Covid-19 counterfactual evidence.

Estimating the effects of school closures

Marco Giuliani

ONLINE APPENDIX

Datasets and code are available on Harvard Dataverse at https://doi.org/10.7910/DVN/JSSIVX

Contents

1 European comparison

1.1 Variables, measurements and sources

Table A.1 Codebook

Variable	Measurement	Source	Link
Incidence of new cases	Weekly number of new COVID-19 cases per million persons	Our World in Data on COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)	https://github.com/owid/covid-19-data
School closing scale	Weekly average of the School closing 0-3 scale (weighted for possible subnational differentiation)	Oxford COVID-19 Government Response Tracker (OxCGRT)	https://github.com/OxCGRT/covid-policy-tracker
School closing index	Weekly average of the index measuring the 0-100 percentage of strictness of regulation relative to its theoretical maximum (0-100)	Oxford COVID-19 Government Response Tracker (OxCGRT)	https://github.com/OxCGRT/covid-policy-tracker
School closing 75 dummy	Dummy variable for the school closing policy (=1 if the school closing index is higher than 75%, and =0 otherwise)	Oxford COVID-19 Government Response Tracker (OxCGRT)	https://github.com/OxCGRT/covid-policy-tracker
Other containment and closure policies	Weekly average of the index computed on the other 7 containment and closure policies (weighted for possible subnational differentiation)	Oxford COVID-19 Government Response Tracker (OxCGRT)	https://github.com/OxCGRT/covid-policy-tracker
Other containment 75 policy dummy	Dummy variable for the other containment and closure policies (=1 if the other containment index is higher than 75%, and =0 otherwise)	Oxford COVID-19 Government Response Tracker (OxCGRT)	https://github.com/OxCGRT/covid-policy-tracker
Fully vaccinated	Cumulated percentage of fully vaccinated population	Data on COVID-19 by Our World in Data	https://github.com/owid/covid-19-data

Deaths per million	Total number of certified COVID deaths per million persons	Data on COVID-19 by Our World in Data	https://github.com/owid/covid-19-data
Positivity rate	Pct of positive COVID test on the total number of tests (weekly)	Data on COVID-19 by Our World in Data	https://github.com/owid/covid-19-data
Urbanization	Pct of population residing in urban areas 2020 estimate	UN World urbanization prospects	https://population.un.org/wup/Download/
Log population	Logarithm of population	Our World in Data	https://github.com/owid/covid-19-data
Density	Population per squared kilometre	World Bank World Development Indicators	https://github.com/owid/covid-19-data
HDI	UN 2019 Human Development index	Human development report	https://hdr.undp.org/en/data
Rule of law	Rule of law World Governance indicator 2019	The World Bank - Data Catalog	https://datacatalog.worldbank.org/home
Regulatory quality	Regulatory quality World Governance indicator 2019	The World Bank - Data Catalog	https://datacatalog.worldbank.org/home
Trust in government	Pct of survey respondents who trust national government in 2020	Eurobarometer Oecd	https://europa.eu/eurobarometer/surveys/browse/all https://data.oecd.org/gga/trust-in-government.htm

Regarding the treatment, the school closing dummy variable takes the value of 1 whenever the school closing index, computed according to the formula provided by the same proposers of the Stringency index (Hale et al., 2021), is higher than 75%. That threshold can be overcome only when the most stringent level of the original ordinal scale is reached, that is, when school closings are required at all levels, at most moderated by the fact that the mandatory requirement applied only to some parts of the country (which happened only to 6% of the observations in the sample). Furthermore, because the sample is based on weekly observations, that threshold corresponds to weeks in which at least 5 days out of 7 have complete school closings at all levels.

