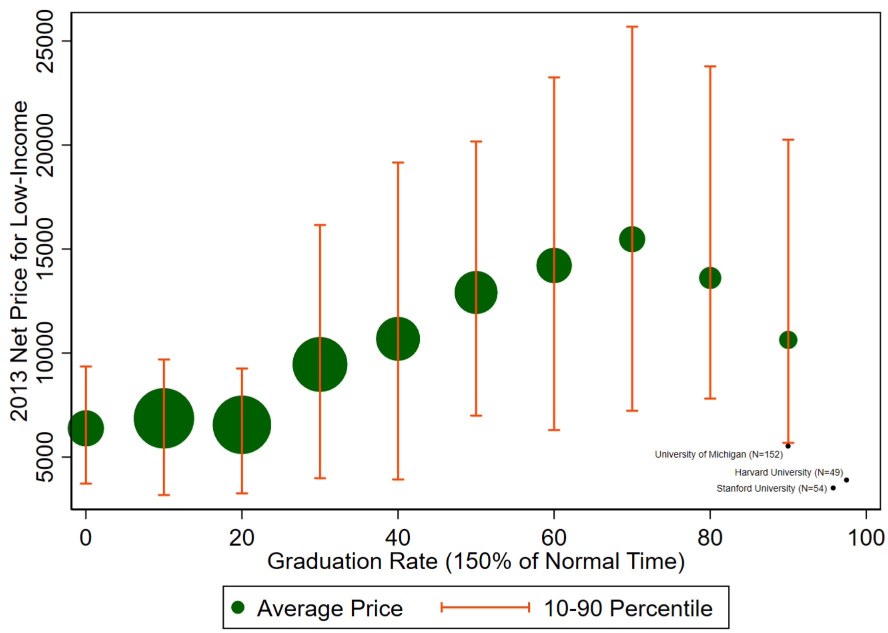
Supplemental Appendix
Increasing Degree Attainment Among Low-Income Students: The Role of Intensive Advising and College Quality
Andrew Barr and Benjamin Castleman

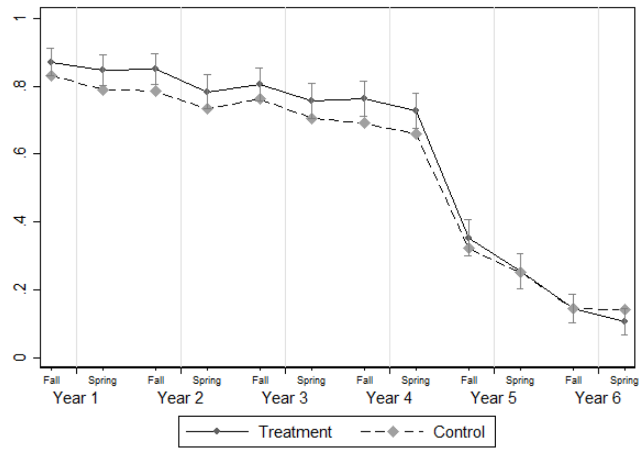
Appendix A: Supplemental Figures and Tables

Figure A1: Net Prices and Graduation Rates among Institutions Attended by Low-Income Students

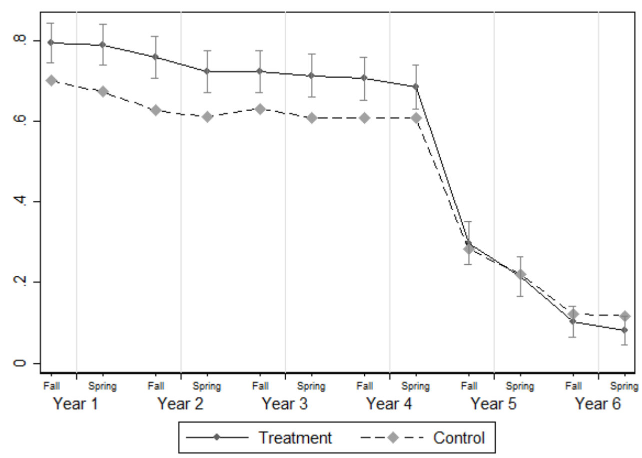


Note: Figure illustrates the variation in graduation rates and average net prices of colleges attended by low-income students. Each bubble indicates the average net price for low-income students within each 10 percentage point graduation rate bin (i.e., 0-10, 10-20, etc.). Bubbles sizes correspond to the fraction of low-income students falling within each graduation rate bin. Graduation rates indicate the percentage of students graduating within 150% of the standard time (e.g., six years for a four-year degree). Net prices are the average net price paid by individuals from the bottom quintile of the income distribution in 2013. The error bars (capturing the 10th and 90th percentile) indicate the range of net prices paid by low-income students within each graduation rate bin. Harvard University, Stanford University, and the University of Michigan are included as individual data points illustrating each institution's net price, graduation rate, and number of low-income students (N). Estimates produced using Chetty Mobility Report Card Data.

Figure A2: College Enrollment Over Time: First Cohort



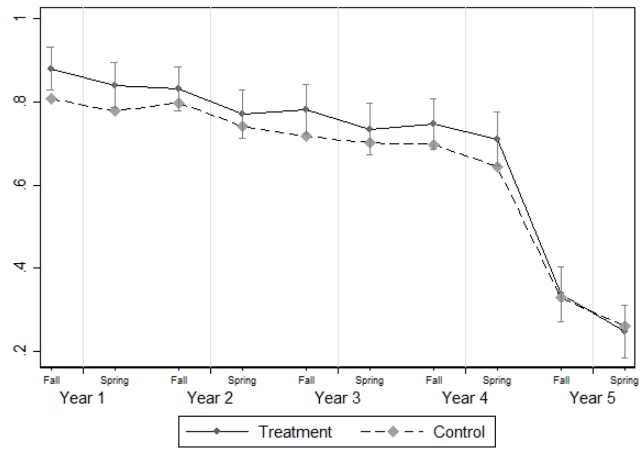
(a) Enrolled



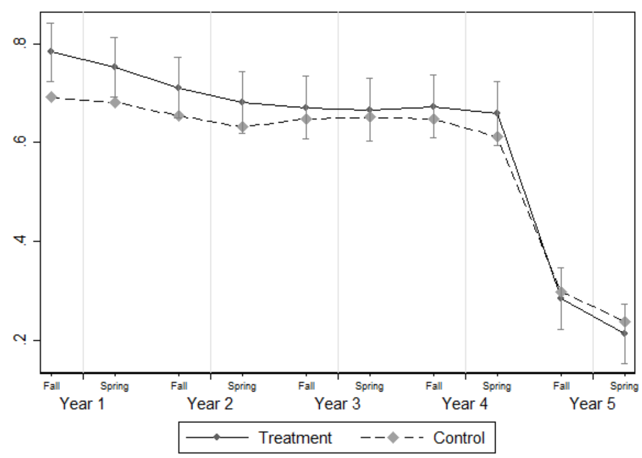
(b) Enrolled Four-Year

Note: Figure illustrates the share of control and treated students enrolled in college in each semester. Year 1 is the academic year beginning in the fall after each high school class's senior year in high school. Treatment line points add estimated treatment effect from equation (1) to the control means. A student is considered enrolled in the fall if enrollment begins after June but by October and lasts at least 30 days. A student is considered enrolled in the spring if enrollment begins after November but by February and lasts at least 30 days. Panel B restricts this variable to enrollment in four-year institutions. Sample restricted to the first experimental cohort. Statistics derived from BL and NSC data.

Figure A3: College Enrollment Over Time: Second Cohort



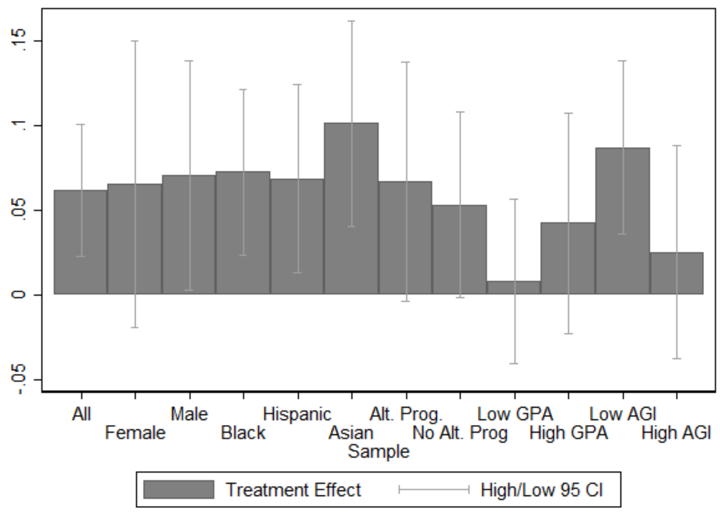
(a) Enrolled



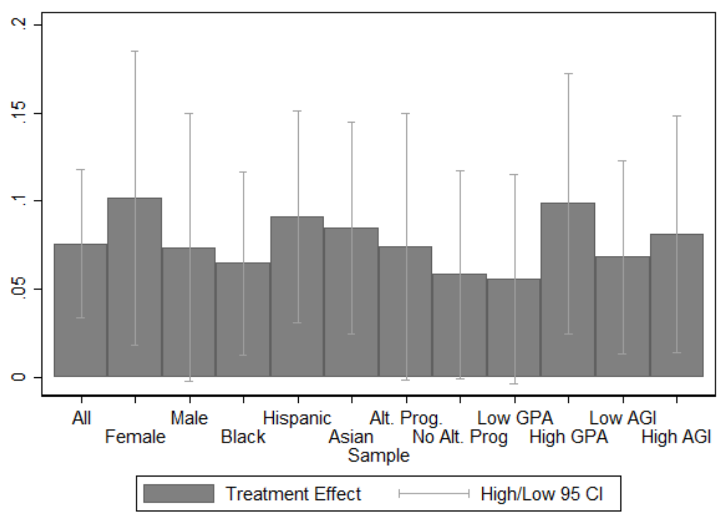
(b) Enrolled Four-Year

Note: Figure illustrates the share of control and treated students enrolled in college in each semester. Year 1 is the academic year beginning in the fall after each high school class's senior year in high school. Treatment line points add estimated treatment effect from equation (1) to the control means. A student is considered enrolled in the fall if enrollment begins after June but by October and lasts at least 30 days. A student is considered enrolled in the spring if enrollment begins after November but by February and lasts at least 30 days. Panel B restricts this variable to enrollment in four-year institutions. Sample restricted to the second experimental cohort. Statistics derived from BL and NSC data.

Figure A4: Bachelor's Degree Attainment: Heterogeneity



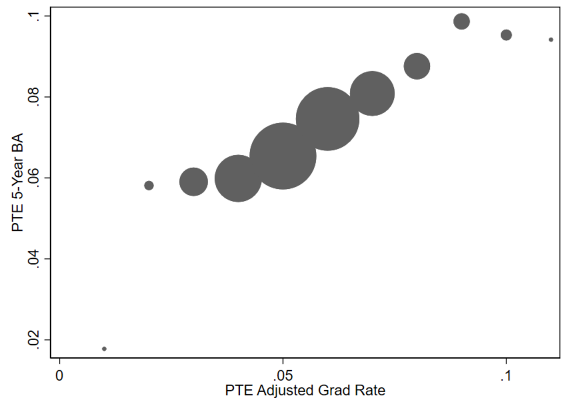
(a) 4 Year Attainment



(b) 5 Year Attainment

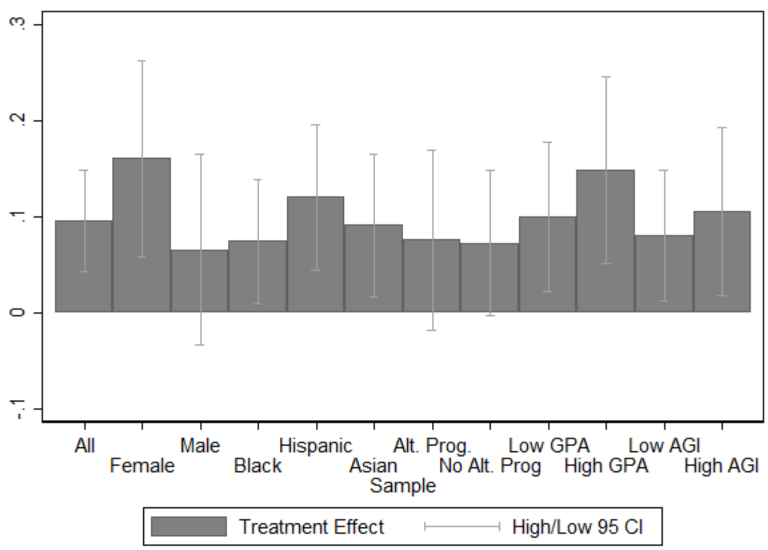
Note: Figure illustrates estimated treatment effects on bachelor's degree attainment by subgroup. Treatment effect estimates provided by estimating equation (1) within each subgroup. Bachelor's degree attainment defined as being awarded a bachelor's degree by X (i.e., 4 or 5) years from July 1 of the each high school class's senior year. Statistics derived from BL and NSC data.

