Supplemental Material

CBE—Life Sciences Education

Rodenbusch *et al.*

Supplemental online materials for *Early Engagement in Course-based Research Increases Graduation Rates and Completion of Science, Engineering, and Mathematics Degrees*

Stacia E. Rodenbusch, Paul R. Hernandez, Sarah L. Simmons, and Erin L. Dolan*

*Corresponding author. Email: edolan@austin.utexas.edu

This supplement contains the following:

Table S1. Summary of Socio-demographic Characteristics as a Function of FRI Status

Variable	*FRI (n = 2,449)*			*Matched Control (n = 2,449)*		
	%	*M*	*SD*	*%*	*M*	*SD*
Female	49.98			49.37		
Race						
Foreign	0.98			1.14		
Hispanic (not foreign)	23.27			22.83		
Black (not Hispanic)	5.84			4.98		
Native American	0.37			0.16		
Asian	33.69			34.22		
White	35.85			36.67		
First Generation	14.13			13.96		
Pell Grant Eligibility	28.75			28.83		
Parental Income		4.44	1.94		4.43	1.95
Parental Education		4.63	1.64		4.66	1.65
SAT (ACT Equivalent)		1272.92	160.87		1272.26	152.72
Propensity Score		0.50	0.04		0.50	0.06

Notes: Parental Income was measured on an 8-point scale (1 = less than $20,000 per year, 2 = $20,000-$39,999 per year, 3 = $40,000-$59,999 per year, 4 = $60,000-$79,999 per year, 5 = $80,000-$99,999 per year, 6 = $100,000-$149,999 per year, 7 = $150,000-$199,999 per year, 8 = more than $200,000 per year); Parental education measured the highest educational level achieved by either parent on a 7-point scale (0 = no high school, 1 = some high school, 2= High school or equivalent, 3 = some college, 4 = Associates degree, 5 = Baccalaureate/4-year degree, 6 = Graduate or professional degree).

Propensity Score Matching Methods. Random assignment to condition is the gold standard for making causal claims, but quasi-experimental designs are often used in cases where random assignment is unfeasible. Selection bias can occur however, as assignment to the treatment group is often correlated with attributes related to the outcomes of interest (43, 44). To control for selection bias, we made use of a propensity score matching procedure to identify an appropriate control group of non-participating students. Specifically, we matched FRI participants and non-FRI participants using a 1:1 nearest neighbor approach, without replacement, with caliper set to 0.20 standard deviations of the logit of the estimated propensity score, and with participants from both FRI and the non-FRI groups excluded if they were outside the area of common support (31). The propensity score matching procedure resulted in two groups of equal size (i.e., FRI group n = 2,449 & Matched control group n = 2,449). In order to check the quality of the match (i.e., covariate balance and reduction in bias), we plotted propensity score distributions (i.e., Q-Q, histogram, and distribution overlap), which indicated that the matching process resulted in groups nearly identical in their propensity to be in FRI (see Supplemental Figure 1). The estimate of selection bias, which is the difference in mean propensity scores for FRI and matched control groups, was reduced from 0.32 prior to matching to 0.01 after matching.