1.2 Descriptive information

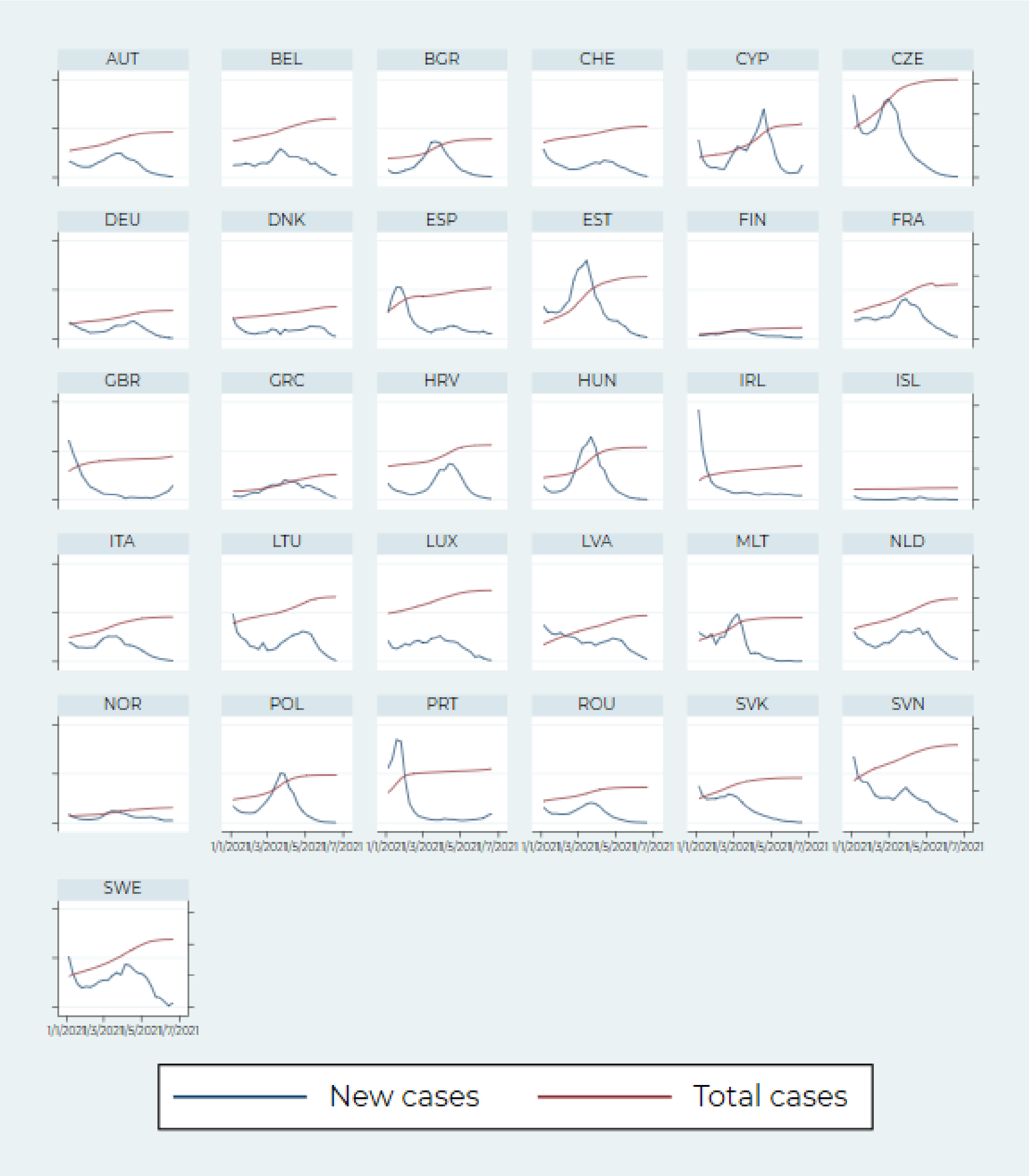

Figure A.1 Incidence of new and total cases per million population – 31 European countries (January-June 2021)

Table A.2 Descriptive statistics of the major variables used in the comparative analysis

Variable	Mean	Std. dev.	Min	Max
New cases (1000000)	1586.02	1497.60	11.72	9260.41
Total Fully vaccinated (avg pct week)	9.41	10.94	0.00	70.18
Total deaths (1000000)	1207.75	663.60	84.98	3103.41
Tests (weekly 1000)	57.74	87.32	4.77	592.15
School index (avg week)	58.14	24.16	0.00	100.00
Avg other constraints indices (avg week)	59.08	15.45	0.00	88.10

1.3 Further analyses

If the matching procedure was perfect, it would be possible to avoid including in the regression the control variables already involved in the balancing procedure. In this case, and keeping in the right-hand side of the equation only the lagged dependent variable for the reasons highlighted in the article, this would produce the following results.

