

Codebook for HERI Data Files:
ajps11.dta, abortiondata6.dta, gaydata6.dta, racedata6.dta, and nonaff.dta

All descriptive statistics and figures in the first subsection are for the sample of affluent students included in the primary model of taxation of the wealthy. A second subsection provides descriptive statistics for the College and Beyond sample of affluent students included in the replication model (Table 2 of the paper) and the models of cross-class contact effects. We also provide selected descriptive statistics for non-affluent students.

Variables were coded based on their distributions and conceptual definitions. For control dummy variables, we collapsed categories that had similar effects when this made no difference to estimates of interest.

Unless otherwise noted, all variables are measured using data from the HERI freshman survey with descriptive statistics based on the sample of affluent students present in the two-wave panel data, which links responses to the TFS and CSS (ajps11.dta).

HERI Variables

Independent Variables and Year and School Indicators	Coding summary (Distributional information provided elsewhere)	<i>Variable Name in Data File</i>
Proportion affluent in respondent's freshman cohort (including the preceding cohort)	Raw proportion affluent (continuous 0 to 1) Less than 37% affluent = 1; 0 = otherwise 37-49% affluent = 1; 0 = otherwise 49-59% affluent = 1; 0 = otherwise More than 59% affluent = 1; 0 = otherwise	<i>aff90_r1pre</i> <i>affpre_1</i> <i>affpre_2</i> <i>affpre_3</i> <i>affpre_45</i>
Year Freshman year	Numeric year student took the TFS	<i>YEAR_TFS</i>
Senior year	Numeric year student took CSS	<i>YEAR</i>
Freshman year cohort	Unique numeric identifier for school and freshman year. (All students entering a particular	<i>ace_year</i>

	school in the same year take the same value.)	
School	Unique numeric identifier for school in data set	<i>ACERECODE</i>
Case Study School	Indicator if respondent attends the school used in the case study 1= yes; 0 otherwise	<i>case_study</i>

Main Dependent Variable	Coding	Distributional Information	Variable Name in Data File
Tax the wealthy (ajps11.dta and nonaff.dta only) (TFS/CSS)	Agree strongly=0, Agree somewhat=0.33, Disagree somewhat=0.66, Disagree strongly=1	Freshman year (Lagged DV): 4-point scale, 0 to 1; Mean = 0.48; SD = 0.33	<i>taxes_tfs1</i>
<p>“Mark <u>one</u> in each row:”</p> <p>(“Wealthy people should pay a larger share of taxes than they do now”)</p> <p>Agree strongly (1); Agree somewhat (2); Disagree somewhat (3); Disagree strongly (4)</p>		Senior year: 4-point scale, 0 to 1; Mean = 0.50; SD = 0.32	<i>taxes_css1</i>

Placebo Dependent Variables	Coding	Distributional Information	Variable Name in Data File
Abortion (abortiondata6.dta only), Homosexual Relationships (gaydata6.dta), and Racial Discrimination (racedata6.dta) (TFS/CSS)	<i>Abortion</i> : Agree strongly=0, Agree somewhat=0.33, Disagree somewhat=0.66, Disagree strongly=1	Abortion: Freshman year (Lagged DV): 4-point scale, 0 to 1; Mean = 0.47; SD = 0.41 / Senior year: 4-point scale, 0 to 1; Mean = 0.41; SD = 0.39	<i>abortion_tfs1 (freshman)</i> ; <i>abortion_css1 (senior)</i>
<p>“Mark <u>one</u> in each row:”</p> <p>(“Abortion should be legal”; “It is important to have laws prohibiting homosexual relationships”; “Racial discrimination is no longer a major problem in America.”)</p> <p>Agree strongly (1); Agree somewhat (2); Disagree somewhat (3); Disagree strongly</p>	<p><i>Homosexual Relationships</i>: Agree strongly=0, Agree somewhat=0.33, Disagree somewhat=0.66,</p>	Homosexual Relationships: Freshman year (Lagged DV): 4-point scale, 0 to 1; Mean = 0.29; SD = 0.33 /	<i>gay_tfs1 (freshman)</i> ; <i>gay_css1 (senior)</i>

