

素。由于五类课程的评教得分都呈现左偏厚尾，不服从正态分布，表明评教数据难以满足 OLS 估计的基本假设，因此本文尝试采用 GLS 回归的估计方法。

由于 S 大学的评教指标主要评价教师的教学态度、教学方法、教学内容等，因此本文结合教师的个人特质、学生参与评教等方面设计多元回归模型。根据相关研究文献，本文设计以下实证模型：

$$\text{Score} = \beta_0 + \beta_1 \text{title} + \beta_2 \text{sex} + \beta_3 \text{zb} + \beta_4 \text{zx} + \beta_5 \text{gb} + \beta_6 \text{gx} + \beta_7 \text{people} + \beta_8 \text{scale}$$

其中 score 表示评教得分；title 代表教师的职称，0 代表助教、讲师和高级讲师，1 表示副教授，2 表示教授；sex 代表教师性别，0 表示女性教师，1 表示男性教师；zb、zx、gb、gx、sy 分别表示课程类型中的专必、专选、公必、公选和实验课程，本文采取虚拟变量的计量方式；people 代表课程学生人数；scale 代表课程参评率。评教得分为实证模型中的因变量；自变量包括课程性质（5 类课程性质，用虚拟变量），教师性别，教师职称（是否高级职称），课程学生人数，课程参评率等。

首先，对全部评教数据进行怀特检验，结果在 1% 置信水平下显著，由怀特检验结果可知评教数据存在显著异方差现象。因此，本文的实证模型采用加权最小二乘法（WLS）进行估计，并得到表 3 的结果。

表 3 总体数据的估计结果

	Coef (系数)	P 值
title (职称)	0.4037	0.000 **
sex (性别)	-0.5879	0.000 **
zb (专必)	-2.4611	0.000 **
zx (专选)	0.0228	0.923
gb (公必)	0.7532	0.000 **
gx (公选)	-0.3345	0.012 *
people (课程学生数)	-0.0014	0.000 **
scale (课程参评率)	-3.3993	0.000 **
cons	95.6206	0.000 **

注：** 表示在 1% 置信水平下显著，* 表示在 5% 置信水平下显著。

根据估计的实证结果，本文发现：

(1) 教师职称对教师评教得分影响显著，教师职称越高，评教得分越高，教师职称每高一级，评教得分则提高 0.4 分。这一结果一定程度反映教师职称越高，其教学质量越得到学生的认可，评教得分与教师的科研能力及对课堂的控制能力正相关，反映中、低职称的教师对大课堂的掌控能力相对较弱。

(2) 课程类型中只有专业选修课不显著影响评教得分，专业必修、公共必修、公共选修、实验四类课程对评教得分的影响显著。学生评教打分不仅是教师的教学质量，还包括学生对课程设置的意见和学生自己的学业期望是否与教师及课程冲突等综合性的因素。从本文数据分析结果可以发现，专选课程得分最低，波动性最大，同时也最不显著。这是由于其他类课程中教师和课程的选择具有固定性和强制性，因此学生对课程设置的意见较小，对课程期望较弱，因此给分普遍较高，从而导致了评教得分过高且影响显著的结果。而专选类课程多是学生根据自己的兴趣爱好自主选择的，具有较强的主观期望和学习目的，当学生的课程设置和自己的学业期望相冲突时，会给较低的分数，且打分波动性变大，造成专选课程对教师评教得分影响不显著。

(3) 课程学生人数显著影响评教得分。课程的学生人数越少，教师评分越高，学生人数减少一人，评教得分将提高 0.0014 分。这可能是由于学生人数多的班级授课，教师在多样化课堂教学活动、批改作业和师生互动上存在较大困难，因而导致授课教师难以顾及每个学生的需求，课堂效果不好；而对于学生人数少的班级，教师可以根据每个学生的学习基础、学习能力给予个别指导和调整教学的难易程度，课堂效果较好，从而评教得分较高。

(4) 性别对教师评教得分影响显著，女性教师比男性教师的评教得分多 0.5879。这是由于男性教师和女性教师在表达方式和教学风格等方面存在差别。男性教师更注重信息量的输入与授课的权威性，从而容易带有说教性质；而女性教师的课堂互动性相对较强，比较注重交流和沟通，因此，学生偏好女性教师的授课方式，从而给女性教师更高的得分。

(5) 课程参评率对评教得分影响显著。课程参评率上升 1%，教师评教得分下降 3.3993%。这一结论说明此次评教数据可能存在偏差，由于个别学生放弃评教而对评教结果产生影响，因此评教指标体系仍需改进。

五、结论与反思

基于 S 大学评教数据的统计分析，本文认为学生评教的客观性总体上比较欠缺。首先，学生评教具有一定的局限性。学生社会阅历较少，知识层面也较浅，对评教的看法和理解难免片面，导致评教结果容易失之客观。其次，学生评教具有主观性。由于评教通常采用匿名方式，学生在评教时容易受情绪影响，