DOCUMENTOS
DE TRABAJO

The Effectiveness of School Closures and Other Pre-Lockdown COVID-19 Mitigation Strategies in Argentina, Italy, and South Korea

Claudio Neidhofer y Guido Neidhofer

Documento de Trabajo Nro. 266
Agosto, 2020
ISSN 1853-0168
www.cedlas.econo.unlp.edu.ar

Cita sugerida: Neidhofer, C. y Neidhofer, G. (2020). The Effectiveness of School Closures and Other Pre-Lockdown COVID-19 Mitigation Strategies in Argentina, Italy, and South Korea. Documentos de Trabajo del CEDLAS Nº 266, Agosto, 2020, CEDLAS-Universidad Nacional de La Plata.

The Effectiveness of School Closures and Other Pre-Lockdown COVID-19 Mitigation Strategies in Argentina, Italy, and South Korea

Claudio Neidhöfer

Institute of Medical Microbiology, Immunology and Parasitology, University Hospital Bonn

Guido Neidhöfer*

ZEW - Leibniz Centre for European Economic Research, Mannheim

This version: July 6, 2020†

Abstract The rapid spread of COVID-19 forced policy-makers to swiftly find solutions to reduce infection rates and keep mortality as low as possible. Empirical analyses on the effectiveness of control measures are hereby of primary importance. School closures were among the earliest measures enacted by the governments of most countries. However, while schools are now reopening in many countries, the impact of school closures on the course of the epidemic is still an open question. Adopting parametric and non-parametric synthetic control methods we estimate the effectiveness of pro-active school closures, and other early social distancing interventions, in three countries that reacted relatively early during the course of the pandemic. Our findings suggest that these interventions were effective at reducing the mortality rate of COVID-19, especially when enacted early.

Keywords: COVID-19, policy evaluation, school closures, mitigation strategies.
JEL: I18, J18

*Corresponding author.

†We gratefully acknowledge Achim Hörauf, Carlos Burton, Friedhelm Pfeiffer, Leonardo Gasparini, Maria Massi, Ralf Wilke, and Sarah McNamara for fruitful discussions and comments on earlier drafts. Any remaining errors are solely ours. Claudio Neidhöfer, Institute of Medical Microbiology, Immunology and Parasitology, University Hospital Bonn, Venusberg Campus 1, 53127 Bonn, Germany. claudio.neidhoefer@ukbonn.de. Guido Neidhöfer, ZEW - Leibniz Centre for European Economic Research, Mannheim. L 7, 1, 68161 Mannheim, Germany, and visiting researcher at CEDLAS (National University of La Plata). guido.neidhoefer@zew.de. We have no conflict of interest to declare.

1 Introduction

Recent evidence shows that mitigation strategies and social distancing policies, particularly the national lockdowns enacted in many countries worldwide, have contributed substantively to reducing the spread of COVID-19 [1, 2, 3, 4]. Furthermore, the moderate impact of the disease in places like Singapore and Hong Kong has been attributed to early government action and social distancing measures [5]. Indeed, past implementation of non-pharmaceutical interventions, like pro-active school closures, has significantly reduced the spread and associated mortality rates of other diseases, especially when these measures were enacted in the early phase of epidemics [6, 7]. Quantifying the effectiveness of early, less stringent interventions is of crucial importance. First, in order to understand whether relaxing these measures may lead to another dramatic rise in COVID-19 infections and deaths. And, second, to help evaluating which measures should be enacted as early as possible in case of further increases in infection rates.

We measure the effect of pro-active school closures and other interventions in three countries with very different characteristics, namely Argentina, Italy and South Korea. What these countries have in common, however, is a relatively early reaction to the spread of COVID-19. Furthermore, in the three countries schools were closed before the national lockdown and, hence, a separate evaluation of their impact is possible. Applying parametric and nonparametric synthetic control methods, we construct counterfactual scenarios for the shape of these countries' respective epidemic curves in the absence of interventions. We build different counterfactual scenarios based on the observed development of the epidemic in countries where the interventions were enacted later or not at all. Hereby, we also account for the growth rate of cases and the number of deaths before the intervention, demographic characteristics, differences in health systems, GDP per capita, and mobility patterns reported by Google.

Our findings show that early interventions, including nationwide school closures, had a substantial and significant impact in the three countries analysed here; reducing the total amount of COVID-19 deaths and flattening the epidemic curve. Although it is not possible to completely abstract the effect of school closures from the effects of other contemporaneous measures, as other studies have highlighted, we minimize the bias derived from the impact of the national lockdown by taking into account the incubation time of the disease and the timing from infection to death. Our preferred estimates - those that in the main analysis are obtained with the smallest root mean squared prediction error - indicate that the interventions prevented 84%, 29%, and 91% COVID-19 deaths in Argentina, Italy, and South Korea, respectively, in comparison to a counterfactual projection. These results are robust across different specifications and show that the effectiveness increases the earlier interventions are enacted.

The remainder of the paper is structured as follows: Section 2 discusses the trade-off concerning school closures. Section 3 and 4 describe the applied synthetic control method and the data, respectively. Section 5 shows the results and describes the performed sensitivity analyses. Section 6 discusses the main implications of our findings.

2 Background

By April 10, 2020, 188 countries had closed their schools and early childcare educational facilities over the entire national territory due to the COVID-19 pandemic. This situation affected more than 1.5 billion learners worldwide. However, the evidence on the direct effectiveness of school closures to curb the spread of COVID-19 is surprisingly rather scant [8]. Hence, in times of COVID-19, school closures persist as an object of debate [5]. A few countries did not close their schools at all during the pandemic, while several others are currently discussing a gradual reopening.

