A Causal Analysis of School Counselor Caseload Size on College Enrollment

Dhathri Chunduru

Johns Hopkins University

March 7. 2019

Abstract

An inability to infer causality about the impact of resources due to the endogenous relationship between students and schools has limited school effects literature. Using data from the 2009 High School Longitudinal Study, which follows 9th graders until 2016, as well as the American Community Survey and the Common Core of Data, this study offers a causal estimate of the effect of school counselor caseloads on college enrollment using propensity score weights. A more robust analysis which adjusts for confounders finds equivalent results. Finally, a stratification analysis in which students are divided into likelihood of college attendance groups finds that students who are most likely to benefit from caseload size are those who are neither the most nor least likely to attend college, but those on the margins of college enrollment. In sum, this study finds strong evidence of the causal effect of caseload size on college enrollment.

Objectives

(1) To offer a causal method that can measure variation in the type and quantity of resource allocation (Sorenson & Hallinan, 1977; Sorenson & Morgan, 2000)

(2) To determine whether caseload size affects all students

(3) To explore heterogeneity of the treatment effect by dividing sample into strata of propensity for treatment (Bryan, Moore-Thomas, Day-Vines, & Holcomb-McCoy, 2010; Erickson, 1975; Farmer-Hinton & Adams, 2006; Stanton-Salazar, 1995)

Research Questions

- RQ1: What is the effect of counselors' caseload size on students' college enrollment
- RQ2: Does this treatment effect of counselor caseload vary between students, based on their likelihood of attending college?

Study Background

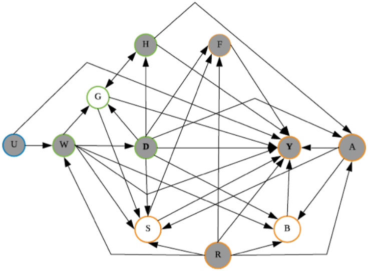
Perna's (2006) conceptual framework for college-going.

- Layer 1: *Social capital* and demographic characteristics Bourdieu (1977)
- Layer 2: School and community context (includes school resource allocation; and racial, ethnic, and socioeconomic composition)
- Layer 3: Context of higher education in student's sphere (including student's location as well as college recruitment efforts)

School counselors critical for college-going per qualitative literature (Farmer-Hinton and Adams 2006)

- School counselors influence college-going differently for students based on their backgrounds (Choy 2001; Cooperative 2007; Mahoney Jr. & Merritt 1993; Stanton-Salazar 1995).
- Counselors convey information and assist students in completing key actions to enter college (Bryan et al. 2010; Farmer-Hinton & Adams 2006; Stephen & Rosenbaum 2013).
- Counselors serve as gatekeepers to college (Erickson 1975b; Rosenbaum, Miller and Krei 1996; Stanton-Salazar & Dornbusch 1995).

Analytic Strategy



(a) DAG

Table 8: Legend for Figure 1's Causal Diagram

Symbol	Description	
○	Unobserved	
●	Observed	
—— (green)	School-level variable	
—— (orange)	Individual-level student variable	
—— (blue)	Community & district-level variable	
Node	**Variable(s)**	**Source**
D	Counselor caseload size (Caseload size)	HSLS
Y	Student goes to college (Outcome variable)	HSLS
F	Intent to apply for financial aid	n/a
H	Counselor percentage of time spent on college; Goal prescience	n/a
A	GPA; Math achievement in 11th grade	n/a
R	SES; Parent is college educated; Member of underrepresented race	HSLS
U	District urbanity; Percent of district that is poor; Percent of district that has a bachelor's degree	ACS
W	Free-and-reduced lunch percentage of school; White & Asian percentage	CCD
G	School's overall college-going rate	n/a
S	Knowledge and value of college	n/a
B	Expected benefits and costs	n/a

(b) Key

Figure 1: A Directed Acyclic Graph (DAG) of the Causal Relationship between Caseload Size and College-Going

Analytic Strategy

Critical covariates required for blocking back-door paths include:

- School context (W), as measuerd by school SES and racial composition
- District context (U), as measured by neighborhood poverty, urbanicity, and education
- Student background (R), as measured by SES, parental education, and race Can only estimate ATT because of...
- The recursive relationship between a school's overall college-going rate (G - unobserved) and counseling (H), which is represented by a dual arrow (Morgan & Winship 2015)
- Descendants of causal variable are not regressed upon

Selection Bias

Two sources of selection bias

- Selection on unobserved variables
  - Consider the role of unobserved confounders that influence both the treatment and outcome, e.g. residential segregation and limited school choice
  - Non-voluntary selection from residential segregation, choice, and mobility
- Selection on the effects of the treatment, itself
  - Students who are most likely to benefit from small caseloads may also be the ones most likely to attend college $\rightarrow$ positive bias

Data

- HSLS: Nationally representative; complex sampling design; 3 waves between 2009 - 2013, beginning in 9th grade
- ACS: 5-year estimates from 2009-2013 by school district
- CCD: 2009-2013 school-level data regarding funding and school demographics
- I link this data using district NCES and school IDs to create a large-scale longitudinal dataset with cost-adjusted variables.