Table A.3. Estimating the effect of closing schools on the incidence of new cases with the balanced samples and without the control variables involved in the balancing procedure

	(1) PCSE	(2) FE	(3) FE-PCSE
L. New cases (1000000)	0.94***	1.03***	1.03***
	(0.03)	(0.03)	(0.05)
L. School closing dummy	-171.46***	-246.18***	-246.18***
	(53.83)	(55.80)	(54.89)
Tests (1000)	-0.03	-0.05	-0.05
	(0.21)	(0.72)	(0.45)
Constant	140.34**	109.03	109.03
	(54.34)	(162.62)	(129.14)
Country fixed effects		✔	✔
N	360	360	360
Countries	31	31	31
R-squared	0.87	0.91	0.91

Note: Standard errors in parentheses *p < 0.10 **p < 0.05 ***p < 0.01

Even in this “minimalist” version, under the supposition of an almost perfectly balanced sample, the results presented in the article are confirmed.

In Table A.4, I have tested a series of variables to try to uncover the country fixed effects included in the article.

Table A.4. Estimating the effect of closing schools on the incidence of new cases with the original and balanced samples (PCSE)

	(1) Original	(2) Balanced
L. New cases (1000000)	0.88***	0.93***
	(0.03)	(0.04)
L. School closing dummy	-87.28	-219.16***
	(63.70)	(55.16)
L. Other policies	-126.98	-47.49
	(86.89)	(61.20)
L. Fully vaccinated (100)	-9.83**	-22.90***
	(3.81)	(7.32)
Total deaths (1000000)	-0.08	-0.04
	(0.05)	(0.06)
Tests (1000)	0.26	0.12
	(0.40)	(0.23)
Urbanization	0.67	-3.34
	(1.98)	(2.14)
Log population	24.97	13.81
	(21.83)	(18.19)
Density	0.13	0.29*
	(0.08)	(0.15)
Human development index	538.80	239.35
	(1425.36)	(1448.40)
Rule of law	-33.38	47.12
	(142.84)	(93.77)
Regulatory quality	-40.10	-241.67*
	(166.21)	(146.22)
Trust in government	-1.57	4.28*
	(1.29)	(2.25)
Constant	-484.83	120.51
	(1034.13)	(1344.46)
N	775	360
Countries	31	31
R-squared	0.85	0.88

Note: Standard errors in parentheses *p < 0.10 **p < 0.05 ***p < 0.01

More specifically, in addition to the time-varying factors already included in the previous models, I have added:

- Three demographic variables such as the size of the population (logged), its density and a measure of urbanization. While there are no precise expectations regarding the association between the size of the country and pandemic dynamics, density and urbanization have often been quoted as favouring the spread of the infections.
- The human development index, synthesizing three different dimensions: wealth (GNI/pc), education (mean years of schooling completed and expected years of schooling upon entering the education system) and health (life expectancy).
- Two world governance indicators – rule of law and regulatory quality – capturing the institutional capacity of a political system, which could be supposed to be directly or indirectly associated with a greater capacity to limit infections.
- Trust in government is often supposed to be associated with a more positive attitude of citizens towards the need to comply with government regulations and recommendations.

Most of these expectations are not confirmed by the empirical analysis of Table A.4, both in the original and in the matched sample. The only variables that follow the expectations are the density of the population, which is positively associated with the spread of the pandemic, and the regulatory quality, which is negatively associated with it. However, the two relationships are statistically only weakly significant.

Furthermore, even in these models, it is possible to notice that the coefficient for the school closing variable is not relevant in the raw model, while it becomes highly significant in the matched model.

As a robustness test, I replicated the same models presented in Table 1-2-3 in the article, but this time using the whole 0-100 scale of the school and containment indices.