Arguably, a governmental decision to close schools and universities is less demanding and politically controversial than other measures to limit interpersonal contact, like a public transportation shut-down and other mitigation strategies with greater direct societal and economic implications. On the other hand, as highlighted by several studies on the economic impact of school closures in the case of other diseases, this measure is costly on several dimensions [9, 10, 11]. For instance, among other costs, prolonged closures cause parental absence from work, a learning loss for the children, bear risks for children's physical and mental health, and may contribute to less economic growth and increased social inequality in the long run [12, 13, 14, 15, 16]. Furthermore, they may also lead to a reduced availability of medical personnel due to child-care obligations [17]. Hence, the governmental decision to institute a nationwide school closure can be seen as a signal. It shows that the information available to governments at that moment in time led them to believe that this would be the right moment to intervene, and that, despite the associated disadvantages, closing schools had become a (first) cost-effective measure to reduce disease transmission [18]. In this sense, it is worthwhile to empirically assess whether the timing of this first set of interventions, including school closures, has been effective at curbing the impact of the disease.

Little is yet known about the progression and transmission of COVID-19 in children and adolescents. Besides isolated cases of critical neonatal coronavirus pneumonia [19, 20, 21], children have not been prominently featured in COVID-19 case statistics [22, 23]. While children of all age groups have been described as being susceptible in China [24], evidence from Iceland and Italy rather suggests that they might be less suscetible than adults [25, 26]. Although estimates of the case fatality rates of COVID-19 for infants, children, adolescents, and individuals younger than 30 are very low, it is expected that the number of infected among these groups are considerable [27].[1] As was the case in the SARS-COV and MERS-CoV epidemics, children appeared to display a much milder and shorter course of infection than adults [30, 31, 32, 33, 34, 35] which might further account for their underrepresentation in the confirmed case count [36, 37]. Per current knowledge, infected children seem to carry a similar viral load to infected adults and are, therefore, likely equally responsible for the transmission of COVID-19 [38, 39]. In contrast, other studies found that children with COVID-19 rarely caused outbreaks within their households, and it has been argued that children are unlikely to be the main drivers of the pandemic [40].

The effects of school closures are twofold, working through the social distancing of children and their parents who have to stay home to take care of them. Past evidence

[1]In addition, recent evidence highlights that there could be a relationship between the SARS-COV-2 virus and a serious hyper-inflammatory shock reaction observed in an increasing number of children [28, 29].

shows that school-closures have a substantial decreasing effect on the incidence rates of other illnesses [41, 42, 43, 44], with a more prominent effect when enforcement is timely [45, 46] and when children play a central role in initial community transmission [47, 48, 49]. School closures are, for instance, commonly implemented for mitigating influenza pandemics; lowering peak attack rates as well as the total number of cases [50, 51]. For the current pandemic, while most studies agree on the effectiveness of national lockdowns to reduce the transmission of COVID-19, the effectiveness of school closures and other early interventions seems to remain an open question. Part of the scant existing evidence shows that early interventions, including nationwide school closures, contributed to reducing the spread of COVID-19 in China and Germany [1, 2]. In contrast, the findings for 11 European countries attribute a large transmission reduction effect to the national lockdown, with a substantially and significantly smaller effect attributed to other interventions, like the banning of public events and school closures [3]. The existing studies agree that it does not seem to be fully possible to disentangle the effect of school closures from other contemporaneous mitigation measures.

3 Method

Our aim is to assess the effectiveness of school closures for mitigating the effects of COVID-19. For this purpose, we quantify the impact of this intervention in three countries that closed their schools on a nationwide level relatively early during the course of the pandemic: Argentina, Italy, and South Korea. For each of these countries these measures were not enacted on the same day as the national lockdown, which took place with some delay. Since reported case numbers of the disease are heavily influenced by testing and heterogeneous reporting across countries, we focus on the effect on the number of reported COVID-19 deaths, which suffer less from these issues, both in absolute terms and relative to population size. To measure the effect, we compare the shape of the epidemic curves in these countries after the nationwide school closure with an hypothetical curve that shows the situation in absence of the intervention.

Consistent with other methodologies for measuring the effectiveness of an intervention, our setup is also based on the assumption of a counterfactual scenario. Hereby, the fundamental issue is the definition of a suitable control unit allowing for a consistent comparison. Comparing the country in which the intervention was enacted (treated unit) with just one other country (control unit) would result in a very imperfect measure of the effect, unless the countries are almost perfectly comparable with respect to demographic characteristics of their population, health system capacity, and course of the epidemic. In order to provide a more suitable counterfactual, we apply synthetic control methods [52, 53]. This method enables us to construct control units based on the disposable information on characteristics of countries as well as on the COVID-19 spread before the day of the intervention. In this set up, the synthetic control unit is the weighted average of the epidemic curves of all countries within the so called *donor pool* that best mimics counterfactual trends for the treated unit (Argentina, Italy, or South Korea) in the absence of treatment (nationwide school closures). Synthetic control units estimated with this methodology provide a more appropriate comparison to the affected unit than any other unaffected unit taken individually [54].

Once suitable control units are established, the effectiveness of school closures to reduce the fatal impact of COVID-19 can be measured. To quantify the number of deaths avoided due to the intervention, comparing the three countries under analysis with their respective synthetic control unit, we apply a *Difference in Differences* set up (DID). DID has a long tradition in epidemiological and public health studies since the investigation on the 1854 outbreak of cholera by John Snow [55, 56]. In modern econometrics, DID is a widely used and established policy evaluation technique for causal inference; in particular when extended by synthetic control methods [57]. Formally, we estimate the following empirical model

$$d_{it} = \alpha + \eta_i + \lambda_t + \delta S_{it} + \varepsilon_{it} \tag{1}$$

where d is the average number of daily COVID-19 deaths in country i (the treated unit or the respective synthetic control unit) at time t. Hereby, the model is estimated separately for each treated-control unit pair.[2] For the treated unit (Argentina, Italy, or South Korea) t is 1 from the day of school closures on, and 0 before that date. For the control unit, t equals 1 from the day when the countries reported at least the same number of aggregate total COVID-19 deaths as the respective treated unit on the day of nationwide school closure. For instance, measuring the impact for Italy, t is 0 before the countries reported 80 total deaths, which in Italy is equivalent to the day of nationwide school closure, and 1 afterwards. The number of aggregate COVID-19 deaths in the two other treated units is 2 in Argentina, and 22 in South Korea.

