

Online Appendix

Gender Peer Effects on Students' Academic and Noncognitive Outcomes: Evidence and Mechanisms

Jie Gong, Yi Lu, and Hong Song

Appendix Table 1: Gender Peer Effect on Academic Self-assessment Scores

	(1)	(2)	(3)	(4)
Proportion female peers	0.325 (0.205)	0.153 (0.183)	0.134 (0.181)	0.136 (0.179)
Female student	0.179*** (0.019)	0.188*** (0.019)	0.188*** (0.019)	0.187*** (0.019)
Subject FE	Yes	Yes	Yes	Yes
School-grade FE	Yes	Yes	Yes	Yes
Student controls	No	Yes	Yes	Yes
Teacher controls	No	No	Yes	Yes
Peer ability controls	No	No	No	Yes
Observations	22,914	22,914	22,914	22,914
R-squared	0.009	0.049	0.050	0.050

Notes: The dependent variable is students' self-assessment of academic performance. Specifically, students were asked to report whether they have difficulty in learning each subject on a scale from 1 (*a lot*) to 4 (*not at all*). The rating is normalized by subject, grade, and school to obtain a mean of zero and standard deviation of one. Student controls include student's age, baseline noncognitive measurements, mother's education, father's education, and dummy variables indicating minority, local residence, rural residence, only child in family, attended kindergarten, repeated a grade, and skipped a grade in primary school. Teacher controls include gender, age, years of schooling, experience, professional job title, and dummy variables indicating marital status and graduated from a normal college. Peer ability controls include male and female peers' average baseline academic performance in primary school, including whether repeated a grade and whether skipped a grade. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

Appendix Table 2. Robustness Check: Gender Peer Effects on Individual Noncognitive Factors

	Mental stress				Social acclimation and general satisfaction			
	Depressed	Blue	Unhappy	Pessimistic	Fulfilling life	Confident abt future	Public enrichment	Private recreation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Proportion female peers	0.281 (0.260)	0.038 (0.234)	-0.042 (0.244)	-0.100 (0.277)	0.537** (0.267)	-0.072 (0.178)	0.341 (0.288)	0.938*** (0.236)
Female student	0.156*** (0.026)	-0.031 (0.025)	0.042* (0.024)	-0.073*** (0.022)	0.092*** (0.025)	-0.030 (0.023)	0.048** (0.021)	0.028 (0.022)
Subject FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School-grade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Peer ability controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,616	7,616	7,616	7,616	7,418	7,418	7,418	7,418
R-squared	0.079	0.075	0.084	0.060	0.098	0.139	0.238	0.239

Notes: Dependent variables are normalized to have a mean of zero and standard deviation of one. Student controls include student age, baseline noncognitive measurements, mother's education, father's education, and dummy variables indicating minority, local residence, rural residence, only child in family, attended kindergarten, and repeated a grade in primary school. Teacher controls include gender, age, years of schooling, experience, professional job title, and dummy variables indicating marital status and graduated from a normal college. Peer ability controls include male and female peers' average baseline academic performance in primary school, including whether repeated a grade and whether skipped a grade. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

Appendix Table 3. Comparison of Student Characteristics by Gender

Predetermined characteristics:	Male			Female			Difference	
	Mean	SD	Obs	Mean	SD	Obs	Difference	<i>p</i> -value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age	13.99	1.35	4462	13.88	1.34	4280	0.11	0.000***
Minority	0.11	0.31	4558	0.11	0.32	4333	0	0.367
Local residence	0.8	0.4	4487	0.81	0.39	4248	-0.01	0.306
Only child	0.54	0.5	4570	0.48	0.5	4338	0.06	0.000***
Preschool attendance	0.81	0.39	4519	0.82	0.38	4316	-0.01	0.151
Repeater	0.13	0.34	4572	0.1	0.29	4338	0.03	0.000***
Skip grade	0.02	0.14	4554	0.01	0.11	4335	0.01	0.006***
Pre noncognitive measure	3.08	0.69	4277	3.05	0.65	4138	0.03	0.04**
Mother college attendance	0.18	0.38	4572	0.18	0.39	4338	-0.17	0.306
Father college attendance	0.2	0.4	4572	0.22	0.41	4338	-0.02	0.033**

Notes: This table presents summary statistics for predetermined characteristics for male and female students separately. Column 7 presents the difference in coefficients, and column 8 presents the corresponding *p*-value.