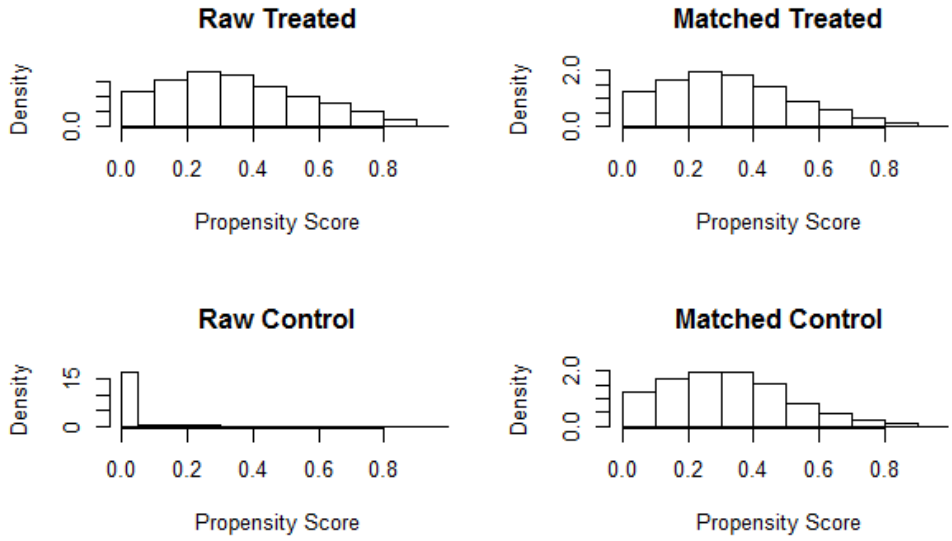

Figure S1. Propensity-score distributions of FRI and non-FRI Control groups before and after matching. "Raw" refers to the total sample before matching. "Matched" refers to those cases where matches were identified in areas of common support. "Treated" refers to FRI group, and "Control" refers to non-FRI group.

Table S2. Summary of differences between FRI and non-FRI groups on covariates used in matching process before and after propensity score matching. *d* = Cohen's d standardized mean difference.

Matching Variable	Pre-Match (N = 52,619)					Matched (N = 4,898)				
	M_{FRI}	SD_{FRI}	$M_{Control}$	$SD_{Control}$	d	M_{FRI}	SD_{FRI}	$M_{Control}$	$SD_{Control}$	d
HSTRMath	50.92	8.12	46.48	8.88	0.52	50.69	8.08	50.84	8.10	-0.02
HSTRScience	49.56	12.22	40.74	10.59	0.77	48.76	13.92	49.16	11.78	-0.03
Pell Grant eligible Yes/No	1.71	0.46	1.75	0.43	-0.10	1.71	0.45	1.71	0.46	-0.01
SAT Total (ACT equivalent)	1273.94	162.54	1215.14	160.11	0.36	1272.26	152.72	1272.92	160.87	0.00
Participant in Texas Interdisciplinary Program	1.81	0.40	1.97	0.16	-0.56	1.85	0.36	1.84	0.37	0.04
FirstMode=Out-of-state High school	0.03	0.18	0.06	0.24	-0.13	0.03	0.18	0.04	0.19	-0.01
FirstMode=TX senior college	0.00	0.06	0.12	0.33	-0.50	0.01	0.09	0.00	0.07	0.05
FirstMode=OUT-OF-STATE SENIOR COLLEGE	0.00	0.02	0.01	0.11	-0.15	0.00	0.02	0.00	0.02	0.00
FirstMode=TEXAS JUNIOR COLLEGE	0.00	0.02	0.03	0.17	-0.25	0.00	0.06	0.00	0.02	0.07
FirstMode=OUT-OF-STATE JUNIOR COLLEGE	0.00	0.02	0.00	0.03	-0.01	0.00	0.02	0.00	0.02	0.00