S_{it} is a binary variable that is 1 for the treated unit after school closure and 0 for the treated unit before school closure, as well as for the control unit. The estimate of the coefficient δ measures the difference between the number of COVID-19 deaths before and after school closure in the treated unit, as compared to the increase in deaths in the control unit. η and λ are country and time fixed effects. α is a constant and ε the error term assumed to be independent and identically distributed across observations.

The key assumption of the identification strategy is that the increase in the number of deaths in the treated unit would have followed a parallel trend to the one observed in the control unit in absence of the implemented school closure. The difference in country differences between the two time periods is:

$$DD = (d_{Treated,t=1} - d_{Treated,t=0}) - (d_{Control,t=1} - d_{Control,t=0}) \tag{2}$$

and inserting the first in the second equation yields:

$$DD = \delta + (\varepsilon_{Treated,t=1} - \varepsilon_{Treated,t=0}) - (\varepsilon_{Control,t=1} - \varepsilon_{Control,t=0}) \tag{3}$$

If the epidemic curve in both units would have followed a common trend in absence of the school closure, then $(\varepsilon_{Treated,t=1} - \varepsilon_{Treated,t=0}) - (\varepsilon_{Control,t=1} - \varepsilon_{Control,t=0}) = 0$, and the coefficient δ gives an unbiased estimate of the effect of school closures in this natural experiment set up. It is not possible to test the parallel trends assumption directly, because it is based on a counterfactual scenario. Analyzing the trends before the event (when $t < 0$) is a second best alternative to verify it. Based on a set of covariates (described below), the synthetic control method makes sure that the treated and control units, as well as their outcome-trends before the event, are as comparable as possible

[2] Hence, the three countries under analysis may belong to the donor pool if they are not analysed as the treated unit.

to each other. Following the standard approach, we let the model define the weights to construct the synthetic control unit based on parametric procedures. In a sensitivity analysis, we define the control unit based on a flexible nonparametric construction of the weights [58].

For the method to create a suitable control unit for each treated unit, we choose criteria to restrict the donor pool. First, we include only those countries that did not close their schools on a nationwide level, or that had reported more accumulated deaths than the treated unit on the day of school closure. Then, the weights to construct the synthetic control group are estimated based on the following country-level characteristics (*predictor variables*): total population, population density, median age, population share over the age of 65, GDP per capita, hospital beds per 100,000 inhabitants, public health expenditures, average number of reported COVID-19 deaths before day zero, growth rate of reported COVID-19 cases with respect to the day before[3], and mobility patterns retrieved from Google Mobility Reports. The inclusion of these predictor variables ensures that even in case of very low and stable pre-treatment number of deaths in the treated unit, the predicted control unit is highly comparable and best mimics the development of the epidemic curve in absence of the intervention. The amount of COVID-19 deaths of the synthetic control unit is the weighted average of the observed deaths in the control group countries. We estimate four distinct synthetic control units, successively including the aforementioned variables, as well as one nonparametric synthetic control unit. Hereby, one of the scenarios controls for cultural differences by including Asian countries only in the donor pool for the comparison with South Korea, while for the comparison with Argentina and Italy only non-Asian countries are included. Table 1 shows the variables included in each scenario, the countries included to form the synthetic control units and their respective weights, as well as the root mean square prediction error (RMSPE) of the pre-trends analysis.

However, the effects of the intervention might not be immediately measurable from the day of school closure, because of the time that must pass between infection and death. Studies for China have estimated a median incubation period for COVID-19 of around 5 days [59, 60, 37] and around 16-18 days from symptom onset to death [27, 61]. A report from the Italian National Institute of Health reported the time between symptoms and death to be 10-13 days [62]. Hence, a measurable reduction in the number of deaths might be observable only after 15 days from the intervention. Hence, to obtain a conservative lower bound of the effect, we show also estimates setting $t = 0$ for all days before the 15th day since the day of school closure (or the respective day zero in the control unit).

Furthermore, since our aim is to estimate the effect of school closures and early interventions, we have to take into account that other events might have mitigated the outbreak of COVID-19 besides the closure of schools. While it is not possible to disentangle the effect of mitigation measures that were enacted on the same date, we are able to abstract from events that happened after the day of school closure. For instance, on March 9 (5 days after the nationwide closure of schools), the Italian government imposed a national quarantine with the objective to induce social distancing among the population in order to reduce the outbreak. Interestingly, the same amount of days lie between school closures and national lockdown in Argentina as well (March 16 and March 20). These events could bias the size of the estimate. The magnitude of δ would not just show the

[3] The growth rate is measured as $log((cases_{day=T} - cases_{day=T-1})/cases_{day=T-1})$

effect of the nationwide school closure, but an overall effect of both mitigation strategies taken together. To get a sense of the pure effect of the pre-lockdown interventions, we restrict the time frame of analysis to avoid bias in our estimates from the effects of the national quarantine. To this end, we truncate the analysis 14 days after the beginning of the national quarantine (because on the 15th day, the effects of the quarantine on the number of deaths should begin to be measurable).[4] For Argentina and Italy, this is equivalent to 18 days after the day of nationwide school closures. To warrant comparison we also chose the same conservative time window for South Korea, although their national lockdown was enacted much later.