(4)	Disagree strongly=1 <i>Racial Discrimination:</i> Agree strongly=0, Agree somewhat=0.33, Disagree somewhat=0.66, Disagree strongly=1	Senior year: 4-point scale, 0 to 1; Mean = 0.21 SD = 0.30 Racial Discrimination: Freshman year (Lagged DV): 4-point scale, 0 to 1; Mean = 0.23; SD = 0.25 / Senior year: 4-point scale, 0 to 1; Mean = 0.21; SD = 0.24	<i>race_tfs1 (freshman);</i> <i>race_css1 (senior)</i>
-----	------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------

Individual Level Controls or Moderators	Coding	Distributional Information	Variable Name in Data File
SAT/ACT score ¹	<i>High standardized test score:</i> Score 1360 or above; yes=1; otherwise or missing=0	Score 1360 or above: 15%	<i>testtop15</i>
Self-reported high school GPA	<i>High H.S. GPA:</i> A- or above; yes=1; otherwise=0	A- or above: 58%	<i>Aminorabove</i>
College aspirations “In deciding to go to college, how important to you was each of the following reasons?” (“To be able to make more money”; “To learn more about things that interest me”) Not important (1); Somewhat important (2);	<i>Attend to make money:</i> Very important=1; otherwise=0 <i>Attend to gain knowledge:</i> Very important; yes=1; otherwise=0	Attend to make money: Very Important: 62% Attend to gain knowledge: Very important: 80%	<i>make_money_vimp</i> <i>get_knowledge_vimp</i>

¹ We measure achievement test scores using a joint measure of ACT and SAT performance based on the College Board’s SAT-ACT concordance table. We replace missing SAT scores with ACT scores where possible. The SAT re-centered the score scales in April 1995. This process reestablished the mean score to about 500 for college-bound seniors, the midpoint the on 200-800 score scale. Self-reported scores pre-recentering are adjusted, for the 1989-1995 cohorts, with the assumption that individuals took the SAT in the calendar year prior to the year they complete the TFS.

Very important (3)			
Gender “Your sex:” Male; Female	Female=1; Male=0 (reference)	Female: 58%	<i>female</i>
Race “Are you:” White/Caucasian; African American/Black; American Indian; Asian American/Asian; Mexican American/Chicano; Puerto Rican; Other Latino; Other	1. <i>White or Asian</i> : yes=1; otherwise=0 (reference) 2. <i>Black</i> : yes=1; otherwise=0 3. <i>Latino</i> : yes=1; otherwise=0 4. <i>Other race</i> (American Indian, Two or more race/ethnicity, Other): yes=1; otherwise=0	White or Asian: 92%; Black: 2% Latino: 2% Other race: 4%	<i>white; asian (omitted)</i> <i>black</i> <i>hispanic</i> <i>other_race</i>
Religious affiliation “Current religious preference:” Baptist; Buddhist; Eastern Orthodox; Episcopal; Islamic; Jewish; LDS (Mormon); Lutheran; Methodist; Presbyterian; Quaker; Roman Catholic; Seventh Day Adventist; United Church of Christ; Other Christian; Other Religion; None “Do you consider yourself a born-again Christian?” Yes; No	1. <i>Roman Catholic or Mainline Protestant</i> : yes=1; otherwise=0 (reference) (Includes: Congregational, Eastern Orthodox, Episcopal, Lutheran, Methodist, Presbyterian, Quaker, and other Christians who indicate they are not born again Christians.) 2. <i>Evangelical</i> : yes=1; otherwise=0 (Includes: Baptist, Seventh Day Adventist, and other Christians who indicate they are born again Christians.) 3. <i>Jewish</i> : yes=1; otherwise=0 4. <i>Other or no religion</i> : yes=1; otherwise=0	Roman Catholic or Mainline: 74% Evangelical: 11% Jewish: 2% Other or no religion: 14%	<i>romancatholic; mainline (omitted)</i> <i>evangelical</i> <i>jewish</i> <i>otherornorelig</i>

