业都招收理科生,样本覆盖更广。

第二个层次是教师特征。包括教师性别、教师职称和入选国家级人才项目次数。其中,教师性别为一个虚拟变量,取1代表男性。教师职称为一个排序变量,从1到4分别代表初级职称(助教、技士等),中级职称(讲师、工程师、实验师等),副高级职称(副教授、副研究员、高级工程师等)和正高级职称(教授、研究员、主任医师等)。自2015—2016学年起,该校设立了“本科优秀教学业绩奖”评选,^②这一奖项每年评选一次,以激发教师的教学积极性、提高教学质量。我们用教师是否获得这一奖项来控制教学的质量;入选国家级人才项目的次数是指入选两院院士、人文社科资深教授、国家杰出青年科学基金项目、长江学者奖励计划和国家高层次人才特殊支持计划(万人计

① 学生专业按所属学院划分,该校学生共来自39个不同的学院。其中,人文专业的学科代码包含哲学(01)、文学(05)、历史学(06)和艺术学(13),社科专业的学科代码包含经济学(02)、法学(03)、教育学(04)和管理学(12),理科专业的学科代码为理学(07),信息科学专业包括电子信息(0807)、计算机(0809)、测绘(0812),工科专业的学科代码为信息专业之外的工学(08)、医科专业的学科代码为医学(10)。该校不含农学、军事学专业。

② 该奖的评选方式是个人申报加单位推荐。评选依据主要包括职业道德、教学工作量和教学质量三方面。其中,职业道德方面要求“近三年无教学事故、无学术失范行为”;教学工作量要求“独立主讲本科课程或指导学生学科竞赛,工作量饱满”;教学质量的评价方式由“学生评价、专家评价和单位评价等方式相结合”。由于学生评价在教师给分之前,一定程度上规避了教师为赢得更高评教分数而可以放水的现象。

划)等人才项目的次数之和。

第三个层次是课程特征。课程类别包括五个虚拟变量(以通识教育选修为基准):专业教育必修、专业教育选修、公共基础必修、公共基础选修和通识教育必修。课程难度根据课程代码的后三位计算得到。根据课程代码编制规则,前若干位数字代表开课学院、课程性质等信息,而最后三位数字越大,代表课程越高级。我们根据前若干位数字进行分组,然后将最后三位数字做标准化处理为一个 0 到 1 的难度指标。此外,我们还用文本分析法计算了另一个课程难度的指标,作为稳健性检验。为了更精准地刻画教师在评分过程中的裁量空间,本文基于教学大纲(Syllabus)的信息,构建了“课程主观性”综合指标。具体维度如表 1 所示。

表 1 课程主观性指标构建

评分维度	0 分(低主观性)	1 分(高主观性)
学科属性	自然科学、工学、医学、逻辑学类	人文、社科类
考核形式	闭卷考试、标准化试题	论文、报告、小组作业
平时成绩权重	低于 40%,或仅含考勤	大于等于 40%

根据上述构建规则,我们得到了一个 0—3 的主观性指标。其中,主观性最强的课程得分为 3,主观性最低的课程得分为 0。此外控制变量还包括该课程的开课学期,学分数,上课时间(是否为早课、周几上课),课程内容(是否含上机课、是否含实践课、是否含实验课)。核心变量的描述性统计报告在表 2。因篇幅所限,其他变量的描述性统计结果未在正文中报告,如需要可向作者索取。

表 2 核心变量的描述性统计

变量		观测数	均值	标准差	最小值	最大值
被解释变量	选课率	4137	0.822	0.257	0.0125	3.076
	平均成绩	4137	84.588	5.928	41.29	97.92
核心解释变量	不及格率	4119	2.930	5.131	0	57.14
	满绩率	4120	40.846	25.244	0	100

(三)模型设定

本文使用的数据是“教师—课程”层面的截面数据,设定如下基准回归模型:

$$enroll_rate_{ij} = \alpha + \beta_1 score_{ij} + \Gamma X_{ij} + \epsilon_{ij}$$

(10)

其中,下标 i 代表教师; j 代表课程。被解释变量 $enroll_rate_{ij}$ 为 i 教师

开设 j 课程的选课率；核心解释变量 $score_{ij}$ 为该教师在该课程上的往届成绩情况； X_{ij} 是一组控制变量，包括学生特征、教师特征和课程特征； ϵ_{ij} 是随机扰动项。

