The Unintended Effects of an Intensive Margin Reform to Student Loans on Educational Attainment

Online Appendix

Pinjas Albagli
London School of Economics

Andrés García-Echalar
Universidad de los Andes

February, 2023

Contents

A Replication of Main Results Excluding Year 2008

This appendix replicates our analysis in the main text excluding cohort 2008 for immediate enrollment and cohorts 2007–2008 for two-year enrollment and second-year dropout. Our main results remain virtually unchanged while the evidence supporting the parallel trends assumption is stronger. The numbering of tables and figures replicates that of the main text to facilitate comparisons.

List of Figures

List of Tables

Table 3: Immediate Enrollment

	HES			Universities			Vocational		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Eligible × exposed	-0.001	-0.002	-0.003	0.025***	0.025***	0.023***	-0.026***	-0.027***	-0.027***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
Exposed	0.062***	0.071***	0.078***	-0.012***	-0.035***	-0.031***	0.074***	0.105***	0.108***
	(0.003)	(0.007)	(0.006)	(0.001)	(0.007)	(0.007)	(0.003)	(0.004)	(0.004)
Eligible	0.261***	0.261***	0.243***	0.291***	0.291***	0.271***	-0.029***	-0.029***	-0.028***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)
Cohort effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Control variables	No	No	Yes	No	No	Yes	No	No	Yes
Observations	1,353,980	1,353,980	1,353,980	1,353,980	1,353,980	1,353,980	1,353,980	1,353,980	1,353,980
Control group size	508,298	508,298	508,298	508,298	508,298	508,298	508,298	508,298	508,298
Outcome mean	0.536	0.536	0.536	0.355	0.355	0.355	0.182	0.182	0.182

Notes: Clustered standard errors at the class level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. School level control variables include indicators of school type, rural area and geographical region. Student level control variables include gender, attendance rate, district and number of family members at different levels in the education system. Control group size accounts for the number of ineligible individuals in the exposure period, while Outcome mean refers to the mean of the dependent variable of those individuals.

Figure 1: Dynamics of Immediate Enrollment

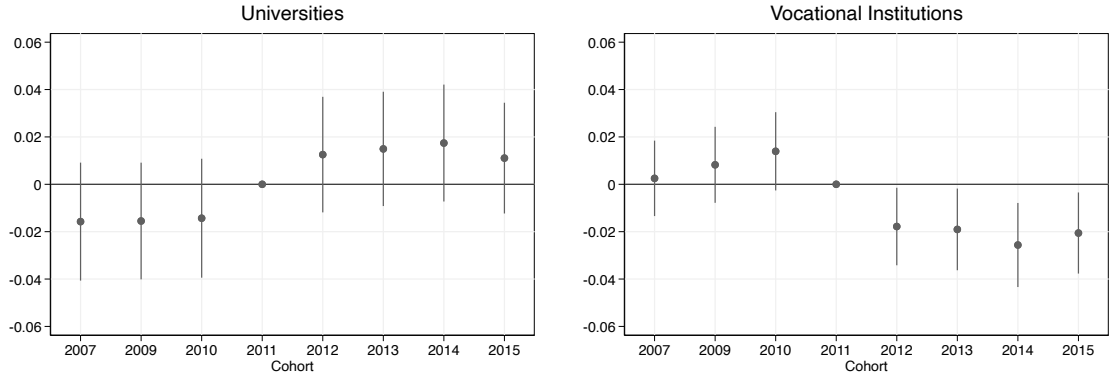


Table 4: Two-Year Enrollment

	HES			Universities			Vocational		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Eligible × exposed (2nd year)	0.015***	0.014***	0.012***	0.023***	0.023***	0.021***	-0.011***	-0.012***	-0.012***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.003)	(0.003)	(0.003)
Exposed (2nd year)	0.035***	0.054***	0.057***	-0.005***	-0.009	-0.007	0.039***	0.063***	0.065***
	(0.004)	(0.006)	(0.006)	(0.001)	(0.006)	(0.006)	(0.003)	(0.004)	(0.004)
Eligible	0.283***	0.283***	0.261***	0.270***	0.270***	0.250***	0.010***	0.010***	0.008***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
Cohort effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Control variables	No	No	Yes	No	No	Yes	No	No	Yes
Observations	1,035,551	1,035,551	1,035,551	1,035,568	1,035,568	1,035,568	1,037,137	1,037,137	1,037,137
Control group size	263,400	263,400	263,400	263,402	263,402	263,402	263,959	263,959	263,959
Outcome mean	0.474	0.474	0.474	0.310	0.310	0.310	0.151	0.151	0.151

