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 2: Descriptive Statistics

	Cohort									
	2007	2008	2009	2010	2011	2012	2013	2014	2015	Pooled
Immediate Enrollment	0.464		0.464	0.475	0.494	0.521	0.547	0.552	0.549	0.510
by Institution										
University	0.299		0.281	0.279	0.290	0.304	0.301	0.301	0.302	0.294
Vocational	0.167		0.183	0.197	0.204	0.218	0.247	0.252	0.248	0.216
by Gender										
Females	0.478		0.469	0.477	0.493	0.518	0.546	0.556	0.557	0.513
Males	0.453		0.459	0.474	0.494	0.524	0.548	0.549	0.543	0.507
by High School										
Public	0.503		0.492	0.504	0.521	0.551	0.572	0.574	0.571	0.538
Voucher	0.416		0.425	0.435	0.453	0.468	0.503	0.514	0.513	0.465
Two-Year Enrollment			0.406	0.412	0.425	0.454	0.473	0.478		0.441
by Institution										
University			0.248	0.247	0.255	0.267	0.262	0.264		0.257
Vocational			0.145	0.153	0.155	0.170	0.194	0.198		0.169
by Gender										
Females			0.407	0.408	0.419	0.447	0.469	0.477		0.438
Males			0.405	0.416	0.431	0.459	0.477	0.479		0.445
by High School										
Public			0.437	0.444	0.456	0.486	0.502	0.504		0.473
Voucher			0.364	0.368	0.379	0.397	0.422	0.433		0.392
Second-Year Dropout			0.122	0.131	0.138	0.129	0.135	0.132		0.131
by Institution										
University			0.113	0.111	0.120	0.120	0.127	0.117		0.118
Vocational			0.209	0.222	0.237	0.220	0.215	0.212		0.219
by Gender										
Females			0.129	0.143	0.150	0.136	0.141	0.140		0.140
Males			0.116	0.122	0.126	0.122	0.129	0.126		0.124
by High School										
Public			0.110	0.118	0.124	0.118	0.122	0.120		0.119
Voucher			0.141	0.153	0.161	0.151	0.160	0.156		0.154
Eligible	0.755		0.768	0.772	0.767	0.769	0.781	0.794	0.814	0.778
PSU	475.759		475.638	473.877	476.538	475.305	476.784	477.304	479.135	476.311
GPA	5.567		5.582	5.584	5.579	5.593	5.609	5.641	5.681	5.606
Female	0.540		0.536	0.531	0.526	0.534	0.531	0.532	0.528	0.532
Public School	0.442		0.422	0.420	0.405	0.360	0.362	0.364	0.365	0.391
Observations	140,142	0	167,166	175,526	180,774	167,409	173,111	173,168	176,684	1,353,980

Notes: Cohort 2015 is not considered for two-year enrollment and second-year dropout as discussed in the main text. For the same reason, the pooled-sample statistics for these variables are computed excluding cohort 2015.

Table 3: Immediate Enrollment

	HES			University			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)
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)
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)
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

