For Online Publication

Appendices for "Teacher Expectations Matter"

by: Nicholas W. Papageorge, Seth Gershenson and Kyungmin Kang

Additional results are organized into three appendices. In Appendix A, we discuss sample selection. In Appendix B we provide estimates using different functional forms and additional outcomes and also show additional summary statistics. We also provide additional parameter estimates and simulations from the measurement error model presented in the main text of the paper. In Appendix C, we provide estimates from alternative specifications of the measurement error model, including an alternative specification of bias (Appendix C.1); a version of the measurement error model that allows teacher biases for the same student to be correlated (Appendix C.2); a simplified version of the model that relaxes parametric assumptions and does not use data on test scores but instead relies on parameter restrictions for identification and uses years of education as a continuous outcome variable (Appendix C.3); and a model where the outcome variable is years of education, but which also uses data on test scores (Appendix C.4). In Appendix D, we show formal identification arguments for the measurement error model.

Appendix A Sample Selection

Table S1 shows how we arrive at our analytic sample, i.e., which variables lead us to drop observations. Table S2 repeats the main analysis in Table 5, but uses the maximum number of observations for each specification. Table S3 shows mean differences in some key variables between our analytic sample and the sample of individuals (whom we drop in our main analysis) for whom expectations data are missing. According to the table, once we condition on scores, GPA and school fixed effects, most variables are statistically indistinguishable across these two samples.

Appendix Table S1: Sample Selection

	Observations
Initial Number of Observations	16200
Dropped Due to Missing Educational Outcome	2950
Dropped Due to Missing Expectations	5340
Dropped Due to Missing Teacher Controls	570
Dropped Due to Missing Student SES Controls	750
Dropped Due to Missing Grades	530
Analytic Sample	6060

Student socio-economic status (SES) controls include indicators for household income and mother's educational attainment as well as indicators for student race, sex, and if a language other than English is spoken at home. Teacher controls include teacher race and gender dummies, years of experience, and whether or not the teacher majored in the subject he or she teachers. Grades include 9th grade GPA, and math and reading assessment scores. All sample sizes are rounded to the nearest 10 in accordance with NCES regulations for restricted data.

Appendix Table S2: OLS Estimates of Effect of Expectations on Educational Attainment: Maximum Sample Size

	All Students									White	Black
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
ELA Teacher Exp.	0.47***		0.31***	0.30***	0.24***	0.15***	0.14***	0.17***	0.16***	0.14***	0.17**
	(0.01)		(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.08)
Math Teacher Exp.		0.47***	0.32***	0.31***	0.25***	0.16***	0.13***	0.15***	0.13***	0.14***	0.11
		(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.07)
Teacher Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student SES	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9th Grade GPA	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
School FE	No	No	No	No	No	No	Yes	No	Yes	Yes	Yes
Teacher Dyad FE	No	No	No	No	No	No	No	Yes	No	No	No
Observations	9410	9860	7910	7220	6500	6060	6060	3600	3600	3970	610
R^2	0.22	0.22	0.29	0.30	0.34	0.37	0.45	0.59	0.46	0.48	0.65
Adjusted R^2	0.22	0.22	0.29	0.30	0.34	0.37	0.38	0.18	0.37	0.39	0.31

Note: $*p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$. The dependent variable is a binary indicator equal to one if the student completed a four-year college degree or more, and zero otherwise. Parentheses contain standard errors that are robust to clustering at the school level. These are OLS estimates of equation (2). Student socioeconomic status (SES) controls include indicators for household income and mother's educational attainment as well as indicators for student race, sex, and if a language other than English is spoken at home. Teacher controls include teacher race and gender dummies, years of experience, and whether or not the teacher majored in the subject he or she teaches. School FE refers to school fixed effects and Teacher Dyad FE refers to Math-ELA teacher pair fixed effects. Estimates in column (9) are from a school fixed effects model, but estimated on the subsample of students for whom teacher dyad fixed effects are identified.

Appendix Table S3: In-Sample and Out-of-Sample Means

	(1) Analytic Sample Mean	(2) Alternative Sample Mean	(3) Residual Difference
HH Income < 20K	0.11	0.19	0.00
HH Income 20K - 35K	0.16	0.21	-0.02
HH Income 35K - 75K	0.40	0.36	0.01
HH Income 75K - 100K	0.15	0.12	0.01
HH Income > 100K	0.18	0.12	-0.00
Father Did Not Finish HS	0.10	0.19	0.00
Father Has HS Diploma	0.26	0.29	-0.06**
Father Has Some College	0.27	0.25	0.05*
Father Has a Bachelor's or More	0.37	0.28	0.01
Mother Did not Finish HS	0.09	0.19	-0.02
Mother Has HS Diploma	0.25	0.28	0.00
Mother Has Some College	0.35	0.31	0.03
Mother Has a Bachelor's or More	0.31	0.23	-0.01
Student Is American Indian	0.00	0.02	-0.01**
Student Is Asian	0.08	0.16	-0.01
Student Is Black	0.10	0.19	-0.02
Student Is Hispanic	0.12	0.20	0.03**
Student Is Multiple Race	0.04	0.05	0.00
Student Is White	0.66	0.38	0.00*
9th Grade GPA	2.92	2.55	.
Reading Test Standardized Score	52.82	48.27	.
Math Test Standardized Score	53.01	48.54	.
Complete 4 Yr College	0.45	0.32	0.01
ELA Teacher Exp	0.64	.	.
Math Teacher Exp	0.63	.	.
Observations	6060	2440	

Note: $*p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$. Students are the unit of analysis. HH is household income. HS is high school. ELA is English language and arts. 9th-grade GPAs are on a 4.0 scale. Math and reading assessment scores are on a 0-100 scale. The alternative sample is the subsample of students in the data whose teacher expectations are not observed and thus not included in our analytic sample. Residual difference reports differences after controlling for 9th grade GPA, reading and math test scores, and school fixed effects. All sample sizes are rounded to the nearest 10 in accordance with NCES regulations for restricted data.

Appendix B Additional Results

Table S4 repeats the main analysis, but removes high school dropouts or individuals who received a graduate degree. Table S5 repeats main analyses using a logit specification and contains both logit coefficients along with marginal effects evaluated at sample means of all variables. Table S6 repeats the analyses using a probit specification. Table S7 provides estimates of the impact of teacher expectations on additional outcomes, such as employment. Summary statistics on variables used to show that disagreements are conditionally random and for IVs are found in Table S8. Table S9 contains "first stage" estimates of instrumental variables on teacher expectations. Table S10 reports the impact of instrumented expectations on outcomes using different sets of instruments. Additional parameter estimates for the structural model presented in the main text are found in Table S11. Figure S1 visualizes teacher bias using a heat map. Figure S2 explores counterfactual bias for white students.