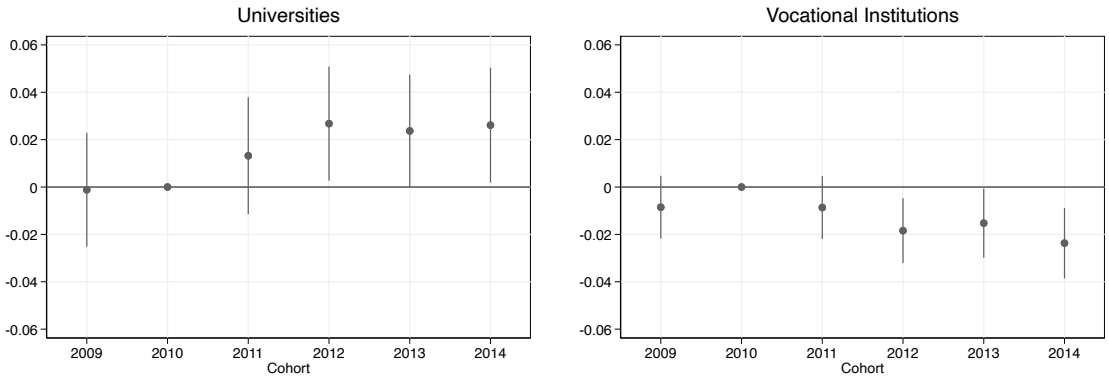
Notes: Clustered standard errors at the class level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. School level control variables include indicators of school type, rural area and geographical region. Student level control variables include gender, attendance rate, district and number of family members at different levels in the education system. Control group size accounts for the number of ineligible individuals in the exposure period, while Outcome mean refers to the mean of the dependent variable of those individuals.

Table 5: Second-Year Dropout

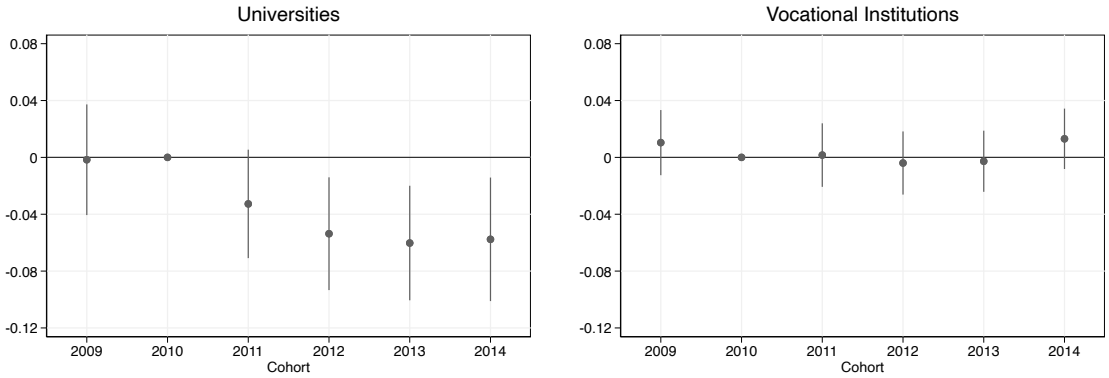
	HES			Universities			Vocational		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Eligible × exposed (2nd year)	-0.014***	-0.014***	-0.014***	-0.049***	-0.049***	-0.050***	-0.003	-0.003	-0.008
	(0.005)	(0.005)	(0.004)	(0.009)	(0.009)	(0.009)	(0.005)	(0.005)	(0.005)
Exposed (2nd year)	0.018***	0.023***	0.034***	0.057***	0.056***	0.070***	0.006	0.007	0.037***
	(0.005)	(0.005)	(0.005)	(0.010)	(0.010)	(0.010)	(0.005)	(0.006)	(0.006)
Eligible	-0.182***	-0.182***	-0.125***	-0.221***	-0.221***	-0.156***	-0.143***	-0.143***	-0.118***
	(0.004)	(0.004)	(0.004)	(0.008)	(0.008)	(0.007)	(0.004)	(0.004)	(0.004)
Cohort effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Control variables	No	No	Yes	No	No	Yes	No	No	Yes
Observations	526,147	526,147	521,391	301,826	301,826	297,440	224,857	224,857	224,480
Control group size	139,280	139,280	139,280	90,748	90,748	90,748	48,669	48,669	48,669
Outcome mean	0.103	0.103	0.103	0.102	0.102	0.102	0.180	0.180	0.180

Notes: Clustered standard errors at the class level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. School level control variables include indicators of school type, rural area and geographical region. Student level control variables include gender, attendance rate, district and number of family members at different levels in the education system. Program characteristics include duration, annual fee, and an indicator for accreditation. Control group size accounts for the number of ineligible individuals in the exposure period, while Outcome mean refers to the mean of the dependent variable of those individuals.

Figure 2: Dynamics of Persistence and Retention



(a) Two-Year Enrollment



(b) Second-Year Dropout

Table 6: Difference-in-Discontinuities Design: Immediate Enrollment

	All students			GPA < 5.3		
	HES	Universities	Vocational	HES	Universities	Vocational
	(1)	(2)	(3)	(4)	(5)	(6)
Difference	0.011*	0.023***	-0.006	0.003	0.025**	-0.024**
	(0.006)	(0.006)	(0.006)	(0.013)	(0.010)	(0.012)
Exposed	0.074***	0.127***	-0.048***	0.062***	0.084***	-0.024***
	(0.004)	(0.005)	(0.005)	(0.009)	(0.007)	(0.009)
Unexposed	0.063***	0.104***	-0.042***	0.059***	0.059***	-0.001
	(0.005)	(0.004)	(0.004)	(0.009)	(0.007)	(0.007)
Bandwidth						
Exposed	51.257	36.629	41.201	48.882	47.259	43.601
Unexposed	50.574	42.220	52.172	42.668	45.042	53.310
Observations						
Exposed	117,087	84,280	94,582	27,136	26,254	24,260
Unexposed	113,523	95,218	116,986	27,528	29,019	34,149

Notes: Optimal bandwidths separately selected by exposure. Triangular kernel is used for local linear regressions. SUEST standard errors clustered at the class level in parentheses. *** p< 0.01, ** p< 0.05, * p< 0.1.

