高等教育学业表现同群效应研究

——基于校园卡大数据

邱磊菊 李 寅 段月姣*

摘 要：本文使用校园卡大数据构建社交网络，创新地识别出具有“社会乘数”效用的大学生学业表现同群效应。研究证实，社交频率越高、关系强度越大，正向的同群效应越显著，但社交关系维持的时长则没有显著差异影响。进一步结合中国高等教育学会学情调查分析发现，外在学习投入和内在情感投入是同群效应的主要作用途径。本文基于规模和硬件升级等外延式投入成本不变的条件，探索高等教育可持续内涵式发展的推进策略。

关键词：大学生学业表现；同群效应；校园卡大数据
DOI：10.13821/j.cnki.ceq.2025.04.03

一、引 言

如何提升教育投入质效是经济转型阶段亟待解决的重要问题。社会交互所产生的同群效应（peer effect），因具有“社会乘数”（social multiplier）作用，可有效放大低成本环境下的政策及措施实施效果，因而在高等教育管理提质中可能发挥节约资源投入的重要作用。同群效应指的是，同一群体中个体的行为和表现会受到其他成员行为和表现的内生影响，进而产生具有政策放大作用的乘数效应。在高等教育管理中，如果能够促进学生积极互动，高效发挥同群效应，则相对降低了学业表现对教育资源的过度依赖，强化了教育成效。当前，同群效应的研究虽较为广泛，但是仍存在群体界定与社会互动识别两大难题。所以，本文为识别社交互动创新地引入了校园卡大数据，从而将群体的界定延展到了具有时间和空间两重维度的社交网络。

具体而言，现有群体界定文献大多基于社区邻里、学校、年级、班级、宿舍等物理空间概念，或者基于性别、种族、移民身份等社会人口学特征进行划分（如，Frijters et al.，2019；赵颖，2019；Horrace et al.，2021；宋泽和邹红，2021；吴娜等，2022），并假设个体在上述特定且固定的群体中互动。然而，此类同伴群体（peer group）的划分与假设，使得同

* 邱磊菊，中央财经大学中国互联网经济研究院；李寅，中央财经大学信息学院；段月姣，南开大学金融学院。通信作者及地址：段月姣，天津市海河教育园区同砚路 38 号南开大学金融学院，300350；电话：13502008704；E-mail：dyj417@163.com。作者感谢三位匿名审稿人专业且具有重要指导意义的意见，以及中央财经大学教育教学改革基金 2022 年度课题（2022XYJG80）、国家自然科学基金青年项目（72203115）和中央高校基本科研业务费的资助。感谢南开大学前沿交叉学科研究院的支持。

中国知网 https://www.cnki.net

群效应的研究存在严重的镜像问题[①](Manski，1993；Bramoullé et al.，2009)，无法真正识别出具有社会乘数效应的同群效应。

应用社交网络作为群体划分标准更贴近现实情况。现有社交网络的实证分析多集中在少数公开可用数据集，如著名的美国青少年健康纵向调查(Add Health)，在该问卷中，受访学生至多提及五位朋友(Calvó-Armengol et al.，2009)；或者基于班内座位邻近关系来认定朋友关系，识别班级内部更小的同伴群体(王春超和钟锦鹏，2018；吴贾等，2020；谭娅等，2021)。然而，基于主观问卷和实验数据容易出现样本偏差问题。因为不同受访者对朋友的定义与理解具有差异性，或者因为某些隐情(如，恋爱关系等)使得被调查者可能不愿相互认定彼此是朋友；而座位不邻近的学生之间朋友关系在实验中则容易被忽略。此外，通过问卷调查等技术收集现实世界中社交网络数据，对提问方法具有敏感性(Kogovšek and Ferligoj，2005)，而且在实践过程中，特别是测度大规模的社交网络时需要花费极高的成本。

本文聚焦校园卡食堂交易大数据，挖掘学生的“共现”频数作为同伴(peer)关系的判别依据，进而构建社交网络。相较于班级、年级等特定群体，本文识别出了更加紧密且动态变化的同伴关系，并定义为“朋友”。相较于主观调查问卷，应用每天实际用餐发生的交易数据推测出的社交网络更加客观。特别地，这一社交网络的构建，可与具有网络同群效应的检验模型有效结合(Bramoullé et al.，2009)以克服镜像问题。因此，有别于传统社交群体划分，本文建立的具有同群效应的社交网络中，首先放松了个体所处的参照群(同伴群体)仅在相同群内互动的假设；其次，结合类似于求解联立方程组式的识别策略，有效解决镜像问题(Bramoullé et al.，2009；Patacchini and Venanzoni，2014)。值得一提的是，数字化时代背景下，如校园卡一样的大数据覆盖范围甚广，本文的识别与构建方法为全国高校，乃至不同行业探究学生、员工等人力资本要素的社交网络，提供了数据依据和科学手段，为高质量转型发展提供助力。

基于校园卡大数据构建的动态演化社交网络，本文证实大学生学业表现具有正向积极的同群效应。有别于选取同一群体视为同伴群的研究(如，同一班级、同一种族等)，本文更清晰地捕捉到群体内的实际互动，从而可以进一步探查群体异质性，以及该群体对同伴群体的差异影响。结合理论推断，群体中社交关系的强弱具有差异影响，强关系更能促进信任、合作和社会规范，而弱关系则更可能提供获得新信息和资源的途径(Granovetter，1973)。本文实证证实，社交频率越高、关系强度越大，正向的同群效应越显著，但社交关系维持的时长则没有显著差异影响，此类精细的差异化分析对教育管理更具内涵性指导意义。

① Manski(1993)指出在同一群体中的个体具有相似行为表现是由内生效应(endogenous effect)、外生效应(exogenous effect)、关联效应(correlated effect)所导致的。内生效应是指同一群体中个体行为表现会受到同伴行为表现的影响；外生效应是指个体行为表现由外生特征决定，因而群体中具有相似外生特征的人会产生相似的行为表现；关联效应则是群体中的个体因某些未能观测的相似特征而产生的相似行为表现。上述三种效应具有不同的政策内涵，内生效应对教育相关政策具有放大作用(Berlinski，2023)，而外生效应和关联效应则不产生这种“社会乘数”效应，因此 Manski(1993)明确指出同群效应指内生效应，其他两种效应需要被剔除。然而，实际研究中很难区分，这就是著名的镜像问题。