Table A.5 presents the results for the full sample, Table A.6 presents the information regarding the matching process, and Table A.7 presents the results for the balanced sample, which after CEM includes only 136 observations for 26 countries (which is one of the reasons for preferring the model with the dummy variables instead of the continuous scales).

Table A.5 Estimating the effect of closing schools on the incidence of new cases

	(1) PCSE	(2) FE	(3) FE-PCSE
L. New cases (1000000)	0.88***	0.83***	0.83***
	(0.03)	(0.02)	(0.04)
L. School index	-2.77**	-2.97**	-2.97**
	(1.18)	(1.29)	(1.37)
L. Other policies index	-0.54	-5.36	-5.36
	(1.98)	(3.26)	(4.03)
L. Fully vaccinated (100)	-10.42***	-9.31***	-9.31**
	(3.50)	(3.39)	(3.72)
Total deaths (1000000)	-0.02	-0.53***	-0.53***
	(0.03)	(0.09)	(0.16)
Tests (1000)	0.21	1.72***	1.72**
	(0.38)	(0.51)	(0.73)
Constant	399.67***	989.15***	989.15**
	(107.72)	(299.35)	(399.50)
Country fixed effects		✔	✔
N	775	775	775
Countries	31	31	31
R-squared	0.85	0.86	0.86

Note: Standard errors in parentheses *p < 0.10 **p < 0.05 ***p < 0.01

Table A.6 Imbalances in the raw and in the matched sample

Raw sample	$\mathcal{L}_1$	Avg	Min	Med	Max
L. New cases (1000000)	0.28	1030.80	84.98	794.7	788.16
L. Other policies index	0.50	14.76	17.52	16.67	3.57
L. Fully vaccinated (100)	0.27	-5.50	0.00	-3.20	-39.35
Total deaths (1000000)	0.30	-59.90	0.00	-254.05	-347.30
Multivariate	0.92				
Balanced sample	$\mathcal{L}_1$	Avg	Min	Med	Max
L. New cases (1000000)	0.09	-26.33	121.12	-111.86	271.5
L. Other policies index	0.16	-0.80	0.34	-1.87	0.00
L. Fully vaccinated (100)	0.24	-0.96	0.00	-1.25	-3.05
Total deaths (1000000)	0.26	-39.72	282.41	-48.01	140.17
Multivariate	0.78				

Table A.7 Estimating the effect of closing schools on the incidence of new cases using the balanced sample

	(1) PCSE	(2) FE	(3) FE-PCSE
L. New cases (1000000)	0.99***	0.99***	0.99***
	(0.05)	(0.06)	(0.07)
L. School index	-5.55***	-6.66***	-6.66***
	(1.98)	(2.27)	(1.56)
L. Other policies index	-3.51	-16.70	-16.70
	(4.94)	(13.04)	(15.30)
L. Fully vaccinated (100)	-10.96	-50.64**	-50.64***
	(8.83)	(20.87)	(16.01)
Total deaths (1000000)	-0.05	0.59	0.59
	(0.05)	(0.43)	(0.38)
Tests (1000)	-0.43	-0.56	-0.56
	(0.36)	(0.93)	(0.44)
Constant	824.50**	1482.08	1482.08
	(366.83)	(1200.72)	(1220.16)
Country fixed effects		✔	✔
N	136	136	136
Countries	26	26	26
R-squared	0.88	0.93	0.93

Note: Standard errors in parentheses *p < 0.10 **p < 0.05 ***p < 0.01

The results confirm the findings of the article, although this time the improvement between the original and matched sample (in addition to the explained variance) mostly concerns the level of statistical significance of the main covariate of interest (from p=0.030 to p<0.0001).

A second robustness test again uses the dummy variables of the school and containment policies but includes in the matching procedure and in the regression models another time-varying variable, namely, the positivity rate. The higher the positivity rate in the week before the observation, the more likely the further spread of the pandemic.

Table A.8 presents the results for the full sample, Table A.9 presents the information regarding the matching process, and Table A.10 presents the results for the balanced sample. The more demanding matching requirements further reduce the balanced sample to just 194 observations and just 28 countries.

