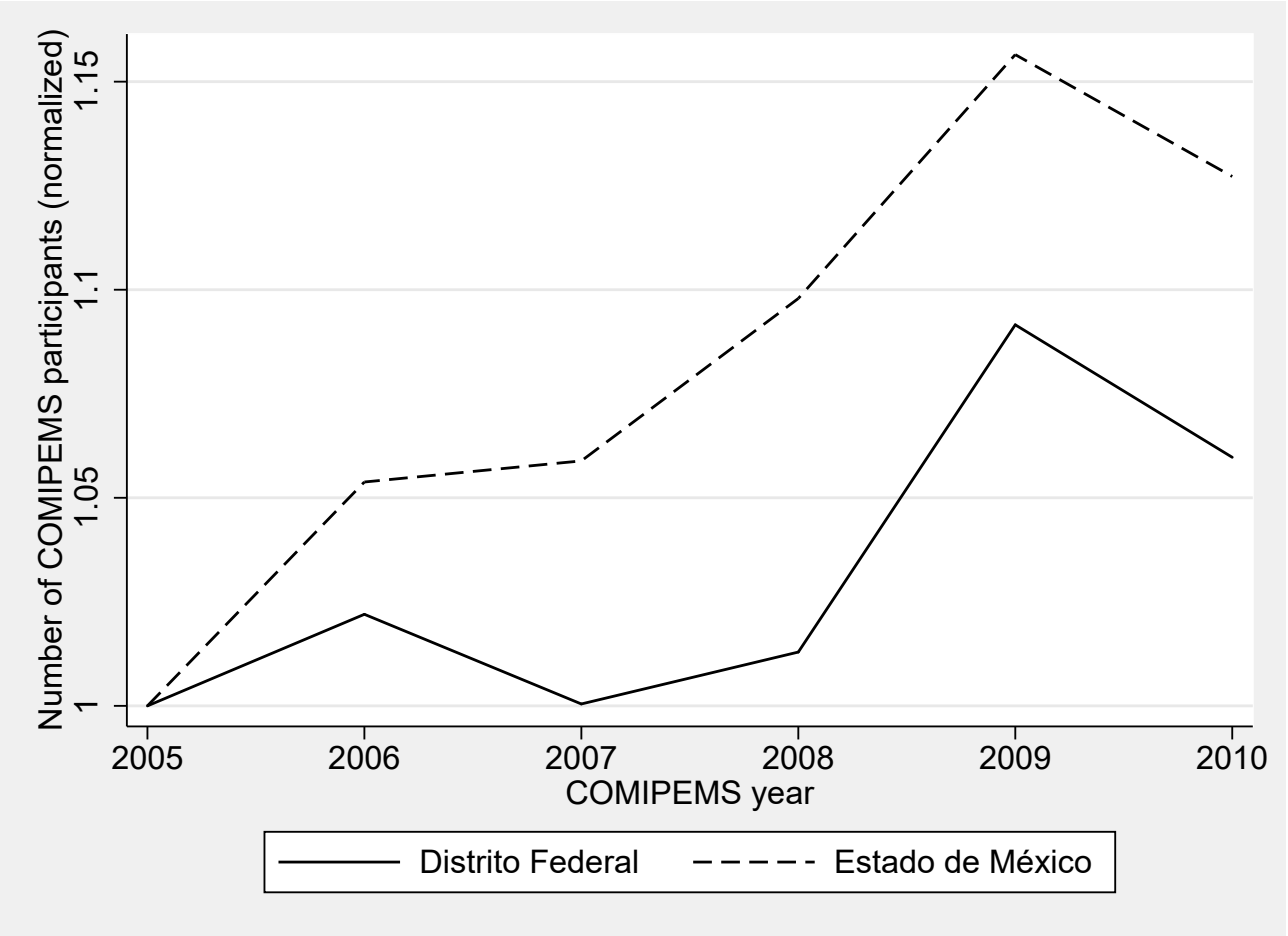
Online appendix to accompany
"Can large, untargeted conditional cash transfers increase urban high school graduation rates? Evidence from Mexico City's Prepa Sí"
by Andrew Dustan

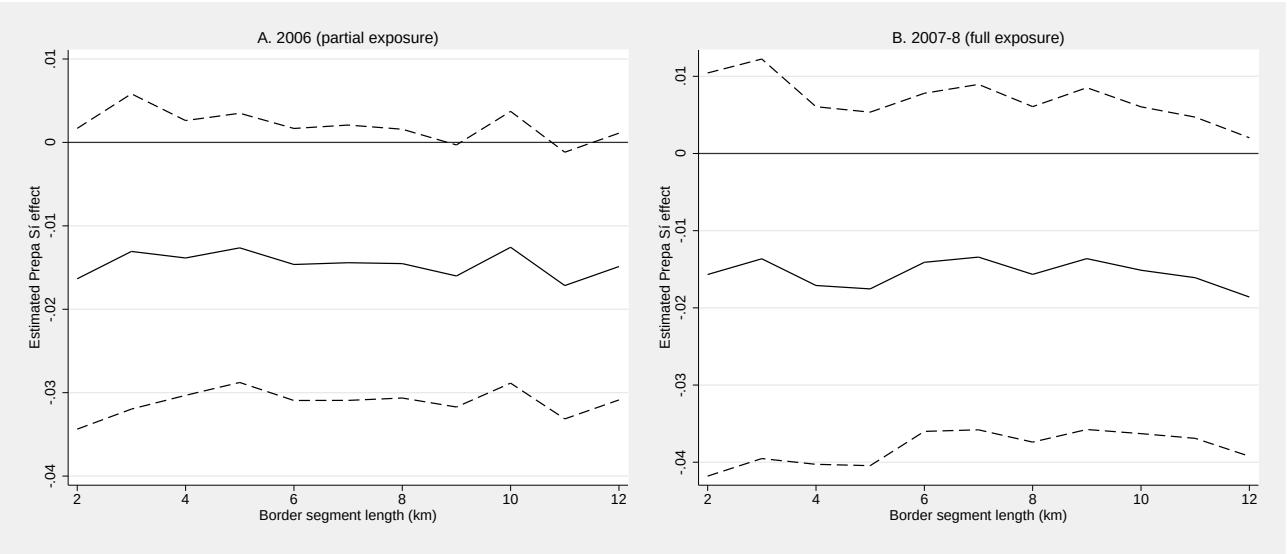
Appendix A. Additional figures and tables

Figure A.1: Number of COMIPEMS takers by year and state of residence, normalized



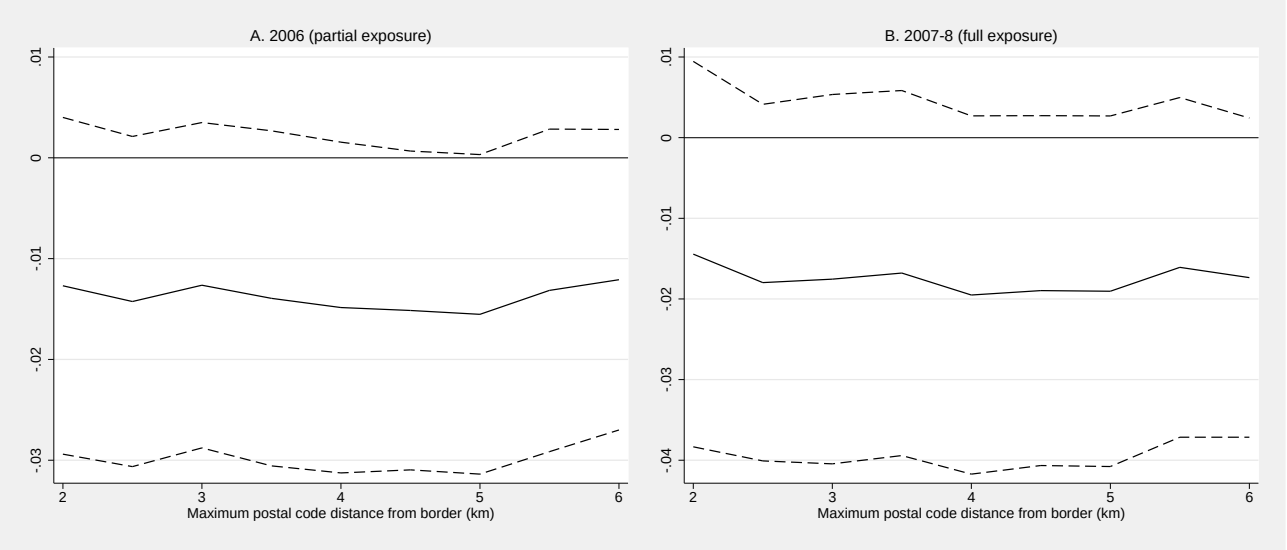
Note: Sample limited to public middle school students. The number of takers in normalized to be 1 in 2005, separately by state.