中国知网　https://www.cnki.net

本文进一步将学生学业投入（school engagement），特别是情感投入（emotional engagement）纳入机制研究中。学业投入包括外在投入（behavioral engagement），如时间投入、课堂参与度等，以及内在情感投入，如对学习环境、学习状态是否满意等心理活动。社会认知理论认为，同伴互动能够对学生信念产生巨大影响，特别是自我效能感的提升（Bandura，1986；Micari and Pazos，2016），学业投入直接影响学生的学业表现（Furrer and Skinner，2003；Fredricks et al.，2004；Schmidt et al.，2017）。具有紧密关系的同伴的成功，更能增强学生完成类似任务的信念，进而影响其学习动机和情感投入。本文将上述投入因素纳入同群效应的机制分析，并结合教育部中国高等教育学会的大学生学情调查问卷数据（以下简称“教育部学情调查数据”）进行实证检验。研究显示，同伴的学业表现会增加主体学生的学习时间、课堂参与等学习投入和心理满足感等情感投入，从而对其学业表现产生积极作用。此外，本文应用了同伴的二阶项（即同伴的同伴的特征）工具变量，以及前一期社交网络影响等多种方法，进行了内生性和稳健性检验。在通过大数据构建社交网络克服镜像问题基础上，进一步地减少其他内生性影响。最终证实，同群效应对大学生学业表现存在着显著且稳健的积极正向影响，且社交关系强度越强，同群效应越大。

二、研究假说

现有同群效应研究的对象多以中小学生居多。Add Health的问卷调查对象主要是以走读制为主的高中生，因为其每天回家所以受家庭影响较大。Presler（2022）使用餐厅就餐排队位置的临近关系识别小学生群体中的朋友关系，但小学生就餐排队顺序容易受到老师的影响。大学生大多离家就学，住在校园内，更加独立，大学生的就餐数据更能体现他们真实的社交互动（Martin et al.，2020）[①]。因此，本文关注大学生群体。

高等教育阶段是学生离开家庭后独自面对的教育经历。有别于义务教育阶段和高中教育阶段，学生在大学校园里接触最多的不是家长和老师，而是与之年纪相仿的同龄人。因此，大学生受同伴的影响可能比中小学生更显著。校园里，大学生一起生活、学习，参加各种活动，互动频繁，从而形成各自的社交网络，大学生就餐数据更能体现他们真实的社交互动（Martin et al.，2020）。学业作为大学的主要任务，在社交中被普遍关注，因此，学生在日常互动中对学习成绩、方法、习惯等情况进行频繁交流，从而相互帮助、模仿或者竞争，产生学业表现的同群效应。所以，本文提出第一个研究假说：

假说1 在大学生社交网络中，存在学业表现的同群效应。

进一步，在教育学研究中，学业投入被认为是学业表现的主要影响因素。学业投入是指学生在学习过程中所展现的积极参与和投入程度，主要包括学习投入和情感投入两个方面。学习投入通常指学生在学习活动中的外在行为表现，如时间投入、课堂参与度等具体行为（Furrer and Skinner，2003；Fredricks et al.，2004）。情感投入则关注学生在学习过程中的情感反应，如对学习环境、学习状态的满意度、兴趣、愉悦感等心理活动

① Martin et al.（2020）利用餐厅入口门禁数据来识别大学生的朋友关系，认为经常同时进入餐厅的学生为朋友，但是在实证研究中并没有构建社交网络，仅探究了班里存在至少一位朋友是否会提高学生的成绩。

中国知网 https://www.cnki.net

(Fredricks et al., 2004)。积极的情感投入，通常能够增强学生对学校活动的参与，从而提高他们的学业成绩(Ainley and Ainley, 2011; Conner and Pope, 2013; Schmidt et al., 2017)。负面的情感体验，如焦虑或不满，可能导致学生对学术任务的回避，进而影响他们的学习效果。情感投入的测量通常通过评估学生对学校和教师的情感反应来实现。例如，Schnitzler et al.(2021, p.635)使用了以下问题来测度情感投入："In the preceding lesson, I was happy to be there."(在前一节课上，我很高兴在那里)，结果表明情感投入对学业成绩的显著正向作用。

现有同群效应研究中，大多侧重竞争或激励机制，究其根本也是通过对学生学业投入的影响从而产生作用。首先，同伴群体内的学生相互交流，互相帮助，模仿彼此学习习惯(Fafchamps and Mo, 2018)，会对群内个体的学习时间、学习方式等学习投入行为产生直接的影响，进而影响其学业表现；其次，当同伴表现出较高学业水平时也会产生竞争环境(谭娅等，2021)，适度的竞争有益于增加学生的学习投入。除此之外，适度竞争对学生学业情感投入方面也会产生积极作用。优秀的同伴会激发学生的自我要求，增加其学习投入，促进其学业表现(Wentzel and Wigfield, 1998)，这是因为当人们感到幸福、满意等积极情感时，所产生的积极心理，会促进更好的发挥和表现(Connell and Wellborn, 1991)。但是，在高度竞争环境中，相对落后的学业表现，也可能让学生产生消极情感，拉大同伴间的差距。然而，上述同伴群体的界定更多是同一班级、学校等，却鲜有识别出这些学生更加紧密的社交网络，进而较少关注学生在更紧密的社交网络中产生的学业投入的变化。因此，科学考察群体中大学生的积极情感投入与消极情感投入是重要且必要的。

本文通过校园卡大数据构建了学生社交网络，在这种关系更紧密的群体中，朋友的好成绩更能为学生个体创造出积极正向的学习氛围和环境，进而对学生的情感投入产生积极影响。社会认知理论认为同伴朋友互动能够对学生信念产生巨大影响，特别是对自我能力的信念，即自我效能感。因此，学生看到与他关系紧密的朋友成功，也更有可能相信自己能够完成类似的任务，从而影响其学习动机和情感投入(Bandura, 1986; Micari and Pazos, 2016)。综上，处在同一社交网络中的学生更倾向于相互模仿，相互激励，良性竞争，朋友的学业表现直接或间接影响着其他个体的学习投入和情感投入，据此本文提出第二个研究假说：