<i>In secondary models:</i>			
Distance from college “How many miles is this college from your permanent home?” (1) 10 or less (2) 11 to 50 (3) 51 to 100 (4) 101 to 500 (5) Over 500	<i>Lives within 100 mi:</i> yes=1; otherwise=0	Lives within 100 mi: 30%	<i>Constructed in “ajpsreplication.R” from variable FARHOME (coded.</i>
Close to home “How important was each reason in your decision to come here” (“I wanted to live near home.”) Not important (1); Somewhat important (2); Very important (3)	<i>Attend because close to home: very</i> <i>Important= 1;</i> <i>otherwise=0</i>	Attending because close to home: 10%	<i>Constructed in “ajpsreplication.R” from variable CHOOSE09_TFS.</i>
Business major (Intended) “Below is a list of different undergraduate major fields grouped into general categories. Mark only <u>one</u> oval to indicate your probable field of study.” ²	<i>Intended business major: yes=1;</i> <i>otherwise=0</i>	Intended business major: 15%	<i>busintend</i>
Business major (Actual) (CSS) “Below is a list of different major fields. Mark only <u>one</u> in each	Actual business major: ⁴ <i>1. Never intended or was business major-</i>	Never business: 76%	<i>neverbus (omitted)</i>

² List of majors available to students available: www.heri.ucla.edu/researchersToolsCodebooks.php

⁴ Business major includes: Business Administration (general), Finance, International Business, Marketing, Management, other Business

column” (“Undergraduate major (final or most recent)”)³	<p>yes=1; otherwise=0 (reference)</p> <p>2. <i>Changed into business major</i>- yes=1; otherwise=0</p> <p>3. <i>Changed from business major</i>- yes=1; otherwise=0</p> <p>4. <i>Always business major</i>- yes=1; otherwise=0</p>	<p>Changed into business: 7%</p> <p>Change from business: 5%</p> <p>Always business: 7%</p>	<p><i>changedintobus</i></p> <p><i>changedfrombus</i></p> <p><i>alwaysbus</i></p>
<p>Social embeddedness (CSS)</p> <p>Frequent socializer “During the past year, how much time did you spend during a typical week doing the following activities?” (“Socializing with friends”)</p> <p>None (1); Less than 1 hour (2); 1 to 2 hours (3); 3 to 5 hours (4); 6 to 10 hours (5); 11 to 15 hours (6); 16 to 20 hours (7); Over 20 hours (8)</p> <p>Greek life “Since entering college have you:” (“Joined a fraternity or sorority”)</p> <p>Yes (1); No (2)</p>	<p><i>Frequent socializer</i>: Socializing with friends more than 20 hours per week; yes=1 otherwise=0</p> <p><i>Greek life</i>: Joined a fraternity or sorority; yes=1; otherwise=0</p>	<p>Socializing more than 20 hours: 25%</p> <p>Joined fraternity or sorority: 26%</p>	<p><i>highsocialize</i></p> <p><i>joinedfrat</i></p>
<p>Frequency of political discussion over the past year (TFS/CSS)</p> <p>“Indicate which of the activities you did during the past year” (“Discussed politics”)</p> <p>Not at all (1); Occasionally (2); Frequently (3)</p>	<p><i>Political discussion</i>: not at all=0, occasionally= 0.5, frequently=1</p>	<p>Freshman year: 3-point scale, 0 to 1; Mean = 0.55; SD = 0.32; Range = 0, 1; Senior year: 3-point scale, 0 to 1; Mean = 0.50; SD = 0.31</p>	<p><i>Constructed from ACT08_TFS (freshman) and GENACT07 (senior) in “ajpssupplemental.R”</i></p>

³ Ibid.

Cohort Level Controls, Mediators or Moderators ⁵	Coding	Distributional Information	Variable Name in Data File
Selectivity	<p><i>Proportion high standardized test score:</i> Proportion first-time full-time freshmen with test scores of 1360 or above</p> <p><i>Proportion high H.S. GPA:</i> Proportion of students with self-reported high school GPA A- or above</p>	<p>Proportion high standardized test score: continuous, 0 to 1; Mean = 0.12; SD = 0.14; Range = 0.00, 0.85</p> <p>Proportion high H.S. GPA: continuous, 0 to 1; Mean = 0.52; SD = 0.20; Range = 0.03, 0.95</p>	<p><i>testtop15pre</i></p> <p><i>Aminorabovepre</i></p>
College aspirations	<p><i>Proportion attending to make money:</i> Proportion Indicating “Very important”;</p> <p>Treated as categorical in the moderation specification based on the affluent panel distribution:</p> <p>i. <i>Less than 58 perc. attend to make money.</i> Less than 58% =1; otherwise=0 (reference)</p> <p>ii. <i>58-70 perc. attend to make money:</i> 58-70% =1; otherwise=0</p> <p>iii. <i>More than 70 perc. attend to make</i></p>	<p>Proportion attending to make money: continuous, 0 to 1; Mean = 0.63; SD = 0.10; Range = 0.17, 0.89</p> <p>Less than 58 perc. attend to make money: Bottom 20% of distribution.</p> <p>58-70 perc. attend to make money: Middle 60% of the distribution</p> <p>More than 70 perc. attend to make money: Top 20% of</p>	<p><i>make_money_vimppre</i></p> <p><i>make_money_vimppre1</i></p> <p><i>make_money_vimppre234</i></p> <p><i>make_money_vimppre5</i></p>