4 Data

Our main source of data is the harmonized statistics retrieved from the project *Our World in Data*.[5] Here, daily harmonized statistics on COVID-19 cases and deaths are retrieved from the European Center for Disease Prevention and Control (ECDC). Other country-level aggregate statistics that we use are included in the data set and are derived from different sources.[6] We complement this data with information on school closures that we retrieve from the UNESCO.[7] For each country we compute and report the days since country-wide school closure. Finally, we include country-level information on population mobility patterns retrieved from the Google Mobility Reports.[8]

5 Results

5.1 Synthetic Control Evidence

Table 2 shows the estimates of η (Treated unit=1), λ (Post=1), and δ (Treated unit=1 $\times$ Post=1) obtained by linear regression analysis. The estimates are obtained separately for Argentina, Italy, and South Korea as the treated unit. As mentioned, the time series are truncated 14 days after the day of national lockdown to diminish the bias in the estimates of the effect of school closures and other contemporaneous early measures. All regressions include a polynomial time trend of the third degree as control variables. Eight different regressions are estimated and their results displayed in different columns. Each row displays the estimated coefficients of the regressions.[9] The variables and countries

[4]Recognizing that the effect of the interventions on deaths is delayed, our correction for this relies on the delay between infection and death being 15 days. Since there might be considerable dispersion in the time to death in what follows we also show the resulting epidemic curves to allow for a visual inspection of the differences between treated and control unit over the entire period after school closure to 15 days after the start of the national quarantine.

[5]Retrieved on May 10, 2020 from https://ourworldindata.org/coronavirus. We obtain number of cases and deaths on missing dates by linear interpolation.

[6]See the project website https://ourworldindata.org/ for more information on the single sources for each item.

[7]Retrieved on April 10, 2020 from https://en.unesco.org/covid19/educationresponse.

[8]Retrieved on April 27, 2020 from https://www.google.com/covid19/mobility/

[9]The statistical significance of the results is robust when we compute standard errors by the heteroskedasticity robust (White-Huber) correction or by bootstrapping with N replications (N is the number of observations).

included to estimate the counterfactual synthetic control units in the scenarios I-IV are indicated in Table 1. The estimates of η show the mean difference in the average daily number of deaths between the treated unit and the synthetic control unit before day zero, while the estimate of λ shows the mean change in the average daily deaths in the control unit from before to after day zero. The first four columns shows the baseline estimates, namely from the days of school closures. The next four columns show the estimated effect taking into account incubation period and time from the first symptoms to death; hence, starting 15 days after the intervention.

The estimates of δ (Treated unit=1×Post=1) show the effects of the nationwide school closure, and of the other contemporaneous measures, on the daily average number of reported COVID-19 deaths. We observe that, in all scenarios and for all three countries, δ is negative and statistically significant. This shows in the three countries considered here the enacted early measures have been effective in reducing the number of COVID-19 deaths.

For the ease of comparison, and to give a clearer picture of the external validity of these findings, the effect size of the point estimates relative to the average daily deaths in the counterfactual scenario is also indicated in the last row of Table 2. We observe that the estimated effect of the interventions ranges from a 63% to a 90% reduction in daily average deaths in Argentina, from 21% to 35% in Italy, and from 72% to 96% in South Korea. The baseline models with the lowest RMSPE yield hereby an effect size of 84%, 29%, and 91% for Argentina, Italy, and South Korea, respectively.

To visualize these effects, Figure 1 shows the epidemic curves of total COVID-19 deaths for Argentina, Italy, and South Korea from 18 days before to 18 days after the day of nationwide school closure. In the synthetic control units, day 0 is set as the day in which at least the same amount of aggregate total COVID-19 deaths were reported as in the respective treated unit. The second vertical line indicates day 13, i.e. 14 days after the day of the intervention. As is evident, the shape of the epidemic curves is much steeper in all synthetic control units compared to the treated unit.

5.2 Sensitivity Analysis

5.2.1 Nonparametric Construction of Counterfactual

We perform a sensitivity analysis extending the prediction of the synthetic control unit by nonparametric estimation of the weights. Although both methods have been proven to provide a small prediction error, nonparametric estimation may slightly outperform the parametric one [58]. The method uses a kernel function with pre-fixed bandwidth to predict the counterfactual. Following a cross validation approach, which is usual for nonparametric estimation, the bandwidth is chosen that minimizes the root mean squared prediction error (RMSPE) in the pre-intervention period.

We estimate the model with nonparametric estimation including the combination of predictor variables that minimizes the RMSPE. For comparison purposes, we estimate the same model with the classical synthetic control method. Figure 2 shows the comparison of the pretrends, i.e. trends of the epidemic curves before day 0, and the RMSPE using both methods. The visual inspection shows rather similar patterns, especially in the pre-trend period close to the intervention, while the RMSPE indicates that by nonparametric estimation of the counterfactual we obtain a slightly more consistent control unit. Table

3 shows that the results obtained adopting this method are qualitatively similar for Argentina and South Korea, while lower, and in one application not significantly different from zero, for Italy. The baseline estimates indicate an effect size of 92%, 11%, and 98% for Argentina, Italy, and South Korea, respectively. Figure 3 shows the resulting epidemic curve of accumulated deaths.

5.2.2 Estimates per 100,000 inhabitants

Especially at the beginning of the epidemic, the absolute number of deaths, rather than relative to the population size, seemed more informative for influencing governmental interventions and measuring the effectiveness of these measures. Nevertheless, to estimate the effect of the interventions relative to the population size of the countries is an interesting further sensitivity analysis. Table 4 and 5 show the estimated effects of the intervention on average daily deaths relative to the population size of the country (per 100,000 inhabitants). All results basically confirm the existence of a substantial and significant reduction in the number of deaths as a consequence of the interventions.[10]

5.2.3 Placebo

To test the validity of our analysis and of the parallel trends assumptions we perform *Placebo-tests* for each scenario. We test for differences in the number of daily COVID-19 deaths in the treated and control unit before the closure of schools. Table 6 shows the results of this exercise. All analyses show no statistically significant differences before the intervention (i.e. the estimate of δ is not significantly different from zero).