Figure A5: The Relationship Between Personalized Treatment Effects (PTEs) on Quality of Initial Enrollment Choice and 5-Year Bachelor's Degree Receipt



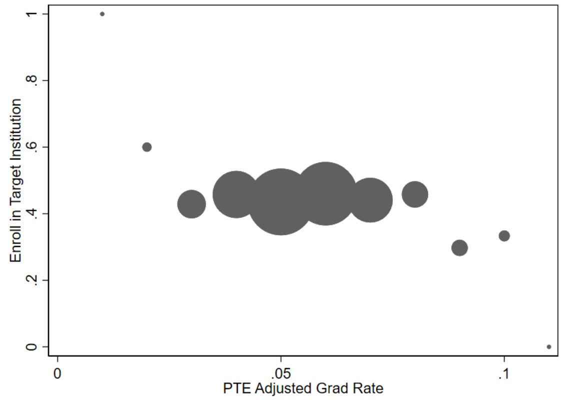
Note: Figure illustrates the relationship between the personalized treatment effects (PTEs) on quality of initial enrollment choice (i.e., the adjusted graduation rate) and 5-year bachelor's degree receipt. For enrolled students, we construct the adjusted graduation rate equal to the 2015 graduation rate (within 150 percent of normal time) for the school the student was enrolled in the fall after their senior year of high-school. For students who were not observed enrolled in the fall after their senior year, we estimate the 6-year bachelor's degree attainment rate of students in the control group who were unenrolled in the fall after their senior year but who subsequently enrolled and completed a bachelor's degree, and assign that number as the adjusted graduation rate. Additional discussion is provided in the text. PTEs are produced using covariates included in main specification. See text and Appendix B for additional explanation regarding the causal forest approach and details of our implementation. We collapse the individual PTEs for bachelor's degree attainment into means within bins of size 0.01. The size of each dot reflects the sample size represented. Statistics derived from BL and NSC data.

Figure A6: Bachelor's Degree Attainment (6 Years): Heterogeneity, First Cohort



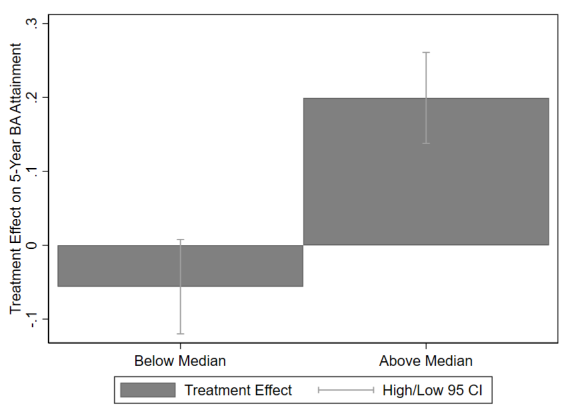
Note: Figure illustrates estimated treatment effects on bachelor's degree attainment by subgroup. Treatment effect estimates provided by estimating equation (1) within each subgroup. Bachelor's degree attainment defined as being awarded a bachelor's degree by 6 years from July 1 of the each high school class's senior year. Sample restricted to the first experimental cohort. Statistics derived from BL and NSC data.

Figure A7: The Relationship Between Personalized Treatment Effects on High Quality Enrollment and Enrollment in a Target Institution



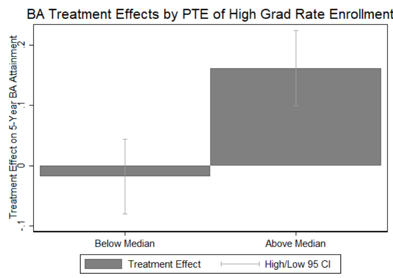
*Note:*Figure illustrates the relationship between the personalized treatment effects (PTEs) on quality of initial enrollment choice (i.e., the adjusted graduation rate) and enrollment in a target institution. For enrolled students, we construct the adjusted graduation rate equal to the 2015 graduation rate (within 150 percent of normal time) for the school the student was enrolled in the fall after their senior year of high school. For students who were not observed enrolled in the fall after their senior year, we estimate the 6-year bachelor's degree attainment rate of students in the control group who were unenrolled in the fall after their senior year but who subsequently enrolled and completed a bachelor's degree, and assign that number as the adjusted graduation rate. Additional discussion is provided in the text. PTEs are produced using covariates included in main specification. See text and Appendix B for additional explanation regarding the causal forest approach and details of our implementation. We collapse the individual PTEs for bachelor's degree attainment into means within bins of size 0.01. The size of each dot reflects the sample size represented. Statistics derived from BL and NSC data.

Figure A8: Effects on 5-Year Bachelor's Degree Attainment: Heterogeneity by Adjusted Graduation Rate PTE



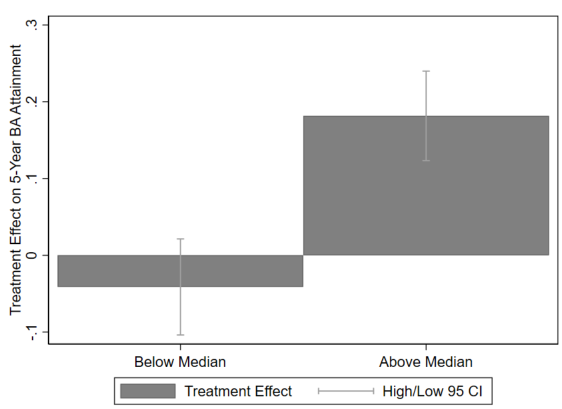
Note: Figure provides estimated effect of treatment on 5-year bachelor's degree attainment by above or below median PTE on adjusted graduation rate. PTEs produced using covariates included in main specification. See prior figure note or text for additional details. Statistics derived from BL data.

Figure A9: Effects on 5-Year Bachelor's Degree Attainment: Heterogeneity by High-Graduation Rate Enrollment PTE



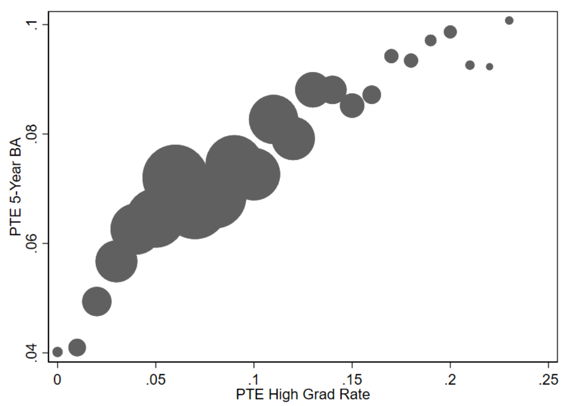
Note: Figure provides estimated effect of treatment on 5-year bachelor's degree attainment by above or below median PTE on initial enrollment in a high-graduation rate (above median) college. PTEs produced using covariates included in main specification. See prior figure note or text for additional details. Statistics derived from BL data.

Figure A10: Effects on 5-Year Bachelor's Degree Attainment: Heterogeneity by Four-Year Enrollment PTE



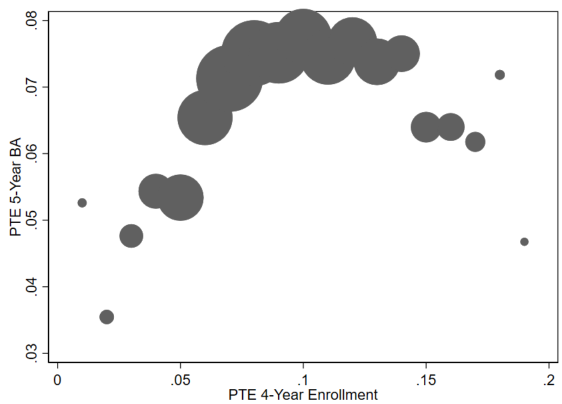
Note: Figure provides estimated effect of treatment on 5-year bachelor's degree attainment by above or below median PTE on initial enrollment in a four-year college. PTEs produced using covariates included in main specification. See prior figure note or text for additional details. Statistics derived from BL data.

Figure A11: The Relationship Between Personalized Treatment Effects on Enrollment in a High-Graduation Rate College and 5-Year Bachelor's Degree Receipt



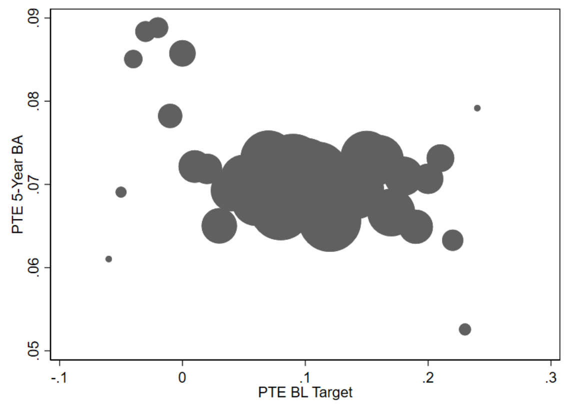
Note: Figure illustrates the relationship between the personalized treatment effects (PTEs) on initial enrollment in a high-graduation rate (above median) college and 5-year bachelor's degree receipt. For enrolled students, we construct the adjusted graduation rate equal to the 2015 graduation rate (within 150 percent of normal time) for the school the student was enrolled in the fall after their senior year of high school. For students who were not observed enrolled in the fall after their senior year, we estimate the 6-year bachelor's degree attainment rate of students in the control group who were unenrolled in the fall after their senior year but who subsequently enrolled and completed a bachelor's degree, and assign that number as the adjusted graduation rate. Additional discussion is provided in the text. PTEs are produced using covariates included in main specification. See text and Appendix B for additional explanation regarding the causal forest approach and details of our implementation. We collapse the individual PTEs for bachelor's degree attainment into means within bins of size 0.01. The size of each dot reflects the sample size represented. Statistics derived from BL and NSC data.

Figure A12: The Relationship Between Personalized Treatment Effects on Four-Year Enrollment and 5-Year Bachelor's Degree Receipt



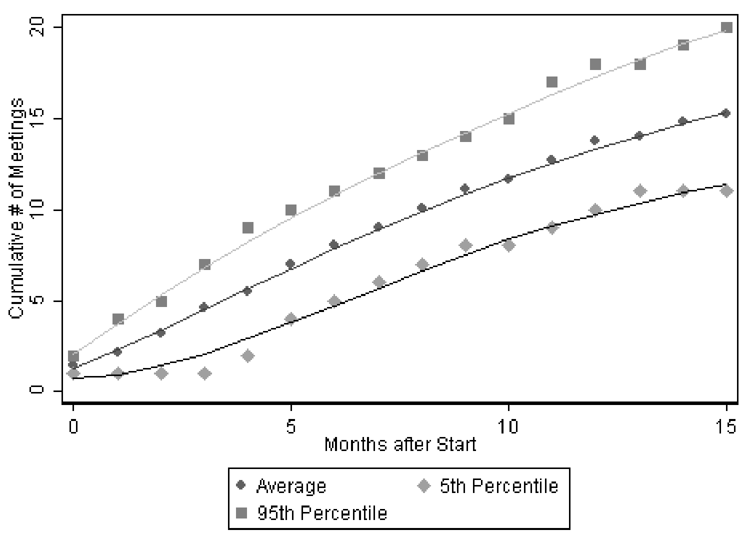
Note: Figure illustrates the relationship between the personalized treatment effects (PTEs) on quality of initial enrollment in a four-year college and 5-year bachelor's degree receipt. For enrolled students, we construct the adjusted graduation rate equal to the 2015 graduation rate (within 150 percent of normal time) for the school the student was enrolled in the fall after their senior year of high school. For students who were not observed enrolled in the fall after their senior year, we estimate the 6-year bachelor's degree attainment rate of students in the control group who were unenrolled in the fall after their senior year but who subsequently enrolled and completed a bachelor's degree, and assign that number as the adjusted graduation rate. Additional discussion is provided in the text. PTEs are produced using covariates included in main specification. See text and Appendix B for additional explanation regarding the causal forest approach and details of our implementation. We collapse the individual PTEs for bachelor's degree attainment into means within bins of size 0.01. The size of each dot reflects the sample size represented. Statistics derived from BL and NSC data.