⁵ For some variables, the number or proportions of students at a school do not vary substantially by year. In particular, school size and schools that are “mostly black” or “mostly female” (defined in this section) remain in these categories across all TFS years in our sample. We therefore treat these three variables (school size, mostly black, and mostly female) as school-level variables.

	<p><i>money</i>: 70% or more =1; otherwise=0</p> <p><i>Proportion attending to gain knowledge</i>: Proportion Indicating “Very important”</p>	<p>the distribution</p> <p>Proportion attending to gain knowledge: continuous, 0 to 1; Mean = 0.79; SD = 0.06; Range = 0.58, 0.97</p>	<p><i>get_knowledge_vimppre</i></p>
<p>Race (note: <i>Mostly Black</i> treated as school-level variable due to lack of variation between cohorts)</p>	<p>Proportion Black first-time full-time freshmen—<i>Mostly Black</i>: More than 89 percent Black (99th Percentile); yes=1; otherwise=0</p> <p>Proportion Latino first-time full-time freshmen—<i>Proportion Latino</i>: 5% or more Latino (85th Percentile); yes=1; otherwise=0</p> <p>Proportion other race first-time full-time freshmen—<i>Proportion other race</i>: 8% or more other race (85th Percentile); yes=1; otherwise=0</p>	<p>Mostly Black: 1%</p> <p>More than 5% Latino: Top 15% of sample</p> <p>More than 8% other race: Top 15% of sample</p>	<p><i>mostlyblack</i></p> <p><i>hisptop15</i></p> <p><i>otherracetop15</i></p>
<p>Female (note: treated as school-level variable due to lack of variation between cohorts)</p>	<p>Proportion Female first-time full-time freshmen—<i>Mostly female</i> (more than 95 percent); yes=1; otherwise=0</p>	<p>Mostly female: 2%</p>	<p><i>allfemale</i></p>
<p>Religious Affiliation</p>	<p><i>Proportion Evangelical</i>: More than 40% Evangelical (90th percentile); yes=1; otherwise=0.</p> <p><i>Proportion Jewish</i>: More than 5% Jewish (85th percentile); yes=1; otherwise=0</p> <p><i>Proportion other or no religion</i>: More than 30% other or no religion (85th percentile); yes=1; otherwise=0</p>	<p>More than 40% Evangelical: Top 10% of sample</p> <p>More than 5% Jewish: Top 15% of sample</p> <p>More than 30% other or no religion: Top 15% of sample</p>	<p><i>evangelicaltop10</i></p> <p><i>jewishtop15</i></p> <p><i>otherornoreligtop15</i></p>

	percentile); yes=1; otherwise=0		
<i>In secondary models</i>			
Cohort opposition to taxation	Mean cohort response to tax the wealthy item in freshman survey.	continuous, 0 to 1; Mean = 0.42; SD = 0.06; Range = 0.19, 0.60	<i>ctaxes_tfs1pre</i>
Business majors	<i>Proportion of cohort intending a business major</i>	continuous, Mean = 0.13; SD = 0.07; Range = 0 to 0.864	<i>busintendpre</i>

School-Level Controls	Coding	Distributional Information	Variable Name in Data File
School Size ⁶	Number of first-time full-time (FTFT) freshmen at school— <i>Large student body</i> : Greater than 1482 FTFT Freshmen; yes=1; otherwise=0	Large student body: Top 20% of sample	<i>sizepre_5</i>
School Type	<i>Public college or university</i> : public = 1; private = 0	Public: 9%	<i>is_public</i>
School region ⁷	<i>1. Midwest or West</i> :	Midwest or West: 43%	<i>mwest; west (omitted)</i>