Appendix Table S4: OLS Estimates of the Effect of Expectations on Education, Restricted Sample

	All Students									White	Black
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
ELA Teacher Exp.	0.42***		0.27***	0.27***	0.21***	0.14***	0.13***	0.15***	0.15***	0.12***	0.17**
	(0.01)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.08)
Math Teacher Exp.		0.42***	0.28***	0.27***	0.22***	0.14***	0.12***	0.14***	0.12***	0.13***	0.09
		(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.07)
Teacher Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student SES	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9th Grade GPA	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
School FE	No	No	No	No	No	No	Yes	No	Yes	Yes	Yes
Teacher Dyad FE	No	No	No	No	No	No	No	Yes	No	No	No
Observations	5320	5320	5320	5320	5320	5320	5320	2960	2960	3460	550
R^2	0.18	0.18	0.24	0.25	0.29	0.32	0.41	0.57	0.42	0.45	0.65
Adjusted R^2	0.18	0.18	0.24	0.24	0.29	0.31	0.32	0.06	0.31	0.34	0.26

Note: $*p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$. The dependent variable is a binary indicator equal to one if the student completed a four-year college degree or more, and zero otherwise. Parentheses contain standard errors that are robust to clustering at the school level. These are OLS estimates of equation (2) for the restricted sample of students who completed at least a high school degree but did not earn a graduate degree. Student socioeconomic status (SES) controls include indicators for household income and mother's educational attainment as well as indicators for student race, sex, and if a language other than English is spoken at home. Teacher controls include teacher race and gender dummies, years of experience, and whether or not the teacher majored in the subject he or she teaches. School FE refers to school fixed effects and Teacher Dyad FE refers to Math-ELA teacher pair fixed effects. Estimates in column (9) are from a school fixed effects model, but estimated on the subsample of students for whom teacher dyad fixed effects are identified.

Appendix Table S5: Logit Estimates of Effect of Expectations on Educational Attainment

	All Students							White	Black
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ELA Teacher Coefficient.	2.32***		1.63***	1.61***	1.37***	0.86***	0.89***	0.84***	3.30***
	(0.08)		(0.08)	(0.08)	(0.09)	(0.09)	(0.12)	(0.16)	(1.13)
Math Teacher Coefficient		2.30***	1.62***	1.61***	1.41***	0.86***	0.80***	0.88***	1.21*
		(0.07)	(0.08)	(0.08)	(0.08)	(0.09)	(0.11)	(0.15)	(0.67)
ELA Teacher APE	0.45***		0.29***	0.28***	0.22***	0.13***	0.12***	0.11***	0.33***
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.10)
Math Teacher APE		0.44***	0.29***	0.28***	0.23***	0.13***	0.11***	0.12***	0.12*
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.07)
Teacher Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Student SES	No	No	No	No	Yes	Yes	Yes	Yes	Yes
9th Grade GPA	No	No	No	No	No	Yes	Yes	Yes	Yes
School FE	No	No	No	No	No	No	Yes	Yes	Yes
Teacher Dyad FE	No	No	No	No	No	No	No	No	No
Observations	6060	6060	6060	6060	6060	6060	5660	3550	300

Note: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is a binary indicator equal to one if the student completed a four-year college degree or more, and zero otherwise. Parentheses contain standard errors that are robust to clustering at the school level. These are logit coefficient estimates and corresponding average partial effects (APE) of equation (2). Student socioeconomic status (SES) controls include indicators for household income and mother's educational attainment as well as indicators for student race, sex, and if a language other than English is spoken at home. Teacher controls include teacher race and gender dummies, years of experience, and whether or not the teacher majored in the subject he or she teaches. School FE refers to school fixed effects.

Appendix Table S6: Probit Estimates of Effect of Expectations on Educational Attainment

	All Students							White	Black
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ELA Teacher Coefficient	1.39***		0.97***	0.97***	0.84***	0.59***	0.57***	0.52***	1.97***
	(0.04)		(0.048)	(0.048)	(0.050)	(0.052)	(0.067)	(0.09)	(0.58)
Math Teacher Coefficient		1.39***	0.97***	0.97***	0.85***	0.60***	0.51***	0.56***	1.18**
		(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.09)	(0.49)
ELA Teacher APE	0.55***		0.38***	0.38***	0.33***	0.23***	0.22***	0.12***	0.33***
	(0.04)		(0.05)	(0.05)	(0.05)	(0.05)	(0.07)	(0.02)	(0.11)
Math Teacher APE		0.54***	0.38***	0.38***	0.33***	0.23***	0.20***	0.13***	0.20**
		(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.02)	(0.09)
Teacher Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Student SES	No	No	No	No	Yes	Yes	Yes	Yes	Yes
9th Grade GPA	No	No	No	No	No	Yes	Yes	Yes	Yes
School FE	No	No	No	No	No	No	Yes	Yes	Yes
Pseudo R^2	0.17	0.17	0.23	0.23	0.28	0.31	0.37	0.39	0.53
Observations	6060	6060	6060	6060	6060	6060	5660	3550	300

Note: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is a binary indicator equal to one if the student completed a four-year college degree or more, and zero otherwise. Parentheses contain standard errors that are robust to clustering at the school level. These are probit coefficient estimates and corresponding average partial effects (APE) of equation (2). Student socioeconomic status (SES) controls include indicators for household income and mother's educational attainment as well as indicators for student race, sex, and if a language other than English is spoken at home. Teacher controls include teacher race and gender dummies, years of experience, and whether or not the teacher majored in the subject he or she teaches. School FE refers to school fixed effects.

Appendix Table S7: OLS Estimates of Effect of Expectations on Additional Outcomes

	Emp, Full (1)	Emp, Part or Full (2)	Married (3)	Ever Married (4)	Child (5)	Assis. (6)	Own Home (7)
ELA Teacher Exp.	0.04*	0.05***	-0.03*	-0.05***	-0.10***	-0.07***	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)
Math Teacher Exp.	0.00	-0.00	0.02	0.02	-0.03*	-0.02	0.04*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Mean of Dep. Var.	0.76	0.86	0.29	0.34	0.28	0.14	0.28
Teacher Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student SES	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9th Grade GPA	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Dyad FE	No	No	No	No	No	No	No
Observations	6020	4530	4530	6020	5960	5470	4310
R^2	0.16	0.15	0.19	0.21	0.29	0.27	0.20
Adjusted R^2	0.05	0.04	0.09	0.11	0.20	0.17	0.07
Joint Significance	0.19	0.02	0.12	0.01	0.00	0.00	0.21

Note: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variables for columns (1)-(7) are employed full-time, employed part-time or full-time, being married, ever having been married, having at least one biological child, receiving public assistance, and owning a home 12 years from the initial year survey, respectively. Joint significance is the p-value from a joint significance test of the two teacher expectations. These are OLS estimates of equation (2). Student socioeconomic status (SES) controls include indicators for household income and mother's educational attainment as well as indicators for student race, sex, and if a language other than English is spoken at home. Teacher controls include teacher race and gender dummies, years of experience, and whether or not the teacher majored in the subject he or she teaches. School FE refers to school fixed effects.