假说 2　大学生学业表现的同群效应，通过外在学习投入和内在情感投入两个渠道产生影响。

三、数据说明、社交网络构建与研究设计

(一) 数据来源与社交网络构建

我国高等教育的寄宿制度，使得大学生用餐地点相对集中固定，校园卡使用频率较高。基于上述特性，本文收集整理大学生校园卡在餐厅交易的大数据构建研究所依赖的社交网络，并进一步结合学生特征、成绩信息以及教育部学情调查数据，进行学业表现的同群效应研究以及相关机制检验。

中国知网　https://www.cnki.net

1. 高校及数据源说明

本文数据来自某财经类高等院校2015级至2017级本科生的信息数据、教育部学情调查数据，以及2015年9月至2018年1月校园一卡通就餐记录数据库。将上述三个数据库进行合并得到涵盖2015—2016学年两个学期、2016—2017学年两个学期，以及2017—2018学年第一学期，共计5个学期（2015—2018）的学生年级、学院、专业、户口、性别、生源地、父母受教育水平、高考分数和各门课程成绩等特征信息，以及校园卡餐厅交易数据信息，其中每笔交易含有交易地点和时间。

具体而言，样本中2015—2016学年两个学期中的学生样本为2015级的大一学生；2016—2017学年两个学期的学生样本包括2015级的大二学生，以及2016级的大一新生；2017—2018学年第一学期的学生样本则涵盖了2015级的大三学生、2016级的大二学生，以及2017级的大一学生。此外，结合教育部学情调查数据，本文整合了该校学生的学习投入时间、师生交流、学校满意度等信息，构成了本文的研究样本[①]，并对样本进行了如下处理：(1)基于校园卡餐厅交易数据构建学生社交网络；(2)按照各省市当年高考情况对高考分数进行了名次排位转化；(3)标准化处理学生课程成绩；(4)剔除军事理论、体育课、外语课等全校讲座课；(5)剔除部分相关数据缺失的样本。最终的研究样本含有6 160个有效观测值。

2. 社交网络构建说明

大学校园中朋友间一起吃饭是常见的社交活动，在社会学研究中称之为“共现”，即两个人在同一时间出现在同一地点。“共现”现象是衡量朋友关系的重要指标（Eagle et al.，2009），两人之间存在朋友关系的概率随着“共现”次数的增多而增加（Crandall et al.，2010）。因此，本文通过对校园卡餐厅交易数据进行处理得到两位学生的“共现”频数，进而推断其朋友关系。

依据对食堂进行调查统计，学生及其朋友通常选择同一楼层餐厅，可能在不同档口进行消费，前后存在时间差，大约5分钟，如果在同一档口消费，通常一起排队不超过5分钟。餐厅同一楼层的所有档口为一个“统计单元”，基于校园卡的每一条刷卡记录时间，统计此笔交易记录往后5分钟内的交易记录，将其记作两两“共现”一次（具体算法详见附图A1[②]）。在稳健性检验中，对时间窗口延长至10分钟进行再次检验。

基于上述算法，加总所有“共现”次数，得到15 763 881人次的“共现”记录。依据统计学“小概率事件”发生概率不超过5%的计算法则，本文将每学期潜在朋友列表中前5%“共现”记录保留，提取样本中五个学期“共现”频数的最大值（28），并将该次数记作数据集最低阈值，从而获取“共现”频数≥28次的记录，也即是将频数小于28次的记录视为陌生人巧遇，而非朋友“共现”。最终保留“共现”记录仅占总记录的0.71%，即只有0.71%的记录被认定为朋友“共现”，余下的99.29%的记录被判断为陌生人偶遇而予以筛除。具体记录如附表A1所示。

① 学生姓名和学号等信息均已脱敏处理。

② 限于篇幅，附录未在正文列示，感兴趣的读者可在《经济学》（季刊）官网（https://ceq.ccer.pku.edu.cn）下载。

中国知网 https://www.cnki.net

（二）变量定义与模型设计

借鉴社交网络中同群效应理论模型(Bramoullé et al., 2009)，本文构建大学生学业表现同群效应模型如下：

$$Score_{i,j,k,t}=a+\beta \boldsymbol{GP_Score}_{-i,j,k,t}+\boldsymbol{\gamma Control}_{i,j,k,t}+\boldsymbol{\delta GP_Control}_{-i,j,k,t}+u_j+v_k+w_t+\varepsilon_{i,j,k,t}\ , \tag{1}$$

其中，$Score_{i,j,k,t}$ 是学院 j 年级 k 学生 i 在学期 t 的成绩。本文称该学生为研究的主体学生，取其成绩代理学业表现，是主要的被解释变量；$\boldsymbol{GP_Score}_{-i,j,k,t}$ 为主体学生朋友的平均成绩，为本文关键的解释变量，下标 $-i,j,k,t$ 表示主体学生 i 的朋友，为区别于主体学生，称其朋友为“朋友学生”。$\boldsymbol{G}$ 为 $n\times n$ 的邻接矩阵(n 为网络中学生数量)，在此邻接矩阵中，如果学生 j 是学生 i 的朋友，则有 $G_{ij}=1/n_i$，否则为零。学生朋友群体被定义为社交网络中单个学生节点 i 周围的一阶邻接学生节点。系数 β 反映了朋友的平均成绩对其行为的影响(同群效应)，是本文重点关注的参数；$\boldsymbol{Control}_{i,j,k,t}$ 是主体学生的特征矩阵(如年龄、性别等)，参数 $\boldsymbol{\gamma}$ 反映了学生个体层面特征对其行为的影响(个体效应)；$\boldsymbol{GP_Control}_{-i,j,k,t}$ 是其朋友的特征矩阵，其中 $\boldsymbol{G}$ 依旧为邻接矩阵，参数 $\boldsymbol{\delta}$ 反映了朋友的平均特征对其行为的影响(外生效应)；本文进一步控制了学院(u_j)、年级(v_k)、学期(w_t)层面的效应以控制关联效应，a 是常数项，ε 为误差项。

由于在不同学院、班级、课程中的学生成绩存在差异性，所以本文对学业成绩变量 $Score$ 经过标准差标准化处理，即首先让每学期、每个教学班的成绩符合标准正态分布，然后计算每个学生每学期的标准分数的算术平均值。该方法使得学生成绩之间具有可比性，减少了任课教师和课程因素的影响，在一定程度上控制了关联效应。