由于数据中存在大量的某位教师同时开设多门课程，或者某门课程由多位教师分别开设的样本，这种交叠的结构使我们可以基于模型(10)的基础上进一步纳入教师固定效应和课程固定效应，以控制那些不随课程变化的教师个体异质性(如教学风格、教师声誉等)和不随教师变化的课程固有属性(如课程重要性、课程难度等)，以缓解遗漏变量问题。如模型(11)所示：

$$enroll_rate_{ij} = \beta_1 score_{ij} + \Gamma X'_{ij} + \mu_i + \gamma_j + \epsilon_{ij} \quad (11)$$

其中， μ_i 是教师固定效应， γ_j 是课程固定效应。通过该设定，能够更准确地利用教师在不同课程间或同一课程在不同教师间表现的差异，识别出往届成绩对选课率的净影响。

五、实证分析结果

(一)基准回归

表 3 汇报了考虑各类型课程、以选课率为因变量的回归结果，估计了往届给分情况对学生选课行为的影响。为保证选课率是学生自发选课的结果，我们剔除了那些只有一位老师开设，并且是必修课的课程。例(1)和(2)通过逐步纳入课程特征、教师特征、学生群体特征等控制变量，确保了核心解释变量系数估计的稳健性。回归结果清晰地表明，往届的给分情况是学生选课决策的一个显著影响因素。在仅包含往届平均成绩的列(1)中，该变量在 1% 的显著性水平上呈现出正向影响。随着控制变量的加入，这一影响仍然稳健。在列(3)和(4)中，我们分别纳入课程固定效应和教师固定效应。列(5)同时纳入了课程固定效应和教师固定效应，控制了所有不随教师/课程变化的因素后，结果依然保持高度稳健，且具有显著的经济意义：往届平均成绩每提高 10 分，选课率将显著提升 14 个百分点。这有力地支持了命题 1 中学生在选课过程中存在的分数偏好。

需要注意的是，这一结果并不意味着学生完全不在乎授课质量。尽管很难直接衡量课程的质量，但在控制了足够多控制变量的基础上，是否获得本科优秀教学业绩奖能够大致代理课程的质量和教师声誉等因素，如列(2)和(3)所示，其系数都显著为正。

表 3 基准回归结果

	选课率	选课率	选课率	选课率	选课率
	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	OLS+ 课程固定效应	OLS+ 教师固定效应	OLS+ 双向固定效应
往届平均成绩	0.004*** (0.001)	0.007*** (0.001)	0.013*** (0.002)	0.004*** (0.002)	0.014*** (0.005)
获得教学业绩奖		0.023** (0.010)	0.040*** (0.014)		
课程特征		YES		YES	
其他教师特征		YES	YES		
学生特征		YES	YES	YES	YES
课程固定效应			YES		YES
教师固定效应				YES	YES
R-squared	0.007	0.185	0.543	0.605	0.664
N	3711	2695	2119	1934	1714

注：第(1)列和第(2)列括号内为异方差稳健标准误；固定效应模型的括号内为聚类稳健标准误，聚类在教师层面。***、**、* 分别表示在 1%、5% 和 10% 的水平上通过显著性检验。下表同。

(二)稳健性检验

为确保基准回归结果的可靠性，本文进行了一系列稳健性检验，结果汇报于表 4。检验表明，往届平均成绩对选课行为的正向影响在不同设定下均保持稳健，证实了核心结论的可靠性。首先，列(1)进一步将给分结构区分为满绩率和不及格率，发现学生不仅追求高分，还会规避挂科风险。具体而言，往届满绩率对选课率有显著的正向拉动作用，而往届不及格率则带来更大幅度的负面影响。

其次，我们替换了被解释变量的度量方式与估计模型。在列(2)中，考虑到选课率本质上源于实际的选课人数，我们将因变量从“选课率”替换为“选课人数”，并使用 OLS 进行估计。往届平均成绩的系数在 1%水平上依然显著为正，表明给分越高，绝对选课人数也越多。考虑到选课人数是计数变量，列(3)进一步采用负二项回归模型进行估计，结果同样稳健。列(4)则以“课程是否选满”这一虚拟变量作为被解释变量，用 Probit 模型进行估计。结果显示往届平均成绩越高，该课程选满的概率显著提升。列(5)使用了分数 Logit 模型，核心解释变量的显著性和正向关系仍未改变。这些不同模型与变量测度