Appendix Table S8: Summary Statistics - Identification

	Mean	Standard Deviation	Count
Ever Bullied	0.2061	(0.4046)	5820
Got in Fight	0.1051	(0.3068)	5810
Participated in Science Fair	0.1532	(0.3602)	5800
S Finds Class Interesting	0.5917	(0.4916)	5800
S Ever in College Prep	0.2193	(0.4138)	5740
P Thinks S Has Disability	0.0938	(0.2915)	5700
Passive (ELA)	0.1184	(0.3231)	5960
Passive (Math)	0.1108	(0.3139)	5980
Never Attentive (ELA)	0.0079	(0.0886)	5940
Rarely Attentive (ELA)	0.0364	(0.1873)	5940
Sometimes Attentive (ELA)	0.1655	(0.3716)	5940
Mostly Attentive (ELA)	0.4645	(0.4988)	5940
Never Attentive (Math)	0.0064	(0.0795)	5980
Rarely Attentive (Math)	0.0387	(0.1928)	5980
Sometimes Attentive (Math)	0.1600	(0.3666)	5980
Mostly Attentive (Math)	0.4546	(0.4980)	5980
Strongly Agree Reading Is Fun	0.1624	(0.3688)	4850
Agree Reading Is Fun	0.3538	(0.4782)	4850
Disagree Reading Is Fun	0.3509	(0.4773)	4850
Strongly Agree Math Is Fun	0.0780	(0.2682)	4800
Agree Math Is Fun	0.2661	(0.4419)	4800
Disagree Math Is Fun	0.4673	(0.4990)	4800
Hours Spent on Homework in School (Math)	2.4384	(2.8428)	5680
Hours Spent on Homework out of School (Math)	2.7791	(2.9072)	5700
Hours Spent on Homework in School (ELA)	1.9719	(2.6221)	5540
Hours Spent on Homework out of School (ELA)	2.6089	(2.9771)	5590
Total Hours on Homework (Math)	5.2169	(4.7048)	5640
Total Hours on Homework (ELA)	4.5816	(4.5987)	5520

Summary table for variables used to check the exogeneity of teacher bias in section 4.3. S is student and P is parent.

Appendix Table S9: The Teacher Expectation Production Function, with School FE

Instruments:	Avg. Expectations for Other Students		Transitory Factors e.g., Passiveness		Both Sets of Instruments	
	English	Math	English	Math	English	Math
	(1)	(2)	(3)	(4)	(5)	(6)
Teacher's Average Expectations (ELA)	-0.83***	-0.06			-0.28***	-0.07
	(0.27)	(0.10)			(0.08)	(0.05)
Teacher's Average Expectations (Math)	-0.03	-1.08***			-0.00	-0.27***
	(0.12)	(0.27)			(0.05)	(0.07)
Passive (ELA)			-0.13***	-0.02	-0.12***	-0.02
			(0.02)	(0.02)	(0.03)	(0.03)
Passive (Math)			-0.02	-0.07***	-0.05*	-0.05
			(0.02)	(0.02)	(0.03)	(0.03)
Never Attentive (ELA)			-0.29***	-0.06	-0.22***	-0.12*
			(0.06)	(0.06)	(0.06)	(0.07)
Rarely Attentive (ELA)			-0.33***	-0.05	-0.35***	-0.06
			(0.04)	(0.03)	(0.05)	(0.04)
Sometimes Attentive (ELA)			-0.31***	-0.04**	-0.26***	-0.02
			(0.02)	(0.02)	(0.04)	(0.03)
Mostly Attentive (ELA)			-0.10***	-0.03**	-0.07***	-0.01
			(0.01)	(0.01)	(0.02)	(0.02)
Never Attentive (Math)			-0.04	-0.34***	-0.05	-0.65***
			(0.07)	(0.10)	(0.16)	(0.19)
Rarely Attentive (Math)			-0.09***	-0.35***	-0.15***	-0.29***
			(0.03)	(0.03)	(0.06)	(0.06)
Sometimes Attentive (Math)			-0.08***	-0.32***	-0.06*	-0.33***
			(0.02)	(0.02)	(0.03)	(0.04)
Mostly Attentive (Math)			-0.01	-0.09***	0.01	-0.09***
			(0.01)	(0.01)	(0.02)	(0.02)
Strongly Agree Reading is Fun			0.07***	-0.03	0.10***	-0.07**
			(0.02)	(0.02)	(0.03)	(0.03)
Agree Reading is Fun			0.05**	-0.02	0.07**	-0.06**
			(0.02)	(0.02)	(0.03)	(0.03)
Disagree Reading is Fun			0.06***	0.00	0.06**	-0.04
			(0.02)	(0.02)	(0.03)	(0.03)
Strongly Agree Math is Fun			-0.07***	0.06**	-0.12***	0.07**
			(0.02)	(0.02)	(0.03)	(0.03)
Agree Math is Fun			-0.05***	0.05**	-0.06**	0.09***
			(0.02)	(0.02)	(0.03)	(0.03)
Disagree Math is Fun			-0.04**	0.00	-0.04*	0.04*
			(0.02)	(0.02)	(0.02)	(0.02)
R^2	0.56	0.62	0.59	0.59	0.61	0.63
Adjusted R^2	0.48	0.55	0.52	0.52	0.52	0.54
N	1450	1450	4420	4420	2120	2120
F-test	5.34	8.07	19.70	23.35	9.00	9.71

Note: $*p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$. The dependent variable is a binary indicator equal to one if the teacher expects the student to complete a four-year college degree or more, and zero otherwise. All specifications include student socioeconomic (SES) controls, teacher controls, and 9th grade GPA. Student socioeconomic status (SES) controls include indicators for household income and mother's educational attainment as well as indicators for student race, sex, and if a language other than English is spoken at home. Teacher controls include teacher race and gender dummies, years of experience, and whether or not the teacher majored in the subject he or she teaches. School FE refers to school fixed effects.

Appendix Table S10: 2SLS Estimates of Effect of Expectations on Educational Attainment

