Does Undue Preference Lead to Unfairness? The Impact of Teacher Favoritism on Teacher Treatment and Student Achievement

Yung-Yu Tsai[1]

Abstract

This study investigates the impact of teacher favoritism (i.e., systematically favoring students of a specific gender, ethnicity, or high previous class ranking) on teacher treatment and student achievement. Based on the practice of random class grouping within schools, a school fixed effect model is used to estimate the causal effect of teacher favoritism. The results suggest that teacher favoritism toward students with a higher previous class rank increases the likelihood of these students consulting with their teacher on academic, emotional, and friendship issues. Also, female students are more likely to discuss friendships and emotional matters with pro-female teachers. However, the evidence of teacher favoritism's impact on student achievement is relatively weak. *JEL Codes: J15, J16, I24.*

*Keywords—*Teacher favoritism, Teacher bias, Teacher-student relationship, Student achievement gap

1   Introduction

Teacher favoritism refers to teachers granting preference to a specific group of students [1–3]. Previous studies suggest that teachers sometimes give preferential treatment to students of a specific gender [4,5], race/ethnicity [6,7], and ability [8–10]. A substantial body of literature identifies that teacher favoritism widens the student achievement gap [11–15]. Existing studies on gender favoritism find an effect size on the gender achievement gap ranging from 0.03 to 0.12 [15–17]. Regarding ethnicity favoritism, relevant literature concludes that effect size ranges from 0.04 to 0.25 [11,12,18]. There is minimal research directly examining achievement favoritism. In a relevant study, Hornstra et al. [19] find that teachers' negative attitude toward dyslexia widens the achievement gap between students with and without dyslexia by a 0.45 standard deviation.

This study examines the impact of teacher favoritism on teacher treatment and student achievement. I address whether and to what extent teachers' preference toward a specific group

---

[1] University of Missouri-Columbia. ytsai@mail.missouri.edu

leads to disparity in how students are treated and whether this disparity amplifies the student achievement gap. I also compare whether favoritism based on different criteria (gender, ethnicity, and previous achievement) differs in patterns or magnitudes toward how it affects teacher treatment and student achievement.

I utilize a longitudinal survey of Taiwanese junior high school students and their homeroom teachers to examine the research questions. The primary identification strategy is based on the random class grouping practice *within* schools in Taiwan. I use a school fixed effect model to identify the causal effect of teacher favoritism.

This study reveals three key findings. First, the impact of favoritism varies according to the criteria. Female students assigned to a pro-female teacher are more likely to reach out to the teacher to consult on emotional and friendship issues. High achievement students have a higher probability of consulting with their teacher on academic issues if the teacher shows favoritism toward them. Evidence of ethnicity favoritism is relatively weak, aligning with Taiwan's less pronounced ethnic disparity. These differences suggest that teacher favoritism is more likely to affect students when merged with cultural context.

Second, while there is some evidence that teachers will potentially hold higher expectations of students in a favored group, the evidence is not as solid as the impact of the teacher-student relationship. The positive expectation gap is not significant when using a more robust estimation (leave-one-out approach). Furthermore, no evidence exists that students will internalize teachers' expectations into their self-aspirations.

Third, despite teacher favoritism leading to disparity in teacher-student relationships, the impact on achievement is less evident. The baseline estimates suggest that teacher gender favoritism affects students' science scores and that achievement-based favoritism impacts students' overall scores. However, the leave-one-out approach, which removes the potential reverse causality of students' traits affecting teachers' favorability, suggests no significant estimation for these effects. The finding indicates that the actual impact of teacher favoritism on achievement may be weak and less prominent than the impact of teacher-student relationships.

The study contributes to the previous studies in the following ways: First, this study uses a more straightforward and comprehensive approach to measure teacher favoritism. Existing literature usually conceptualizes teacher favoritism in terms of grading bias. These studies quantify the gap between the blind and non-blind grading of test scores and determine that the grading bias

2

negatively affects the unfavored group's academic performance [13,15,16]. However, grading bias is only a "proxy" of favoritism [16]; not all favoritism results in grading bias. For example, research finds that teacher prejudice toward ethnic minority students can adversely impact students' learning even if a grading bias does not exist [20,21]. Therefore, using grading bias to measure favoritism may fail to capture some of the negative impacts of such practice.

This study focuses on a more intuitive approach to measuring favoritism—quantifying the extent a teacher grants preference to one group over another. Based on previous studies [2,22,23], I use the *partial correlation coefficient* to quantify the connection intensity between student group (i.e., gender, ethnicity, or previous achievement) and teacher favorability, conditional on observed student characteristics. Specifically, if two students possess the same qualities (e.g., behavior and personality), but the teacher still prefers one student to the other, it is favoritism. The advantage of using the correlation between teacher favorability and student traits is that it captures a more comprehensive and straightforward concept of favoritism than grading bias.

Secondly, this study investigates and compares favoritism based on different criteria and explores the potential synergistic effects when one type of favoritism is combined with another. Few studies directly examine this relationship. However, there is evidence suggesting that different stereotypes may produce interaction effects. For instance, Ambady et al. [24] find that female Asian-American students improve their math performance when the researchers activate their ethnic identity rather than gender identity. However, the experiment does not examine when both ethnicity and gender identity are stimulated.

This study compares the effects of gender, ethnicity, and achievement favoritism. The findings suggest that the effect size of various types of favoritism integrate with the social context. The evidence of achievement favoritism is more prominent likely due to the pro-achievement culture in Asian countries. The impact of ethnicity favoritism is null, aligning with the social context in Taiwan. Furthermore, this study explores the potential synergistic effects when students benefit from multiple types of favoritism. While most results are noisy, the point estimates present interesting patterns that suggest synergistic effects and are worth further research.

The remainder of this paper is organized as follows. Section 2 discusses the educational context and cultural background in Taiwan. Section 3 presents the data, measurements, and summary statistics. Section 4 explains the identification strategy. Section 5 presents the empirical results. Finally, Section 6 discusses the results and provides concluding remarks.

3

2   The Educational Context in Taiwan

2.1   The Educational System and Student Allocation

In Taiwan's educational system, primary schools cover grades one through six, junior high is grades seven through nine, and senior high covers the tenth to twelfth grades. Schooling up to the ninth grade is compulsory. Students receive comprehensive primary and junior high school education but are streamed into academic or vocational tracks starting in the tenth grade.

This study focuses on junior high school students. Typically, seventh-grade students are around 13 years old. Most junior high schools are public, although around 10% are private. Admission to public schools is based on students' place of residence and school catchment area. Admission to private schools varies by school and may be based on an admission exam.

While school admission in Taiwan is not random, class assignment within the schools is random (at least quasi-random). Taiwan's Ministry of Education (MOE) requires all junior high schools to implement *heterogeneous class groupings*, which requires the random assignment of students to different classes to ensure student performance is equally distributed.[2]

2.2   Gender and Ethnicity Disparity and Pro-educational Culture

Before the 2000s, Taiwanese women had lower educational attainment, but the gap has recently been closing [25,26]. In addition to gender disparity, stereotypes regarding gender talent also persist. Studies in Taiwan have found that males are more likely to be interested in and choose STEM majors, while females prefer humanities or languages [27,28]. However, evidence from international student assessments suggests that Taiwanese students' gender gap in math and science is relatively small compared to other countries [29–31].

Taiwan has four primary ethnicities: Mainlanders, Holo, Hakka, and Indigenous. Indigenous people are those whose ancestors arrived in Taiwan in ancient times, while the ancestors of Holo and Hakka immigrated to Taiwan in the 17th century. Mainlanders arrived in Taiwan after the end of World War II. The original Mainlanders retreated to Taiwan with the Republic of China (ROC) government and have since occupied the nation's privileged positions [32]. However, since the 1950s, this ethnic educational disparity has gradually been bridged [33].

---

[2] Schools can adopt two approaches to reach the heterogeneous grouping. First, they can randomly assign students to different classes. Second, they can administer a test at enrollment and assign students to classes based on a sigmoid curve (i.e., the top-ranked student to class 1, the second to class 2, etc., and then reverse the order after the first round of assignments).

Like other Asian countries, Taiwan has a pro-educational culture in which parents and educators highly value academic achievement [34]. This mindset is deeply rooted in Taiwanese education and the teacher-student relationship. For example, test scores are highly valued, and it is firmly believed that without good scores, students cannot have a successful future [35,36].

3 Data, the Sample, and Measurements

3.1 Data and the Sample

This study uses data from the Taiwan Youth Project (TYP) [37], which targeted grade seven students in Taipei City, Taipei County, and Yilan County, Taiwan, in 2000. The survey sampled several junior high schools and then randomly chose two to three classes in each school. The respondents include all class students, their parents, and homeroom teachers.

3.2 Measurements and Variables

The key treatment variable in this study is teacher favoritism, measured by the teacher-reported favorability toward each student. Teachers provided information on whether they liked a specific student based on a four-point Likert scale. The survey was conducted in the Spring semester of grade seven when teachers had known their students for at least half a year.

I calculate the *partial correlation* between teacher favorability and student groups for *each class*, conditional on observed students' traits.[3] The index captures the extent to which a homeroom teacher favors a specific group after accounting for the observed traits. Correlation is calculated within classes rather than teachers. As this study only includes one cohort and each homeroom teacher is assigned to only one class in a specific year, I cannot construct teacher-level favoritism across multiple classes. Due to the small cluster size, there might be concern about measurement error. The measurement error could bias the estimation toward zero [38].

This study measures teacher favoritism in terms of gender (favors females), ethnicity (favors Mainlanders), and achievement (favors students with a higher rank in primary school). For gender, I use students' self-reported gender.[4] For ethnicity, I categorize students into Mainlanders (students with a Mainlander parent) and non-Mainlanders. For achievement, I use the self-reported

---

[3] These traits include gender, ethnicity, previous achievement, parental education, socioeconomic status, family income, and personality. Personalities are constructed with factor analysis. See Table A2 in Appendix A for details.
[4] The questionnaire uses the Mandarin word "xing-bie," a neutral word referring to either sex or gender when adding the prefix "biological" or "social." As it is self-reported, and in the previous relevant literature, "gender bias" and "gender gap" are more common terms than "sex," throughout the paper, I use the term "gender favoritism."

class rank in primary school. Table A1 in Appendix A includes the questionnaire wording.

The *partial correlation* between teacher favorability and student groups denotes the direction and intensity of the favoritism. A positive value means that with all other (observed) things equal, the teacher favors females over males, Mainlanders over non-Mainlanders, and high achievers over low achievers, or vice versa for a negative value. The partial correlation ranged from –1 to +1, with a value of –1 or +1 considered partial and a value of 0 considered neutral.

An underlying concern is that partial correlation only captures the tendency for students to possess unobserved "appealing" characteristics. To address the concern, I utilize the feature of teacher change to inspect measurement validity. In the sample, approximately 25% of the classes experienced a change in homeroom teachers in grades eight or nine. If the favoritism index only captures unobserved student characteristics, it should remain the same regardless of whether the teachers changed. In contrast, if the index captures the teachers' preference, a higher correlation should exist for the classes where homeroom teachers have not changed than for those that have.

Figure 1 illustrates the relationship between teacher favoritism in grades seven and nine by teacher change status. In classes with the same teachers, favoritism in the seventh grade is positively correlated with the ninth grade. Conversely, there was no positive correlation for classes experiencing a teacher change, suggesting that the teacher favoritism index is primarily driven by teachers' preferences rather than the students' (unobserved) characteristics.

To further address this concern, this study adopts the leave-one-out approach. I measure teacher favoritism for each individual but exclude the individuals themselves. This approach rules out bias due to individual unobserved characteristics. However, the leave-one-out approach produces a negative correlation between the measurement and student traits [16,39,40]. A female more favored by a teacher will receive a lower measured value of gender favoritism than her peers. This negative correlation could lead to a downward bias of the treatment effect. Therefore, I report results from both measurements, with the one based on all students as the baseline approach.