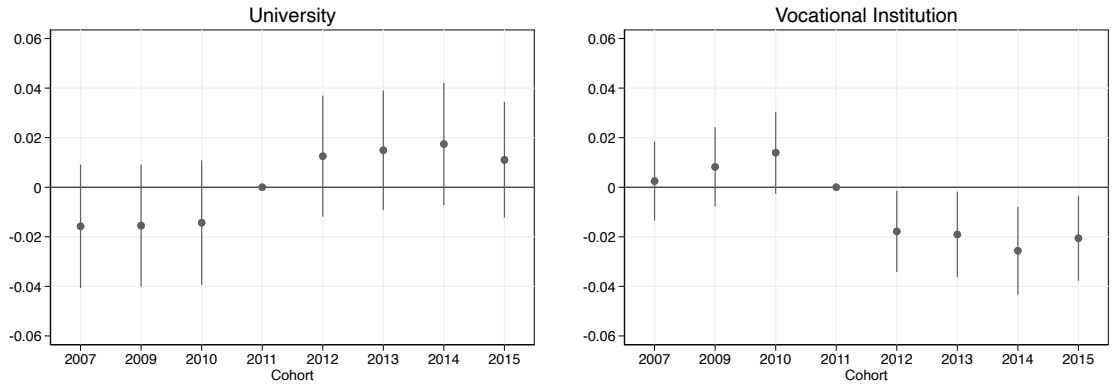


Table 4: Two-year Enrollment

	HES			University			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

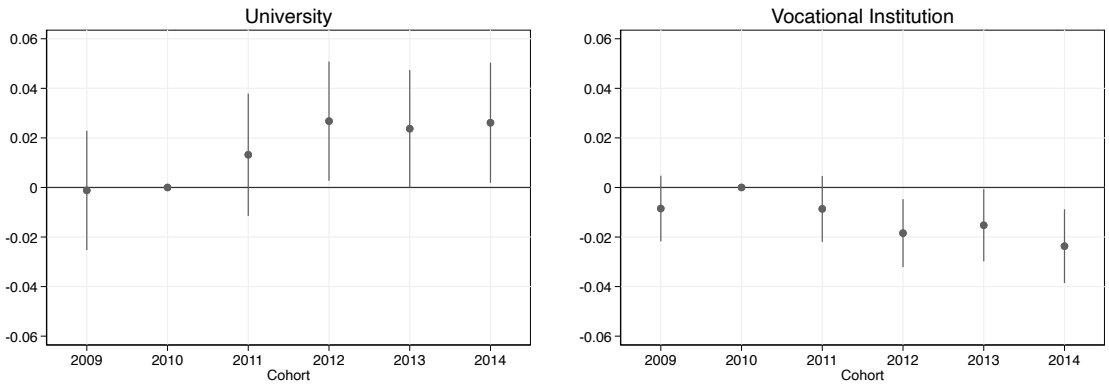
Notes: Clustered standard errors at the class level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. School level control variables include indicators of school type, rural area and geographical region. Student level control variables include gender, attendance rate, district and number of family members 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

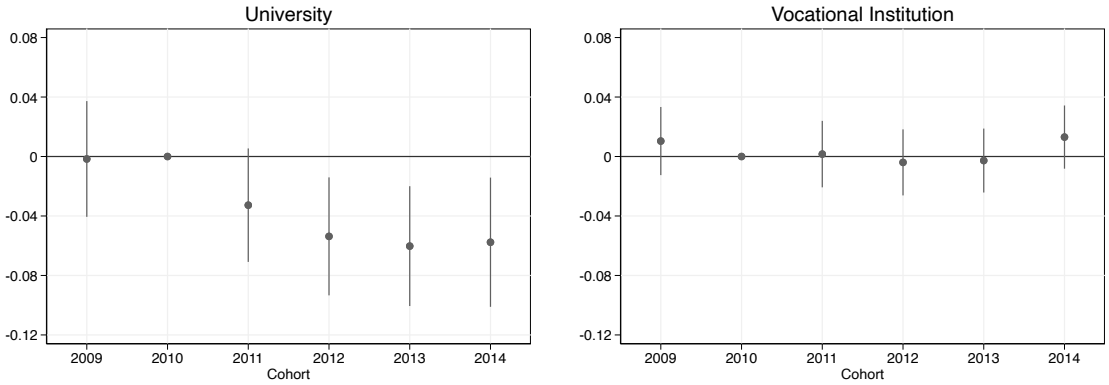
	HES			University			Vocational		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Eligible × exposed (2nd year)	-0.014***	-0.014***	-0.014***	-0.049***	-0.049***	-0.050***	-0.003	-0.003	-0.008
	(0.005)	(0.005)	(0.004)	(0.009)	(0.009)	(0.009)	(0.005)	(0.005)	(0.005)
Exposed (2nd year)	0.018***	0.023***	0.034***	0.057***	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	139,397	139,397	139,397	90,804	90,804	90,804	48,730	48,730	48,730
Outcome mean	0.103	0.103	0.103	0.102	0.102	0.102	0.181	0.181	0.181

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		
	(1)	(2)	(3)	(4)	(5)	(6)
	HES	University	Vocational	HES	University	Vocational
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)
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)
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)
Bandwidth						
Unexposed	50.574	42.220	52.172	42.668	45.042	53.310
Exposed	51.257	36.629	41.201	48.882	47.259	43.601
Observations						
Unexposed	113,523	95,218	116,986	27,528	29,019	34,149
Exposed	117,087	84,280	94,582	27,136	26,254	24,260