Table 7: Difference-in-Discontinuities Design: Two-Year Enrollment

	All students			GPA < 5.3		
	HES	Universities	Vocational	HES	Universities	Vocational
	(1)	(2)	(3)	(4)	(5)	(6)
Difference	0.019**	0.015**	0.004	0.041**	0.028**	0.006
	(0.007)	(0.007)	(0.007)	(0.016)	(0.011)	(0.013)
Exposed	0.076***	0.107***	-0.038***	0.080***	0.072***	0.000
	(0.004)	(0.004)	(0.005)	(0.008)	(0.006)	(0.008)
Unexposed	0.057***	0.092***	-0.042***	0.040***	0.044***	-0.006
	(0.006)	(0.005)	(0.005)	(0.014)	(0.010)	(0.010)
Bandwidth						
Exposed	58.077	37.934	38.461	64.812	51.634	48.703
Unexposed	53.155	50.305	48.858	35.814	44.375	45.506
Observations						
Exposed	133,494	88,264	89,627	38,607	31,115	29,424
Unexposed	61,536	58,343	56,759	11,691	14,464	14,832

Notes: Optimal bandwidths separately selected by exposure. Triangular kernel is used for local linear regressions. SUEST standard errors clustered at the class level in parentheses. *** p< 0.01, ** p< 0.05, * p< 0.1.

Table 8: Difference-in-Discontinuities Design: Second-Year Dropout

	All students			GPA < 5.3		
	HES	Universities	Vocational	HES	Universities	Vocational
	(1)	(2)	(3)	(4)	(5)	(6)
Difference	-0.009	0.000	-0.025**	-0.025	-0.019	-0.051**
	(0.007)	(0.013)	(0.011)	(0.017)	(0.031)	(0.025)
Exposed	-0.008*	-0.017**	0.002	-0.029***	-0.009	-0.038**
	(0.004)	(0.008)	(0.006)	(0.010)	(0.018)	(0.016)
Unexposed	0.001	-0.018	0.026***	-0.004	0.010	0.013
	(0.006)	(0.011)	(0.009)	(0.014)	(0.025)	(0.019)
Bandwidth						
Exposed	54.348	51.297	46.782	50.644	54.499	31.156
Unexposed	50.402	49.227	50.556	41.813	38.712	50.574
Observations						
Exposed	69,669	30,248	32,749	15,517	6,386	5,968
Unexposed	28,947	13,368	15,330	6,267	2,471	4,341

Notes: Optimal bandwidths separately selected by exposure. Triangular kernel is used for local linear regressions. SUEST standard errors clustered at the class level in parentheses. *** p< 0.01, ** p< 0.05, * p< 0.1.

Table 9: Heterogeneity of Main Results by Student Sex

	HES			Universities			Vocational		
	Female	Male	Difference	Female	Male	Difference	Female	Male	Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Immediate Enrollment	-0.011**	0.001	-0.012**	0.021***	0.022***	-0.001	-0.032***	-0.021***	-0.011**
	(0.004)	(0.005)	(0.006)	(0.005)	(0.007)	(0.007)	(0.004)	(0.004)	(0.005)
	[720,112]	[633,868]	[1,353,980]	[720,112]	[633,868]	[1,353,980]	[720,112]	[633,868]	[1,353,980]
	{0.51}	{0.51}	{0}	{0.30}	{0.29}	{.01}	{0.21}	{0.22}	{-.01}
Two-Year Enrollment	0.005	0.018***	-0.013*	0.019***	0.022***	-0.003	-0.018***	-0.007*	-0.011**
	(0.005)	(0.007)	(0.007)	(0.006)	(0.008)	(0.009)	(0.004)	(0.004)	(0.005)
	[550,288]	[485,263]	[1,035,551]	[550,302]	[485,266]	[1,035,568]	[551,073]	[486,064]	[1,037,137]
	{0.45}	{0.44}	{.01}	{0.26}	{0.25}	{.01}	{0.17}	{0.17}	{0}
Second-Year Dropout	-0.013**	-0.016***	0.003	-0.041***	-0.057***	0.016	-0.005	-0.011*	0.007
	(0.006)	(0.006)	(0.008)	(0.013)	(0.013)	(0.018)	(0.007)	(0.007)	(0.010)
	[276,309]	[245,082]	[521,391]	[158,711]	[138,729]	[297,440]	[117,898]	[106,582]	[224,480]
	{0.12}	{0.14}	{-.02}	{0.11}	{0.13}	{-.02}	{0.21}	{0.23}	{-.02}
Cohort effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

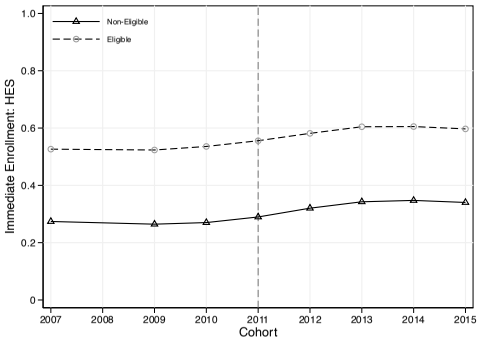
Notes: SUEST standard errors clustered at the class level in parentheses. Sample sizes in square brackets. Outcome sample means in curly braces. *** p<0.01, ** p<0.05, * p<0.1. School level control variables include indicators of school type, rural area and geographical region. Student level control variables include attendance rate, district and number of family members at different levels in the education system.