Appendix Table 4. Gender Peer Effect by Subject

	Chinese	Math	English
	(1)	(2)	(3)
Proportion female peers	0.808*** (0.303)	1.299*** (0.392)	1.207*** (0.298)
Female student	0.580*** (0.026)	0.148*** (0.031)	0.539*** (0.026)
School-grade fixed effects	Yes	Yes	Yes
Student controls	Yes	Yes	Yes
Teacher controls	Yes	Yes	Yes
Peer ability controls	Yes	Yes	Yes
Observations	7,400	7,312	7,305
R-squared	0.129	0.064	0.132

Notes: This table presents the gender peer effect on test score by subject. Test score is normalized by subject, grade, and school to obtain a mean of zero and standard deviation of one. Student controls include student's age, baseline noncognitive measurements, mother's education, father's education, and dummy variables indicating minority, local residence, rural residence, only child in family, attended kindergarten, and repeated a grade in primary school. Teacher controls include gender, age, years of schooling, experience, professional job title, and dummy variables indicating marital status and graduated from a normal college. Peer ability controls include male and female peers' average baseline academic performance in primary school, including whether repeated a grade and whether skipped a grade. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

Appendix Table 5. Balancing Test of Teacher Characteristics

VARIABLES	Coefficient	SE	Observations
<i>Head teacher</i>	(1)	(2)	(3)
Female	0.737	(0.642)	208
Education	0.139	(0.414)	207
Graduation from normal college	-0.025	(0.305)	207
Professional title	-0.052	(1.100)	206
Experience	0.455	(1.206)	205
Tenure status	-0.251	(0.290)	208
Teaching main subject	-0.423	(0.639)	208
<i>Subject teacher</i>			
Female	0.009	(0.159)	617
Education	0.311	(0.235)	613
Graduation from normal college	0.066	(0.175)	614
Professional title	0.372	(0.515)	614
Experience	-0.022	(0.390)	611
Tenure status	0.138*	(0.081)	607

Notes: Each cell represents a separate regression that regresses the corresponding teacher characteristic on peer female proportion, controlling for school-grade FE. Specifications for subject teacher also control for subject fixed effects. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

Appendix Table 6. Robustness Check: Sample Attrition

	Test score	Self- assessment	Depressed	Blue	Unhappy	Pessimistic	Fulfilling life	Confident abt future	Public enrichment	Private recreation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Prop female peers	0.046 (0.065)	-0.018 (0.058)	-0.021 (0.067)	-0.016 (0.068)	-0.022 (0.067)	-0.012 (0.068)	-0.045 (0.057)	-0.045 (0.056)	-0.055 (0.065)	-0.046 (0.062)
School-grade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	26,730	26,730	8,910	8,910	8,910	8,910	8,910	8,910	8,910	8,910
R-squared	0.060	0.042	0.044	0.047	0.043	0.044	0.044	0.046	0.047	0.047

Notes: The dependent variable in each cell is a dummy that equals one if the corresponding measurement is missing. All regressions include school-grade fixed effects. Regressions in column 1 and column 2 further include subject fixed effects.

Appendix Table 7. Robustness Check: Related Behavioral Outcome Variables

	Late for school	Drop classes
	(1)	(2)
Proportion female peer	-0.325*	-0.251
	(0.175)	(0.173)
Female student	-0.094***	-0.100***
	(0.029)	(0.025)
Subject FE	Yes	Yes
School-grade FE	Yes	Yes
Student controls	Yes	Yes
Teacher controls	Yes	Yes
Observations	7,758	7,752
R-squared	0.090	0.085

Notes: Dependent variables are normalized to have a mean of zero and standard deviation of one. Student controls include student age, baseline noncognitive measurements, mother's education, father's education, and dummy variables indicating minority, local residence, rural residence, only child in family, attended kindergarten, and repeated a grade in primary school. Teacher controls include gender, age, years of schooling, experience, professional job title, and dummy variables indicating marital status and graduated from a normal college. Peer ability controls include male and female peers' average baseline academic performance in primary school, including whether repeated a grade and whether skipped a grade. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

Appendix Table 8A: Mechanism Analysis by Gender: Female Peer Proportion and Channels