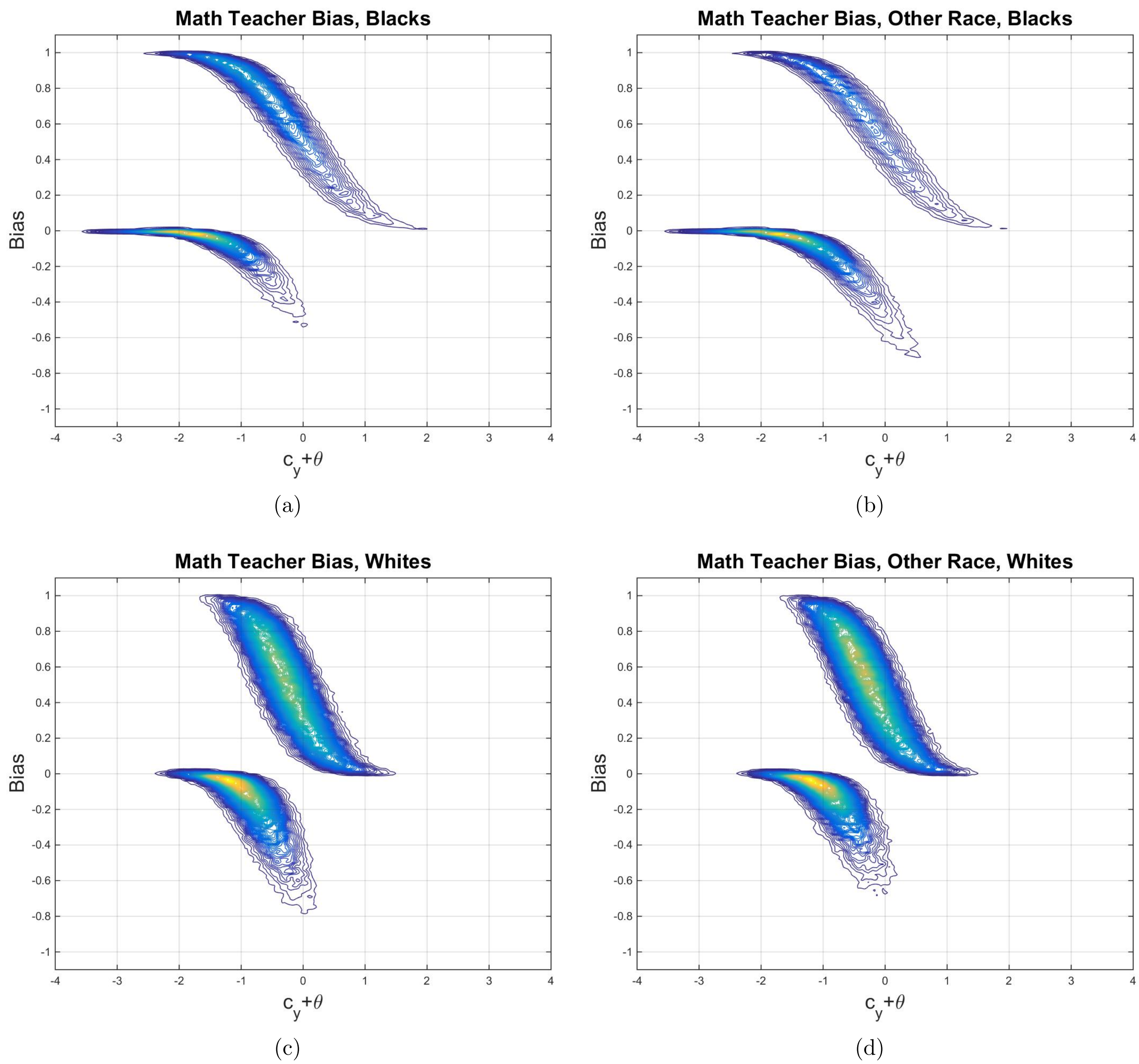
Instruments:	Avg. Expectations for Other Students		Transitory Factors e.g., Passiveness		Both Sets of Instruments	
English	Math	English	Math	English	Math	
Panel A: School FE	(1)	(2)	(3)	(4)	(5)	(6)
Expects College or More (OLS)	0.14***	0.13***	0.14***	0.13***	0.14***	0.13***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Expects College or More (2SLS)	0.27	0.14*	0.22***	0.21***	0.20***	0.16***
	(0.19)	(0.08)	(0.06)	(0.06)	(0.08)	(0.09)
Hausman Test	0.11		0.95		0.53	
Control Function Test	0.45		0.05		0.58	
1st Stage F-test	40.99		18.89		12.11	
Over-identification Test	.		0.51		.	
R^2	0.24		0.26		0.26	
Adjusted R^2	0.11		0.14		0.10	
N	1450		4420		2890	
Panel B: Teacher-dyad FE	(1)	(2)	(3)	(4)	(5)	(6)
Expects College or More (OLS)	0.17***	0.15***	0.17***	0.15***	0.17***	0.15***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Expects College or More (2SLS)	0.17***	0.13***	0.28***	0.18*	0.09**	0.17***
	(0.04)	(0.04)	(0.10)	(0.10)	(0.04)	(0.04)
Hausman Test	1.00		1.00		1.00	
Control Function Test	0.30		0.48		0.10	
1st Stage F-test	1512.90		4.06		73.83	
Over-identification Test	.		0.95		0.91	
R^2	0.48		0.69		0.60	
Adjusted R^2	.		.		.	
N	1450		3160		2140	

Note: $^*p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is a binary indicator equal to one if the student completed a four-year college degree or more, and zero otherwise. Parentheses contain standard errors that are robust to clustering at the school level. Panel A presents OLS and 2SLS estimates of equation (2) that condition on school FE for the analytic sample for which all instruments are observed. Panel B presents analogous results, but instead conditions on teacher-dyad FE, which reduces the sample size as it is limited to students for whom at least one other student experienced the same teacher-dyad. Instruments in the first two columns are the average of each teachers' expectations for their other students. Instruments in the second two columns include the transitory factors described in the main text (e.g., teacher reports that a student is "passive in class" or "likes math"). Instruments in the third column include all instruments from the first two sets of columns. Regressions in both panels control for student race, sex, 9th-grade GPA, math and ELA scores, household income, an indicator for single-parent family, and mother's educational attainment. Regressions in panel A control for teacher race, sex, and educational attainment; these variables are colinear with the teacher-dyad FE. Column (1) restricts the sample to students whose teachers have 5 or more students in the sample. Column (3) restricts the sample to students whose teachers have 3 or more students in the sample.

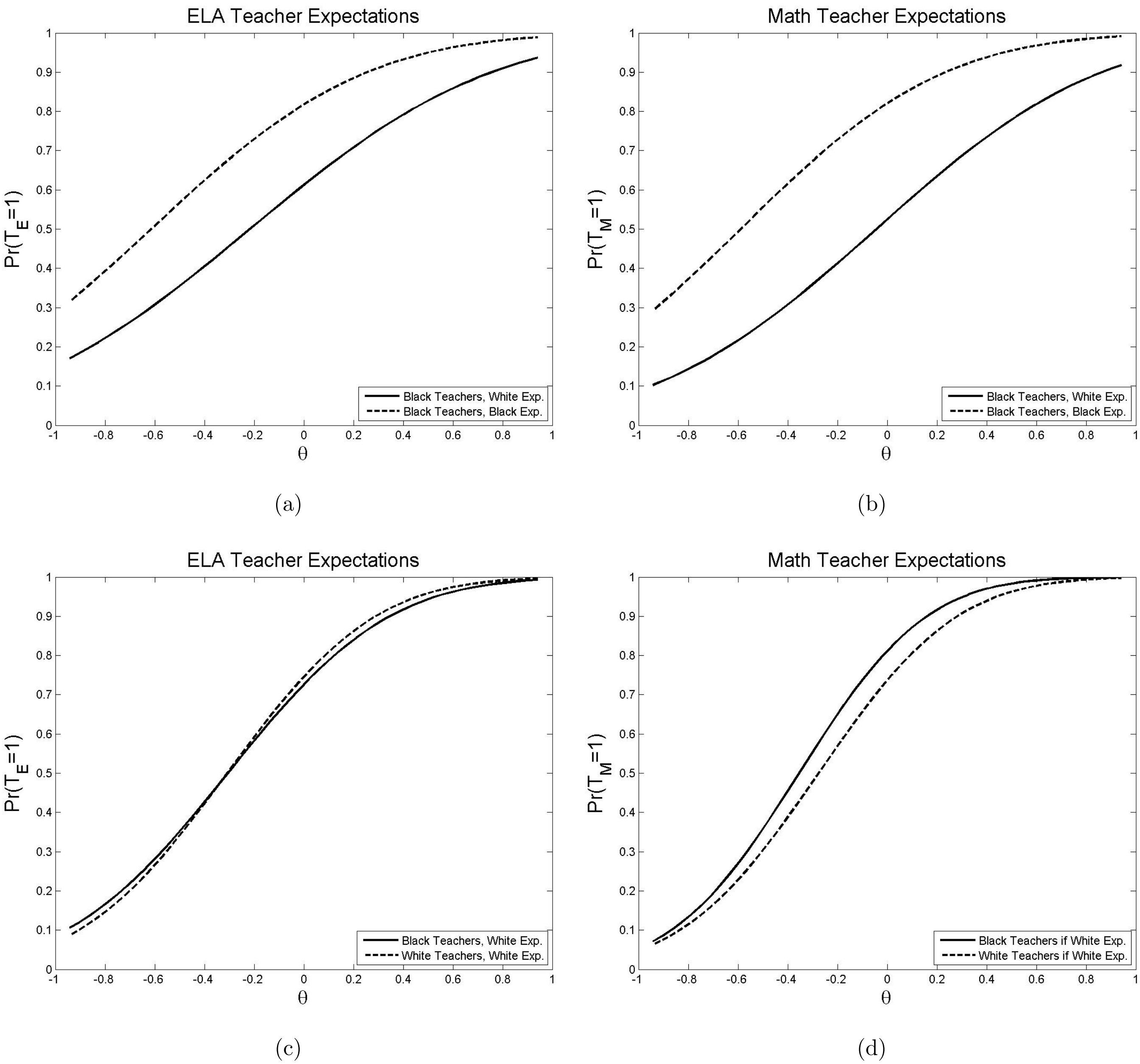
Appendix Table S11: Parameter Estimates - Additional Measures