Figure A13: The Relationship Between Personalized Treatment Effects on Target College Enrollment and 5-Year Bachelor's Degree Receipt



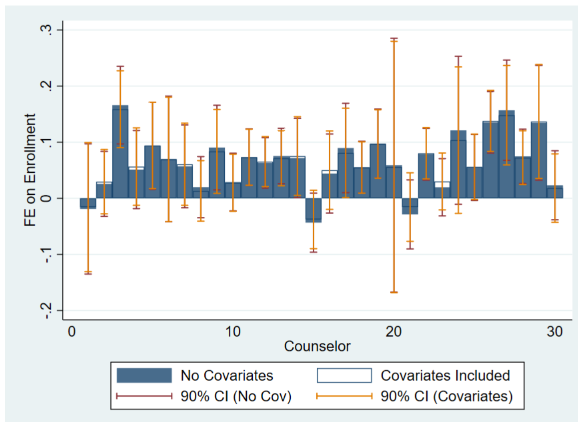
Note: Figure illustrates the relationship between the personalized treatment effects (PTEs) on initial enrollment at a BL target institution and 5-year bachelor's degree receipt. For enrolled students, we construct the adjusted graduation rate equal to the 2015 graduation rate (within 150 percent of normal time) for the school the student was enrolled in the fall after their senior year of high school. For students who were not observed enrolled in the fall after their senior year, we estimate the 6-year bachelor's degree attainment rate of students in the control group who were unenrolled in the fall after their senior year but who subsequently enrolled and completed a bachelor's degree, and assign that number as the adjusted graduation rate. Additional discussion is provided in the text. PTEs are produced using covariates included in main specification. See text and Appendix B for additional explanation regarding the causal forest approach and details of our implementation. We collapse the individual PTEs for bachelor's degree attainment into means within bins of size 0.01. The size of each dot reflects the sample size represented. Statistics derived from BL and NSC data.

Figure A14: Advisor Interaction Patterns over Time

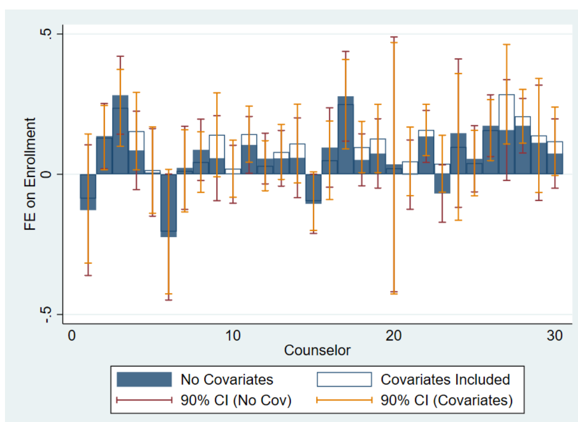


Note: Statistics derived from BL data. Figure illustrates the average cumulative number of meetings with an advisor over time. Month 0 is May of each high school class's junior year. Statistics derived from BL data.

Figure A15: Advisor Fixed Effect Estimates



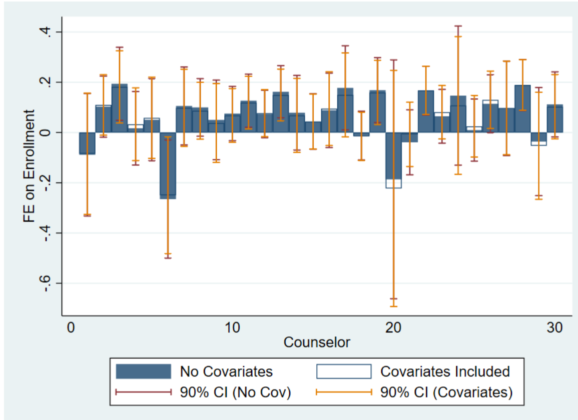
(A) Adjusted Graduation Rate



(B) Enrolled in High Grad Rate College

Note: Figure plots individual advisor effects on measures of initial enrollment quality. Estimates derived from estimating basic specification (see equation 1 in the text) but replacing the treatment indicator with advisor fixed effects. Statistics derived from BL and NSC data.

Figure A16: Advisor Fixed Effect Estimates on 5-Year BA Attainment



Note: Figure plots individual advisor effects on measures of initial enrollment quality. Estimates derived from estimating basic specification (see equation 1 in the text) but replacing the treatment indicator with advisor fixed effects. Statistics derived from BL and NSC data.

Table A1: Target Colleges

College Names	Graduation Rate	Tuition and Fees	Net Price (0-48K)
Bentley University	84.1	41110	20544
Boston College	92.2	45622	16196
Boston University	83.9	44910	23573
Bridgewater State University	54.4	8053	14680
Buffalo State SUNY	48.1	7022	8021
CUNY Hunter College	45.7	6129	5258
CUNY John Jay College of Criminal Justice	43.1	6059	3993
CUNY Lehman College	34.9	6108	3297
CUNY New York City College of Technology	13.6	6069	5220
CUNY York College	25.6	6096	4590
Clark University	79.8	39550	18293
College of the Holy Cross	92.9	44272	15607
Fitchburg State University	50.8	8985	9013
Fordham University	81	43577	23352
Framingham State University	51.5	8080	12515
MCPHS University	66.4	28470	29807
Northeastern University	78.5	41686	20140
SUNY at Albany	64.4	8040	11019
Saint Joseph's College-New York	67.5	21878	10292
Salem State University	45.4	8130	11800
St Francis College	51.9	20700	9448
State University of New York at New Paltz	72.7	7083	9844
Suffolk University	55.9	31716	22900
The Sage Colleges	51.8	28000	14834
University of Massachusetts-Amherst	70.4	13258	12437
University of Massachusetts-Boston	37.9	11966	8084
University of Massachusetts-Dartmouth	49.9	11681	12581
University of Massachusetts-Lowell	53.8	12097	10258
Wentworth Institute of Technology	64	29200	25754
Worcester Polytechnic Institute	83.5	42778	27224
Worcester State University	51	8157	10907
Mean	59.6	20854	13919

Table A2: Treatment and Control Assignments

	Boston	New York	Worcester	Total
Control	193	450	92	735
Treatment	860	582	245	1,687

Table A3: Advisor Interaction Patterns

	Mean
Ever Interact with Student (proportion):	0.97
Office Meeting	0.95
Interactions per Student (number):	13.07
By Medium:	
Office Meeting	8.81
Phone Meeting	0.42
Text or Email	0.28
By Subject:	
First Meeting	2.13
Second Meeting	1.37
Application Meeting	3.47
Financial Aid Meeting	2.03
Missed Meetings	0.59
Estimated Contact Time per Student (hours):	10-15

Note: Table provides summary information on the interaction of students assigned to treatment with advisors. Sample for rows (1)-(3) includes all students assigned to treatment and has a sample size of 1687. Remaining rows are restricted to the 97.2 percent of students assigned to treatment who had any post-assignment interaction with BL. Sample size for these rows is 1639. Statistics derived from BL data.

Table A4: Survey Response Balance

	Full Sample	Respondents	
	Control Mean	Control Mean	Treatment
	(1)	(2)	(3)
Female	0.697	0.756	-0.053
			(0.035)
Black	0.302	0.297	-0.027
			(0.036)
Hispanic	0.325	0.337	-0.021
			(0.037)
Asian	0.246	0.265	0.033
			(0.036)
Other Race	0.094	0.085	0.009
			(0.023)
Citizen	0.787	0.768	-0.071**
			(0.034)
Verified GPA	3.264	3.340	0.020
			(0.042)
Parent AGI	22520	22988	-815
			(1453)
Household Size	4.26	4.32	-0.012
			(0.13)
Mom Employed	0.641	0.545	-0.019
			(0.040)
Mom Employed (missing)	0.144	0.171	-0.039
			(0.028)
Dad Employed	0.435	0.353	0.066*
			(0.039)
Dad Employed (missing)	0.446	0.492	-0.038
			(0.040)
First Generation	0.811	0.821	-0.049
			(0.032)
Sibling College	0.389	0.366	0.012
			(0.039)
Sibling College (missing)	0.059	0.05	-0.002
			(0.017)
Sibling Bottom Line	0.075	0.053	0.017
			(0.019)
Sibling Bottom Line (missing)	0.074	0.065	0.009
			(0.021)
Other Program	0.444	0.412	-0.048
			(0.038)
Observations	2422	814	

Note: Table illustrates the comparability of survey respondents with the overall sample as well as the similarity of treated and control respondents. Column 1 contains control group means for the full sample. Column 2 contains control group means for survey respondents. Each cell in column 3 contains a coefficient from a separate regression of the observed characteristics on a treatment indicator variable, controlling for site by cohort (i.e., risk set) indicators. Response rates did not differ significantly between control and treatment groups. Statistics derived from BL data. Robust standard errors in parentheses. * (p<0.10) **(p<0.05), ***(p<0.01).

Table A5: Effects on Degree Attainment by Site

	Boston Area		NYC	
	Control Mean (1)	Treatment (2)	Control Mean (3)	Treatment (4)
BA Degree (5 Years)	0.418	0.074** (0.030)	0.504	0.072** (0.031)
Observations		1390		1032

Note: Table provides estimated treatment effects on bachelor's degree attainment. Unless otherwise indicated, bachelor's degree attainment is measured within 5 years of July 1 after an individual's senior year of high school. Column 1 contains control group means. Each cell in column 2 contains a coefficient from a separate regression of each outcome variable (indicated by the row title) on a treatment indicator variable, controlling for site by cohort (i.e., risk set) indicators as well as the covariates indicated in Table 1. Robust standard errors in parentheses. * ($p<0.10$) **($p<0.05$), ***($p<0.01$).

Table A6: Share of Students Discussing College and Financial Aid Topics with Different Sources of Support

	Which Apply (1)	How Apply (2)	College Cost (3)	How Apply Aid (4)	Aid Letter (5)	Which Enroll (6)
Parent	-0.0584	-0.0342	0.0233	-0.0915**	-0.0382	0.0265
	(0.0424)	(0.0427)	(0.0411)	(0.0446)	(0.0451)	(0.0407)
Control Mean	0.651	0.387	0.693	0.604	0.514	0.693
Other Family	0.00496	-0.0635	-0.0195	-0.0651	-0.0558	-0.0540
	(0.0447)	(0.0406)	(0.0442)	(0.0427)	(0.0412)	(0.0436)
Control Mean	0.429	0.311	0.410	0.363	0.316	0.425
Guidance Counselor	-0.0420	-0.0917**	0.00488	-0.0801**	-0.0237	-0.0281
	(0.0408)	(0.0368)	(0.0392)	(0.0382)	(0.0417)	(0.0428)
Control Mean	0.703	0.802	0.717	0.778	0.684	0.637
Teacher	0.0363	-0.0380	-0.0191	-0.0354	-0.0301	-0.0117
	(0.0448)	(0.0435)	(0.0440)	(0.0439)	(0.0424)	(0.0444)
Control Mean	0.458	0.434	0.509	0.425	0.363	0.443
Friends	-0.0338	-0.0435	-0.0498	-0.0609	-0.0248	-0.0368
	(0.0433)	(0.0443)	(0.0439)	(0.0436)	(0.0431)	(0.0440)
Control Mean	0.632	0.519	0.571	0.505	0.415	0.594
BL (Treatment Only)	0.905	0.916	0.912	0.914	0.884	0.865
Observations	642	642	642	642	642	642

Note: Table provides treatment effect estimates for the share of students that discuss different college and financial aid topics with different sources of support. Each cell contains a coefficient from a separate regression of each variable (indicated by the row title) on a treatment indicator variable, controlling for site by cohort (i.e., risk set) indicators as well as the covariates indicated in Table 1. Statistics derived from survey of individuals in cohort 1. See text for additional details. Robust standard errors in parentheses. (p<0.10) **(p<0.05), ***(p<0.01).