The primary outcome variable is the test scores of the Basic Competency Test (BCT), a grade nine standardized examination. This examination covers five subjects— reading (Mandarin), English, math, science, and social science—graded on a 0–60-point scale. The exam is multiple-choice and graded by computers; hence, it has blind grading and is free from favoritism. The exam score variables were derived from copies of the class exam results provided by the homeroom teacher. Thus, this variable is likely reliable and not subject to student or teacher response bias.
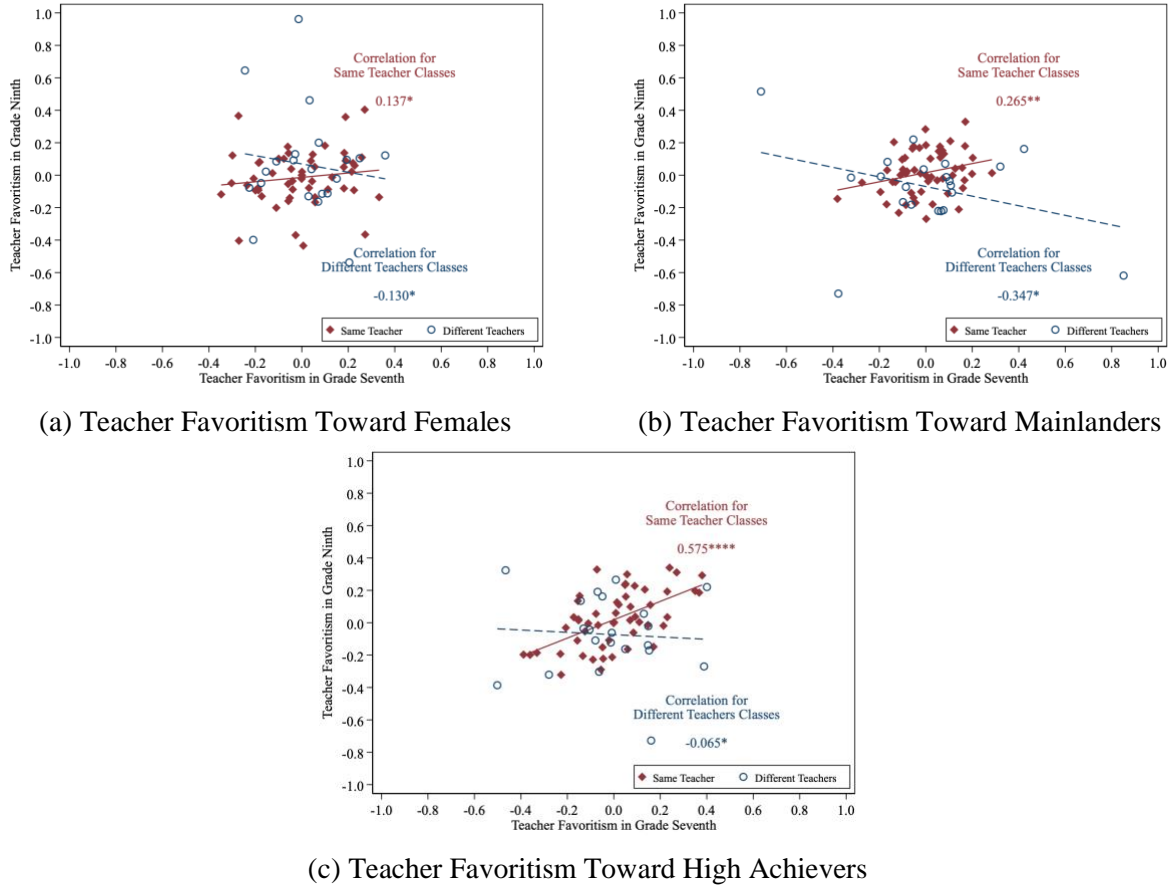
(a) Teacher Favoritism Toward Females

(b) Teacher Favoritism Toward Mainlanders

(c) Teacher Favoritism Toward High Achievers

Figure 1: The Correlation Between Grades Seven and Nine Teacher Favoritism

Note: The horizontal axis specifies teacher favoritism in grade seven, while the vertical axis represents teacher favoritism in grade nine. Each symbol represents one class. The diamond symbols denote classes with no change in teachers, and the solid line is the fitted line. The hollow circles indicate classes that changed teachers, and the dashed line is the fitted line.

3.3   Summary Statistics

Table 1 presents the summary statistics. The final sample comprises 2,463 students nested in 75 classes from 38 schools. The average number of students per class is 33.5, ranging from 14 to 41. The proportion of females is 49%, and that of the Mainlanders is 15%. The proportion of high achievers (self-reported as ranking in the top five of their primary school class) was 18%, while the proportion of low achievers was 22% (self-reported as being ranked 21st or after).[5]

---

[5] Primary school classes, on average, have 34 students. Hence, ranking in the top five is roughly the top 15%, and ranking 21st or after is the bottom 40%. The patterns show that students tend to overestimate their performance, which may cause bias in measuring teacher favoritism. Students favored by teachers may overestimate their performance, while those unfavored may underestimate themselves. Thus, the pattern might exaggerate the correlation between teacher favorability and students' self-reported past achievements, causing a downward bias in the treatment effect.

Table 1: Summary Statistics

Variables	Mean	SD	Within-school SD	Min	Max
Demographic Variables					
Female	0.488	0.500	0.502	0.000	1.000
Ethnicity[†]					
Mainlander	0.145	0.352	0.332	0.000	1.000
Holo	0.849	0.358	0.347	0.000	1.000
Haka	0.126	0.332	0.319	0.000	1.000
Indigenous	0.028	0.165	0.132	0.000	1.000
Previous Achievement[‡]					
High Achiever	0.177	0.382	0.376	0.000	1.000
Middle Achiever	0.605	0.489	0.486	0.000	1.000
Low Achiever	0.218	0.413	0.405	0.000	1.000
Teacher Favoritism					
Baseline Approach					
Gender Favoritism (Favors Female)	0.016	0.250	0.136	–0.566	0.634
Ethnicity Favoritism (Favors Mainlanders)	–0.014	0.189	0.098	–0.523	0.440
Achievement Favoritism (Favors High Achievers)	0.148	0.223	0.113	–0.265	0.599
Leave-One-Out Approach					
Gender Favoritism (Favors Female)	0.015	0.255	0.153	–0.674	0.731
Ethnicity Favoritism (Favors Mainlanders)	–0.014	0.195	0.114	–0.780	0.604
Achievement Favoritism (Favors High Achievers)	0.149	0.229	0.136	–0.640	0.795
Outcomes Variable: Standardized Exam Score					
Total Score	148.757	58.446	54.721	0.000	300.000
Reading	29.765	12.857	12.254	0.000	60.000
English	30.074	14.119	13.144	0.000	60.000
Math	29.633	12.070	11.370	0.000	60.000
Science	29.622	12.143	11.455	0.000	60.000
Social Science	29.662	12.216	11.483	0.000	60.000
Average Number of Students in a Class	33.502	4.145	0.946	14.000	41.000
Number of Schools	38				
Number of Classes	75				
Number of Students	2,463				

[†]Ethnicity is defined individually by each parent; thus, students may have multiple ethnicities. Hence, the sum of all ethnic categories does not equal one.
[‡]High achievers self-report as in the top five in their primary school class. Middle achievers self-report as ranked between 6th to 20th in their class. Low achievers self-report as ranked at or after 21st in their class.

The average teacher favoritism of the baseline approach was 0.016 for gender, –0.014 for ethnicity, and 0.148 for achievement. Column (2) reports the overall standard deviation (SD), while Column (3) displays the within-school SD. The within variation is about half less than the overall variation. The statistics for the leave-one-out approach have a similar mean and variation to the baseline approach but possess wider distributions.

4 Research Design

4.1 Empirical Strategy

The main identification challenge in estimating teacher effect is that students and teachers are sorted between and within schools [41–43]. The identification strategy in this study is primarily based on random class assignments *within* a school. The key assumption is that conditional on school fixed effects, students assigned to different classes are comparable.

Students typically remain in the same class and with the same teacher from grades seven to nine. However, classes can be assigned a new teacher,[6] or students may change classes. While the initial teacher assignment is random, the following changes may not be. This study measured teacher favoritism in grade seven regardless of whether the teachers and students later changed classes. Thus, the estimations in this study are the effects of the initial teacher assignment.[7]

Based on within-school variation in teacher favoritism, I compare students in the same schools with different teachers and varying levels of favoritism. I estimate the following equation:

$$Y_{ijk} = \alpha_0 + \beta_1 Favoritism_{jk} + \beta_2 Favoritism_{jk} \times Group_{ijk}^* + X_{ijk}$$
$$+ X_{ijk} \times Group_{ijk}^* + Z_{jk} + Z_{jk} \times Group_{ijk}^* + \theta_k + \varepsilon_{ijk} \quad (1)$$

Where Y_{ijk} is the outcome for student i assigned to teacher j in school k. $Favoritism_{jk}$ is the level of gender (ethnicity or achievement) favoritism. $Group_{ijk}^*$ is a dummy variable indicating whether student i is female (Mainlanders or high-achievers). $Favoritism_{jk} \times Group_{ijk}^*$ is the interaction term between teacher favoritism and the student group. X_{ijk} is a vector of student characteristics.[8] Z_{jk} is a vector of teacher characteristics.[9] The equation also includes the

---

[6] Potential reasons for a new teacher include the existing teacher's retirement or leave. Also, parents may request, or the school decides, to rearrange teachers due to academic or non-academic considerations.

[7] In the sample, approximately 25% of the classes were assigned a new teacher in grade eight or nine and approximately 5% of the students left their original classes (or schools).

[8] Student characteristics include gender (categorical variable of male or female), age (in years), whether living with father/mother (dummy variables), whether parents are married (dummy variable), parental ethnicity (categorical variables of Mainlanders, Holo, Haka, Indigenous, and Missing), parental education level (categorical variables of illiterate, primary school, junior high school, senior high school, vocational school, associate degree, bachelor degree, graduate degree, and not applicable), parental socioeconomic status (measured with the International Socio-Economic Index [ISEI] as proposed by Ganzeboom and Treiman [44]), log family income (retrieved from the parental questionnaire reported in NT thousand [1,000 NTD is roughly 30 USD]), the number of siblings (continuous), and class rank in primary school (see Table A1 in Appendix A for the measurement).

[9] Teacher characteristics include gender (categorical variable of male or female), teacher/student gender match (dummy variable), age (in years), normal university/college graduate (institutions specifically for training teachers)

interaction terms between the student group with student covariates ($X_{ijk} \times Group^*_{ijk}$) and teacher covariates ($Z_{jk} \times Group^*_{ijk}$). θ_k denotes the school fixed effect. ε_{ijk} is the error term.

In Equation (1), β_1 identifies the effect of teacher favoritism on the reference groups (males, non-mainlanders, and low achievers). The key estimate is β_2, which denotes the effect of favoritism on the *difference* in outcomes between female and male (mainlander and non-mainlander, or low-achieving and high-achieving) students.

This paper further considers the synergistic effects of multiple favoritism types by interacting all three types of favoritism in a regression model. Students experiencing both gender and ethnic favoritism should benefit from a more substantial impact than those receiving only one type of favoritism. To simplify the analysis, I use a dummy variable indicating whether teachers have positive or negative favoritism toward a specific student instead of using a continuous scale of favoritism measurement. This study estimates the following model:

$$
\begin{aligned}
Y_{ijk} = \alpha_0 + &\beta Favoritism.Gender^+_{jk} \times Favoritism.Ethnicity^+_{jk} \\
&\times Favoritism.Achievement^+_{jk} \\
&+ \gamma Favored.Gender_{ijk} \times Favored.Ethnicity_{ijk} \\
&\times Favored.Achievement_{ijk} + X_{ijk} + Z_{jk} + \theta_k + \varepsilon_{ijk}
\end{aligned}
\tag{2}
$$

In Equation (2), $Favoritism.Gender^+_{jk}$ is a dummy variable indicating that teacher j in school k has positive gender favoritism (favors females). Similarly, $Favoritism.Ethnicity^+_{jk}$ denotes positive ethnicity favoritism (favors Mainlanders), and $Favoritism.Achievement^+_{jk}$ indicates positive achievement favoritism (favors high achievers).