5.2.4 Exponential growth

It could be argued that the applied linear models are not suitable for measuring the effect on non-linear epidemic curves. Hence, we also run the synthetic control method on a log-linear model. Hereby, we estimate the effects of the intervention on the *log* aggregate number of deaths, rather than the daily change in deaths as done so far, because of days in which no new COVID-19 deaths were reported. For the same reason, we cannot rely on such a long pre-trend period as before and must restrict the analysis to three days before the intervention in Italy and South Korea. For Argentina, we are not able to measure logarithmic deaths consistently, because the country closed its schools almost immediately after reporting the first COVID-19 death. Furthermore, since measuring the effect size on the aggregate number of deaths would not be meaningful in this application, we rely only on the visual inspection of the shape of the logarithmic epidemic curves, shown in Figure 4. The application for all parametric and non-parametric synthetic control units shows that the epidemic curves in Italy and South Korea have a less steep shape than the counterfactual without the intervention.

[10]The only exception are the estimates obtained in scenario III for South Korea. As mentioned, this scenario includes only Asian countries in the donor pool, namely China and Indonesia. These two countries have a much higher population size than South Korea. Hence, the counterfactual predicted from this application is a poor match and yields inconsistent estimates. For uniformity with respect to the main analysis, we still opt to report these results.

5.2.5 Excess mortality

As mentioned previously, the accuracy of reported COVID-19 cases in official statistics has been questioned, as it should depend, among other factors, on the quantity of tests performed. To further prove the robustness of our estimates we run the same empirical models described above on an alternative data source.

Instead of reported COVID-19 deaths, we retrieve weekly data on excess mortality - i.e. the difference between overall deaths, reported by national official sources in a particular time interval, and the median value of deaths in the equivalent period from 2015 to 2019 - collected by the Financial Times.[11] While not all of these deaths are necessarily attributable to COVID-19, many of them caused by other diseases might depend on the overuse of the health care system due to the pandemic. Hence, the dataset offers a valid alternative to verify the results obtained so far.[12] Unfortunately, among the countries under analysis only Italy is included in this data.

Figure 5 shows the trends in excess mortality for Italy and the synthetic counterfactual units.[13] The timing of the intervention is measured in analogous way as in the preceding analyses: 0 is the week of nationwide school closure in Italy, and in the other countries the week in which at least 80 aggregate COVID-19 deaths were reported. Again, to avoid the effects of the national lockdown biasing our estimates, we restrict the observation period to the third week after school closure, and choose a symmetric time window to measure the trends in the period before the intervention. The predictor variables of the estimated scenarios are the same as before.

Since the excess mortality is measured at the end of the week, we already observe an effect in the week of nationwide school closure. Table 7 shows the estimated effect sizes. They measure the effect of school closures and other early interventions in Italy between 27% and 43% of the number of excess deaths in the synthetic control unit. Due to the low number of observations, the coefficients are not statistically significant. Nevertheless, they support the results obtained in the main analysis.

5.3 Contemporaneous social isolation measures

As higlighted by previous studies, to assess which part of the estimated effects can be attributed to proactive school closures, rather than to the other contemporaneous interventions, is not directly feasible when observing single countries. In this sense, our analysis has the advantage of measuring the effect in multiple countries which experienced very different development paths of the epidemic. Hence, it is informative to observe which other social isolation measures were already in place in Argentina, Italy, and South Korea, on the day of school closures, and which one were enacted on the same day or shortly after.

We retrieve this information from the Oxford COVID-19 Government Response Tracker [63]. The data set collects government responses to the pandemic, codes these into indi-

[11] Retrieved on June 12, 2020 from https://github.com/Financial-Times/coronavirus-excess-mortality-data . Sources of the mortality data for single countries can be consulted there.

[12] Although the correlation across countries between reported COVID-19 deaths and excess mortality is very high (.95), the reported excess mortality is usually higher, even by several thousand deaths; up to almost three times the number of weekly COVID-19 deaths in Italy.

[13] In this application, the countries in the donor pool for which data on excess deaths is available, are Spain, France, the UK, Sweden, and the US.

cators, and creates a composite measure, the *Stringency Index*, by summing the scores of these indicators. Figure 6 shows the variation in the responses of the governments in Argentina, Italy, and South Korea, measured by the Stringency Index, which varies on a scale from 0 to 100.

We observe that on the day of school closures the three countries show rather different values of the Stringency Index: in Argentina 41.67, in Italy 74.54, and in South Korea 55.56. The measures shared by the three countries on the day of nationwide school closure are, besides the school closure itself, the banning of public events, the restriction of international flights, contact tracing, and public information campaigns; in contrast, public transport was not shut down in neither of them. No other common pattern of government intervention is observable; neither on advice, requirements for workplace closure and stay-home measures, nor on internal movement controls. As mentioned before, and taken into account when we restrict the time window of our empirical evaluation, few days after the school closure all three countries adopted more stringent measures (evident by a drastic increase in the index score in Figure 6).

5.4 Later school closures

Our results show that school closures and other early interventions had a substantial effect. Now, we would like to establish whether these measures might not have the expected effect if enacted later during the course of the epidemic. For this purpose, we apply the model estimated before (including the predictor variables used in scenario II) on several other countries. Hereby, we order the countries by the number of reported COVID-19 deaths on the day of school closures.

Table 8 shows the results of this application, which mainly confirm the findings obtained so far. The later schools were closed nationwide during the course of the pandemic, the lower the effectiveness of this measure. Indeed, in Brazil, France, the UK, and Spain, we do not detect any reducing effect of the early interventions on the number of deaths. Surprisingly, in Germany we also find only a very small and not statistically significant effect, although the country closed its schools nationwide relatively early during the epidemic course, measured by the number of reported deaths.[14] Of course, the uncovered effects in these countries might be upward biased by other parallel social distancing measures, or downward biased by particular incidents, like superspreading events. Hence, the evidence presented here for these additional countries should be interpreted cautiously.