⁶ Schools size could not be accurately measured by aggregating freshman survey data, so we use data from the IPEDS database, a repository for college statistics: <https://nces.ed.gov/ipeds/datacenter/>

⁷ We use regions defined by IPUMS-CPS: Northeast Region (New England Division: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont; Middle Atlantic Division: New Jersey, New York, Pennsylvania); Midwest (East North Central Division: Illinois, Indiana, Michigan, Ohio, Wisconsin; West North Central Division: Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota); South Region (South Atlantic Division: Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia; East South Central Division: Alabama, Kentucky, Mississippi, Tennessee; West South Central Division: Arkansas, Louisiana, Oklahoma, Texas);

Local economic context	<i>County-level inequality:</i> Gini coefficient in the county where a student's campus is located, measured with the 2000 census.	County-level inequality: continuous; Mean =0.45; SD = 0.04; Range = 0.37, 0.59	<i>GINI2000</i>
	<i>County-level poverty:</i> Percent living below the poverty line in the county where a student's campus is located, measured with the 2000 census.	County-level poverty: continuous; Mean =0.12; SD = 0.05; Range = 0.03, 0.33	<i>PERPOV2000</i>
	<i>Proportion cohort affluent based on respondent home being in zip code with median household level income in top 10 percent of region</i>	continuous, 0 to 1	<i>divper90affpre</i> , (individual indicator if respondent is affluent by this threshold: <i>divper90aff</i> : 1= Yes; 0= otherwise)
	<i>Proportion cohort affluent based on respondent home being in zip code with median household level income in top 10 percent of nation</i>	continuous, 0 to 1	<i>natper90affpre</i> , (individual indicator if respondent is affluent by this threshold: <i>natper90aff</i> : 1= Yes; 0= otherwise)

College and Beyond (CB) Variables (affluentcb2.dta)

The Mellon Foundation's College and Beyond (CB) dataset consists of students' freshman HERI surveys from 1989 merged with a College and Beyond follow-up survey administered in 1997, four years post-graduation.⁸ The CB data is unique in providing self-reported data from up to four schools to which students applied, as well as whether or not they were accepted to these schools.⁹ We use HERI data to

⁸ The schools analyzed include: Princeton University, Wesleyan University, Xavier University, Morehouse College, Penn State University, Bryn Mawr College, Wellesley College, Miami University – Ohio, Oberlin College, Stanford University, University of Pennsylvania, Williams College, Kenyon College, University of North Carolina, and Vanderbilt University.

⁹ Respondents are asked, "Back when you were applying to undergraduate schools, which school did you most want to attend, that is, what was your first choice school?" Students who indicated this was "another school" besides where they attended named this school and whether they applied to and were accepted by the school. Students were then asked: "In rough order of preference, please list the other

measure percent affluent for a subset of these schools where possible, allowing us to control for the median percent affluent at the schools to which students applied, excluding the school attended by the student.

The full 1989 CB survey consists of 9,549 respondents at 21 schools. The primary mode of surveying was through the mail, with respondents who did not respond to mailings contacted by phone. Seventy-six percent of matriculants who had entered these schools in the 1989 cohort responded to the survey. Six of the 21 schools did not participate in HERI's freshman survey and are excluded from this analysis. This leaves 5,671 respondents at 15 schools. Of these respondents, 2,969 are identified as affluent using a threshold of \$50,000. This threshold corresponds to the 77th percentile of the national household income distribution, which is lower than the 90th percentile threshold we use in the HERI analysis. We use this lower threshold to increase the sample size and statistical power. Percent affluent at the schools was measured as the percentage of students from families making over \$60,000 (84th percentile).

We were able to measure percent affluent for at least one unattended school to which the student submitted an application for 1,634 affluent students (55% of the affluent students who participated in the HERI and CB surveys). Of these students, 10% were missing data one or more variables used on our models, leaving us with a sample of 1,469 affluent students.