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		
	(1)	(2)	(3)	(4)	(5)	(6)
	HES	University	Vocational	HES	University	Vocational
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)
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)
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)
Bandwidth						
Unexposed	53.155	50.305	48.858	35.814	44.375	45.506
Exposed	58.077	37.934	38.461	64.812	51.634	48.703
Observations						
Unexposed	61,536	58,343	56,759	11,691	14,464	14,832
Exposed	133,494	88,264	89,627	38,607	31,115	29,424

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		
	(1)	(2)	(3)	(4)	(5)	(6)
	HES	University	Vocational	HES	University	Vocational
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)
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)
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)
Bandwidth						
Unexposed	50.402	49.227	50.556	41.813	38.712	50.574
Exposed	54.348	51.297	46.782	50.644	54.499	31.156
Observations						
Unexposed	28,947	13,368	15,330	6,267	2,471	4,341
Exposed	69,669	30,248	32,749	15,517	6,386	5,968

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			University			Vocational		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Female	Male	Difference	Female	Male	Difference	Female	Male	Difference
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

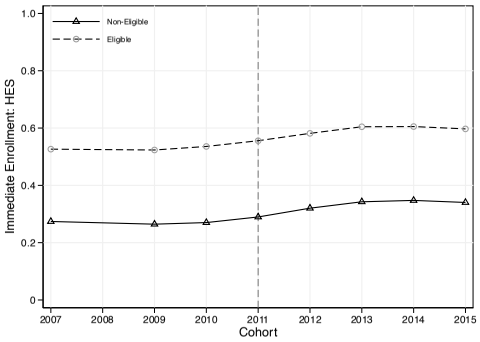
Notes: SUEST clustered standard errors at the class level in parentheses. Sample sizes in square brackets. Outcome sample means in curly braces. *** p<0.01, ** p<0.05, * p<0.1. School level control variables include indicators of school type, rural area and geographical region. Student level control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

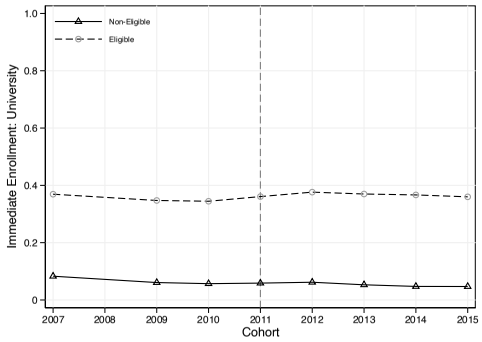
Table 10: Heterogeneity of Main Results by School Type

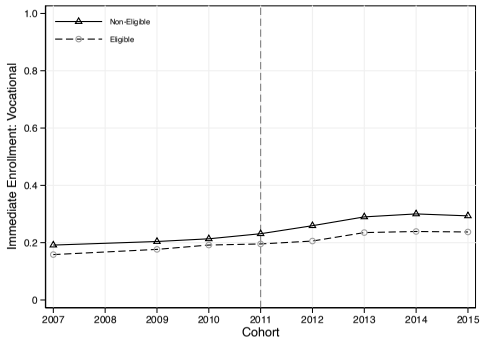
	HES			University			Vocational		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Public	Voucher	Difference	Public	Voucher	Difference	Public	Voucher	Difference
Immediate Enrollment	0.002	-0.005	0.007	0.010	0.028***	-0.018*	-0.008	-0.033***	0.025***
	(0.006)	(0.004)	(0.007)	(0.008)	(0.005)	(0.010)	(0.005)	(0.004)	(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 clustered standard errors 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.