Table A7: Estimated Effects on Bachelor's Degree Attainment and Program Costs

					5-Year BA		6-Year BA		Ever BA		BA Effect per
Program Name	Program Years	Cost	Cost Year	Strategy	Estimate	Mean	Estimate	Mean	Estimate	Mean	$1,000
BL Full	HS class of 2015 and 2016	$4161	2016	RCT	0.076	0.471	0.096	0.528			0.0231
BL Access	HS class of 2015 and 2016	$2145	2016	RCT	0.076	0.471	0.096	0.528			0.0448
Class Size (STAR)	K-3rd grade 1985-1989	$12000	2007	RCT			0.009	0.124			0.0006
School Spending (MI)	4th graders 1994-1999	$5000	2012	IV					0.023[a]	0.201[a]	0.0044
School Spending	HS class of 1990-2010	$839	2013	IV			0.000	0.36			0.0000
Preschool (Boston)	1997-2001 Preschool Cohorts	$9750	2020	IV					0.035	0.297	0.0039
Pell	College entrants 2002-2014	$1000	2008	RD			0.004[a]	0.49[a]			0.0036
Wisconsin Scholars	HS class of 2009-2015	$9512	2012	RCT			0.015	0.534			0.0015
Kalamazoo Promise	HS class of 2003-2013	$17620	2007	DiD			0.074	0.30			0.0036
Dell Scholars	HS class of 2009-2010	$25000	2010	RD			0.158	0.633			0.0057
Buffett Scholars	HS class of 2012-2016	$32250	2014	RCT			0.084	0.629			0.0026
GA Merit (revisited)	Early/mid 1990s	$1677	1993	DiD					0.009	0.33	0.0032
Cal Grant-Income	HS class of 1998-2000	$8115	1999	RD					0.03	0.672	0.0026
Cal Grant-GPA	HS class of 1998-2000	$4311	1999	RD					0.046	0.462	0.0074
WV Promise	HS class of 2002-2003	$10000	2003	RD	0.045	0.37					0.0034
GA/AR Merit	Early/mid 1990s	$1677	1993	DiD					0.03	0.33	0.0108

Note: Table provides information on estimated effects and program costs for programs with estimated effects on bachelor's degree attainment using nationwide (e.g., National Student Clearinghouse) data. Costs per degree are calculated by dividing the estimated effect on bachelor's degree attainment at the longest window available by the cost of the program inflated to $2016. We include all studies published in top general interest or field journals since 2009 that evaluated U.S. interventions from the last 50 years; contained estimated treatment effects on bachelor's degree attainment; and included cost measures. Estimated effects for intensive advising (BL) are included for easy comparison. Estimates and costs obtained from Angrist, Autor, and Pallais (2021); Bettinger et al. (2019); Bartik, Hershbein, and Lachowska (2019); Dynarski (2008); Eng and Matsudsaira (2021); Page et al. (2019); Scott-Clayton (2009); Lafortune, Rothstein, and Schanzenbach (2018); Hyman (2017); Dynarski, Hyman, and Schanzenbach (2013); Gray-Lobe, Pathak, and Walters (2023). [a] Denotes estimates for which the measure of degree attainment was program completion or any degree.

Appendix B: Causal Forest Approach

We apply recently developed causal forest methods to investigate the extent to which we can attribute increases in bachelor's degree attainment to intensive advising's impact on students attending higher-quality institutions. The basic intuition of our approach is that if the effects on bachelor's degree attainment are operating (at least to some degree) through shifts in college choice, then we should see stronger treatment effects on bachelor's degree attainment among individuals who are more likely to attend higher-quality institutions as a result of intensive advising. Because we cannot recover individual treatment effects from the RCT, we use causal forest methods and our extensive set of baseline covariates to generate a personalized treatment effect (PTE) for each individual in the sample. We then explore the relationship between PTEs on the college choice and degree attainment margins to determine whether the effects on college choice likely mediate the degree attainment effects we observe.

Causal forests are a machine learning-driven enhancement to traditional methods for investigating treatment heterogeneity (Athey and Imbens 2016; Athey and Wager 2019). Rather than rely on researchers to specify subgroups to include in heterogeneity analysis, causal forest methods use machine learning to investigate the complex relationships between baseline covariates and treatment effects. The intuition behind the method is to create "trees" that iteratively split the sample in ways that maximize treatment effect heterogeneity. For example, one tree might split on gender and then GPA below or above 3.0, and so on. Individuals in the resulting "leaves" of the tree share similar combinations of baseline covariates that are predictive of the treatment effect associated with their group. To prevent overfitting, the treatment effect associated with these "leaves" is estimated using a holdout sample that was not used in the production of the tree.[1] Using a new random training (and hold out) sample each time, the algorithm produces a series of these trees that are referred to as a "forest". This "forest" can then be used to generate personalized treatment effects (PTEs) for each individual by averaging across the treatment estimate associated with each "leaf" that an individual falls into in each tree. These PTEs indicate the expected magnitude of impact the treatment has for an individual (on a particular outcome) based on their observable characteristics. Researcher use of causal forests to enhance treatment heterogeneity analysis has been limited to date, particularly for educational interventions. We are unaware of any that use the method to explore mechanisms in the fashion we pursue in this paper, although it is possible that examples exist in literatures for which we have less familiarity.

An immediate question is why this method is any better than other options. Why would causal forests be better than pursuing traditional approaches of exploring heterogeneity within researcher-specified subgroups or by interacting baseline covariates with the treatment indicator? These approaches, while useful (and employed in the paper), have limitations. They tend to limit the scope of investigation by focusing on the construction of simple subgroups defined by one or perhaps two demographic characteristics. But perhaps it is the interaction of multiple attributes that most strongly predicts the magnitude of a treatment effect. For example, information on the availability of a benefit may only be useful for those who are (1) capable of making use of a benefit, and (2) unlikely to be aware of it in the absence of the information. In the context of college choice advising, it may be the case that the individuals with scope to make high-quality college choices are those able to gain admittance to better colleges (for example, those with higher GPAs or test scores). However, it may also be the case that many students with higher GPAs and test scores

[1]Using this holdout sample rather than the sample used to fit the trees to estimate the treatment effects makes this approach "honest" (as in an "honest causal tree"). Further refinements that pre-split the data before estimating the honest causal forest produce similar results in our context. This approach maintains a second holdout sample used for neither training nor estimation of treatment effects within leaves.

are already aware of these tradeoffs (for example, those with higher incomes or with parents or siblings with college experience). As a result, there may be little heterogeneity in effects across GPA/test score or family income/first generation measures, but significant heterogeneity when we examine those with higher GPAs *and* lower family incomes.[2] Further, standard approaches to heterogeneity often have continuous measures that enter linearly (when interacted with a treatment indicator) or are used to split the sample into subsets (e.g., when a researcher explores effects among individuals with above- and below-median GPAs). But perhaps it is individuals in the middle of some distribution (e.g., GPA) that benefit from treatment, while the top and bottom do not. Or perhaps it is individuals in the middle of the GPA distribution *and* from families at a particular segment in the income distribution. The causal forest algorithm is not confined to a narrowly defined set of "expected" dimensions of heterogeneity.

One concern then is that the causal forest overfits the data by allowing for the formation of unusual subgroups defined by the interaction of covariates. The approach overcomes this concern though by splitting the sample into one sample to construct subgroups (i.e., leaves) and another to estimate treatment effects. If the algorithm overfits in the process of choosing covariate splits and interactions to define subgroups, this overfitting will be reflected in the estimation of treatment effects out of sample within subgroups defined by the same characteristics. This is more disciplined than the standard approach of using full sample data to estimate heterogeneity within an "expected" set of subgroups, where the researchers get to subjectively decide what goes in the "expected" set.

Returning to our context, we use this method to better understand the contribution of shifts in institution quality to increases in bachelor's degree attainment. As noted in the text, it could be the case that shifting students to attend higher-quality institutions results in the higher rates of degree attainment. But it is also possible that the students induced to attend higher-quality colleges and universities were inframarginal to bachelor's degree attainment and that intensive advising is affecting degree completion among a different population of students (e.g., through the ongoing Success advising provided to many students).

To investigate whether intensive advising's impacts on enrollment quality are resulting in higher rates of degree attainment, we estimate separate PTEs for two outcomes: (1) enrollment at a four-year institution with above-median graduation rates, and (2) bachelor's degree attainment within five years.

For each outcome we estimate 5,000 trees. To ensure that we are capturing real heterogeneity in treatment effects, we use "honest" estimation, beginning each of the 5,000 iterations by splitting the sample in half. The first subsample serves as a holdout sample in which we will estimate the treatment effect within each terminal leaf (i.e., set of covariate partition interactions) of the estimated tree. The second subsample serves as the training sample in which we estimate the tree structure. We use the full set of measures listed in Table 1 as potential covariates to form trees. Unless noted otherwise, we use the default parameters suggested by Athey and Wager (2019). This includes a minimum node (terminal leaf) size of 5, which requires a minimum of 5 individuals in the treatment group and 5 individuals in the control group in each node (i.e., set of covariate partition interactions.)

The algorithm chooses covariate splits so as to maximize the difference in treatment effects in the resulting leaves of the tree. Once the tree structure is established within the training sample, treatment effects are estimated with equivalently defined leaves in the holdout sample. For example, it might be determined within the training sample that the optimal tree structure to maximize heterogeneity in treatment effects results in splitting the sample into (1) individuals

[2] A researcher could use traditional analytic methods to interact GPA and family income with a treatment indicator to explore heterogeneity in this instance, but only if there was a strong a priori reason for doing so. Increasingly complex or unexpected interactions would naturally be met with suspicion and concerns about overfitting the data.

with GPAs above 3.0 and incomes above the poverty line, (2) individuals with GPAs above 3.0 and incomes below the poverty line, and (3) individuals with GPAs below 3.0.

The estimation of treatment effects within these subgroups would then be estimated within the holdout sample. Each individual in the full sample would then be assigned the treatment effect corresponding to the leaf into which they fall based on their covariates. An individual's personalized treatment effect (PTE) is calculated as the simple average of treatment effects from each terminal leaf in which they fall across the 5,000 trees. We then use these PTEs to explore relationships between treatment effects on mediating and final outcome measures as discussed in the paper.

The method we have outlined here can also be applied to address concerns about the confounding role of other potential mediators. For example, our earlier analysis suggests that shifting students to enroll at higher-quality institutions may be an important driver of advising's impact on bachelor's degree attainment. However, this analysis does not address the possibility of other factors that might also make students offered advising more likely to obtain bachelor's degrees that are correlated with their likelihood of being shifted into higher-quality institutions.[3] In our context, one possibility is that the impact of in-college "Success" advising on bachelor's degree attainment may be larger among students more likely to be shifted into higher-quality institutions as a result of the initial advising. We find little evidence to support this: There is minimal relationship between the likelihood that an individual is shifted into a higher-quality institutions and that an individual *attends* a target institution where treated students would be eligible to receive in-college advising (Appendix Figure A7).[4]

Another way to examine the potential role of ongoing in-college advising is to look at individuals with above- versus below-median PTEs on our measure of initial college quality. While individuals with above- versus below-median PTEs on our quality enrollment measure enroll in target institutions at similar rates (around 45 percent), those with above-median PTEs on enrollment quality experience large (∼19 percentage point) treatment effects on bachelor's degree attainment, while there is no effect on attainment for students with below-median PTEs on enrollment quality.[5]

Another possibility is that the ongoing in-college advising contributes to higher rates of degree attainment, but only for those who are shifted into target institutions as a result of the initial advising. In contrast to the relationship between the PTE on adjusted graduation rate and the PTE on bachelor's degree attainment, however, the relationship between the PTE on target institution enrollment and the PTE on bachelor's degree attainment is small and not significantly different from zero.[6]

[3]In mediation analyses, these factors are sometimes referred to as post-treatment confounders. The presence of post-treatment confounders would violate the exclusion restriction required to interpret the estimated effects of quality on degree attainment as causal.

[4]The coefficient from a corresponding regression is 0.171 (se 0.757), confirming the weak and statistically insignificant relationship between an individual's probability of being shifted into a higher-quality institution and their likelihood of attendance at a target institution.

[5]The lack of effect for students with below-median PTEs on enrollment quality occurs despite these students having similar access to Success advisors as students with above-median PTEs on enrollment quality (Appendix Figure A8). We see similar results when splitting the sample by the PTE on high-graduation rate or four-year enrollment (Appendix Figures A9 and A10).

[6]See Appendix Figure A13. The coefficient from a regression of the PTE on 5-year BA attainment on the PTE for enrollment at a target institution is -0.025 (se 0.014).