Weighted Regression

- ATT demonstrated by:

$$E[\delta|D=1] = E[Y^1 - Y^0|D=1]$$

- Propensity score weights constructed as follows:

$$For\ d_i = 1 : w_i, ATT = 1;\ For\ d_i = 0 : w_i, ATT = p/(1-p_i)$$

- To construct these weights, I use a "kitchen-sink" method to create weights that balance the key covariates between the control and treatment group
- To study effect heterogeneity, I conduct weighted regressions over propensity score strata, in addition to overall analysis
- Because it is assumed that students in the control group did not select into default schools, voluntarily, the ATC is not identified

Measures

- Outcome variable
  - Dichotomous measure of college-going
- Causal variable
  - Dichotomous indicator of counselor caseload, < 250 students per counselor (ASCA recommendations)
- Student-level covariates
  - Dichotomous indicator of race and parental education, and scaled family SES
- School-level covariates
  - Free-and reduced price lunch percentage in schools
  - Racial composition

Descriptive Statistics

Table 1: Descriptions, means, and standard deviations of outcome variable and pre-treatment covariates by treatment group, weighted by analytic weights

	D_i=1		D_i=0	
Variable	Mean	SD	Mean	SD
Outcome: College-going (12^{th})	0.65	0.48	0.55	0.50
Student: Parent attended college	0.30	0.46	0.34	0.47
Student: Underrepresented race	0.35	0.48	0.42	0.49
Student: Socioeconomic status	-0.12	0.69	-0.12	0.74
School: Percent FRL	0.47	0.23	0.43	0.21
School: Percent of White / Asian in school	0.64	0.33	0.59	0.29
District: Urban area*	0.23	0.42	0.29	0.45
District: Percent poor*	0.15	0.08	0.16	0.07
District: Percent with bachelor's degree*	0.26	0.12	0.26	0.07

N = 10,950 (9,698 in treatment and 1,325 in control), HSLS provided weights are used

Unless otherwise noted, all student- and school-level variables are drawn from 9^{th} grade student data, collected in 2009

*Variables are drawn from the American Community Survey's 2011-2015 five-year estimates. We extrapolate that the values are indicative of area statistics for the relevant years of students' 12^{th} grade year, which is 2013. All other variables are derived from the HSLS 2009-2013 survey.

Balance Demonstration

Table 9: Balance of Means and SDs for Matching Covariates with

	ATT Weights				ATC Weights			
	$D_i = 1$		$D_i = 0$		$D_i = 1$		$D_i = 0$	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Parent attended college	0.30	0.46	0.29	0.46	0.29	0.46	0.34	0.47
Underrep race	0.35	0.48	0.35	0.48	0.50	0.50	0.42	0.49
SES	-0.12	0.69	-0.12	0.73	-0.22	0.72	-0.12	0.74
Pct FRL	0.47	0.23	0.47	0.22	0.45	0.23	0.43	0.21
Pct White & Asian	0.64	0.33	0.63	0.30	0.51	0.36	0.59	0.29
Urban area*	0.23	0.42	0.24	0.43	0.27	0.44	0.29	0.45
Pct poor*	0.15	0.08	0.15	0.07	0.18	0.09	0.16	0.07
Pct with bach degree*	0.26	0.12	0.26	0.12	0.24	0.13	0.26	0.11
Mean Abs Diff	0.02 (0.06)				0.15 (0.10)			

n = 15,751 for $D_i = 1$ and $D_i = 0$

Estimate of the ATT

Table 3: Weighted Logistic Regression Estimates of the ATT

Method	Coefficient (SE) ATT weight
Model 1: Weighted regression*	
Estimate	0.22 (0.11)
Treatment = 0	0.60 (0.01)
Treatment = 1	0.65 (0.02)
Model 2: Weighted regression & doubly robust*	
Estimate	0.26 (0.12)
Treatment = 0	0.60 (0.01)
Treatment = 1	0.65 (0.02)
Model 3: Common support°^	
Estimate	0.22 (0.11)
Treatment = 0	0.60 (0.01)
Treatment = 1	0.65 (0.02)
Model 4: Common support & doubly robust^	
Estimate	0.26 (0.12)
Treatment = 0	0.60 (0.01)
Treatment = 1	0.65 (0.02)

*n=7,203 for Di = 1 and n=10,545 for Di=0

^n = 7,188 for Di=1 and n = 10.532 for Di=0

° 38 cases fell below the lower bound of (~0.008) and 0 cases fell above the upper bound of (~0.722) the predicted probability of treatment ($\hat{p}$) and are, therefore, dropped

Note: Model 1 regresses college-going on caseload size. Model 2 repeats this regression but also includes the covariates from the conditioning set used earlier. The estimates do not vary significantly. While the log odds increase from 0.22 to 0.26, the predicted probabilities remain the same. Thus, a doubly robust estimate does little to change the effect of treatment. In models 3 and 4, the sample is limited to common support. In sum, 38 cases are dropped as they fall outside of the lower and upper bounds of the propensity score, $\hat{p}$. This confirms the overlap between the original treatment and control groups discussed earlier. Furthermore, model 4, which uses a doubly robust estimator just as model 2 did yields similar estimates to model 2. The ATT weight in this regression is constructed from a datamined weight from Equation 1a and 1b.