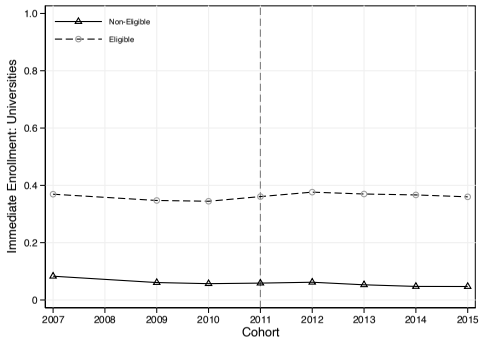
Table 10: Heterogeneity of Main Results by School Type

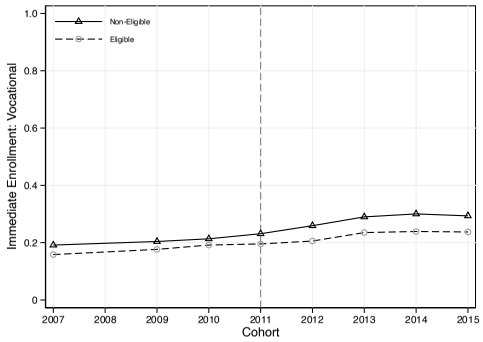
	HES			Universities			Vocational		
	Public	Voucher	Difference	Public	Voucher	Difference	Public	Voucher	Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Immediate Enrollment	0.002	-0.005	0.007	0.010	0.028***	-0.018*	-0.008	-0.033***	0.025***
	(0.006)	(0.004)	(0.007)	(0.008)	(0.005)	(0.010)	(0.005)	(0.004)	(0.007)
	[530,018]	[823,962]	[1,353,980]	[530,018]	[823,962]	[1,353,980]	[530,018]	[823,962]	[1,353,980]
	{0.46}	{0.54}	{-.08}	{0.24}	{0.33}	{-.09}	{0.23}	{0.21}	{.02}
Two-Year Enrollment	0.007	0.016***	-0.009	0.010	0.025***	-0.015	-0.004	-0.013***	0.009
	(0.008)	(0.005)	(0.009)	(0.010)	(0.006)	(0.012)	(0.005)	(0.004)	(0.006)
	[402,810]	[632,741]	[1,035,551]	[402,815]	[632,753]	[1,035,568]	[403,430]	[633,707]	[1,037,137]
	{0.40}	{0.48}	{-.08}	{0.21}	{0.29}	{-.08}	{0.18}	{0.17}	{.01}
Second-Year Dropout	-0.007	-0.020***	0.013	-0.055***	-0.050***	-0.005	-0.001	-0.012*	0.011
	(0.008)	(0.006)	(0.009)	(0.017)	(0.011)	(0.020)	(0.008)	(0.006)	(0.011)
	[185,603]	[335,788]	[521,391]	[94,366]	[203,074]	[297,440]	[91,470]	[133,010]	[224,480]
	{0.15}	{0.12}	{.03}	{0.13}	{0.11}	{.02}	{0.23}	{0.21}	{.02}
Cohort effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: SUEST standard errors clustered at the class level in parentheses. Sample sizes in square brackets. Outcome sample means in curly braces. *** p<0.01, ** p<0.05, * p<0.1. School level control variables include indicators of school type, rural area and geographical region. Student level control variables include attendance rate, district and number of family members at different levels in the education system.

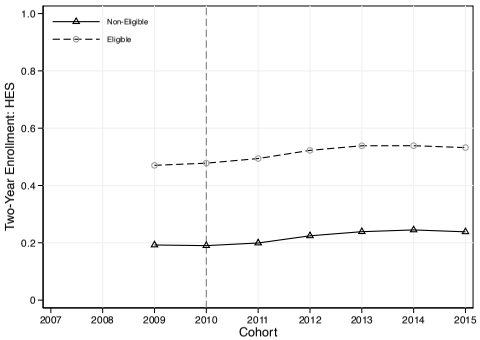
Figure A.1: Outcomes over Time by Eligibility