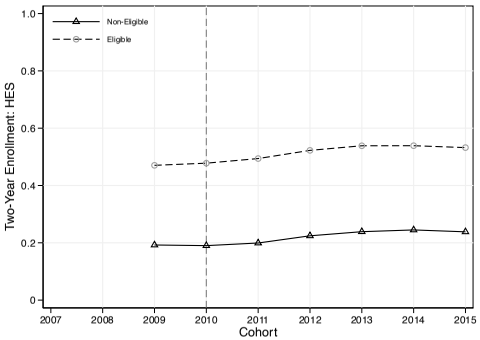
Figure A.1: Outcomes over Time by Eligibility

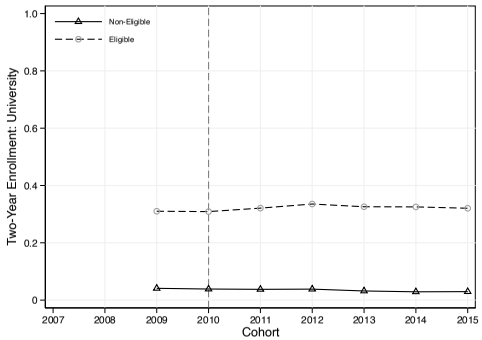


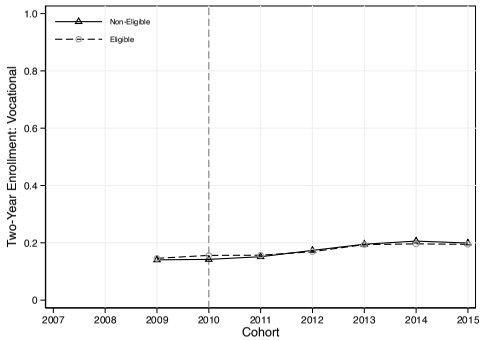




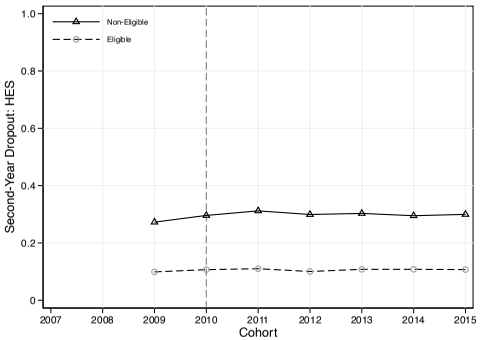
Panel A: Immediate Enrollment

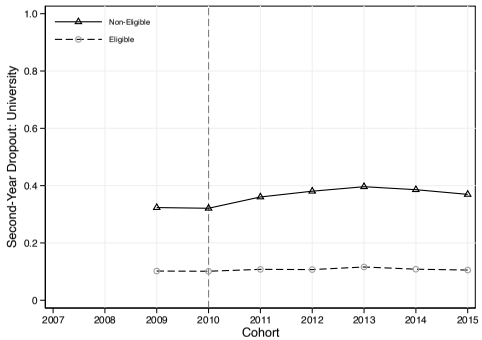


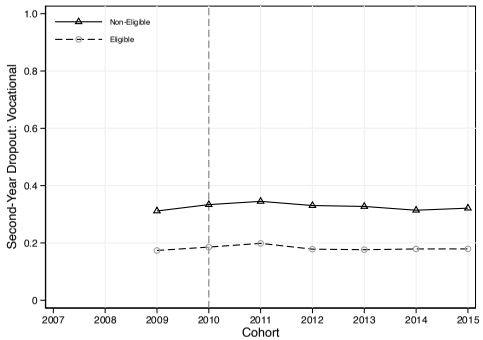




Panel B: Two-Year Enrollment







Panel C: Second-Year Dropout

Table A.1: Dynamics of Immediate Enrollment

	University		Vocational	
	(1)	(2)	(3)	(4)
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
	(0.010)	(0.010)	(0.006)	(0.006)
Eligible × cohort 2010	-0.014	-0.015	0.014**	0.015**
	(0.010)	(0.010)	(0.006)	(0.006)
Eligible × cohort 2012	0.013	0.011	-0.018***	-0.016***
	(0.009)	(0.009)	(0.006)	(0.006)
Eligible × cohort 2013	0.015	0.012	-0.019***	-0.018***
	(0.009)	(0.009)	(0.007)	(0.007)
Eligible × cohort 2014	0.017*	0.016	-0.026***	-0.025***
	(0.010)	(0.010)	(0.007)	(0.007)
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***
	(0.007)	(0.007)	(0.004)	(0.004)
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***
	(0.003)	(0.003)	(0.006)	(0.005)
Cohort 2010	-0.002	0.000	-0.018***	-0.019***
	(0.003)	(0.003)	(0.006)	(0.005)
Cohort 2012	0.003	0.005	0.028***	0.031***
	(0.003)	(0.003)	(0.005)	(0.005)
Cohort 2013	-0.006**	-0.005	0.059***	0.059***
	(0.002)	(0.003)	(0.006)	(0.005)
Cohort 2014	-0.012***	-0.010***	0.069***	0.071***
	(0.002)	(0.003)	(0.006)	(0.006)
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
Control variables	No	Yes	No	Yes
Observations	1,353,980	1,353,980	1,353,980	1,353,980