Exactly measuring the effect in each of them is beyond the scope of this work, but remains a subject of great interest for future research. For instance, the differential impact of early school closures in different countries might depend on the characteristics and universality of the health care system. Furthermore, it might be influenced by other social distancing policies enacted earlier or at the same time. To get a first, stylized picture of these mechanisms, we observe the Stringency Index on the day of school closures also for the countries included in Table 8 and do not find support for the latter hypothesis. As shown in the left graph in Figure 7, excluding the extreme values of the UK, no systematic

[14] This result is not contradicting the evidence showing a decreasing spreading rate for Germany associated to several interventions [1]. Rather, they confirm that the German context is very specific. On the one hand, the policies enacted before, e.g. the cancellation of large public events and information campaigns, might have been effective to curb the spread. On the other, the characteristics of the health system might have contributed to reduce the associated mortality.

pattern in the relationship between the size of the effect and the stringency of measures can be observed. In contrast, the right graph in Figure 7 shows a suggestive correlation between the effect size of early interventions and public health expenditures.

This preliminary evidence suggests a possible interaction between the timing of intervention and the health care system. Hence, in countries with less extensive public health care an early response to an epidemic seems fundamental to limit the disease. A deeper analysis would exceed the framework of this evaluation, but these, and other possible mechanisms should be analysed in detail in future research.

6 Discussion

The COVID-19 pandemic hit the world unexpectedly forcing policy makers to take actions quickly. Most countries implemented large-scale control strategies to limit the spread of COVID-19, such as travel restrictions, school closures, or national quarantine. When these decisions were made most facts about the characteristics of the virus, the course of the disease, its basic reproduction number, and many other relevant factors were unknown. This information is necessary, however, for governments to justify their actions, for instance on the choice of containment measures. All of these measures bear substantive economic and social costs that urge governments to constantly revisit their strategies, evaluating costs and benefits of the enacted measures. Hence, empirical evidence on the effectiveness of the implemented measures is fundamental.

Applying synthetic control methods, an alternative methodology to epidemiological modelling based on observational data, we provided a novel quantitative assessment about the effectiveness of early non-pharmaceutical interventions to reduce the fatal impact of COVID-19. In our application, we evaluated the mortality of COVID-19 before and after these interventions in Argentina, Italy, and South Korea constructing counterfactual scenarios that mimic the country specific trend in reported deaths while accounting for demographic characteristics, population densitiy, GDP per capita, health care systems, and mobility patterns. The early measures that the three countries analysed here have in common are proactive school closures, banning of public events, the restriction of international flights, contact tracing, and public information campaigns. Although, as previous studies on this subject highlighted, it is not entirely possible to disentangle the single effect of each of these early measures, we do our best to minimize the potential bias deriving from later measures. Our findings confirm that the early, pre-lockdown measures were effective, and suggest increased effectiveness the earlier these measures were enacted during the course of the epidemic. We find a sizeable and robust impact on reducing the number of reported COVID-19 deaths - not only in absolute terms, but also per 100,000 inhabitants, and on a logarithmic scale - as well as on reducing excess mortality reported by statistical offices.

Among the studied interventions, the most stringent one is, arguably, the proactive closure of schools. Its effects work through different channels, namely the social distancing of children and their parents, as well as a signaling value to increase attentiveness among the population. It should be the pre-lockdown intervention affecting the largest share of the population and, hence, with the strongest effect on reducing disease transmission. Indeed, it has been shown that other interventions, for instance even very restrictive travel limitations, may only have a modest effect on reducing the transmission of COVID-19 if not accompanied by other measures [64].

We hope our analysis will help to inform policy-makers about the effectiveness of the implemented early measures, and to motivate further research in this direction. Research on this topic is of great importance, especially in light of the fact many countries are gradually re-opening their schools. Policy makers should hereby take several aspects into account when making future decisions, evaluating the trade-off between mortality risk and social cost. On the one hand, it is clear that the detrimental side effects of prolonged school closures on learning losses, educational and gender inequality, children's health, and family well-being will constitute a major challenge for our society. On the other, a scientific evaluation of the effectiveness associated with school closures is necessary to understand the costs of reopening schools during the course of the pandemic. As our analysis suggests, with early interventions, including the proactive closures of schools, governments have been able to reduce the fatal impact of COVID-19. If relaxing these measures would again induce a dramatic rise in infections and deaths, governments might again see themselves forced to adopt more stringent measures, like curfews and economic shutdowns. In comparison, keeping schools closed until more information is available on how to reduce the risk of infection might be more cost-effective, at least in the short run.

Surely, even when prolonging school closures might be useful to save lives, it should only be a viable option if accompanied by a proper and well-developed plan for home-schooling and online learning that reaches all children and adolescents, especially the most vulnerable. And, clearly, further evidence is needed for a more comprehensive picture of the effectiveness of school closures for reducing the impact of COVID-19, as well as the short and long term costs attached to this mitigation policy.

References

[1] Dehning J, Zierenberg J, Spitzner FP, Wibral M, Neto JP, Wilczek M, et al. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. Science. 2020;Available from: https://science.sciencemag.org/content/early/2020/05/14/science.abb9789.

[2] Zhang J, Litvinova M, Liang Y, Wang Y, Wang W, Zhao S, et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. Science. 2020;Available from: https://science.sciencemag.org/content/early/2020/05/04/science.abb8001.

[3] Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature. 2020;p. 1–8.

[4] Qiu Y, Chen X, Shi W. Impacts of social and economic factors on the transmission of coronavirus disease 2019 (COVID-19) in China. Journal of Population Economics. 2020;p. 1.

[5] Anderson RM, Heesterbeek H, Klinkenberg D, Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? The Lancet. 2020;395(10228):931 – 934. Available from: http://www.sciencedirect.com/science/article/pii/S0140673620305675.