Appendix C: Mediation Analysis

In the text, we present results from an analysis that attempts to quantify the mediating role of initial school quality in explaining the effects of advising on eventual degree attainment. This analysis relies on several assumptions that we lay out in greater detail in this appendix. Following the potential outcomes framework of Heckman, Pinto, and Savalyev (2013), we can characterize the relationship between degree attainment (Y_z) and individual pre-treatment characteristics (X) and our key mediator of interest ($Quality$) as follows:

$$Y_z = \alpha_z + \lambda_z Quality_z + \beta_z X + \varepsilon_z, \quad z \in \{0, 1\} \qquad (C1)$$

where assignment to treatment is given by $z = 1$ and assignment to control group is given by $z = 0$. X is a vector of pre-program variables unaffected by the treatment (listed in Table 1), α_z is the intercept, and ε_z is an error term. With the exception of X, all variables and coefficients in equation (C1) are allowed to depend on treatment assignment. In particular, α_z captures the effect of experimentally induced changes in determinants of Y that are unobserved (i.e., unmeasured mediators). Experimental variation allows us to credibly estimate the effects of advising on initial enrollment quality and eventual degree attainment. It does not allow us to estimate the *causal* relationship between the enrollment quality and the degree attainment. As a result, additional assumptions (or instruments) are required to separately identify the contributions of α_z (i.e., the unmeasured mediators) and $Quality$ (the measured mediator) to the overall treatment effect on degree attainment. A common approach is to make assumptions on (1) the independence of observed and unobserved factors in the no-treatment state, and (2) the structural invariance of the contributions of mediators and baseline characteristics to outcomes (e.g., Heckman, Pinto, and Savelyev 2013; Carlana, Ferreira, and Pinotti 2022). The first assumption, which is untestable, implies that the OLS estimate of the relationship between degree attainment and initial enrollment quality represents the causal relationship. The second assumption, implies that coefficients of equation C1 do not vary with treatment assignment (i.e., $\lambda_0 = \lambda_1$ and $\beta_0 = \beta_1$). Given these assumptions, we can decompose the effect of advising:

$$E(Y_1 - Y_0) = \lambda E(Quality_1 - Quality_0) + (\alpha_1 - \alpha_0) \qquad (D2)$$

where $E(Y_1 - Y_0)$ is the average treatment effect and $E(Quality^1 - Quality^0)$ is the average change induced in initial enrollment quality, and λ is the estimated effect of initial enrollment quality on degree attainment, and $(\alpha_1 - \alpha_0)$ is the effect due to other factors.

This approach suggests that initial college quality explains 78 percent of the overall effect on bachelor's degree attainment. This approach relies heavily on the assumption that the OLS relationship between degree attainment and initial enrollment quality represents the causal relationship. In particular, there must not be unobserved factors that confound the relationship between the quality of a student's initial college choice and eventual degree attainment. This assumption is violated if students more likely to attend a high-quality college are also more likely to obtain a degree for some other reason, such as greater aptitude, academic preparation, or motivation. This is a strong assumption and untestable.

The basic idea of our approach is to use variation across types of students in the extent to which advising shifts initial college quality ($Quality$). To fix ideas, we can think about the personalized treatment effect (PTE) on adjusted graduation rate as used in the text. We use the interaction of this measure with treatment as an instrument for $Quality_z$. This approach relies on what we view as a significantly weaker assumption, the independence of ($PTE * Z = 1$) from observed and unobserved factors that might influence degree attainment. By construction, the instrument is independent of observed (X) and unobserved baseline characteristics. The primary concern then is that the instrument might be correlated with unobserved mediators that contribute to the treatment effect (i.e., $\alpha_1 - \alpha_0$). While this assumption is untestable, we view it as a significantly weaker assumption than the assumption of independence between initial college quality and eventual degree attainment.

How much of the overall treatment effect on degree attainment is accounted for via increases in the quality of initial college enrollment? We incorporate the PTEs from our causal forest approach into a more formal mediation analysis. To overcome the endogeneity of school choice, we use the interaction of PTEs and treatment to instrument for observed enrollment quality.[7] Due to the random assignment of treatment, this interaction is orthogonal to baseline characteristics of students. The approach thus overcomes concerns about the role of "pre-treatment" confounders (e.g., greater aptitude, academic preparation, motivation, or family resources) in a standard mediation approach while providing much greater explanatory power than linear covariates interacted with treatment. As noted above though, there may still be other mediators that are shifted by treatment in a similar way as initial enrollment quality (i.e., post-treatment confounders). We proceed with that caveat in mind.

Table 5 compares the results from this exercise with those from a standard mediation analysis. The instrumented effect of shifts in *initial enrollment quality* is similar, but smaller, than the relationship observed using the standard mediation analysis approach (0.94 vs 1.02). Using this simple approach, shifts in the quality of initial enrollment explain roughly 72 percent of the overall treatment effect, confirming the important role of college quality in mediating bachelor's degree attainment.

We can also incorporate shifts into *target institutions* into our mediation analysis approach to more formally disentangle the contribution of the two factors. To do so we use the interaction of treatment assignment with the PTE for attending a target institution as a second instrument. When we incorporate both instruments into the mediation approach simultaneously, we see that initial enrollment quality continues to generate a significant increase in degree attainment. The implied role of college quality in explaining degree attainment (0.89) is similar to the estimate produced when instrumenting solely for initial enrollment college quality (0.94). In contrast, attending a target institution does not generate a significant increase in degree attainment, although the point estimate is non-trivial in magnitude. If we take the point estimate (0.061) in column 4 of Table 5 at face value, it implies that attending a BL target institution increases the likelihood of degree attainment by 6.1 percentage points *for those experimentally induced into target institutions by advising.* Given the treatment effect on attendance at a target institution (10 percentage points from Table 3), this implies that advising may have increased degree attainment by 0.6 percentage points through this channel. This is roughly 8 percent of the overall treatment effect.

There is little evidence that Success advising benefited inframarginal target institution enrollees. Estimating this effect directly is hampered somewhat by the changing selection of students into target schools. While treatment is randomly assigned, it is not randomly assigned *conditional on enrolling in a target institution.* That said, the extent of selection on observables appears to be relatively modest. If we estimate the effect of assignment to treatment among students attending BL target institutions, the estimated effect is -0.002 (se 0.037). The estimate attenuates upon inclusion of baseline covariates (coefficient 0.0128, se 0.0361), but remains close to zero, suggesting that the effect of ongoing advising on degree attainment was likely minimal. Even if we make the generous assumption that students who would have attended target institutions in the absence of

[7] Interacting treatment with baseline covariates would similarly overcome some endogeneity concerns, but this simpler approach produces a partial F-statistic around 1 when using the full set of baseline covariates presented in Table 1; as we show earlier the interaction of PTEs and treatment produce a first-stage partial F-statistic of 316.

treatment would benefit as much from ongoing advising as those who were induced to attend (from Table 6), the availability of ongoing advising via BL target institutions would account for at most a third of the overall treatment effect.[8] Even under optimistic assumptions on the contribution of Success advising to degree attainment, it plays a much less meaningful role than initial enrollment quality.

[8]An upper bound for the contribution of ongoing advising to eventual degree attainment is provided by the multiplication of the share of students who would have attended a target institution anyways with the estimated effect of attending a target institution for those induced to attend combined with the increase in attainment among those induced to attend.

Appendix D: Aggregation of Information from Baseline College Preferences

In the main text, we explore the extent to which treatment effects vary with the intensity of suboptimal baseline college preferences. We quantify the intensity of suboptimal baseline college preferences with an index that aggregates information from the various proxies for "sub-optimal" preferences listed previously.[9] Summary statistics for these characteristics are contained in Appendix Table D1. These statistics indicate that many students are interested in sub-optimal institutions at baseline. In the main text, we aggregate information from these measures into an equally weighted index because it is simple and transparent. We also show that the effects are robust to weighting the index components by the inverse covariance matrix, which adjusts for potentially redundant information provided by correlated measures.[10] In this appendix we discuss the tradeoffs with other approaches and demonstrate the robustness of our results to these approaches. We discuss four potential approaches here: (1) the equally weighted standardized index, (2) an index weighted by the inverse covariance matrix, (3) Principal Component Analysis (PCA), and (4) factor analysis. We use the first two approaches in the main text.

The standardized index approach that we use in the text (and which is used in Kling, Liebman and Katz 2007 and other recent papers, e.g., Bailey et al. 2023) is simple, transparent, and makes minimal assumptions about the relationships of the underlying variables. On the other hand, there is no theoretical reason to support the equal weighting of standardized variables. Relatedly, the index does not take into account the correlations of variables. If some of the variables are highly correlated, it is *possible* that the equally weighted index is in some sense double-counting the same information and thus underweighting other information. It is also possible, though, that the variables represent different dimensions of "mistakes" that *happen* to be correlated. Adjusting for the correlation essentially assumes that the correlation implies double-counting, but we don't know with certainty that correlated variables are contributing redundant information to the index.

If one is confident that the correlations between measures imply double-counting, it makes sense to adjust for the covariance structure of the data in weighting the variables in an index. This is the second approach we use in the text. It is the approach popularized by Anderson (2008) (and used in other recent papers, e.g., Haushofter and Shapiro, 2016), who weights standardized variables by the inverse covariance matrix in producing an index. When two component measures are highly

[9] Specifically, the intensity of suboptimal baseline college preferences is captured by an index that combines information from 12 different proxies for suboptimal baseline preferences. The proxies include listing a school with a below median 6-year graduation rate (53.8%) as well as the share of institutions listed with low graduation rates; listing a school with a below median 6-year graduation rate and a high net price (over $10,000) as well as the share of institutions listed with low graduation rates and high net prices; listing a school for which the residual from a regression of graduation rate on individual GPA and predicted SAT falls in the bottom 25 percent of the residual distribution as well as the share of institutions listed meeting those conditions and the opposite of the average residual across baseline schools listed; listing a school in Barron's categories 4-6 (i.e., "Less Competitive", "Non-competitive", "Special"); listing a for-profit institution, listing a school with median earnings for individuals from low-income from low-income families ($0-$30,000) below the bottom quartile in MA or NY as well as the share listed meeting those conditions; and not listing a single "match" institution (i.e., an institution for which one's predicted SAT fell between the 25th and 75th percentile).

[10] Appendix Figures D1 and D2 illustrate that the interaction effects obtained using this index are largely insensitive to the inclusion of specific measures. These figures contain the distribution of interaction effect estimates generated when dropping any three measures. For the four-year enrollment and BA within five years outcomes, the interaction effect estimates are fairly tightly clustered around the effects reported in Table 4; this is also true for the adjusted graduation rate outcome with the equally weighted index. The interaction effect estimates for the adjusted graduation rate with the inverse-covariance weighted index are somewhat more dispersed, with a minority of estimates suggesting very small (less than 0.5 percentage points) negative interaction effects.

correlated, this approach downweights the contributions of these measures to the overall index. The approach *assumes* these correlated measures are providing redundant information. Information from less correlated measures is effectively viewed as "new information". As illustrated in the text, the interaction effect estimates using the inverse-covariance weighted index are somewhat smaller than the equally weighted index, suggesting that highly correlated measures (e.g., listing a low graduation rate school and listing a school with a low graduation rate relative to similar students) interact more strongly with treatment. If these measures are essentially capturing the same thing, it is appropriate to weight them less heavily in the overall index.

A third approach is to use Principal Component Analysis (PCA). The advantage of this approach is that it provides a hands-off method of extracting components from the data that maximize explained variation. This approach was taken by Autor et al. (2017) in constructing an index measuring socioeconomic status (SES) to investigate how gender gaps in school performance vary by SES. As in the equally-weighted index approach, the PCA approach will often weight highly correlated measures heavily as they will load onto the same component that is used to construct the index. If we believe that these measures are essentially providing the same information, this is again in a sense double counting, and will show up in the extracted component. A distinction of PCA from the above approaches is that it might suggest multiple orthogonal components that explain the data. While the equally-weighted and inverse weighted index approaches assume that all index variables enter the index, PCA might highlight that some variables contribute more to one dimension of variation and other variables to another. This could be helpful if there is variation in "mistakes" along different dimensions that could be interpreted and leveraged (e.g., quality versus price mistakes). On the other hand, because it is unclear how to interpret PCA components, if we were to focus on only the first principal component, we would effectively be underweighting information on certain "mistakes" simply because those mistakes have lower correlation with whatever is captured in the first component.[11]

A related approach to PCA is factor analysis. Factor analysis assumes that there are latent variables that drive variation in the various proxy measures (i.e., variables in the initial index) we identified. The approach is intuitively appealing because it identifies the extent to which the identified factor(s) explain overall variation as well as how each "mistake" measure loads onto each factor. This aids in interpretability of the identified factors. A key distinction between PCA and factor analysis is that factor analysis assumes that proxy measures are measured with error, whereas PCA assumes that all variance is common variance. Practically, there is little difference in the factors/components produced by a standard factor analysis and PCA when the underlying variables are correlated.

