Appendices

A Additional tables & figures

A.1 Timeline and description of college application process in Chile

The academic year in Chile begins in March and ends in December. First-year applicants apply for financial aid in November of the year prior to starting school. If a student plans on starting school in March 2014, she should apply in the Fall of 2013. Below is a timeline of events, beginning with applications for the CAE. If not otherwise noted, all dates listed were those published for the 2013 school year.[A.1]

Table A.1: Typical timeline of college and loan applications, admissions and decisions

June 12 - July 13, 2012	Students register to take the PSU.
October 3–17, 2012	Continuing students apply for CAE, FSCU and for federal grants by completing the FUAS application online (Extended to October 24).
October 29 November 21, 2012	First–year students apply for CAE, FSCU and for federal grants by completing the FUAS application online (Extended to November 29).
December 3–4, 2012	PSU exams administered.
January 2, 2013 (11:00 pm)	PSU exam results published.
January 3–5, 2013	Application period.
January 11, 2013	List of students ”preselected” for CAE are published.
January 13, 2013	Application results published on DEMREs website, along with information on how to register.
January 14–16, 2013	Registration period for CRUCH schools and some private universities. Each university sets its own days and times for registration within this window.
January 17–23, 2013	Registration period for waitlisted students.
January 14–23, 2013	Period to withdraw registration.
January 14–23, 2013	Students who have applied for federal loans or grants must present proof of income at their schools.
February 18–20, 2013	Results of grant and FSCU applications are published.

[A.1] Dates from DEMRE’s 2013 admissions calendar and from Ingresa’s 2013 application timeline.

A.2 Timeline and description of college application process in Chile

Table A.2 shows how family background, academic performance, and school characteristics vary with poverty status of high school. The first panel presents high school characteristics for all high schools from 2008, since this is the most recent year for which we have high school poverty ratings from the Ministry of Education (Mineduc). The second panel presents individual characteristics for the sample of students who were invited to participate in the survey. Each high school is identified by a High School Identification Number (RBD).

Table A.2: Demographics by High School Poverty Ratings

School-level Characteristics	A	B	C	D	E	All
Municipal HS	67.3%	53.5%	15.9%	3.7%	0.0%	31.0%
Voucher HS	32.7%	46.5%	83.5%	85.1%	6.6%	52.8%
Private HS	0.0%	0.0%	0.6%	11.2%	93.4%	16.2%
Fraction of families in lowest income quintile	76.9%	64.8%	52.1%	38.8%	22.8%	53.3%
Fraction of families in top two income quintiles	6.0%	11.6%	22.4%	39.9%	66.0%	26.1%
Total Observations	453	637	524	430	363	2,407

Individual Characteristics	A	B	C	D	E	All
Took entrance exam	85.2%	92.3%	97.9%	99.4%	99.8%	94.2%
Matriculated to higher education 2013	61.6%	66.4%	73.5%	75.5%	76.5%	70.0%
Entrance exam score (conditional on taking exam)	438	469	523	577	625	508
	(78)	(79)	(81)	(79)	(75)	(95)
One parent with at least some post-secondary	9.6%	16.8%	36.9%	66.1%	92.3%	35.8%
Both parents with at least some post-secondary	2.4%	4.3%	13.7%	36.8%	72.8%	16.7%
One parent with a college degree	3.3%	4.6%	12.8%	32.8%	69.7%	15.6%
Both parents with a college degree	0.6%	0.7%	2.6%	9.6%	32.4%	4.5%
Total Observations	16,784	50,411	52,060	27,368	7,083	164,798

Standard deviations are in parentheses. High Schools are identified by an identification number (RBD). RBD poverty ratings are from Mineduc, A being the highest poverty rating and E being the lowest. Fraction of families in income quintiles are constructed from our observed tax data in the tax authority (SII), based on income cutoffs defined by Formulario Unico de Acreditacin Socioeconmica (FUAS). Higher education matriculation records are from Mineduc. Entrance exam is the Prueba de Selecion Universitaria (PSU) and the value presented is the average of the Math and Language components. Parent's education is reported by students in standardized tests administered, known as Sistema Nacional de Medicin de la Calidad de la Educacin (SIMCE).

Table A.3: Potential Earnings Gains & Tuition Changes

	Pooled	Low-SES	High-SES
Earnings Gains from Switching to Suggested Institution			
Mean	267,566	289,420	255,128
P90	607,000	679,000	545,000
P75	385,000	392,000	385,000
P50	203,000	231,000	190,000
P25	71,000	84,000	60,000
P10	0	4,000	0
% Earnings Gains from Switching to Suggested Institution			
Mean	85.87%	126.45%	57.55%
P90	169.13%	259.71%	120.07%
P75	75.08%	112.18%	54.95%
P50	33.09%	44.08%	29.37%
P25	11.20%	16.60%	8.66%
P10	0.00%	0.00%	0.00%
Change in Tuition from Switching to Suggested Institution			
Mean	-2,068	376	-3,493
P90	26,430	26,430	26,691
P75	7,719	8,635	7,719
P50	0	0	0
P25	-10,099	-7,841	-14,674
P10	-31,369	-21,797	-36,588
% Change in Tuition from Switching to Suggested Institution			
Mean	0.20%	2.00%	-0.78%
P90	33.98%	43.88%	28.75%
P75	12.45%	17.99%	10.50%
P50	0.00%	0.00%	0.00%
P25	-17.14%	-17.17%	-17.14%
P10	-33.02%	-36.36%	-32.15%

Distribution of earnings and cost changes from search prompts provided to treatment group. Units are CLP. See Section 4 for details.

A.3 Validating the random assignment of treatment

Table A.4 shows balance of baseline characteristics across treatment and control groups.

Table A.4: Balance of Baseline Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Pooled			Low SES			High SES	
	Control	Treated	P-Value Diff	Control	Treated	P-Value Diff	Control	Treated	P-Value Diff
Graduated from municipal HS	0.338	0.326	0.819	0.594	0.621	0.708	0.197	0.178	0.950
	(0.473)	(0.469)	[49,025]	(0.491)	(0.485)	[16,594]	(0.398)	(0.382)	[29,848]
Female	0.584	0.565	0.137	0.573	0.578	0.866	0.592	0.561	0.088
	(0.493)	(0.496)	[49,166]	(0.495)	(0.494)	[16,594]	(0.492)	(0.496)	[29,850]
Net Value first-choice degree	733,096	736,867	0.832	602,295	596,068	0.270	822,640	828,682	0.407
	(494,209)	(499,756)	[40,806]	(454,020)	(444,197)	[13,298]	(498,946)	(508,165)	[25,394]
Listed multiple fields	0.566	0.564	0.988	0.637	0.642	0.376	0.520	0.520	0.851
	(0.496)	(0.496)	[44,964]	(0.481)	(0.480)	[14,925]	(0.500)	(0.500)	[27,669]
Absolutely certain about enrollment plans	0.340	0.337	0.409	0.348	0.340	0.086	0.328	0.326	0.765
	(0.474)	(0.473)	[49,166]	(0.477)	(0.474)	[16,594]	(0.470)	(0.469)	[29,850]
No first-choice degree earnings/cost info.	0.176	0.170	0.156	0.213	0.197	0.002	0.152	0.151	0.755
	(0.381)	(0.376)	[49,155]	(0.409)	(0.398)	[16,589]	(0.359)	(0.358)	[29,846]
P-Value of joint significance			0.1914			0.4052			0.5842
			[37,356]			[11,984]			[23,631]
P-Value of joint significance conditional on matriculation			0.1936			0.2576			0.4526
			[23,297]			[7,299]			[14,973]

Balance on predetermined variables for samples listed in columns. 'Pooled' is full sample. 'Low-SES' and 'High SES' samples are as defined by high school income ratings. See text for details. Means with standard deviations in parentheses are presented in columns 1, 2, 4, 5, 7, 8. Columns 3, 6, and 9 with the exception of the final row, present the p-value of treatment when the demographic variable in each row was regressed on treatment using an OLS regression that controls for randomization blocks and has robust standard errors clustered at the high school of graduation. The randomization blocks for 2012 high school graduates were based on RBD characteristics, and randomization was done at the high school level. The randomization blocks for older high-school graduates were based on previous entrance exam scores, with randomization at the individual level. Sample sizes are in brackets in Columns 3, 6, 9. "Net Value first-choice degree" is the Net Value at the degree a student listed as most-preferred. See Section 3 for a discussion of the Net Value variable. "Listed multiple fields" is an indicator defined if the student listed more than one field of study in their degree choices, out of a maximum of three, listed in Q2 of the survey. "Absolutely certain about enrollment plans" is an indicator defined if the student responded in Q3 that they were absolutely certain that they would be applying to their listed degree choices in Q2. "No first-choice degree earnings/cost info." is an indicator defined if the student responded "I don't know" to all survey questions about earnings and cost beliefs (Q4 and Q5). The last rows are the p-value of an F-test for joint significance in an OLS regression of treatment on each of the demographic controls listed, controlling for randomization blocks and using robust clustered standard errors (at the high school level for 2012 high school graduates, individual level for 2009-11 graduates). Sample sizes for regressions corresponding to each p-value are in brackets. Different sample sizes for each demographic variable are due to incomplete administrative data.

Table A.5: Balance of Possible Outcome Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Pooled			Low SES			High SES	
	Control	Treated	P-Value Diff	Control	Treated	P-Value Diff	Control	Treated	P-Value Diff
Entrance exam score	535.6	537	0.566	481.3	485.4	0.116	571.1	569.6	0.992
	(93.41)	(92.61)	[47,568]	(78.93)	(77.84)	[15,606]	(83.62)	(84.39)	[29,628]
Loan recipient	0.752	0.749	0.365	0.655	0.667	0.448	0.826	0.815	0.0538
	(0.432)	(0.433)	[49,166]	(0.475)	(0.471)	[16,594]	(0.379)	(0.389)	[29,850]
Matriculate to degree with age 26 earnings	0.836	0.835	0.890	0.827	0.825	0.898	0.844	0.844	0.909
	(0.370)	(0.371)	[37,747]	(0.378)	(0.380)	[12,300]	(0.363)	(0.363)	[23,448]

Tests of treatment effects on intermediate outcomes determined after treatment for samples listed in columns. 'Pooled' is full sample. 'Low-SES' and 'High SES' samples are as defined by high school income ratings. See text for details. Means with standard deviations in parentheses are presented in columns 1, 2, 4, 5, 7, 8. Columns 3, 6, and 9 with the exception of the final row, present the p-value of treatment when the demographic variable in each row was regressed on treatment using an OLS regression that controls for randomization blocks and has robust standard errors clustered at the high school of graduation. The randomization blocks for 2012 high school graduates were based on RBD characteristics, and randomization was done at the high school level. The randomization blocks for older high-school graduates were based on previous entrance exam scores, with randomization at the individual level. Sample sizes are in brackets in Columns 3, 6, 9. Entrance exam (PSU) scores are averages of the Math and Language components. "Loan recipient" is an indicator defined if the student is a loan recipient from the two sources of government funded loans: CAE and FSCU. See section 2 for details. Sample sizes for regressions corresponding to each p-value are in brackets. Different sample sizes for each demographic variable are due to incomplete administrative data.

Table A.6: Exploded Logits

	1	2	3
Observed earnings	0.1158**		0.0630
	(0.0347)		(0.0450)
Observed costs	-0.4992**		-0.6187**
	(0.0494)		(0.0675)
Earnings beliefs		0.3260**	0.3543**
		(0.0453)	(0.0467)
Cost beliefs		-0.0823*	0.0478
		(0.0373)	(0.0407)
N	57740	42558	42483

Coefficient estimates from exploded logits. Sample: students' elicited preferences (up to three per student). Dependent variable is listed rank. "Observed earnings" and "Observed Costs" are authors' calculations of observed values for past graduates. See section 5 for details. "Earnings beliefs" and "Cost beliefs" are students' reported own-earnings and own-cost beliefs from survey data. Standard errors cluster at student level.

A.4 Additional outcome tables and summary statistics

Table A.7 presents Table 3 results with no controls and the dependent variable defined as the change in the outcome of interest. Change in outcome variables is defined as the difference between the matriculated value and first-choice value of the variable of interest.

Table A.7: Impact of treatment on change in outcome variables (no controls)

	Pooled	Low-SES	High-SES
Matriculation to higher education in 2013	0.004	0.003	0.001
	(0.005)	(0.008)	(0.006)
All Students			
Net Value	8,402	20,356*	3,151
	(7,128)	(9,169)	(9,819)
Earnings gains	8,983	21,505*	3,615
	(7,905)	(10,029)	(10,904)
Monthly debt	85.5	615	-91.4
	(736)	(959)	(1,023)
Conditional on Matriculation			
Net Value	10,544*	22,909**	6,914
	(4,583)	(7,884)	(5,879)
Earnings gains	11,841*	24,207**	8,349
	(4,921)	(8,466)	(6,312)
Monthly debt	489	1,104	203
	(469)	(778)	(622)
Degree average earnings at age 26	6,342+	13,703*	3,135
	(3,371)	(5,534)	(4,405)
Monthly payment (conditional on enrollment)	767	1,248	473
	(490)	(854)	(596)
Degree graduation rate ('00-'05 cohorts)	0.004	0.000	0.005
	(0.003)	(0.005)	(0.004)
Expected length of matric. degree ('00-'05 cohorts)	0.016	0.019	0.012
	(0.019)	(0.034)	(0.023)

Change in outcome variables is defined as the difference between matriculated value and first-choice value of the variable of interest (variables of interest are reported in the first column). Low-SES is defined as the lowest two income quintiles as defined by their high school (RBD) poverty ratings defined by Mineduc in 2010, 2008, 2006, 2003, 2001; high-SES is the top three income quintiles. Table reports coefficients on Treatment from a regression of the dependent variable (row) on treatment. Clustered standard errors are in parentheses. This is the same output as Table 3, with no controls for first-choice degree or randomization block. Regression results in the second panel combine extensive and intensive margins; values of the outcome variables are set to zero if the respondent didn't matriculate anywhere in 2013. The third, fourth, and fifth panels report intensive margin effects, set to missing the outcome variable of interest if the respondent didn't matriculate to a higher education degree in 2013. Net Value, earnings gains, and monthly debt are the values for degrees as exhibited in our experiment. Degree average earnings and monthly payment values are based on OLS predictions at condition on student observables. See section 3 for details of earnings and cost variable construction. Degree graduation rate is the average degree graduation rate for the 2000-2005 freshmen cohorts. Expected length of degrees are calculated based on the average amount of time enrolled in the degrees for 2000-2005 freshmen cohorts. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.8: Changes in enrollment by institution-major, major, and broad field