The equation includes the three-way interaction to examine various combinations of favoritism. $Favored.Gender_{ijk}$ equals one if student i benefits from gender favoritism. $Favored.Ethnicity_{ijk}$ denotes student benefits from ethnicity favoritism. $Favored.Achievement_{ijk}$ indicates that students benefit from achievement favoritism. γ represents the coefficients of all interaction terms, using the group that does not benefit from any favoritism as the reference group.

Throughout the analysis, the primary treatment variable—teacher favoritism—is a generated variable. Therefore, this study computes the standard error with a two-stage bootstrap

---

(dummy variable), possesses a graduate degree (dummy variable), experience (in years), whether the teacher is married (dummy variable), number of children (continuous variable), subject taught (categorical variables of Chinese Mandarin, English, Math, Physics and Chemistry, History, Geography, Civic Studies, and Others), attitudes toward physical punishment, strictness, role of the homeroom teacher, evening self-study session, cram school, and high school admission policy (measured with a 4-point Likert scale from 'strongly disagree' to 'strongly agree').

method. First, I resample observations from each class with replacement. Then, I recalculate favoritism with the new sample and re-estimate the main equations. This process is repeated 1,000 times, and the standard error is computed as the standard deviation of all estimates.

## 4.2   The Validity of the Empirical Strategy

Ideally, the causal effect of teacher favoritism should be estimated by comparing each student's real outcome with the *potential outcome* if assigned to a neutral teacher. However, the potential outcome is unobserved. Therefore, the key identification assumption is that after accounting for the school fixed effect, students who were assigned to different teachers possess the same potential outcomes and can serve as counterfactuals for each other.

As the true potential outcome is unobserved, this study estimates the predicted outcome as the function of a series of student exogenous variables.[10] If the identification assumption holds, the predicted student achievement should not correlate with teacher favoritism. Also, the predicted achievement gap should be the same regardless of the intensity of the teacher favoritism.

Figure 2 illustrates the relationship between favoritism and the actual and predicted outcomes (i.e., test scores). The symbols denote the average (predicted) test scores by favoritism level, grouped into twelve bins of equal sample sizes. I subtract the school means from the test scores to account for the school fixed effect. The fitted lines represent the linear relationship between favoritism and test scores, and the shaded areas refer to the 95% CI. The left panel shows the predicted outcomes, and the right shows the actual outcomes.

Figure 2a shows the relationship between gender favoritism and predicted exam scores. The two lines representing male and female are almost parallel, indicating that the predicted gender gap maintains a similar level throughout the favoritism distribution. This finding supports the empirical assumption that the potential gender gap should be the same across students assigned to different teachers. Figure 2b illustrates the case for the actual scores. When teachers have negative gender favoritism (i.e., favor male students), the gap between genders is small. However, as teacher gender favoritism increases (i.e., favors females), the gender achievement gap increases slightly.

---

[10] The predictors include gender, age, living with father/mother, whether parents are married, parental ethnicity, parental education level, parental socioeconomic status, log family income, number of siblings, class rank in primary school, and the interaction terms between students' groups and the above variables.
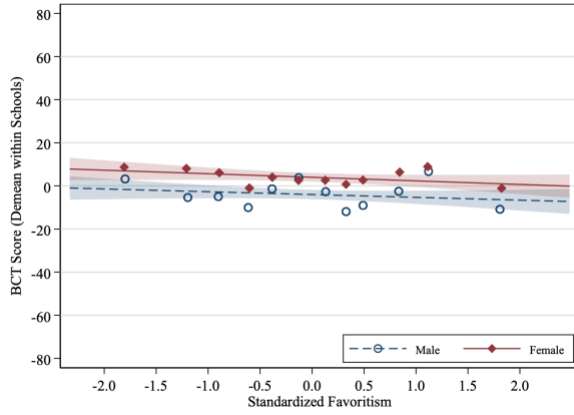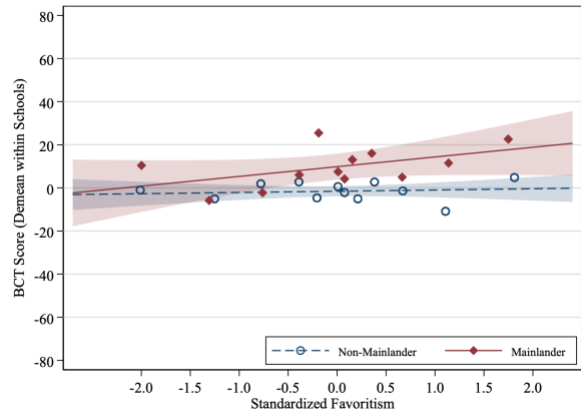
(a) Gender Favoritism and Predicted Test Scores
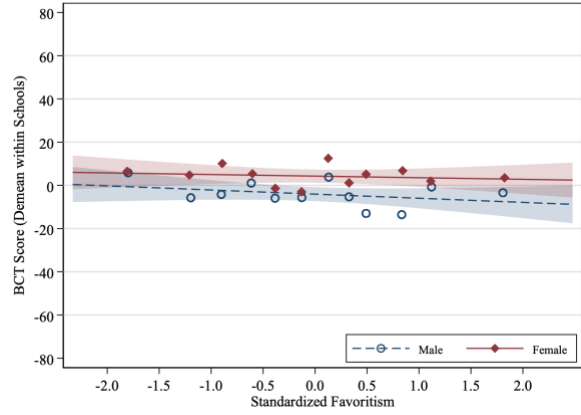
(b) Gender Favoritism and Actual Test Scores
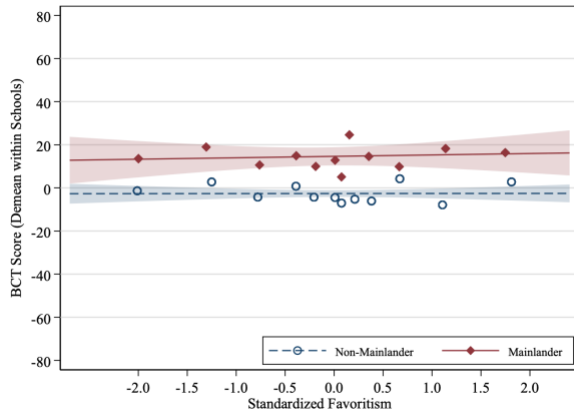
(c) Ethnic Favoritism and Predicted Test Scores

(d) Ethnic Favoritism and Actual Test Scores
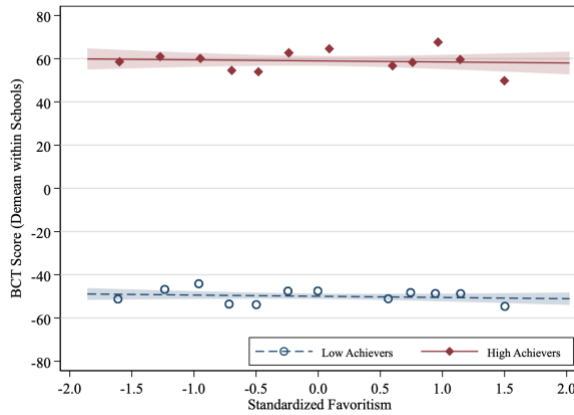
(e) Achievement Favoritism and Predicted Test Scores
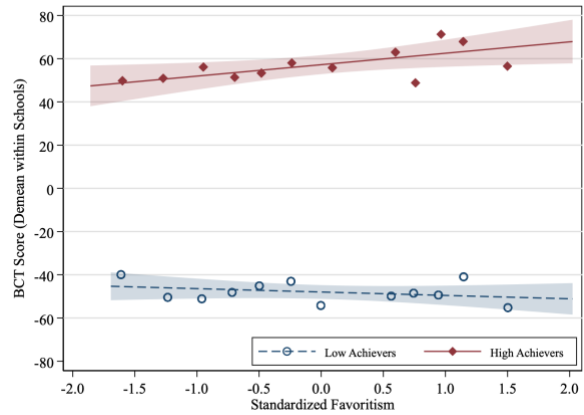
(f) Achievement Favoritism and Actual Test Scores

Figure 2: The Relationship Between Teacher Favoritism and (Predicted) Test Scores

Note: Each point stands for a bin of equal sample size. The (predicted) BCT scores are demeaned within schools and are predicted using the interaction terms between student groups (gender, ethnicity, or previous achievement) and student characteristics.

Figure 2c illustrates the relationship between ethnic favoritism and the predicted test scores. The fitted lines are parallel, suggesting that the ethnic gap predicted with the observed characteristics is not correlated with teacher favoritism. Conversely, Figure 2d displays the results of actual test scores. When teachers strongly favor non-mainlanders, there is almost no gap in achievement. However, if teachers favor mainlanders, the gap increases.

Figure 2e presents the relationship between teacher favoritism toward high achievers and predicted exam scores. The lines for high- and low-achievers are parallel, suggesting that the expected achievement gap is not associated with favoritism. This reveals that after accounting for school fixed effects, the student achievement gap predicted with exogenous covariates does not systematically correlate with teacher favoritism. Figure 2f illustrates the results of the actual scores, showing that the gap widens when teachers favor students with better prior achievements.

5 Empirical Results

5.1 The Effect of Teacher Favoritism on Teacher Treatment

Before examining the impact of teacher favoritism on student achievement, this section first presents the results on whether teacher favoritism leads to disparate teacher treatment toward favored and unfavored groups.

5.1.1. Teacher expectation and student self-expectation

This section first explores the impact of teacher favoritism on teachers' expectations toward students. I also examine the "Pygmalion effect," which states that teachers' high expectations of students change students' self-perceptions [45,46]. Table 2 reports the estimation results. Columns (1) to (4) use the baseline approach, and columns (5) to (8) apply the leave-one-out approach. All estimates are standardized coefficients.

Columns (1) and (5) report whether a teacher expects a student will attend high school (instead of vocational school or not continue their education). Columns (2) and (6) report the effect of teachers' expected years of education on students. Table A1 in Appendix A provides the questionnaire wording. The coefficient of $Favoritism$ represents the impact of teacher favoritism on the baseline group (i.e., male, non-mainlander, and low-achievers), and the coefficient of $Favoritism \times Female$, $Favoritism \times Mainlander$, and $Favoritism \times HighAchiever$ shows the impact of teacher favoritism on the gap between the favored and unfavored groups.