	Whites	Blacks
c_G	0.15***	-0.63***
	(0.02)	(0.04)
c_{S_E}	0.19***	-0.63***
	(0.02)	(0.04)
c_{S_M}	0.17***	-0.75***
	(0.01)	(0.04)
ϕ_G	1.22***	0.67***
	(0.13)	(0.12)
ϕ_{S_E}	1.53***	0.96***
	(0.16)	(0.17)
ϕ_{S_M}	1.53***	1.01***
	(0.16)	(0.18)
σ_G	0.77***	0.83***
	(0.01)	(0.03)
σ_{S_M}	0.47***	0.46***
	(0.01)	(0.03)
σ_{S_R}	0.54***	0.52***
	(0.01)	(0.02)
N	3970	610

Note: $*p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$. Parameter estimates of equation (15) are reported. Standard errors are computed by constructing the Hessian of the likelihood function using the outer product measure. To compute the outer product measure, we calculate two-sided numerical derivatives of the likelihood function for each estimated parameter. In each direction, the derivative is calculated by perturbing each parameter and then computing the likelihood. Standard errors for the average partial effects (APE) are calculated using the delta method.



Appendix Figure S1: Distribution of Bias. Panel 1(a) shows contour plot of black math teachers' bias regarding black students. Panel 1(b) shows the distribution of non-black math teacher bias regarding black students. Panel 1(c) and 1(d) show contour plots of white and non-white teachers' bias, respectively, regarding white students.



Appendix Figure S2: Teacher Expectations Regarding White Students. Panel 2(a) shows how teacher expectations change when white students face the same expectation production function from black ELA teachers as black students. Panel 2(b) shows how the expectations change in the counterfactual scenario for black math teachers. Panels 2(c) and 2(d), respectively, compare white and black ELA and math teachers' expectation for white students with given θ_i.

Appendix C Robustness Checks and Alternative Specifications

Appendix C.1 Alternative Definitions of Bias

There are different ways to define bias in our setup. For results reported in the main text, we define bias as

$$b_{ji} \equiv T_{ji} - \Phi(c + \theta_i + G_i\beta), \tag{1}$$

which is equation (13) in the main text. An alternative is to define bias as:

$$b_{ji} = \Phi(c_j + \phi_j\theta_i + G_i\beta_j) - \Phi(c + \theta_i + G_i\beta). \tag{2}$$

The problem with this alternative definition is that the two teacher expectations, on average, are close to each other, even after introducing nonlinearity by using a probit specification. Therefore, we run into a multicollinearity problem. Thus, we define bias as above but modify the outcome equation (equation (10) in the main text) to be

$$Pr(y_i = 1) = \Phi\left(c + \theta_i + G_i\beta + \frac{1}{2}(b_{Ei} + b_{Mi})\gamma\right). \tag{3}$$

Parameter estimates are in Appendix Table S12. Remaining equations are the same as in the model in the main text. The results are qualitatively similar to the main results. Lastly, we report parameter estimates where

$$b_i = T_i. \tag{4}$$

That is, we assume that teacher expectation itself, and not bias, enters into the education production function. The parameter estimates are in Appendix Table S13.

Appendix Table S12: Parameter Estimates – Alternative Definition of Bias

Variable	All	Whites	Blacks
γ	0.33***	0.38***	0.47
	(0.11)	(0.13)	(0.54)
c	-0.39***	-0.31***	-0.91**
	(0.08)	(0.09)	(0.48)
b_y	0.42***	0.44***	0.24
	(0.04)	(0.05)	(0.17)
σ_θ	0.64***	0.61***	0.84***
	(0.04)	(0.05)	(0.16)
c_E	0.52***	0.59***	0.25***
	(0.02)	(0.03)	(0.08)
c_M	0.50***	0.59***	0.06
	(0.02)	(0.03)	(0.08)
ϕ_E	1.15***	1.35***	0.74***
	(0.10)	(0.15)	(0.19)
ϕ_M	1.30***	1.56***	0.85***
	(0.10)	(0.16)	(0.21)
b_E	0.53***	0.53***	0.45***
	(0.03)	(0.04)	(0.07)
b_M	0.49***	0.48***	0.38***
	(0.03)	(0.04)	(0.08)
c_G	0.02**	0.15***	-0.63***
	(0.01)	(0.02)	(0.04)
c_{S_M}	-0.00	0.17***	-0.75***
	(0.01)	(0.01)	(0.04)
c_{S_R}	-0.00	0.19***	-0.63***
	(0.01)	(0.02)	(0.04)
ϕ_G	1.04***	1.03***	0.64***
	(0.07)	(0.09)	(0.13)
ϕ_{S_M}	1.37***	1.27***	0.91***
	(0.09)	(0.11)	(0.17)
ϕ_{S_R}	1.30***	1.27***	0.96***
	(0.08)	(0.11)	(0.18)
σ_G	0.80***	0.77***	0.83***
	(0.01)	(0.01)	(0.03)
σ_{S_M}	0.49***	0.48***	0.47***
	(0.01)	(0.01)	(0.03)
σ_{S_R}	0.56***	0.55***	0.52***
	(0.01)	(0.01)	(0.02)

Note: $*p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Parameter estimates of equations (10)-(15) using an alternative specification in equations (2) and (3) in Appendix C.1 are reported. Standard errors are computed by constructing the Hessian of the likelihood function using the outer product measure. To compute the outer product measure, we calculate two-sided numerical derivatives of the likelihood function for each estimated parameter. In each direction, the derivative is calculated by perturbing each parameter and then computing the likelihood.