Table A.8 Estimating the effect of closing schools on the incidence of new cases

	(1) PCSE	(2) FE	(3) FE-PCSE
L. New cases (1000000)	0.84***	0.77***	0.77***
	(0.03)	(0.03)	(0.04)
L. School closing dummy	-95.96	-102.88	-102.88
	(59.73)	(67.03)	(71.61)
L. Other policies index	-70.31	-168.77**	-168.77*
	(85.47)	(85.25)	(101.40)
L. Fully vaccinated (100)	-7.93**	-5.64*	-5.64
	(3.69)	(3.02)	(4.09)
L. Positivity rate	15.46***	22.71***	22.71***
	(4.95)	(7.12)	(8.12)
Total deaths (1000000)	-0.04	-0.46***	-0.46***
	(0.03)	(0.09)	(0.16)
Tests (1000)	0.64	1.90***	1.90***
	(0.41)	(0.52)	(0.71)
Constant	184.26***	406.31**	406.31
	(61.22)	(192.88)	(270.52)
Country fixed effects		✔	✔
N	775	775	775
Countries	31	31	31
R-squared	0.85	0.86	0.86

Note: Standard errors in parentheses   *p < 0.10 **p < 0.05 ***p < 0.01

Table A.9 Imbalances in the raw and in the matched sample

Raw sample	$\mathcal{L}_1$	Avg	Min	Med	Max
L. New cases (1000000)	0.28	1030.80	84.98	794.7	788.16
L. Other policies	0.34	0.34	0.00	0.00	0.00
L. Fully vaccinated (100)	0.27	-5.50	0.00	-3.20	-39.35
L. Positivity rate	0.18	-3.3	0.24	2.31	-7.67
Total deaths (1000000)	0.30	-59.90	0.00	-254.05	-347.3
Multivariate	0.92				
Balanced sample	$\mathcal{L}_1$	**Avg**	**Min**	**Med**	**Max**
L. New cases (1000000)	0.14	44.99	84.98	-118.60	-119.18
L. Other policies	0.00	-0.00	0.00	0.00	0.00
L. Fully vaccinated (100)	0.23	-0.85	0.00	-1.27	2.04
L. Positivity rate	0.03	0.06	0.23	-0.02	-0.29
Total deaths (1000000)	0.04	-22.34	0.00	-56.97	49.7

Multivariate	0.53

Table A.10 Estimating the effect of closing schools on the incidence of new cases using the balanced sample

	(1) PCSE	(2) FE	(3) FE-PCSE
L. New cases (1000000)	0.94***	1.06***	1.06***
	(0.02)	(0.05)	(0.06)
L. School index	-171.85**	-130.23**	-130.23**
	(83.87)	(60.06)	(53.40)
L. Other policies index	-23.09	-113.04	-113.04
	(69.79)	(120.89)	(110.95)
L. Fully vaccinated (100)	-18.72***	-11.24	-11.24
	(6.04)	(7.67)	(8.81)
L. Positivity rate	0.70	-16.52	-16.52
	(5.03)	(14.08)	(17.80)
Total deaths (1000000)	-0.09*	-0.55***	-0.55***
	(0.05)	(0.18)	(0.19)
Tests (1000)	0.35	1.81*	1.81**
	(0.27)	(0.96)	(0.86)
Constant	231.94***	326.35	326.35
	(82.71)	(286.83)	(330.84)
Country fixed effects		✔	✔
N	194	194	194
Countries	28	28	28
R-squared	0.92	0.95	0.95