Table 2: The Effect of Teacher Favoritism on Teacher's and Student's (Self-)Expectations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline Approach				Leave-one-out Approach			
	Teacher Expectation		Student Self-expectation		Teacher Expectation		Student Self-expectation	
	Attend HS	Year of Ed.	Attend HS	Year of Ed.	Attend HS	Year of Ed.	Attend HS	Year of Ed.
Panel A: Teacher Favoritism towards Females								
Favoritism	−0.003	−0.243	0.014	0.008	−0.002	−0.203	0.016	0.019
	(0.059)	(0.077)	(0.062)	(0.060)	(0.059)	(0.077)	(0.062)	(0.060)
	[0.476]	[0.001]	[0.412]	[0.449]	[0.487]	[0.004]	[0.397]	[0.376]
Favoritism × *Female*	0.040	0.042	0.009	−0.010	0.005	0.003	0.005	−0.020
	(0.043)	(0.040)	(0.059)	(0.058)	(0.043)	(0.040)	(0.059)	(0.058)
	[0.181]	[0.149]	[0.439]	[0.431]	[0.450]	[0.470]	[0.464]	[0.363]
Panel B: Teacher Favoritism towards Mainlanders								
Favoritism	−0.028	−0.142	0.004	0.041	−0.008	−0.092	−0.002	0.073
	(0.065)	(0.089)	(0.067)	(0.056)	(0.065)	(0.089)	(0.067)	(0.056)
	[0.336]	[0.055]	[0.476]	[0.233]	[0.452]	[0.149]	[0.488]	[0.095]
Favoritism × *Mainlander*	0.059	0.119	0.123	0.019	−0.044	0.041	0.071	0.006
	(0.079)	(0.075)	(0.092)	(0.104)	(0.079)	(0.075)	(0.092)	(0.103)
	[0.229]	[0.056]	[0.092]	[0.425]	[0.290]	[0.293]	[0.222]	[0.477]
Panel C: Teacher Favoritism towards High Achievers								
Favoritism	−0.072	−0.019	0.055	0.072	−0.041	−0.006	0.008	0.073
	(0.065)	(0.096)	(0.095)	(0.077)	(0.065)	(0.096)	(0.095)	(0.077)
	[0.132]	[0.424]	[0.282]	[0.176]	[0.265]	[0.477]	[0.467]	[0.172]
Favoritism × *MiddleAchiever*	−0.021	0.012	−0.089	−0.024	−0.056	−0.031	−0.056	−0.022
	(0.057)	(0.053)	(0.089)	(0.070)	(0.057)	(0.053)	(0.089)	(0.070)
	[0.353]	[0.410]	[0.160]	[0.364]	[0.163]	[0.278]	[0.266]	[0.377]
Favoritism × *HighAchiever*	0.095	0.153	−0.023	−0.063	0.025	0.071	−0.009	−0.063
	(0.057)	(0.088)	(0.100)	(0.096)	(0.057)	(0.088)	(0.100)	(0.096)
	[0.049]	[0.042]	[0.408]	[0.255]	[0.329]	[0.210]	[0.464]	[0.258]
Observations	2,463	2,463	2,313	2,313	2,463	2,463	2,313	2,313

Note: Bootstrap standard errors retrieved from 1,000 resamples within class clusters are reported in parentheses. P-values are reported in squared brackets. The standardized coefficients for β_i are retrieved from Equation (1). All columns include the school fixed effect, student characteristics, teacher characteristics, and interaction terms between the student group and student/teacher characteristics. The number of observations differs due to missing outcomes.

Positive point estimates for expectation gaps exist for all types of favoritism. However, only the baseline approach shows some significant estimates (for ethnic favoritism on expected years of education and achievement favoritism on both educational expectations). The leave-one-out approach provides no solid evidence of the expectation gaps of teachers.

The next issue is whether students internalize teacher bias into self-expectation. Columns (3), (4), (7), and (8) take students' self-expectation as the outcomes. I use the change in variables from grades seven to nine as outcomes. In most cases, the point estimates on students' self-expectation do not align with the findings regarding teacher expectations. Some even report negative coefficients on the self-aspiration gap (i.e., favoritism lower the favored group's self-aspiration). For example, the baseline approach suggests positive and significant effects of teacher achievement favoritism on the teacher expectation gap (see Panel C of Columns (1) and (2)). However, this teacher expectation does not transfer into students' self-aspirations (see Panel C of Columns (3) and (4)). No evidence supports the Pygmalion effect results from teacher favoritism.

5.1.2. The Teacher-Student Relationship

This section considers the impact of teacher favoritism on the teacher-student relationship. When teachers favor specific students, the students will be more willing to seek help from their teachers [47]. Table 3 provides the estimated results. The outcomes demonstrate students' willingness to seek help from teachers on school life difficulties, coursework, emotional, or friendship issues. Table A1 in Appendix A lists the questionnaire wording.

Columns (1) to (4) show the results from the baseline approach, and Columns (5) to (8) are based on the leave-one-out approach. Panel A reveals that female students assigned to pro-female teachers are more likely to consult with teachers regarding emotional and friendship issues. The standardized coefficients range from 0.154 to 0.171 for the baseline approach and about two-thirds of the magnitude for the leave-one-out approach. Both approaches show significant estimates.

Panel B examines the impact of ethnicity favoritism. Both the baseline and leave-one-out approaches provide no significant results, and some point estimates are negative.

Panel C investigates the effect of achievement favoritism. Both the baseline and leave-one-out approaches suggest significant or marginally significant effects of teacher favoritism on students' frequency of reaching out to teachers for academic, emotional, and friendship issues. The magnitude of the standardized coefficients is 0.176 to 0.302 for the baseline approach and around two-thirds of the magnitude for the leave-one-out approach.

15

Table 3: The Effect of Teacher Favoritism on the Teacher-Student Relationship

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Seeks Help from the Homeroom Teacher for the Following:							
	Baseline Approach				Leave-one-out Approach			
	Difficulties	Academic	Emotional	Friend	Difficulties	Academic	Emotional	Friend
Panel A: Teacher Favoritism towards Females								
Favoritism	0.049	0.042	–0.021	–0.019	0.037	0.059	–0.000	–0.004
	(0.067)	(0.102)	(0.102)	(0.110)	(0.067)	(0.102)	(0.102)	(0.110)
	[0.234]	[0.341]	[0.420]	[0.431]	[0.291]	[0.282]	[0.500]	[0.487]
Favoritism × Female	0.010	0.113	0.171	0.154	0.017	0.046	0.108	0.094
	(0.058)	(0.052)	(0.053)	(0.057)	(0.058)	(0.052)	(0.053)	(0.057)
	[0.430]	[0.015]	[0.001]	[0.003]	[0.385]	[0.191]	[0.021]	[0.050]
Panel B: Teacher Favoritism towards Mainlanders								
Favoritism	0.034	0.124	0.185	0.272	0.044	0.150	0.219	0.288
	(0.057)	(0.117)	(0.111)	(0.120)	(0.057)	(0.117)	(0.111)	(0.120)
	[0.272]	[0.144]	[0.048]	[0.012]	[0.217]	[0.099]	[0.024]	[0.008]
Favoritism × Mainlander	–0.007	0.082	–0.012	0.021	0.012	–0.062	–0.129	–0.118
	(0.082)	(0.076)	(0.073)	(0.072)	(0.082)	(0.076)	(0.074)	(0.072)
	[0.464]	[0.141]	[0.433]	[0.383]	[0.440]	[0.207]	[0.040]	[0.050]
Panel C: Teacher Favoritism towards High Achievers								
Favoritism	–0.178	–0.093	–0.094	–0.072	–0.165	–0.047	–0.041	–0.059
	(0.089)	(0.121)	(0.114)	(0.126)	(0.089)	(0.121)	(0.114)	(0.126)
	[0.023]	[0.221]	[0.206]	[0.284]	[0.032]	[0.348]	[0.358]	[0.320]
Favoritism × MiddleAchiever	0.079	0.124	0.077	0.099	0.053	0.063	0.020	0.058
	(0.084)	(0.077)	(0.071)	(0.074)	(0.084)	(0.077)	(0.071)	(0.074)
	[0.175]	[0.053]	[0.140]	[0.092]	[0.265]	[0.206]	[0.387]	[0.217]
Favoritism × HighAchiever	0.149	0.302	0.263	0.176	0.118	0.204	0.184	0.143
	(0.103)	(0.107)	(0.093)	(0.090)	(0.103)	(0.108)	(0.093)	(0.090)
	[0.074]	[0.002]	[0.002]	[0.026]	[0.126]	[0.029]	[0.023]	[0.057]
Observations	2,313	2,457	2,457	2,449	2,313	2,457	2,457	2,449

Note: Bootstrap standard errors retrieved from 1,000 resamples within class clusters are reported in parentheses. P-values are reported in squared brackets. The standardized coefficients of β_i are retrieved from Equation (1). All columns include the school fixed effect, student characteristics, teacher characteristics, and interaction terms between the student group and student/teacher characteristics. The number of observations differs due to missing outcomes.

Overall, gender and achievement favoritism increase the likelihood of students seeking help. The impacts are most pronounced for achievement favoritism. The evidence on ethnicity favoritism is weak.

5.2   The Effect of Teacher Favoritism on Student Achievement

Previous sections reveal that teacher favoritism (specifically gender and achievement) increases the gap between favored and unfavored groups in teacher-student relationships. This section focuses on the impact on student achievement.

5.2.1. Gender Favoritism

Table 4 estimates the effect of gender favoritism on student achievement. All estimates are standardized coefficients. I progressively include different sets of covariates to check the robustness. Columns (1) to (3) report the results from the baseline approach, while Columns (4) to (6) display the results from the leave-one-out approach. Panel A shows the impact on the total score. The coefficient of $Favoritism \times Female$ denotes the impact of gender favoritism on the gender achievement gap. The baseline approach produces positive estimates ranging from 0.02 to 0.04. However, the leave-one-out approach estimates are near zero. Overall, there is no solid evidence of the impact of gender favoritism on total scores.

Previous studies suggest the gender favoritism effect is domain-specific [15,17,48]. In Table 4, Panels B to D estimate the favoritism effects on different subjects, showing a more substantial point estimate for science. The baseline approach suggests that a one SD increase in teacher favoritism leads to a marginally significant 0.06 SD increase in the science achievement gap (see Panel D Column (3)). The leave-one-out approach (serving as the lower bound) estimates an insignificant effect of 0.03. The magnitude of estimates falls in the range of previous studies regarding gender bias (ranging from 0.03 to 0.12) [15–17].

5.2.2. Ethnicity Favoritism

Table 5 examines the effect of teachers' ethnicity favoritism. The coefficient of $Favoritism \times Mainlander$ denotes the impact of ethnic favoritism on the achievement gap. The baseline approach with full covariates (Column (3)) suggests that a one SD increase in favoritism leads to a 0.61 SD insignificant increase in the ethnic achievement gap. The effect size is within the range of previous studies (ranging from 0.04 to 0.25) [11,12,18]. However, the leave-one-out approach reports a negative estimate (Column (6)). The effects on different subjects are similar (see Panels B to D). No conclusive evidence supports the existence of an ethnic favoritism effect.

17

Table 4: The Effect of Teachers' Gender Favoritism on Student Achievement

	Baseline Approach			Leave-one-out Approach		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Total Score						
Favoritism	–0.059	–0.061	–0.050	–0.052	–0.051	–0.043
	(0.050)	(0.040)	(0.049)	(0.050)	(0.040)	(0.049)
	[0.119]	[0.062]	[0.155]	[0.148]	[0.100]	[0.192]
Favoritism × Female	0.020	0.040	0.034	–0.006	0.010	0.003
	(0.047)	(0.036)	(0.038)	(0.047)	(0.036)	(0.038)
	[0.336]	[0.132]	[0.185]	[0.451]	[0.391]	[0.466]
Panel B: Reading Score						
Favoritism	–0.041	–0.043	–0.017	–0.034	–0.035	–0.016
	(0.048)	(0.040)	(0.046)	(0.048)	(0.040)	(0.046)
	[0.200]	[0.140]	[0.356]	[0.239]	[0.194]	[0.364]
Favoritism × Female	–0.002	0.017	0.007	–0.032	–0.017	–0.027
	(0.047)	(0.038)	(0.041)	(0.047)	(0.038)	(0.041)
	[0.483]	[0.330]	[0.435]	[0.249]	[0.326]	[0.254]
Panel C: Math Score						
Favoritism	–0.032	–0.033	–0.039	–0.027	–0.025	–0.030
	(0.052)	(0.045)	(0.055)	(0.052)	(0.045)	(0.055)
	[0.267]	[0.229]	[0.241]	[0.304]	[0.291]	[0.294]
Favoritism × Female	0.002	0.020	0.024	–0.016	–0.003	–0.001
	(0.047)	(0.039)	(0.042)	(0.047)	(0.039)	(0.042)
	[0.479]	[0.307]	[0.286]	[0.366]	[0.472]	[0.495]
Panel D: Science Score						
Favoritism	–0.075	–0.076	–0.076	–0.066	–0.063	–0.064
	(0.050)	(0.041)	(0.052)	(0.050)	(0.041)	(0.052)
	[0.066]	[0.030]	[0.069]	[0.093]	[0.059]	[0.107]
Favoritism × Female	0.046	0.062	0.057	0.020	0.033	0.026
	(0.049)	(0.039)	(0.041)	(0.049)	(0.039)	(0.041)
	[0.176]	[0.055]	[0.083]	[0.345]	[0.201]	[0.267]
Observations			2,323			
School Fixed Effect	✓	✓	✓	✓	✓	✓
Student Characteristics		✓	✓		✓	✓
Teacher Characteristics			✓			✓

Note: Bootstrap standard errors retrieved from 1,000 resamples within class clusters are reported in parentheses. P-values are reported in squared brackets. The standardized coefficients of β_i are retrieved from Equation (1). Student characteristics include gender, age, living with father/mother, parent marital status, parental ethnicity, parental education level, parental socioeconomic status, log family income, number of siblings, class rank in primary school, and the interaction terms of gender and these variables. Teacher characteristics include gender, teacher/student gender match, age, university/college graduate, possession of a graduate degree, experience (in years), marriage status, number of children, subjects taught, attitudes toward physical punishment, strictness, the role of a homeroom teacher, evening self-study session, cram school, high school admission policy, and the interaction terms of student gender and these variables.