的检验共同表明，基准结果并非由特定的函数形式或变量构建方式所驱动。

此外，我们调整样本范围来检验结果的稳健性。列(6)剔除了往届平均成绩最高和最低的 5% 极端样本，防止结果被异常值所扭曲，核心系数的显著性与方向均未发生变化。列(7)将样本限制在选修课范围内，结果仍稳健。

表 4 稳健性检验

	替换解释变量	替换被解释变量			更换模型	剔除极端值	仅选修课
	选课率	选课人数	选课人数	是否选满	选课率	选课率	选课率
	OLS	OLS	NBR	Probit	FLR	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
往届平均成绩		0.379*** (0.051)	0.006*** (0.001)	0.028*** (0.005)	0.039*** (0.005)	0.004*** (0.001)	0.011*** (0.002)
满绩率	0.001*** (0.000)						
不及格率	-0.003*** (0.001)						
控制变量	YES	YES	YES	YES	YES	YES	YES
R-squared	0.182	0.846				0.141	0.177
N	2695	2695	2695	2694	2378	2403	964

注：第(2)至第(4)列还控制了课程容量。

(三)调节效应分析

在基准回归的基础上，表 5 的列(1)加入了同一课程其他教师平均给分这一变量，发现其他教师平均给分越高，这门课程的选课率越低。说明学生会在不同课程之间进行权衡，从给分低的课程流向给分高的课程，这验证了理论模型关于学生在课程间替换的基本假设。进一步引入交互项，以探究学生群体特征与课程属性之间是否存在协同效应，从而更深入地理解选课决策的复杂机制。表 5 第(2)列至第(5)列汇报了引入交互项的回归结果。

表 5 学生能力与目标的调节效应

	选课率	选课率	选课率	选课率	选课率
	(1)	(2)	(3)	(4)	(5)
往届平均成绩	0.009*** (0.001)	0.007*** (0.001)	-0.107**① (0.051)	0.003* (0.002)	0.006*** (0.001)

① 由于交互项的引入，主效应系数代表了调节变量取值为 0 时的影响。由于录取分数线远远大于 0，这一系数并不具实际的经济学解释价值。

	续表				
	选课率	选课率	选课率	选课率	选课率
	(1)	(2)	(3)	(4)	(5)
保研率	−0.077** (0.033)	−0.020 (0.039)	−0.073** (0.032)	−1.124*** (0.352)	−0.065** (0.032)
课程难度	−0.028** (0.013)	0.019 (0.027)	−0.023* (0.013)	−0.025** (0.012)	−0.034*** (0.013)
录取分数线	−0.210*** (0.058)	−0.154*** (0.054)	−1.649** (0.672)	−0.174*** (0.055)	−0.135** (0.055)
同一课程其他 教师平均给分	−0.003** (0.001)				
保研率× 课程难度		−0.141* (0.080)			
录取分数线× 平均成绩			0.018** (0.008)		
保研率× 平均成绩				0.013*** (0.004)	
课程主观性× 平均成绩					0.004*** (0.001)
控制变量	YES	YES	YES	YES	YES
R-squared	0.179	0.186	0.186	0.187	0.194
N	2370	2695	2695	2695	2673

列(2)引入了保研率与课程难度的交互项。结果显示，该交互项的系数在10%的显著性水平上为负。这表明，保研率的提高可能并没有缓解竞争压力，反而使之加剧了(陈涛等，2024)。对于保研率低的学院，很多学生很早就转向就业、考研、出国等出路；而保研率的升高，使得原先处于边缘位置的学生参加到保研竞争中，促使他们更显著地规避高难度课程。对于志在保研的学生而言，维持极高的平均绩点是其核心目标，因此他们对可能导致绩点下滑的高风险(高难度)课程尤为敏感，“求稳”动机超过了通过挑战难课来展示能力的动机。列(3)考察了录取分数线与往届平均成绩的交互作用，交互项系数显著为正。这意味着，学生的能力放大了往届成绩的正向作用。录取分数更高的学生(可视为能力更强的群体)在选课时，比那些能力较低的学生更加注重给分的好坏。这与理论模型的命题2的情形2相一致。这通常是因为他