Institution-majors					Majors				Fields			
Inst. Name	Major name	Fraction enrollment change	Monthly earnings	Monthly costs	Name	Fraction enrollment change	Monthly earnings	Monthly costs	Name	Fraction enrollment change	Monthly earnings	Monthly costs
Losses					*Losses*							
IP DR. VIRGINIO GOMEZ G.	Técnico Otros de Salud	-0.112	164831	24292	Técnico Dental y Asistente de Odontología	-0.0603	212508	23279	Educación	-0.0136	457008	53711
CFT SANTO TOMAS	Técnico Laboratorista Dental	-0.110	164264	23832	Técnico Asistente del Educador Diferencial	-0.0563	185566	19447	Humanidades	-0.0124	502398	66370
IP AIEP	Técnico Dental y Asistente de Odontología	-0.0878	172206	25519	Técnico en Sonido	-0.0472	341307	28837	Arte y Arquitectura	-0.0119	558391	81459
CFT DEL MEDIO AMBIENTE	Técnico Otros de Salud	-0.0833	113020	27334	Técnico Asistente del Educador de Párvulos	-0.0440	200837	17749	Salud	-0.00690	780950	87596
IP DR. VIRGINIO GOMEZ G.	Técnico Dental y Asistente de Odontología	-0.0725	174438	17187	Técnico en Gastronomía y Cocina	-0.0438	232447	25209	Ciencias Básicas	-0.00675	763616	86526
IP DUOC UC	Técnico en Sonido	-0.0687	293111	29265	Técnico en Turismo y Hotelería	-0.0343	287145	22507	Agropecuaria	-0.00484	705220	91502
IP DUOC UC	Técnico Dental y Asistente de Odontología	-0.0669	205709	25926	Técnico Otros de Salud	-0.0326	250424	23672	Administración y Comercio	0.000725	574718	38349
CFT SANTO TOMAS	Técnico en Gastronomía y Cocina	-0.0635	210440	25611	Técnico Jurídico	-0.0230	312598	19488	Ciencias Sociales	0.00174	872852	85728
CFT SANTO TOMAS	Técnico Asistente del Educador Diferencial	-0.0563	185566	19447	Artes y Licenciatura en Artes	-0.0209	417037	79483	Derecho	0.00809	1.602e+06	100502
IP SANTO TOMAS	Técnico Otros de Tecnología	-0.0553	315020	24486	Técnico en Deporte, Recreación y Preparación Física	-0.0208	295710	21521	Tecnología	0.0125	1.004e+06	60494
Gains					*Gains*							
IP LA ARAUCANA	Ingeniería en Prevención de Riesgos	0.0491	881497	28650	Ingeniería Civil Metalúrgica	0.0199	1.955e+06	112721		.	.	.
CFT INACAP	Técnico en Instrumentación, Automatización y Control Industrial	0.0506	781965	19522	Ingeniería Civil en Minas	0.0214	2.558e+06	111522		.	.	.
UNIVERSIDAD DE LOS LAGOS	Técnico en Prevención de Riesgos	0.0513	603416	17314	Técnico en Mantenimiento Industrial	0.0227	670315	23653		.	.	.
UNIVERSIDAD IBEROAMERICANA DE CIENCIAS Y TECNOLOGIA, UNICYT	Enfermería	0.0562	1.308e+06	78232	Técnico en Electrónica y Electrónica Industrial	0.0248	707745	23062		.	.	.
UNIVERSIDAD DE LA SERENA	Ingeniería en Minas y Metalurgia	0.0583	2.070e+06	64850	Ingeniería en Prevención de Riesgos	0.0249	992835	44487		.	.	.
UNIVERSIDAD CATOLICA DE LA SANTISIMA CONCEPCION	Ingeniería en Prevención de Riesgos	0.0595	1.030e+06	29893	Ingeniería Mecánica	0.0250	1.291e+06	62512		.	.	.
IP LA ARAUCANA	Contador Auditor	0.0603	899292	31549	Construcción Civil	0.0271	1.088e+06	65963		.	.	.
CFT CAMARA DE COMERCIO DE SANTIAGO	Técnico en Administración de Empresas	0.0610	589027	14845	Técnico en Instrumentación, Automatización y Control Industrial	0.0287	753668	25162		.	.	.
IP SANTO TOMAS	Ingeniería en Prevención de Riesgos	0.0649	1.158e+06	34006	Ingeniería en Electricidad	0.0288	1.355e+06	58914		.	.	.
CFT MAGNOS	Técnico en Enfermería	0.0677	597682	20030	Técnico en Contabilidad General	0.0315	490699	20576		.	.	.
UNIVERSIDAD DE LAS AMERICAS	Ingeniería Civil Industrial	0.0695	1.897e+06	70420	Ingeniería en Minas y Metalurgia	0.0385	1.611e+06	49077		.	.	.

Percent changes in enrollment for top and bottom ten institution-majors combinations and institutions (left and center panels, respectively) and enrollment changes by field (right). Changes reflect differences between full treatment and no-treatment counterfactuals. See section 6 for simulation details.

Table A.9: Selected logit coefficient estimates

	All	Low SES	High SES	Low PSU	High PSU	Low-SES, low-PSU	All	Low SES	High SES	Low PSU	High PSU	Low-SES, low-PSU
Treat × earn	0.0897+	0.187*	0.0354	0.0921	0.107	0.15	0.0875	0.185*	0.0332	0.092	0.1	0.15
	(0.054)	(0.088)	(0.068)	(0.070)	(0.082)	(0.098)	(0.054)	(0.088)	(0.068)	(0.070)	(0.082)	(0.098)
Treat × cost	-0.0315	-0.104	0.00826	-0.0741	0.0458	-0.0917	-0.0324	-0.107	0.00801	-0.0734	0.044	-0.0885
	(0.049)	(0.078)	(0.064)	(0.062)	(0.083)	(0.086)	(0.049)	(0.079)	(0.064)	(0.062)	(0.083)	(0.086)
Treat × rec'd							0.124	0.191	0.0911	-0.0506	0.27	-0.239
							-0.155	-0.264	-0.192	-0.236	-0.207	-0.318
Earnings	0.283***	0.462***	0.271***	0.416***	0.109	0.571***	0.283***	0.460***	0.271***	0.414***	0.113	0.569***
	(0.072)	(0.118)	(0.089)	(0.104)	(0.105)	(0.134)	(0.072)	(0.118)	(0.089)	(0.104)	(0.105)	(0.134)
Costs	0.689***	0.745***	0.447***	1.167***	0.192+	0.983***	0.693***	0.750***	0.449***	1.168***	0.197+	0.984***
	(0.071)	(0.120)	(0.084)	(0.112)	(0.102)	(0.143)	(0.071)	(0.120)	(0.084)	(0.112)	(0.102)	(0.143)
Earn × low SES	0.0951			0.120+	0.178		0.0936			0.120+	0.176	
	(0.058)			(0.071)	(0.110)		(0.058)			(0.071)	(0.110)	
Cost × low SES	-0.382***			-0.446***	-0.267*		-0.382***			-0.446***	-0.267*	
	(0.052)			(0.063)	(0.104)		(0.052)			(0.064)	(0.104)	
Mean PSU	0.0255***	0.00931	0.0305***	0.102***	0.0631***	0.0674***	0.0254***	0.0092	0.0304***	0.101***	0.0630***	0.0672***
	(0.004)	(0.007)	(0.004)	(0.010)	(0.006)	(0.014)	(0.004)	(0.007)	(0.004)	(0.010)	(0.006)	(0.014)
Pref'd region	0.664***	0.675***	0.644***	0.570***	1.126***	0.635***	0.665***	0.676***	0.646***	0.571***	1.128***	0.636***
	(0.067)	(0.089)	(0.103)	(0.074)	(0.155)	(0.095)	(0.067)	(0.089)	(0.103)	(0.074)	(0.155)	(0.095)
Pref'd area	1.272***	1.244***	1.283***	1.327***	1.197***	1.297***	1.287***	1.260***	1.296***	1.340***	1.212***	1.309***
	(0.033)	(0.053)	(0.042)	(0.045)	(0.048)	(0.060)	(0.033)	(0.054)	(0.042)	(0.045)	(0.049)	(0.060)
Pref'd major	2.306***	2.354***	2.285***	2.162***	2.446***	2.196***	2.288***	2.334***	2.269***	2.145***	2.429***	2.180***
	(0.032)	(0.057)	(0.038)	(0.049)	(0.043)	(0.067)	(0.032)	(0.057)	(0.039)	(0.049)	(0.044)	(0.068)
Pref'd inst.	1.819***	1.827***	1.817***	1.797***	1.877***	1.761***	1.819***	1.828***	1.817***	1.797***	1.877***	1.761***
	(0.033)	(0.057)	(0.041)	(0.049)	(0.045)	(0.066)	(0.033)	(0.057)	(0.041)	(0.049)	(0.045)	(0.066)
Rec'd							-0.361***	-0.430*	-0.307*	-0.290+	-0.380*	-0.175
							(0.121)	(0.209)	(0.149)	(0.174)	(0.168)	(0.230)
N	13112133	4518344	8593789	7139390	5972739	3427651	13112133	4518344	8593789	7139390	5972739	3427651

Selected coefficient estimates from equation (5). Observations are at student-option level for degree programs in each student's choice set. Columns define student sub-samples. Columns 1-6 are baseline specifications. Columns 7-12 add an indicator variable equal to one if a degree program was in the narrow major recommended to the student by our choice application (or, for control students, the degree program that would have been recommended had the student been treated). 'Treat' is dummy for treatment. 'Earn' and 'cost' are log monthly earnings and costs, respectively. 'Mean PSU' is mean of combined math and reading scores for admitted students at program. 'Preferred region' is an indicator for programs in same region as a student's top choice region. SES gap is the absolute value of the gap between a program's high-SES share and the high-SES share at a student's top choice. 'Pref'd area,' 'Pref'd major,' and 'Pref'd inst.' are dummies equal to one if a degree is in a student's most-preferred broad area, narrow major, or institution, respectively. 'Rec'd' is a dummy equal to one if a degree was recommended to a treated student (or would have been recommended to a control student). Standard errors clustered at student level. +: $p < 0.10$, *: $p < 0.05$, **: $p < 0.01$ ***: $p < 0.001$.

B Survey design and materials

B.1 Sample Selection

We divide the sample universe into two groups: current high school seniors (as of 2012) and past high school graduates. We sent survey invitations to all students who had completed the *Formulario Único de Acreditación Socioeconómica* [A.2] (FUAS) for the 2013 school year. Invites were sent from a Mineduc email account associated with the FUAS application (fuas2012@mineduc.cl). The email requested applicants to participate in a brief survey that would be used by Mineduc to make decisions about higher education, and assured recipients that survey responses would not affect their FUAS applications in any way.

B.1.1 Current High School Seniors

The base sample for current high school was the population of high school seniors in 2012. Schools were randomized into groups based on school type (public, voucher, private), PSU performance in the past two years (2010, 2011), school size, and PSU test-taking rate among current high school seniors. A total of 3,136 high schools were divided into sub-groups based on private, voucher or public school types. The schools were divided into further sub-groups based on PSU score distribution among senior students within the school in the past two years. Next, we stratified schools into three groups by number of classes to reflect school size. Finally, schools were divided into two groups based on the fraction of the 2012 graduating class who registered to take the PSU. This resulted in a total of 61 randomization groups.

Within each randomization group, 50% of the schools were treated and 50% were assigned to the control group. Students in both the treatment and control groups were invited to take an online survey, conditional on applying to FUAS. However, only those in the treatment group were treated with an information sheet at the end of the survey, and provided with a link to the searchable database at the end of the survey questionnaire.

B.1.2 Past Students

The base sample for older students was those registered for PSU 2013 and not currently in high school as of 2012. Randomization for older students was based on past PSU score groups of 50 points. Those taking the test for the first time were placed in a separate group. Within each randomization group, 50% of the students were assigned to treatment and 50% of the students were assigned to control groups. Students in both the treatment and control groups were invited

[A.2] The Formulario Único de Acreditación Socioeconómica (FUAS), or Single Form for Socioeconomic Accredition, is analogous to the Free Application for Federal Student Aid (FAFSA) used in the United States.

to take an online survey, conditional on applying to FUAS. However, only those in the treatment group were treated with an information sheet at the end of the survey.

B.2 Survey Administration

We administered the survey online. The survey was titled 'Project 3E 2012,' and implemented using a custom programmed survey on the dedicated website `www.3e2012.cl`. Beginning in mid-November 2012, students in the sample described above received email invitations from Mineduc to participate in the survey, and could click through to log in. Waves of invitations and reminders were sent out before the 2012 PSU exams were administered in December, with reminders sent up though early February 2013, by which time institutions affiliated with the Consejo de Rectores de las Universidades Chilenas (CRUCH) had informed students of the results of their applications.

We piloted a longer version of the survey in 2011. The 2012 version was simplified to six questions. These questions asked for specific enrollment plans (the top three degree enrollment choices), certainty of application plans, expectations about earnings upon completing each degree, expectations about degree costs, and expected performance on the PSU. Section B.3 provides Spanish Language and English translations of the questions. Section B.5 shows screen shots of what the survey looked like to the respondents.

Upon logging in, a single question appeared on each screen of the survey, and participants were routed through the appropriate survey questions, depending on their prior responses. The website did not allow students to change their prior responses. If a student logged off and logged back in, they would return to the last unanswered question.

Upon completing their survey, treated students were shown an information sheet. The first part of the information sheet presented information on the degrees respondents identified as their top choices. The information sheet provided information about the monthly increase in earnings a graduate could expect to earn and the expected monthly loan payment needed to cover the cost of the degree, both calculated over a 15 year horizon. The info sheet also provided the difference between the two values, which we term the 'Net Value.'

The info sheet also provided treated students with two search prompts. The first informed respondents if there were other institutions offering their first-choice major with a higher Net Value. The second let respondents know if there were other institution-major combinations in their first-choice broad field which offered a higher Net Value. We chose these degree programs from the set of degree programs that in the past had admitted students with with PSU scores similar to those admitted at the respondent's first-choice program. To do this, the web program accessed a back-end database containing information on test score distributions and Net Values for the set of degree programs.[A.3] Students were shown the potential monthly gain in pesos for switching, but not the names of the specific degree programs. This was intended to communicate the importance of searching on the subsequent page. If the student's own choices already offered the highest Net Value within the academic area or major, they were shown the gains from switching to the highest Net Value degree in any of the ten academic areas. If the student's own choices already offered the highest Net Value across all academic areas, they received a message communicating this to them. Sections B.3 and B.5 provide text and screen shots of the information page.

Treated students still needed to click through to a next screen to see their confirmation code. This next screen was the searchable database. The career search application prompted individuals to enter a career of interest (from a dropdown menu), and a PSU score. It then populated a table of degrees they could likely get into, sorted by Net Value in descending order. If earnings gains were missing (for example the degree had no graduates), the degree was placed under those with present net value, and all such degrees were sorted in ascending order by tuition. The Net Value was replaced with a *! and students could see a note if they hovered over it that stated that there were not enough graduates from this degree to calculate expected earnings gains.

If a student in the treated group completed the survey and performed a search, when they logged back in they would go directly to the search page and their prior table would be saved for them. They could save up to 10 different search tables in their survey cache. All 10 would appear when they logged in and could be collapsed to facilitate viewing and comparison. The web program recorded all logins, responses, and search activity.

B.3 Survey Questions

Here is the text in Spanish (original) with English translations for each question in Proyecto 3E 2012.

0. Full text of Mineduc's informed consent ("terms and conditions") box:

   El Ministerio de Educación está realizando una encuesta con el objeto de entender cómo los jóvenes chilenos se informan y toman decisiones sobre educación superior. Las respuestas a la encuesta son confidenciales y sólo se utilizarán con fines estadísticos para la investigación académica y para informar a las polticas públicas.

[A.3] The way we operationalize this is the following. We take the 25th, 50th and 75th percentile historic PSU values of the student's reported first-choice degree. We find all degrees in the same major for which one of the degree's 25th, 50th or 75th percentiles of PSU scores for enrollees at the first-choice school falls within the the 5th-95th percentile range of historic enrollees. Historic PSU scores are defined as the most recent PSU scores for enrollees in each degree during the 2011-2012 school year. We then search for the highest Net Value degree satisfying both the score restriction and major or broad area restriction. Broad academic area was defined by CINE-UNESCO areas of study.

Como en toda encuesta, no hay respuestas correctas ni incorrectas. Estamos interesados en tus experiencias y opiniones con el propósito de poder recabar mayores antecedentes y as alimentar la investigación para mejorar las polticas pblicas en educación superior. Tus respuestas no afectarán de ninguna forma tu postulación.