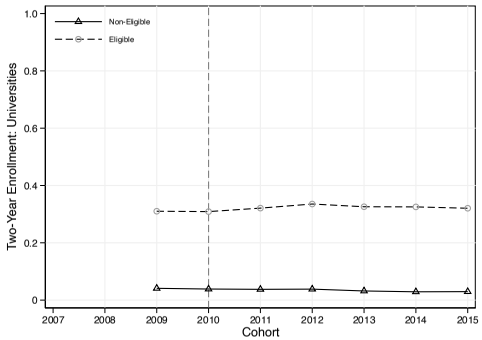


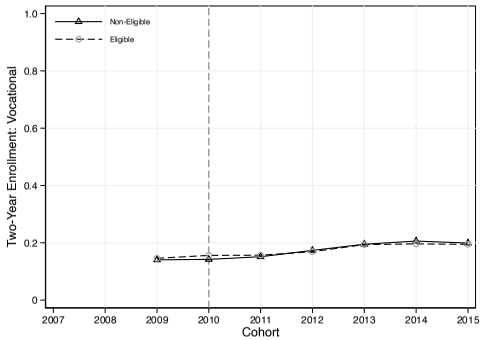




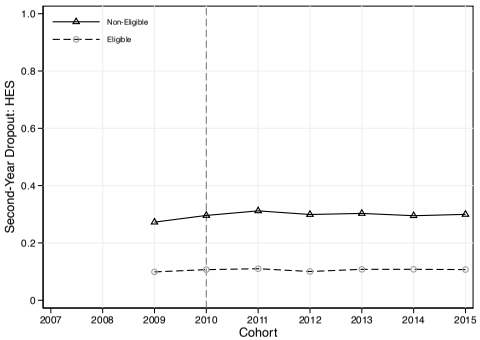
Panel A: Immediate Enrollment

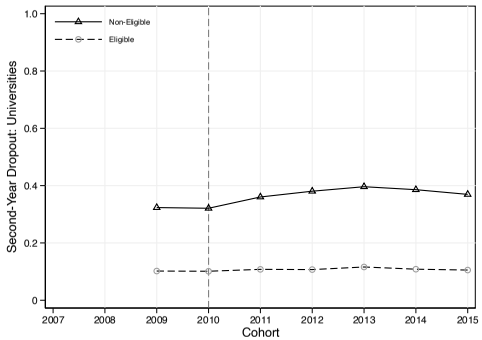


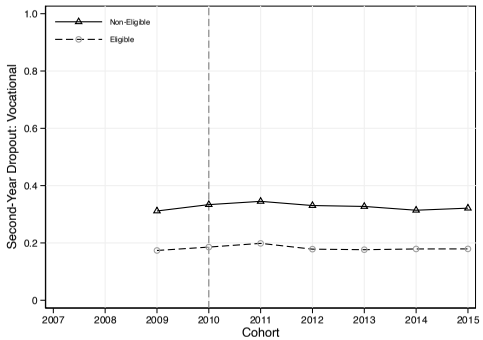




Panel B: Two-Year Enrollment







Panel C: Second-Year Dropout

Table A.1: Dynamics Excluding Year 2008

	Immediate Enrollment				**Two-Year Enrollment**				**Second-Year Dropout**			
	Universities		Vocational		Universities		Vocational		Universities		Vocational	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Eligible × cohort 2007	-0.016	-0.017*	0.003	0.004								
	(0.010)	(0.010)	(0.006)	(0.006)								
Eligible × cohort 2009	-0.015	-0.015	0.008	0.008	-0.001	-0.001	-0.009*	-0.009*	-0.002	-0.003	0.010	0.010
	(0.010)	(0.010)	(0.006)	(0.006)	(0.009)	(0.009)	(0.005)	(0.005)	(0.015)	(0.015)	(0.009)	(0.009)
Eligible × cohort 2010	-0.014	-0.015	0.014**	0.015**								
	(0.010)	(0.010)	(0.006)	(0.006)								
Eligible × cohort 2011					0.013	0.014	-0.009*	-0.009*	-0.033**	-0.049***	0.002	-0.002
					(0.010)	(0.010)	(0.005)	(0.005)	(0.015)	(0.015)	(0.009)	(0.008)
Eligible × cohort 2012	0.013	0.011	-0.018***	-0.016***	0.027***	0.026***	-0.018***	-0.018***	-0.054***	-0.057***	-0.004	-0.007
	(0.009)	(0.009)	(0.006)	(0.006)	(0.009)	(0.009)	(0.005)	(0.005)	(0.015)	(0.015)	(0.009)	(0.008)
Eligible × cohort 2013	0.015	0.012	-0.019***	-0.018***	0.024***	0.021**	-0.015***	-0.015***	-0.060***	-0.057***	-0.003	-0.010
	(0.009)	(0.009)	(0.007)	(0.007)	(0.009)	(0.009)	(0.006)	(0.006)	(0.016)	(0.015)	(0.008)	(0.008)
Eligible × cohort 2014	0.017*	0.016	-0.026***	-0.025***	0.026***	0.025***	-0.024***	-0.024***	-0.058***	-0.038**	0.013	0.006
	(0.010)	(0.010)	(0.007)	(0.007)	(0.009)	(0.010)	(0.006)	(0.006)	(0.017)	(0.017)	(0.008)	(0.008)
Eligible × cohort 2015	0.011	0.009	-0.021***	-0.020***								
	(0.009)	(0.009)	(0.007)	(0.007)								
Eligible	0.302***	0.283***	-0.036***	-0.035***	0.270***	0.250***	0.014***	0.012***	-0.220***	-0.155***	-0.148***	-0.122***
	(0.007)	(0.007)	(0.004)	(0.004)	(0.007)	(0.007)	(0.004)	(0.004)	(0.011)	(0.011)	(0.006)	(0.006)
Cohort 2007	0.024***	0.023***	-0.040***	-0.043***								
	(0.003)	(0.004)	(0.006)	(0.005)								
Cohort 2009	0.002	0.002	-0.027***	-0.028***	0.002	0.001	-0.002	-0.001	0.002	0.004	-0.022**	-0.026***
	(0.003)	(0.003)	(0.006)	(0.005)	(0.002)	(0.003)	(0.004)	(0.004)	(0.015)	(0.015)	(0.009)	(0.008)
Cohort 2010	-0.002	0.000	-0.018***	-0.019***								
	(0.003)	(0.003)	(0.006)	(0.005)								
Cohort 2011					-0.001	-0.003	0.009**	0.010**	0.039***	0.066***	0.011	0.022***
					(0.002)	(0.003)	(0.004)	(0.004)	(0.015)	(0.015)	(0.009)	(0.008)
Cohort 2012	0.003	0.005	0.028***	0.031***	-0.000	0.001	0.031***	0.036***	0.060***	0.066***	-0.003	0.002
	(0.003)	(0.003)	(0.005)	(0.005)	(0.002)	(0.003)	(0.004)	(0.004)	(0.016)	(0.015)	(0.009)	(0.009)
Cohort 2013	-0.006**	-0.005	0.059***	0.059***	-0.007***	-0.008**	0.053***	0.054***	0.075***	0.087***	-0.006	0.019**
	(0.002)	(0.003)	(0.006)	(0.005)	(0.002)	(0.003)	(0.005)	(0.004)	(0.016)	(0.016)	(0.009)	(0.008)
Cohort 2014	-0.012***	-0.010***	0.069***	0.071***	-0.010***	-0.009***	0.064***	0.067***	0.065***	0.061***	-0.020**	0.008
	(0.002)	(0.003)	(0.006)	(0.006)	(0.002)	(0.003)	(0.005)	(0.005)	(0.017)	(0.017)	(0.009)	(0.008)
Cohort 2015	-0.012***	-0.010***	0.062***	0.063***								
	(0.002)	(0.003)	(0.006)	(0.005)								
Student district fixed effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Control variables	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	1,353,980	1,353,980	1,353,980	1,353,980	1,035,568	1,035,568	1,037,137	1,037,137	301,826	297,440	224,857	224,480
Pre-trends p-value	0.295	0.260	0.126	0.113	0.900	0.955	0.098	0.072	0.913	0.858	0.243	0.241