[6] Hatchett RJ, Mecher CE, Lipsitch M. Public health interventions and epidemic intensity during the 1918 influenza pandemic. Proceedings of the National Academy of Sci-

ences. 2007;104(18):7582–7587. Available from: https://www.pnas.org/content/104/18/7582.
[7] Markel H, Lipman HB, Navarro JA, Sloan A, Michalsen JR, Stern AM, et al. Nonpharmaceutical Interventions Implemented by US Cities During the 1918-1919 Influenza Pandemic. JAMA. 2007 08;298(6):644–654. Available from: https://doi.org/10.1001/jama.298.6.644.
[8] Viner RM, Russell SJ, Croker H, Packer J, Ward J, Stansfield C, et al. School closure and management practices during coronavirus outbreaks including COVID-19: a rapid systematic review. The Lancet Child Adolescent Health. 2020;Available from: http://www.sciencedirect.com/science/article/pii/S235246422030095X.
[9] Smith RD, Keogh-Brown MR, Barnett T, Tait J. The economy-wide impact of pandemic influenza on the UK: a computable general equilibrium modelling experiment. Bmj. 2009;339:b4571.
[10] Lempel H, Epstein JM, Hammond RA. Economic cost and health care workforce effects of school closures in the US. PLoS currents. 2009;1.
[11] Adda J. Economic Activity and the Spread of Viral Diseases: Evidence from High Frequency Data *. The Quarterly Journal of Economics. 2016 02;131(2):891–941. Available from: https://doi.org/10.1093/qje/qjw005.
[12] UNESCO. Adverse consequences of school closures;. Available from: https://en.unesco.org/covid19/educationresponse/consequences.
[13] Lancker] WV, Parolin Z. COVID-19, school closures, and child poverty: a social crisis in the making. The Lancet Public Health. 2020;Available from: http://www.sciencedirect.com/science/article/pii/S2468266720300840.
[14] Wang G, Zhang Y, Zhao J, Zhang J, Jiang F. Mitigate the effects of home confinement on children during the COVID-19 outbreak. The Lancet. 2020;395(10228):945 – 947. Available from: http://www.sciencedirect.com/science/article/pii/S014067362030547X.
[15] Lee J. Mental health effects of school closures during COVID-19. The Lancet Child & adolescent health. 2020 Apr;p. S2352–4642(20)30109–7. 32302537[pmid]. Available from: https://pubmed.ncbi.nlm.nih.gov/32302537.
[16] Fuchs-Schündeln N, Kuhn M, Tertilt M, et al. The Short-Run Macro Implications of School and Child-Care Closures. Institute of Labor Economics (IZA); 2020.
[17] Bayham J, Fenichel EP. Impact of school closures for COVID-19 on the US health-care workforce and net mortality: a modelling study. The Lancet Public Health. 2020;Available from: http://www.sciencedirect.com/science/article/pii/S2468266720300827.
[18] Hollingsworth TD, Klinkenberg D, Heesterbeek H, Anderson RM. Mitigation Strategies for Pandemic Influenza A: Balancing Conflicting Policy Objectives. PLOS Computational Biology. 2011 02;7:1–11. Available from: https://doi.org/10.1371/journal.pcbi.1001076.
[19] Wang S, Guo L, Chen L, Liu W, Cao Y, Zhang J, et al. A Case Report of Neonatal 2019 Coronavirus Disease in China. Clinical Infectious Diseases. 2020 03;Ciaa225. Available from: https://doi.org/10.1093/cid/ciaa225.
[20] Zeng L, Xia S, Yuan W, Yan K, Xiao F, Shao J, et al. Neonatal Early-Onset Infection With SARS-CoV-2 in 33 Neonates Born to Mothers With COVID-19 in Wuhan, China. JAMA Pediatrics. 2020 03;Available from: https://doi.org/10.1001/jamapediatrics.2020.0878.

[21] Wang L, Shi Y, Xiao T, Fu J, Feng X, Mu D, et al. Chinese expert consensus on the perinatal and neonatal management for the prevention and control of the 2019 novel coronavirus infection (First edition). Annals of translational medicine. 2020 Feb;8(3):47–47. Available from: https://pubmed.ncbi.nlm.nih.gov/32154287.

[22] Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72â¯314 Cases From the Chinese Center for Disease Control and Prevention. JAMA. 2020 02;Available from: https://doi.org/10.1001/jama.2020.2648.

[23] Lu X, Zhang L, Du H, Zhang J, Li YY, Qu J, et al. SARS-CoV-2 Infection in Children. New England Journal of Medicine. 0;0(0):null. Available from: https://doi.org/10.1056/NEJMc2005073.

[24] Dong Y, Mo X, Hu Y, Qi X, Jiang F, Jiang Z, et al. Epidemiology of COVID-19 Among Children in China. Pediatrics. 2020 Apr;p. e20200702. Available from: http://pediatrics.aappublications.org/content/early/2020/03/16/peds.2020-0702.1.abstract.

[25] Gudbjartsson DF, Helgason A, Jonsson H, Magnusson OT, Melsted P, Norddahl GL, et al. Spread of SARS-CoV-2 in the Icelandic population. New England Journal of Medicine. 2020;.

[26] Lavezzo E, Franchin E, Ciavarella C, Cuomo-Dannenburg G, Barzon L, Del Vecchio C, et al. Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Voâ. Nature. 2020;.

[27] Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet Infectious Diseases. 2020;Available from: http://www.sciencedirect.com/science/article/pii/S1473309920302437.

[28] Viner RM, Whittaker E. Kawasaki-like disease: emerging complication during the COVID-19 pandemic. The Lancet. 2020 jun;395(10239):1741–1743. Available from: https://doi.org/10.1016%2Fs0140-6736%2820%2931129-6.

[29] Toubiana J, Poirault C, Corsia A, Bajolle F, Fourgeaud J, Angoulvant F, et al. Kawasaki-like multisystem inflammatory syndrome in children during the covid-19 pandemic in Paris, France: prospective observational study. BMJ. 2020;369. Available from: https://www.bmj.com/content/369/bmj.m2094.

[30] Al-Tawfiq JA, Kattan RF, Memish ZA. Middle East respiratory syndrome coronavirus disease is rare in children: An update from Saudi Arabia. World journal of clinical pediatrics. 2016 Nov;5(4):391–396. Available from: https://pubmed.ncbi.nlm.nih.gov/27872828.