Finalmente, todas las respuestas serán almacenadas en un servidor anónimo y seguro, que garantiza la confidencialidad de todos los participantes de la encuesta. Si tienes alguna pregunta acerca de este estudio, puedes comunicarte con los supervisores de este proyecto por correo electrónico a FUAS2012@mineduc.cl.

Al hacer clic en el botón de abajo "Aceptar", junto con manifestar tu intención de participar en este estudio, certificas que has ledo la información anterior.

Te reiteramos que tu participación en la encuesta (o tu decisión de no responderla) no afectará tu proceso de postulación. Haz clic en "Aceptar" para empezar.

The Ministry of Education is conducting a survey with the goal of understanding how young Chileans inform themselves and make decisions about higher education. The answers to the survey are confidential and will only be used for statistical purposes for academic research and to inform public policy.

As in all surveys, there are no right or wrong answers. We are interested in your experiences and opinions, with the goal of being able to gather more information and to strengthen research in order to improve public policy regarding higher education. Your answers will not affect your application in any way.

Finally, all of the answers will be stored on an anonymous and secure server, which guarantees the confidentiality of all survey participants. If you have any questions about the study, you can contact the project supervisors by email at FUAS2012@mineduc.cl.

Upon clicking the button below that says Accept, along with expressing your intention of participating in this study, you are certifying that you have read the information above.

We remind you that your participation in the survey (or your decision not to participate) will not affect your application process.

Click on Accept to begin.

1. ¿Cuál de las siguientes opciones te describe mejor? Por favor elige una.
   Which of the following options best describes you? Please choose one.

   (a) Estoy postulando a una institución de educación superior por primera vez.
   I am applying to an institute of higher education for the first time.

   (b) Estoy estudiando actualmente en una institución de educación superior, y voy a quedarme en la misma carrera en la misma institución en el año que viene.

I am currently enrolled at an institute of higher education and will stay in the same major at the same school next year.

(c) Estoy estudiando actualmente en una institución de educación superior, pero pienso cambiarme a otra carrera en la misma institución en el año que viene.
I am currently enrolled at a higher education institution but I plan to enroll at a different career at the same institution next year)

(d) Estoy estudiando actualmente en una institución de educación superior, pero pienso cambiarme a otra institución con la misma carrera en el año que viene.
I am currently enrolled at a higher education institution, but I plan to go to a different institution next year.

(e) Estoy estudiando actualmente en una institución de educación superior, pero pienso cambiarme a otra institución en el año que viene, cambiendo también de carrera.
I am currently enrolled at a higher education institution, but I plan to go to a different institution next year, and also plan to change my major.

(f) Ya empecé mis estudios de educación superior, pero tuve que dejarlos un tiempo, y estoy empezando nuevamente.
I began my higher education studies, but I had to leave for a while, and am starting again.

2. Imagina que hoy debes postular a las instituciones y a las carreras de educación superior que deseas. Por favor elige hasta 3 opciones de acuerdo a tus preferencias.
Imagine that today you have to apply to the higher education schools and majors that you would like. Please list up to 3 options according to your preferences.

Si te encuentras actualmente estudiando en educación superior, indica también tus preferencias.
If you are currently enrolled studying in higher education also list your preferences.

[A series of dropdown boxes are shown, with the headings below]

Opción 1	Institution Type → Nivel → Institución → Carrera
1st Choice	*Institution Type → Level → Institution → Major*
Opción 2	Institution Type → Nivel → Institución → Carrera
2nd Choice	*Institution Type → Level → Institution → Major*
Opción 3	Institution Type → Nivel → Institución → Carrera
3rd Choice	*Institution Type → Level → Institution → Major*
Opción 4	Institution Type → Nivel → Institución → Carrera
4th Choice	*Institution Type → Level → Institution → Major*
Opción 5	Institution Type → Nivel → Institución → Carrera
5th Choice	*Institution Type → Level → Institution → Major*

3. ¿Qué tan seguro(a) estás que las opción(es) que mencionaste anteriormente serán a las que efectivamente postules el próximo año?
How sure are you about the option(s) you listed will be the ones to which you will apply next year.

1 No estoy para nada seguro(a). *I am not sure at all.*
2 Estoy un poco seguro(a). *I am a little sure.*
3 Estoy medianamente seguro(a). *I am fairly sure.*
4 Estoy bastante seguro(a). *I am pretty sure.*
5 Estoy absolutamente seguro(a). *I am absolutely sure.*

4. ¿Considerando los costos de matrcula y arancel, aproximadamente cuáles crees que serán los costos ANUALES de estudiar la(s) carrera(s) en la(s) institución(es) elegida(s) anteriormente?
Considering the costs of registration and tuition, approximately how much do you think the ANNUAL costs are for studying in the institution(s) previously selected?

Opción 1 de pregunta 2	Slider - No sé - Checkbox	
Respondent's 1st choice from question 2	Slider to choose values	Don't know - Checkbox
Opción 2 de pregunta 2	Slider - No sé - Checkbox	
Respondent's 2nd choice from question 2	Slider to choose values	Don't know - Checkbox
Opción 3 de pregunta 2	Slider - No sé - Checkbox	
Respondent's 3rd choice from question 2	Slider to choose values	Don't know - Checkbox

5. ¿Cuánto crees que será TU sueldo mensual al comenzar a trabajar, una vez titulado(a), con trabajo estable de tiempo completo? Responde a continuación en la columna izquierda.
What do you think YOUR monthly salary will be once you graduate and start to work in a stable, full-time job? Please respond below in the left-hand column.

¿Cuánto crees que será el sueldo mensual de UN GRADUADO TÍPICO al comenzar de trabajar una vez titulado(a) con trabajo estable de tiempo completo? Responde a continuación en la columna derecha.
What do you think the monthly salary will be FOR A TYPICAL GRADUATE once s/he graduates and starts to work in a stable, full-time job? Please respond below in the right-hand column.

[1ª carrera de pregunta 2] de [1ª institución de pregunta 2]
[1st choice major from question 2] from [1st choice institution from question 2]

Mi ingreso mensual será	No sé	El ingreso mensual de un titulado tipo	No sé
SLIDER	Checkbox	será SLIDER	Checkbox

[2ª carrera de pregunta 2] de [2ª institución de pregunta 2]
[2nd choice major from question 2] from [2nd choice institution from question 2]

Mi ingreso mensual será SLIDER	No sé Checkbox	El ingreso mensual de un titulado tipo será SLIDER	No sé Checkbox

[3ª carrera de pregunta 2] de [3ª institución de pregunta 2]
[3rd choice major from question 2] from [3rd choice institution from question 2]

Mi ingreso mensual será SLIDER	No sé Checkbox	El ingreso mensual de un titulado tipo será SLIDER	No sé Checkbox

6. ¿Con qué puntaje PSU en matemáticas y en lenguaje crees que postularás a la educación superior? Por favor responde sinceramente.
   With which PSU score in mathematics and language do you think you will apply to higher education? Please respond sincerely.

   Matemáticas ___________
   Mathematics ___________

   Lenguaje ___________
   Language ___________

   = FIN / *END* =

B.4 Constructing the components of Net Value

As described in Section 3, we worked with Mineduc to compute the measures of earnings returns and costs presented in the information intervention. Policymakers imposed several constraints on the types of information we could display:

1. We could display earnings conditional on graduation, not enrollment. This reflected both a policy preference for Mineduc and a data restriction. At the time of our intervention in 2012, enrollment data was not available prior to 2007. This time frame was not long enough to compute earnings outcomes at most programs. Later Mineduc changed their approach to focus on enrollment rather than graduation outcomes (Beyer et al. 2015). They collected further historic data from higher education institutions to permit us to calculate a full panel of enrollment, matriculation and graduation for cohorts from 2000 through 2013. See section 3 for a discussion of how we use these enrollment records to evaluate the effects of the intervention.

2. We could display only cell means, not regression-adjusted predictions. In designing a subsequent accountability policy, Mineduc later chose to adopt loan caps based on regression-adjusted expected earnings, to prevent institutions from manipulating ratings using selective admission (Beyer et al. 2015).

3. We were initially (i.e., in the planning phase of the project) restricted to simply present mean earnings over the first five years of experience. We demonstrated to Mineduc that this would overstate the benefits of two year and less-selective degrees (leading to jobs with low wage profiles) and understate benefits of longer and/or more selective degrees in which career growth and wage growth continues after the first few years of experience. We were allowed to estimate growth rates from our data by subject area and degree level to correct for this problem.

B.4.1 Earnings Gains

We compute earnings gains as described in section 3.4. Here we provide some additional detail. Our earnings data span tax years 2005-2012. We use these data to compute two types of statistics: earnings for students who never enroll in college (μ_{0t} in equation 1) and earnings for graduates of different degree programs j (mu_{jt} in equation 1).

We compute earnings for non-enrollees using data on high school graduates who did not appear as matriculating into any higher education degree within four years of high school completion. We use data on high school graduates from 2006 and 2007 to compute these statistics. We choose these years because the earliest college matriculation data we had at the time of the survey was for the 2007 academic year. This meant we could not verify non-enrollment for pre-2006 cohorts of high school graduates.

We calculated earnings for college graduates using data on the population of graduates for the years 2001 and 2005-2009. These were the only graduation records available at the time of our survey. These data cover graduates from all higher education degree programs in Chile. We first use the data to compute degree-level mean earnings gains for experience years 1-5. Gains here are defined relative to earnings for non-enrollees.

To account for long-term gains in enrolling in a higher education degree, we project earnings through experience year 15. We were allowed to do this by field and degree level, though not by the selectivity level of the degree within field and degree level. Each degree falls within one of ten academic areas (Business Administration, Agriculture, Art & Architecture, Basic Sciences, Social Sciences, Law, Education, Humanities, Health, or Technology) and within a type of degree: professional or technical. We calculate earnings growth rates for each academic area-type combination using a simple linear regression of average earnings gains for degree j in experience year e on dummy variables for experience year. We find that the dummy variables imply a linear relationship between earnings and experience year. We estimate the linear relationship and use the estimates to project earnings by area and degree level.

We take a present discounted value of mean earnings through year 15 post-graduation relative to mean earnings through the same time period for those not going to college and convert it into a monthly payment equivalent using a 2% discount rate. We choose the 15 year period because that is the repayment period for student loans, and the 2% discount rate because that is the interest rate on student loans. Note that, as per Mineduc's request, we do not include opportunity cost of earnings while in school in our Net Value calculation.

B.4.2 Costs of Graduating from a Degree

Chilean students pay two types of direct costs: one-time matriculation fees and annual tuition. Our cost measures include both. We calculate annual tuition values for all degrees that students matriculated to in 2012 using data reported to Mineduc. Where nonreports led to missing tuition values, we fill in using inflation adjusted values from prior years. Our historical tuition data extend back to 2008. We use the same process to calculate degree level matriculation fees. For those degrees that are still missing matriculation fees, we fill in the average matriculation fee by major and institution type (Technical Institute, Professional Institute, Private University, or CRUCH University). Matriculation fees do not vary significantly at this level.

To calculate a monthly payment for the costs of enrolling in a degree, we use the duration of the degree as reported in our matriculation files from the Ministry of Education. Degree duration is reported in semesters, so we divide these values by two and round up if there are durations with half-year increments. To construct the total amount that would be paid for the entirety of the degree, tuition values are brought forward with interest (which is set to 2% APR per college loan rates) for the appropriate number of years to completion of the degree. For example, the annual tuition for the 2nd year of enrollment in a degree with duration of 4 years is brought forward with interest for 2 years. The matriculation fees are also brought forward with interest to time of graduation from the degree. The total amount of tuition is then amortized as a monthly amount over 180 periods (15 years $\times$ 12 months).

B.4.3 Net Value

The monthly Net Value displayed to the treatment group in our survey is simply the monthly present-value of earnings gains less the monthly present-value of the degree costs.

B.5 Screen shots of survey and treatment

Figure A.1: Introduction Page

Figure A.2: Consent Page

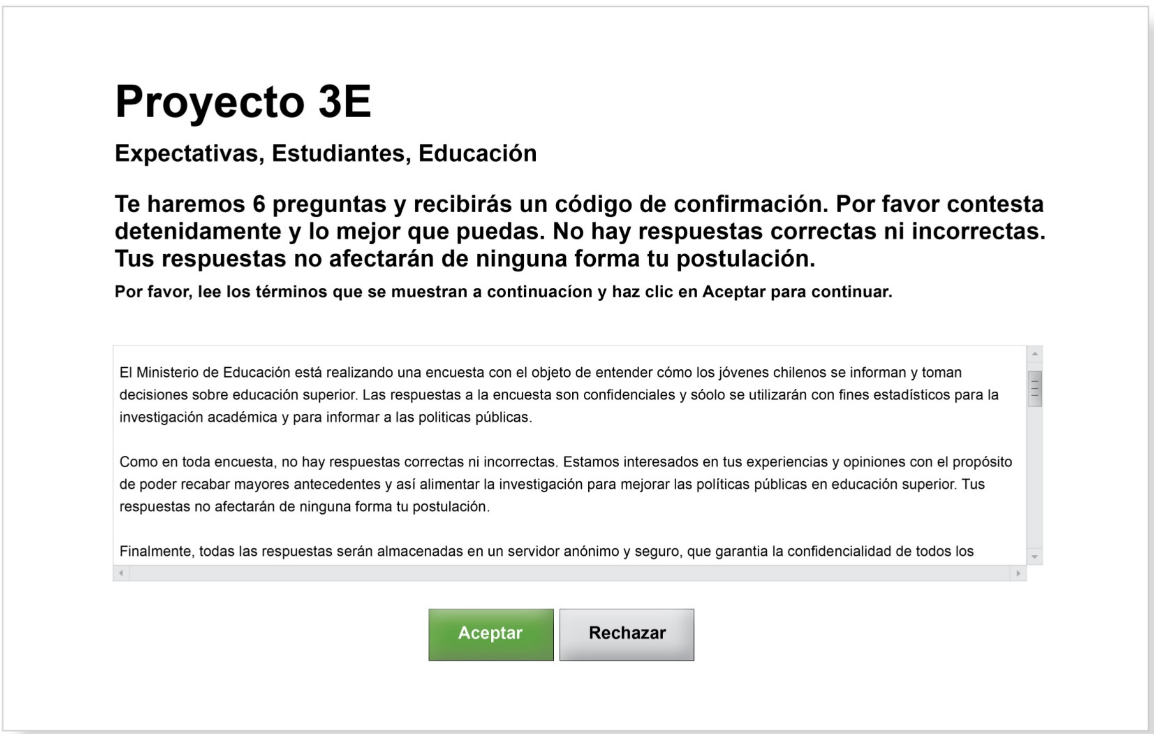
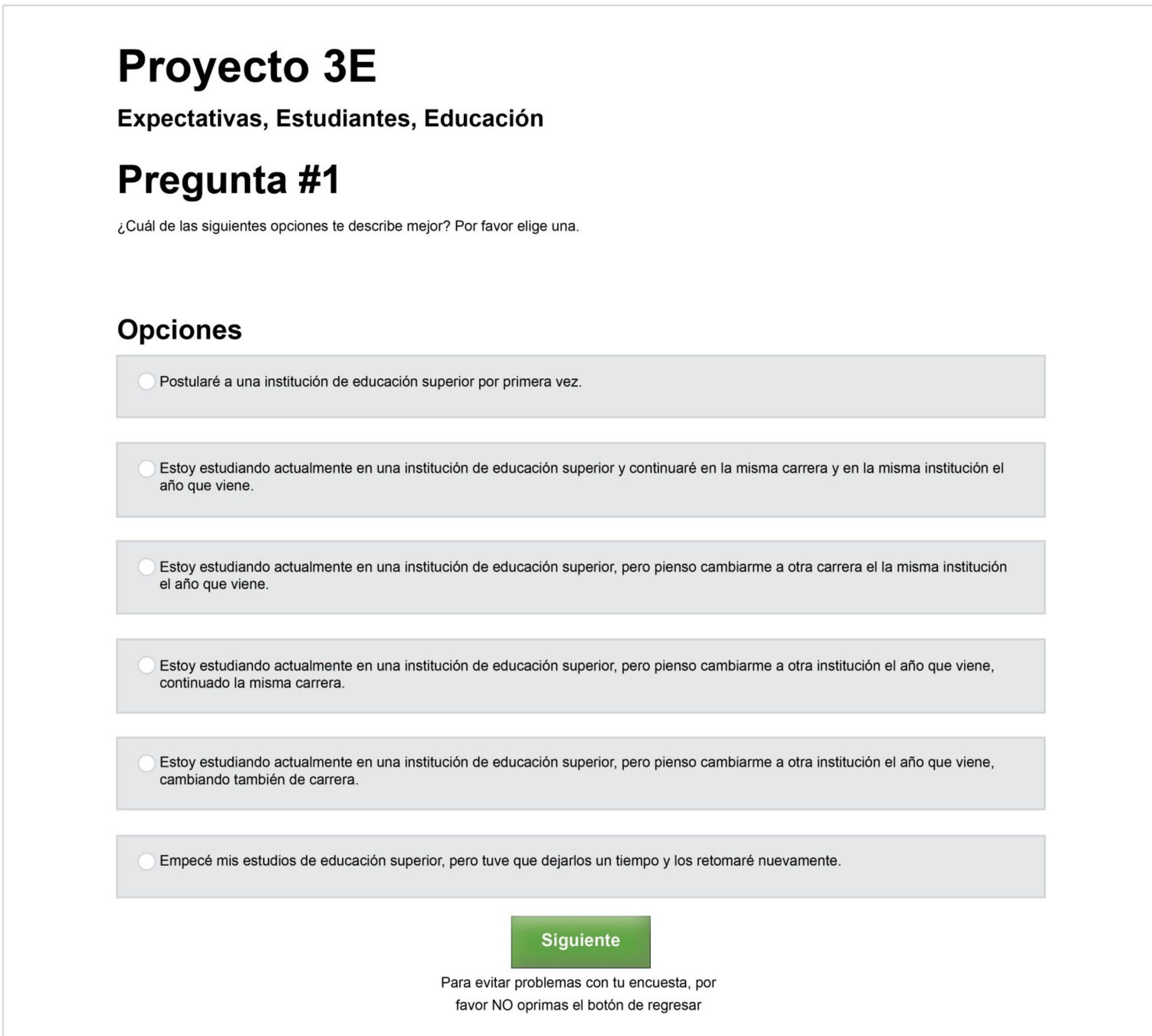