Notes: Clustered standard errors at the class level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. School level control variables include indicators of school type, rural area and geographical region. Student level control variables include gender, attendance rate, district and number of family members at different levels in the education system. Control group size accounts for the number of ineligible individuals in the exposure period, while Outcome mean refers to the mean of the dependent variable of those individuals.

Table A.2: Placebo Reform Excluding Year 2008

	HES			Universities			Vocational		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Immediate Enrollment									
Eligible × exposed (placebo)	0.006	0.006	0.007	0.000	0.000	0.002	0.006	0.006	0.005
	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.006)	(0.006)	(0.006)
Observations	307,308	307,308	307,308	307,308	307,308	307,308	307,308	307,308	307,308
Two-Year Enrollment									
Eligible × exposed (placebo)	0.002	0.002	0.003	-0.004	-0.004	-0.003	0.006	0.006	0.005
	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.005)	(0.005)	(0.005)
Observations	306,858	306,858	306,858	306,859	306,859	306,859	307,307	307,307	307,307
Second-Year Dropout									
Eligible × exposed (placebo)	0.002	0.002	0.008	0.009	0.009	0.021	-0.007	-0.007	-0.001
	(0.008)	(0.008)	(0.008)	(0.015)	(0.015)	(0.015)	(0.009)	(0.009)	(0.009)
Observations	143,421	143,421	140,619	89,088	89,088	86,634	54,507	54,507	54,152
Cohort effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Control variables	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Clustered standard errors at the class level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Control variables are the same as in Tables 3, 4, and 5.

Table B.1: IV-DiD Regressions for Two-Year Outcomes

	HES			Universities			Vocational		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Two-Year Enrollment									
Eligible × exposed (2nd year)	0.014***	0.013***	0.011**	0.022***	0.022***	0.021***	-0.011***	-0.013***	-0.012***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.003)	(0.003)	(0.003)
Cohort effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Control variables	No	No	Yes	No	No	Yes	No	No	Yes
Observations	1,035,551	1,035,551	1,035,551	1,035,568	1,035,568	1,035,568	1,037,137	1,037,137	1,037,137
Cragg-Donald	10,761,691	10,763,467	10,576,213	10,762,107	10,763,884	10,576,618	10,783,649	10,785,436	10,597,726
Second-Year Dropout									
Eligible × exposed (2nd year)	-0.014***	-0.014***	-0.014***	-0.049***	-0.049***	-0.050***	-0.003	-0.003	-0.008
	(0.005)	(0.005)	(0.005)	(0.010)	(0.010)	(0.009)	(0.005)	(0.005)	(0.005)
Cohort effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Control variables	No	No	Yes	No	No	Yes	No	No	Yes
Observations	526,147	526,147	521,391	301,826	301,826	297,440	224,857	224,857	224,480
Cragg-Donald	40,887,766	40,887,893	38,270,537	11,742,481	11,741,672	10,985,932	22,570,275	22,567,295	21,934,238