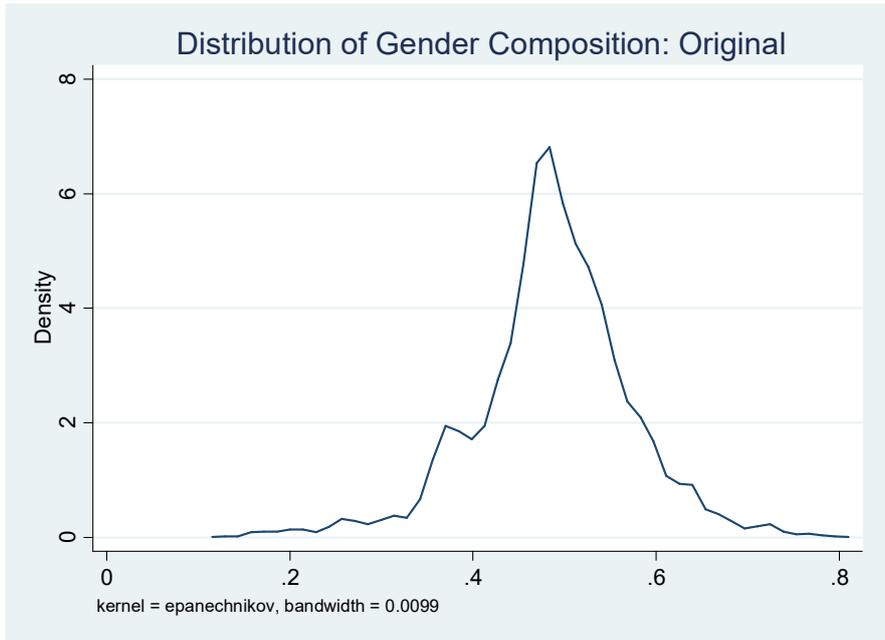
	Discussion in lecture	Teacher effort	“Teacher is responsible and patient”	Classroom environment	Student effort
<i>Female sample</i>					
	(1)	(2)	(3)	(4)	(5)
Proportion female peers	0.897 (0.583)	0.379* (0.219)	0.675* (0.357)	0.163 (0.328)	1.511*** (0.308)
Observations	555	459	11,055	11,305	10,632
R-squared	0.333	0.623	0.177	0.191	0.256
<i>Male sample</i>					
Proportion female peers	0.897 (0.583)	0.379* (0.219)	0.762* (0.422)	0.298 (0.348)	1.174*** (0.230)
Observations	555	459	10,909	11,490	10,862
R-squared	0.333	0.623	0.164	0.149	0.201
School-grade FE	Yes	Yes	Yes	Yes	Yes
Student controls	Yes	Yes	Yes	Yes	Yes
Teacher controls	Yes	Yes	Yes	Yes	Yes
Peer ability controls	NA	NA	Yes	Yes	Yes

Notes: Estimates are obtained from specifications that regress each channel on proportion of female peers, controlling for school-grade FE and other controls. Estimates for “Discussion in lecture” and “Teacher effort” are the same for male and female samples, because they are obtained from teacher-level analysis. Student controls include student age, baseline noncognitive measurements, mother’s education, father’s education, and dummy variables indicating minority, local residence, rural residence, only child in family, attended kindergarten, and repeated a grade in primary school. Teacher controls include gender, age, years of schooling, experience, professional job title, and dummy variables indicating marital status and graduated from a normal college. Peer ability controls include male and female peers’ average baseline academic performance in primary school, including whether repeated a grade and whether skipped a grade. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

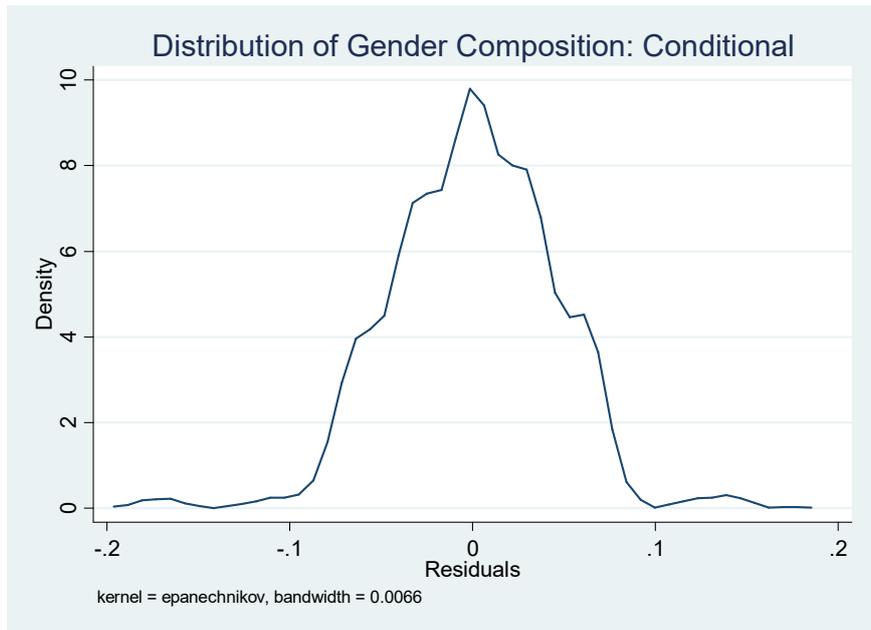
Appendix Table 8B: Percentage Each Channel Attributed to Test Score, by Gender