Note: Standard errors in parentheses   *p < 0.10 **p < 0.05 ***p < 0.01

2 Italian subnational analysis

2.1 Variables, measurements and sources

Table A.11 Codebook

Variable	Measurement	Source	Link
Incidence of new/total cases	Weekly number of new COVID-19 cases per 100000 persons	Dipartimento Protezione Civile - Dati COVID-19 Italia	https://github.com/pcm-dpc/COVID-19
Incidence active cases	Weekly average of the active COVID-19 cases per 100000 persons	Dipartimento Protezione Civile - Dati COVID-19 Italia	https://github.com/pcm-dpc/COVID-19
Tests	Weekly number of COVID-19 tests per 1000 persons	Dipartimento Protezione Civile - Dati COVID-19 Italia	https://github.com/pcm-dpc/COVID-19
Positivity rate	Weekly ratio between new cases and tests	Dipartimento Protezione Civile - Dati COVID-19 Italia	https://github.com/pcm-dpc/COVID-19
Reproduction rate Rt	Regional reproduction rate	Sole 24 Ore	https://lab24.ilsole24ore.com/coronavirus/#
Population	Provincial and regional population	Dipartimento Protezione Civile - Dati COVID-19 Italia	https://github.com/pcm-dpc/COVID-19
Student population	Percentage of student population (from pre-school to high-school)	ISTAT	http://dati.istat.it/Index.aspx?DataSetCode=DCIS_SCUOLE#
Density	Population per squared kilometre	Dati ISTAT 1 Gennaio 2021	https://www.tuttitalia.it/