个人特征变量包括高考排名($Rank$)、户口($Birth$)、性别($Gender$)、生源地区(东部、中部、东北、西部四大区)、父母受教育程度(Edu)。借鉴 Lépine and Estevan(2021)使用 SAT 成绩作为学生认知能力的代理变量，本文使用高考成绩来衡量学生认知能力。由于不同省份高考试卷存在差异，本文选取高考排名来衡量，值越小表示排名越靠前，学生认知能力越高。本文同时控制了主体学生以及朋友学生的上述特征变量。此外，将学院(u_j)作为固定效应变量，以控制不同学院对学生入学要求、培养方案、课程体系和日常教学管理模式等无法观测到的因素；将年级(v_k)作为固定效应变量，以排除入学年份或在校时长等不可观测因素的影响；将学期(w_t)作为固定效应变量，减少学生朋友关系和学业成绩因学期不同而产生的影响。具体含义与描述性统计见表 1 与表 2。

表 1　变量定义表

变量类型	变量名	变量定义
被解释变量：主体学生成绩	*Score*	主体学生的学业表现代理变量，即 ln(10＋主体学生学期标准分数)
解释变量：朋友学生成绩	*GP_Score*	主体学生的朋友的成绩平均值，即 ln(10＋朋友学生学期标准分数)

中国知网　https://www.cnki.net

（续表）

变量类型	变量名	变量定义
主体学生控制变量	*Rank*	学生高考排名，即学生分数排名/考生总人数
	Birth	学生户口所在地，城镇为 1，否则为 0
	Gender	学生性别，男性为 1，女性为 0
	East	东部地区生源学生取值为 1，西部地区为 0
	Middle	中部地区生源学生取值为 1，西部地区为 0
	Northeast	东北地区生源学生取值为 1，西部地区为 0
	Edu	父母受教育程度，父母中有本科及以上学历者为 1，否则为 0
朋友学生控制变量	*GP_Rank*	朋友学生高考排名均值
	GP_Birth	朋友学生户口所在地均值
	GP_Gender	朋友学生性别均值
	GP_East	朋友学生东部地区生源地均值
	GP_Middle	朋友学生中部地区生源地均值
	GP_Northeast	朋友学生东北地区生源地均值
	GP_Edu	朋友学生父母受教育程度均值
固定效应变量	u_j	学院虚拟变量
	v_k	年级虚拟变量
	w_t	学期虚拟变量

表 2 描述性统计

变量名	样本量	均值	标准差	P25	P50	P75
Score	6 160	0.156	0.625	−0.180	0.254	0.590
GP_Score	6 160	0.249	0.326	0.165	0.301	0.395
Rank	6 160	0.014	0.023	0.004	0.008	0.016
Birth	6 160	0.806	0.396	1	1	1
Gender	6 160	0.319	0.466	0	0	1
East	6 160	0.370	0.483	0	0	1
Middle	6 160	0.233	0.422	0	0	0
Northeast	6 160	0.076	0.265	0	0	0
Edu	6 160	0.451	0.498	0	0	1
GP_Rank	6 160	0.013	0.013	0.008	0.012	0.015
GP_Birth	6 160	0.776	0.218	0.689	0.783	1.000
GP_Gender	6 160	0.365	0.293	0.125	0.336	0.500
GP_East	6 160	0.353	0.245	0.250	0.341	0.444
GP_Middle	6 160	0.246	0.221	0.100	0.238	0.312

中国知网 https://www.cnki.net

（续表）

变量名	样本量	均值	标准差	P25	P50	P75
GP_Northeast	6 160	0.069	0.136	0.000	0.036	0.083
GP_Edu	6 160	0.416	0.266	0.283	0.379	0.500

四、实证结果与分析

（一）基准结果

表 3 展示了学生学业表现同群效应的基准回归结果。第(1)列仅考虑社交网络中朋友的成绩的影响，系数为 0.387，并在 1%水平上显著为正，表明朋友成绩对主体学生成绩具有正向影响作用，存在正向的同群效应。第(2)、(3)列依次加入主体学生特征控制变量(高考排名、户口、性别、生源地区、父母受教育程度)，以及朋友特征变量和固定效应(学院、年级、学期)，结果表明在考虑个体特征、朋友特征、固定效应等外生和关联效应的情况下，朋友成绩对主体学生成绩仍具有显著正向影响。控制变量中，高考排名的系数显著为负，说明高考排名越靠前，即认知能力越强的学生学业表现越好；样本中城镇户口、女大学生成绩相对更好；相较于西部地区的学生，东部、中部、东北生源区的学生成绩更好一些。而学生父母的受教育程度对其大学成绩影响不显著，部分印证了大学期间的成绩受父母影响较小，突出了大学生群体相对中小学生作为样本研究同群效应的优势。综上，大学生群体中，朋友学业表现对主体学生学业具有显著正向同群效应，证实了假说 1 的推断。

表 3　学业表现的同群效应

变量名	社交网络主体学生学业表现：*Score*		
	(1)	(2)	(3)
GP_Score	0.387***	0.298***	0.304***
	(0.024)	(0.024)	(0.025)
Rank		−0.353***	−0.353***
		(0.035)	(0.035)
Birth		0.006***	0.006***
		(0.002)	(0.002)
Gender		−0.029***	−0.028***
		(0.002)	(0.002)
East		0.016***	0.016***
		(0.002)	(0.002)
Middle		0.012***	0.013***
		(0.002)	(0.002)

中国知网　https://www.cnki.net

（续表）

变量名	社交网络主体学生学业表现：*Score*		
	(1)	(2)	(3)
Northeast		0.008**	0.008**
		(0.003)	(0.003)
Edu		0.000	0.000
		(0.002)	(0.002)
Constant	1.415***	1.627***	1.613***
	(0.056)	(0.056)	(0.059)
控制朋友特征变量	否	否	是
固定效应	否	是	是
N	6 160	6 160	6 160
R^2	0.040	0.115	0.115