Figure A.2: Estimated effect of Prepa Sí on ENLACE-taking rate, varying length of border segment



Note: Solid lines are estimated coefficients corresponding to Table 4, Column 6 for the indicated border segment length definition. Dashed lines are the 95% confidence intervals.

Figure A.3: Estimated effect of Prepa Sí on ENLACE-taking rate, varying maximum allowed distance from DF border



Note: Solid lines are estimated coefficients corresponding to Table 4, Column 6 for the indicated maximum distance from DF border allowed for inclusion in the estimation sample. Dashed lines are the 95% confidence intervals.

Table A.1: Summary statistics by place of residence and location of assigned high school, unrestricted by distance to border, 2005-2008 cohorts

	(1) DF resident, DF high school	(2) EdoMex resident, DF high school	(3) EdoMex resident, EdoMex high school	(4) p-value, (1) - (2)	(5) p-value, (2) - (3)
Normalized COMIPEMS score	0.12	0.37	−0.17	0.00	0.00
	[0.87]	[0.93]	[0.82]		
Middle school GPA	7.97	8.23	8.10	0.00	0.00
	[0.80]	[0.84]	[0.79]		
Male	0.49	0.56	0.45	0.00	0.00
	[0.50]	[0.50]	[0.50]		
Parental education (years)	10.42	10.30	9.57	0.00	0.00
	[3.24]	[3.22]	[3.18]		
Income (pesos/month)	4409.45	4494.40	3792.79	0.00	0.00
	[3352.96]	[3343.86]	[3014.29]		
Number of siblings	2.03	2.14	2.35	0.00	0.00
	[1.35]	[1.38]	[1.50]		
Annual fees, assigned school (pesos)	3105.89	2561.22	3759.12	0.00	0.00
	[2243.08]	[2139.67]	[2176.62]		
Distance to assigned school (km)	6.00	12.42	5.34	0.00	0.00
	[4.60]	[7.80]	[4.89]		
Student has matched ENLACE result	0.48	0.56	0.60	0.00	0.00
	[0.50]	[0.50]	[0.49]		
Normalized ENLACE math score	0.04	0.37	−0.02	0.00	0.00
	[1.02]	[1.06]	[0.94]		
Normalized ENLACE Spanish score	0.01	0.20	0.02	0.00	0.00
	[1.01]	[1.00]	[0.96]		
Annual fees, first choice school (pesos)	1141.25	1404.30	2773.61	0.00	0.00
	[1848.67]	[1775.88]	[2455.73]		
Distance to first choice school (km)	6.41	12.14	7.48	0.00	0.00
	[4.56]	[7.46]	[6.59]		
Elite school as first choice	0.18	0.39	0.01	0.00	0.00
	[0.38]	[0.49]	[0.12]		
Normalized cutoff score of first choice	0.70	0.59	−0.15	0.00	0.00
	[1.06]	[1.01]	[1.31]		
Observations	166069	65367	248404		

Note: Standard deviations in brackets. Exchange rate from Mexican pesos to US dollars was approximately 10.8 in 2005.

Table A.2: Effect of Prepa Sí on ENLACE-taking rates in DF high schools, excluding Oportunidades recipients

	(1) Took ENLACE	(2) Took ENLACE	(3) Took ENLACE	(4) Took ENLACE	(5) Took ENLACE	(6) Took ENLACE	(7) Took ENLACE	(8) Took ENLACE
DF × partial exposure	−0.013	−0.018	−0.019	−0.017	−0.013	−0.018	−0.019	−0.017
	(0.0118)	(0.0115)	(0.0122)	(0.0128)	(0.0118)	(0.0115)	(0.0122)	(0.0128)
DF × full exposure	−0.011	−0.013	−0.011	−0.008				
	(0.0089)	(0.0081)	(0.0089)	(0.0101)				
DF × 2007 cohort					−0.011	−0.015	−0.014	−0.010
					(0.0110)	(0.0108)	(0.0112)	(0.0125)
DF × 2008 cohort					−0.012	−0.010	−0.009	−0.005
					(0.0105)	(0.0086)	(0.0102)	(0.0110)
Normalized COMIPEMS score		0.044***	0.040***			0.044***	0.040***	
		(0.0034)	(0.0041)			(0.0034)	(0.0041)	
Middle school GPA		0.210***	0.212***			0.210***	0.212***	
		(0.0046)	(0.0046)			(0.0046)	(0.0046)	
Male		0.043***	0.043***			0.043***	0.043***	
		(0.0050)	(0.0053)			(0.0050)	(0.0053)	
Distance to assigned school (km)		−0.004***	−0.004***			−0.004***	−0.004***	
		(0.0006)	(0.0007)			(0.0006)	(0.0007)	
Elite school as first choice		0.013**	0.007			0.013**	0.007	
		(0.0066)	(0.0071)			(0.0066)	(0.0071)	
Normalized cutoff score of first choice		0.004	0.004			0.004	0.004	
		(0.0026)	(0.0029)			(0.0026)	(0.0029)	
Parental education (years)			0.005***				0.005***	
			(0.0007)				(0.0007)	
Income (1000 pesos/month)			0.003***				0.003***	
			(0.0006)				(0.0006)	
Number of siblings			−0.003**				−0.003**	
			(0.0015)				(0.0015)	
Covariate-by-year interactions				X				X
Observations	79746	78858	68727	68727	79746	78858	68727	68727
Adjusted R^2	0.065	0.162	0.164	0.164	0.065	0.162	0.164	0.164
Baseline ENLACE-taking rate (2005, DF)	0.51	0.51	0.52	0.52	0.51	0.51	0.52	0.52
2005 conditional DF-EdoMex difference	−.02	−.02	−.02	−.02	−.02	−.02	−.02	−.02

Note: Models correspond to Equation 1 and include assigned high school-by-cohort, border segment-by-cohort, postal code, and middle school fixed effects. Sample is restricted to students in postal codes within 3 kilometers of the DF-EdoMex border who did not report being Oportunidades recipients in their COMIPEMS demographic survey. The "2005 conditional DF-EdoMex difference" is the coefficient on DF in a regression of the outcome on DF residence, high school and border segment fixed effects, for the 2005 cohort.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the high school and middle school levels in parentheses.

Table A.3: Effect of Prepa Sí eligibility on ENLACE-taking rates in DF high schools, including middle school-by-cohort fixed effects

	(1) Took ENLACE	(2) Took ENLACE	(3) Took ENLACE	(4) Took ENLACE	(5) Took ENLACE	(6) Took ENLACE	(7) Took ENLACE	(8) Took ENLACE
DF × partial exposure	−0.004	−0.003	−0.007	−0.006	−0.004	−0.003	−0.007	−0.006
	(0.0195)	(0.0180)	(0.0198)	(0.0200)	(0.0195)	(0.0180)	(0.0198)	(0.0200)
DF × full exposure	−0.004	0.001	0.002	0.004				
	(0.0142)	(0.0130)	(0.0124)	(0.0130)				
DF × 2007 cohort					−0.009	−0.004	−0.006	−0.003
					(0.0188)	(0.0180)	(0.0173)	(0.0181)
DF × 2008 cohort					0.000	0.005	0.009	0.011
					(0.0154)	(0.0138)	(0.0139)	(0.0142)
Normalized COMIPEMS score		0.043***	0.038***			0.043***	0.038***	
		(0.0035)	(0.0041)			(0.0035)	(0.0041)	
Middle school GPA		0.211***	0.213***			0.211***	0.213***	
		(0.0045)	(0.0045)			(0.0045)	(0.0045)	
Male		0.043***	0.042***			0.043***	0.042***	
		(0.0052)	(0.0055)			(0.0052)	(0.0055)	
Distance from home (km)		−0.004***	−0.004***			−0.004***	−0.004***	
		(0.0007)	(0.0007)			(0.0007)	(0.0007)	
elite1		0.011	0.003			0.011	0.003	
		(0.0067)	(0.0071)			(0.0067)	(0.0071)	
Normalized cutoff score of first choice		0.004	0.004			0.004	0.004	
		(0.0026)	(0.0029)			(0.0026)	(0.0029)	
Parental education (years)			0.005***				0.005***	
			(0.0007)				(0.0007)	
Income (1000 pesos/month)			0.003***				0.003***	
			(0.0006)				(0.0006)	
Number of siblings			−0.003*				−0.003*	
			(0.0016)				(0.0016)	
Covariate-by-year interactions				X				X
Observations	79173	78293	68176	68176	79173	78293	68176	68176
Adjusted R^2	0.064	0.161	0.163	0.163	0.064	0.161	0.163	0.163
Baseline ENLACE-taking rate (2005, DF)	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52
2005 conditional DF-EdoMex difference	−.02	−.02	−.02	−.02	−.02	−.02	−.02	−.02

Note: Models correspond to Equation 1 and include assigned high school-by-cohort, border segment-by-cohort, postal code, and middle school-by-cohort fixed effects. Sample is restricted to students in postal codes within 3 kilometers of the DF-EdoMex border. The "2005 conditional DF-EdoMex difference" is the coefficient on DF in a regression of the outcome on DF residence, high school and border segment fixed effects, for the 2005 cohort.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the high school and middle school levels in parentheses.