们的目标分数更高，而此时努力的边际收益已经很低，从而陷入内卷的困境。列(4)说明保研率同样放大了往届平均成绩的正向作用。这一发现与列(2)相互印证，反映了越是追求保研的学生，其在选课决策中对“往届平均成绩”这一变量的依赖程度就越高，其选课策略高度服务于保研这一终极目标。列(5)则引入了课程主观性和往届平均成绩的交互项，其系数显著为正。说明对于主观性更强的那些课程，学生更加看重往届平均成绩。这验证了命题 3 的结论。这通常是由于对于考核标准较为抽象、主观性较强的课程(如人文社科类课程、以论文为考核依据的课程)，师生之间存在更高的信息不对称性，从而学生会更加依赖往届平均成绩以降低对评分公正性和不确定性的担忧。

(四)异质性分析

表 6 汇报了异质性分析的结果。首先，从学科来看，无论是理工科、医科还是人文社科，往届平均成绩都对选课率有显著的正向影响。其次，从年级分布来看，对成绩的追求贯穿了本科学习的全过程。需要说明的是，考虑到普通四年制专业在大四学年以毕业实习和毕业论文为主，基本不设常规课程，第(7)列的样本仅包含长学制(建筑学、医学五年制和八年制等)专业学生。回归系数依然显著为正，且效应与大三学生接近。最后，从课程类别来看，无论是通识课还是专业课，给分好坏都是影响选课的重要因素。点估计显示，通识课样本中的回归系数大于专业课样本。为进一步验证这一系数差异的统计显著性，本文进行了 Suest 检验。检验结果拒绝了两个样本系数相等的原假设($\chi^2=10.25$, $p=0.001$)。这说明，学生对通识课的给分好坏更为敏感。这主要是因为通识课选择的自由度更高、替代性更强，且通识课被认为是刷分的策略性工具(吕林海和汪霞，2012)，选择的主要标准表现出更强的分数偏好。

表 6 异质性分析结果

	分专业			分年级				分课程	
	理工科	医科	人文社科	大一	大二	大三	大四	通识课	专业课
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
往届平均成绩	0.006*** (0.001)	0.008*** (0.002)	0.009*** (0.002)	0.009*** (0.002)	0.005*** (0.001)	0.008*** (0.001)	0.009*** (0.003)	0.016*** (0.004)	0.005*** (0.001)
控制变量	YES	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.205	0.120	0.180	0.227	0.177	0.183	0.551	0.383	0.179
N	1983	400	773	486	944	1003	190	424	2643

六、结论与政策建议

(一)主要结论

本文借鉴委托代理理论的核心逻辑,通过构建一个旨在捕捉激励结构失衡的学生选课决策模型,为理解大学生选课行为提供了一个全新的微观分析框架。在此基础上,利用某研究型大学的微观选课数据进行实证检验,验证了理论模型的相关命题,揭示了教学评价体系中激励相容机制失效的根本原因。其一,在以成绩为核心的评价体系下,学生(代理人)与教师(委托人)之间存在明显的目标偏离。其二,教师给分宽松度的差异性,提供了学生向给分宽松课程转移的条件。这种制度性错配不仅造成了“水课”现象的持续存在,也使高等教育的价值导向出现偏移。主要结论如下。

第一,基于委托代理框架的模型分析发现,现有评分制度存在明显的激励扭曲。教师希望学生通过努力完成学习任务,然而这一实际努力很难被准确评估,导致了“主观成绩”的存在。在信息不对称情境中,教师的给分宽松度改变了学生的努力决策。模型结果表明:较高的给分宽松度会降低学生实现目标分数所需的最小努力水平(引理1),从而吸引学生向“水课”转移(命题1)。且这一负向效应在学生能力较高时(命题2)和课程主观性更强时(命题3)更为显著。这是因为高能力学生的目标分数更高,且分数的边际增长递减。这一结果从理论上揭示了教学评价体系中激励失衡的机制,为理解“水课”现象提供了制度层面的解释。

第二,实证证据显示,学生的选课行为呈现出典型的代理人道德风险特征。在契约形成前的课程选择阶段,学生作为信息优势方,倾向于选择考核要求宽松、给分较高、不及格率低的课程,形成对“水课”的需求(命题1)。

第三,异质性分析进一步揭示了不同学生群体在选课行为上的结构性差异。学生对课程给分的敏感性在保研竞争激烈、高能力群体中更为突出(命题2),且在主观性较强的课程中更为明显(命题3),印证了信息不对称的关键作用。