Figure A.3: Question 1: Confirming current education enrollment status


Proyecto 3E

Expectativas, Estudiantes, Educación

Pregunta #1

¿Cuál de las siguientes opciones te describe mejor? Por favor elige una.

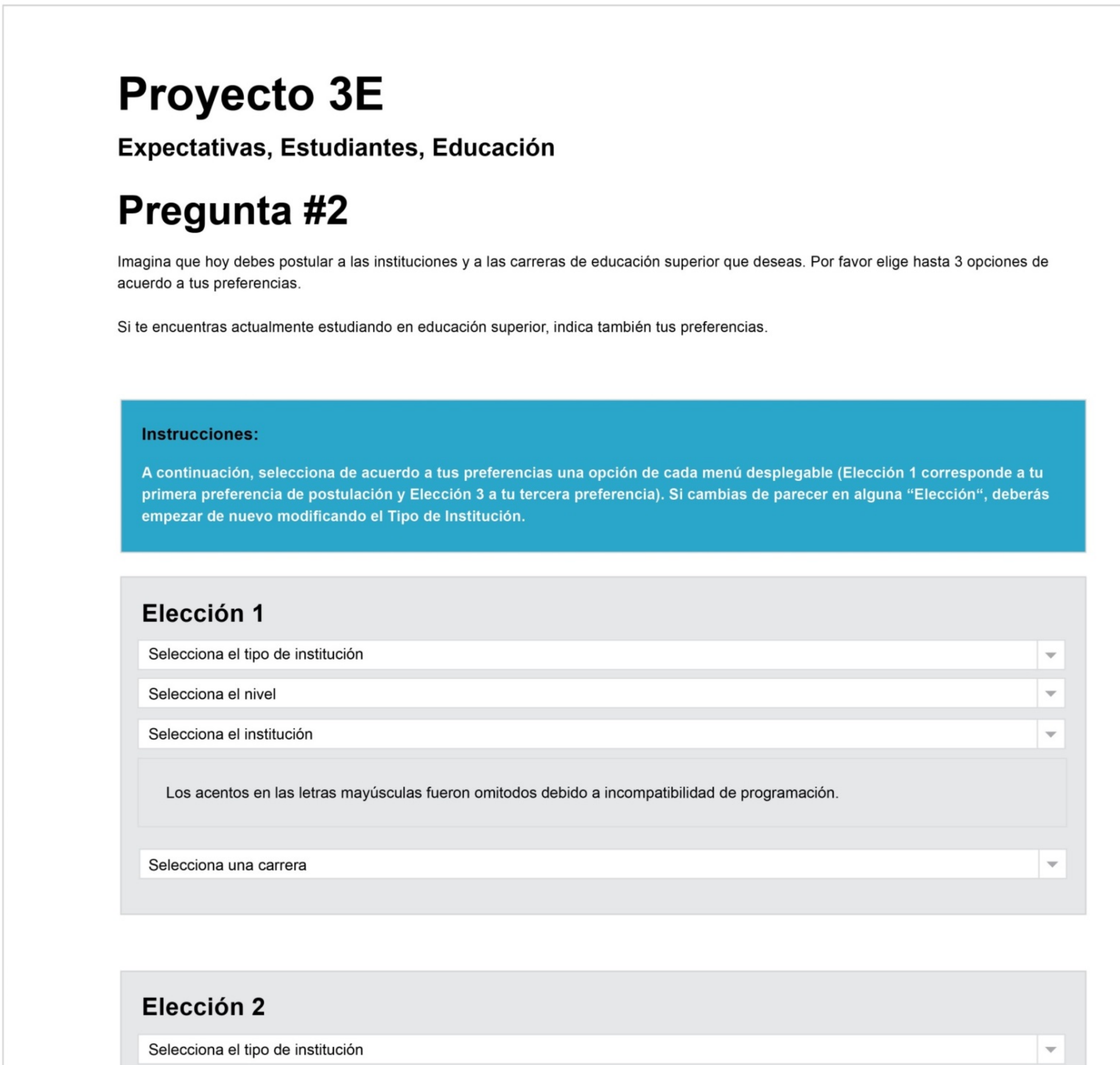
Opciones

- Postularé a una institución de educación superior por primera vez.
- Estoy estudiando actualmente en una institución de educación superior y continuaré en la misma carrera y en la misma institución el año que viene.
- Estoy estudiando actualmente en una institución de educación superior, pero pienso cambiarme a otra carrera el la misma institución el año que viene.
- Estoy estudiando actualmente en una institución de educación superior, pero pienso cambiarme a otra institución el año que viene, continuado la misma carrera.
- Estoy estudiando actualmente en una institución de educación superior, pero pienso cambiarme a otra institución el año que viene, cambiando también de carrera.
- Empecé mis estudios de educación superior, pero tuve que dejarlos un tiempo y los retomaré nuevamente.

Siguiente

Para evitar problemas con tu encuesta, por favor NO oprimas el botón de regresar

Figure A.4: Question 2: Planned enrollment choices (requires at least one choice)



Proyecto 3E

Expectativas, Estudiantes, Educación

Pregunta #2

Imagina que hoy debes postular a las instituciones y a las carreras de educación superior que deseas. Por favor elige hasta 3 opciones de acuerdo a tus preferencias.

Si te encuentras actualmente estudiando en educación superior, indica también tus preferencias.

Instrucciones:

A continuación, selecciona de acuerdo a tus preferencias una opción de cada menú desplegable (Elección 1 corresponde a tu primera preferencia de postulación y Elección 3 a tu tercera preferencia). Si cambias de parecer en alguna "Elección", deberás empezar de nuevo modificando el Tipo de Institución.

Elección 1

Selecciona el tipo de institución

Selecciona el nivel

Selecciona el institución

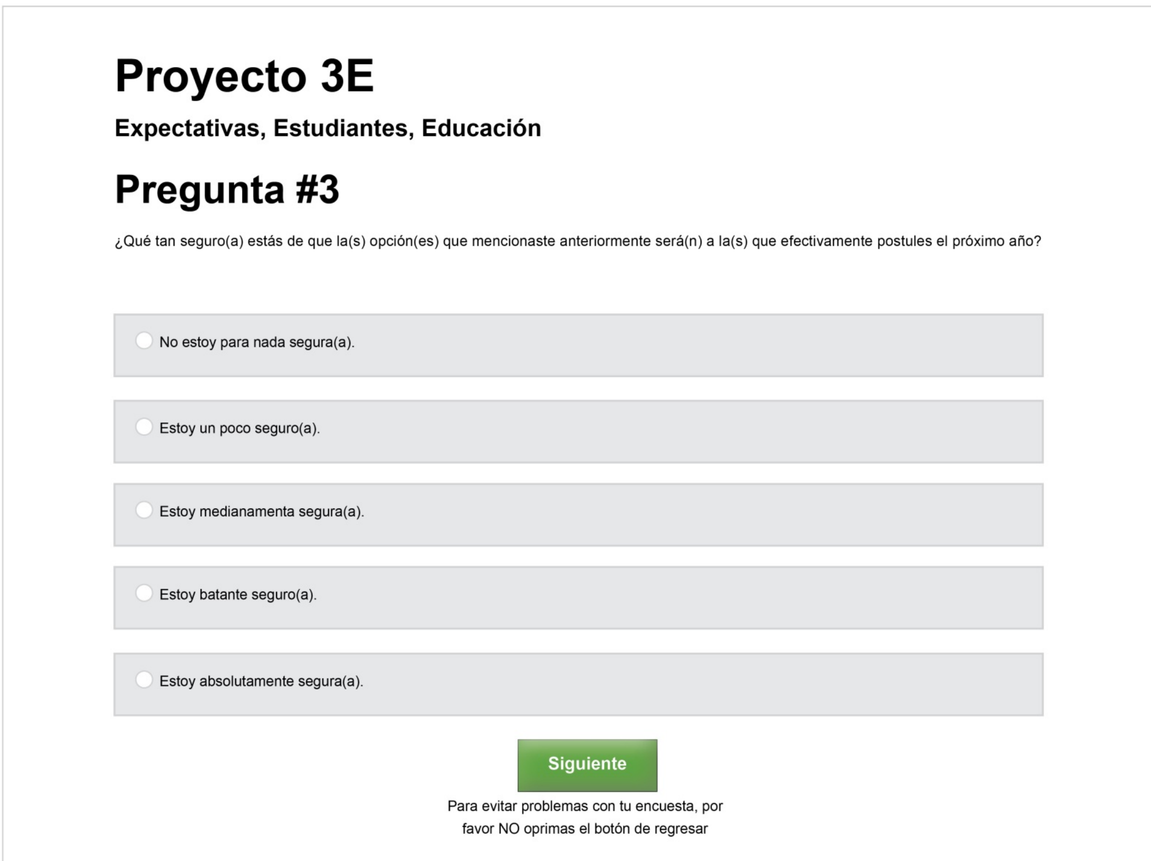
Los acentos en las letras mayúsculas fueron omitodos debido a incompatibilidad de programación.

Selecciona una carrera

Elección 2

Selecciona el tipo de institución

Figure A.5: Question 3: Certainty of enrollment plans



Proyecto 3E

Expectativas, Estudiantes, Educación

Pregunta #3

¿Qué tan seguro(a) estás de que la(s) opción(es) que mencionaste anteriormente será(n) a la(s) que efectivamente postules el próximo año?

- No estoy para nada segura(a).
- Estoy un poco seguro(a).
- Estoy medianamenta segura(a).
- Estoy batante seguro(a).
- Estoy absolutamente segura(a).

Siguiente

Para evitar problemas con tu encuesta, por favor NO oprimas el botón de regresar

Figure A.6: Question 4: Expected annual cost of tuition and fees at up to top three choices (piped in from question 2)