Table A.4: Effect of Prepa Sí on ENLACE-taking rates in DF high schools, excluding CONALEP and IPN schools

	(1) Took ENLACE	(2) Took ENLACE	(3) Took ENLACE	(4) Took ENLACE	(5) Took ENLACE	(6) Took ENLACE	(7) Took ENLACE	(8) Took ENLACE
DF × partial exposure	−0.024	−0.029**	−0.031**	−0.029*	−0.024	−0.029**	−0.031**	−0.029*
	(0.0146)	(0.0143)	(0.0142)	(0.0146)	(0.0146)	(0.0143)	(0.0142)	(0.0146)
DF × full exposure	−0.008	−0.010	−0.007	−0.003				
	(0.0142)	(0.0130)	(0.0133)	(0.0152)				
DF × 2007 cohort					−0.012	−0.017	−0.015	−0.008
					(0.0163)	(0.0163)	(0.0164)	(0.0179)
DF × 2008 cohort					−0.004	−0.004	0.001	0.001
					(0.0152)	(0.0128)	(0.0135)	(0.0154)
Normalized COMIPEMS score		0.037***	0.032***			0.037***	0.032***	
		(0.0051)	(0.0065)			(0.0051)	(0.0065)	
Middle school GPA		0.202***	0.203***			0.202***	0.203***	
		(0.0052)	(0.0052)			(0.0052)	(0.0052)	
Male		0.040***	0.040***			0.040***	0.040***	
		(0.0070)	(0.0072)			(0.0070)	(0.0072)	
Distance to assigned school (km)		−0.005***	−0.005***			−0.005***	−0.005***	
		(0.0009)	(0.0011)			(0.0009)	(0.0011)	
Elite school as first choice		0.021***	0.017**			0.021***	0.017**	
		(0.0065)	(0.0064)			(0.0065)	(0.0064)	
Normalized cutoff score of first choice		0.005*	0.004			0.005*	0.004	
		(0.0029)	(0.0030)			(0.0029)	(0.0030)	
Parental education (years)			0.004***				0.004***	
			(0.0007)				(0.0007)	
Income (1000 pesos/month)			0.003***				0.003***	
			(0.0008)				(0.0008)	
Number of siblings			−0.005**				−0.005**	
			(0.0022)				(0.0022)	
Covariate-by-year interactions				X				X
Observations	43155	42717	36992	36992	43155	42717	36992	36992
Adjusted R^2	0.049	0.140	0.141	0.141	0.049	0.140	0.141	0.141
Baseline ENLACE-taking rate (2005, DF)	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
2005 conditional DF-EdoMex difference	−.02	−.02	−.02	−.02	−.02	−.02	−.02	−.02

Note: Models correspond to Equation 1 and include assigned high school-by-cohort, border segment-by-cohort, postal code, and middle school fixed effects. Sample is restricted to students in postal codes within 3 kilometers of the DF-EdoMex border who were not assigned to CONALEP or IPN high schools. The "2005 conditional DF-EdoMex difference" is the coefficient on DF in a regression of the outcome on DF residence, high school and border segment fixed effects, for the 2005 cohort.

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the high school and middle school levels in parentheses.

Table A.5: Differential impacts of Prepa Sí eligibility on ENLACE-taking rates with respect to high school fees and distance, low parental education subsample

	(1) Took ENLACE	(2) Took ENLACE	(3) Took ENLACE	(4) Took ENLACE	(5) Took ENLACE
Annual fees (1000 pesos) × partial exposure	0.000	0.001			
	(0.0022)	(0.0024)			
Annual fees (1000 pesos) × full exposure	−0.004	−0.004			
	(0.0035)	(0.0033)			
Distance from home (km) × partial exposure			0.001	0.001	−0.000
			(0.0012)	(0.0014)	(0.0014)
Distance from home (km) × full exposure			0.001	0.002	−0.001
			(0.0011)	(0.0014)	(0.0013)
Distance from home (km)	−0.003***	−0.003***	−0.004***	−0.005***	−0.003**
	(0.0005)	(0.0005)	(0.0010)	(0.0012)	(0.0011)
Postal code FE	X		X		
Postal code-by-cohort FE		X		X	X
High school FE	X	X	X	X	
High school-by-cohort FE					X
Middle school FE	X		X		
Middle school-by-cohort FE		X		X	X
Cohort FE	X		X		
Observations	69829	69308	69829	69308	69308
Adjusted R^2	0.138	0.138	0.138	0.138	0.144
Baseline ENLACE-taking rate (2005)	0.47	0.47	0.47	0.47	0.47

Note: Models correspond to Equation 3 and include student-level controls for normalized COMIPEMS score, middle school GPA, and elite school as first choice, along with the fixed effects indicated. Sample is limited to students with DF residence.
* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the high school and middle school levels in parentheses.

Table A.6: Differential impacts of Prepa Sí eligibility on ENLACE-taking rates with respect to high school fees and distance, low income subsample

	(1) Took ENLACE	(2) Took ENLACE	(3) Took ENLACE	(4) Took ENLACE	(5) Took ENLACE
Annual fees (1000 pesos) × partial exposure	−0.000	−0.000			
	(0.0027)	(0.0031)			
Annual fees (1000 pesos) × full exposure	−0.004	−0.004			
	(0.0040)	(0.0037)			
Distance from home (km) × partial exposure			0.001	0.001	−0.001
			(0.0013)	(0.0016)	(0.0015)
Distance from home (km) × full exposure			0.002	0.002	0.001
			(0.0011)	(0.0015)	(0.0014)
Distance from home (km)	−0.004***	−0.004***	−0.005***	−0.005***	−0.004***
	(0.0006)	(0.0006)	(0.0010)	(0.0012)	(0.0012)
Postal code FE	X		X		
Postal code-by-cohort FE		X		X	X
High school FE	X	X	X	X	
High school-by-cohort FE					X
Middle school FE	X		X		
Middle school-by-cohort FE		X		X	X
Cohort FE	X		X		
Observations	54677	54152	54677	54152	54152
Adjusted R^2	0.139	0.138	0.139	0.138	0.144
Baseline ENLACE-taking rate (2005)	0.48	0.48	0.48	0.48	0.48

Note: Models correspond to Equation 3 and include student-level controls for normalized COMIPEMS score, middle school GPA, and elite school as first choice, along with the fixed effects indicated. Sample is limited to students with DF residence.
* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the high school and middle school levels in parentheses.

Table A.7: Effect of Prepa Sí eligibility on ENLACE test scores in DF high schools, by cohort

Panel A. Math scores	(1)	(2)	(3)	(4)	(5)	(6)
	All students	Male	Female	Parental education $<$ HS	Family income $<$ 3000 MXN/mo	Middle school GPA $\geq$ 8
DF $\times$ partial exposure	−0.018	−0.019	−0.034	0.052	0.006	−0.029
	(0.0219)	(0.0336)	(0.0279)	(0.0396)	(0.0430)	(0.0264)
DF $\times$ 2007 cohort	−0.005	−0.014	−0.014	0.014	0.034	0.002
	(0.0250)	(0.0386)	(0.0292)	(0.0399)	(0.0455)	(0.0297)
DF $\times$ 2008 cohort	0.018	0.037	−0.008	0.032	0.039	−0.001
	(0.0203)	(0.0305)	(0.0309)	(0.0360)	(0.0456)	(0.0257)
Observations	40515	19625	20676	17935	13540	25562
Adjusted R^2	0.553	0.562	0.509	0.519	0.515	0.568
Baseline mean (2005, DF)	0.17	0.39	−0.04	0.02	−0.01	0.40
2005 conditional DF-EdoMex difference	−.04	−.04	−.03	−.06	−.06	−.04

Panel B. Spanish scores	(1)	(2)	(3)	(4)	(5)	(6)
	All students	Male	Female	Parental education $<$ HS	Family income $<$ 3000 MXN/mo	Middle school GPA $\geq$ 8
DF $\times$ partial exposure	−0.018	−0.022	0.001	−0.001	−0.026	−0.038
	(0.0261)	(0.0420)	(0.0362)	(0.0420)	(0.0461)	(0.0287)
DF $\times$ 2007 cohort	0.002	−0.044	0.042	0.029	0.019	0.012
	(0.0269)	(0.0387)	(0.0329)	(0.0396)	(0.0471)	(0.0383)
DF $\times$ 2008 cohort	0.016	0.017	0.028	0.029	0.005	−0.007
	(0.0249)	(0.0416)	(0.0298)	(0.0403)	(0.0452)	(0.0271)
Observations	40578	19667	20696	17950	13557	25583
Adjusted R^2	0.416	0.423	0.412	0.403	0.397	0.410
Baseline mean (2005, DF)	0.08	0.04	0.11	−0.06	−0.07	0.30
2005 conditional DF-EdoMex difference	−.02	0.01	−.06	−.03	−.04	−.01

Note: Column headers represent sample restrictions. Models correspond to Equation 1 and include assigned high school-by-cohort, border segment-by-cohort, postal code, and middle school fixed effects, as well as student-level controls for normalized COMIPEMS score, middle school GPA, distance to assigned school, and elite school as first choice. Sample is restricted to students in postal codes within 3 kilometers of the DF-EdoMex border. The "2005 conditional DF-SoM difference" is the coefficient on DF in a regression of the outcome on DF residence, high school and border segment fixed effects, for the 2005 cohort.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the high school and middle school levels in parentheses.

