Online Appendix
The Effect of Education on Mortality and Health: Evidence from a Schooling Expansion in Romania
Ofer Malamud, Andreea Mitrut, and Cristian Pop-Eleches

Appendix Table 1: Specification Checks

Outcome	Density	Female	Romanian	Hungarian	Ethnic Roma	Ethnic Other	Born Bucharest
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Full samp	ple						
AFTER	0.105***	-0.024***	0.018***	-0.016***	-0.001	-0.001**	-0.012***
AFTER	[0.012]	[0.004]	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]
A ETTED #TDE A T	0.010	-0.018***	-0.004	0.005**	-0.000	-0.001	0.002
AFTER*TREAT	[0.016]	[0.005]	[0.003]	[0.002]	[0.001]	[0.001]	[0.002]
Sample size	3,240	3,240	3,240	3,240	3,240	3,240	3,240
\mathbb{R}^2	0.229	0.203	0.341	0.392	0.087	0.068	0.185
Panel B: 7-day don	nut						
AFTER	0.071***	-0.010***	0.011***	-0.012***	0.002***	-0.001	-0.008***
MILK	[0.005]	[0.003]	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]
AFTER*TREAT	-0.003	-0.010***	-0.003	0.003	0.000	-0.000	0.001
THIER TREAT	[0.007]	[0.004]	[0.003]	[0.002]	[0.001]	[0.001]	[0.002]
Sample size	3,114	3,114	3,114	3,114	3,114	3,114	3,114
\mathbb{R}^2	0.289	0.096	0.308	0.368	0.158	0.046	0.152
Panel C: 14-day de	onut						
AFTER	0.072***	-0.001	0.008***	-0.010***	0.003***	-0.000	-0.006***
AFIEK	[0.006]	[0.003]	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]
A ETED*TDE A T	-0.004	-0.014***	-0.002	0.002	0.001	-0.001	0.000
AFTER*TREAT	[0.008]	[0.004]	[0.003]	[0.002]	[0.001]	[0.001]	[0.002]
Sample size	2,988	2,988	2,988	2,988	2,988	2,988	2,988
R^2	0.229	0.053	0.270	0.332	0.177	0.038	0.126

Appendix Table 2: 2SLS Estimates of Years of Schooling on Socio-Economic Outcomes

Outcome	Employment	Occupational Skill	Fertility	Spouse's schooling					
	(1)	(2)	(3)	(4)					
Panel A: Full sample									
Years of schooling	0.023*	0.113***	-0.170**	0.315**					
	[0.012]	[0.023]	[0.071]	[0.124]					
Sample size R ²	3,240	3,240	3,240	3,240					
	0.639	0.784	0.370	0.926					
Panel B: 7-day don	Panel B: 7-day donut sample								
Years of schooling	0.016	0.062	-0.125	-0.011					
	[0.021]	[0.041]	[0.123]	[0.229]					
Sample size R ²	3,114	3,114	3,114	3,114					
	0.625	0.730	0.363	0.895					
Panel C: 14-day donut sample									
Years of schooling	0.017	0.022	0.101	-0.686					
	[0.029]	[0.065]	[0.193]	[0.505]					
Sample size R ²	2,988	2,988	2,988	2,988					
	0.632	0.650	0.199	0.744					

Notes: Individual observations for cohorts born from mid-1944 to mid-1953 in the 1992 Census collapsed to the day-of-birth level. Outcome variables are defined as in Table 1. 2SLS regressions instrument for years of schooling with the interaction of AFTER*TREAT, control for AFTER, and include saturated controls for day-of birth interacted with AFTER and TREAT. Columns (1)-(8) report standard 2SLS estimates while columns (9)-(12) use report estimates from two-sample 2SLS. All specifications use a bandwidth of 180 days. Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level respectively.

Appendix Table 3: Outcomes from LSMS household surveys

Outcome	Years of schooling	Employment	Income	Smoking	Chronic conditions
	(1)	(2)	(3)	(4)	(5)
AFTER	0.0609	-0.0188	0.0105	0.0031	0.0032
TH TEN	[0.2240]	[0.0130]	[0.0438]	[0.0058]	[0.0047]
AFTER*TREAT	0.6913**	0.0415***	0.0195	0.0086	0.0083
AFIERTIKEAI	[0.2997]	[0.0044]	[0.0229]	[0.0089]	[0.0074]
~	10.015	20.240	20.122	0.4.000	0.4.000
Sample size	10,315	30,340	30,133	84,032	84,032
\mathbb{R}^2	0.034	0.020	0.009	0.006	0.001

Notes: Individual observations for cohorts born from mid-1944 to mid-1953 in the 1992 Census collapsed to the day-of-birth level. Outcome variables are defined as in Table 1. AFTER is an indicator that equals 1 for individuals born after January 1. TREAT is an indicator that equals 1 for cohorts who experienced the education expansion during 1945-1950. Regressions include saturated linear controls for month-of birth interacted with AFTER and TREAT. All specifications use a bandwidth of 90 days. Heteroskedasticity-robust standard errors are in parentheses. ***, ***, and * indicate statistical significance at the 1, 5, and 10 percent level respectively.