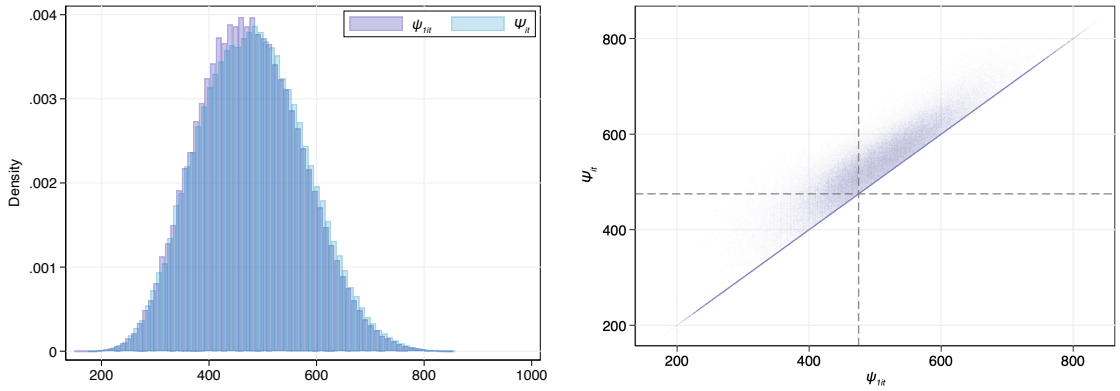
Notes: 2SLS estimates instrumenting eligible_{it} with eligible_{2it}. Clustered standard errors at the class level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Control variables are the same as in Tables 3, 4, and 5.

Table B.2: IV-Diff-in-Disc Design for Two-Year Outcomes

	Two-Year Enrollment						**Second-Year Dropout**					
	All students			GPA < 5.3			All students			GPA < 5.3		
	HES	Universities	Vocational	HES	Universities	Vocational	HES	Universities	Vocational	HES	Universities	Vocational
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Difference	0.024***	0.024***	0.001	0.046**	0.033**	0.005**	-0.009**	-0.000**	-0.025**	-0.025**	-0.020**	-0.051**
	(0.009)	(0.008)	(0.008)	(0.019)	(0.014)	(0.015)	(0.007)	(0.014)	(0.011)	(0.017)	(0.031)	(0.025)
Exposed	0.097***	0.142***	-0.051***	0.095***	0.088***	-0.002***	-0.008**	-0.019**	0.001***	-0.029***	-0.009***	-0.038**
	(0.005)	(0.005)	(0.005)	(0.009)	(0.007)	(0.009)	(0.004)	(0.008)	(0.006)	(0.010)	(0.019)	(0.016)
Unexposed	0.073***	0.117***	-0.052***	0.049***	0.054***	-0.007	0.001***	-0.019*	0.027***	-0.004***	0.010***	0.013
	(0.008)	(0.007)	(0.006)	(0.017)	(0.011)	(0.012)	(0.006)	(0.011)	(0.009)	(0.014)	(0.025)	(0.019)
Bandwidth												
Exposed	58.077	37.934	38.461	64.812	51.634	48.703	54.348	51.297	46.782	50.644	54.499	31.156
Unexposed	53.155	50.305	48.858	35.814	44.375	45.506	50.402	49.227	50.556	41.813	38.712	50.574
Observations												
Exposed	133,494	88,264	89,627	38,607	31,115	29,424	69,669	30,248	32,749	15,517	6,386	5,968
Unexposed	61,536	58,343	56,759	11,691	14,464	14,832	28,947	13,368	15,330	6,267	2,471	4,341

Notes: Optimal bandwidths separately selected by exposure. Triangular kernel is used in local linear regressions. Standard errors clustered at the class level in parentheses. *** p< 0.01, ** p< 0.05, * p< 0.1.

Figure B.1: Immediate and Two-Year-Best PSU Scores



(a) Histograms

(b) Scatter Plot

Table B.3: Immediate vs Two-Year Eligibility

	Two-year eligibility		
	Yes	No	Total
Immediate eligibility			
Yes	77.27%	0%	77.27%
No	0.76%	21.97%	22.73%
Total	78.03%	21.97%	100.00%

Table C.2: Evidence of Female Delay

	Repetition		Improvement	
	All students	Non-enrolled	All students	Non-enrolled
	(1)	(2)	(3)	(4)
Female × exposed	0.008**	0.034***	0.005	-0.013**
	(0.003)	(0.007)	(0.005)	(0.005)
Female	0.029***	0.060***	0.024***	0.006*
	(0.002)	(0.005)	(0.003)	(0.003)
Exposed	-0.009***	0.004	-0.077***	-0.024***
	(0.003)	(0.007)	(0.004)	(0.004)
Observations	1,023,720	452,286	196,854	155,145

Notes: Clustered standard errors at the class level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Repetition and Improvement are indicator variables. All students comprises the sample of students who sat the PSU immediately after high school graduation. Non-enrolled is the subsample of students that did not enroll immediately. Cohort 2015 is excluded because we do not have access to PSU scores for year 2016.

B IV Details

Our IV linear regression model is given by the structural equation

$$\underbrace{y_{it}}_{1\times 1} = \underbrace{\boldsymbol{x}_{it}'}_{K\times 1}\underbrace{\boldsymbol{\lambda}}_{K\times 1} + \underbrace{\eta_{it}}_{1\times 1} \tag{B.1}$$

and the first stage

$$\underbrace{\boldsymbol{x}_{2it}}_{K_2\times 1} = \underbrace{\boldsymbol{\Gamma}}_{K_2\times L}\underbrace{\boldsymbol{z}_{it}}_{L\times 1} + \underbrace{\boldsymbol{\nu}_{it}}_{K_2\times 1} \tag{B.2}$$

where

$$\underbrace{\boldsymbol{x}_{it}}_{K\times 1} = \begin{pmatrix} \underbrace{\boldsymbol{x}_{1it}}_{K_1\times 1} \\ \underbrace{\boldsymbol{x}_{2it}}_{K_2\times 1} \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{z}_{it} = \begin{pmatrix} \underbrace{\boldsymbol{x}_{1it}}_{K_1\times 1} \\ \underbrace{\boldsymbol{z}_{2it}}_{L_2\times 1} \end{pmatrix}$$

whith $K = K_1 + K_2$ and $L = K_1 + L_2 \geq K$.