注：括号内为标准误；显著性水平分别为* $p<0.1$，** $p<0.05$，*** $p<0.01$。下同。

（二）控制内生性问题

本文运用具有社交网络的同群效应回归模型，已经很大程度上克服了镜像问题（Bramoullé et al.，2009），同时也尽可能地控制了个体特征、朋友特征所引起的个体效应和外生效应，并考虑了年级、学院、学期等固定效应来控制关联效应，但仍可能存在潜在偏差。由于学生社交网络并非随机产生，某些未被观测到的特征（如，兴趣、喜好等），可能同时对其朋友关系属性和学业成绩产生影响，这种内生性可能导致所检测到的同群效应是偏误的，因此本文运用工具变量等方法进行检验。

首先，选取学生个体特征的二阶邻接矩阵（$\boldsymbol{G}^2\boldsymbol{X}$）和社交网络的二阶因变量（$\boldsymbol{G}^2\boldsymbol{Y}$）作为工具变量。具体而言，二阶邻接矩阵$\boldsymbol{G}^2\boldsymbol{X}$是指同伴的同伴（peer of peer）的特征（Bailey et al.，2018），其与同伴的个体特征$\boldsymbol{GX}$及同伴的成绩$\boldsymbol{GP_Score}$相关，但并不与主体学生的成绩*Score*直接相关，是较为理想的工具变量（Bramoullé et al.，2009；Bailey et al.，2018）。社交网络的二阶因变量$\boldsymbol{G}^2\boldsymbol{Y}$，可以理解为同伴的同伴的成绩，也与同伴的个体特征及同伴的成绩相关，但不直接与主体学生的成绩相关，也是理想的工具变量。借鉴Boucher and Houndetoungan(2020)，构建社交网络中同伴的同伴数据计算出$\boldsymbol{G}^2\boldsymbol{X}$和$\boldsymbol{G}^2\boldsymbol{Y}$，进行工具变量回归。

表4展示了使用不同工具变量组合的回归结果。第(1)列使用同伴的同伴个体特征$\boldsymbol{G}^2\boldsymbol{X}$作为工具变量，系数为0.649，表明同伴成绩与主体学生成绩呈显著正相关关系；第(2)列使用同伴的同伴的成绩$\boldsymbol{G}^2\boldsymbol{Y}$作为工具变量，系数为0.316，表明同伴成绩与主体学生成绩呈显著正向相关；第(3)列同时使用$\boldsymbol{G}^2\boldsymbol{X}$和$\boldsymbol{G}^2\boldsymbol{Y}$作为工具变量，结果依然显著为正。说明本文的结果是稳健的，大学生学业表现存在正向同群效应。

中国知网 https://www.cnki.net

表 4　工具变量回归结果

变量名	社交网络主体学生学业表现：*Score*		
	IV：$\boldsymbol{G^2X}$	IV：$\boldsymbol{G^2Y}$	IV：$\boldsymbol{G^2X+G^2Y}$
	(1)	(2)	(3)
GP_Score	0.649***	0.316***	0.451***
	(0.093)	(0.081)	(0.072)
Constant	0.810***	1.590***	1.273***
	(0.216)	(0.189)	(0.169)
控制主体特征变量	是	是	是
控制朋友特征变量	是	是	是
固定效应	是	是	是
N	6 160	6 160	6 160
R^2	0.101	0.096	0.100

其次，检验非随机性选择导致的遗漏变量偏差。参考 Oster et al.(2019)，使用自变量系数 β 和拟合优度 R^2 的函数构建调整后的自变量系数 β^*，计算出同群效应的调整参数 $\beta^*=0.245$，与原参数 $\beta=0.304$ 相比，调整后的参数略小，但仍为正数，且数值相对较大。说明在考虑了不可观测变量的影响后，同群效应对学业的影响依然显著存在，且并未发生剧烈变化，结果较为稳健。但是，此方法只能说明朋友关系的内生形成对研究的结论影响不大，并未完全克服朋友关系形成的内生问题，未来需要寻求更高质量的数据或者实验方法进行深入研究。此外，本文借鉴 Martin et al.(2020)，使用前一学期学生形成的社交网络对本学期学生学业表现进行回归，结果如附表 A2 所示，系数依然显著。

（三）异质性影响分析

学生的认知能力、生源地区等方面的个体特征差异会使学业表现的同群效应产生异质性。不同认知能力的学生受他人影响程度有所不同，因此其学业表现的同群效应也不尽相同。社交关系强度、长度等方面也存在差异，这些因素也会直接或间接地影响学业表现的同群效应。强度较高、长度较长的社交关系，往往会带来更多的互动、模仿、竞争。综合研究上述差异因素，可以更全面地分析大学生学业表现的同群效应，为教育管理提供更科学细致的参考依据。本文主要从学生个体特征异质性与社交网络异质性两个维度来进行讨论。

(1) 个体特征异质性影响分析。本文选取高考排名作为衡量学生认知能力的代理变量，按中位数将样本分为低认知能力（高考排名较低）和高认知能力（高考排名较高）两组。如表 5 中第(1)、(2)列所示，低、高认知能力的两组学生学业成绩系数分别为 0.334 和 0.281，说明同群效应在较低认知能力组中的影响更大。作为发展中国家，我国不同区域省份的教育发展仍不均衡。中西部经济欠发达、地理位置偏远地区，存在师资力量短缺、教育设施落后、教育质量较低等方面问题。本文将样本中学生分为中西部地区生源

中国知网　https://www.cnki.net

和其他地区两组，回归结果如表 5 第(3)、(4)列所示。结果显示中西部地区生源系数为 0.265，其他地区为 0.351，表明中西部地区学生受同群效应影响小于其他地区生源的学生。

表 5 考虑个体特征异质性后的同群效应影响

变量名	社交网络主体学生学业表现：*Score*			
	较低认知	较高认知	中西部	其他地区
	(1)	(2)	(3)	(4)
GP_Score	0.334***	0.281***	0.265***	0.351***
	(0.035)	(0.035)	(0.034)	(0.036)
Constant	1.559***	1.646***	1.711***	1.517***
	(0.082)	(0.083)	(0.081)	(0.086)
控制主体特征变量	是	是	是	是
控制朋友特征变量	是	是	是	是
固定效应	是	是	是	是
N	3 080	3 080	3 411	2 749
R^2	0.109	0.124	0.131	0.099