In Table 2 in the paper, we predict students' economic conservatism. Due to the substantial reduction in statistical power we trimmed the controls to the most essential cohort- and school-level variables, while maintaining the full range of individual-level controls. CB did not contain a freshman year version of the economic conservatism DV, so we added freshman-year ideological conservatism and preference for raising taxes to reduce the deficit in order to approximate a lagged DV. We also added a variable measuring the importance of having a career with high earnings, a variable that was not available in the full HERI sample.

Dependent Variable	Coding	Distributional Information	Variable Name in Data File
Economic conservatism "Thinking about your views concerning economic and social issues, where would you place yourself on the scale below:" ("Economic issues" ¹⁰)	1 (Very liberal)=0, 2=0.25, 3=0.5, 4=0.75, 5 (Very conservative)=1	5-point scale, 0 to 1; Mean = 0.51; SD = 0.28	<i>econ.views</i>

undergraduate schools you seriously considered. If there were more than three, list the three of most interest to you." Students were asked whether they applied to and were accepted by the schools.

¹⁰ "Social issues" was then asked about separately.

1 (Very liberal); 2; 3; 4; 5 (Very conservative)			
--------------------------------------------------------	--	--	--

School Level Predictor	Coding	Distributional Information	Variable Name in Data File
Percent affluent	<p>1. <i>Less than 50% affluent:</i> yes=1; otherwise=0 (reference)</p> <p>2. <i>50 - 64% affluent:</i> yes=1; otherwise=0</p> <p>3. <i>More than 64% affluent:</i> yes=1; otherwise=0</p>	<p>Less than 50% affluent: 16%</p> <p>50 - 64% affluent: 45%</p> <p>More than 64% affluent: 39%</p>	<p><i>low.aff</i></p> <p><i>med.aff</i></p> <p><i>high.aff</i></p>
Applied to affluent and non-affluent schools subset (CB)	<p>1. <i>Accepted to non-affluent schools only:</i> yes=1; otherwise=0 (reference)</p> <p>2. <i>Accepted to affluent schools only:</i> yes=1; otherwise=0</p>	<p>Accepted to non-affluent schools only: 58%</p> <p>Accepted to affluent schools only: 42%</p>	<p><i>aff.control</i></p> <p><i>aff.treat</i></p>

Individual Level Controls	Coding	Distributional Information	Variable Name in Data File
Application median percent affluent	<p><i>Application median percent affluent:</i> Median percent affluent at schools applied to by the student, not including the school attended. Students were able to list up to four colleges to which they applied, as well as whether or not they were accepted.</p>	<p>continuous, 0 to 1; Mean = 0.55; SD = 0.10; Range = 0.08, 0.68</p>	<i>med.peraffluent.natt</i>
<p>Ideological conservatism</p> <p>“How would you characterize your political views?”</p> <p>Far right (1); Conservative (2); Middle-of-the-road (3); Liberal (4); Far left</p>	<p>Far left=0, Liberal=0.25, Middle-of-the-road=0.5, Conservative=0.75, Far right=1</p>	<p>5-point scale, 0 to 1; Mean = 0.45; SD = 0.21</p>	<i>conservatism</i>

(5)			
<p>Raise taxes to reduce deficit</p> <p>“Please indicate your agreement with the following statements” (“The federal government should raise taxes to reduce the deficit.”)</p> <p>Agree strongly (1); Agree somewhat (2); Disagree somewhat (3); Disagree strongly (4)</p>	<p><i>Raise tax to reduce deficit:</i> Agree strongly=0, Agree somewhat=0.33, Disagree somewhat=0.66, Disagree strongly=1</p>	<p>4-point scale, 0 to 1; Mean = 0.49; SD = 0.28</p>	<p><i>rataxreddef</i></p>
<p>Important to have a career with high earnings¹¹</p> <p>“Which of the following are important to you in your long-term choice of career occupation?” (“High anticipated earnings”)</p> <p>Not important (1); Somewhat important (2) Very important (3); Essential (4)</p>	<p><i>Important career with high earnings:</i> Not important=0, Somewhat important=0.33, Very important=0.66, Essential=1</p>	<p>4-point scale, 0 to 1; Mean = 0.53; SD = 0.30</p>	<p><i>carearnings</i></p>
SAT/ACT score	<p><i>High standardized test score:</i> Score 1360 or above; yes=1; otherwise or missing=0</p>	<p>Score 1360 or above: 50%</p>	<p><i>testtop15</i></p>
Self-reported high school GPA	<p><i>High H.S. GPA:</i> A- or above; yes=1; otherwise=0</p> <p><i>Missing H.S. GPA</i>¹²: yes=1; otherwise=0</p>	<p>A- or above: 59%</p> <p>Missing H.S. GPA: 13%</p>	<p><i>Aminorabove.yes</i></p> <p><i>Aminorabove.miss</i></p>

¹¹ This variable was not available for the basic HERI analysis.