FirstMode=TRANSIENT	0.00	0.00	0.01	0.08	-0.12	0.00	0.02	0.00	0.00	0.03
FirstMode=PARALLEL ENROLLMENT PROGRAM	0.00	0.00	0.00	0.01	-0.02	-	-	-	-	-
FirstMode=SUMMER FRESHMAN ADMISSION	0.02	0.13	0.03	0.16	-0.06	0.02	0.13	0.02	0.13	-0.01
FirstSemesterCCYY=2007.0	0.09	0.29	0.16	0.36	-0.19	0.10	0.30	0.10	0.30	0.01
FirstSemesterCCYY=2008.0	0.13	0.33	0.14	0.35	-0.04	0.13	0.34	0.13	0.33	0.02
FirstSemesterCCYY=2009.0	0.15	0.36	0.15	0.36	0.00	0.15	0.36	0.16	0.36	-0.01
FirstSemesterCCYY=2010.0	0.17	0.38	0.15	0.36	0.06	0.18	0.38	0.17	0.38	0.01
FirstSemesterCCYY=2011.0	0.18	0.38	0.13	0.34	0.13	0.17	0.38	0.18	0.38	-0.02
FirstSemesterCCYY=2012.0	0.22	0.41	0.14	0.34	0.22	0.20	0.40	0.21	0.41	-0.02
FirstSemesterCCYY=2013.0	0.00	0.02	0.00	0.06	-0.07	0.00	0.00	0.00	0.02	-0.03
FirstSemesterS=Summer	0.04	0.20	0.04	0.21	-0.01	0.03	0.18	0.04	0.20	-0.03
FirstSemesterS=Fall	0.95	0.21	0.92	0.27	0.13	0.96	0.19	0.95	0.21	0.03
Sex=Male	0.51	0.50	0.47	0.50	0.08	0.51	0.50	0.50	0.50	0.01
RaceDescription=Hispanic (not foreign)	0.24	0.43	0.21	0.41	0.08	0.23	0.42	0.23	0.42	-0.01
RaceDescription=Black (not hispanic)	0.06	0.24	0.05	0.21	0.07	0.05	0.22	0.06	0.23	-0.04
RaceDescription=Native American	0.00	0.06	0.00	0.06	-0.01	0.00	0.04	0.00	0.06	-0.04
RaceDescription=Asian	0.33	0.47	0.20	0.40	0.32	0.34	0.47	0.34	0.47	0.01
RaceDescription=White	0.34	0.48	0.53	0.50	-0.38	0.37	0.48	0.36	0.48	0.02
FatherEDLevel=SOME HIGH SCHOOL	0.04	0.21	0.04	0.20	0.02	0.04	0.21	0.04	0.21	0.00
FatherEDLevel=HIGH SCHOOL DIPLOMA OR	0.10	0.31	0.11	0.32	-0.02	0.11	0.31	0.11	0.31	0.01