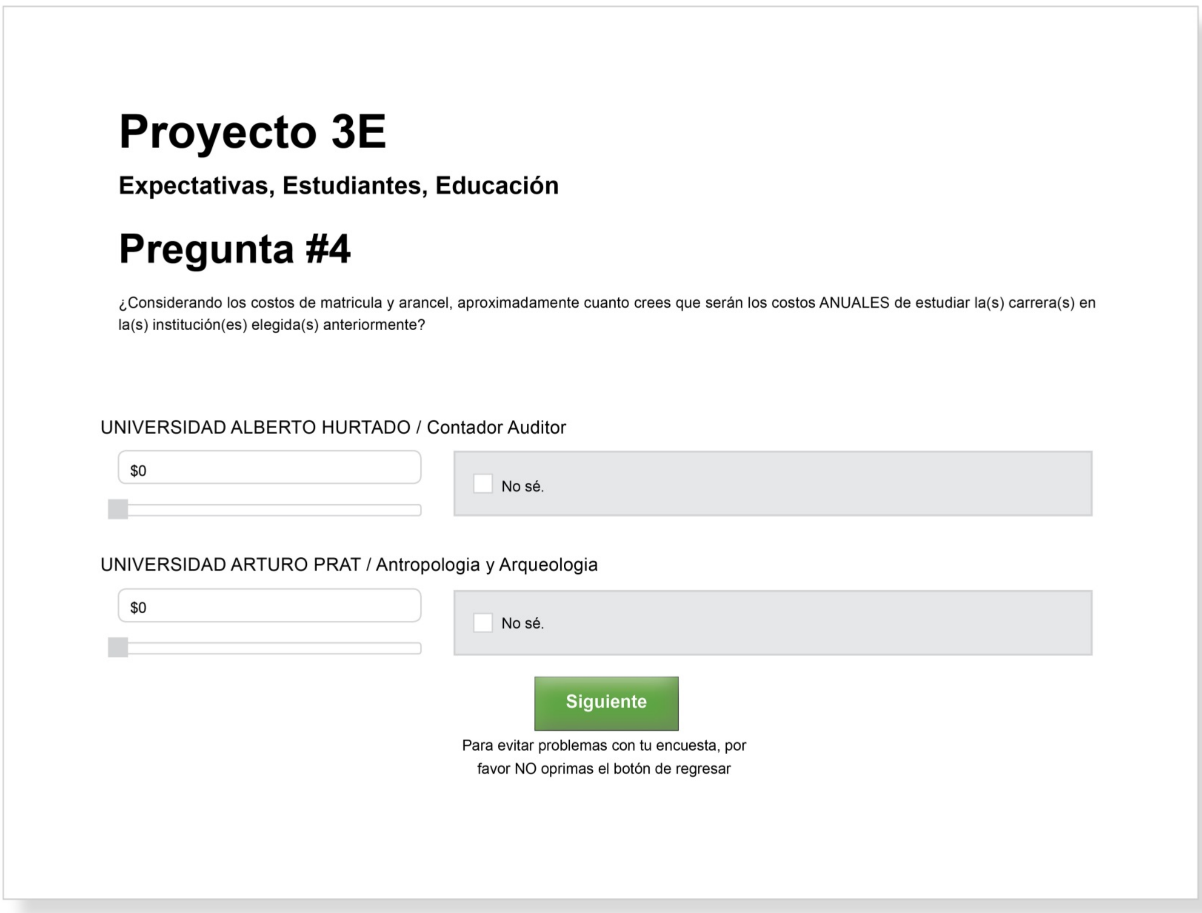
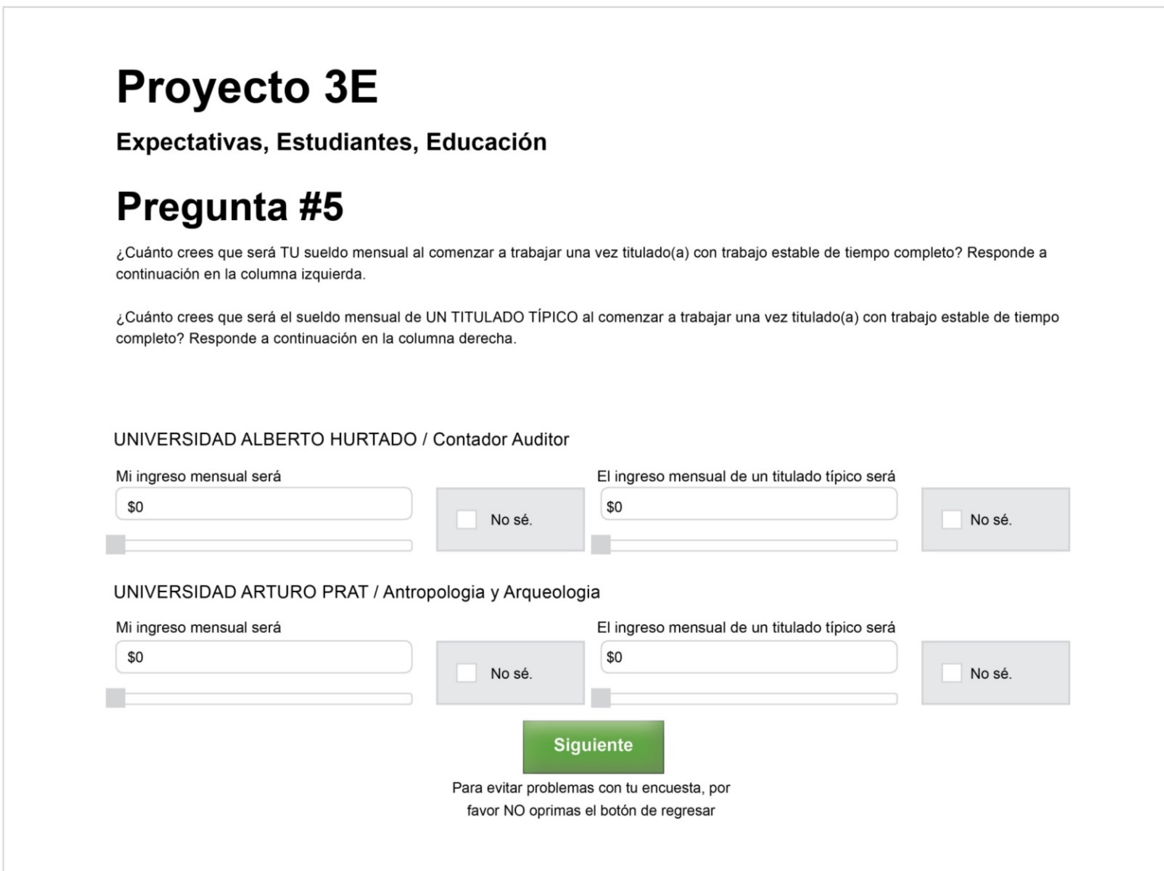


Figure A.7: Question 5: Expected Earnings at up to top three choices (piped in from question 2)



Proyecto 3E

Expectativas, Estudiantes, Educación

Pregunta #5

¿Cuánto crees que será TU sueldo mensual al comenzar a trabajar una vez titulado(a) con trabajo estable de tiempo completo? Responde a continuación en la columna izquierda.

¿Cuánto crees que será el sueldo mensual de UN TITULADO TÍPICO al comenzar a trabajar una vez titulado(a) con trabajo estable de tiempo completo? Responde a continuación en la columna derecha.

UNIVERSIDAD ALBERTO HURTADO / Contador Auditor

Mi ingreso mensual será $0 No sé. El ingreso mensual de un titulado típico será $0 No sé.

UNIVERSIDAD ARTURO PRAT / Antropologia y Arqueologia

Mi ingreso mensual será $0 No sé. El ingreso mensual de un titulado típico será $0 No sé.

Siguiente

Para evitar problemas con tu encuesta, por favor NO oprimas el botón de regresar

Figure A.8: Question 6: Expected PSU Scores in Math and Language

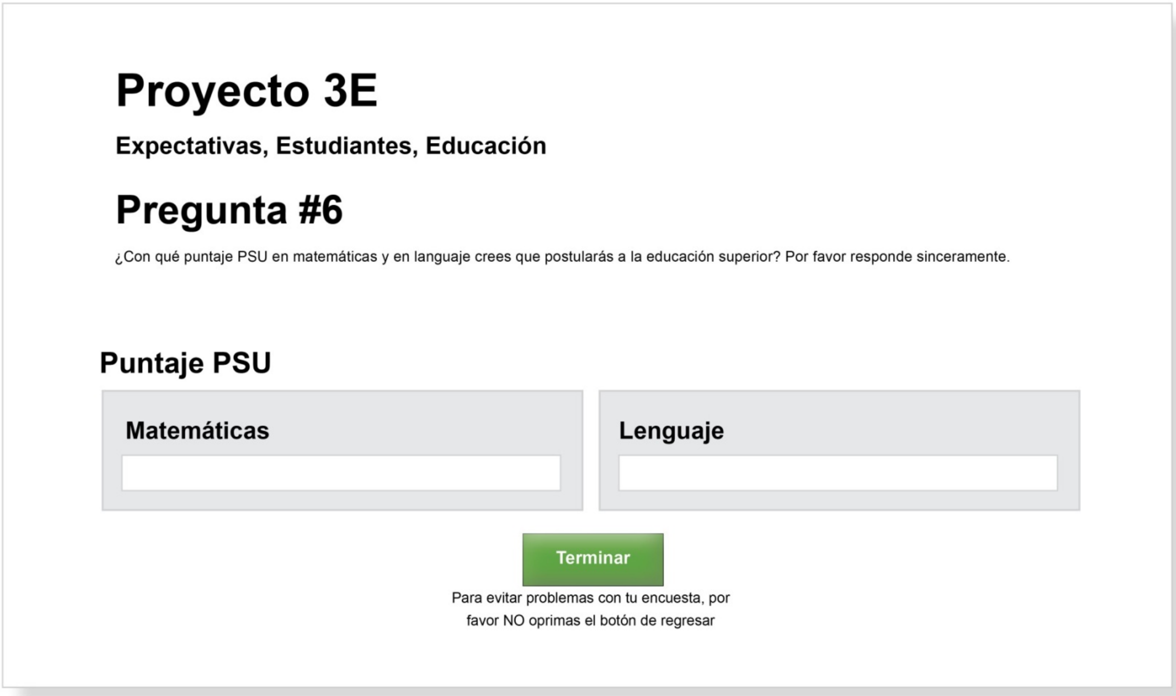
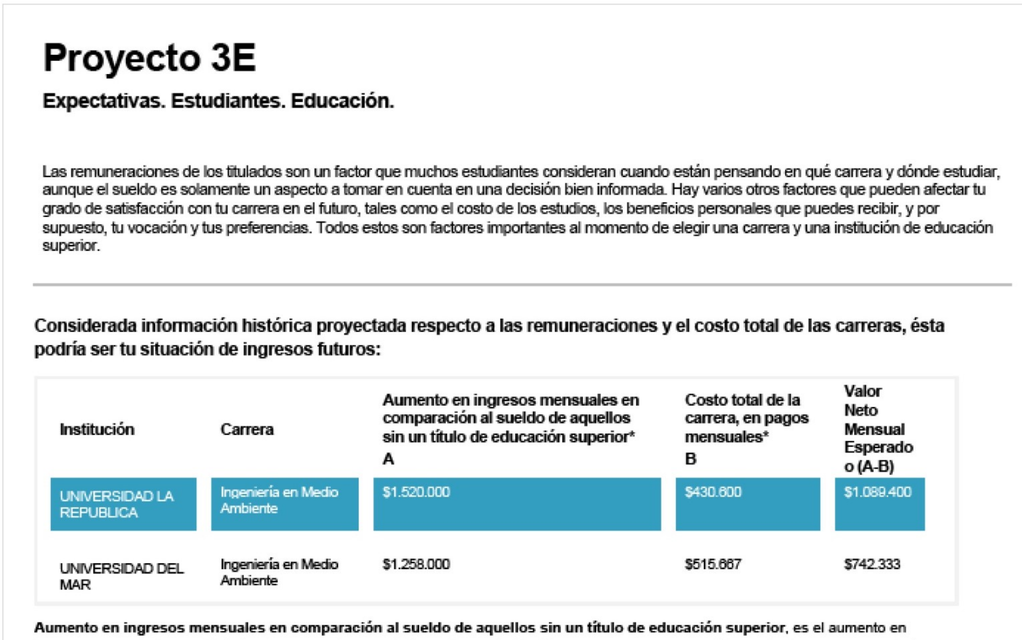


Figure A.9: Final Display Page for Control: Thank you page and confirmation code for control group

Figure A.10: Information and Suggestion Page for Treated Group, with alternate institution and career search nudges



Proyecto 3E

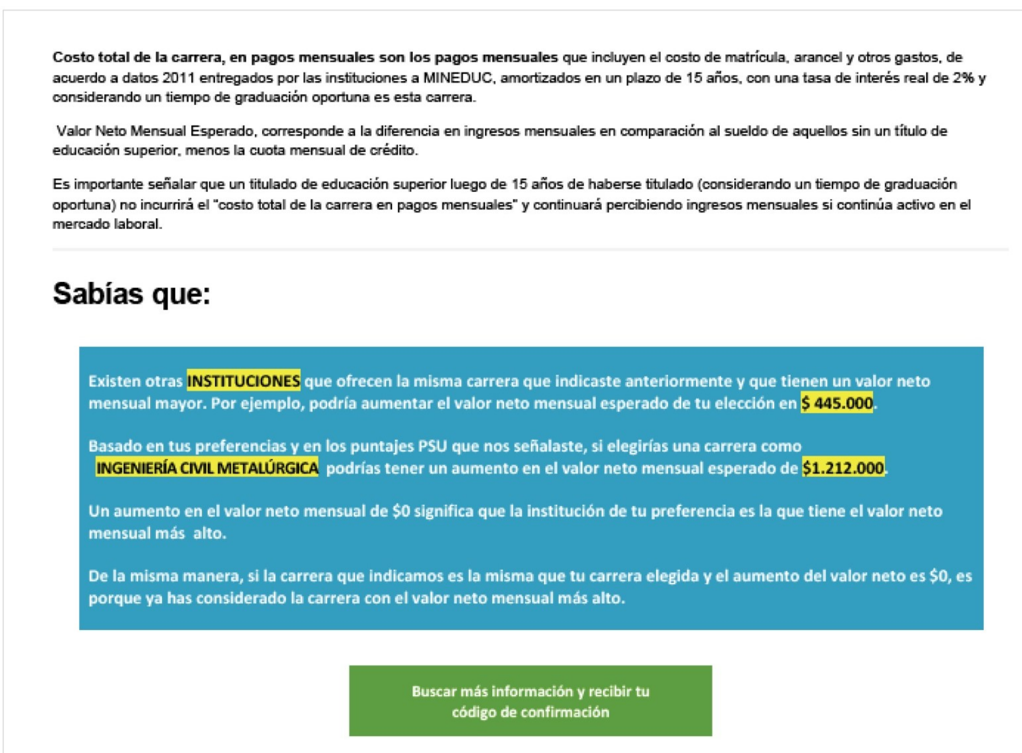
Expectativas. Estudiantes. Educación.

Las remuneraciones de los titulados son un factor que muchos estudiantes consideran cuando están pensando en qué carrera y dónde estudiar, aunque el sueldo es solamente un aspecto a tomar en cuenta en una decisión bien informada. Hay varios otros factores que pueden afectar tu grado de satisfacción con tu carrera en el futuro, tales como el costo de los estudios, los beneficios personales que puedes recibir, y por supuesto, tu vocación y tus preferencias. Todos estos son factores importantes al momento de elegir una carrera y una institución de educación superior.

Considerada información histórica proyectada respecto a las remuneraciones y el costo total de las carreras, ésta podría ser tu situación de ingresos futuros:

Institución	Carrera	Aumento en ingresos mensuales en comparación al sueldo de aquellos sin un título de educación superior* A	Costo total de la carrera, en pagos mensuales* B	Valor Neto Mensual Esperado o (A-B)
UNIVERSIDAD LA REPUBLICA	Ingeniería en Medio Ambiente	$1.520.000	$430.600	$1.089.400
UNIVERSIDAD DEL MAR	Ingeniería en Medio Ambiente	$1.258.000	$515.667	$742.333

Aumento en ingresos mensuales en comparación al sueldo de aquellos sin un título de educación superior, es el aumento en



Costo total de la carrera, en pagos mensuales son los pagos mensuales que incluyen el costo de matrícula, arancel y otros gastos, de acuerdo a datos 2011 entregados por las instituciones a MINEDUC, amortizados en un plazo de 15 años, con una tasa de interés real de 2% y considerando un tiempo de graduación oportuna es esta carrera.

Valor Neto Mensual Esperado, corresponde a la diferencia en ingresos mensuales en comparación al sueldo de aquellos sin un título de educación superior, menos la cuota mensual de crédito.

Es importante señalar que un titulado de educación superior luego de 15 años de haberse titulado (considerando un tiempo de graduación oportuna) no incurrirá el "costo total de la carrera en pagos mensuales" y continuará percibiendo ingresos mensuales si continúa activo en el mercado laboral.

Sabías que:

Existen otras INSTITUCIONES que ofrecen la misma carrera que indicaste anteriormente y que tienen un valor neto mensual mayor. Por ejemplo, podría aumentar el valor neto mensual esperado de tu elección en $ 445.000.

Basado en tus preferencias y en los puntajes PSU que nos señalaste, si elegirías una carrera como INGENIERÍA CIVIL METALÚRGICA podrías tener un aumento en el valor neto mensual esperado de $1.212.000.

Un aumento en el valor neto mensual de $0 significa que la institución de tu preferencia es la que tiene el valor neto mensual más alto.

De la misma manera, si la carrera que indicamos es la misma que tu carrera elegida y el aumento del valor neto es $0, es porque ya has considerado la carrera con el valor neto mensual más alto.