Appendix Table S13: Parameter Estimates – Teacher Expectations in the Education Production Function

Variable	Estimates	s.e.
γ_E	0.55***	(0.05)
γ_M	0.50***	(0.05)
c	-0.90***	(0.05)
b_y	0.31***	(0.03)
σ_θ	0.67***	(0.03)
c_E	0.52***	(0.02)
ϕ_E	0.57***	(0.06)
b_E	0.55***	(0.03)
c_M	0.49***	(0.02)
ϕ_M	1.14***	(0.07)
b_M	0.52***	(0.03)
c_{S_E}	0.00	(0.01)
ϕ_{S_E}	1.32***	(0.07)
σ_{S_E}	0.48***	(0.01)
c_{S_M}	0.00	(0.01)
ϕ_{S_M}	1.24***	(0.06)
σ_{S_M}	0.55***	(0.08)
c_G	0.00	(0.01)
ϕ_G	0.94***	(0.05)
σ_G	0.78***	(0.08)

Note: $*p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$. Parameter estimates of equations (10)-(15) using an alternative bias definition in equation (4) in Appendix C.1 are reported. Standard errors are computed by constructing the Hessian of the likelihood function using the outer product measure. To compute the outer product measure, we calculate two-sided numerical derivatives of the likelihood function for each estimated parameter. In each direction, the derivative is calculated by perturbing each parameter and then computing the likelihood.

Appendix C.2 Correlated Errors

In our main specification discussed in section 4.2 we have assumed that the two teacher error terms are independent. Given our use of 9th grade GPA and test scores as additional measures, identification is still achieved if we allow these to be correlated. The correlation captures the possibility that a student with a given objective probability of college completion may face two teachers that make similar errors. It is worth mentioning that parameter estimates are quite similar to the main model where the correlation is set to 0.

Formally, teacher expectations, denoted T_{ji} for teachers $j \in \{E, M\}$, are given by:

$$T_{ji}^* = c_j + \phi_j\theta_i + G_i\beta_j + D_{ji} \times [c_{j,D} + \phi_{j,D}\theta_i + G_i\beta_{j,D}] + e_{ji}, \tag{5}$$

where T_{ji} is a binary indicator equal to 1 if $T_{ji} \geq 0$ for $j \in \{E, M\}$ and 0 otherwise, and

$$\begin{pmatrix} e_{Ei} \\ e_{Mi} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{EM} \\ \sigma_{EM} & 1 \end{pmatrix}\right].$$

The indicator D_{ji} takes the value of one if student i faces an other-race subject-j teacher, and zero otherwise. This captures how teacher-student racial mismatch can change how teachers form expectations for a given student with a singular objective probability of college completion. Estimates are reported in Table S14, S15, and S16.

Appendix Table S14: Education Production Function Estimates, Correlated Teacher Expectations

	Whites	Blacks
γ_E	0.54***	0.54***
	(0.06)	(0.16)
γ_M	0.57***	0.25*
	(0.06)	(0.15)
β	0.50***	0.27**
	(0.05)	(0.11)
c	-0.47***	-0.84***
	(0.05)	(0.14)
σ_θ	0.51***	0.79***
	(0.05)	(0.14)
APE		
b_E	0.19***	0.15***
	(0.02)	(0.05)
b_M	0.20***	0.07
	(0.02)	(0.04)
Elasticities		
b_E	0.13***	0.20***
	(0.02)	(0.06)
b_M	0.13***	0.09
	(0.02)	(0.05)
N	3970	610

Note: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Parameter estimates of equation (10) for the model in Appendix C.2 are reported. The dependent variable is a binary indicator equal to one if the student completed a four-year college degree or more, and zero otherwise. Standard errors are computed by constructing the Hessian of the likelihood function using the outer product measure. To compute the outer product measure, we calculate two-sided numerical derivatives of the likelihood function for each estimated parameter. In each direction, the derivative is calculated by perturbing each parameter and then computing the likelihood. Standard errors for the average partial effects (APE) and elasticities are calculated using the delta method.

Appendix Table S15: Teacher Expectation Production Function Estimates, Correlated Teacher Expectations

	Whites		Blacks	
	ELA	Math	ELA	Math
	(1)	(2)	(3)	(4)
c	0.56***	0.54***	0.47**	0.58***
	(0.03)	(0.03)	(0.18)	(0.19)
c_D	-0.04	0.20	-0.25	-0.58***
	(0.12)	(0.14)	(0.20)	(0.21)
ϕ	1.32***	1.56***	0.75***	1.18**
	(0.17)	(0.19)	(0.39)	(0.50)
ϕ_D	-0.33	-0.05	-0.07	-0.35
	(0.41)	(0.35)	(0.30)	(0.47)
β	0.58***	0.52***	0.49***	0.29
	(0.03)	(0.03)	(0.16)	(0.21)
β_D	0.19	0.21	0.00	0.13
	(0.17)	(0.13)	(0.17)	(0.23)
σ_{EM}	0.43***		0.44***	
	(0.03)		(0.07)	
APE				
D	-0.02	0.05	-0.10*	-0.25***
	(0.03)	(0.03)	(0.05)	(0.07)
N	3970		610	

Note: $*p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$. Parameter estimates of equation (5) for the model in Appendix C.2 are reported. Standard errors are computed by constructing the Hessian of the likelihood function using the outer product measure. To compute the outer product measure, we calculate two-sided numerical derivatives of the likelihood function for each estimated parameter. In each direction, the derivative is calculated by perturbing each parameter and then computing the likelihood. Standard errors for the average partial effects (APE) are calculated using the delta method.

Appendix Table S16: Parameter Estimates - Additional Measures, Correlated Teacher Expectations

	Whites	Blacks
c_G	0.15***	-0.62***
	(0.02)	(0.04)
c_{S_E}	0.19***	-0.62***
	(0.02)	(0.04)
c_{S_M}	0.17***	-0.75***
	(0.01)	(0.04)
ϕ_G	1.23***	0.67***
	(0.13)	(0.13)
ϕ_{S_E}	1.56***	0.97***
	(0.17)	(0.17)
ϕ_{S_M}	1.57***	1.03***
	(0.17)	(0.18)
σ_G	0.78***	0.84***
	(0.01)	(0.03)
σ_{S_E}	0.54***	0.52***
	(0.01)	(0.03)
σ_{S_M}	0.46***	0.45***
	(0.01)	(0.04)
N	3970	610

Note: $*p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$. Parameter estimates of equation (15) for the model in Appendix C.2 are reported. Standard errors are computed by constructing the Hessian of the likelihood function using the outer product measure. To compute the outer product measure, we calculate two-sided numerical derivatives of the likelihood function for each estimated parameter. In each direction, the derivative is calculated by perturbing each parameter and then computing the likelihood. Standard errors for the average partial effects (APE) are calculated using the delta method.

Appendix C.3 Identification Using Parameter Restrictions

In the main text, we claim that we can identify the impact of bias using data on two teachers' expectations and student outcomes as long as we are willing to make strong functional form assumptions and if we restrict parameters. One of the key reasons we instead opt for using additional data is that it allows us to relax continuity of outcomes. Moreover, we can avoid parameter restrictions. Still, we think it is worth demonstrating that even if we limit ourselves to teacher expectations and student outcomes and relax functional form restrictions, we can still achieve identification of the magnitude of bias, its various sources and its impact on outcomes. Moreover, we are able to demonstrate that our results are similar, which suggests that our main results are not driven by the test score data we use to identify additional model parameters.