Table 5: The Effect of Teachers' Ethnicity Favoritism on Student Achievement

	Baseline Approach			Leave-one-out Approach		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Total Score						
Favoritism	0.036	0.051	0.067	0.049	0.056	0.065
	(0.048)	(0.039)	(0.052)	(0.048)	(0.039)	(0.052)
	[0.223]	[0.094]	[0.099]	[0.153]	[0.075]	[0.105]
Favoritism × Mainlander	0.064	0.053	0.061	0.019	–0.011	–0.033
	(0.072)	(0.058)	(0.071)	(0.072)	(0.058)	(0.071)
	[0.187]	[0.178]	[0.193]	[0.398]	[0.425]	[0.323]
Panel B: Reading Score						
Favoritism	0.043	0.057	0.076	0.053	0.060	0.070
	(0.048)	(0.040)	(0.047)	(0.048)	(0.040)	(0.047)
	[0.185]	[0.078]	[0.053]	[0.137]	[0.065]	[0.067]
Favoritism × Mainlander	0.047	0.033	0.053	–0.002	–0.027	–0.042
	(0.076)	(0.065)	(0.083)	(0.076)	(0.065)	(0.083)
	[0.269]	[0.305]	[0.263]	[0.488]	[0.339]	[0.308]
Panel C: Math Score						
Favoritism	0.035	0.044	0.080	0.041	0.041	0.066
	(0.050)	(0.044)	(0.057)	(0.050)	(0.044)	(0.057)
	[0.238]	[0.157]	[0.080]	[0.205]	[0.172]	[0.124]
Favoritism × Mainlander	0.063	0.065	0.041	0.033	0.012	–0.030
	(0.072)	(0.061)	(0.074)	(0.072)	(0.061)	(0.074)
	[0.193]	[0.143]	[0.290]	[0.322]	[0.420]	[0.341]
Panel D: Science Score						
Favoritism	0.034	0.049	0.036	0.047	0.054	0.041
	(0.047)	(0.038)	(0.053)	(0.047)	(0.038)	(0.053)
	[0.235]	[0.102]	[0.251]	[0.159]	[0.080]	[0.221]
Favoritism × Mainlander	0.072	0.063	0.054	0.009	–0.019	–0.059
	(0.071)	(0.059)	(0.072)	(0.071)	(0.059)	(0.072)
	[0.155]	[0.144]	[0.228]	[0.448]	[0.372]	[0.207]
Observations			2,323			
School Fixed Effect	✓	✓	✓	✓	✓	✓
Student Characteristics		✓	✓		✓	✓
Teacher Characteristics			✓			✓

Note: Bootstrap standard errors retrieved from 1,000 resamples within class clusters are reported in parentheses. P-values are reported in squared brackets. The standardized coefficients of β_i are retrieved from Equation (1). Student characteristics include gender, age, living with father/mother, parent marriage status, parental ethnicity, parental education level, parental socioeconomic status, log family income, number of siblings, class rank in primary school, and the interaction terms of ethnicity and these variables. Teacher characteristics include gender, teacher/student gender match, age, university/college graduate, possession of a graduate degree, experience (in years), marriage status, number of children, subjects taught, attitudes toward physical punishment, strictness, the role of a homeroom teacher, evening self-study session, cram school, high school admission policy, and the interaction terms of student ethnicity and these variables.

5.2.3. Achievement Favoritism

Table 6 reports the estimated effect of teachers' achievement favoritism on the gap between high and low achievers, denoted by the coefficient of $Favoritism \times HighAchiever$. The baseline approach with full covariates (Column (3)) suggests that a one SD increase in favoritism leads to a 0.121 SD increase in the achievement gap. The leave-one-out approach halves the magnitude of the baseline approach, and the estimates become non-significant. Panels B to D present the results by subject. The impacts on reading and science are similar to the main results. However, the impact on math scores is relatively small. Still, none of the estimates from the leave-one-out approach is significant, providing no solid evidence of favoritism effect on achievement.

5.2.4. The Synergistic Effects of Multiple Favoritism

This section examines the synergistic effects of multiple types of favoritism. Figure 3a shows the estimated results from Equation (2) using the baseline approach. The outcome variable is a standardized value, and the independent variable is the interaction terms of the three dummies indicating whether students benefit from a specific type of favoritism. The symbol denotes the point estimates, and the error bars refer to a 95% CI generated with the bootstrap method.

Most estimates are noisy, potentially because the treatment variables are dummy variables. Especially when computing standard errors with the bootstrap method, resampled data might produce a different sign of favoritism if the true value is near zero.

Still, the point estimates show some interesting patterns. When students benefit from only one type of favoritism, they experience a small positive gap from the reference group (those without the benefit of teacher favoritism). In contrast, the point estimates of a dual favoritism combination are always larger than the sum of the independent effects. For example, the point estimates are 0.038 for gender favoritism and 0.035 for achievement favoritism, but the combined effect is 0.138 (greater than 0.038 + 0.035). Students exposed to all three types of favoritism experience a 0.185 SD increase in outcomes, while the sum of the three separate favoritism effects is only 0.143. However, the 95% CI of all estimates overlaps, indicating no evidence that the difference in point estimates differs from the null.

Figure 3b uses the leave-one-out approach. The impacts of single-type favoritism are negative and near zero. The estimates of favoritism combinations are larger than for one favoritism type. The results share the same pattern as the baseline approach, but there is no evidence that the effect of different types of favoritism combinations significantly differ from each other.

20

Table 6: The Effect of Teacher Achievement Favoritism on Student Achievement

	Baseline Approach			Leave-one-out Approach		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Total Score						
Favoritism	−0.018	−0.017	−0.051	0.005	0.003	−0.023
	(0.045)	(0.044)	(0.059)	(0.045)	(0.044)	(0.059)
	[0.343]	[0.346]	[0.192]	[0.452]	[0.472]	[0.350]
Favoritism × Middle Achiever	0.028	0.036	0.038	0.003	0.012	0.008
	(0.048)	(0.046)	(0.053)	(0.048)	(0.046)	(0.053)
	[0.279]	[0.217]	[0.236]	[0.478]	[0.399]	[0.443]
Favoritism × HighAchiever	0.111	0.109	0.121	0.053	0.053	0.056
	(0.063)	(0.059)	(0.068)	(0.063)	(0.059)	(0.068)
	[0.040]	[0.032]	[0.038]	[0.200]	[0.184]	[0.207]
Panel B: Reading Score						
Favoritism	−0.061	−0.048	−0.062	−0.029	−0.022	−0.026
	(0.050)	(0.049)	(0.061)	(0.050)	(0.049)	(0.061)
	[0.111]	[0.164]	[0.155]	[0.279]	[0.330]	[0.333]
Favoritism × Middle Achiever	0.061	0.056	0.060	0.028	0.026	0.020
	(0.053)	(0.052)	(0.061)	(0.053)	(0.052)	(0.061)
	[0.128]	[0.138]	[0.164]	[0.299]	[0.309]	[0.370]
Favoritism × HighAchiever	0.124	0.113	0.128	0.061	0.053	0.056
	(0.067)	(0.062)	(0.073)	(0.067)	(0.062)	(0.073)
	[0.032]	[0.035]	[0.041]	[0.180]	[0.195]	[0.224]
Panel C: Math Score						
Favoritism	0.001	0.004	−0.034	0.013	0.010	−0.023
	(0.046)	(0.046)	(0.061)	(0.046)	(0.046)	(0.061)
	[0.487]	[0.462]	[0.292]	[0.388]	[0.416]	[0.353]
Favoritism × Middle Achiever	0.007	0.013	0.008	−0.007	0.004	−0.004
	(0.048)	(0.048)	(0.055)	(0.048)	(0.048)	(0.055)
	[0.441]	[0.397]	[0.444]	[0.445]	[0.471]	[0.473]
Favoritism × HighAchiever	0.042	0.031	0.020	0.007	0.001	−0.015
	(0.065)	(0.061)	(0.068)	(0.065)	(0.061)	(0.068)
	[0.260]	[0.307]	[0.384]	[0.459]	[0.496]	[0.415]
Panel D: Science Score						
Favoritism	−0.021	−0.010	−0.024	−0.001	0.004	−0.008
	(0.049)	(0.047)	(0.063)	(0.048)	(0.047)	(0.063)
	[0.332]	[0.416]	[0.353]	[0.491]	[0.463]	[0.447]
Favoritism × Middle Achiever	0.039	0.036	0.027	0.017	0.018	0.005
	(0.051)	(0.049)	(0.056)	(0.051)	(0.049)	(0.056)
	[0.225]	[0.230]	[0.311]	[0.372]	[0.357]	[0.464]
Favoritism × HighAchiever	0.116	0.105	0.113	0.062	0.053	0.055
	(0.069)	(0.066)	(0.075)	(0.069)	(0.066)	(0.075)
	[0.046]	[0.056]	[0.065]	[0.185]	[0.210]	[0.230]
Observations			2,323			
School Fixed Effect	✓	✓	✓	✓	✓	✓
Student Characteristics		✓	✓		✓	✓
Teacher Characteristics			✓			✓

Note: Bootstrap standard errors retrieved from 1,000 resamples within class clusters are reported in parentheses. P-values are reported in squared brackets. The standardized coefficients of β_i are retrieved from Equation (1). Student characteristics include gender, age, living with father/mother, parent marriage status, parental ethnicity, parental education level, parental socioeconomic status, log family income, number of siblings, class rank in primary school, and the interaction terms of previous achievement and these variables. Teacher characteristics include gender, teacher/student gender match, age, university/college graduate, possession of a graduate degree, experience (in years), marriage status, number of children, subjects taught, attitudes toward physical punishment, strictness, the role of a homeroom teacher, evening self-study session, cram school, high school admission policy, and the interaction terms of student achievement and these variables.
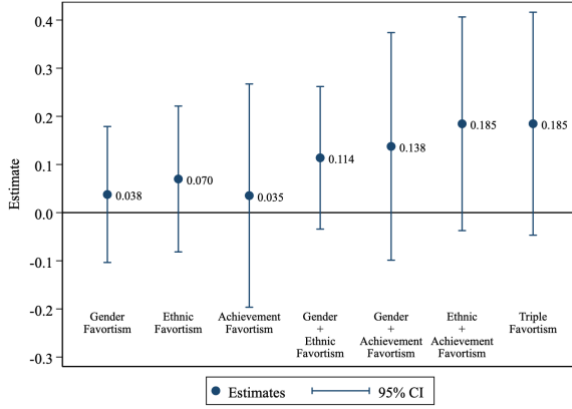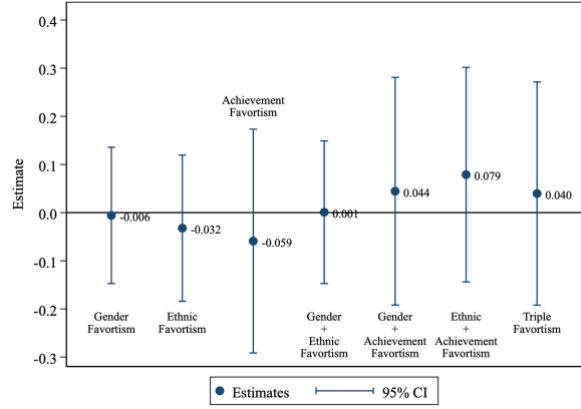
(a) Baseline Approach	(b) Leave-one-out Approach

Figure 3: The Synergistic Effects of Multiple Favoritism

Note: This figure reports the coefficients of γ retrieved from Equation (2). The outcomes variable is a standardized value. The symbol denotes the point estimates, and the error bars refer to the 95% confidence interval. Standard errors are based on the Bootstrap method retrieved from 1,000 times resampling within class clusters reported in parentheses.