Distribution of Covariates in Treatment and Control Groups over Propensity Strata

Table 5: Mean Distribution of Outcome Variable and Covariates over Strata

	Stratum 1 [0 – 0.11)		Stratum 2 [0.11 – 0.13)		Stratum 3 [0.13 – 0.21)		Stratum 4 [0.21 – 0.61)	
	$D_i = 1$	$D_i = 0$	$D_i = 1$	$D_i = 0$	$D_i = 1$	$D_i = 0$	$D_i = 1$	$D_i = 0$
College-going	0.33	0.44	0.51	0.45	0.50	0.46	0.48	0.43
Parent attended college	0.55	0.58	0.46	0.52	0.45	0.44	0.29	0.28
Underrep race	0.76	0.49	0.56	0.41	0.31	0.27	0.26	0.20
SES	-0.48	-0.10	-0.20	-0.12	-0.03	-0.06	-0.12	-0.07
Pct FRL	0.47	0.36	0.41	0.38	0.39	0.40	0.51	0.53
Pct White & Asian	0.36	0.53	0.58	0.65	0.72	0.72	0.72	0.73
Urban area*	0.25	0.33	0.19	0.21	0.12	0.18	0.09	0.17
Pct poor*	0.22	0.18	0.16	0.15	0.14	0.14	0.13	0.15
Pct with bach degree*	0.24	0.25	0.27	0.26	0.26	0.26	0.24	0.25

Estimate of the ATT over Strata

Table 6: Weighted Logistic Estimates of the ATT over Strata

Method	Stratum 1	Stratum 2	Stratum 3	Stratum 4
Model 1: Logistic regression*				
Estimate	0.20 (0.39)	0.19 (0.34)	0.07 (0.17)	0.29 (0.15)
Treatment = 0	0.54 (0.05)	0.58 (0.03)	0.65 (0.02)	0.56 (0.02)
Treatment = 1	0.59 (0.08)	0.72 (0.07)	0.67 (0.03)	0.63 (0.03)
Model 2: Doubly robust logistic regression*				
Estimate	0.32 (0.38)	0.39 (0.33)	-0.01 (0.15)	0.10 (0.16)
Treatment = 0	0.53 (0.03)	0.66 (0.03)	0.67 (0.02)	0.59 (0.02)
Treatment = 1	0.59 (0.07)	0.73 (0.05)	0.66 (0.03)	0.61 (0.03)
Model 3: Common support°^				
Estimate	0.21 (0.41)	0.19 (0.35)	0.07 (0.17)	0.29 (0.15)
Treatment = 0	0.54 (0.05)	0.68 (0.03)	0.65 (0.02)	0.56 (0.02)
Treatment = 1	0.59 (0.09)	0.72 (0.07)	0.67 (0.03)	0.63 (0.03)
Model 4: Common support & doubly robust°^				
Estimate	0.25 (0.38)	0.39 (0.33)	-0.01 (0.15)	0.10 (0.16)
Treatment = 0	0.54 (0.03)	0.66 (0.03)	0.67 (0.02)	0.59 (0.02)
Treatment = 1	0.59 (0.08)	0.73 (0.05)	0.67 (0.03)	0.61 (0.03)

*Sample counts per strata: $n_{stratum1}$ = 2,614; $n_{stratum2}$ = 1,115; $n_{stratum3}$ = 4,415; $n_{stratum4}$ = 2,879

^Sample counts per strata: $n_{stratum1}$ = 1,645; $n_{stratum2}$ = 700; $n_{stratum3}$ = 2,932 ; $n_{stratum4}$ = 1,911

Results

- ATT for total sample: Table 3 indicates that the treatment effects are all within 5%,
  - e.g. students in the treated group are 5% more likely to attend college than students in the control group.
  - Doubly robust and estimate on common support shows similar findings.
- Treatment effect heterogeneity: Results in table 6 show that students in lowest strata have much higher treatment effects, an effect which becomes stronger with conditioning.
  - Control and treatment groups divided into four strata, defined by a difference in means that falls below a p-value of less than 0.001.
  - As strata increases, so does the probability of treatment.

Discussion

The gneral and stratified weighted regression method accounted for the two primary sources of selection: confounder bias and treatment effect heterogeneity.

While caseload size may be influential on students' college-going rates as well as more influential for students who may be less likely to attend college, an exploration of other factors in schools suggests that caseload size is far from the only resource that varies between schools.

The findings of this study confirm policies that aim to concentrate human capital efforts towards students that most need them.

By using weighted regression over strata, this study offers a method to conduct causal analysis of school-level resources.

Limitations

- Caseload size doesn't determine nature of counselor relationships or quality of counseling
- Caseload size may simply be serving as a proxy for the many ways in which human capital is distributed between schools
- Despite adjusting for a variety of factors, many unobservable characterstics remain that likely impact college-going
- This study implores research to consider the influence of a variety of resource allocation mechanisms on college-going