Table A.8: Summary statistics by place of residence, 2005-2010 cohorts (school choice sample)

	(1) DF resident	(2) EdoMex resident	(3) p-value, (1) - (2)
Normalized COMIPEMS score	0.11	−0.08	0.00
	[1.01]	[0.98]	
Middle school GPA	8.05	8.11	0.00
	[0.83]	[0.82]	
Male	0.46	0.46	0.21
	[0.50]	[0.50]	
Parental education (years)	10.71	10.08	0.00
	[3.24]	[3.24]	
Income (pesos/month)	4640.39	4245.29	0.00
	[3528.74]	[3319.58]	
Number of siblings	1.97	2.19	0.00
	[1.33]	[1.43]	
Annual fees, assigned school (pesos)	2795.27	3164.69	0.00
	[2711.60]	[2699.96]	
Distance to assigned school (km)	6.11	6.40	0.00
	[5.05]	[5.54]	
Student has matched ENLACE result	0.38	0.45	0.00
	[0.49]	[0.50]	
Normalized ENLACE math score	−0.05	−0.07	0.00
	[1.02]	[1.00]	
Normalized ENLACE Spanish score	−0.07	−0.08	0.01
	[1.01]	[0.99]	
Annual fees, first choice school (pesos)	995.47	1595.25	0.00
	[1912.83]	[2392.91]	
Distance to first choice school (km)	6.64	7.56	0.00
	[4.84]	[5.25]	
Elite school as first choice	0.33	0.23	0.00
	[0.47]	[0.42]	
Normalized cutoff score of first choice	0.81	0.45	0.00
	[1.00]	[1.22]	
Observations	147767	154068	

Note: Standard deviations in brackets. Sample is restricted to students in postal codes within 3 kilometers of the DF-EdoMex border. Exchange rate from Mexican pesos to US dollars was approximately 10.8 in 2005.

Table A.9: Effect of Prepa Sí introduction on characteristics of first choice and assigned schools, 2005-2010 cohorts, low parental education sample

Panel A. First choice school characteristics	(1) Annual fees	(2) Distance from home	(3) Located in DF	(4) Normalized cutoff score	(5) Annual fees	(6) Distance from home	(7) Located in DF	(8) Normalized cutoff score
Federal District × 2008-10 cohorts	−18.87	−0.22***	−0.01***	0.02	−75.96	−0.26*	0.00	0.03
	(36.604)	(0.074)	(0.004)	(0.014)	(69.836)	(0.133)	(0.008)	(0.029)
Postal code trends					X	X	X	X
Observations	139839	138618	139848	139848	139839	138618	139848	139848
Adjusted R^2	0.111	0.228	0.246	0.185	0.111	0.228	0.247	0.185
Mean of dependent variable (2007, DF)	1322.25	6.80	0.95	0.62	1322.25	6.80	0.95	0.62
SD of dependent variable (2007, DF)	2464.22	4.91	0.23	1.16	2464.22	4.91	0.23	1.16
2007 conditional DF-EdoMex difference	−470.22	−.26	0.27	0.26	−470.22	−.26	0.27	0.26

Panel B. Assigned school characteristics	(1) Annual fees	(2) Distance from home	(3) Located in DF	(4) Normalized cutoff score	(5) Annual fees	(6) Distance from home	(7) Located in DF	(8) Normalized cutoff score
Federal District × 2008-10 cohorts	−146.01*	−0.15**	0.01	0.04***	−202.56*	−0.20	0.01	0.04***
	(81.253)	(0.074)	(0.006)	(0.008)	(116.643)	(0.138)	(0.011)	(0.014)
Postal code trends					X	X	X	X
Observations	116149	115191	116196	139848	116149	115191	116196	139848
Adjusted R^2	0.140	0.120	0.363	0.645	0.142	0.120	0.363	0.645
Mean of dependent variable (2007, DF)	3459.18	6.15	0.86	−0.74	3459.18	6.15	0.86	−0.74
SD of dependent variable (2007, DF)	3663.33	5.15	0.34	0.96	3663.33	5.15	0.34	0.96
2007 conditional DF-EdoMex difference	−93.17	0.08	0.44	0.12	−93.17	0.08	0.44	0.12

Note: Dependent variables in column headers. Models correspond to Equation 4 and include border segment-by-cohort, postal code, and middle school fixed effects; postal code-level linear trends (where indicated); and student normalized COMIPEMS score, middle school GPA, and dummy for male. Sample is restricted to students in postal codes within 3 kilometers of the DF-EdoMex border whose parental education is less than high school. The "2007 conditional DF-EdoMex difference" is the coefficient on DF in a regression of the outcome on DF residence and border segment fixed effects, for the 2007 cohort.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the postal code and middle school levels in parentheses.

Table A.10: Effect of Prepa Sí introduction on characteristics of first choice and assigned schools, 2005-2010 cohorts, low family income sample

Panel A. First choice school characteristics	(1) Annual fees	(2) Distance from home	(3) Located in DF	(4) Normalized cutoff score	(5) Annual fees	(6) Distance from home	(7) Located in DF	(8) Normalized cutoff score
Federal District × 2008-10 cohorts	2.53	−0.14**	−0.01**	0.02*	−33.19	0.10	−0.00	−0.00
	(36.359)	(0.070)	(0.005)	(0.014)	(68.858)	(0.142)	(0.015)	(0.037)
Postal code trends					X	X	X	X
Observations	105418	104524	105424	105424	105418	104524	105424	105424
Adjusted R^2	0.114	0.221	0.254	0.190	0.114	0.221	0.254	0.190
Mean of dependent variable (2007, DF)	1324.39	6.62	0.95	0.63	1324.39	6.62	0.95	0.63
SD of dependent variable (2007, DF)	2504.24	4.85	0.22	1.15	2504.24	4.85	0.22	1.15
2007 conditional DF-EdoMex difference	−548.48	−.33	0.29	0.31	−548.48	−.33	0.29	0.31

Panel B. Assigned school characteristics	(1) Annual fees	(2) Distance from home	(3) Located in DF	(4) Normalized cutoff score	(5) Annual fees	(6) Distance from home	(7) Located in DF	(8) Normalized cutoff score
Federal District × 2008-10 cohorts	−167.41**	0.05	0.01	0.04***	−236.55*	0.28*	0.00	0.02
	(78.260)	(0.081)	(0.006)	(0.009)	(128.695)	(0.155)	(0.015)	(0.016)
Postal code trends					X	X	X	X
Observations	87239	86551	87270	105424	87239	86551	87270	105424
Adjusted R^2	0.141	0.116	0.368	0.648	0.143	0.116	0.369	0.648
Mean of dependent variable (2007, DF)	3502.39	6.00	0.87	−0.75	3502.39	6.00	0.87	−0.75
SD of dependent variable (2007, DF)	3692.39	5.03	0.34	0.96	3692.39	5.03	0.34	0.96
2007 conditional DF-EdoMex difference	−111.63	−.09	0.46	0.15	−111.63	−.09	0.46	0.15

Note: Dependent variables in column headers. Models correspond to Equation 4 and include border segment-by-cohort, postal code, and middle school fixed effects; postal code-level linear trends (where indicated); and student normalized COMIPEMS score, middle school GPA, and dummy for male. Sample is restricted to students in postal codes within 3 kilometers of the DF-EdoMex border whose self-reported family income is below 2,500 pesos per month. The "2007 conditional DF-EdoMex difference" is the coefficient on DF in a regression of the outcome on DF residence and border segment fixed effects, for the 2007 cohort.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the postal code and middle school levels in parentheses.