(2) 社交网络异质性影响分析。样本中包含五个学期的校园卡数据，因而本文将五个学期中的任意两个或以上学期的“共现”频数超过阈值时，定义为具有长期社交关系的“长期朋友”；反之，仅有一个学期记录的社交关系称为“短期朋友”。表 6 第(1)、(2)列显示，长期朋友和短期朋友的成绩系数分别为 0.200 和 0.215，差别不大，说明社交关系长度对大学生学业表现的同群效应的影响差别较小。社交关系的强度可以通过社交频次来衡量(Crandall et al., 2010)。本文根据最终样本中每学期“共现”频次的中位数，将频次数大于中位数的归入高社交强度组，频次数小于等于中位数的归入低强度组，然后对五个学期的不同组别学生重新构建社交网络，使用两个数据集分别进行回归分析，结果如表 6 第(3)、(4)列所示。高强度社交互动的组中，代表同群效应的朋友成绩系数为 0.284，远大于低强度社交组的 0.09，说明社交关系强的朋友产生更大的同群效应。

表 6 考虑社交关系特征异质性后的同群效应影响

变量名	社交网络主体学生学业表现：*Score*			
	长期朋友	短期朋友	高社交强度	低社交强度
	(1)	(2)	(3)	(4)
GP_Score	0.200***	0.215***	0.284***	0.090***
	(0.026)	(0.027)	(0.022)	(0.027)
Constant	1.859***	1.820***	1.661***	2.115***
	(0.061)	(0.063)	(0.053)	(0.064)

中国知网 https://www.cnki.net

（续表）

变量名	社交网络主体学生学业表现：*Score*			
	长期朋友	短期朋友	高社交强度	低社交强度
	(1)	(2)	(3)	(4)
控制主体特征变量	是	是	是	是
控制朋友特征变量	是	是	是	是
固定效应	是	是	是	是
N	4 223	5 457	5 568	5 297
R^2	0.121	0.106	0.123	0.093

（四）稳健性检验[①]

主结论中时间窗口为 5 分钟的设定，可能会导致有一些朋友关系被遗漏的情形。因此，本文延长了“共现”行为被判定为同一社交关系的时长，进而更大范围地捕捉潜在的学生社交关系。基于 10 分钟的时间窗口设定，我们重新构建了学生社交网络，并进一步进行实证检验。通过比较五个学期的“共现”频数，选定最大值 45 次作为样本数据集的最低阈值。缺失值处理后，仅有 0.68％的“共现”记录被纳入研究，其余 99.32％的“共现”记录则被判断为学生偶然的相遇，从而被筛除。对比 5 分钟窗口设定时保留的“共现”记录的 0.71％比例，研究样本“共现”数据的保留比例的变化仅为 0.03％，从而说明本文构建的社交网络相对稳定。表 7 汇报了时间窗口设定为 10 分钟后的社交网络同群效应检验结果，同群效应系数为 0.288，并在 1％水平下显著，证实本文的结论具有较强的稳健性。

表 7　时间窗口设定为 10 分钟的社交网络同群效应影响

变量名	10 分钟窗口设定下的社交网络主体学生学业表现：*Score*		
	(1)	(2)	(3)
GP_Score	0.362***	0.283***	0.288***
	(0.025)	(0.025)	(0.026)
Constant	1.475***	1.663***	1.656***
	(0.059)	(0.059)	(0.061)
控制主体特征变量	否	是	是
控制朋友特征变量	否	否	是
固定效应	否	是	是
N	5 598	5 598	5 598
R^2	0.036	0.113	0.115

① 其他的稳健性检验结果见附录Ⅰ。

中国知网 https://www.cnki.net

五、同群效应的影响机制

本文构建学业表现同群效应的影响机制模型如下：

$$Study_{i,j,k,t} = a + \beta \boldsymbol{GP_Score}_{-i,j,k,t} + \gamma \boldsymbol{Control}_{i,j,k,t} + \delta \boldsymbol{GP_Control}_{-i,j,k,t} + u_j + v_k + w_t + \varepsilon_{i,j,k,t} \quad (2)$$

$$Emo_{i,j,k,t} = a + \beta \boldsymbol{GP_Score}_{-i,j,k,t} + \gamma \boldsymbol{Control}_{i,j,k,t} + \delta \boldsymbol{GP_Control}_{-i,j,k,t} + u_j + v_k + w_t + \varepsilon_{i,j,k,t} \quad (3)$$

上述模型(2)和模型(3)，分别代表外在的学习行为投入与内在的心理情感投入两个影响渠道，其中 $Study_{i,j,k,t}$ 代表主体学生 i 的学习投入，$Emo_{i,j,k,t}$ 代表主体学生 i 的情感投入。

首先，学习投入是影响学生学业表现的重要因素，包括学习时间投入、课程参与度等。教育部学情调查中，包含“每周课内外学习时间”与“做口头报告或公开演讲的数量”的问题，能在很大程度上反映学生的学习投入。参考上述问题，本文设定学习时长(*Stime*)、演讲(*Present*)进行学习投入机制检验。具体而言，这两个变量是对学习时间或报告演讲的次数取对数计算得到。结果汇报于表 8，第(1)、(3)列显示朋友成绩能够很大程度上提升主体学生的学习时间、演讲报告次数等学习投入。第(2)、(4)列为上述不同学习投入变量对主体学生成绩的回归。显而易见，学习投入越大，学习成绩就越高。综上，朋友学业表现越好，主体学生的学习投入越多，学业表现也越好。

其次，对学业的情感投入主要涉及学习满足感等心理活动，通常通过评估学生对学校和教师的情感反应来实现(Schnitzler et al., 2021)。教育部学情调查中，含有“你对所学专业的满意程度”“重新选择时你仍会选择本专业的可能性”的问题，在一定程度上反映了学生在群体中的情感感知，因此本文根据上述问题构建相关情感变量，记 *S-major*。具体来说，当学生对一个问题给予正面评价时，*S-major* 取值为 1，如果都是满意评价时，取值为 2，否则为 0。此外，为了排除转专业需求而提高的成绩，将转专业的学生删除后，将朋友成绩对上述情感投入代理变量进行回归检验，结果汇报于表 8 中第(5)、(6)列，结果均显著为正，说明朋友成绩对学生的情感投入存在积极正向显著影响，而学生对所学专业越满意，成绩就越好。本文还对情感投入机制做了进一步的分析，主要从高低成绩的学生的情感投入的差异性分析进行讨论，具体见附录Ⅱ表Ⅱ1。