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: Dynamics of Two-year Enrollment

	University		Vocational	
	(1)	(2)	(3)	(4)
Eligible × cohort 2009	-0.001	-0.001	-0.009*	-0.009*
	(0.009)	(0.009)	(0.005)	(0.005)
Eligible × cohort 2011	0.013	0.014	-0.009*	-0.009*
	(0.010)	(0.010)	(0.005)	(0.005)
Eligible × cohort 2012	0.027***	0.026***	-0.018***	-0.018***
	(0.009)	(0.009)	(0.005)	(0.005)
Eligible × cohort 2013	0.024***	0.021**	-0.015***	-0.015***
	(0.009)	(0.009)	(0.006)	(0.006)
Eligible × cohort 2014	0.026***	0.025***	-0.024***	-0.024***
	(0.009)	(0.010)	(0.006)	(0.006)
Eligible	0.270***	0.250***	0.014***	0.012***
	(0.007)	(0.007)	(0.004)	(0.004)
Cohort 2009	0.002	0.001	-0.002	-0.001
	(0.002)	(0.003)	(0.004)	(0.004)
Cohort 2011	-0.001	-0.003	0.009**	0.010**
	(0.002)	(0.003)	(0.004)	(0.004)
Cohort 2012	-0.000	0.001	0.031***	0.036***
	(0.002)	(0.003)	(0.004)	(0.004)
Cohort 2013	-0.007***	-0.008**	0.053***	0.054***
	(0.002)	(0.003)	(0.005)	(0.004)
Cohort 2014	-0.010***	-0.009***	0.064***	0.067***
	(0.002)	(0.003)	(0.005)	(0.005)
Student district fixed effects	No	Yes	No	Yes
Control variables	No	Yes	No	Yes
Observations	1,035,568	1,035,568	1,037,137	1,037,137

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.3: Dynamics of Second-year Dropout

	University		Vocational	
	(1)	(2)	(3)	(4)
Eligible × cohort 2009	-0.002	-0.003	0.010	0.010
	(0.015)	(0.015)	(0.009)	(0.009)
Eligible × cohort 2011	-0.033**	-0.049***	0.002	-0.002
	(0.015)	(0.015)	(0.009)	(0.008)
Eligible × cohort 2012	-0.054***	-0.057***	-0.004	-0.007
	(0.015)	(0.015)	(0.009)	(0.008)
Eligible × cohort 2013	-0.060***	-0.057***	-0.003	-0.010
	(0.016)	(0.015)	(0.008)	(0.008)
Eligible × cohort 2014	-0.058***	-0.038**	0.013	0.006
	(0.017)	(0.017)	(0.008)	(0.008)
Eligible	-0.220***	-0.155***	-0.148***	-0.122***
	(0.011)	(0.011)	(0.006)	(0.006)
Cohort 2009	0.002	0.004	-0.022**	-0.026***
	(0.015)	(0.015)	(0.009)	(0.008)
Cohort 2011	0.039***	0.066***	0.011	0.022***
	(0.015)	(0.015)	(0.009)	(0.008)
Cohort 2012	0.060***	0.066***	-0.003	0.002
	(0.016)	(0.015)	(0.009)	(0.009)
Cohort 2013	0.075***	0.087***	-0.006	0.019**
	(0.016)	(0.016)	(0.009)	(0.008)
Cohort 2014	0.065***	0.061***	-0.020**	0.008
	(0.017)	(0.017)	(0.009)	(0.008)
Student district fixed effects	No	Yes	No	Yes
Control variables	No	Yes	No	Yes
Observations	301,826	297,440	224,857	224,480

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 C.2: Evidence of Female Delay

	Repetition		Improvement	
	(1) All students	(2) Non-enrolled	(3) All students	(4) Non-enrolled
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.