Appendix B. Estimation of dropout rates

This appendix describes the process used to estimate the proportion of total dropout that occurs in the first and second years of high school. There is no student-level microdata on grade progression and dropout for this population. The annual school census provides raw counts of new students in each grade, the number of students repeating each grade, and the number of graduates. Restricting the universe of schools to Federal District high schools that participate in COMIPEMS but excluding the UNAM schools, these raw counts (aggregated over all schools) are used to estimate the probability that a student in grade 10 either 1) drops out, 2) repeats the grade, or 3) advances to the next grade. These probabilities are also estimated for grade 11. For grade 12, the relevant probabilities are for the student 1) dropping out, repeating the grade, or 3) graduating. For example, the rate of continuation from 10th to 11th grade in year t is estimated by $\text{cont1011}_t = (\text{new 11th graders})_{t+1}/(\text{all 10th graders})_t$. I estimate the following transition rates, using 2005 as the base year for 10th grade, 2006 for 11th grade, and 2007 for 12th grade:

- Continue from 10th to 11th: 61.9%
- Repeat 10th: 6.7%
- Dropout in 10th: 100% - (61.9% + 6.7%) = 31.3%
- Continue from 11th to 12th: 72.2%
- Repeat 11th: 12.0%
- Dropout in 11th: 100% - (72.2% + 12.0%) = 15.8%
- Graduate: 79.5%
- Repeat 12th: 14.6%
- Dropout in 12th: 100% - (79.5% + 14.6%) = 5.9%

I then generate a synthetic cohort of 1 million 10th grade students. I use these probabilities to simulate an action (continuation, repetition, or dropout) for each student in the first year. This simulation is repeated on the same students for five years, tracking students' status (dropout, grade level, or graduation) across years. Each student is then classified as either a graduate or non-graduate (i.e. dropout), and the percentage of non-graduates who dropped out in the first two years is computed.

This process results in an estimate of 63% of all eventual dropout occurring in the first year and 87% taking place within the first two years. These rates are very similar when using different cohorts for the transition rates.

Appendix C. Using Prepa Sí's grade point average rules to estimate inframarginal effects

This appendix reports the results of a regression discontinuity (RD) analysis that estimates the effects of eligibility for a higher Prepa Sí transfer amount on high school completion. Prepa Sí is designed to give higher payments to students with higher grade point averages (GPAs): students with a GPA of 7.5 or below receive 500 pesos/month, those with GPAs from 7.6 to 9.0 receive 600 pesos/month, and those with GPAs from 9.1 to 10 receive 700 pesos/month. In the Mexican education system, GPAs are rounded down to the first decimal point. The program rules were not explicit with respect to which GPA would be used to determine the transfer amount, which creates some ambiguity in the correct transfer amount. For students in their first year of high school, the most likely case is that high schools report students' middle school GPA to Prepa Sí for the purpose of determining the transfer amount. Thus, under the assumption that all other factors affecting high school completion rates are smooth near the 7.6 and 9.1 middle school GPA cutoffs, an RD design can be used to estimate the completion effects of being eligible for an additional 100 pesos/month transfer. The COMIPEMS microdata provide student-level middle school GPA.

The effect of eligibility for a higher transfer amount (according to middle school GPA) is estimated using the following RD specification, separately for the 7.6 and 9.1 point cutoffs:

$$y_i = \alpha + \delta above_i + \beta_1 \widetilde{gpa_i} + \beta_2 \left(above_i \times \widetilde{gpa_i}\right) + \varepsilon_i,$$

where $above_i$ is a dummy variable equal to 1 if student i's GPA exceeds the cutoff of interest and $\widetilde{gpa_i}$ is i's GPA, centered to equal 0 at the cutoff value. Similar to the issue facing RD estimation using COMIPEMS exam scores, this analysis must account for the discreteness of GPA. The standard deviation of GPA in this sample is approximately 0.8 points. This discreteness has two consequences. First, nonparametric approaches using bandwidth selection procedures such as Calonico, Cattaneo, and Titiunik (2014) are not appropriate. Instead, I report results for a range of bandwidths. Second, the Lee and Card (2008) method of clustering standard errors at the level of the running variable cannot be used to account for the potentially severe specification error caused by the few points of support by using . This problem is explained by Kolesár and Rothe (2018), who propose alternative methods of inference that are robust to specification error arising from a discrete running variable. I apply one of their suggested inference methods, bounded misspecification error (BME), which generates confidence intervals under the assumption that the true misspecification error at the cutoff does not exceed the error observed for any point of support within the

chosen bandwidth.[1] I also show estimates from specifications that add quadratic terms in $\widetilde{gpa}_i$ and its interaction with $above_i$, which in some cases result in tighter BME confidence intervals.

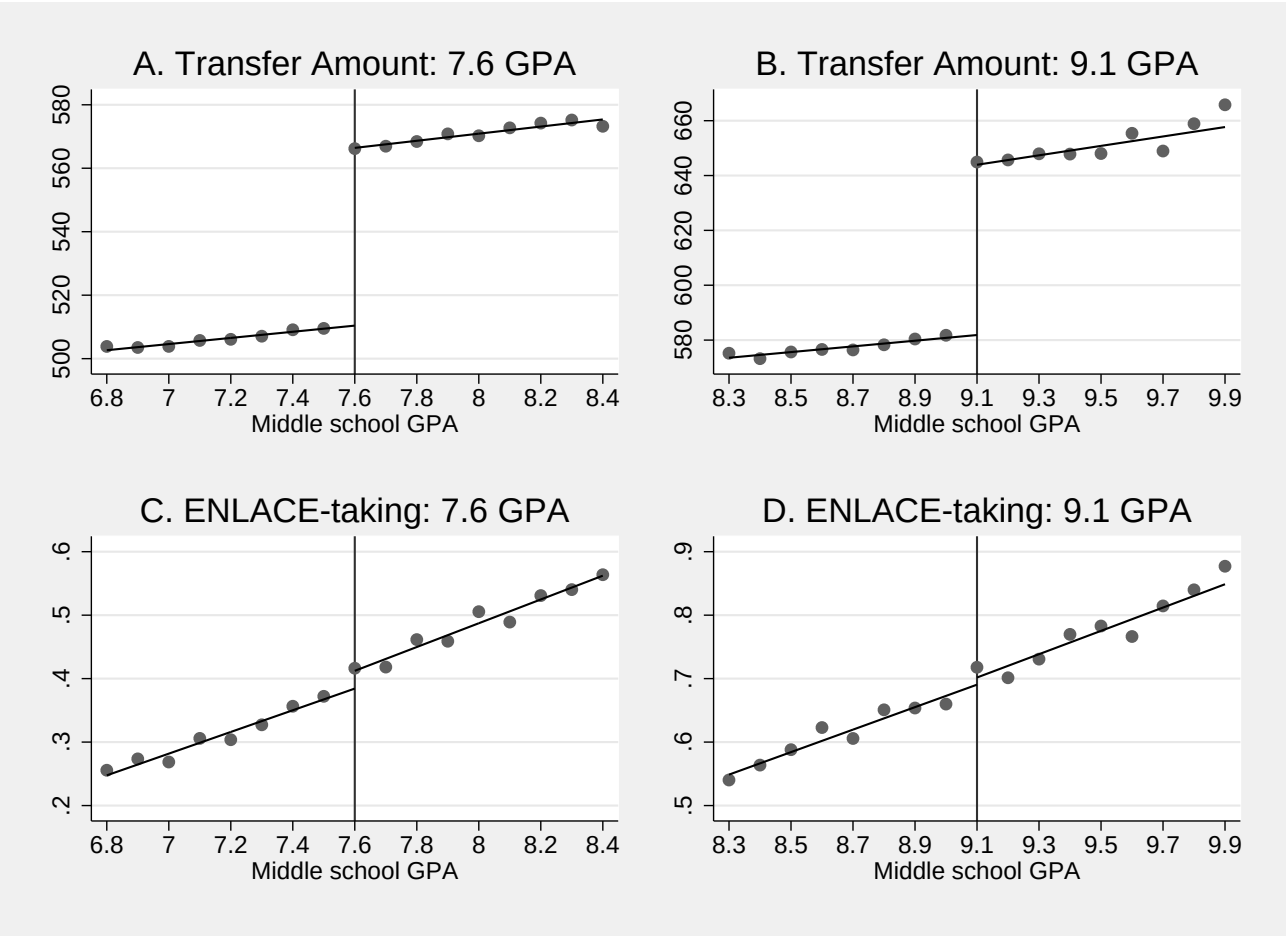
I first estimate the discontinuity in the Prepa Sí monthly transfer amount at each cutoff, using the transfer amounts reported on the 2007 and 2008 beneficiary lists for students from the incoming cohort. These are the years for which I observe the transfer amount for beneficiaries. Appendix Table C.1, Panel A, shows the results for a bandwidth of 0.8 points (about 1 standard deviation), illustrated in Appendix Figure C.1. Meeting the 7.6 GPA cutoff predicts an increase of 56 pesos per month, slightly more than half of the increase that would be expected if middle school GPA was the exclusive determinant of transfer amount. Results are similar for a quadratic polynomial in the running variable, shown in Column 2. Appendix Figure C.2 shows that this estimate is insensitive to bandwidth choice. Columns 3 and 4 show similar results for the 9.1 GPA cutoff: on average, meeting the cutoff results in a 62 peso per month increase in the transfer. While these results show a strong "first stage" effect of meeting the GPA cutoff on transfer amount, I will not estimate an instrumental variables model of the effect of transfer amount on high school completion. This is because transfer amount is only observed for students enrolled in Prepa Sí (which is endogenous to dropout), while sample used in estimating the reduced-form effect on ENLACE-taking is not subject to this limitation.