表 8 同群效应的影响机制研究

变量名	*Stime*	*Score*	*Present*	*Score*	*S-major*	*Score*
	(1)	(2)	(3)	(4)	(5)	(6)
GP_Score	0.227***		0.083**		1.121***	
	(0.050)		(0.037)		(0.356)	
Stime		0.073***				
		(0.006)				

中国知网 https://www.cnki.net

（续表）

变量名	*Stime* (1)	*Score* (2)	*Present* (3)	*Score* (4)	*S-major* (5)	*Score* (6)
Present				0.078***		
				(0.009)		
S-major						0.008***
						(0.001)
Constant	2.094***	2.135***	2.337***	2.130***	−1.810**	2.317***
	(0.118)	(0.018)	(0.088)	(0.023)	(0.840)	(0.006)
控制主体特征变量	是	是	是	是	是	是
控制朋友特征变量	是	是	是	是	是	是
固定效应	是	是	是	是	是	是
N	6 160	6 160	6 160	6 160	5 910	5 910
R^2	0.042	0.113	0.091	0.106	0.043	0.106

六、结论与启示

本文通过大学校园卡大数据拓展了同群效应相关研究，丰富了大学生学业表现的影响研究。基于校园卡交易大数据，设计算法并构建学生社交网络，为类似研究提供了数据获取与处理的方法和启示。基于本文所构建的社交网络数据进行的实证检验，克服了同群效应研究中的镜像问题，识别了具有“乘数效应”的内生效应，弥补了现有高等教育领域同群效应研究中的不足。同群效应和社会乘数是经济政策研究中的重要概念，尤其是在当前经济压力和财政资源紧张的背景下，更加凸显出通过同伴互动来放大政策效果的潜力。社会乘数作用意味着，一个人的行为影响会通过同伴互动传播，带来较大范围的经济和社会效应，这为经济政策提供了放大效应的机会。

教育、人才是全面建设社会主义现代化国家的基础性、战略性支撑，本文为高等教育管理与时俱进的升级过程提供科学参考依据。首先，高等院校管理者在组织学生活动、制定相关制度时，应科学考虑大学生的同群效应影响，有效利用“乘数效应”，实现较低成本投资下的教育质量提升。其次，高校教师在日常教学中，可采取分组或结对方式，促进学生之间的社交互动，引导学习成绩较好的学生发挥带动作用，从而高效提升学生整体的学业表现。最后，充分认识同群效应的异质性与影响机制，引导学生积极参与健康向上、充满正能量的社交活动，通过改善学习投入和情感投入，提高学业表现，促进学生人力资本的提升。

参 考 文 献

[1] Abbas, Q., and J. Foreman-Peck, “Human Capitaland Economic Growth: Pakistan, 1960-2003”, *The Lahore Journal of Economics*, 2008, 13(1), 1-27.

中国知网 https://www.cnki.net

[2] Ainley, M., and J. Ainley, "Student Engagement with Science in Early Adolescence: The Contribution of Enjoyment to Students' Continuing Interest in Learning About Science", *Contemporary Educational Psychology*, 2011, 36(1), 4-12.

[3] Bailey, M., R. Cao, T. Kuchler, and J. Stroebel, "The Economic Effects of Social Networks: Evidence from the Housing Market", *Journal of Political Economy*, 2018, 126(6), 2224-2276.

[4] Bandura, A., "Social Foundations of Thought and Action", *Englewood Cliffs*, NJ 1986, 2, 23-28.

[5] Berlinski, S., M. Busso, and M. Giannola, "Helping Struggling Studentsand Benefiting All: Peer Effects in Primary Education", *Journal of Public Economics*, 2023, 224, 104925.

[6] Boucher, V., and A. Houndetoungan, "Estimating Peer Effects Using Partial Network Data", Working Paper, 2020.

[7] Bramoullé, Y., H. Djebbari, and B. Fortin, "Identification of Peer Effects through Social Networks", *Journal of Econometrics*, 2009, 150(1), 41-55.

[8] Calvó-Armengol, A., E. Patacchini, and Y. Zenou, "Peer Effects and Social Networks in Education", *The Review of Economic Studies*, 2009, 76(4), 1239-1267.

[9] Connell, J. P., and J. G. Wellborn, "Competence, Autonomy, and Relatedness: A Motivational Analysis of Self-System Processes", *Journal of Personality & Social Psychology*, 1991, 65, 43-77.

[10] Conner, J. O., and D. C. Pope, "Not Just Robinson Crusoe: Student Engagement, Autonomy and the Project of an Alternative School", *Educational Policy*, 2013, 27(1), 155-187.

[11] Crandall, D. J., L. Backstrom, D. Cosley, S. Suri, D. Huttenlocher, and J. Kleinberg, "Inferring Social Ties from Geographic Coincidences", *Proceedings of the National Academy of Sciences of the United States of America*, 2010, 107(52), 22436-22441.

[12] Currie, J., and D. Thomas, "Early Test Scores, School Quality and SES: Longrun Effectson Wage and Employment Outcomes", *Research in Labor Economics*, 2001, 20, 103-132.

[13] Eagle, N., A. S. Pentland, and D. Lazer, "Inferring Friendship Network Structure by Using Mobile Phone Data", *Proceedings of the National Academy of Sciences of the United States of America*, 2009, 106(36), 15274-15278.

[14] Fafchamps, M., and D. Mo, "Peer Effects in Computer-Assisted Learning: Evidence from a Randomized *Experiment*", *Experimental Economics*, 2018, 21, 355-382.

[15] Fredricks, J. A., P. C. Blumenfeld, and A. H. Paris, "School Engagement: Potential of the Concept, State of the Evidence", *Review of Educational Research*, 2004, 74(1), 59-109.

[16] Frijters, P., A. Islam, and D. Pakrashi, "Heterogeneity in Peer Effects in Random Dormitory Assignment in a Developing Country", *Journal of Economic Behavior & Organization*, 2019, 163, 117-134.

[17] Fruehwirth, J. C., "Identifying Peer Achievement Spillovers: Implications for Desegregation and the Achievement Gap",*Quantitative Economics*, 2013, 4(1), 85-124.

[18] Furrer, C., and E. Skinner, "Sense of Relatedness as a Factor in Children's Academic Engagement and Performance", *Journal of Educational Psychology*, 2003, 95(1), 148-161.