In what follows, we omit the subscript i. Y is a continuous outcome. T_j are teacher expectations for teacher $j \in \{E, M\}$ about the outcome Y. We have suppressed student indices. b_j are biases about the student for teacher j and will be explained below. We allow teachers to have mean expectations that deviate from each other and also from the true mean, denoted c. Teacher means are denoted c_j. This captures how, on average, teachers can be wrong. Teachers can make a student-specific error, which is denoted e_j. We also allow teachers be wrong about how θ maps to outcomes, which is captured by ϕ.

$$\begin{aligned} Y &= c + \theta + [b_E + b_M]\gamma + e_Y \\ T_E &= c_E + \phi\theta + e_E \\ T_M &= c_M + \phi\theta + e_M \end{aligned} \tag{6}$$

Notice that we have made parameter restrictions on the model in the main text. In particular, $\phi_1 = \phi_2 \equiv \phi$ and $\gamma_1 = \gamma_2 \equiv \gamma$. Further, we ignore the discussion of other-race teacher interactions here. We relax the normality assumption used in the main text and only assume that the disturbances e_j, $j \in \{y, E, M\}$, and θ have mean 0 and positive, finite variances σ_j^2 and σ_θ^2. We rewrite the production of expectations to be:

$$\begin{aligned} T_E &= c + \theta + (c_E - c) + (\phi - 1)\theta + e_E \\ T_M &= c + \theta + (c_M - c) + (\phi - 1)\theta + e_M \end{aligned} \tag{7}$$

Notice teacher expectations are the correct expectations plus a systematic component $c_j - c$, a component that depends on the objective probability θ and an idiosyncratic component.

Bias is defined as follows:

$$\begin{aligned} T_E - c - \theta &\equiv b_E = (c_E - c) + (\phi - 1)\theta + e_E \\ T_M - c - \theta &\equiv b_M = (c_M - c) + (\phi - 1)\theta + e_M \end{aligned} \tag{8}$$

Given the above, we re-write the outcome equation as follows:

$$\begin{aligned} Y &= c + (c_E + c_M - 2c)\gamma \\ &+ \theta(1 + 2\gamma(\phi - 1)) \\ &+ e_E\gamma + e_M\gamma \\ &+ e_Y \end{aligned} \tag{9}$$

Rewrite again as:

$$\begin{aligned} Y &= \bar{c} + \theta\psi + e_E\gamma + e_M\gamma + e_Y \\ T_E &= c_E + \phi\theta + e_E \\ T_M &= c_M + \phi\theta + e_M \\ \bar{c} &= c + (c_E + c_M - 2c)\gamma \\ \psi &= 1 + 2\gamma(\phi - 1) \end{aligned} \tag{10}$$

De-mean, so that $Y - \bar{c} = \tilde{Y}$, $T_E - c_E = \tilde{T}_E$ and $T_M - c_M = \tilde{T}_M$. Next, independence implies the following:

$$\begin{aligned} \text{Cov}(\tilde{T}_E, \tilde{T}_M) &= \phi^2 \text{Var}(\theta) \\ \text{Cov}(\tilde{Y}, \tilde{T}_E) &= \psi\phi \text{Var}(\theta) + \gamma \text{Var}(e_E) \\ \text{Cov}(\tilde{Y}, \tilde{T}_M) &= \psi\phi \text{Var}(\theta) + \gamma \text{Var}(e_M) \\ \text{Var}(\tilde{T}_E) &= \phi^2 \text{Var}(\theta) + \text{Var}(e_E) \\ \text{Var}(\tilde{T}_M) &= \phi^2 \text{Var}(\theta) + \text{Var}(e_M) \end{aligned} \tag{11}$$

Notice

$$\begin{aligned} \text{Var}(e_E) &= \text{Var}(\tilde{T}_E) - \text{Cov}(\tilde{T}_E, \tilde{T}_M) \\ \text{Var}(e_M) &= \text{Var}(\tilde{T}_M) - \text{Cov}(\tilde{T}_E, \tilde{T}_M) \\ \text{Cov}(\tilde{Y}, \tilde{T}_E) - \text{Cov}(\tilde{Y}, \tilde{T}_M) &= \gamma[\text{Var}(e_E) - \text{Var}(e_M)] \end{aligned} \tag{12}$$

Therefore

$$\gamma = \frac{\text{Cov}(\tilde{Y}, \tilde{T}_E) - \text{Cov}(\tilde{Y}, \tilde{T}_M)}{\text{Var}(\tilde{T}_E) - \text{Var}(\tilde{T}_M)} \tag{13}$$

Since we have γ, we can identify ϕ and ψ as follows:

$$\begin{aligned}
\phi[\text{Cov}(\tilde{Y},\tilde{T}_E)-\gamma\text{Var}(e_E)] &= \psi\text{Cov}(\tilde{T}_E,\tilde{T}_M)=\psi\phi^2\text{Var}(\theta) \\
\implies \frac{\phi}{\psi} &= \frac{\text{Cov}(\tilde{T}_E,\tilde{T}_M)}{[\text{Cov}(\tilde{Y},\tilde{T}_E)-\gamma\text{Var}(e_E)]} \\
&= \frac{\text{Cov}(\tilde{T}_E,\tilde{T}_M)}{[\text{Cov}(\tilde{Y},\tilde{T}_E)-\gamma(\text{Var}(\tilde{T}_E)-\text{Cov}(\tilde{T}_E,\tilde{T}_M))]} \\
&\equiv \Lambda
\end{aligned} \tag{14}$$

We also have that

$$\psi = 1+2\gamma(\phi-1) \tag{15}$$

Together, we get that:

$$\psi = \frac{1-2\gamma}{1-2\gamma\Lambda} \tag{16}$$

When we get results, it will sometimes be interesting to decompose the different effects of bias. To make this clear, re-write the outcome equation as follows:

$$\begin{array}{rcll}
Y &=& c+\theta & : \text{Explains } Y \\
&+& (c_E+c_M-2c)\gamma & : \text{Systematic Bias} \\
&+& \theta 2(\phi-1)\gamma & : \text{Bias as a Function of } \theta \\
&+& (e_E+e_M)\gamma & : \text{Idiosyncratic Bias} \\
&+& e_y & : \text{Disturbance} \\
T_E &=& c_E+\phi\theta+e_E & \\
T_M &=& c_M+\phi\theta+e_M &
\end{array} \tag{17}$$

Estimating this model purely on expectations and outcomes data yields $\hat{\gamma}=0.22$, and $s.e.(\hat{\gamma})=0.06$ using expected and actual years of education measures. This estimate is reassuring as it is fairly similar to parameters we estimate in the main analysis. In other words, the additional data we use to relax continuity and to identify additional model parameters do not appear to drive our main results.