Panel B of Appendix Table C.1 shows the reduced-form effect of meeting each GPA cutoff on the probability of taking the ENLACE, using the sample of all 2007 and 2008 cohort DF residents who were assigned to DF high schools. Regardless of the cutoff or polynomial order, estimated effects are statistically insignificant. Appendix Figure C.1 shows that GPA is a strong predictor of ENLACE-taking, while also illustrating why the BME confidence intervals are important in this setting. While the Huber-White robust standard errors in some specifications yield positive estimated effects, it is clear that the increase in ENLACE-taking rate at the cutoff is similar in magnitude to the misspecification error at several points of support that are away from the cutoff. Varying the bandwidth does not result in any statistically significant effects, as shown in C.2. The BME confidence intervals contain both large positive and large negative effects of meeting each GPA cutoff. Thus while middle school GPA did indeed affect average transfer amounts for those enrolled in Prepa Sí, the evidence for the effect of eligibility for higher amounts on high school completion is

[1]BME confidence intervals can be calculated for this GPA analysis because the regression specification does not include additional covariates. This is in contrast to the difference-in-discontinuities analysis in Appendix D, which interacts the admission indicator and linear splines with exposure to Prepa Sí and includes cutoff school-by-year fixed effects. The BME method has not been extended to accommodate such a specification.

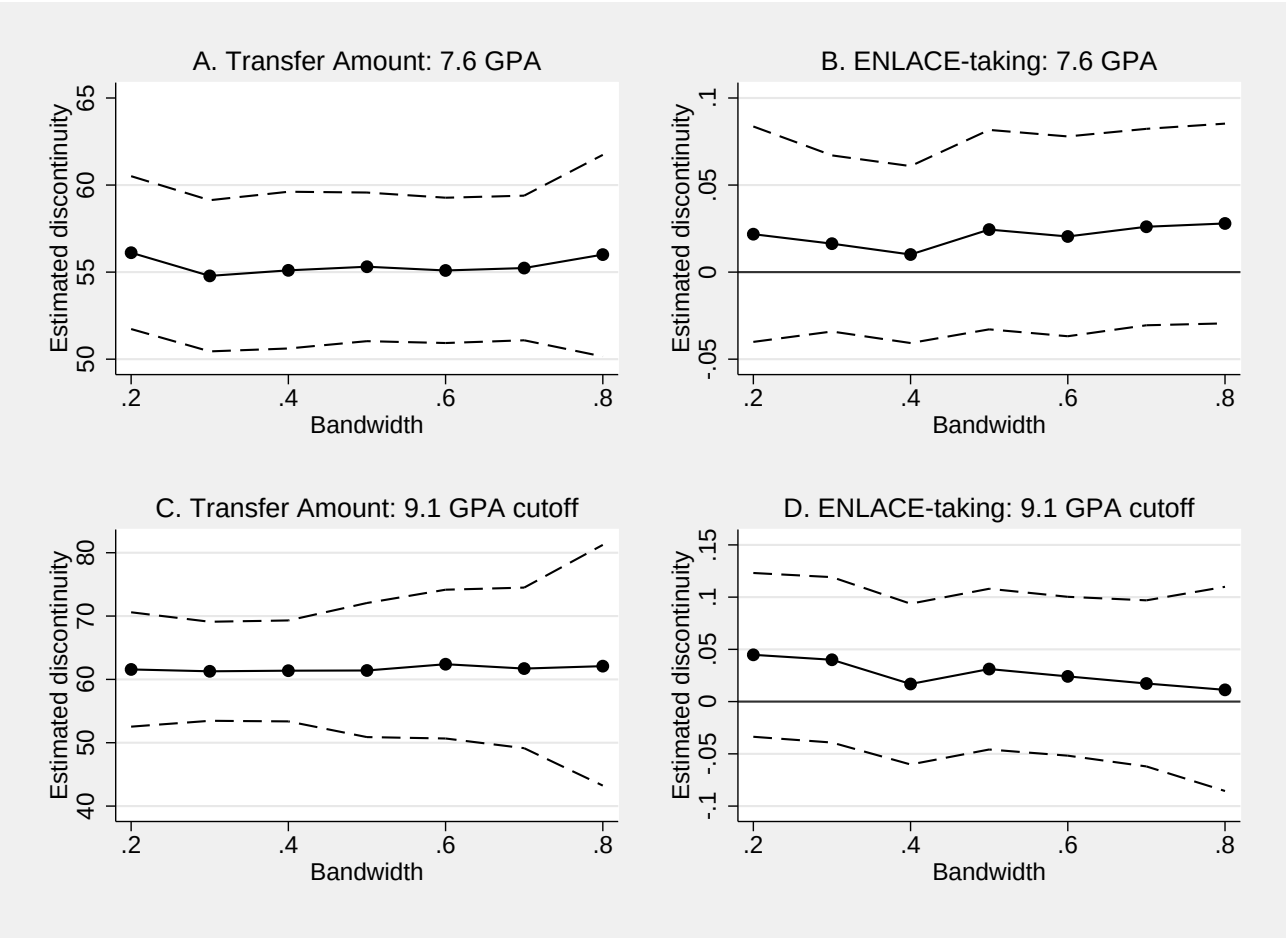
inconclusive.

Figure C.1: Regression discontinuity effects of eligibility for higher Prepa Sí transfer amount



Note: Circles are mean values of the dependent variable for the indicated (discrete) level of middle school GPA. Lines are linear fits with respect to middle school GPA. Sample is restricted to 2007 and 2008 cohort DF residents within 0.8 points (about 1 standard deviation) of the indicated GPA cutoff who were assigned to a DF high school.

Figure C.2: Regression discontinuity effects of eligibility for higher Prepa Sí transfer amount, varying bandwidth



Note: Estimated coefficients (circles) are from local linear regression of the indicated dependent variable on a constant, a dummy for whether middle school GPA meets the cutoff to qualify for the higher Prepa Sí transfer amount, middle school GPA normalized to zero at the cutoff, and an interaction of the latter two variables. Sample is restricted to 2007 and 2008 cohort DF residents within the specified bandwidth of the indicated GPA cutoff. Dashed lines are Kolesár and Rothe (2018) bounded misspecification error (BME) 95% confidence intervals.

Table C.1: Regression discontinuity effects of eligibility for higher Prepa Sí transfer amount

Panel A. Transfer amount	(1) Transfer amount: 7.6 GPA	(2) Transfer amount: 7.6 GPA	(3) Transfer amount: 9.1 GPA	(4) Transfer amount: 9.1 GPA
GPA meets cutoff	56.01	54.23	62.09	61.13
	(2.293)	(2.344)	(2.796)	(2.836)
	[50.16, 61.73]	[49.39, 59.11]	[43.23, 81.24]	[46.73, 75.47]
Observations	70346	70346	43504	43504
Adjusted R^2	0.401	0.401	0.266	0.266
Mean of dependent variable	543.24	543.24	602.39	602.39
Polynomial order	1	2	1	2

Panel B. ENLACE-taking	(1) Took ENLACE: 7.6 GPA	(2) Took ENLACE: 7.6 GPA	(3) Took ENLACE: 9.1 GPA	(4) Took ENLACE: 9.1 GPA
GPA meets cutoff	0.03	0.01	0.01	0.04
	(0.008)	(0.013)	(0.011)	(0.017)
	[-0.03, 0.09]	[-0.05, 0.07]	[-0.09, 0.11]	[-0.05, 0.13]
Observations	115932	115932	60618	60618
Adjusted R^2	0.037	0.037	0.029	0.029
Mean of dependent variable	0.40	0.40	0.65	0.65
Polynomial order	1	2	1	2

Note: Estimated coefficient is from local polynomial regression of the indicated dependent variable on a constant, a dummy for whether middle school GPA meets the cutoff to qualify for the higher Prepa Sí transfer amount, a polynomial of the indicated order in middle school GPA normalized to zero at the cutoff, and an interaction of the polynomial and the meeting the cutoff dummy. Sample is restricted to 2007 and 2008 cohort DF residents within 0.8 points (about 1 standard deviation) of the indicated GPA cutoff who were assigned to a DF high school. Standard errors clustered at the high school and middle school levels in parentheses. Kolesár and Rothe (2018) bounded misspecification error (BME) 95% confidence intervals in brackets.