Partition

$$\underbrace{\boldsymbol{\Gamma}}_{K_2\times L} = \Big[\underbrace{\boldsymbol{\Gamma}_1}_{K_2\times K_1} \quad \underbrace{\boldsymbol{\Gamma}_2}_{K_2\times L_2}\Big] \qquad \text{and} \qquad \underbrace{\boldsymbol{\lambda}}_{K\times 1} = \begin{bmatrix} \underbrace{\boldsymbol{\lambda}_1}_{K_1\times 1} \\ \underbrace{\boldsymbol{\lambda}_2}_{K_2\times 1} \end{bmatrix}$$

and rewrite Equation (B.2) as

$$\underbrace{\boldsymbol{x}_{2it}}_{K_2\times 1} = \Big[\underbrace{\boldsymbol{\Gamma}_1}_{K_2\times K_1} \quad \underbrace{\boldsymbol{\Gamma}_2}_{K_2\times L_2}\Big] \begin{pmatrix} \underbrace{\boldsymbol{x}_{1it}}_{K_1\times 1} \\ \underbrace{\boldsymbol{z}_{2it}}_{L_2\times 1} \end{pmatrix} + \underbrace{\boldsymbol{\nu}_{it}}_{K_2\times 1}. \tag{B.2'}$$

Now, plugging Equation (B.2') into (B.1), we obtain

$$\begin{aligned} \underbrace{y_{it}}_{1\times 1} &= \underbrace{\Big[\boldsymbol{x}_{1it}' \quad \Big(\boldsymbol{x}_{1it}'\boldsymbol{\Gamma}_1' + \boldsymbol{z}_{2it}'\boldsymbol{\Gamma}_2' + \boldsymbol{\nu}_{it}'\Big)\Big]}_{1\times K}\underbrace{\begin{bmatrix}\boldsymbol{\lambda}_1 \\ \boldsymbol{\lambda}_2\end{bmatrix}}_{K\times 1} + \underbrace{\eta_{it}}_{1\times 1} \\ &= \underbrace{\boldsymbol{x}_{1it}'}_{1\times K_1}\underbrace{\Big(\boldsymbol{\lambda}_1 + \boldsymbol{\Gamma}_1'\boldsymbol{\lambda}_2\Big)}_{K_1\times 1} + \underbrace{\boldsymbol{z}_{2it}'}_{1\times L_2}\underbrace{\boldsymbol{\Gamma}_2'\boldsymbol{\lambda}_2}_{L_2\times 1} + \underbrace{\boldsymbol{\nu}_{it}'\boldsymbol{\lambda}_2 + \eta_{it}}_{1\times 1} \\ &\equiv \underbrace{\boldsymbol{x}_{1it}'}_{1\times K_1}\underbrace{\boldsymbol{\beta}_1}_{K_1\times 1} + \underbrace{\boldsymbol{z}_{2it}'}_{1\times L_2}\underbrace{\boldsymbol{\beta}_2}_{L_2\times 1} + \underbrace{\varepsilon_{it}}_{1\times 1}. \end{aligned}$$

Finally, letting

$$\underbrace{\boldsymbol{\beta}}_{L\times 1} \equiv \begin{bmatrix} \underbrace{\boldsymbol{\beta}_1}_{K_1\times 1} \\ \underbrace{\boldsymbol{\beta}_2}_{L_2\times 1} \end{bmatrix},$$

we obtain the reduced form

$$\underbrace{y_{it}}_{1\times 1} = \underbrace{\boldsymbol{z}_{it}'}_{1\times L}\underbrace{\boldsymbol{\beta}}_{L\times 1} + \underbrace{\varepsilon_{it}}_{1\times 1}. \tag{1}$$

Notice that

$$\mathbb{E}\big[\underbrace{\boldsymbol{z}_{it}}_{L\times 1}\underbrace{\boldsymbol{\nu}_{it}'}_{1\times K_2}\big] = \underbrace{\mathbf{0}}_{L\times K_2}$$

by construction since Equation (B.2) is a linear projection. Therefore,

$$\mathbb{E}\big[\boldsymbol{z}_{it}\,\varepsilon_{it}\big] = 0 \implies \mathbb{E}\big[\boldsymbol{z}_{it}\,\eta_{it}\big] = 0$$

since

$$\mathbb{E}\big[\boldsymbol{z}_{it}\,\varepsilon_{it}\big] = \mathbb{E}\big[\boldsymbol{z}_{it}\,\boldsymbol{\nu}_{it}'\big]\,\boldsymbol{\lambda}_2 + \mathbb{E}\big[\boldsymbol{z}_{it}\,\eta_{it}\big]$$

by definition.

In our DiD-IV setup, the parallel trends assumption underlying our main specification—given by Equation (1)—implies that $\mathbb{E}\big[\boldsymbol{z}_{it}\,\varepsilon_{it}\big] = 0$. Thus, by the argument above, the independence/ignorability requirement for a valid instrument is satisfied for our excluded instruments $\boldsymbol{z}_{2it}$ under the parallel trends assumption.