Appendix C.4 Measurement Model with Years of Education

Here, we present estimates of a simplified version of our econometric model where the outcome is years of education, there are no race-teacher interactions and where we omit 9th grade GPA. The aim is to provide a comparison of γ estimates to the model in section Appendix C.3. Specifically, we jointly estimate the following system of equations, where we assume that error terms are normally distributed with mean zero.

$$\begin{aligned} Y &= c + \theta + \gamma_E b_E + \gamma_M b_M + e_Y \\ T_E &= c_E + \phi_E \theta + e_E \\ T_M &= c_M + \phi_M \theta + e_M \\ S_R &= c_{S_R} + \phi_{S_R} \theta + e_{S_R} \\ S_M &= c_{S_M} + \phi_{S_M} \theta + e_{S_M}. \end{aligned} \tag{18}$$

Parameter estimates are given in Table S17 for the full sample and then separately for white and black students. Notice that estimates of γ for the full sample are similar to γ estimated using the restricted model in section Appendix C.3, which does not use test scores for identification. This similarity provides some support for the idea that test scores, used for identification of the main model estimated in section 4, are valid additional measurements of θ that can be used for identification of the non-linear model.

Appendix Table S17: Years of Education

Variable	All	Whites	Blacks
γ_E	0.18***	0.21***	0.11***
	(0.01)	(0.02)	(0.04)
γ_M	0.17***	0.19***	0.11***
	(0.01)	(0.02)	(0.04)
c	14.63***	14.67***	14.40***
	(0.04)	(0.06)	(0.09)
c_E	15.67***	15.79***	14.89***
	(0.03)	(0.04)	(0.10)
c_M	15.52***	15.66***	14.68***
	(0.03)	(0.03)	(0.09)
ϕ_E	1.89***	2.10***	1.59***
	(0.10)	(0.18)	(0.20)
ϕ_M	1.84***	2.03***	1.57***
	(0.09)	(0.17)	(0.21)
c_{S_M}	-0.00	0.17***	-0.75***
	(0.01)	(0.01)	(0.04)
c_{S_R}	0.00	0.19***	-0.63***
	(0.01)	(0.01)	(0.04)
ϕ_{S_M}	1.12***	1.13***	1.00***
	(0.06)	(0.09)	(0.11)
ϕ_{S_R}	1.07***	1.14***	0.96***
	(0.06)	(0.09)	(0.11)
σ_θ	0.77***	0.68***	0.79***
	(0.04)	(0.06)	(0.09)
σ_E	1.65***	1.57***	1.76***
	(0.02)	(0.02)	(0.05)
σ_M	1.50***	1.41***	1.59***
	(0.01)	(0.02)	(0.05)
σ_{S_M}	0.50***	0.50***	0.48***
	(0.01)	(0.01)	(0.03)
σ_{S_R}	0.57***	0.55***	0.53***
	(0.01)	(0.01)	(0.02)
σ_Y	1.48***	1.50***	1.30***
	(0.01)	(0.01)	(0.03)

Note: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table provides parameter estimates for the years-of-education model described in Appendix C.4 for all students and then separately for white students and black students. Standard errors are computed by constructing the Hessian of the likelihood function using the outer product measure. To compute the outer product measure, we calculate two-sided numerical derivatives of the likelihood function for each estimated parameter. In each direction, the derivative is calculated by perturbing each parameter and then computing the likelihood.

Appendix D Identification

Here we argue that the main model discussed in section 4 has enough measurements to estimate the latent factors that we claim to identify. There are several ways to show this, each requiring slightly different assumptions. Here, we use arguments from Hu and Schennach (2008) to show that our model is non-parametrically identified. This implies that the model is also identified if we make parametric assumptions (Matzkin, 2007). In particular, we show that our main model is identified without distributional assumptions on θ_i, ϵ^Y, or e_{ji}, $j \in \{E, M, S_M, S_R, G\}$.

We start by writing down our model as a latent variable model. Let * denote an auxiliary random variable such that, for example, $Y_i = 1$ if $Y_i^* \geq 0$ and 0 otherwise. Then, the model can be written as:

$$\begin{aligned} Y_i^* &= c + \theta_i + G_i\beta + (T_{Ei} - \Phi(c + \theta_i + G_i\beta))\gamma_E + (T_{Mi} - \Phi(c + \theta_i + G_i\beta))\gamma_M + \epsilon_i^Y \\ T_{Ei}^* &= c_E + \phi_E\theta_i + G_i\beta_E + e_{Ei} \\ T_{Mi}^* &= c_M + \phi_M\theta_i + G_i\beta_M + e_{Mi} \\ S_{Mi} &= c_{S_M} + \phi_{S_M}\theta_i + e_{S_Mi} \\ S_{Ri} &= c_{S_R} + \phi_{S_R}\theta_i + e_{S_Ri} \\ G_i &= c_G + \phi_G\theta_i + e_{Gi} \end{aligned} \tag{19}$$

Note that we have assumed the following in section 4:

Assumption 1. $(\epsilon_i^Y, e_{S_Mi}, e_{S_Ri}, e_{Gi}, e_{Ei}, e_{Mi})$ *are jointly independent for all* i.

It follows that

$$S_{Mi} \perp S_{Ri} \perp G_i \perp T_{Ei}^* \perp T_{Mi}^* \perp Y_i^* | \theta_i.$$

Since Y is a function of Y^* alone, we have

$$Y \perp S_M \perp S_R | \theta_i.$$

Intuitively, independence is preserved because if Y^* does not contain information about S_j, $j \in \{M, R\}$ then Y should not contain information about them either. Notice that S_M and S_R are two continuous measures and $Y \in \{0, 1\}$ is a 0-1 dichotomous indicator of the latent variable Y^*.

Next, we note that our model satisfies the definition of a 2.1 model (reproduced below as Definition 1) from Hu (2015). Thereafter, we reproduce a theorem from Hu (2015) stating that a 2.1 model is identified.

Definition 1. *A 2.1-measurement model contains two measurements, X and Z, and a 0-1 dichotomous indicator $Y \in \{0,1\}$ of the latent variable X^* satisfying*

$$X \perp Y \perp Z | X^*.$$

Thus, (X, Y, Z) are jointly independent conditional on X^*.

Theorem 1. *From (Hu and Schennach, 2008): Under regularity assumptions, the 2.1-measurement model in Definition 1 with a continuous X^* is non-parametrically identified in the sense that the joint distribution of the three variables (X, Y, Z), $f_{X,Y,Z}$, uniquely determines the joint distribution of the four variables (X, Y, Z, X^*), f_{X,Y,Z,X^*}, which satisfies*

$$f_{X,Y,Z,X^*} = f_{X|X^*} f_{Y|X^*} f_{Z|X^*} f_{X^*}$$

In our case, Y is a binary indicator equal to one if the student completed a four-year college degree or more, and zero otherwise, $X = S_M$, $Z = S_R$ and $\theta = X^*$. Thus, our model meets the definition of the 2.1-measurement model given. Given that our system of equations satisfies the conditions under which the above theorem applies, we can identify the distribution of θ, ϵ^Y, e_E and e_M as well as the parameter values non-parametrically. It follows that identification is also achieved if we make parametric assumptions.

References

Hu, Yingyao, "Microeconomic Models with Latent Variables: Applications of Measurement Error Models in Empirical Industrial Organization and Labor Economics," Mimeo, Johns Hopkins University 2015.

_ **and Susanne M. Schennach**, "Instrumental Variable Treatment of Nonclassical Measurement Error Models," *Econometrica*, 2008, *76* (1), 195–216.

Matzkin, Rosa L, "Nonparametric Identification," *Handbook of Econometrics*, 2007, *6*, 5307–5368.