Appendix D. Difference-in-discontinuities design using assignment cutoffs

The COMIPEMS assignment mechanism generates sharp discontinuities in the probability of certain students being assigned to DF schools, permitting a complementary empirical approach to estimating the effect of Prepa Sí on high school completion rates that relies on a different set of assumptions.

D.1 Econometric approach

The COMIPEMS assignment mechanism generates sharp discontinuities in the probability of certain students being assigned to DF schools, permitting a complementary empirical approach that relies on a different set of assumptions. As the assignment mechanism proceeds down the entrance exam score-ordered list of students and assigns each student to his most-preferred school with an open seat remaining, schools begin to fill their seat quotas. Because the mechanism ranks students only by exam score, when a student is given the final seat in school j, no student with a lower exam score can be assigned there. Thus each school that fills its seat quota has an (ex post) cutoff score, c_j, equal to the exam score of its lowest-scoring student. Consider, then, a school j that is located in the DF, with cutoff score c_j. There is a set of students who, because of their ranked preferences, would have been assigned to j (and thus to a school in the DF) if they had an exam score of c_j. A subset of these students would have been assigned to schools in EdoMex if they had an exam score of $c_j - 1$. For these students, moving from a score of $c_j - 1$ to c_j increases the probability of DF school assignment from 0 to 1. A sharp regression discontinuity (RD) design can thus be used in this subsample to estimate the effect of DF admission (and, for DF residents in certain cohorts, Prepa Sí eligibility) on high school completion, using exam score as the running variable and the DF high school's cutoff score as the cutoff value. Similarly, for each EdoMex school j that fills its capacity, there may be a subset of students for whom a score of c_j results in assignment to j and a score of $c_j - 1$ results in assignment to a DF school. The same RD approach can be applied to these students, except that the DF assignment "treatment" occurs for students directly below the cutoff.

This sharp RD design is similar to those taken by several other papers using the COMIPEMS data to estimate the effects of elite school admission on academic outcomes (Dustan, de Janvry, and Sadoulet 2017; Estrada and Gignoux 2017) and the effects of school assignments on sibling school choices (Dustan 2018). The present analysis will follow much of these papers' empirical strategies, although some modifications are necessary to suit the research question

at hand. Because there are many high schools that filled their quotas, I use a stacked RD specification that includes all students near the margin of DF admission (whether EdoMex-to-DF or DF-to-EdoMex) and includes cutoff school-by-cohort fixed effects. As Dustan, de Janvry, and Sadoulet (2017) show for the case of elite schools, marginal admission to a DF school could affect dropout for reasons other than Prepa Sí eligibility, for example attending a different type of school, occupying a different position in the ability distribution, or facing a longer commute. Assuming that these other effects are constant over time, a difference-in-discontinuities design suggested by Grembi, Nannicini, and Troiano (2016) can be used to difference out such effects and isolate the Prepa Sí effect by comparing the admission effects for the essentially unexposed cohort with the effects for partially- and fully-exposed cohorts. The regression specification is:

$$\begin{aligned} y_{ijt} = \alpha_{jt} + DFadmit_{ijt}\left(\delta_0 + \delta_1 partial_t + \delta_2 full_t\right) + \tilde{s}_{ijt}\left(\beta_{00} + \beta_{01} partial_t + \beta_{02} full_t\right) \\ + DFcutoff_j \times \tilde{s}_{ijt}\left(\beta_{10} + \beta_{11} partial_t + \beta_{12} full_t\right) + DFadmit_{ijt} \times \tilde{s}_{ijt}\left(\beta_{20} + \beta_{21} partial_t + \beta_{22} full_t\right) \\ + DFcutoff_j \times DFadmit_{ijt} \times \tilde{s}_{ijt}\left(\beta_{30} + \beta_{31} partial_t + \beta_{32} full_t\right) + \varepsilon_{ijt}, \end{aligned} \tag{D.1}$$

where α_{jt} is a cutoff school-by-cohort fixed effect, $\tilde{s}_{ijt} \equiv s_i - c_{jt}$ is student i's exam score centered to equal 0 at j's cutoff score in year t, $DFcutoff_j$ is a dummy variable equal to 1 if cutoff school j is in the DF, and $DFadmit_{ijt}$ is a dummy variable equal to 1 if either of the following conditions is true: (1) $\tilde{s}_{ijt} \geq 0$ and $DFcutoff_j = 1$ or (2) $\tilde{s}_{ijt} < 0$ and $DFcutoff_j = 0$. That is, $DFadmit_{ijt}$ is 1 when a student meets a DF school's cutoff score or when a student fails to meet an EdoMex school's cutoff score, which in both cases increases the probability of DF high school assignment from 0 to 1 at the cutoff. The coefficient δ_0 is the average effect of DF assignment for marginal students in the 2005 cohort. Coefficients δ_1 and δ_2 are, respectively, the differences in this average effect for cohorts partially and fully exposed to Prepa Sí (the difference-in-discontinuities parameters). I follow Grembi, Nannicini, and Troiano (2016) in using a linear spline to control for centered exam score. A non-parametric approach using Calonico, Cattaneo, and Titiunik (2014) or related methods for bandwidth selection is not well-justified given that the running variable is discrete and quite coarse: the COMIPEMS exam score consists of integer values and has a standard deviation of about 18. I instead report the difference-in-discontinuities estimates for a wide range of reasonable bandwidths to show the sensitivity of the results to bandwidth choice. Local linear polynomials are allowed to have different slopes on either side of the cutoff and to vary based on whether the cutoff school is in the DF or EdoMex.[2] The triangular

[2]That the empirical specification pools admission effects across DF and EdoMex cutoff subsamples, each

kernel is used to weight observations with respect to centered exam score and standard errors are clustered at the middle school and assigned high school levels. Because of the discrete support, I use a t-distribution with $G - 1$ degrees of freedom for hypothesis testing, where G is the number of distinct values of the running variable.[3]

Interpretation of the difference-in-discontinuities estimates depends on the subsample of students for which they are estimated. Estimating Equation D.1 for the subsample of DF residents gives the average effect of Prepa Sí eligibility on high school completion for students near the margin of DF or EdoMex school assignment. Limiting the subsample to EdoMex residents provides estimates of the average spillover effects of Prepa Sí on ineligible students in DF high schools, again for students near this particular margin. While the RD design relaxes the assumption of parallel trends between geographically proximate DF and EdoMex students, it does not allow for high school-by-cohort shocks as in the difference-in-differences approach. Sample sizes are also much smaller, particularly for the subsample of DF residents used in estimating the effect of Prepa Sí eligibility.

D.2 Results

The difference-in-discontinuities results fail to find evidence that Prepa Sí eligibility affects high school completion, although the estimated effects have large standard errors compared to the main difference-in-differences estimates. Results for DF residents are presented in Table D.1, Panel A. Columns 1 through 3 indicate, for bandwidths from 5 to 20 entrance exam points (about one standard deviation), that crossing the admission cutoff for DF assignment results in an increase of the probability of assignment to a DF high school from 0 to 1. This is simply a specification check since the sample consists of students near the margin of assignment to the DF.[4] This holds for partially- and fully-exposed cohorts. The sample size is small compared to the difference-in-differences sample. This is because

with their own set of local linear polynomials, presents a further complication in implementing bandwidth selection procedures.

[3]This approach follows guidance in Cameron and Miller (2015) for inference with few clusters. The honest confidence interval approach put forth by Kolesár and Rothe (2018) is not directly implementable here for several reasons, including the clustered standard errors and inclusion of cutoff-by-year fixed effects as controls.