5.3 Placebo Tests

This section discusses a series of placebo tests conducted to ensure that the effects identified in the previous sections do not occur by chance. The central identification strategy is based on random class assignment; hence, I calculate the permutation p-values by reshuffling homeroom teachers to different classes and then re-estimating the main equations, repeated 1,000 times. As the randomly assigned teachers do not teach the classes in this permutation test, the expected value of the estimated effect should be zero.

Figure B1 in Appendix B demonstrates the distribution of the placebo estimates for the outcome variables regarding the teacher-student relationship. For all models, the estimates have a distribution centered at zero, with negative estimates ranging from 46% to 53%. Figure B1b shows the result of the gender favoritism effect on a student's frequency of consulting with the homeroom teacher on emotional issues. Among the placebo estimates, approximately 0.1% are greater than the baseline estimate, and 2.3% are greater than the leave-one-out approach estimate. Figure B1e illustrates the placebo tests for the achievement favoritism effect on students consulting with homeroom teachers on academic issues. Among the placebo estimates, only 0.2% are higher than the baseline estimate, and 2.5% are greater than the leave-one-out approach estimate. Figure B1f shows the result for the achievement favoritism effect on students consulting on emotional issues. The permutation p-value is 0 for the baseline approach and 0.011 for the leave-one-out approach.

These permutation tests confirm that the main results are not likely due to chance.

Figure B2 in Appendix B displays a similar analysis for achievement outcomes. The placebo estimates still have a distribution centering at zero. Figure B2d shows the permutation test for the achievement favoritism effect on students' total scores. Of the placebo estimates, 1.2% are greater than the baseline estimate, and 14.6% are higher than the leave-one-out approach estimate. The statistical inference from the permutation test is aligned with the main results.

Figure B3 in Appendix B depicts another version of the placebo test using favoritism based on one criterion to predict the achievement gap of another group category. The three horizontal rows display the impacts of gender, ethnicity, and achievement favoritism, respectively. The three vertical columns represent impacts on the three types of achievement gaps. The shaded area refers to the favoritism type aligning with the group variable relevant to the achievement gap. These cases should yield a positive effect, and as expected, the point estimates are positive when favoritism and the achievement gap are based on the same variable.

Conversely, the unshaded cells show that the favoritism type is unrelated to the achievement gap group. If the positive estimates in the shaded cells are merely driven by "the random correlation between teacher favorability and student traits" or "some teachers are more likely to generate larger achievement gaps," the placebo estimates would produce positive effects. However, the estimates are negative or near zero throughout all placebo cells.

6 Discussion and Conclusion

This study investigates how homeroom teachers' favoritism based on gender, ethnicity, and achievement affects the disparity in teacher treatment and the student achievement gap among junior high school students in Taiwan. Based on the heterogeneous class grouping practice, this paper identifies the causal effect of favoritism using within-school variations.

The evidence of the teacher favoritism effect varies by type of favoritism. Gender and achievement favoritism both raise the likelihood that the favored group of students to reach out to their teacher when encountering emotional or friendship issues. Students who benefit from achievement favoritism are also more likely to consult with homeroom teachers regarding academic issues.

There is some evidence regarding the favoritism effect on the achievement gap when using the baseline approach. Achievement favoritism exacerbates the achievement gap in grade nine

standardized examinations among low- and high-achievers in all subjects except math. Gender favoritism only yields a significant effect on magnifying the science gender achievement gap. This aligns with previous studies, which find that gender bias produces the largest effect in study fields with gender stereotypes [15,17,48,49]. The effect size of gender favoritism is at the mid-point among the results from previous studies on gender bias [15–17]. There is no consistent evidence indicating that ethnic favoritism increases the achievement gap. The point estimate is near the lower bound of relevant previous studies regarding ethnicity favoritism [11,12,18].

However, the leave-one-out approach provides no evidence supporting the effect of favoritism on achievement. The reason for this might be that as the key variable (favoritism) in this paper is calculated within classes (which are generally small in size), a leave-one-out approach with the two-stage bootstrap process produces some degree of downward bias and significant noisiness. Still, the fact that the leave-one-out approach halved the main estimates and led to insignificant results also suggests that the actual effect of favoritism on achievement is not prominent.

The findings of this study provide three conclusions. First, the effects of favoritism can be linked to a social context. The evidence for ethnicity favoritism is the weakest, which aligns with the cultural context in Taiwan, as ethnic inequalities have gradually decreased. In contrast, the most potent effect on achievement favoritism reflects the pro-educational cultural values of Asian countries. The result that gender favoritism has a greater impact on teacher-student relationships, especially regarding emotional rather than academic consultations, also aligns with the general stereotype of connecting specific gender roles with a caring personality.

Secondly, compared to the Pygmalion theory, the teacher-student relationship theory better explains how teacher favoritism could affect students. Female and high-achieving students experience better interactions with teachers when assigned to pro-female (pro-high achiever) teachers. These students are more willing to seek help when facing difficulties than those in unfavored groups. In contrast, while the baseline approach suggests that favoritism leads to some extent of biased teacher expectations, this does not transfer to students' self-expectations.

Third, even when teacher favoritism is not related to grading and achievement, it does not mean it produces no negative impacts. Previous studies on teacher bias concentrate mainly on teacher "grading bias." However, this study finds that the effects of teacher favoritism on teacher-student relationships exist even when there is little evidence of the impact on the expectation and

achievement gap. The result implies that expressions of teacher favoritism beyond grading bias deserve more attention from researchers and educators.

There are several limitations to these findings. First, this study cannot rule out the confounders from unobserved student characteristics that correlate with teacher favorability and student achievement. The achievement gap's positive impact could be attributed to teachers tending to favor students with more potential. While this study adopts the leave-one-out approach to avoid reverse causality from *individual* characteristics, unobserved *classmate* characteristics may still drive measurement and bias the results. Furthermore, the fact that the leave-one-out approach suggests no detectable effect on the achievement gap implies that the baseline estimates still could not rule out reverse causality. Secondly, while the primary outcome—ninth-grade student achievement—is objectively reported, other intermediate outcomes are self-reported, which may cause bias. For example, if teachers grant the same support to all students but are more impressed when interacting with their favorites, they might report a better relationship with those students. Finally, as discussed in Section 2.2, the cultural context in Taiwan may differ from other countries, impacting generalizability. Still, as the findings align with the cultural background in Taiwan, this implies that the favoritism impact might be more pronounced when interacting with societal context.

Data and Code Availability

The data employed in this paper is available at Survey Research Data Archive (SRDA), Academia Sinica (https://srda.sinica.edu.tw). The codes for data construction and analyses are available at: https://github.com/yungyutsai/teacher_favoritism.

Reference

[1]   M.H. Roy, F.C. Roy, An empirical analysis of favoritism during business training, Industrial and Commercial Training. 36 (2004) 238–242. https://doi.org/10.1108/00197850410556676.

[2]   A. Ali, D.M. Khan, M. Hussain, Causes of teacher's favouritism and its effects on the university students: A case study, Global Social Sciences Review. 3 (2018) 369–384. http://dx.doi.org/10.31703/gssr.2018(III-II).21.

[3]   U. Quraishi, H. Munir, A. Afzal, M. Saeed, A study on teachers' favoritism and its effects on students' academic learning at university level in Multan district, Psychology and Education Journal. 58 (2021) 2304–2311. https://doi.org/10.17762/pae.v58i3.4243.

[4]   C. Einarsson, K. Granström, Gender-biased interaction in the classroom: The influence of gender and age in the relationship between teacher and pupil, Scandinavian Journal of Educational Research. 46 (2002) 117–127. https://doi.org/10.1080/00313830220142155.

[5]   S. Wang, C.M. Rubie-Davies, K. Meissel, The stability and trajectories of teacher expectations: Student achievement level as a moderator, Learning and Individual Differences. 78 (2020) 101819. https://doi.org/10.1016/j.lindif.2019.101819.

[6]   D.N. Figlio, Names, expectations and the black-white test score gap, NBER Working Paper. (2005). https://doi.org/10.3386/w11195.

[7]   M. Sprietsma, Discrimination in grading: Experimental evidence from primary school teachers, Empirical Economics. 45 (2013) 523–538. https://doi.org/10.1007/s00181-012-0609-x.

[8]   E. Babad, Measuring and changing teachers' differential behavior as perceived by students and teachers, Journal of Educational Psychology. 82 (1990) 683. https://doi.org/10.1037/0022-0663.82.4.683.

[9]   D. Siegle, T. Powell, Exploring teacher biases when nominating students for gifted programs, Gifted Child Quarterly. 48 (2004) 21–29. https://doi.org/10.1177/001698620404800103.

[10]   M. Zhu, D. Urhahne, C.M. Rubie-Davies, The longitudinal effects of teacher judgement and different teacher treatment on students' academic outcomes, Educational Psychology. 38 (2018) 648–668. https://doi.org/10.1080/01443410.2017.1412399.

[11]   L. Van den Bergh, E. Denessen, L. Hornstra, M. Voeten, R.W. Holland, The implicit prejudiced attitudes of teachers: Relations to teacher expectations and the ethnic achievement gap, American Educational Research Journal. 47 (2010) 497–527. https://doi.org/10.3102/0002831209353594.

[12]   E.R. Peterson, C. Rubie-Davies, D. Osborne, C. Sibley, Teachers' explicit expectations and implicit prejudiced attitudes to educational achievement: Relations with student achievement and the ethnic achievement gap, Learning and Instruction. 42 (2016) 123–140. https://doi.org/10.1016/j.learninstruc.2016.01.010.

[13] V. Lavy, R. Megalokonomou, Persistency in teachers' grading bias and effects on longer-term outcomes: University admissions exams and choice of field of study, NBER Working Paper. (2019). http://doi.org/10.3386/w26021.

[14] D.M. Merolla, O. Jackson, Structural racism as the fundamental cause of the academic achievement gap, Sociology Compass. 13 (2019) e12696. https://doi.org/10.1111/soc4.12696.

[15] C. Terrier, Boys lag behind: How teachers' gender biases affect student achievement, Economics of Education Review. 77 (2020) 101981. http://dx.doi.org/10.2139/ssrn.2868309.

[16] V. Lavy, E. Sand, On the origins of gender human capital gaps: Short and long term consequences of teachers' stereotypical biases, Journal of Public Economics. 167 (2018) 263–279. https://doi.org/10.1016/j.jpubeco.2018.09.007.

[17] M. Carlana, Implicit stereotypes: Evidence from teachers' gender bias, The Quarterly Journal of Economics. 134 (2019) 1163–1224. https://doi.org/10.1093/qje/qjz008.

[18] H. Cherng, If they think I can: teacher bias and youth of color expectations and achievement, Social Science Research. 66 (2017) 170–186. http://dx.doi.org/10.1016/j.ssresearch.2017.04.001.

[19] L. Hornstra, E. Denessen, J. Bakker, L. Van Den Bergh, M. Voeten, Teacher attitudes toward dyslexia: Effects on teacher expectations and the academic achievement of students with dyslexia, Journal of Learning Disabilities. 43 (2010) 515–529. https://doi.org/10.1177/0022219409355479.

[20] R. Van Ewijk, Same work, lower grade? Student ethnicity and teachers' subjective assessments, Economics of Education Review. 30 (2011) 1045–1058. https://doi.org/10.1016/j.econedurev.2011.05.008.

[21] S. Alan, E. Duysak, E. Kubilay, I. Mumcu, Social Exclusion and Ethnic Segregation in Schools: The Role of Teacher's Ethnic Prejudice, The Review of Economics and Statistics. forthcoming (2021) 1–45. https://doi.org/10.1162/rest_a_01111.