Appendix Table 4: Parametric RD Estimates

Outcome	Y	ears of schooli	ing		Mortality			Density	
Poly. function	Linear (1)	Quadratic (2)	Cubic (3)	Linear (4)	Quadratic (5)	Cubic (6)	Linear (7)	Quadratic (8)	Cubic (9)
AFTER	-0.050***	-0.018	0.032	0.005*	0.008**	0.004	0.059***	0.091***	0.074***
	[0.018]	[0.029]	[0.046]	[0.002]	[0.004]	[0.006]	[0.005]	[0.007]	[0.010]
AFTER*TREAT	0.123***	0.104**	0.177***	0.004	0.003	0.011	-0.005	-0.001	-0.015
	[0.025]	[0.042]	[0.065]	[0.004]	[0.006]	[0.009]	[0.007]	[0.009]	[0.013]
Sample size R ²	3,114	3,114	3,114	3,114	3,114	3,114	3,114	3,114	3,114
	0.920	0.922	0.923	0.779	0.779	0.780	0.167	0.246	0.258

Notes: Individual observations for cohorts born from mid-1944 to mid-1950 in the 1992 Census collapsed to the day-of-birth level. Outcome variables are defined as in Table 1. AFTER is an indicator that equals 1 for individuals born after January 1. AFTER is an indicator that equals 1 for individuals born after January 1. TREAT is an indicator that equals 1 for cohorts who experienced the education expansion during 1945-1950. Local, quadratic, and cubic specifications include saturated controls for day-of birth interacted with AFTER and TREAT. All specifications use a bandwidth of 180 days. Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level respectively.

Appendix Table 5: Un-collapsed RD Estimates

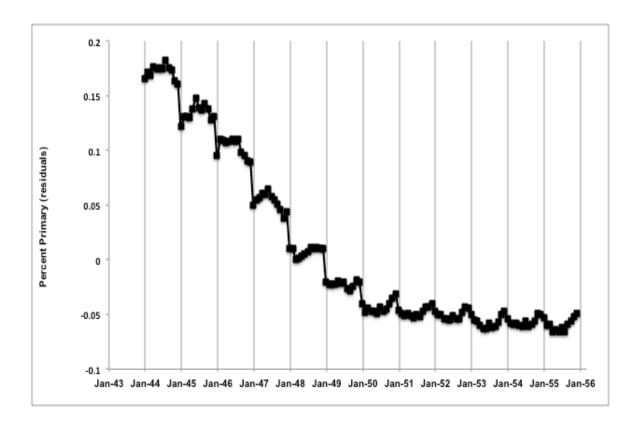
Outcome	Years schooling (1)	Primary (2)	Lower Secondary (3)	Upper Secondary (4)	Higher education (5)	Employment (6)	Occupational skill (7)	Fertility (8)	Spouse's schooling (9)
AFTER	-0.038**	0.004***	-0.009***	0.011***	-0.006***	0.001	-0.015***	0.014	-0.033*
	[0.017]	[0.001]	[0.002]	[0.003]	[0.002]	[0.002]	[0.004]	[0.011]	[0.019]
AFTER*TREAT	0.102***	-0.016***	0.008***	0.007**	0.000	0.001	0.006	-0.012	-0.013
	[0.023]	[0.002]	[0.003]	[0.003]	[0.002]	[0.002]	[0.005]	[0.014]	[0.024]
Sample size R ²	2,466,846	2,466,846	2,466,846	2,466,846	2,466,846	2,474,835	2,155,293	1,240,796	2,114,104
	0.022	0.044	0.001	0.014	0.001	0.002	0.002	0.001	0.017

Notes: Individual observations for cohorts born from mid-1944 to mid-1950 in the 1992 Census. Outcome variables are defined as in Table 1. AFTER is an indicator that equals 1 for individuals born after January 1. Local linear regressions use triangular kernels and include controls for day-of-birth interacted with AFTER. All specifications use a bandwidth of 180 days. Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level respectively.