	Female	Male
	(1)	(2)
Teaching style:	-0.014	-0.008
Discussion in lecture	(0.012)	(0.015)
Teacher effort	0.050	0.140***
	(0.039)	(0.044)
Teacher is responsible and patient	0.109***	0.106***
	(0.018)	(0.021)
Classroom environment	0.096***	0.135***
	(0.022)	(0.022)
Student effort	-0.000	0.110***
	(0.022)	(0.026)
Proportion female peers	0.629**	1.403***
	(0.301)	(0.379)
School-grade FE	Yes	Yes
Student controls	Yes	Yes
Teacher controls	Yes	Yes
Peer ability controls	Yes	Yes
Observations	7,430	7,299
R-squared	0.111	0.105

Notes: Estimates are obtained from specification that regresses test score on the above channels, with controlling for proportion of female peers, school-grade FE, and other controls. Student controls include student age, baseline noncognitive measurements, mother's education, father's education, and dummy variables indicating minority, local residence, rural residence, only child in family, attended kindergarten, and repeated a grade in primary school. Teacher controls include gender, age, years of schooling, experience, professional job title, and dummy variables indicating marital status and graduated from a normal college. Peer ability controls include male and female peers' average baseline academic performance in primary school, including whether repeated a grade and whether skipped a grade. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

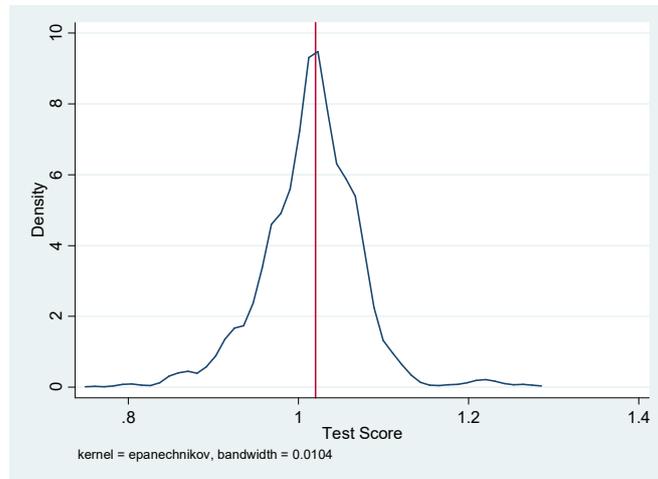


Appendix Figure 1A. Original Distribution of Gender Composition

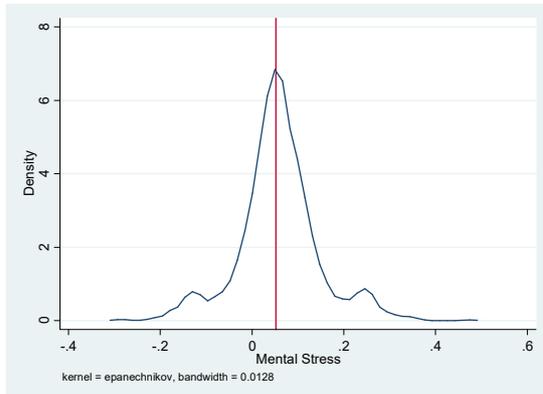


Appendix Figure 1B. Conditional Distribution of Gender Composition

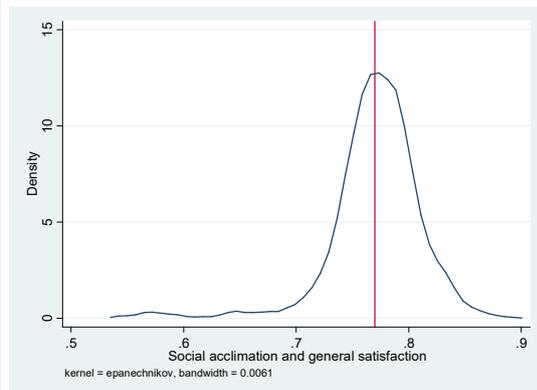
Notes: This figure plots the original and conditional distributions of female peers in class. The conditional distribution is the distribution of residuals obtained by regressing female peer proportion on school-grade FE and controls. The corresponding $(1-R^2)$ from this regression (i.e., that regresses peer female proportion on school-grade fixed effect and all control variables) equals to 0.243.



Appendix Figure 2A



Appendix Figure 2B



Appendix Figure 2C

Appendix Figure 2. Robustness Check:
Distribution of Coefficients after Randomly Dropping Two Schools

Notes: Figures plot the distributions of the coefficients from 2,211 regressions that each time randomly drop two schools from the sample. Vertical lines indicate our baseline estimates for the respective student outcome.