事实上,本文是古德哈特定律(Goodhart's Law)的一个例证。由于评价体系的局限性,当复杂的生产活动(如教与学、科研创新、公共服务)被简化为单一的量化指标(如绩点、论文数、KPI)时,代理人(学生、员工、基层官员)有很大可能会产生策略性行为(如选修给分宽松的课、追求短期绩效),从而损害这个指标的有效性(Goodhart, 1984)。本质上,其根源于评价制度提供了错误的激励信号。而这种策略性行为会对代理人本身的长期发展,乃至社会的整体福利造成何种影响,讨论还不够充分(Courty and Marschke,

2004)。例如，那些投机于给分宽松课程，而非真正学习知识的学生，从长期来看是否会发展得更好？这需要未来研究进一步关注。

(二)政策建议

首先，优化激励机制设计，降低评分的主观性。制度的清晰与稳定是规范委托代理双方行为、降低代理成本的关键。首先，优化成绩构成结构，明确并限制主观性评分(如模糊的“课堂表现”分)在总成绩中的上限比例，提高基于客观标准考试、标准化作业与项目成果的评价权重。其次，建立结构化的平时成绩记录与公示制度，使学生的努力水平和学习效果能够被更准确地识别与量化。具体而言，可将评价依据分解为课堂发言、个人汇报、小组展示等多个可观测、可记录的学习环节，并明确公布各环节的评分标准。这些举措有助于压缩教师主观评分的操作空间，降低成绩评定中的信息不对称与随意性，使最终成绩更能真实反映学生的努力与能力，从而削弱学生通过选择“高主观分”课程进行制度套利的动机。

其次，推行分类评价机制，引导多元学习目标。本文的分析发现，高能力学生和处于高保研竞争压力下的学生，其选课行为对历史给分更为敏感。这反映出“一刀切”的百分制评分体系难以适应学生多元化的能力与发展需求。高校应积极探索并实施分类评价机制(刘云杉，2018；Bizzotto and Vigier, 2025)。例如，为有志于学术研究的拔尖学生设立“荣誉课程”或“挑战性项目”，并在评价中引入基于学术潜力和创新能力的相对评价指标，以满足其高阶发展需求，避免其在常规课程中因目标分数过高而陷入过度努力的陷阱。

最后，强化制度约束与信息供给，治理“水课”生态。在课程建设中，应推动评价标准的系统化与规范化，压缩学生通过策略性选课进行制度套利的空间。具体而言，对于由多位教师共同开设的同一课程，可建立标准化的课程大纲与评价标准，由课程组统一考试和评分，从而削弱因教师个人评分尺度差异所引发的“逆向选择”现象，引导学生将选课动机回归至课程内容与教学质量本身。在信息供给层面，应优化课程信息平台，公布课程的教学目标、培养要求、阅读清单及作业量等实质性信息，并鼓励基于学习体验的质性课程评价，从而帮助学生从“追求高分”转向“追求高价值”，从根本上重塑健康的选课文化与学习生态。

[参考文献]

- 包海芹, 2004:《教育政策执行中的委托代理问题》,《江苏高教》第3期。
- 陈涛、李点石、鞠沁汝, 2024:《谁更容易保研——保研经济的生成逻辑与运行机制》,《湖南师范大学教育科学学报》第1期。
- 哈巍、赵颖, 2019:《教学相“涨”:高校学生成绩和评教分数双重膨胀研究》,《社会学研究》第1期。
- 韩晓敏、孙元涛, 2021:《我国减负政策执行的现实困境及治理策略——基于委托代理理论的视角》,《当代教育科学》第4期。
- 黄海燕、张永婷、沈秋怡, 2025:《超越功利主义:大学生选课行为的异化与矫正——基于“选课秘籍”传播与使用行为的定性研究》,《南京晓庄学院学报》第3期。
- 黄亚苹、张洋磊、王曦影, 2024:《学分绩点制下学优生的策略主义行动及生成逻辑》,《江苏高教》第6期。
- 黄亚婷、潘仁同, 2025:《研究生推免制度的“拔尖”悖论——一项基于整合性学生发展分析框架的质性研究》,《江苏高教》第2期。
- 解德渤, 2020:《如何打造“金课体系”:大学课程改革的框架设计与制度创新》,《四川师范大学学报(社会科学版)》第1期。
- 李琳琳, 2020:《大学生选课与评教行为理性与非理性倾向》,《江苏高教》第3期。
- 李鹏虎、王传毅, 2024:《委托代理理论视域下高校自设交叉学科政策执行研究:偏差、成因及治理机制》,《清华大学教育研究》第2期。
- 李忠伟、赵富学, 2024:《高校公共体育分数膨胀现象的生成与消解》,《体育学刊》第6期。
- 林小英、杨芊芊, 2023:《过度的自我监控:评价制度对拔尖创新人才培养的影响》,《全球教育展望》第4期。
- 刘进、林松月, 2019:《什么是大学“水课”:概念范畴、关键特征与治理方向》,《黑龙江高教研究》第9期。
- 刘云杉, 2018:《拔尖与创新:精英成长的张力》,《清华大学教育研究》第6期。
- 刘振天, 2024:《高等教育评价改革的主要成效、基本经验与完善之策》,《中国高等教育》第Z2期。
- 陆国栋, 2018:《治理“水课” 打造“金课”》,《中国大学教学》第9期。
- 吕林海、汪霞, 2012:《我国研究型大学通识课程实施的学生满意度调研》,《江苏高教》第3期。
- 罗志敏、马浚锋, 2022:《在预期与现实之间:大学生“口碑”背后的选课行为逻辑与动力机制》,《杭州师范大学学报(社会科学版)》第2期。
- 马莉萍、郑翔睿, 2024:《知识诚可贵还是分数价更高——精英大学本科生选课动机研究》,《北京大学教育评论》第4期。
- 隋学礼、周裕、李江城, 2009:《高校委托代理关系中教师与学生博弈模型》,《财经科学》