Buscar más información y recibir tu código de confirmación

Figure A.11: Final Page for Treatment Group: Searchable Database (Example searchable database continues onto next page)

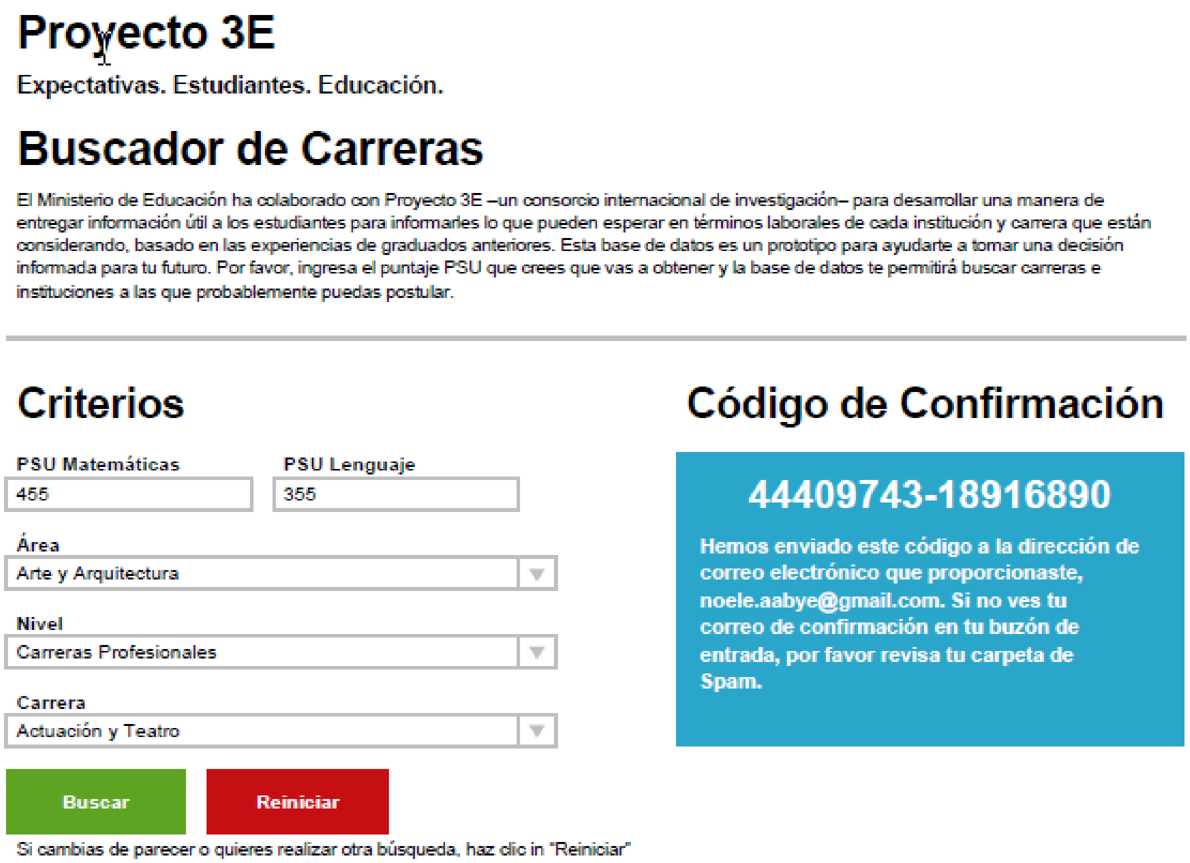
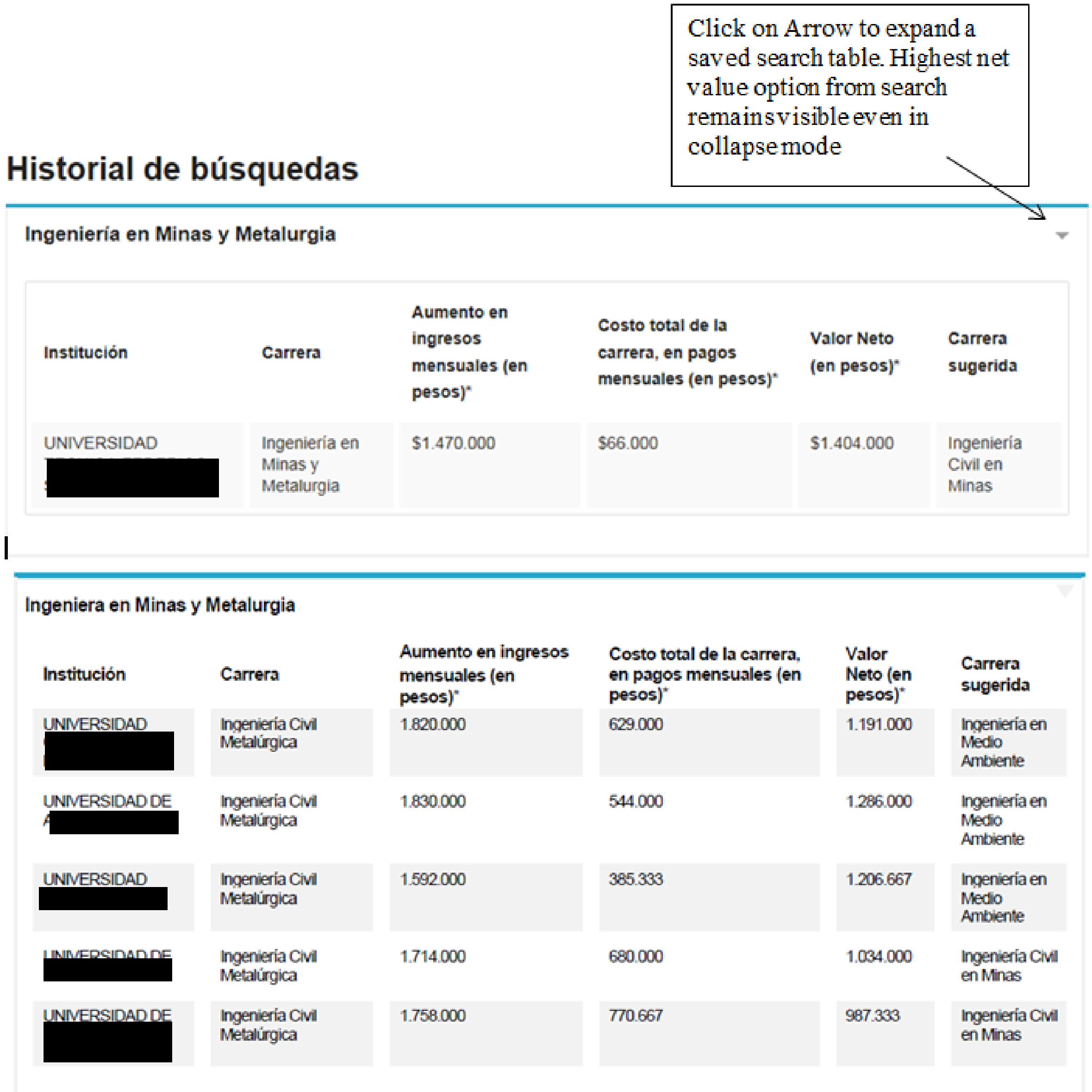


Figure A.12: Search Examples: Examples of Degrees Searched



Click on Arrow to expand a saved search table. Highest net value option from search remains visible even in collapse mode

Historial de búsquedas

Ingeniería en Minas y Metalurgia

Institución	Carrera	Aumento en ingresos mensuales (en pesos)*	Costo total de la carrera, en pagos mensuales (en pesos)*	Valor Neto (en pesos)*	Carrera sugerida
UNIVERSIDAD [redacted]	Ingeniería en Minas y Metalurgia	$1.470.000	$66.000	$1.404.000	Ingeniería Civil en Minas

Ingeniera en Minas y Metalurgia

Institución	Carrera	Aumento en ingresos mensuales (en pesos)*	Costo total de la carrera, en pagos mensuales (en pesos)*	Valor Neto (en pesos)*	Carrera sugerida
UNIVERSIDAD [redacted]	Ingeniería Civil Metalúrgica	1.820.000	629.000	1.191.000	Ingeniería en Medio Ambiente
UNIVERSIDAD DE [redacted]	Ingeniería Civil Metalúrgica	1.830.000	544.000	1.286.000	Ingeniería en Medio Ambiente
UNIVERSIDAD [redacted]	Ingeniería Civil Metalúrgica	1.592.000	385.333	1.206.667	Ingeniería en Medio Ambiente
[redacted]	Ingeniería Civil Metalúrgica	1.714.000	680.000	1.034.000	Ingeniería Civil en Minas
UNIVERSIDAD DE [redacted]	Ingeniería Civil Metalúrgica	1.758.000	770.667	987.333	Ingeniería Civil en Minas

Tus búsquedas serian guardadas para que las puedas ver cuando vuelvas a ingresar, así puedas usar esta información en el futuro. Esperamos que se te sirva la información que te estamos entregando. Es posible que contactemos para solicitar tus opiniones respecto a lala utilidad de esta información una vez que concluye el proceso de postulación.

C Earnings predictions conditional on enrollment

To calculate predicted earnings for current enrollees using the earnings of prior cohorts, we do the following. We use our linked database for freshman enrolling cohorts from 2000 through 2013, the baseline years before the student loan expansion, and link it to tax return data for the tax years of 2005 through 2012.

Let earnings in year t for individual i from entering cohort c who enrolled in institution-major combination (degree) j be a function of individual characteristics, the degree field of study and higher education institution they enrolled in and an idiosyncratic error term.

$$\begin{aligned} y_{ijct} &= X_{ict}\beta_{s(j)} + W_{it}\delta_{m(j)s(j)} + \tau_j + v_{ijct} \\ v_{ijct} &= \mu_{jc} + \epsilon_{ijct} \end{aligned} \tag{A.1}$$

X_{ict} includes dummies for student socioeconomic status, gender, and whether a student took the entrance exam. It also includes linear controls for entrance exam score, years of labor market experience, interactions between labor market experience and student covariates, and tax year dummies. Coefficients on the X_{ict} are permitted to vary over broad area of study[A.4] and five selectivity tiers. $s(j)$ denotes the interaction between broad area and selectivity tier. We allow the effects of gender, SES, test taking and test scores, and labor market experience, denoted W_{ijct}, to vary with major $m(j)$ as well. The τ_j are degree-specific fixed effects, and v_{ijct} is a mean-zero residual. We decompose v_{ijct} into a degree-cohort specific mean residual component μ_{jc} and an idiosyncratic error ϵ_{ijct}.

We estimate equation A.1 by degree and selectivity tier securely within the tax authority in Chile, taking out parameter estimates and freshman-cohort-level mean residuals needed to estimate μ_{jc}. Selectivity tier has five categories: no college enrollment, technical degrees, and three selectivity tiers for professional and university degrees defined by the median entrance exam scores of enrollees over our sample period 2000-2013. Thus the parameter estimates from equation A.1 for a nursing degree, for example, can vary by whether it is a technical nursing degree, an average nursing degree, or a very selective nursing degree.

To obtain the earnings predictions we use in the main text, we estimate these equations in data on the set of student-year observations in which students are between two and seven years removed from schooling completion (either graduation or dropout). We consider a variety of treatments for the mean residual term μ_{jc}. Our main approach is to take the mean of the $\hat{\mu}_{jc}$ across the 2000 through 2005 cohorts and apply this prediction to the 2013 entering cohort. We also consider shrinkage procedures in which we use past values of the $\hat{\mu}_{jc}$ to predict the 2013

[A.4] We use CINE-UNESCO field classifications, which break programs into Business, Agriculture, Art and Architecture, Basic Sciences, Social Sciences, Law, Education, Humanities, Health, and Technology.

value. We find little difference between these two approaches because sample sizes within each institution-degree-cohort cell are on average quite large and value added is stable over time. Using regression estimates and mean residuals, we compute the age 26 predicted earnings measure described in the main text.

In addition to point-in-time earnings at age 26, we are interested in estimating the expected earnings over T years for an individual i enrolling in degree j, $\hat{y}^T_{ijct}$. We use these longer-run estimates to evaluate the long-run effects of the disclosure intervention, as described in Section 5.4 and presented in Table 3. To scale earnings predictions to total earnings through age 30 or age 50, we add to each prediction an expected earnings growth rate. Specifically, we estimate the following equation separately by broad area and selectivity tier.

$$y_{ijct} = \tilde{\tau}_j + X_{it}\beta_{s(j)} + E_{it}\tilde{\gamma}_{s(j)} + \tilde{\epsilon}_{ijct} \tag{A.2}$$

The $\tilde{\tau}_j$ are degree-specific fixed effects, X_{it} includes tax year dummies and individual, time-invariant baseline characteristics (entrance exam score, gender and socio-economic status), and E_{it} is either a vector of labor-market experience year dummies or a linear term for years of experience, depending on the specification. The years of experience in the base data set span one to 12 years of experience. Using our predicted earnings from equation A.1 for the first 7 years of experience as a baseline, we project earnings past seven years linearly using the major-selectivity-tier specific growth rate estimated in equation A.2. We allow earnings to grow linearly through experience year 15 for degrees in the most selective category. For all other degrees, and for the no-college option, we project earnings through experience year 10. We choose these years because the coefficients on experience year dummies from equation A.2 appear linear up to year 10 of experience, at which point they flatten out for lower-selectivity degrees. Earnings continue to rise through at least year 12 for the most selective degrees. These patterns are consistent across areas of study. To obtain the PDV of earnings through some age T, we apply the 2% discount rate for state-backed student loans to our earnings predictions.

Table 3 compares the effects of treatment on earnings predictions to its effects on cost predictions. We calculate costs as the total tuition plus any one-time matriculation fees for enrolling in a degree. We use matriculation fee and tuition data from 2010 (adjusted for inflation) and compute the expected total payment evaluated at the time of enrollment. We calculate expected time of enrollment in degree by taking the average time enrolled in the degree across all of the 2000-2005 first-year cohorts.[A.5] We cap the maximum expected time in degree at 6.5 years.

[A.5] These cohorts are used to construct expected time to completion because these cohorts will have had enough time to graduate in the longer degrees, and will thus include both students who graduated from the degree and students who dropped out of the degree.

D Robustness in beliefs description

D.1 Measurement error in earnings beliefs

In Section 5.1 we show that log beliefs about earnings for typical graduates have a slope of 0.43 in log observed earnings outcomes for past students. We argue that this is consistent with a Bayesian updating model in which students are uncertain about outcomes for past students and shrink beliefs back towards the sample mean. A possible alternate explanation is that our measurements of observed values are noisy and produce downward bias in estimates of the slope of beliefs in earnings. In this section we show that correcting our observed earnings outcomes for measurement has trivial effects on the relationship between beliefs and observed values. We also show that the slope of beliefs in observed values is nearly identical when computed in levels as opposed to logs.

We correct for measurement error as follows. Observed earnings means are given by

$$Y_j = \mu_j + \bar{\epsilon}_j$$

where μ_j is the true underlying mean and $\bar{\epsilon}_j = N_j^{-1} \sum_i \epsilon_{ij}$ is the mean of mean-zero idiosyncratic earnings shocks ϵ_{ij} for individuals graduating from the degree program. The ϵ_{ij} are iid and have variance σ_ϵ^2 while the μ_j have variance σ_μ^2 and mean μ. A precision-weighted expectation of μ_j after observing Y_j and μ is given by

$$\hat{\mu}_j = \frac{\sigma_{\bar{\epsilon}}^{-2}}{\sigma_{\bar{\epsilon}}^{-2} + \sigma_\mu^{-2}} Y_j + \left(1 - \frac{\sigma_{\bar{\epsilon}}^{-2}}{\sigma_{\bar{\epsilon}}^{-2} + \sigma_\mu^{-2}}\right) \mu$$

with $\sigma_{\bar{\epsilon}}^2 = \sigma_\epsilon^2 / N_j$. If ϵ_{ij} and μ_j are normally distributed this is the Bayesian posterior expectation.