Matching Variable	Pre-Match (N = 52,619)					Matched (N = 4,898)				
	M_{FRI}	SD_{FRI}	$M_{Control}$	$SD_{Control}$	d	M_{FRI}	SD_{FRI}	$M_{Control}$	$SD_{Control}$	d
EQUIVALENT										
FatherEDLevel=SOME COLLEGE	0.11	0.32	0.14	0.35	-0.09	0.10	0.30	0.11	0.32	-0.03
FatherEDLevel=BACHELOR'S / 4-YEAR DEGREE	0.29	0.46	0.33	0.47	-0.08	0.30	0.46	0.30	0.46	-0.01
FatherEDLevel=GRAD / PROFESSIONAL DEGREE	0.37	0.48	0.31	0.46	0.12	0.37	0.48	0.36	0.48	0.01
FatherEDLevel=ASSOCIATE'S DEGREE	0.03	0.17	0.02	0.15	0.03	0.03	0.17	0.03	0.17	0.00
MotherEDLevel=SOME HIGH SCHOOL	0.05	0.22	0.04	0.19	0.06	0.05	0.22	0.05	0.22	0.00
MotherEDLevel=HIGH SCHOOL DIPLOMA OR EQUIVALENT	0.13	0.34	0.13	0.33	0.01	0.13	0.33	0.13	0.33	0.00
MotherEDLevel=SOME COLLEGE	0.14	0.34	0.17	0.38	-0.10	0.14	0.35	0.14	0.35	0.01
MotherEDLevel=BACHELOR'S / 4-YEAR DEGREE	0.35	0.48	0.39	0.49	-0.09	0.35	0.48	0.36	0.48	-0.01
MotherEDLevel= GRAD / PROFESSIONAL DEGREE	0.24	0.42	0.20	0.40	0.10	0.23	0.42	0.23	0.42	0.00
MotherEDLevel=ASSOCIATE'S DEGREE	0.05	0.23	0.04	0.19	0.09	0.06	0.23	0.05	0.23	0.00
ParentIncomeLevel=$20,000 - $39,999 yr	0.14	0.35	0.12	0.33	0.05	0.14	0.34	0.14	0.35	-0.01
ParentIncomeLevel=$40,000 - $59,999 yr	0.12	0.32	0.12	0.32	0.01	0.12	0.32	0.12	0.32	0.01
ParentIncomeLevel=$60,000 - $79,999 yr	0.10	0.30	0.10	0.31	0.00	0.10	0.30	0.11	0.31	-0.02
ParentIncomeLevel=$80,000 - $99,999 yr	0.29	0.45	0.38	0.48	-0.20	0.30	0.46	0.29	0.46	0.01
ParentIncomeLevel=$100,000 - $149,999 yr	0.13	0.34	0.08	0.28	0.15	0.12	0.33	0.13	0.33	-0.01
ParentIncomeLevel=$150,000 - $199,999 yr	0.07	0.26	0.05	0.21	0.12	0.07	0.26	0.07	0.26	0.01
ParentIncomeLevel=More than $200,000 yr	0.07	0.26	0.08	0.27	-0.02	0.07	0.26	0.07	0.26	-0.01
FirstSch=EDUCATION	0.00	0.02	0.04	0.19	-0.28	0.00	0.05	0.00	0.02	0.05
FirstSch=ENGINEERING	0.01	0.11	0.16	0.37	-0.54	0.01	0.11	0.01	0.12	-0.02
FirstSch=FINE ARTS	0.00	0.02	0.03	0.18	-0.26	0.00	0.05	0.00	0.02	0.05
FirstSch=PHARMACY	0.00	0.00	0.00	0.05	-0.06	-	-	-	-	-
FirstSch=ARCHITECTURE	0.00	0.00	0.01	0.08	-0.12	0.00	0.03	0.00	0.00	0.04
FirstSch=COMMUNICATION	0.00	0.02	0.08	0.27	-0.41	0.00	0.06	0.00	0.02	0.07
FirstSch=NATURAL SCIENCES	0.96	0.21	0.22	0.41	2.24	0.94	0.24	0.95	0.21	-0.05
FirstSch=GEOSCIENCES	0.00	0.00	0.01	0.09	-0.14	0.00	0.03	0.00	0.00	0.05
FirstSch=LIBERAL ARTS	0.02	0.13	0.25	0.43	-0.73	0.02	0.15	0.02	0.13	0.03
FirstSch=NURSING	0.00	0.02	0.02	0.12	-0.17	0.00	0.03	0.00	0.02	0.02
FirstSch=SOCIAL WORK	0.00	0.02	0.01	0.08	-0.11	0.00	0.00	0.00	0.02	-0.03
FirstSch=UNDERGRADUATE STUDIES	0.01	0.07	0.07	0.26	-0.35	0.01	0.08	0.01	0.08	0.01