[22] H.H. Davidson, G. Lang, Children's perceptions of their teachers' feelings toward them related to self-perception, school achievement and behavior, The Journal of Experimental Education. 29 (1960) 107–118. https://doi.org/10.1080/00220973.1960.11010675.

[23] H.A. Davis, A.M. Lease, Perceived organizational structure for teacher liking: The role of peers' perceptions of teacher liking in teacher–student relationship quality, motivation, and achievement, Social Psychology of Education. 10 (2007) 403–427. https://doi.org/10.1007/s11218-007-9031-1.

[24] N. Ambady, M. Shih, A. Kim, T.L. Pittinsky, Stereotype susceptibility in children: Effects of identity activation on quantitative performance, Psychological Science. 12 (2001) 385–390. https://doi.org/10.1111/1467-9280.00371.

[25] Y.-H. Luo, K.-H. Chen, Education expansion and its effects on gender gaps in educational attainment and political knowledge in Taiwan from 1992 to 2012, International Journal of Educational Development. 60 (2018) 88–99. https://doi.org/10.1016/j.ijedudev.2017.10.014.

[26] D.-F. Chang, Effects of higher education expansion on gender parity: a 65-year trajectory in Taiwan, Higher Education. 76 (2018) 449–466. https://doi.org/10.1007/s10734-017-0219-9.

[27] K. Yu-Chen, S.-J. Sheu, The impact of mathematics background and gender on the choice of major in Taiwan's senior high school, Taiwan Economic Review. 39 (2011) 541. http://doi.org/10.6277/ter.2011.394.4.

[28] E.M. Evans, H. Schweingruber, H.W. Stevenson, Gender differences in interest and knowledge acquisition: The United States, Taiwan, and Japan, Sex Roles. 47 (2002) 153–167. https://doi.org/10.1023/A:1021047122532.

[29] OECD, The ABC of Gender Equality in Education Aptitude, Behaviour, Confidence, Organisation for Economic Cooperation and Development, 2015. https://doi.org/10.1787/9789264229945-en.

[30] OECD, Girls' and boys' performance in PISA, in: PISA 2018 Results (Volume II) : Where All Students Can Succeed, Organisation for Economic Cooperation and Development, 2018. https://www.oecd-ilibrary.org/sites/f56f8c26-en/index.html?itemId=/content/component/f56f8c26-en.

[31] IEA, TIMSS 2019 International Results in Mathematics and Science, International Association for the Evaluation of Educational Achievement, 2019. https://timss2019.org/reports/trends-in-achievement-by-gender-s8/.

[32] K.-H. Su, W.-H. Yu, When social reproduction fails: Explaining the decreasing ethnic gap in Taiwan, Taiwanese Journal of Sociology. 39 (2007) 1–63. https://www.tsatw.org.tw/page.php?menu_id=82\&new_id=1062.

[33] T.-S. Lin, The effects of family background on tracking of secondary education in Taiwan—A study of the distinction between "academic/vocational" and "public/private" tracking, Soochow Journal of Sociology. 8 (1999) 35–77. http://fguir.fgu.edu.tw:8080/handle/039871000/5588.

[34] A. Liu, Y. Xie, Why do Asian Americans academically outperform Whites?—The cultural explanation revisited, Social Science Research. 58 (2016) 210–226. https://doi.org/10.1016/j.ssresearch.2016.03.004.

[35] K.-C. Liu, Diplomaism, school campus life, and cram schooling: A group of seventh grade student's curriculum perspectives, Journal of Education Studies. 47 (2013) 73–98. https://doi.org/10.3966/199044282013104702004.

[36] H.-C. Yen, Education and social order: An analysis of the entrance examination, Journal of Educational Practice and Research. 27 (2014) 121–144. https://academic.ntue.edu.tw/ezfiles/7/1007/img/41/27-1-5.pdf.

[37] C.-C. Yi, Taiwan Youth Project: J1W1S (Public Access Data) (C00176_4) [data file], Available from Survey Research Data Archive, Academia Sinica, 2014. https://doi.org/10.6141/TW-SRDA-C00176_4-1.

[38] J.A. Hutcheon, A. Chiolero, J.A. Hanley, Random measurement error and regression dilution bias, Bmj. 340 (2010). https://doi.org/10.1136/bmj.c2289

[39] A. Cools, R. Fernández, E. Patacchini, Girls, boys, and high achievers, National Bureau of Economic Research, 2019. https://doi.org/10.3386/w25763.

[40] M.K. Humlum, M. Thorsager, The importance of peer quality for completion of higher education, Economics of Education Review. 83 (2021) 102120. https://doi.org/10.1016/j.econedurev.2021.102120.

[41] D. Kalogrides, S. Loeb, Different teachers, different peers: The magnitude of student sorting within schools, Educational Researcher. 42 (2013) 304–316. https://doi.org/10.3102/0013189X13495087.

[42] D. Kalogrides, S. Loeb, T. Béteille, Systematic sorting: Teacher characteristics and class assignments, Sociology of Education. 86 (2013) 103–123. https://doi.org/10.1177/0038040712456555.

[43] D. Goldhaber, L. Lavery, R. Theobald, Uneven playing field? Assessing the teacher quality gap between advantaged and disadvantaged students, Educational Researcher. 44 (2015) 293–307. https://doi.org/10.3102/0013189X15592622.

[44] H.B.G. Ganzeboom, D.J. Treiman, International stratification and mobility file: Conversion tools, (2019). http://www.harryganzeboom.nl/ismf/.

[45] R.K. Merton, The self-fulfilling prophecy, The Antioch Review. 8 (1948) 193–210. https://doi.org/10.2307/4609267.

[46] A. Friedrich, B. Flunger, B. Nagengast, K. Jonkmann, U. Trautwein, Pygmalion effects in the classroom: Teacher expectancy effects on students' math achievement, Contemporary Educational Psychology. 41 (2015) 1–12. https://doi.org/10.1016/j.cedpsych.2014.10.006.

[47] J.E. Brophy, T.L. Good, Teachers' communication of differential expectations for children's classroom performance: Some behavioral data, Journal of Educational Psychology. 61 (1970) 365. https://doi.org/10.1037/h0029908.

[48] M. Nürnberger, J. Nerb, F. Schmitz, J. Keller, S. Sütterlin, Implicit gender stereotypes and essentialist beliefs predict preservice teachers' tracking recommendations, The Journal of Experimental Education. 84 (2016) 152–174. https://doi.org/10.1080/00220973.2015.1027807.

[49] D. Xu, Q. Li, Gender achievement gaps among Chinese middle school students and the role of teachers' gender, Economics of Education Review. 67 (2018) 82–93. https://doi.org/10.1016/j.econedurev.2018.10.002.

A Questionnaire and Measurement

Table A1: Variables and Questionnaire Wording

Variable	Respondent	Survey Year (Student Age)	Coding	Question Item(s)
Panel A: Independent Variable				
Gender	Student	2000 (13)		Gender: Male/Female
Ethnicity	Student	2000 (13)	Coded student as a Mainlander if either father or mother is a Mainlander	What is your father's (mother's) ethnicity? (Holo/Hakka/Mainlander/Indigenous/Other)
Previous achievement	Student	2000 (13)	Coded top 5 as high achievers, after 21 as low achievers, and others as middle achievers	What were your grades when you graduated from primary school? (Class rank: top 5/6–10/11–20/after 21)
Teacher favorability towards students	Teacher	2000 (13)	1–4 coding (4 as strongly likes)	How is the student getting along with you? Does your relationship fit the following description? I like him/her (Strongly agree to strongly disagree)
Panel B: Outcome Variable (Achievement)				
Test scores of standardized exams	Teacher	2002 (15)	Covers five subjects (reading (Mandarin), English, math, science, and social science), each graded on a 0–60-point scale	What are the scores for this student's Basic Competency Test: Reading/English/Math/Science/Social Science/Total (Please provide a copy of the class exam results)
Panel C: Intermediate Outcome (Mechanism Variable)				
Teacher expectation				
Attend high school	Teacher	2000 (13)	Coded as 1 if selected upper secondary high school, 0 otherwise	Based on your understanding of this student, which path is most suitable for him/her upon graduating from lower secondary school? (Senior high school/vocational school/vocational college, not suitable for remaining in the educational system)
Expected Years of Education	Teacher	2000 (13)	Coded as the number of years of education	If there are no restrictions, what level of education do you think this student should be able to reach? (Junior high school/senior high school/associated degree/bachelor's degree/master's degree/doctoral degree)
Student Self-expectation				
Attend high school	Student	Difference in 2000 (13) and 2002 (15)	Coded as 1 if selected upper secondary high school, 0 otherwise	What is your goal upon graduation? (Go to upper secondary high school/go to vocational school or college/work full-time/become an apprentice/other)
Expected Years of Education	Student	Difference in 2000 (13) and 2002 (15)	Coded as the number of years of education	Given your current environment and abilities, what level of education do you think you can reach? (Lower secondary school/upper secondary school/associated degree/bachelor's degree/master's degree/doctoral degree)
Seek help from/consult with teacher:				
School life difficulties	Student	2002 (15)	Coded as 1 if marked homeroom teacher, 0 otherwise.	When you encounter difficulties in school life, who do you most often turn to for help? (Classmates/Homeroom teacher/School counselors /Parents/No one/Others)
Academic	Teacher	2000 (13)	Coded as 1–4 (4 as strongly agree)	Does this student get along with you? Do you think your relationship fits the following description? (Strongly agree to strongly disagree) He/She will talk to me if he/she has a problem with coursework.
Emotional	Teacher	2000 (13)	Coded as 1–4 (4 as strongly agree)	He/She will talk to me if he/she has emotional troubles.
Friendship	Teacher	2000 (13)	Coded as 1–4 (4 as strongly agree)	He/She will seek help from me if he/she has friendship troubles.

Note: The data of this study comes from the Taiwan Youth Project (TYP) [45]. The targeted sample is grade 7 students in Taipei City, Taipei County, and Yilan County, Taiwan, in 2000.

Table A2: Factor Analysis of Students' Personality

Questionnaire Items	Factor Loading	Cronbach's α
Factor 1: Positive Virtues		0.843
Patient	0.358	
Responsible for work	0.320	
Able to care for myself	0.457	
Friendly to classmates, friends, and other children	0.621	
Stand in other's shoes	0.697	
Get along with others	0.730	
Humble	0.636	
Keep promises	0.598	
Factor 2: Outgoing and Egocentric		0.823
I like to tell jokes or funny stories	0.477	
I like to decide what I want to do	0.455	
I like to share things with my friends	0.499	
I am willing to follow tradition so that people I care about do not think I am unconventional	0.413	
I want to go wherever I love to go	0.428	
I like to be in the spotlight of a group	0.586	
I like to be free, do whatever I like	0.409	
I like to follow instructions and do what others want me to do	0.414	
I like to be close with my friends	0.505	
I would rather do things with friends than do them alone	0.463	
I like to join a warm and friendly group	0.313	
When I plan, I hope to get opinions from people I respect	0.310	
When I am in public, I like people to pay attention to my appearance	0.598	
I like to say things that others think are smart and witty	0.556	
When deciding on the actions of the group, I am willing to accept the leadership of others	0.389	
I like to do things my way, regardless of what others think	0.367	
Factor 3: Obedience and Manners		0.832
Respect and obey parents	0.499	
Keep clothes clean	0.413	
Be polite to adults	0.635	
Obey parents and teachers	0.617	
Keep things neat	0.469	
Be a good student	0.577	
Behave appropriately	0.531	
Do not be pushy	0.376	
Living frugally	0.450	
Factor 4: Curious and Creative		0.694
Curious	0.531	
Imaginative	0.520	
Likes to explore why things happen	0.539	
Factor 5: Motivated		0.771
Attending college is important	0.600	
Working hard to get good grades is important	0.633	