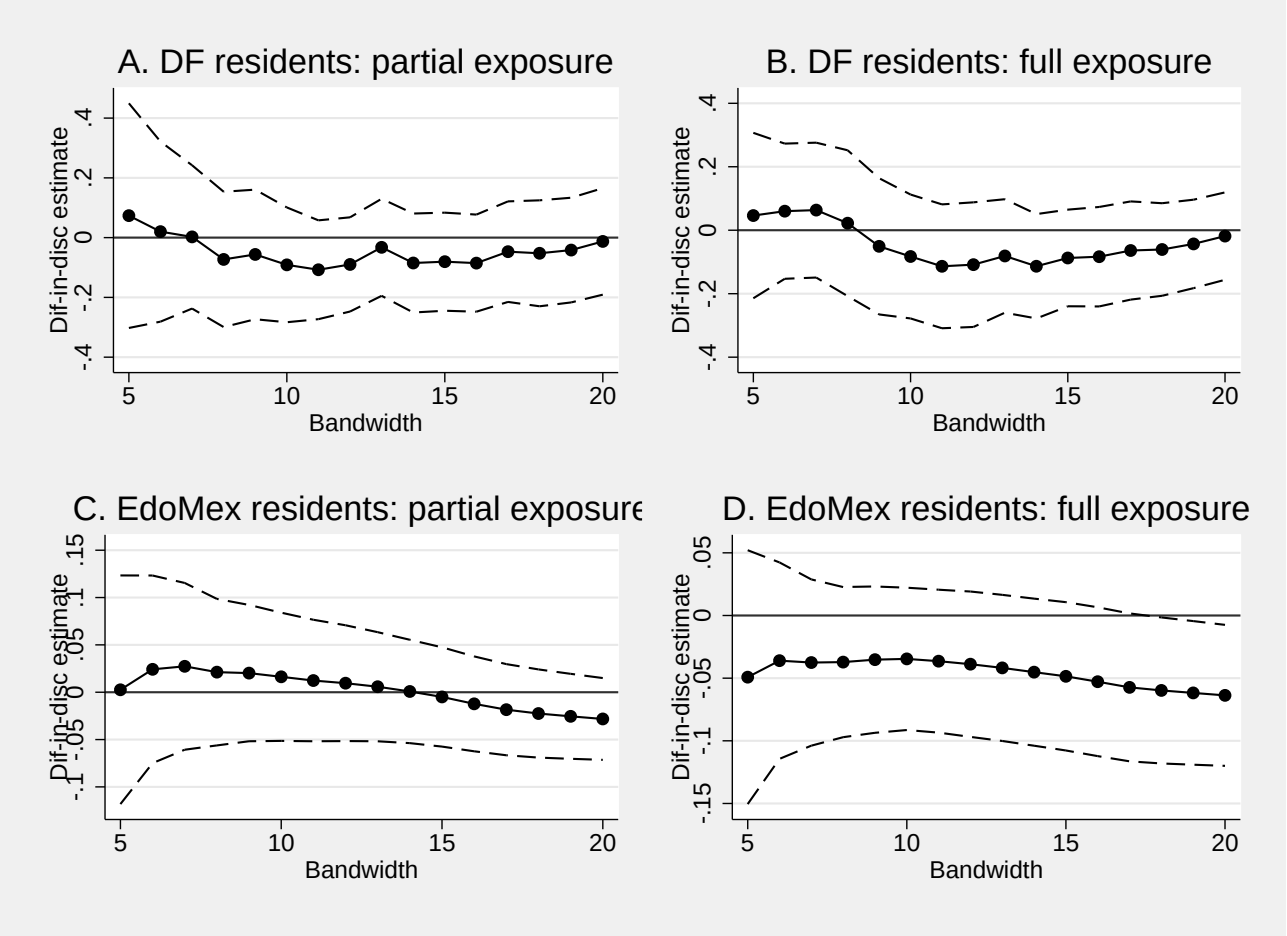
[4]The coefficient estimate is not identical to 1 because a small number of students are in the cutoff sample for a DF (EdoMex) high school but would be assigned to an EdoMex (DF) school for some point value above the cutoff score. For example, suppose that a student's stated school choices were such that for scores below 50, he would be assigned to School A in EdoMex, for scores 50 to 53 he would be assigned to the cutoff school, School B in the DF, and for scores 54 and above he would be assigned to School C in EdoMex. Suppose that his score was 55. Then even though he scored high enough for School B (and thus DF) admission, he was actually assigned to School C (EdoMex). The linear splines in the difference-in-discontinuities regression do not perfectly fit this declining probability of DF assignment away from the cutoff.

relatively few DF students apply to EdoMex high schools and, among those who do, few have combinations of reported preferences and exam scores that place them near a cutoff between DF and EdoMex schools.

Columns 4 through 6 of Panel A show the difference-in-discontinuities estimates of the effect of Prepa Sí on ENLACE-taking, again for bandwidths 5, 10, and 20. The baseline effect of DF admission for these marginal students is imprecisely estimated and sensitive to bandwidth choice. In all cases the estimated effect is insignificant at conventional levels. The differential effects for partially- and fully-exposed students are also imprecise, statistically insignificant, and sensitive to bandwidth choice. Even for the widest reported bandwidth, the 95% confidence interval contains effects for fully-exposed cohorts between -16 and 12 percentage points. Figure D.1, Panels A and B plot out these difference-in-discontinuities coefficients for every bandwidth from 5 to 20, showing that all estimates are statistically insignificant at the 5% level and, for most bandwidths, are slightly negative.

Panel B of Table D.1 shows the estimated spillover effects of Prepa Sí for EdoMex residents assigned to DF high schools, accompanied by Panels C and D in Figure D.1. The sample size is larger than for DF residents because many EdoMex students list both DF and EdoMex high schools in their preferences. The estimated baseline effect of DF school assignment on ENLACE-taking is negative, ranging from -7.5 to -4.0 percentage points for bandwidths of 10 and 20, respectively. These DF assignment penalties could be caused by characteristics of the assigned schools themselves or factors related to student-school match, for example due to a smaller pre-existing social network at the DF school. The difference-in-discontinuities estimates, while more precise than for the DF resident sample, are still sensitive to bandwidth choice and, in most cases, 95% confidence intervals include zero. The estimates for partial exposure are small and positive for small bandwidths, but small and negative for bandwidths above 15. In all cases, we cannot rule out a zero effect at the 5% confidence level. For fully-exposed cohorts, point estimates are negative, near -5 percentage points, and become statistically significant at the 5% level for bandwidths near 20. This provides some suggestive evidence that, for marginal EdoMex students most exposed to Prepa Sí, spillovers were if anything *negative*, contrary to the hypothesis that Prepa Sí freed up institutional resources that were then targeted toward EdoMex students. This is consistent with the small, negative, statistically insignificant estimated effect of DF high school assignment for EdoMex students presented in the difference-in-differences analysis.

Figure D.1: Difference-in-discontinuities effects of DF high school assignment on ENLACE-taking rates for different bandwidths



Note: Estimated coefficients correspond to estimates from Equation D.1 for the bandwidth on the horizontal axis. Coefficients represent difference in estimated RD effect of DF high school assignment on ENLACE-taking between the cohort with the indicated exposure level and the base cohort (2005). Dashed lines correspond to 95% confidence intervals.

Table D.1: Difference-in-discontinuities effects of DF high school assignment on ENLACE-taking rates

Panel A. DF residents	(1) Assigned to DF high school	(2) Assigned to DF high school	(3) Assigned to DF high school	(4) Took ENLACE	(5) Took ENLACE	(6) Took ENLACE
Admitted to DF HS	1.019***	1.039***	1.080***	−0.140	0.028	0.014
	(0.0138)	(0.0234)	(0.0320)	(0.1272)	(0.0902)	(0.0731)
Admitted to DF HS × partial exposure	0.016	0.011	−0.042	0.074	−0.091	−0.013
	(0.0213)	(0.0300)	(0.0300)	(0.1630)	(0.0914)	(0.0880)
Admitted to DF HS × full exposure	−0.005	−0.004	−0.029	0.046	−0.083	−0.018
	(0.0189)	(0.0241)	(0.0311)	(0.1130)	(0.0929)	(0.0680)
Observations	2935	5519	6783	2935	5519	6783
Adjusted R^2	0.957	0.889	0.770	0.084	0.088	0.106
Mean of dependent variable	0.54	0.54	0.55	0.49	0.49	0.47
Bandwidth	5	10	20	5	10	20
Panel B. EdoMex residents	(1) Assigned to DF high school	(2) Assigned to DF high school	(3) Assigned to DF high school	(4) Took ENLACE	(5) Took ENLACE	(6) Took ENLACE
Admitted to DF HS	1.022***	1.074***	1.022***	−0.063*	−0.075***	−0.040*
	(0.0081)	(0.0224)	(0.0418)	(0.0367)	(0.0235)	(0.0234)
Admitted to DF HS × partial exposure	−0.008	−0.035**	−0.013	0.003	0.016	−0.028
	(0.0096)	(0.0156)	(0.0123)	(0.0524)	(0.0322)	(0.0213)
Admitted to DF HS × full exposure	−0.004	−0.043***	0.003	−0.049	−0.035	−0.064**
	(0.0062)	(0.0119)	(0.0164)	(0.0439)	(0.0270)	(0.0278)
Observations	21058	44141	82460	21058	44141	82460
Adjusted R^2	0.935	0.838	0.716	0.069	0.058	0.046
Mean of dependent variable	0.50	0.46	0.42	0.55	0.54	0.53
Bandwidth	5	10	20	5	10	20

Note: Estimated coefficients are from difference-in-discontinuities specification in Equation D.1. Interaction terms correspond to the difference-in-discontinuity parameters, compared to the base cohort (2005). Sample is restricted to students in the speicified centered COMIPEMS score bandwidth. The p-values are based on a t-distribution with $G - 1$ degrees of freedom, where G is the number of distinct values of the running variable. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the high school and middle school levels in parentheses.