2.2 Descriptive information

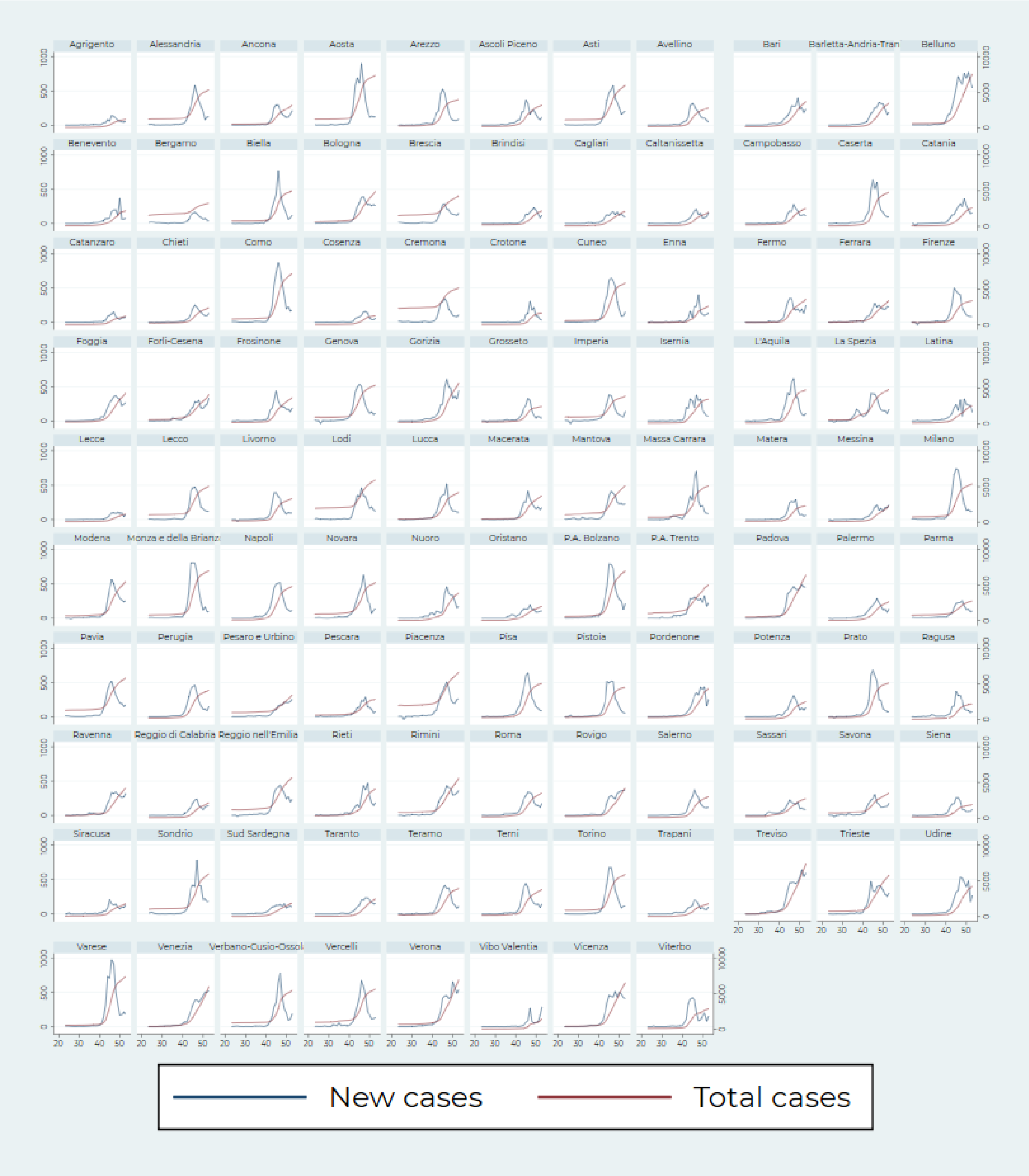

Figure A.2 Incidence of new and total cases per 100000 population – 31 European countries (June-December 2020)

Table A.12 Descriptive statistics of the major variables used in the within-Italy analysis

Variable	Mean	Std. dev.	Min	Max
New cases (100000)	96.12	148.11	0.00	973.15
Total cases (100000)	1090.28	1216.82	27.62	7586.21
Active cases (100000)	371.40	516.42	0.37	2177.94
Population	553809.00	602707.80	81918.00	4227588.00
Density	265.31	375.95	36.00	2560.00
Student population	14.04	1.17	11.54	17.58
Tests (weekly 1000)	12.44	8.71	0.00	45.96
Positivity rate	0.052	0.06	0.00	0.23
Reproduction rate Rt	.9510582	.3696576	0.00	2.37

2.3 Further analyses

Preliminary panel regressions to test the appropriate factors to include in the identification strategy for the synthetic counterfactual.

Table A.13 Panel regressions on the weekly incidence of new cases in Italy (June-December 2020)

	(1)	(2)	(3)
Lag Total cases (100000)	0.005		0.003
	(0.007)		(0.007)
Lag Active cases (100000)	0.016		0.018
	(0.015)		(0.015)
Tests (1000)	5.835 ***		5.778 ***
	(0.547)		(0.538)
Positivity rate	1457.391 ***		1466.838 ***
	(91.390)		(89.187)
Reproduction rate Rt	-12.996		-13.997
	(7.476)		(7.450)
Log Population		6.134 **	-4.214 *
		(2.261)	(1.657)
Density		0.031 **	0.023 **
		(0.012)	(0.008)
Student population (100)		-9.404 ***	-0.470
		(2.243)	(1.166)
Constant	-50.738 ***	140.635 **	6.244
	(10.653)	(48.631)	(25.0533)
Provinces	107	107	107
Observations	3210	3317	3210

Note: Panel corrected standard errors in parentheses. *p < 0.05 **p < 0.01 ***p < 0.001

The synthetic control method used in the article is often considered an alternative to difference-in-difference (DID) models (O'Neill et al., 2016). For robustness, I also fitted two DID models for cross-sectional data with panel and time fixed effects to explain the impact of school openings on the incidence of new COVID cases. The first one uses only the covariates of interest, whereas the second one also introduces the set of control variables used in the article: the number of total and active cases per 100000 inhabitants, the number of tests per thousand persons, the positivity rate and the reproduction rate R_t.

Table A.14 Average effects of early school-openings in Italy (June-December 2020)

```
Difference-in-differences regression                    Number of obs = 3,317
Data type: Longitudinal

                              (Std. err. adjusted for 107 clusters in province)
```

newnew	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
ATET						
school (1 vs 0)	**83.08866**	**11.0505**	**7.52**	**0.000**	**61.17997**	**104.9973**

Note: ATET estimate adjusted for panel effects and time effects.

```
Difference-in-differences regression                    Number of obs = 3,317
Data type: Longitudinal

                              (Std. err. adjusted for 107 clusters in province)
```

newnew	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
ATET						
school (1 vs 0)	**16.71952**	**6.884668**	**2.43**	**0.017**	**3.069993**	**30.36904**

Note: ATET estimate adjusted for covariates, panel effects, and time effects.