Notes: HSTRMath = number of high school math courses taken; HSTRScience = number of high school science courses taken; FirstMode = mode of acceptance into UT Austin; FirstSemesterCCYY = Year of enrollment into UT Austin; d = standardized mean difference between FRI and non-FRI (control) group

Table S3. Summary of parameter estimates from the final step of the logistic regression analysis with STEM degree as the outcome (*n* = 1,624).

				98.3% C.I. for EXP(B)	
Predictors	*B*	*S.E.*	*EXP(B)*	*Lower*	*Upper*
Constant	0.90	0.20	2.45		
HSTRMath	0.02	0.01	1.02	1.00	1.05
HSTRSci	0.00	0.01	1.00	0.99	1.01
SAT Total	0.00	0.00	1.00	1.00	1.00
PellElig(1)	0.10	0.22	1.11	0.65	1.89
TIP(1)	-0.21	0.19	0.81	0.52	1.26
Sex(1)	-0.48	0.13	0.62	0.45	0.85
RaceDescription_r(1)	-0.40	0.18	0.67	0.44	1.03
RaceDescription_r(2)	0.29	0.27	1.34	0.71	2.53
RaceDescription_r(3)	0.31	0.16	1.37	0.92	2.03
FatherEDLevel_r(1)	0.19	0.18	1.21	0.79	1.86
FatherEDLevel_r(2)	-0.09	0.16	0.91	0.63	1.32
MotherEDLevel_r(1)	-0.02	0.17	0.98	0.66	1.46
MotherEDLevel_r(2)	-0.18	0.17	0.84	0.56	1.26
ParentIncomeLevel_r(1)	-0.02	0.24	0.98	0.55	1.73
ParentIncomeLevel_r(2)	0.34	0.20	1.41	0.88	2.25
ParentIncomeLevel_r(3)	0.31	0.24	1.36	0.76	2.44
FirstSch_d(1)	2.59	0.31	13.26	6.40	27.49
FirstMode_d(1)	0.34	0.25	1.41	0.77	2.58
FirstSemesterCCYYS_r(1)	-0.92	0.26	0.40	0.21	0.75
FirstSemesterCCYYS_r(2)	-0.41	0.25	0.66	0.36	1.22
FirstSemesterCCYYS_r(3)	-0.56	0.28	0.57	0.29	1.11
FirstSemesterCCYYS_r(4)	-0.99	0.35	0.37	0.16	0.85
FirstSemesterCCYYS_r(5)	-0.50	0.46	0.61	0.20	1.79
FirstSemesterCCYYS_r(6)	-0.70	0.86	0.50	0.06	3.90
Course 1	-0.37	0.22	0.69	0.41	1.16
Course 1 & 2	0.31	0.17	1.37	0.92	2.04
Course 1, 2 & 3	1.81	0.21	6.08	3.65	10.12

Notes: All continuous variables centered at their mean (i.e., HSTRMath, HSTRScience, & SAT Total); key predictors (Course 1, Courses 1 & 2, and Courses 1, 2 & 3) were entered into the model using dummy coding (reference group was the control). . All categorical control variables were entered into the model using Helmert contrast coding. In some cases, the levels of some control variables had to be collapsed for model convergence (due to low cell sizes or empty cells). The categorical control variables and levels were as follows: PellElig (1 = yes, 2 = no);

TIP (1 = yes, 2 = no); Sex (1 = female; 2 = male); RaceDescription (0 = Hispanic, 1 = foreign or Black or Native American, 2 = Asian, 3 = White); FatherEDLevel (0 = some college or less, 1 = college degree [2 or 4 year], 2 = advanced/professional degree); MotherEDLevel (0 = some college or less, 1 = college degree [2 or 4 year], 2 = advanced/professional degree); ParentIncomeLevel (0 = $39,999 or Less per year, 1 = $40,000 - $79,999 per year, 2 = $80,000 - $99,999 per year, 3 = More than $100,000 per year);First School (0 = STEM college [e.g., natural sciences], 1 = non-STEM college [e.g., Education]); FirstMode (0 = TX high school, 1 = other method of entry); FirstSemesterCCYYS (2006 [fall entry], 2007 [spring or fall entry], 2008 [summer, spring or fall entry], 2009 [summer, spring or fall entry], 2010 [summer, spring or fall entry], 2011 [summer, spring or fall entry], 2012 [summer, spring or fall entry], 2013 [summer entry]), $EXP^{(B)}$ = exponentiated B parameter estimate.

Table S4. Summary of parameter estimates from the final step of the logistic regression analysis with graduation within six years from any major as the outcome (*n* = 990).