[19] Granovetter, M. S., "The Strength of Weak Ties", *American Journal of Sociology*, 1973, 78(6), 1360-1380.

[20] Hanushek, E. A., and L. Woessmann, "School Resourcesand Student Achievement: A Review of Cross-Country Economic Research", *Cognitive Abilities and Educational Outcomes: A Festschrift in Honour of Jan-Eric Gustafsson*, 2017, 149-171.

[21] Hedges, L. V., R. D. Laine, and R. Greenwald, "An Exchange: Part I: Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes", *Educational Researcher*, 1994, 23(3), 5-14.

[22] Horrace, W. C., H. Jung, J. Presler, and A. E. Schwartz, "What Makes a Classmate a Peer? Examining Which

中国知网 https://www.cnki.net

Peers Matter in NYC Elementary Schools", Working Paper, 2021.

[23] Kogovšek, T., and A. Ferligoj, "Effects on Reliability and Validity of Egocentered Network Measurements", *Social Networks*, 2005, 27(3), 205-229.

[24] Lépine, A., and F. Estevan, "Do Ability Peer Effects Matter for Academic and Labor Market Outcomes?", *Labour Economics*, 2021, 618, 102022.

[25] Lewbel, A., X. Qu, and X. Tang, "Social Networks with Misclassified or Unobserved Links", *Journal of Political Economy*, 2023, 131(4), 898-946.

[26] Li, Y., J. V. Lerner, and R. M. Lerner, "Personal and Contextual Factorsin School Engagement: Examining the Relationship Between Self-Regulation and School Engagement", *Journal of Youth and Adolescence*, 2021, 50(2), 353-365.

[27] Lin, X., "Utilizing Spatial Autoregressive Models to Identify Peer Effects among Adolescents", *Empirical Economics*, 2015, 49, 929-960.

[28] Manski, C. F., "Identification of Endogenous Social Effects: The Reflection Problem", *The Review of Economic Studies*, 1993, 60(3), 531-542.

[29] Martin, D. D., A. C. Wright, and J. M. Krieg, "Social Networks and College Performance: Evidence from Dining Data", *Economics of Education Review*, 2020, 79, 102063.

[30] Micari, M., and P. Pazos, "Fitting in and Feeling Good: The Relationships Among Peer Alignment, Instructor Connectedness, and Self-Efficacy in Undergraduate Satisfaction with Engineering", *European Journal of Engineering Education*, 2016, 41(4), 380-392.

[31] Oster, E., "Unobservable Selection and Coefficient Stability: Theory and Evidence", *Journal of Business & Economic Statistics*, 2019, 37(2), 187-204.

[32] Osiobe, E. U., "A Literature Reviewof Human Capital and Economic Growth", *Business and Economic Research*, 2019, 9(4), 179-196.

[33] Patacchini, E., E. Rainone, and Y. Zenou, "Heterogeneous Peer Effects in Education", *Journal of Economic Behavior & Organization*, 2017, 134, 190-227.

[34] Patacchini, E., and G. Venanzoni, "Peer Effects in the Demand for Housing Quality", *Journal of Urban Economics*, 2014, 83, 6-17.

[35] Presler, J., "You Are Who You Eat With: Evidence on Academic Peer Effects from School Lunch Lines", *Journal of Economic Behavior & Organization*, 2022, 203, 43-58.

[36] Schmidt, J. A., L. Shumow, and H. Z. Kackar-Cam, "Does Mindset Intervention Predict Students' Daily Experience in Classrooms? A Comparison of Seventh and Ninth Graders' Trajectories", *Journal of Youth and Adolescence*, 2017, 46, 582-602.

[37] Schnitzler, K., D. Holzberger, and T. Seidel, "All Better Than Being Disengaged: Student Engagement Patterns and Their Relations to Academic Self-Concept and Achievement", *European Journal of Psychology of Education*, 2021, 36, 627-652.

[38] 宋泽、邹红，"增长中的分化：同群效应对家庭消费的影响研究"，《经济研究》，2021 年第 1 期，第 74—89 页。

[39] 谭娅、封世蓝、张庆华、龚六堂，"同群压力还是同群激励？——高中合作小组的同群效应研究"，《经济学》(季刊)，2021 年第 2 期，第 533—556 页。

[40] Tilak, J. B., "The Dilemmaof Reforms in Financing Higher Education in India", *Higher Education Policy*, 1997, 10(1), 7-21.

[41] 王春超、钟锦鹏，"同群效应与非认知能力——基于儿童的随机实地实验研究"，《经济研究》，2018 年第 12 期，第 177—192 页。

[42] Wentzel, K. R., and A. Wigfield, "Academic and Social Motivational Influences on Students' Academic Performance", *Educational Psychology Review*, 1998, 10(2), 155-175.

中国知网　https://www.cnki.net

[43] 吴贾、刘姝、周翔，“空间距离和朋友网络的同群效应：来自小学随机实验的证据”，《世界经济》，2020 年第 8 期，第 75—99 页。

[44] 吴娜、白雅馨、安毅，“主动模仿还是被动反应：商业信用同群效应研究”，《南开管理评论》，2022 年第 2 期，第 149—161 页。

[45] 赵颖，“同群效应如何影响学生的认知能力”，《财贸经济》，2019 年第 8 期，第 33—49 页。

The Peer Effects of Academic Performance in Higher Education: Based on Campus Card Big Data

QIU Leiju LI Yin

(Central University of Finance and Economics)

DUAN Yuejiao*

(Nankai University)

Abstract: Using the transaction data of campus card, we develop a social network to identify the peer effects with the “social multiplier” effect. We find a positive peer effect on the academic performance of college students. The higher the social interaction frequency and strength of relationships among college students, the more significant the peer effects, while there is no significant difference in the duration of social interactions. With the survey data of college students' academic situation, we verify the operation channels: external behavioral engagement and internal emotional engagement.

Keywords: academic performance in higher education; peer effects; campus card big data

JEL Classification: I23, J24, N35

* Corresponding Author: DUAN Yuejiao, School of Finance, Nankai University, No. 38 Tongyan Road, Tianjin Haihe Education Park, Tianjin 300350, China; Tel: 86-13502008704; E-mail: dyj417@163.com.

中国知网 https://www.cnki.net