Table 4: Effects of Educational Expansion on Mortality and Health Outcomes

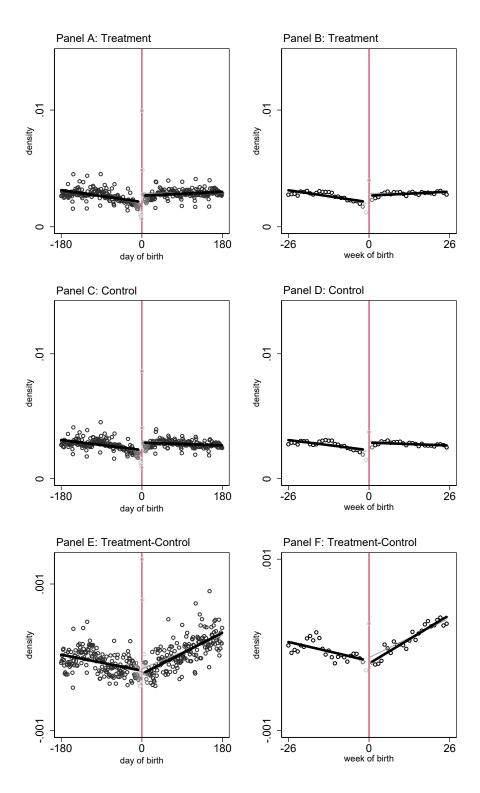
Outcome	International Migration	Attrition
	(1)	(2)
Panel A: Full sample		
AFTER	0.011***	-0.007
AFIEK	[0.003]	[0.006]
AFTER*TREAT	-0.001	-0.006
AFIEKTREAT	[0.003]	[0.008]
Sample size	3,240	3,240
R ²	0.563	0.218
Panel B: 7-day donut	sample	
AFTER	0.009***	-0.007
AFIER	[0.003]	[0.006]
AFTER*TREAT	-0.000	-0.004
	[0.004]	[0.008]
Sample size	3,114	3,114
R^2	0.576	0.227
Panel C: 14-day donu	t sample	
AFTED	0.008***	-0.004
AFTER	[0.003]	[0.007]
A ETED*TDE A T	0.002	-0.007
AFTER*TREAT	[0.004]	[0.009]
C 1	2.000	2.000
Sample size R ²	2,988 0.581	2,988 0.224
IX	0.501	U.227

Notes: Individual observations for cohorts born from mid-1944 to mid-1953 in the 1992 Census collapsed to the day-of-birth level. Outcome variables are defined as in Table 1. AFTER is an indicator that equals 1 for individuals born after January 1. TREAT is an indicator that equals 1 for cohorts who experienced the education expansion during 1945-1950. Local linear regressions use triangular kernels and include saturated controls for day-of birth interacted with AFTER and TREAT. All specifications use a bandwidth of 180 days. Heteroskedacticity-robust standard errors are in parentheses. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level respectively.



Appendix Figure 1: Primary school completion (residualized) for cohorts born 1943-1955

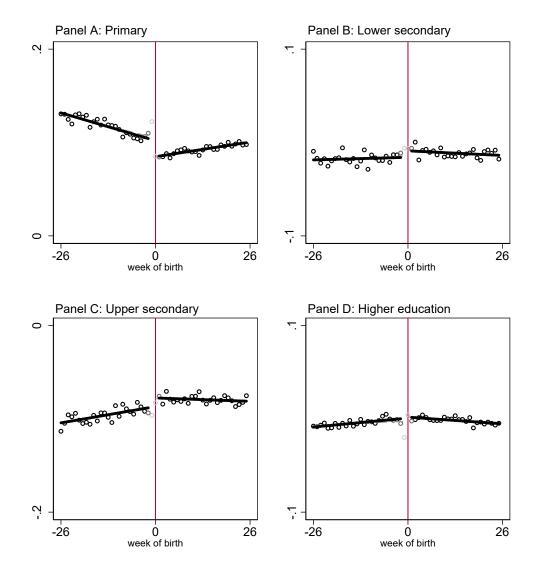
Notes: This figure plots the percent of individuals born between 1943 and 1955 who completed only primary education by their month of birth, which are based on residuals. Source: 1992 Romanian Census (complete count).