第 6 期。

田华文、李艳萍, 2020:《博弈视角下高校“水课”现象的产生机理及治理对策》,《山东科技大学学报(社会科学版)》第 4 期。

汪雅霜、郝龙飞、钱蕾, 2019:《谁往课程里面掺了“水”——高校“水课”形成影响因素的质性研究》,《重庆高教研究》第 4 期。

张洋磊、黄亚苹, 2022:《“内圈”竞争:本科生学习投入内卷化表征及成因——基于学生评价视角》,《高等工程教育研究》第 4 期。

赵颖、哈巍, 2020:《分数膨胀的影响因素与学生选课策略》,《清华大学教育研究》第 6 期。

郑佳然、张煜, 2022:《大学生选课动机及在线学习行为机制研究》,《中国青年社会科学》第 5 期。

郑伦楚、王金华、郑旭东, 2017:《大学生选课动机的调查研究与改进策略建议——以华中师范大学为例》,《中国大学教学》第 3 期。

Babcock, P., 2010, “Real Costs of Nominal Grade Inflation? New Evidence from Student Course Evaluations”, *Economic Inquiry*, 48(4): 983—996.

Bai, C. E., W. Chi and X. Qian, 2014, “Do College Entrance Examination Scores Predict Undergraduate GPAs? A Tale of Two Universities”, *China Economic Review*, 30(9): 632—647.

Bar, T., V. Kadiyali and A. Zussman, 2009, “Grade Information and Grade Inflation: The Cornell Experiment”, *Journal of Economic Perspectives*, 23(3): 93—108.

Bar, T., V. Kadiyali and A. Zussman, 2012, “Putting Grades in Context”, *Journal of Labor Economics*, 30(2): 445—478.

Bizzotto, J. and A. Vigier, 2025, “Sorting and Grading”, *Journal of the European Economic Association*, 23(2): 682—704.

Boleslavsky, R. and C. Cotton, 2015, “Grading Standards and Education Quality”, *American Economic Journal: Microeconomics*, 7(2): 248—279.

Bowen, R. S. and M. M. Cooper, 2022, “Grading on a Curve as a Systemic Issue of Equity in Chemistry Education”, *Journal of Chemical Education*, 99(1): 185—194.

Chan, W., L. Hao and W. Suen, 2007, “A Signaling Theory of Grade Inflation”, *International Economic Review*, 48(3): 1065—1090.

Courty, P. and G. Marschke, 2004, An Empirical Investigation of Gaming Responses to Explicit Performance Incentives. *Journal of Labor Economics*, 22(1): 23—56.

Goodhart, C. A. E., 1984, *Monetary Theory and Practice*, London: Palgrave.

Gregor, M., 2021, “Electives Shopping, Grading Policies and Grading Competition”, *Economica*, 88(350): 364—398.

Jensen, M. C. and W. H. Meckling, 1976, “Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure”, *Journal of Financial Economics*, 3(4):

305—360.