In both cases, the average treatment effect for those provinces that opened the schools earlier is positive and highly significant. Adjusting the estimate for all control variables, ATET is almost 17 more new weekly cases per 100000 inhabitants. In any case, given the demanding assumptions of DID models (Cunningham, 2021), the synthetic control method should be preferred in these circumstances.

The synthetic research method uses a linear combination of observations from a donor pool to build a counterfactual that minimizes the error of prediction before the treatment. In the article, the best synthetic Bolzano was a weighted mix among the provinces of Pescara, Cosenza and Reggio Calabria. Recently, Cerulli (2020: 845) suggested that “relaxing the linearity assumption by providing a nonparametric estimation of the weights may somehow improve their estimation” and better distribute them across a larger number of non-treated units. As a further robustness test, I ran a non-parametric synthetic control analysis whose results are reported in the online appendix. Using this approach, it was possible to slightly improve (i.e., reduce) the pre-treatment prediction error (RMSPE=4.40) while distributing the weights across a much larger number of provinces, which makes the counterfactual less dependent on some local trend. However, the alternative specification does not modify, but instead reinforces, the conclusions regarding the systematic increase in infections in correspondence with the earlier opening of schools.

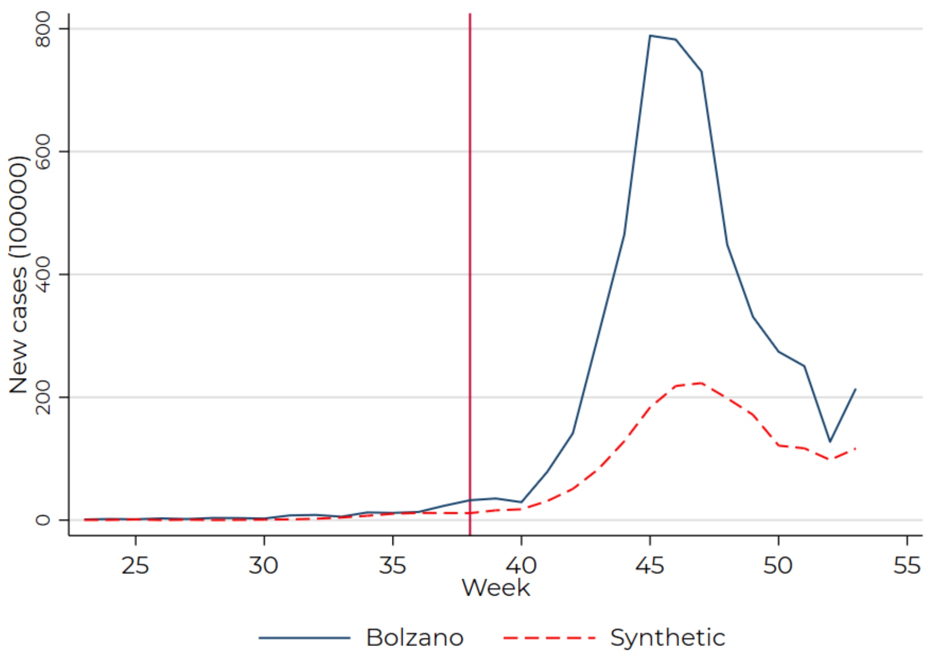

Figure A.3 Incidence of new cases in Bolzano and in its synthetic counterfactual

3 References

Cerulli, G. (2020) 'Nonparametric synthetic control using the npsynth command', *The Stata Journal: Promoting communications on statistics and Stata* 20(4): 844-865.

Cunningham, S. (2021) *Causal Inference. The mixtape.* New Haven: Yale University Press.

Hale, T., Angrist, N., Goldszmidt, R., et al. (2021) 'A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker)', *Nature Human Behaviour* 5(4): 529-538.

O'Neill, S., Kreif, N., Grieve, R., et al. (2016) 'Estimating causal effects: considering three alternatives to difference-in-differences estimation', *Health Services and Outcomes Research Methodology* 16(1-2): 1-21.