Appendix Figure 2: Regression discontinuity (RD) plots of the density

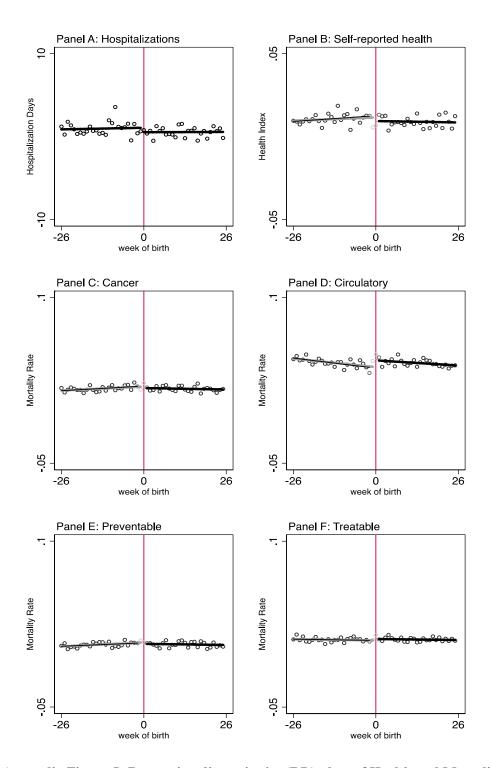
Notes: Panels A and B are restricted to individuals born in the treatment years (1944-1950). Panels C and D are restricted to individuals born in the control years (1950-1953). Panels E and F are restricted to individuals born in both treatment and control years (1944-1953). The open circles indicate the mean of

the outcome by day of birth (panels A, C and E) or week of birth (panels B, D and F). The solid lines are based on linear spline regressions. The light gray lines and circles show our preferred 7-day donut specification (when we drop individuals born within 7 days of January 1 in order to be symmetric around the cutoff). Source: 1992 Romanian Census (complete count).



Appendix Figure 3: Regression discontinuity (RD) plots for specific schooling categories

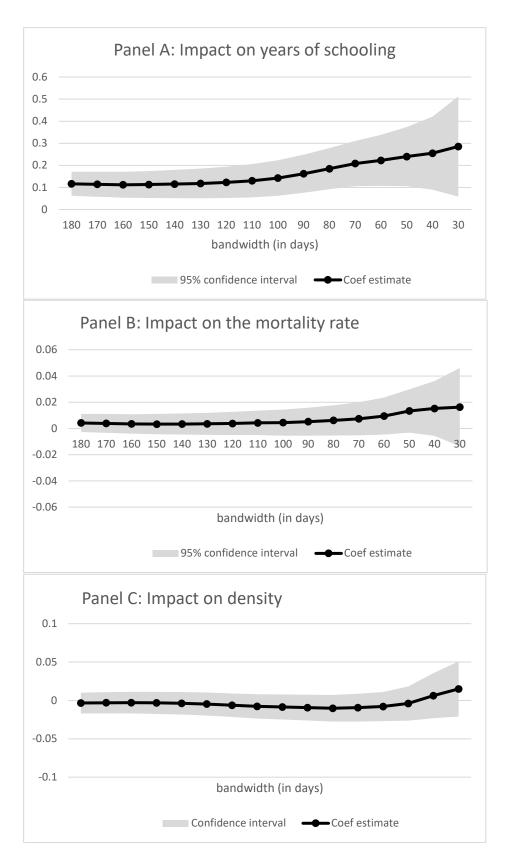
Notes: All Panels are restricted to individuals born in both treatment and control years (1944-1953). The open circles indicate the mean of the outcome by week of birth. The solid lines are based on linear spline regressions. The light gray lines and circles show our preferred 7-day donut specification (when we drop individuals born within 7 days of January 1 in order to be symmetric around the cutoff). Source: 1992 Romanian Census (complete count).



Appendix Figure 5: Regression discontinuity (RD) plots of Health and Mortality by Cause

Notes: All Panels are restricted to individuals born in both treatment and control years (1944-1953). The open circles indicate the mean of the outcome by week of birth. The solid lines are based on linear spline regressions. The light gray lines and circles show our preferred 7-day donut specification (when we drop

individuals born within 7 days of January 1 in order to be symmetric around the cutoff). Source: 2011 Romanian Census (complete count) and 1994-2011 Vital Statistics Mortality files.



Appendix Figure 6: RD estimates for alternative bandwidths

Notes: This figure plots RD estimates for three outcomes (years of schooling, mortality rate and density) over a broad range of alternative bandwidths between 180 and 30 days using the 7-day donut sample. Source: 1992 Romanian Census (complete count), Vital Statistics Mortality files from 1994 to 2011.