¹² One school in the College and Beyond dataset was missing High School GPA for all its students, so we insert a dummy for missingness on this variable to avoid dropping this school from the analysis.

College aspirations (See HERI section for wording)	<p><i>Attend to make money:</i> Very important = 1; otherwise=0</p> <p><i>Attend to gain knowledge:</i> Very important; yes=1; otherwise=0</p>	<p>Attend to make money: 48%</p> <p>Attend to gain knowledge: 85%</p>	<p><i>make.money.vimp</i></p> <p><i>get.knowledge.vimp</i></p>
Gender	<i>Female</i> = 1; Male = 0 (reference)	Female: 51%	<i>female</i>
Race	<p>1. <i>White or Asian:</i> yes=1; otherwise=0 (reference)</p> <p>2. <i>Black:</i> yes=1; otherwise=0</p> <p>3. <i>Latino:</i> yes=1; otherwise=0</p> <p>4. <i>Other race</i> (American Indian, Two or more race/ethnicity, Other): yes=1: otherwise=0</p>	<p>White or Asian: 88%</p> <p>Black: 7%</p> <p>Latino: 2%</p> <p>Other race: 3%</p>	<p>(reference, omitted)</p> <p><i>black</i></p> <p><i>hispanic</i></p> <p><i>other.race</i></p>
Religious affiliation	<p>1. <i>Roman Catholic or Mainline Protestant:</i> yes=1; otherwise=0 (reference) (Includes: Congregational, Eastern Orthodox, Episcopal, Lutheran, Methodist, Presbyterian, Quaker, and other Christians who indicate they are not born again Christians.)</p> <p>2. <i>Evangelical:</i> yes=1; otherwise/NA=0 (Includes: Baptist, Seventh Day Adventist)</p> <p>3. <i>Jewish:</i> yes=1; otherwise=0</p> <p>4. <i>Other or no religion:</i> yes=1; otherwise=0</p>	<p>Roman Catholic or Mainline: 51%</p> <p>Evangelical: 12%</p> <p>Jewish: 10%</p> <p>Other or no religion: 27%</p>	<p>(reference, omitted)</p> <p><i>evangelical</i></p> <p><i>jewish</i></p> <p><i>otherornorelig</i></p>
<i>In secondary models:</i>			

Cross-class contact (CB)	<i>Knew 2+ poorer students well:</i> yes=1; otherwise=0	Knew 2+ poorer students well: 66%	<i>met.poor2</i>
“Did you get to know 2 or more of these students <u>well</u> while in school?”	<i>Knew 2+ wealthier student well:</i> yes=1; otherwise=0	Knew 2+ wealthier student well: 72%	<i>met.wealthy2</i>
(“From a family much poorer than yours?”; “From a family much wealthier than yours?”)			
Yes (1); No (0)			

School Level Controls	Coding	Distributional Information	Variable Name in Data File
Selectivity	<i>Proportion High standardized test score:</i> Proportion first-time full-time freshmen with test scores of 1360 or above <i>Proportion High H.S. GPA:</i> Proportion of students with self-reported high school GPA A- or above	Proportion High test score: continuous, 0 to 1; Mean = 0.41; SD = 0.24 Range = 0.00, 0.69 Proportion High H.S. GPA: continuous, 0 to 1; Mean = 0.72; SD = 0.16; Range = 0.22, 0.96	<i>per.testtop15</i> <i>per.Aminorabove</i>
Race	<i>All Black:</i> HBCU=1; otherwise=0	All Black: 2%	<i>mostlyblack</i>
Female	<i>All female:</i> All female=1; otherwise=0	All female: 5%	<i>allfemale</i>
Institutional Identifier	Numeric value indicating data is from a unique school	15 schools in the analysis	<i>inst</i>