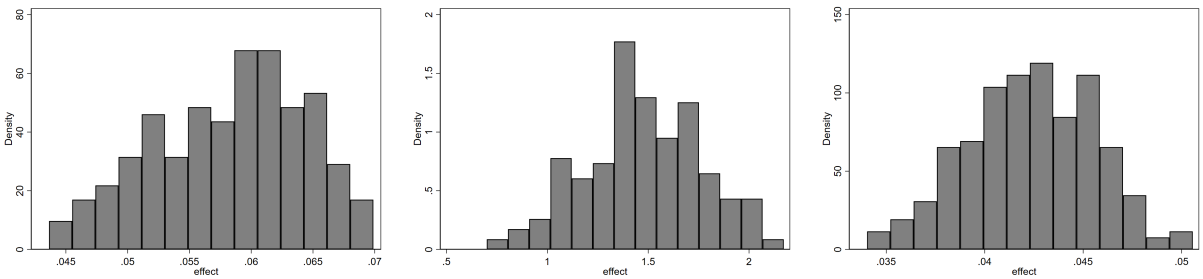
The optimal approach to information aggregation in our context depends to some extent on assumptions about the underlying "mistake" measures and the correlations between them as well as preferences for transparency and simplicity. We find the equally weighted index intuitively appealing given the transparency and simplicity of the approach. The inverse-covariance weighted index for potential double-counting implied by correlated measures. For completeness, we show in Table D2 that the estimates on the interaction terms are similar across approaches. Broadly, the evidence is supportive of the conclusion that individuals who appear less informed based on their pre-experiment college preferences (particularly in terms of expressing interest in low-quality

[11] Reinforcing this point, Kling, Liebman, and Katz (2007) motivate their focus on the equally weighted index by noting that PCA would underemphasize effects on certain components of health, saying "... we do not believe that hypertension is less important than, say, asthma simply because it has lower correlation with self-reported overall health and with physical limitations (and consequently, with the first principal component of physical health); therefore, we do not adopt the principal component approach."

institutions) seem to be helped more by access to intensive advising.[12]

[12]In Appendix Tables D3 and D5, we show the principal components (identified by PCA) and the factors identified by factor analysis. As can be seen from the tables, the first component/factor has a high Eigenvalue and explains much of the overall variance. In the case of factor analysis, the first factor explains roughly 69% of the common variation in the measures. In looking at the variable loadings in Tables D4 and D6, we can see that nearly all of the measures load positively on the first component/factor, consistent with the notion that the proxies we selected (and the resulting index we focus on) are doing a good job representing some underlying or unmeasured notion of being uninformed. The second component/factor is much more marginal in its explanation of the variation in the data and there is no clear pattern in the loadings. As a result, we use the first factor/component as our index in Table D2.

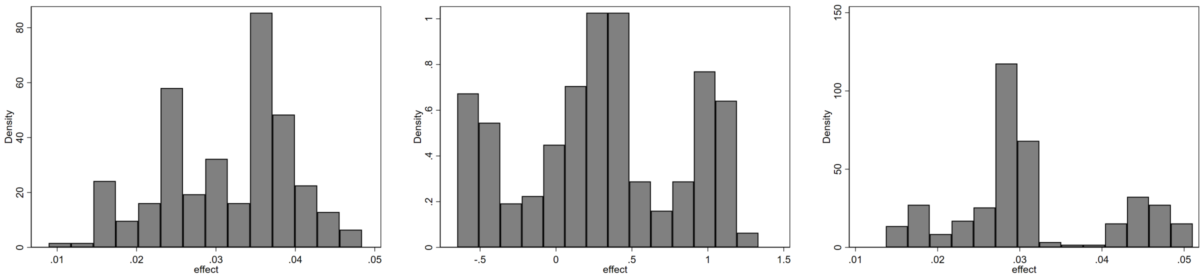
Figure D1: Robustness of Treatment*Index Estimates



(a) 4-Year College (b) Adjusted Graduation Rate (c) BA Degree (5-Years)

Note: Each figure plots the distribution of treatment* suboptimal baseline preferences index interaction coefficients for all possible permutations based on dropping any three index components. For each figure, the estimates come from 220 separate regressions of a dependent variable on treatment, treatment*suboptimal baseline preference index, site by cohort indicators, as well as the covariates indicated in Table 1. The potential index components include listing a school with a below median 6-year graduation rate (53.8%) as well as the share of institutions listed with low graduation rates; listing a school with a below median 6-year graduation rate and a high net price (over $10,000) as well as the share of institutions listed with low graduation rates and high net prices; listing a school for which the residual from a regression of graduation rate on individual GPA and predicted SAT falls in the bottom 25 percent of the residual distribution as well as the share of institutions listed meeting those conditions and the opposite of the average residual across baseline schools listed; listing a school in Barron's categories 4-6 (i.e., "Less Competitive", "Non-competitive", "Special"); listing a for-profit institution, listing a school with median earnings for individuals from low-income families ($0-$30,000) below the bottom quartile in MA or NY as well as the share listed meeting those conditions; and not listing a single "match" institution (i.e., an institution for which one's predicted SAT fell between the 25th and 75th percentile).

Figure D2: Robustness of Treatment*ICW Index Estimates



(a) 4-Year College (b) Adjusted Graduation Rate (c) BA Degree (5-Years)

Note: Each figure plots the distribution of treatment* suboptimal baseline preferences inverse covariance weighted (ICW) index interaction coefficients for all possible permutations based on dropping any three index components. For each figure, the estimates come from 220 separate regressions of a dependent variable on treatment, treatment*suboptimal baseline preference index, site by cohort indicators, as well as the covariates indicated in Table 1. The potential index components include listing a school with a below median 6-year graduation rate (53.8%) as well as the share of institutions listed with low graduation rates; listing a school with a below median 6-year graduation rate and a high net price (over $10,000) as well as the share of institutions listed with low graduation rates and high net prices; listing a school for which the residual from a regression of graduation rate on individual GPA and predicted SAT falls in the bottom 25 percent of the residual distribution as well as the share of institutions listed meeting those conditions and the opposite of the average residual across baseline schools listed; listing a school in Barron's categories 4-6 (i.e., "Less Competitive", "Non-competitive", "Special"); listing a for-profit institution, listing a school with median earnings for individuals from low-income families ($0-$30,000) below the bottom quartile in MA or NY as well as the share listed meeting those conditions; and not listing a single "match" institution (i.e., an institution for which one's predicted SAT fell between the 25th and 75th percentile).

Table D1: Measures of Sub-optimal Baseline College Preferences

Variable	**Mean**	**Std. Dev.**
List School with Low Grad Rate	0.66	0.47
List School with Low Grad Rate and High Price	0.41	0.49
List School with Low Grad Rate Relative to Similar Students	0.64	0.48
List School in Barrons Categories 4-6	0.78	0.41
List No Match Institutions	0.92	0.27
List a For-Profit Institution	0.07	0.25
List School with Low Earnings	0.08	0.27
Share Listed with Low Grad Rate	0.34	0.32
Share Listed with Low Grad Rate Relative to Similar Students	0.26	0.28
Share Listed with Low Grad Rate and High Price	0.14	0.34
Share Listed with Low Earnings	0.02	0.09
Average Negative Grad Rate Residual (flipped)	9.37	8.96

Note: The table contains summary statistics (the mean and standard deviation) for the proxy measures of sub-optimal baseline college preferences discussed in the text. The proxies include listing a school with a below median 6-year graduation rate (53.8%) as well as the share of institutions listed with low graduation rates; listing a school with a below median 6-year graduation rate and a high net price (over $10,000) as well as the share of institutions listed with low graduation rates and high net prices; listing a school for which the residual from a regression of graduation rate on individual GPA and predicted SAT falls in the bottom 25 percent of the residual distribution as well as the share of institutions listed meeting those conditions and the opposite of the average residual across baseline schools listed; listing a school in Barron's categories 4-6 (i.e., "Less Competitive", "Non-competitive", "Special"); listing a for-profit institution, listing a school with median earnings for individuals from low-income families ($0-$30,000) below the bottom quartile in MA or NY as well as the share listed meeting those conditions; and not listing a single "match" institution (i.e., an institution for which one's predicted SAT fell between the 25th and 75th percentile).

Table D2: Robustness of Treatment*Index Interaction Effect to other Aggregation Methods

	4-Year College	Adjusted Grad Rate	BA Degree (5-Years)
	(1)	(2)	(3)
Treatment*Index	0.0599***	1.547	0.0438*
	(0.0222)	(1.1230)	(0.0246)
Treatment*Inv. Cov. Weighted Index	0.0352*	0.9540	0.0298*
	(0.0205)	(0.8440)	(0.0174)
Treatment*First Component	0.0426*	1.2690	0.0404*
	(0.0218)	(1.0880)	(0.0243)
Treatment*First Factor	0.0438**	1.3140	0.0378
	(0.0219)	(1.0840)	(0.0243)

Note: The table illustrates the robustness of the estimates that interact treatment with the intensity of suboptimal baseline college preferences. In addition to the equally weighted index discussed in the main text, we also present estimates for the interaction effect of treatment with an index weighted by the inverse covariance matrix; the first principal component from Principal Component Analysis (PCA); and the first factor from factor analysis. Each of these measures are standardized to have mean zero and standard deviation one. Each cell contains estimates from a separate regression of a dependent variable (in columns) on a treatment indicator variable and treatment indicator variable interacted with a measure of the intensity of suboptimal baseline college preferences, controlling for site by cohort (i.e., risk set) indicators as well as the covariates indicated in Table 1. The intensity of suboptimal baseline college preferences is captured by one of the four approaches that combine information from 12 different proxies for suboptimal baseline preferences. The proxies include listing a school with a below median 6-year graduation rate (53.8%) as well as the share of institutions listed with low graduation rates, listing a school with a below median 6-year graduation rate and a high net price (over $10,000) as well as the share of institutions listed with low graduation rates and high net prices, listing a school for which the residual from a regression of graduation rate on individual GPA and imputed SAT falls in the bottom 25 percent as well as the share listed meeting those conditions and the opposite of the average residual across baseline schools listed, listing a school in Barron's categories 4-6 (i.e., "Less Competitive", "Non-competitive", "Special"), listing a for-profit institution, listing a school with median earnings for individuals from low-income families ($0-$30,000) below the bottom quartile in MA or NY as well as the share listed meeting those conditions, and not listing a single "match" institution (i.e., an institution for which one's predicted SAT fell between the 25th and 75th percentile). Robust standard errors in parentheses. * ($p<0.10$), ** ($p<0.05$), *** ($p<0.01$).

Table D3: Principal Component Analysis Results

Component	**Eigenvalue**	**Difference**	**Proportion**	**Cumulative**
Comp1	5.27	3.43	0.4394	0.4394
Comp2	1.84	0.73	0.1533	0.5927
Comp3	1.11	0.15	0.0922	0.6849
Comp4	0.95	0.25	0.0795	0.7644
Comp5	0.71	0.02	0.0590	0.8234
Comp6	0.68	0.15	0.0570	0.8804
Comp7	0.53	0.15	0.0441	0.9246
Comp8	0.38	0.13	0.0314	0.9560
Comp9	0.25	0.14	0.0205	0.9765
Comp10	0.11	0.01	0.0092	0.9858
Comp11	0.10	0.04	0.0087	0.9944
Comp12	0.07	–	0.0056	1.0000

Note: The table presents the results from Principal Component Analysis (PCA). The 'Eigenvalue' column shows the amount of variance explained by each principal component, with higher values indicating greater importance. 'Difference' represents the change in eigenvalue from the previous component, highlighting where the variance contribution declines. 'Proportion' denotes the fraction of total variance explained by each component, while 'Cumulative' shows the cumulative variance explained. The underlying measures include listing a school with a below median 6-year graduation rate (53.8%) as well as the share of institutions listed with low graduation rates, listing a school with a below median 6-year graduation rate and a high net price (over $10,000) as well as the share of institutions listed with low graduation rates and high net prices, listing a school for which the residual from a regression of graduation rate on individual GPA and imputed SAT falls in the bottom 25 percent as well as the share listed meeting those conditions and the opposite of the average residual across baseline schools listed, listing a school in Barron's categories 4-6 (i.e., "Less Competitive", "Non-competitive", "Special"), listing a for-profit institution, listing a school with median earnings for individuals from low-income families ($0-$30,000) below the bottom quartile in MA or NY as well as the share listed meeting those conditions, and not listing a single "match" institution (i.e., an institution for which one's predicted SAT fell between the 25th and 75th percentile).