- Jewell, R. T. , M. A. McPherson and M. A. Tieslau, 2013, “Whose Fault Is It? Assigning Blame for Grade Inflation in Higher Education”, *Applied Economics*, 45 (9): 1185—1200.
- Kostal, J. W. , N. R. Kuncel and P. R. Sackett, 2016, “Grade Inflation Marches On: Grade Increases from the 1990s to 2000s”, *Educational Measurement: Issues and Practice*, 35(1): 11—20.
- Ma, B. , M. Lu, Y. Taniguchi and S. Konomi, 2021, “Investigating Course Choice Motivations in University Environments”, *Smart Learning Environments*, 8(1): 31.
- Ostrovsky, M. and M. Schwarz, 2010, “Information Disclosure and Unraveling in Matching Markets.” *American Economic Journal: Microeconomics*, 2(2): 34—63.
- Pattison, E. , E. Grodsky and C. Muller, 2013, “Is the Sky Falling? Grade Inflation and the Signaling Power of Grades”, *Educational Researcher*, 42(5): 259—265.
- Ross, S. A. , 1973, “The Economic Theory of Agency: The Principal’s Problem”, *American Economic Review*, 63(2), 134—139.
- Sabot, R. and J. Wakeman-Linn, 1991, “Grade Inflation and Course Choice”, *Journal of Economic Perspectives*, 5(1): 159—170.
- Stroebe, W. , 2016, “Why Good Teaching Evaluations May Reward Bad Teaching: On Grade Inflation and Other Unintended Consequences of Student Evaluations”, *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 11(6): 800—816.
- Tippin, G. K. , K. D. Lafreniere and S. Page, 2012, “Student Perception of Academic Grading: Personality, Academic Orientation, and Effort”, *Active Learning in Higher Education*, 13(1): 51—61.
- Velazquez, L. , B. Atenas, N. Cruz Hernández, J. C. Castro Palacio and J. A. Monsoriu, 2025, “Diagnosing Grade Inflation: A Curriculum Analytics Approach to Quality Assurance in Higher Education”, *Studies in Higher Education*, 49(3): 1—30.
- Wilhelm, W. B. , 2004, “The Relative Influence of Published Teaching Evaluations and Other Instructor Attributes on Course Choice”, *Journal of Marketing Education*, 26(1): 17—30.
- Yallow, A. , H. Juusola, I. Ahmad and S. Törmälä, 2018, “Exploring Principal-Agent Theory in Higher Education Research”, *Journal of Research and Innovation in Higher Education*, 3(1): 78—98.
- Zubrickas, R. , 2015, “Optimal Grading”, *International Economic Review*, 56 (3): 751—776.

Chasing the Grade: Grade Preference in University Course Selection

ZHANG You-yang¹, HAN Pei-yan²

(1. Economics and Management School, Wuhan University;

2. Institute of International and Comparative Education, Beijing Normal University)

Abstract: Against the backdrop of the New Era's educational evaluation reform, student assessment reform has garnered increasing attention. This paper constructs a theoretical model of students' course selection behavior to explore the formation mechanism and impact of grade preferences in college students' course selection. Within a framework of goal inconsistency and information asymmetry, the model draws on the logic of Principal-Agent theory, designating teachers as the principals of teaching activities and students as agents pursuing their target grades. The model demonstrates that the more lenient a teacher's grading, the lower the student's optimal effort level, leading students to select courses with a history of high marks. This effect is more pronounced when students possess higher abilities or when the course content is more subjective. Using course selection data and historical grade information from over 3,000 undergraduate courses at a research university, the core propositions of the theoretical model are empirically validated. The findings indicate that students' course selection behavior exhibits significant moral hazard characteristics. Historical grading trends in courses, such as average grades, the proportion of top grades, and failure rates, significantly influence course selection rates. Students tend to prefer courses with lenient assessment criteria and higher grades. This grade preference is more pronounced among high-achieving students with higher rates of postgraduate recommendations and is particularly evident in humanities and social sciences courses, general education courses, and elective courses. This study uncovers incentive distortions in the current teaching evaluation system from the perspective of student decision-making mechanisms, providing a theoretical foundation and practical insights for addressing the issue of "grade-inflated courses", optimizing educational evaluation systems, and enhancing the quality of higher education. It also enriches the application scenarios of the principal-agent theory in the field of higher education.

Keywords: course selection; grade preference; the principal-agent problem; reform of education evaluation

(责任编辑: 刘泽云 责任校对: 胡咏梅)