				98.3% C.I.for EXP(B)	
Predictors	*B*	*S.E.*	*EXP(B)*	*Lower*	*Upper*
Constant	0.67	0.25	1.96		
HSTRMath	0.03	0.01	1.03	1.00	1.07
HSTRSci	0.00	0.01	1.00	0.98	1.02
SAT Total	0.00	0.00	1.00	1.00	1.00
PellElig(1)	-0.09	0.28	0.92	0.47	1.78
TIP(1)	0.13	0.22	1.13	0.67	1.92
Sex(1)	0.28	0.17	1.32	0.89	1.97
RaceDescription_r(1)	-0.28	0.21	0.76	0.46	1.24
RaceDescription_r(2)	-0.29	0.29	0.75	0.38	1.49
RaceDescription_r(3)	0.10	0.22	1.10	0.66	1.85
FatherEDLevel_r(1)	-0.43	0.22	0.65	0.38	1.10
FatherEDLevel_r(2)	0.04	0.21	1.05	0.63	1.73
MotherEDLevel_r(1)	-0.08	0.20	0.93	0.57	1.51
MotherEDLevel_r(2)	0.16	0.23	1.18	0.69	2.02
ParentIncomeLevel_r(1)	0.33	0.35	1.38	0.59	3.22
ParentIncomeLevel_r(2)	1.07	0.40	2.93	1.13	7.60
ParentIncomeLevel_r(3)	1.70	0.68	5.50	1.08	27.90
FirstSch_d(1)	-1.10	0.47	0.33	0.11	1.03
FirstMode_d(1)	0.15	0.29	1.16	0.58	2.31
FirstSemesterCCYYS_r(1)	-0.42	0.22	0.65	0.39	1.10
FirstSemesterCCYYS_r(2)	0.31	0.20	1.37	0.85	2.19
FirstSemesterCCYYS_r(3)	0.15	0.21	1.16	0.71	1.91
Course 1	-0.47	0.24	0.63	0.36	1.10
Course 1 & 2	0.07	0.21	1.07	0.64	1.78
Course 1, 2 & 3	0.89	0.25	2.43	1.34	4.43

Notes: All continuous variables centered at their mean (i.e., HSTRMath, HSTRScience, & SAT Total); key predictors (Course 1, Courses 1 & 2, and Courses 1, 2 & 3) were entered into the model using dummy coding (reference group was the control). All categorical control variables were entered into the model using Helmert contrast coding. In some cases, the levels of some control variables had to be collapsed for model convergence (due to low cell sizes or empty cells). The categorical control variables and levels were as follows: PellElig (1 = yes, 2 = no); TIP (1 = yes, 2 = no); Sex (1 = female; 2 = male); RaceDescription (0 = Hispanic, 1 = foreign or Black or Native American, 2 = Asian, 3 = White); FatherEDLevel (0 = some college or less, 1 = college degree [2 or 4 year], 2 = advanced/professional degree); MotherEDLevel (0 = some

college or less, 1 = college degree [2 or 4 year], 2 = advanced/professional degree); ParentIncomeLevel (0 = $39,999 or Less per year, 1 = $40,000 - $79,999 per year, 2 = $80,000 - $99,999 per year, 3 = More than $100,000 per year);First School (0 = STEM college [e.g., natural sciences], 1 = non-STEM college [e.g., Education]); FirstMode (0 = TX high school, 1 = other method of entry); FirstSemesterCCYYS (2006 [fall entry], 2007 [spring or fall entry], 2008 [summer, spring or fall entry], 2009 [summer, spring or fall entry], 2010 [summer, spring or fall entry], 2011 [summer, spring or fall entry], 2012 [summer, spring or fall entry], 2013 [summer entry]) , $EXP^{(B)}$ = exponentiated B parameter estimate.

Table S5. Multiple regression analysis with mid-point GPA as the outcome.

		Mid-College Cumulative GPA (n = 3,686)		
Step	Predicator variable	R^2	ΔR^2	$\Delta F(df)$
1	Controls[a]	.173	.173	31.94 (24)***
2	FRI Courses	.187	.012	20.49 (3)***

[a] The list of control variables is described in the Methods but was omitted here for simplicity.