Table D4: Principal Component Analysis - Eigenvectors

Variable	**Comp1**	**Comp2**	**Comp3**	**Comp4**	**Comp5**	**Comp6**	**Comp7**	**Comp8**	**Comp9**	**Comp10**	**Comp11**	**Comp12**	**Unexplained**
List Low Grad Rate	0.3670	-0.1546	-0.0484	0.2671	0.0100	-0.0623	-0.2960	-0.2911	-0.5033	0.4228	0.1775	0.3575	0.00
List Low Grad Rate and High Price	0.3134	0.0078	0.2077	0.0979	-0.3791	-0.2490	-0.6248	0.3374	0.3480	-0.0162	0.0059	-0.1325	0.00
List Low Grad Rate Resid	0.3554	-0.1533	-0.1453	0.2408	0.2367	0.0073	-0.1039	-0.5810	0.1861	-0.4294	-0.1251	-0.3633	0.00
List Barrons Categories 4-6	0.2610	-0.1576	-0.1472	0.3776	-0.1600	-0.5467	0.6062	0.2052	0.0854	0.0208	0.0152	0.0282	0.00
List No Match Institutions	-0.1240	0.1147	0.4751	0.7320	0.1637	0.3878	0.0766	0.1321	0.0886	0.0372	0.0071	0.0056	0.00
List a For-Profit Institution	0.1513	0.3227	0.4305	-0.1907	0.6388	-0.4757	-0.0273	0.0612	-0.0873	-0.0528	0.0220	0.0401	0.00
List Low Earnings	0.1724	0.6202	-0.2466	0.0923	-0.0910	0.0457	0.0075	-0.0422	-0.0010	0.1724	-0.6808	0.1040	0.00
% with Low Grad Rate	0.3749	-0.0928	0.1029	-0.0893	-0.0884	0.2589	0.0762	0.3952	-0.5744	-0.4835	-0.1652	-0.0667	0.00
% with Low Grad Rate Residual	0.3794	-0.1224	-0.0667	-0.1622	0.2000	0.3014	0.0851	0.1030	0.4743	-0.1125	0.0448	0.6496	0.00
% with Low Grad Rate and High Price	0.2126	0.1163	0.5917	-0.2549	-0.4732	0.0616	0.3115	-0.4418	0.0480	0.0555	0.0252	0.0140	0.00
% with Low Earnings	0.1697	0.6148	-0.2706	0.0644	-0.0948	0.1159	0.0513	0.0120	-0.0163	-0.1579	0.6763	-0.0903	0.00
Avg. Neg. Grad Rate Residual (flipped)	0.3822	-0.1034	-0.0370	-0.1853	0.2253	0.2853	0.1562	0.1911	0.1035	0.5726	0.0346	-0.5241	0.00

Note: This table presents the principal component loadings (eigenvectors) from the Principal Component Analysis (PCA). Each row represents a variable, and each column corresponds to a principal component. The values indicate the extent to which each original variable contributes to the respective principal component, with higher absolute values representing stronger influence. The 'Unexplained' column shows the portion of variance not captured by the principal components. The variables include listing a school with a below median 6-year graduation rate (53.8%) as well as the share of institutions listed with low graduation rates, listing a school with a below median 6-year graduation rate and a high net price (over $10,000) as well as the share of institutions listed with low graduation rates and high net prices, listing a school for which the residual from a regression of graduation rate on individual GPA and imputed SAT falls in the bottom 25 percent as well as the share listed meeting those conditions and the opposite of the average residual across baseline schools listed, listing a school in Barron's categories 4-6 (i.e., "Less Competitive", "Non-competitive", "Special"), listing a for-profit institution, listing a school with median earnings for individuals from low-income families ($0-$30,000) below the bottom quartile in MA or NY as well as the share listed meeting those conditions, and not listing a single "match" institution (i.e., an institution for which one's predicted SAT fell between the 25th and 75th percentile).

Table D5: Factor Analysis Results

Factor	**Eigenvalue**	**Difference**	**Proportion**	**Cumulative**
Factor1	4.997	3.406	0.6880	0.6880
Factor2	1.590	1.058	0.2189	0.9069
Factor3	0.532	0.070	0.0732	0.9801
Factor4	0.462	0.321	0.0637	1.0438
Factor5	0.141	0.101	0.0194	1.0632
Factor6	0.040	0.052	0.0055	1.0687
Factor7	-0.012	0.012	-0.0017	1.0670
Factor8	-0.024	0.068	-0.0033	1.0637
Factor9	-0.092	0.003	-0.0127	1.0510
Factor10	-0.095	0.032	-0.0130	1.0380
Factor11	-0.127	0.022	-0.0175	1.0205
Factor12	-0.149	–	-0.0205	1.0000

Note: This table presents the results from Factor Analysis, which identifies latent constructs underlying the observed variables. The 'Eigenvalue' column indicates the variance explained by each factor, with higher values suggesting greater explanatory power. 'Difference' represents the drop in eigenvalue from the preceding factor, helping to determine the number of factors to retain. 'Proportion' shows the share of total variance explained by each factor, while 'Cumulative' captures the cumulative variance explained. The underlying measures include listing a school with a below median 6-year graduation rate (53.8%) as well as the share of institutions listed with low graduation rates, listing a school with a below median 6-year graduation rate and a high net price (over $10,000) as well as the share of institutions listed with low graduation rates and high net prices, listing a school for which the residual from a regression of graduation rate on individual GPA and imputed SAT falls in the bottom 25 percent as well as the share listed meeting those conditions and the opposite of the average residual across baseline schools listed, listing a school in Barron's categories 4-6 (i.e., "Less Competitive", "Non-competitive", "Special"), listing a for-profit institution, listing a school with median earnings for individuals from low-income families ($0-$30,000) below the bottom quartile in MA or NY as well as the share listed meeting those conditions, and not listing a single "match" institution (i.e., an institution for which one's predicted SAT fell between the 25th and 75th percentile).

Table D6: Factor Analysis - Factor Loadings

Variable	**Factor1**	**Factor2**	**Factor3**	**Factor4**	**Factor5**	**Factor6**	**Uniqueness**
List Low Grad Rate	0.836	-0.188	-0.272	0.213	-0.025	0.038	0.144
ListLow Grad Rate and High Price	0.672	0.003	0.077	0.327	-0.088	-0.091	0.420
List Low Grad Rate Relative to Similar Students	0.809	-0.182	-0.345	0.001	0.182	-0.002	0.161
List Barrons Categories 4-6	0.540	-0.134	-0.206	0.105	-0.077	-0.023	0.631
List No Match Institutions	-0.245	0.074	0.004	0.224	0.140	0.119	0.850
List a For-Profit Institution	0.302	0.255	0.221	0.131	0.204	-0.033	0.735
List Low Earnings	0.391	0.831	-0.107	-0.031	-0.014	-0.005	0.145
% Low Grad Rate	0.846	-0.119	0.213	0.016	-0.154	0.103	0.190
% Low Grad Rate Relative to Similar Students	0.875	-0.159	0.107	-0.297	0.058	-0.044	0.103
% Low Grad Rate and High Price	0.434	0.080	0.374	0.233	0.041	-0.007	0.609
% Low Earnings	0.385	0.824	-0.091	-0.090	-0.048	0.022	0.154
Avg. Neg. Gr ad Rate Residual (flipped)	0.881	-0.138	0.161	-0.282	0.038	0.035	0.097

Note: This table presents the factor loadings from the factor analysis, indicating how strongly each variable is associated with each extracted factor. Higher absolute values suggest stronger relationships between a variable and a factor. The 'Uniqueness' column represents the portion of each variable's variance that is not explained by the extracted factors, with higher values indicating weaker alignment with the factor structure. The underlying measures include listing a school with a below median 6-year graduation rate (53.8%) as well as the share of institutions listed with low graduation rates, listing a school with a below median 6-year graduation rate and a high net price (over $10,000) as well as the share of institutions listed with low graduation rates and high net prices, listing a school for which the residual from a regression of graduation rate on individual GPA and imputed SAT falls in the bottom 25 percent as well as the share listed meeting those conditions and the opposite of the average residual across baseline schools listed, listing a school in Barron's categories 4-6 (i.e., "Less Competitive", "Non-competitive", "Special"), listing a for-profit institution, listing a school with median earnings for individuals from low-income families ($0-$30,000) below the bottom quartile in MA or NY as well as the share listed meeting those conditions, and not listing a single "match" institution (i.e., an institution for which one's predicted SAT fell between the 25th and 75th percentile).

Appendix E: Qualitative Response Coding

To better understand how intensive advising may have affected students we attempt to extract additional content from a free-response question that we asked in the survey: “What other comments would you like to make about your college decisions or the application process as a whole?” We began by asking an expert in natural language processing (NLP) to suggest an automated method to extract useful information from the free-response text. This expert suggested that we pursue two approaches: (1) sentiment analysis, and (2) topic identification.[13] For the sentiment analyses, we categorize each response as containing positive, neutral, or negative sentiment. For the topic analysis, the NLP process first identified 34 clustered groups containing common topical elements, which our outside expert aggregated into 6 topic areas as follows:

- Gratitude: Students sharing gratitude about their experiences, particularly to their family members or advisors (e.g., “XXX helped me a lot”)
- Difficulty: Students remarking on how difficult, long, and stressful the process was in general (e.g., “It was very stressful”)
- Regrets: Students articulating their regrets about the process, or things they wish they'd done differently (e.g., “I wish I applied to other school, but I didn't know how”)
- Advice: Students sharing advice they'd received or that they offered to future students applying to college (e.g., “It's better to start early and take it one day at a time”)
- Financial Aid: Students discussing specific topics about financial aid (e.g., “The decision was made easily when I received the financial aid award letters from both colleges”)
- Plans: Students remarking in general terms about their enrollment plans (e.g., “I chose XXX as my commitment, but I am still waiting for YYY, if I get accepted from YYY, I will change it”)

Following our conclusion that the NLP had identified meaningful topics, and that significant differences existed between the sentiment and topical content of the responses of treated and control students, we decided to pursue human coding of the responses to corroborate the findings of the NLP approach. We first reviewed the student survey responses ourselves to ensure that the topics identified by the NLP approach seemed to capture the key themes we observed in the text. Following this process, we decided to add multiple potential topics to those identified via the NLP approach: “importance of adult assistance”, “lacked support”, “planning and time management”, and “paying for college”. We then hired two individuals to review each response and code it for sentiment and topical content. We describe this process and our instructions to coders below. The results from the associated analyses are contained in Appendix Table D1, where we show control means and treatment effects for the first coder, second coder, the intersection of the two coders, the union of the two coders, and the NLP approach.

E.1 NLP Approach

Our outside expert first prepared the survey response data for analysis by differentiating substantive responses from non-substantive responses that should not be incorporated into ensuing text analysis steps, and by separating each student's response into individual sentences, so that both

[13] We describe each process in greater detail below.

the sentiment analysis and topic modeling algorithms could be conducted at the sentence-level. 205 out of 230 responding students provided substantive responses, which averaged 1.67 sentences in length (342 total sentences for analysis).

E.1.1 Sentiment Modeling

For the sentiment modeling, the outside expert then used a neural network "transformer" algorithm (similar to those used by large language trained to identify the positivity of a given sentence (Barbieri et al., 2020). The algorithm identifies positivity of a given sentence by first assessing the sentence for meaning and then inferring the likely positivity, neutrality, or negativity of the meaning of that sentence based on a training set of textual data (tens of millions of tweets) tagged for sentiment. To facilitate student-level analysis, the expert then created a count variable for each sentence-level sentiment type and a binary flag at the student level equal to one if that student's response included at least one sentence of a given sentiment type.

E.1.2 Topical Modeling

For the topic modeling, the outside expert first leveraged a natural language processing technique known as "sentence embeddings" to take a written sentence and roughly quantify its meaning (Reimers Gurevych, 2019). By the end of this process, a model can "read" the words in a sentence, examine their relationship to one another, and then characterize the **overall** meaning of the sentence by giving it a value of 0 to 1 on a broad array of abstract characteristics. In our specific model's case, there are 768 different measures of "meaning" in the array that the algorithm assigns to each sentence it reads. Sentences which are mathematically closer together on these indices are also intended to be closer together in meaning. For instance, the sentence "I walked in the park" would have an extremely similar set of measures on these indices to "I strolled in the park." Conversely, those sentences would have measures very dissimilar to, or far away from, a sentence like "Ants often live in colonies." Sentence embedding models are specifically trained to group together sentences with similar meanings according to human judgment. The outside expert then used a topic modeling algorithm to identify discrete clusters of sentences with similar meaning ("topics") in our data (Grootendorst, 2022). Because the indices themselves are not interpretable, the outside expert then needed to read the grouped sentences to interpret what the unifying meaning of a topic actually is by human judgment.