True mean μ and variance terms σ_ϵ^2 and σ_μ^2 are not observed. We therefore create sample analogs using data on degree mean earnings and within-degree earnings variances. In levels specifications, we take the sample mean observed past earnings, $\bar{Y}$, as an estimate of μ . We take the within-j variance of earnings averaged across degrees j as our estimate of σ_ϵ^2 . Defining the total variance of observed degree-level earnings σ_Y^2, we compute $\hat{\sigma}_\mu^2 = \hat{\sigma}_Y^2 - \hat{\sigma}_{\bar{\epsilon}}^2$ as our estimate of σ_μ^2, using sample analogues of the latter two terms.

Log specifications (the focus of the main text) require additional analysis because we do not have access to estimates of within-degree variance of log earnings. Our data extract from the Chilean tax authority included only means and variances by degree in level terms. Sampling error in levels induces both sampling error and a mean shift in logs. Specifically, let $y_j = \log(Y_j) = \log(\mu_j + \bar{\epsilon}_j)$. Taking a Taylor expansion around μ_j , we have $E[y_j] = \log(\mu_j) - \frac{\sigma_{\bar{\epsilon}}^2}{2\mu_j^2}$ and $Var(y_j) = \frac{\sigma_{\bar{\epsilon}}^2}{\mu_j^2}$. An assumption here is that the Y_j are bounded away from zero. This assumption is satisfied in our data. To address sampling error in log specifications we compute corrected degree-specific mean log earnings as $y_j^c = \log(Y_j) + \frac{\hat{\sigma}_{\bar{\epsilon}}^2}{2\hat{\mu}_j^2}$ using the sample analogues described above. We take $\frac{\hat{\sigma}_{\bar{\epsilon}}^2}{\hat{\mu}_j^2}$ as our estimate of sampling variance, and compute shrinkage estimates using a procedure analogous to the levels case above.

Results indicate that the variance of measurement error is small relative to the variance of mean outcomes, so sampling error has little effect on our estimates of the relationship between beliefs and means. As reported in Table A.10, the mean earnings calculation in our beliefs dataset uses 491 student-year earnings observations. Means are precisely estimated relative to the dispersion of outcomes across programs. The square root of the mean sampling variance of log earnings (earnings levels) is 0.066 (53,000), relative to a cross degree standard deviation of 0.59 (459,000).

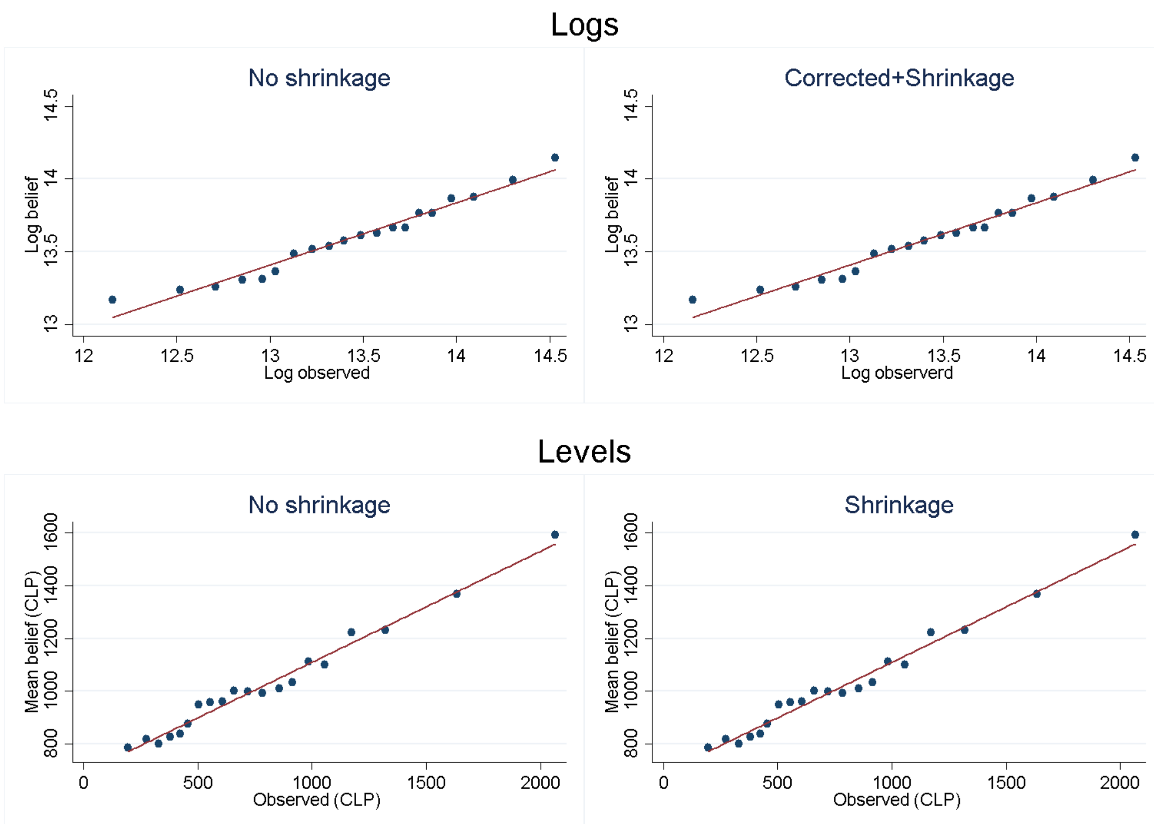
Table A.10: Measurement error in observed earnings values

	Logs	Levels
Statistics		
Obs per. Degree	491	491
Effect SD σ_μ	0.590	459
Sampling SD $\sigma_{\bar{\epsilon}}$	0.066	53
Slopes		
Uncorrected	0.427	0.421
	(0.005)	(0.008)
Corrected	0.437	0.439
	(0.005)	(0.008)

'Statistics' panel: 'Obs per degree' is the mean count of observations used to compute earnings values for degree programs for which respondents report beliefs. Effect SD σ_μ is the square root of estimated cross-program variance in mean earnings. Sampling SD $\sigma_{\bar{\epsilon}}$ is the square root of the mean of sampling error across programs. Levels column is measured in 1000s of CLP. 'Slopes' panel: slope estimates from regressions of earnings beliefs on observed earnings. 'Uncorrected' row uses reported beliefs and observed earnings. 'Corrected' row shrinks measures of observed earnings towards the sample mean as described in the text. In the log specification this row also applies the bias correction described in the text.

Because correction terms are small, regression results that apply shrinkage and bias corrections are very similar to the baseline specifications. We report these results in the lower panel of Table A.10. Slopes in belief levels are nearly identical to slopes in logs. Figure A.13 shows that the linear relationship between mean beliefs and mean observed earnings holds in levels and logs specifications with and without corrections for measurement error.

Figure A.13: Beliefs vs. observed earnings



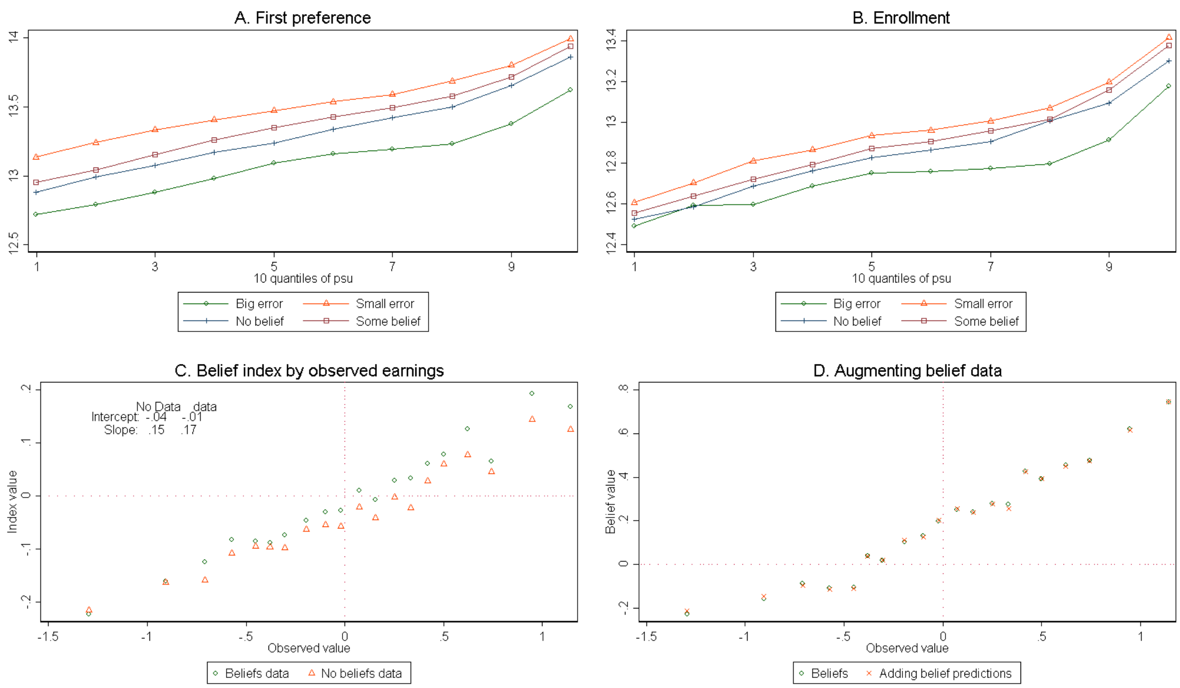
Means of earnings beliefs by ventile of observed earnings outcomes in logs (upper panel) and levels (lower panel). 'Corrected+shrinkage' panel in log specification shrinks observed earnings values towards sample mean and applies bias correction as described in text. 'Shrinkage' panel in levels specification shrinks observed earnings values towards the sample mean.

D.2 Belief nonreports

This section examines how belief nonreports (i.e., students choosing the "I don't know" option on their belief survey) affect our analysis of belief errors for students who do report. The key point is that, conditional on elicited preferences, students who do not report beliefs are similar to students who do in terms of their observable characteristics and enrollment outcomes. Including beliefs for non-reports would likely flatten the slope of beliefs in earnings.

Panels A and B of Figure A.14 show that conditional on PSU score, students who do not report beliefs have preferences and enrollment choices that are fairly similar to those for students who do report beliefs in terms of predicted earnings. The gap in log earnings at the listed first preference for students with big and small belief errors, as defined in section 5.3, is equal to 35 log points. In contrast, students who do report beliefs enroll in degree programs where predicted earnings are only 7 log points higher than students who do not report beliefs. Students with small belief errors enroll in degree programs where their average earnings are 21 log points higher than students with large belief errors, while the gap in predicted earnings by belief non-report status is only 4 log points. Beliefs are an input into the processes of preference formation and enrollment choice. That reported preferences and choices are similar for reporters and non-reporters conditional on admissions score suggests that beliefs may be similar as well.

Figure A.14: Assessing belief nonreports



Panel A: Vertical axis is mean observed log earnings at most-preferred degree program by listed subgroup. 'Big error' and 'small error' refer to earnings belief errors, as described in Section 5.3. 'Some belief' and 'No belief' groups are students who do and do not report earnings beliefs at their most-preferred program, respectively. Horizontal axis: deciles of PSU admissions score. Panel B: As in panel A, but the vertical axis observed earnings outcomes at the degree programs in which students enroll. Panel C: Predicted log earnings beliefs (vertical axis) by ventile of observed log earnings value (horizontal axis). 'Beliefs data' is sample of students who report beliefs, 'No beliefs data' is sample of students who do not. Values normalized relative to mean of observed distribution. Predictions reflect linear index based on student observables and enrollment behavior. See text for details. Panel D: Mean log belief (vertical axis) by ventile of observed log earnings distribution (horizontal axis). 'Beliefs' sample is students for whom we observe beliefs data. 'Adding belief predictions' fills in data for non-reporters using predicted belief values from Panel C. Horizontal and vertical axes normalized so observed value mean is zero.

Panels C and D of Figure A.14 quantify the relationship between belief nonreports and observable predictors of earnings beliefs. Working within the sample of students who do report beliefs, we regress beliefs on student covariates separately within each ventile of the distribution of observed earnings outcomes at the first choice degree. The covariates we use are gender, SES, a quadratic in PSU score, and log predicted earnings at the degree program in which a student enrolls. We use the regression results to predict beliefs in the sample of reporters and non-reporters. Panel C of Figure A.14 plots the predictions for each sample by ventile of the observed earnings distribution. Differences in predicted beliefs by report status are small. The slope of predicted values is slightly steeper in the sample of belief reporters (0.17) than non-reporters (0.15), but this difference is not statistically significant.

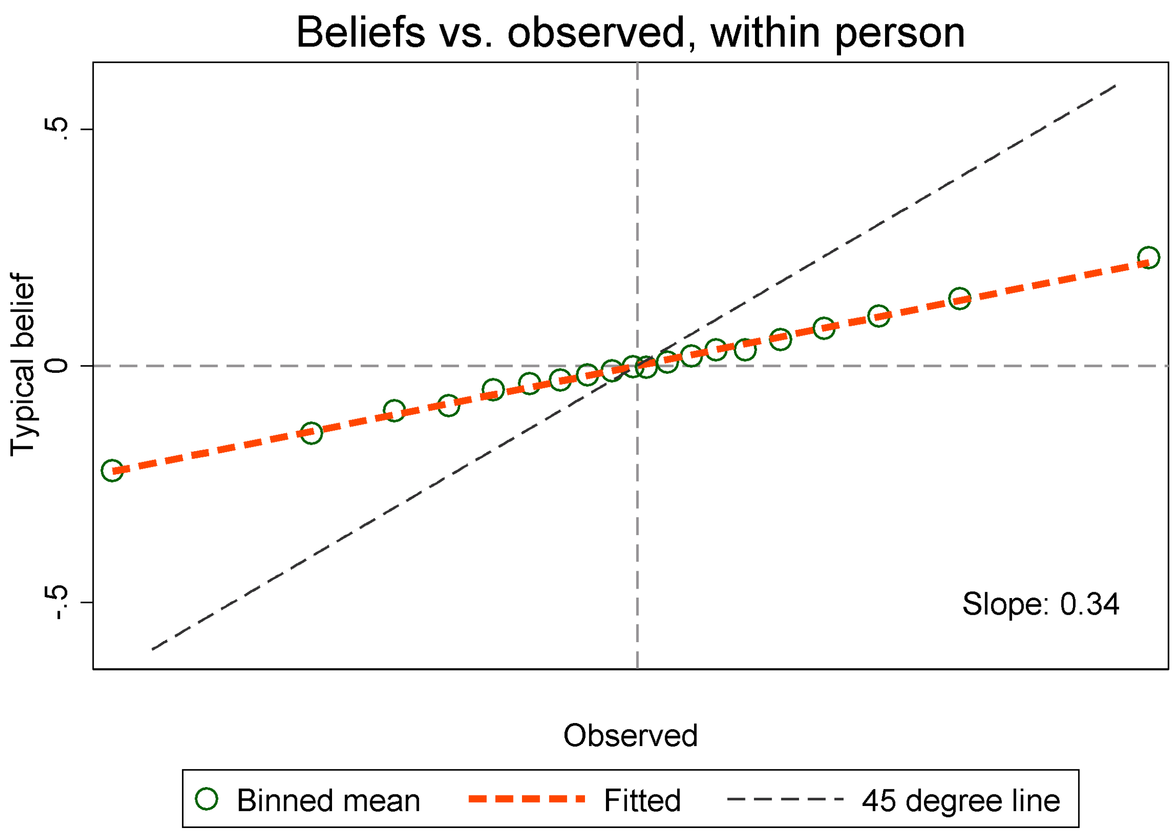
Panel D of Figure A.14 reports mean beliefs for each ventile of the observed earnings distribution for belief reporters, and for the full sample of students. In the full sample, beliefs for non-reporters are imputed using the predictions from Panel C. Differences between the two plots are minimal. The slope of beliefs in the sample that includes imputations is 0.414, compared to 0.427 in the sample of observed data. Our findings here suggest that even though rates of belief nonreport are high, average beliefs among nonreporters are likely fairly similar to those among reporters conditional on preferences, and therefore will not substantively alter the conclusion of Section 5.1. To the extent nonreports do affect our findings, they seem likely to slightly decrease the slope of beliefs in observed values.