*** $p \leq 0.001$

Table S6. Summary of parameter estimates from the final step of the multiple regression analysis with cumulative GPA at graduation as the outcome (n = 1,510*).*

				98.3% C.I.	
Predictors	*β*	*b*	*S.E.*	*Lower*	*Upper*
Intercept		3.39	0.02	3.33	3.44
HSTRMath	0.00	0.00	0.00	0.00	0.00
HSTRSci	0.01	0.00	0.00	0.00	0.00
SAT Total	0.00	0.00	0.00	0.00	0.00
Pell Grant eligible Yes/No	0.01	0.01	0.01	-0.01	0.03
Participant in Texas Interdisciplinary Program	0.01	0.01	0.01	-0.01	0.03
Sex of student	-0.03	-0.03	0.01	-0.04	-0.01
First School dummy coded into STEM vs non-STEM	0.00	0.01	0.01	-0.02	0.04
First Mode of entry into UT dummy coded into	0.00	0.00	0.01	-0.03	0.03
FirstSemesterCCYYS_r=2007.0	0.00	0.00	0.01	-0.03	0.02
FirstSemesterCCYYS_r=2008.0	0.01	0.01	0.01	-0.02	0.03
FirstSemesterCCYYS_r=2009.0	0.01	0.01	0.01	-0.02	0.03
FirstSemesterCCYYS_r=2010.0	0.03	0.03	0.01	0.00	0.06
FirstSemesterCCYYS_r=2011.0	0.01	0.01	0.01	-0.02	0.03
FirstSemesterCCYYS_r=2012.0	0.00	0.01	0.03	-0.05	0.07
RaceDescription_r=Other	-0.01	-0.02	0.01	-0.05	0.01
RaceDescription_r=Asian	-0.01	-0.01	0.01	-0.03	0.01
RaceDescription_r=White	0.00	0.00	0.01	-0.02	0.02
FatherEDLevel_r=College degree (2 or 4 year)	0.00	0.00	0.01	-0.02	0.02
FatherEDLevel_r=Advanced degree	0.00	0.00	0.01	-0.02	0.02
MotherEDLevel_r=College degree (2 or 4 year)	0.00	0.00	0.01	-0.01	0.02
MotherEDLevel_r=Advanced degree	0.01	0.01	0.01	-0.01	0.03
ParentIncomeLevel_r=$40,000 yr - $79,999 yr	-0.02	-0.02	0.01	-0.04	0.00
ParentIncomeLevel_r=$80,000 yr - $99,999 yr	-0.02	-0.02	0.01	-0.05	0.00
ParentIncomeLevel_r=More than $100,000 yr	-0.01	-0.01	0.01	-0.04	0.01
Course 1	-0.01	-0.01	0.01	-0.04	0.01
Courses 1 & 2	-0.01	-0.01	0.01	-0.03	0.01
Courses 1, 2, & 3	-0.02	-0.02	0.01	-0.03	0.00
Mid-point GPA	0.97	0.93	0.01	0.91	0.94

Notes: All continuous variables centered at their mean (e.g., HSTRMath); all categorical control variables were entered into the model using dummy-coding; key predictors (Course 1) were entered into the model using dummy coding. In some cases, the levels of some control variables had to be collapsed for model convergence (due to low cell sizes or empty cells). The categorical control variables and levels were as follows: PellElig (1 = yes, 2 = no); TIP (1 = yes, 2 = no); Sex (1 = female; 2 = male); RaceDescription (0 = Hispanic, 1 = Other, 2 = Asian, 3 = White); FatherEDLevel (0 = some college or less, 1 = college degree [2 or 4 year], 2 = advanced / professional degree); MotherEDLevel (0 = some college or less, 1 = college degree [2 or 4 year],

2 = advanced/professional degree); ParentIncomeLevel (0 = $39,999 or Less per year, 1 = $40,000 - $79,999 per year, 2 = $80,000 - $99,999 per year, 3 = More than $100,000 per year);First School (0 = STEM college [e.g., natural sciences], 1 = non-STEM college [e.g., Education]); FirstMode (0 = TX high school, 1 = other method of entry); FirstSemesterCCYYS (2006 [fall entry], 2007 [spring or fall entry], 2008 [summer, spring or fall entry], 2009 [summer, spring or fall entry], 2010 [summer, spring or fall entry], 2011 [summer, spring or fall entry], 2012 [summer, spring or fall entry], 2013 [summer entry]).