E.2 Human Coding

Following our conclusion that the NLP had identified meaningful topics, and that significant differences existed between the sentiment and topical content of the responses of treated and control students, we decided to pursue human coding of the responses to corroborate the findings of the NLP approach. We first reviewed the responses ourselves to ensure that the topics identified by the NLP approach seemed to capture the key themes we observed in the text. Following this process, we decided to add multiple topics to those identified via the NLP approach: "importance of adult assistance", "lacked support", "planning and time management", and "paying for college". We then hired two individuals to independently review each response and code it for (1) sentiment, and (2) topical content. We provided each individual with a spreadsheet containing the full set of responses as well as instructions for how to code for sentiment and topic alongside several (fake) examples. The instructions were as follows:

- Sentiment Instructions: For each response, enter a 1 in the relevant column to indicate if the response is positive, negative, or neutral. Enter a 1 in only one of the three columns. Enter a 0 in the other columns.

- Topic Instructions: For each response, enter a 1 in any column that applies to the response and a 0 in any column that does not apply. Coders were also allowed to enter "other" topics or "no specific topic".

In Appendix Table E1, we show the control means and treatment effects for the first coder, second coder, the intersection of the two coders, and the union of the two coders.

Table E1: Human Coder and NLP Assessment of Student Free-Form Responses about the College Application Process

	Coder 1 (1)	Coder 2 (2)	Both Coders (3)	Either Coder (4)	NLP (5)
Sentiment Analyses:					
Positive	0.299*** (0.0590)	0.309*** (0.0661)	0.314*** (0.0574)	0.294*** (0.0673)	0.190*** (0.0762)
Control Mean	0.100	0.230	0.0900	0.240	0.411
Negative	-0.101 (0.0642)	-0.179*** (0.0660)	-0.130** (0.0630)	-0.151** (0.0665)	0.006 (0.0704)
Control Mean	0.340	0.370	0.310	0.400	0.411
Topical Analyses:					
Gratitude	0.300*** (0.0648)	0.276*** (0.0562)	0.286*** (0.0545)	0.290*** (0.0658)	0.168** (0.0733)
Control Mean	0.190	0.0900	0.0800	0.200	0.233
Difficulty	-0.0873 (0.0700)	-0.135** (0.0681)	-0.0931 (0.0673)	-0.129* (0.0690)	-0.112 (0.0737)
Control Mean	0.400	0.480	0.360	0.520	0.433
Advice	-0.0443 (0.0557)	-0.0813* (0.0483)	-0.103** (0.0463)	-0.0235 (0.0572)	-0.0135 (0.0453)
Control Mean	0.253	0.212	0.212	0.253	0.156
Adult Assistance	0.126** (0.0531)	0.140** (0.0598)	0.117** (0.0520)	0.150** (0.0606)	NA
Control Mean	0.120	0.160	0.110	0.170	
Lack Assistance	-0.143*** (0.0428)	-0.106*** (0.0393)	-0.106*** (0.0373)	-0.142*** (0.0447)	NA
Control Mean	0.150	0.110	0.100	0.160	
Observations	230	230	230	230	205

Note: Table illustrates the robustness of the effects in Panel B of Table 6, which provides estimated treatment effects on the sentiment and topical content of student free-form responses when asked to provide additional information about their "college decisions or the application process as a whole." See the text and Appendix E for additional discussion and a detailed description of how these measures were extracted from the textual data using natural language processing (NLP) and human coders. Here we present estimated treatment effects based on coder 1, coder 2, the intersection of coders (i.e, both coders), the union of coders (i.e., either coder), and a NLP method of coding responses. Each cell contains a coefficient from a separate regression of each variable (in rows) and coding procedure (in columns) on a treatment indicator variable, controlling for site by cohort (i.e., risk set) indicators as well as the covariates indicated in Table 1. Robust standard errors in parentheses. (p<0.10) **(p<0.05), ***(p<0.01).

Appendix F: Quasi-random Advisor Assignment

One question related to mechanisms and to the potential scalability of BL is whether the large observed treatment effects on enrollment, persistence, and degree attainment vary across advisors.[14] In our preliminary work we demonstrated that there was little causal relationship between advisor demographics or behaviors and advisor effectiveness in increasing enrollment or four-year enrollment (Barr and Castleman 2019).[15] In Appendix Figures A15 and A16, we show that this is similarly true of effects on measures of college quality and bachelor's attainment. We similarly find no significant relationship between advisor gender or race and student success (Appendix Table F2).This is further evidence of the potential scalability of the BL model, since it suggests that a combination of coherent organizational leadership, successful staff recruitment and training, and effective curriculum are driving the results we observe, rather than a handful of particularly strong advisors who may be hard to identify and recruit in other contexts.[16]

We explore the notion of random assignment of students to advisors more formally by conducting a set of randomization tests. In Table F1, we explore the relationship between a number of advisor characteristics and baseline student characteristics. Formally, we estimate the following specification:

$$C_i = \alpha + \beta X_i + \sum_j \gamma_j l_{ij} + \varepsilon_i \tag{1}$$

where C_i are observable demographic characteristics of the advisors and measures of the extent to which a advisor meets with his or her assigned students, and X_i includes baseline demographic student characteristics. The l_{ij} are site by cohort fixed effects which control for site by cohort variation in the pool of students randomized across advisors.

The advisor interaction measures (in columns 5 through 8), indicate the average number of meetings of each type that a advisor holds over the course of the program. For example, the dependent variable in column 5 is the average number of meetings about applications that a advisor has had with each of his or her students. We follow a leave-one-out procedure to eliminate the possibility that a particular student could influence his or her advisor's score via their own behavior; thus, our variable of interest takes the form $X_{-i,s}$. The estimates in Table F1 suggest little relationship between advisor observables characteristics (or behavior) and baseline individual student characteristics, supporting the argument that advisors are as good as randomly assigned. F tests for the joint significant of all the pre-determined variables are generally insignificant, illustrating that particular types of students do not appear to be assigned to particular types of advisors.[17] Similarly, columns 6-9 indicate that particular types of students do not appear to be assigned to advisors who exhibit different advising tendencies. This suggests that students are as good as randomly assigned to advisors. There is some suggestive evidence that assigning a student to an advisor who tends to have more application meetings with their assignees may increase the likelihood that that student goes to college, suggesting that BL's increased focus on this aspect of advising may

[14] If only advisors with certain characteristics/behaviors are effective and advisors with these characteristics are in short supply, it would be more difficult to scale the program.

[15] That paper, published in AEA Papers and Proceedings, discussed in greater depth the quasi-random assignment of advisors to students. This notion is explored more formally later in the appendix.

[16] We note that it is, of course, possible that the quality of advisors falls with expansion, but argue that the consistency of effects across current advisors suggests the importance of the BL model to ensuring consistent delivery of services.

[17] The lone exception is for white advisors, a result that appears to be driven by white advisors adjusting verified GPAs rather than non-random assignment. If we exclude verified GPA from the regression, the remaining variables are not predictive of having a white advisor.

be important. It is clear that advisors who have more application meetings have somewhat larger effects on enrollment, but the results merely suggest that the extent of interaction is causing the higher enrollment rates. It may be that advisors who have more application meetings have some other characteristics that leads their advisees to be more likely to attend four-year colleges.

Table F1: Tests of Random Advisor Assignment

Couns. Chars.	(1) Female	(2) Black	(3) White	(4) Hispanic	(5) # of App.	(6) # of Fin. Aid	(7) # of Office	(8) # of Contacts
Baseline Covariates:								
Parent AGI	-0.000	0.000	-0.000	-0.000	-0.000	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Household Size	0.006	-0.002	0.012**	0.001	0.005	-0.001	0.006	-0.002
	(0.007)	(0.006)	(0.006)	(0.007)	(0.005)	(0.005)	(0.010)	(0.015)
Verified GPA	-0.008	0.014	-0.022	-0.014	-0.021	-0.002	-0.057*	-0.120***
	(0.020)	(0.018)	(0.018)	(0.020)	(0.014)	(0.013)	(0.030)	(0.045)
Female	0.013	-0.006	0.032	-0.030	0.030*	0.018	0.072**	0.087
	(0.024)	(0.021)	(0.021)	(0.023)	(0.016)	(0.016)	(0.036)	(0.053)
White or Asian	-0.006	-0.037	0.023	0.034	0.030	0.025	0.057	0.108
	(0.042)	(0.038)	(0.038)	(0.042)	(0.029)	(0.028)	(0.063)	(0.094)
Black	-0.062	0.027	0.015	-0.016	-0.038	0.013	-0.021	-0.014
	(0.041)	(0.036)	(0.036)	(0.040)	(0.028)	(0.026)	(0.061)	(0.090)
Hispanic	-0.053	0.002	0.007	-0.008	-0.018	-0.006	-0.051	-0.036
	(0.041)	(0.037)	(0.037)	(0.041)	(0.028)	(0.027)	(0.062)	(0.091)
Observations	1,596	1,596	1,596	1,596	1,596	1,596	1,596	1,596
R-squared	0.007	0.008	0.013	0.005	0.010	0.004	0.010	0.009
Prob>F	0.362	0.262	0.0208	0.702	0.0912	0.728	0.106	0.155
Mean	0.727	0.228	0.281	0.282	3.591	2.121	9.151	13.58

Note: Table illustrates the relationship between advisor and student characteristics. Each column contains a regression of a different advisor characteristic on the full set of covariates, controlling for site by cohort indicators. The average # of meetings variables are constructed using a leave one out procedure, so that each individual is assigned the average number of meetings occurring between every other student with the same advisor. Robust standard errors in parentheses. * (p<0.10) **(p<0.05), ***(p<0.01).

Table F2: Relationship Between Advisor Characteristics and Outcomes

	Enrolled		Enrolled 4-Year		Enrolled High Grad		Adjusted Grad Rate		BA 5 Years	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Counselor Characteristics										
Female	0.011	0.011	-0.006	-0.012	0.004	0.000	-0.002	-0.004	0.010	-0.001
	(0.027)	(0.027)	(0.035)	(0.034)	(0.042)	(0.041)	(0.021)	(0.020)	(0.043)	(0.042)
Black	-0.041	-0.042	-0.045	-0.048	-0.023	-0.013	-0.019	-0.016	-0.010	-0.007
	(0.029)	(0.029)	(0.037)	(0.037)	(0.045)	(0.044)	(0.022)	(0.022)	(0.046)	(0.045)
White	-0.033	-0.032	-0.034	-0.032	-0.018	-0.010	-0.023	-0.018	-0.012	-0.003
	(0.029)	(0.029)	(0.037)	(0.037)	(0.045)	(0.044)	(0.022)	(0.022)	(0.046)	(0.045)
Hispanic	-0.020	-0.019	-0.024	-0.023	-0.008	-0.010	-0.013	-0.013	0.010	0.018
	(0.028)	(0.028)	(0.036)	(0.035)	(0.043)	(0.042)	(0.021)	(0.021)	(0.044)	(0.043)
Application Meetings	0.064**	0.066**	0.055	0.054	-0.009	-0.027	0.038	0.033	0.005	-0.008
	(0.032)	(0.032)	(0.040)	(0.040)	(0.049)	(0.048)	(0.024)	(0.024)	(0.050)	(0.049)
Financial Aid Meetings	-0.033	-0.037	-0.043	-0.050	-0.037	-0.034	-0.012	-0.011	-0.016	-0.028
	(0.044)	(0.044)	(0.055)	(0.055)	(0.067)	(0.066)	(0.033)	(0.033)	(0.069)	(0.068)
Covariates		X		X		X		X		X

Note: Table illustrates the relationship between advisor characteristics and student outcomes. Each column contains estimates from a separate regression of a dependent variable (in columns) on a set of advisor characteristics. Application meetings and financial aid meetings variables provide a measure of the average number of meetings of each type per student for each advisor. The variable is constructed using a leave one out procedure, so that each individual is assigned the average number of meetings occurring between every other student with the same advisor. Robust standard errors in parentheses. * ($p<0.10$) **($p<0.05$), ***($p<0.01$).

References

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. 2019. "Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes." *American Economic Journal: Applied Economics*, 11(3): 338â81.

Bailey, Martha J, Hilary Hoynes, Maya Rossin-Slater, and Reed Walker. 2023. "Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence From the Food Stamps Program." *The Review of Economic Studies*, 91(3): 1291–1330.

Barbieri, Francesco, José Camacho-Collados, Leonardo Neves, and Luis Espinosa Anke. 2020. "TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification." *CoRR*, abs/2010.12421.

Grootendorst, Maarten. 2022. "BERTopic: Neural topic modeling with a class-based TF-IDF procedure."

Haushofer, Johannes, and Jeremy Shapiro. 2016. "The Short-term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya*." *The Quarterly Journal of Economics*, 131(4): 1973–2042.

Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review*, 103(6): 2052â86.

Reimers, Nils, and Iryna Gurevych. 2019. "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks." *CoRR*, abs/1908.10084.