D.3 Within-person belief errors

Consistent with the Bayesian belief formation framework that we present in section 5.1, our descriptive finding that the slope of beliefs in observed values is less than one holds in a within-person analysis as well as in the cross section. Our model of belief formation is based on the idea that students have some mean belief about the returns to college overall and then receive noisy signals about earnings at particular degree programs. However, the cross sectional analysis we present in section 5.1 is also consistent with an alternate model in which belief errors are constant across degree programs for each student. The relationship we observe could then arise if the lower-skilled students expressing interest in low-earning degrees are systematically optimistic about all degree programs, while the higher-skilled students expressing interest in high-earning degrees are pessimistic about all programs.

A within-person analysis of the relationship between beliefs and observed earnings values for past students shows that this is not the case. This analysis amounts to a regression of beliefs about earnings for the typical student on observed earnings outcomes that includes individual fixed effects. This regression is feasible because our survey elicits beliefs about multiple programs from each student. We present results from this regression in Figure A.15. This graph plots deviations from within-student mean observed earnings against deviations from within-student mean beliefs. As in the cross-sectional graph, we observe a linear relationship with a slope of less than one. This is consistent with the idea that belief errors are idiosyncratic across degree programs and inconsistent with a model in which beliefs rise one-for-one with true values after accounting for student-specific bias. The within-person slope is 0.34, slightly smaller than the slope of 0.43 we observe in the cross-section. This slope corresponds to a noise-to-signal ratio of 1.94, compared to 1.34 in the cross-sectional estimation.

Figure A.15: Within person earnings beliefs by observed values



Horizontal axis: deviations of observed earnings values from within-individual mean. Vertical axis: deviations of beliefs about earnings for typical graduate from within-person mean. Points are binned means of beliefs within ventiles of distribution of observed values. 45 degree line plotted for reference. Sample is the set of surveyed individuals reporting earnings beliefs for multiple degree programs.

E Validating OLS earnings predictions using regression discontinuities

E.1 Overview

This section describes the estimation strategy and empirical implementation of the regression discontinuity benchmarking exercise for which we report results in Section 7.1. We also present additional regression tables that show the robustness of our findings to alternate prediction approaches. The goal of the benchmarking exercise is to test the mapping between regression predictions of earnings outcomes and causal effects for particular students or groups of students. Broadly, our approach is to compare observed earnings effects for students shifted quasi-randomly from one degree program to another to the expected change based on our predictions.

E.2 Econometric model

Suppressing variation across admissions cohorts and outcome years for brevity, let Y_{ij} denote the potential earnings outcome for individual i enrolling in degree program j. Outcomes are determined by program mean effects θ_j, individual skill levels τ_i, and degree match effects ϕ_{ij}, so that

$$Y_{ij} = \theta_j + \mu_i + \phi_{ij}$$

Let $Y^*_{ij} = E[Y_{ij}|X_{ij}]$ be our prediction of earnings for an individual with i's characteristics enrolling at j. We may then rewrite potential outcomes as the sum of the predicted and unpredicted components:

$$Y_{ij} = Y^*_{ij} + v_{ij}$$

where $v_{ij} = (\theta_j - E[\theta_j|X_{ij}]) + (\tau_i - E[\tau_i|X_{ij}]) + (\phi_{ij} - E[\phi_{ij}|X_{ij}])$. Prediction errors may arise from a number of sources, including the failure of the regression prediction to adequately account for selection into degree programs on the basis of absolute advantage τ_i or on individual-specific match effects ϕ_{ij}.

In our admissions data, we observe students applying to some program j. Applications are evaluated based on an index of grades and test scores, and students are admitted to j if they cross an admissions threshold. Marginally rejected students attend a mix of next-choice degree programs. HNZ describe the setting in the detail and show that the conditions for regression discontinuity estimation are satisfied.

Let $\Delta Y_{ijn(i,j)} = Y_{ij} - Y_{in(i,j)}$ be the difference between earnings at some target degree program j and i's next option $n(i,j)$. $n(i,j)$ is the degree to which i would be admitted if he were rejected in his application to j. Let $\Delta Y_j = E[\Delta Y_{ijn(i,j)}|i \in M_j]$ be the expected change in earnings from admission to j for marginal applicants to j. M_j denotes the set of marginal applicants to degree program j. This is what we observe (in expectation) in a regression discontinuity estimate. Note that $\Delta Y_{ijn(i,j)} = \theta_j - \theta_n + \phi_{ij} - \phi_{in}$. Individual absolute advantage τ_i is differenced out across the threshold, so these values reflect the causal effects of admission to j on earnings for marginal students.

Our benchmarking exercise compares the observed cross-threshold-changes to predicted cross-threshold changes. We now explain how to interpret results from this exercise. Decompose the overall change in earnings into a predicted component and an unpredicted component. Specifically, $\Delta Y_{ijn(i,j)} = \Delta Y^*_{ijn(i,j)} + \Delta v_{ijn(i,j)}$, where $\Delta Y^*_{ijn(i,j)} = Y^*_{ij} - Y^*_{in}$ and $\Delta v_{ijn(i,j)} = v_{ij} - v_{in}$. Taking expectations over marginal applicants to j, we may write

$$\Delta Y_j = \Delta Y^*_j + \Delta v_j$$

where $\Delta Y^*_j = E[\Delta Y^*_{ijn(i,j)}|i \in M_i]$ and $\Delta v_j = E[\Delta v_{ijn(i,j)}|i \in M_i]$. The ΔY^*_j are what one would recover in expectation in a regression discontinuity analysis for degree j that takes predicted earnings as the outcome of interest.

Intuitively, predicted cross-threshold earnings changes will differ from observed values in expectation when the predicted values are on average incorrect for the subset of marginal applicants. To see this consider a regression of the form

$$\Delta Y_j = \tilde{\lambda}_0 + \tilde{\lambda}_1 \Delta Y^*_j + \tilde{e}_j$$

The expectation of the slope term is given by

$$E[\tilde{\lambda}_1] = \frac{Cov(\Delta Y_j, \Delta Y^*_j)}{Var(\Delta Y^*_j)} = \frac{Cov(\Delta Y^*_j + \Delta v_j, \Delta Y^*_j)}{Var(\Delta Y^*_j)} = 1 + \frac{Cov(\Delta v_j, \Delta Y^*_j)}{Var(\Delta Y^*_j)}$$

$$E[\tilde{\lambda}_0] = E[\Delta Y_j] - \tilde{\lambda}_1 E[\Delta Y^*_j] = E[\Delta Y_j](1 - \tilde{\lambda}_1) + \tilde{\lambda}_1 E[\Delta v_j]$$

If our earnings predictions accurately capture most of the variation in earnings outcomes that is known to students at the time of choice, $Cov(\Delta v_j, \Delta Y^*_j)$ and $E[\Delta v_j]$ will be small, so $\tilde{\lambda}_1$ will be close to one and $\tilde{\lambda}_0$ close to zero.

Alternatively, if our predictions are systematically wrong for some types of degree programs or for some types of applicants, bias terms may be large. For example, consider an extreme case where there are no degree effects: $\theta_j = \phi_{ij} = 0, \forall i, j$. Differences in predicted values across degree programs emerge only from our failure to control adequately for selection on individual absolute advantage, τ_i. In this setting, the prediction error becomes more negative one-for-one as the predicted value grows larger.

$$\frac{Cov(\Delta v_j, \Delta Y_j^*)}{Var(\Delta Y_j^*)} = -1, \qquad \text{and} E[\tilde{\lambda}_1] = 0$$

.

In practice, the regression using ΔY_j and ΔY_j^* is not feasible because we do not observe these values. Instead, we observe noisy estimates $\hat{\Delta} Y_j$ and $\hat{\Delta} Y_j^*$, respectively. The sampling error associated with estimates may be substantial even if our earnings predictions are themselves precisely estimated. This is because sampling error also arises from the discontinuity design, which relies exclusively on small samples of students close to admissions cutoffs.

To address this, we take a Bayesian approach and shrink noisily estimated values of $\hat{\Delta} Y_j^*$ back towards the sample mean. This follows a literature on teacher effects and school effects (Kane and Staiger 2008; Chetty et al. 2014a,b; Deming 2014). It provides a measurement-error adjusted estimate of ΔY_j^* for each degree program j, which allows for visual inspection of the relationship between predicted and observed effects that accounts for measurement error. It also permits heterogeneity in measurement error across degree programs. This is realistic in a setting in which counts of marginal applicants vary across programs.

Assume that predicted earnings gains ΔY_j^* are drawn from a normal distribution with mean $\bar{\Delta} Y^*$ and variance $\sigma^2_{Y^*}$. We observe an unbiased but noisy signal $\hat{\Delta} Y_j^*$ about the value of each ΔY_j^*; the variance of the normally distributed error term is σ_j^2, with σ_j^2 independent of ΔY_j^* and Δv_j. In this setting, the posterior expectation of ΔY_j^* is given by $\tilde{\Delta} Y_j^* = \alpha_j \hat{\Delta} Y_j^* + (1 - \alpha_j)\bar{Y}^*$, with $\alpha_j = \frac{\sigma_Y^2}{\sigma_Y^2 + \sigma_j^2}$. This a standard precision-weighting formula. Noisier estimates are weighted more towards the sample mean. We then estimate specifications of the form

$$\hat{\Delta} Y_j = \lambda_0 + \lambda_1 \tilde{\Delta} Y_j^* + \tilde{e}_j$$

taking λ_0 and λ_1 as the coefficients of interest.

There is also sampling error in the outcome variable. To address this, we weight each observation by the inverse of the estimated sampling variance.

E.3 Implementation

We implement these specifications as follows. For cohorts of applicants beginning in 1982 and ending in 2005, we have access to regression discontinuity estimates of the earnings effects associated with crossing admissions thresholds at each selective CRUCH degree program. These estimates come from HNZ, which estimates the change in earnings attributable to crossing the threshold for admission to each target degree within the CRUCH system.

We first reconstruct the HNZ applicant database, and merge on enrollment data from our enrollment panel. Based on enrollment outcomes, we predict earnings in tax years 2005-2013, the tax years used to estimate the regression discontinuity earnings gains estimates in HNZ, for each applicant in the sample. We construct predictions using the estimates from equation 2 in the main text. We focus on students who are a minimum of seven years past the year of application, or approximately age 25. We then use our earnings predictions for the sample of marginal CRUCH applicants to estimate regression discontinuity specifications that use the OLS earnings predictions as the dependent variable, rather than the realized earnings used in HNZ. As in HNZ, we use standard regression discontinuity specifications of the form

$$Y_{ij} = f_j(d_{ij}) + \Delta_j Z_{ij} + \epsilon_{ij} \tag{A.3}$$

where Y_{ij} here is the average of predicted earnings values six or more years after application for individual i applying to degree j. This exercise parallels equation (1) in HNZ, but with predicted earnings as the outcome of interest. The running variable, d_{ij}, is the difference between the admissions score assigned to i's application to program p and the cutoff score for admission to that program in the year i applies, $f_j(d_{ij})$ is a smooth function of the score difference (which can change on either side of the cutoff), $Z_{ij} = 1(d_{ij} \geq 0)$ is an indicator variable equal to one if i's application to degree p is above the cutoff score (i.e., a dummy for acceptance to j), and ϵ_{ij} is an error term. We estimate equation A.3 separately for every degree in the system using data within a score window of 50 points around the cutoff score.

These estimates provide the $\hat{\Delta}Y_j^*$ we use for the analysis described in section 7.1. We compare the predicted RD estimates we obtain here to two sets of RD estimates $\hat{\Delta}Y_j$ obtained using observed earnings data from HNZ. The first consists of marginal applicants for the years 2000 through 2005. The second focuses on older cohorts of students who applied to college between 1982 and 1993. To the extent that we fail to reject the null of no bias in the 2000-2005 applicant sample, we will interpret our findings as evidence that OLS predictions are at least on average reasonable predictors of causal effects for marginal CRUCH applicants. If we also fail to reject these claims for the 1982 through 1993 sample, we will interpret our findings as suggestive evidence that degree effects may be relatively stable over time.

One additional empirical question is how to make earnings predictions for students who are enrolled in school at the time of observation. This is important to consider because our predicted measures focus on early career outcomes, and students who take a long time to complete school are included in outcome regressions for recent cohorts. Our main specifications take these values as missing when computing predictions. We also consider alternative approaches that set earnings for students enrolled in college to different fractions of the annual minimum wage when computing predicted earnings. Our findings are not sensitive to this choice: we obtain similar findings to our main results when we set earnings predictions for college enrollees to 20%, 50%, or 100% of the annual minimum. Here, we report findings from the 100% specification.

E.4 Results

Table A.11 reports our findings. The upper panel reports estimates obtained using the shrinkage procedure. We cannot reject a slope of one or an intercept of zero in any specification. The '2000-2005, pre-2000 predictions' and '1982-1993' columns replicate results from Table 6 in the main text. The remaining columns show that changing the set of cohorts used to make earnings predictions and the treatment of earnings for students still enrolled in school does not affect our outcomes. With no correction for measurement error, slope estimates are smaller, as expected. We present these results in the lower panel. We interpret these findings as providing support for the validity of our earnings predictions over the short run, and also as suggesting that degree effects are relatively stable across application cohorts.

Table A.11: Observed vs. predicted earnings changes

	2000-2005 all prediction years	2000-2005 pre-2000 predictions	1982-1993	2000-2005 +MW	1982-1993 +MW
Corrected					
Slope	0.761	0.716	0.778	1.153	0.772
	[0.417,1.104]	[0.332,1.100]	[0.415,1.142]	[0.707,1.598]	[0.218,1.326]
Intercept	-18.85	28.90	149.9	-111.2	176.0
	[-188.3,150.6]	[-129.6,187.4]	[-95.82,395.6]	[-301.0,78.56]	[-76.22,428.3]
Uncorrected					
Slope	0.529	0.466	0.180	0.638	0.124
	[0.371,0.686]	[0.324,0.608]	[0.0895,0.271]	[0.474,0.801]	[0.0136,0.234]
Intercept	71.95	134.7	341.9	42.68	351.7
	[-86.77,230.7]	[-26.77,296.2]	[116.6,567.2]	[-115.2,200.5]	[124.9,578.5]
N cutoffs	501	500	501	501	501

Point estimates and 95% CIs from regressions of RD estimates of observed earnings changes on RD estimates of predicted earnings changes as described in equation 6. Observations are at degree program level with results weighted by inverse sampling variance of dependent variable. "Corrected" panel reports results that account for sampling error in the independent variable by shrinking values towards the sample mean. "Uncorrected" panel reports results that do not account for sampling error. '2000-2005, all prediction years' column takes outcomes for 2000-2005 entering cohorts as those of interest, and computes predicted earnings using all entering cohorts. '2000-2005, pre-2000 predictions' is the same but uses only pre-2000 entering cohorts to make earnings predictions. '1982-1993' column uses outcomes for 1982-1993 cohorts and predictions for all entering cohorts. The '+MW' columns repeat the analyses from columns 1 and 3 but assign students still enrolled in school the minimum annual wage rather than a missing value. These values are treated as missing in the left three columns.