Note: This data comes from the Taiwan Youth Project (TYP) [45]. The personality variables are derived from factor analysis based on the 38 items in the table and categorized into five latent factors.
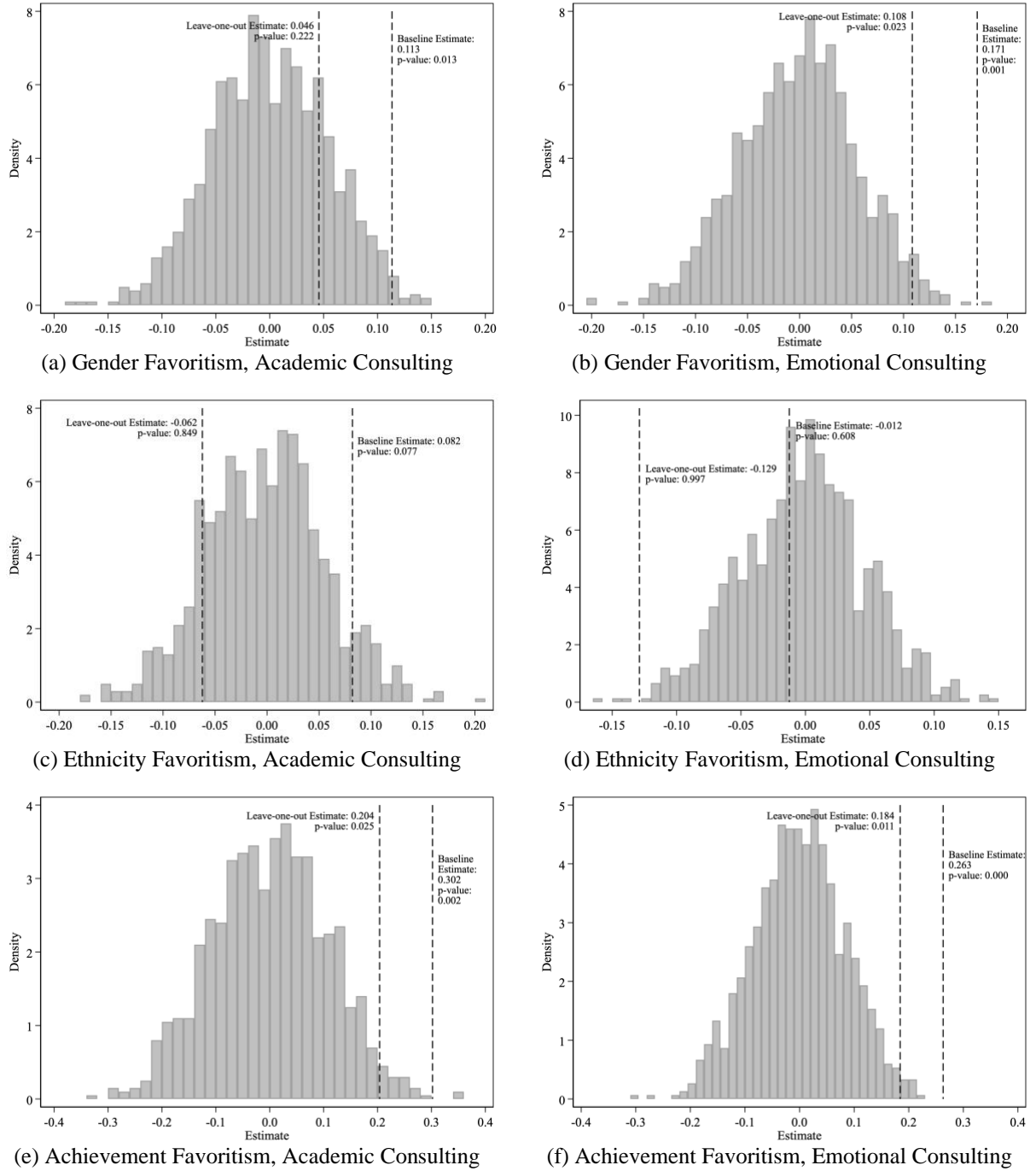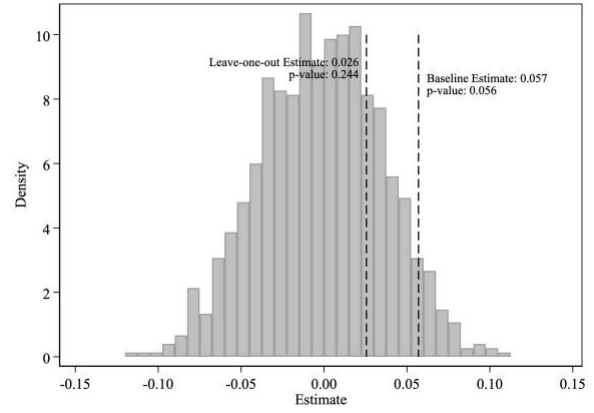
B   Placebo Tests

(a) Gender Favoritism, Academic Consulting
(b) Gender Favoritism, Emotional Consulting
(c) Ethnicity Favoritism, Academic Consulting
(d) Ethnicity Favoritism, Emotional Consulting
(e) Achievement Favoritism, Academic Consulting
(f) Achievement Favoritism, Emotional Consulting

Figure B1: Permutation Test: Teacher-Student Relationship

Note: This figure shows the distribution of the placebo estimates of β_2 from equation (1) from a 1,000 times permutation test of random reshuffling of teachers to classes. The histogram stands for the placebo estimates. The vertical line denotes the actual estimates from the baseline and leave-one-out approaches. The permutation p-value is calculated as the proportion of the placebo estimate greater than the actual estimate.
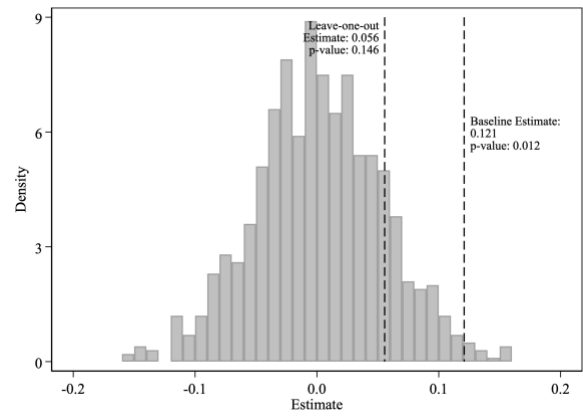
33
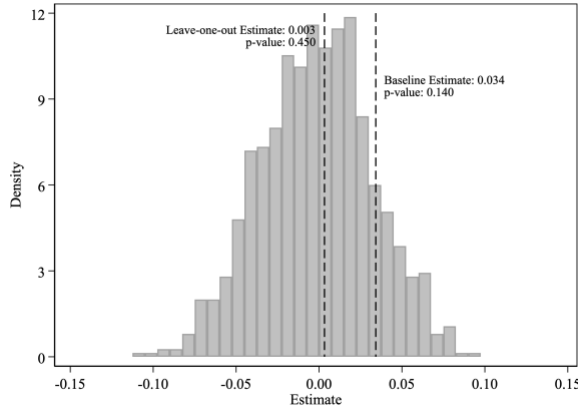
(a) Gender Favoritism, Total Scores

(b) Gender Favoritism, Science Scores
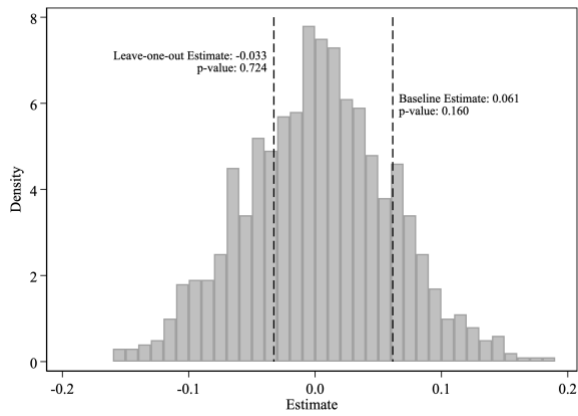
(c) Ethnicity Favoritism, Total Scores

(d) Achievement Favoritism, Total Scores

Figure B2: Permutation Test: Student Achievement

Note: This figure shows the distribution of the placebo estimates of β_2 from equation (1) from a 1,000 times permutation test of random reshuffling of teachers to classes. The histogram stands for the placebo estimates. The vertical line denotes the actual estimates from the baseline and leave-one-out approaches. The permutation p-value is calculated as the proportion of the placebo estimate greater than the actual estimate.
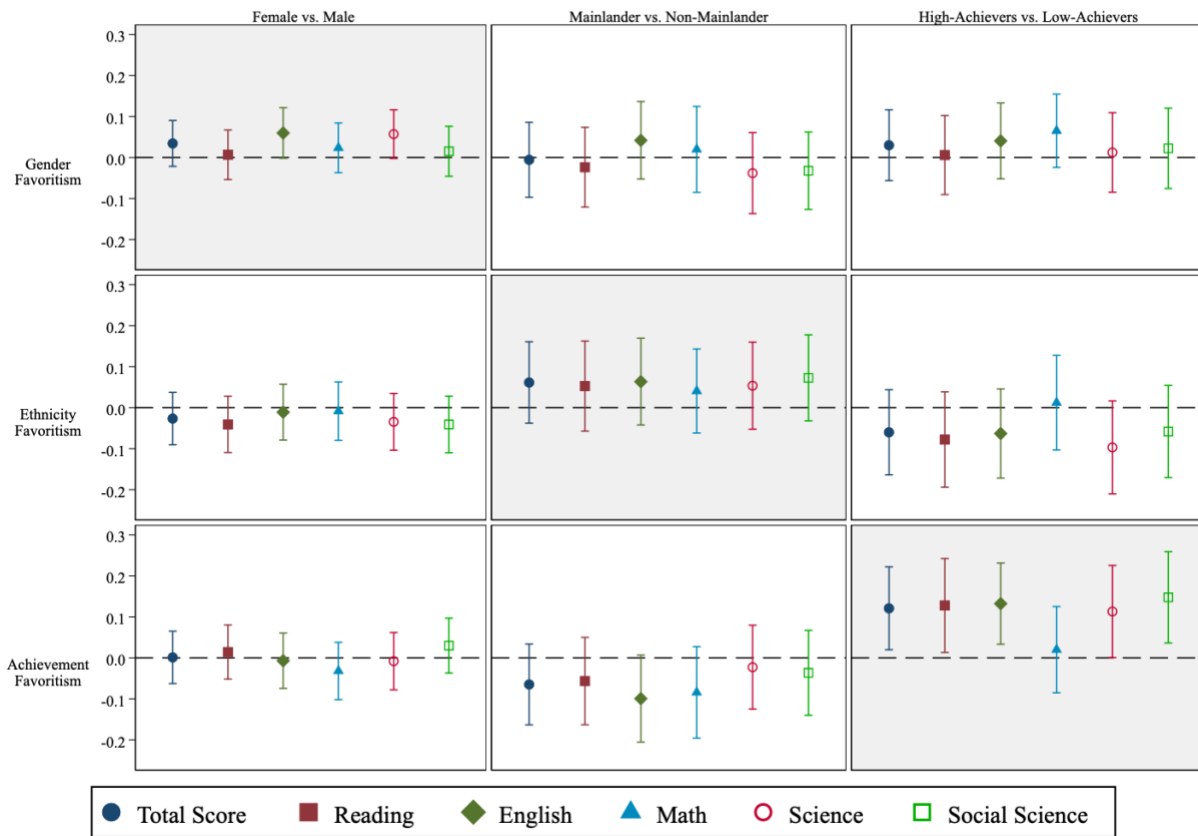
Figure B3: Placebo Test: Using Favoritism Based on One Criteria to Predict the Achievement

Gap in Another

Note: This figure shows a placebo test using gender, ethnicity, and achievement favoritism to predict the achievement gap. The horizontal rows represent impacts of gender, ethnicity, and achievement favoritism, respectively. The vertical columns represent impacts on the achievement gap relevant to gender, ethnicity, and previous achievement. The symbols denote the estimated effect on the achievement gap between the favored and unfavored groups. The error bars indicate a 95% CI. The shaded area is the type of favoritism matched with the category of the achievement gap, where positive estimates should be expected.