[31] Denison MR. Severe Acute Respiratory Syndrome Coronavirus Pathogenesis, Disease and Vaccines: An Update. The Pediatric Infectious Disease Journal. 2004;23(11). Available from: https://journals.lww.com/pidj/Fulltext/2004/11001/Severe_Acute_Respiratory_Syndrome_Coronavirus.5.aspx.

[32] Kelvin AA, Halperin S. COVID-19 in children: the link in the transmission chain. The Lancet Infectious Diseases. 2020 2020/04/07;Available from: https://doi.org/10.1016/S1473-3099(20)30236-X.

[33] Ruan S. Likelihood of survival of coronavirus disease 2019. The Lancet Infectious Diseases. 2020;Available from: http://www.sciencedirect.com/science/article/pii/S1473309920302577.

[34] Ludvigsson JF. Systematic review of COVID-19 in children shows milder cases and a better prognosis than adults. Acta Paediatrica. 2020 Jun;109(6):1088–1095. Available from: https://doi.org/10.1111/apa.15270.

[35] Zimmermann P, Curtis N. Coronavirus Infections in Children Including COVID-19: An Overview of the Epidemiology, Clinical Features, Diagnosis, Treatment and Prevention Options in Children. The Pediatric infectious disease journal. 2020 May;39(5):355–368. Available from: https://pubmed.ncbi.nlm.nih.gov/32310621.

[36] Qiu H, Wu J, Hong L, Luo Y, Song Q, Chen D. Clinical and epidemiological features of 36 children with coronavirus disease 2019 (COVID-19) in Zhejiang, China: an observational cohort study. The Lancet Infectious Diseases. 2020 Mar;Available from: http://www.sciencedirect.com/science/article/pii/S1473309920301985.

[37] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel CoronavirusâInfected Pneumonia. New England Journal of Medicine. 2020;382(13):1199–1207. Available from: https://doi.org/10.1056/NEJMoa2001316.

[38] Jones TC, Mühlemann B, Veith T, Biele G, Zuchowski M, Hoffmann J, et al. An analysis of SARS-CoV-2 viral load by patient age. medRxiv. 2020;Available from: https://www.medrxiv.org/content/early/2020/06/09/2020.06.08.20125484.

[39] Kawasuji H, Takegoshi Y, Kaneda M, Ueno A, Miyajima Y, Kawago K, et al. Viral load dynamics in transmissible symptomatic patients with COVID-19. medRxiv. 2020;Available from: https://www.medrxiv.org/content/early/2020/06/04/2020.06.02.20120014.

[40] Ludvigsson JF. Children are unlikely to be the main drivers of the COVID-19 pandemic–a systematic review. Acta Paediatrica. 2020;.

[41] Halloran ME, Ferguson NM, Eubank S, Longini IM, Cummings DAT, Lewis B, et al. Modeling targeted layered containment of an influenza pandemic in the United States. Proceedings of the National Academy of Sciences. 2008;105(12):4639–4644. Available from: https://www.pnas.org/content/105/12/4639.

[42] Ali ST, Cowling BJ, Lau EHY, Fang VJ, Leung GM. Mitigation of Influenza B Epidemic with School Closures, Hong Kong, 2018. Emerging infectious diseases. 2018 November;24(11):2071â2073. Available from: https://europepmc.org/articles/PMC6200008.

[43] Sypsa V, Hatzakis M. School closure is currently the main strategy to mitigate influenza A(H1N1)v: a modeling study. Euro Surveil. 2009;14.

[44] Markel H, Lipman HB, Navarro JA, Sloan A, Michalsen JR, Stern AM, et al. Nonpharmaceutical Interventions Implemented by US Cities During the 1918-1919 Influenza Pandemic. JAMA. 2007 08;298(6):644–654. Available from: https://doi.org/10.1001/jama.298.6.644.

[45] Halder N, Kelso JK, Milne GJ. Developing guidelines for school closure interventions to be used during a future influenza pandemic. BMC infectious diseases. 2010 Jul;10:221–221. Available from: https://pubmed.ncbi.nlm.nih.gov/20659348.

[46] Ebrahim SH, Ahmed QA, Gozzer E, Schlagenhauf P, Memish ZA. Covid-19 and community mitigation strategies in a pandemic. BMJ. 2020;368. Available from: https://www.bmj.com/content/368/bmj.m1066.

[47] Bi Q, Wu Y, Mei S, Ye C, Zou X, Zhang Z, et al. Epidemiology and Transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts.

medRxiv. 2020;Available from: https://www.medrxiv.org/content/early/2020/03/27/2020.03.03.20028423.

[48] Glass K, Barnes B. How much would closing schools reduce transmission during an influenza pandemic? Epidemiology. 2007;18(5):623–628. Available from: doi:10.1097/EDE.0b013e31812713b4.

[49] Group WHOW, Bell D, Nicoll A, Fukuda K, Horby P, Monto A, et al. Non-pharmaceutical interventions for pandemic influenza, international measures. Emerging infectious diseases. 2006 Jan;12(1):81–87. Available from: https://pubmed.ncbi.nlm.nih.gov/16494722.

[50] Cauchemez S, Ferguson NM, Wachtel C, Tegnell A, Saour G, Duncan B, et al. Closure of schools during an influenza pandemic. The Lancet Infectious diseases. 2009 Aug;9(8):473–481. Available from: https://pubmed.ncbi.nlm.nih.gov/19628172.

[51] Ferguson NM, Cummings DAT, Fraser C, Cajka JC, Cooley PC, Burke DS. Strategies for mitigating an influenza pandemic. Nature. 2006;442(7101):448–452. Available from: https://doi.org/10.1038/nature04795.

[52] Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies: Estimating the effect of Californiaâs tobacco control program. Journal of the American statistical Association. 2010;105(490):493–505.

[53] Abadie A, Gardeazabal J. The economic costs of conflict: A case study of the Basque Country. American economic review. 2003;93(1):113–132.

[54] Athey S, Imbens GW. The state of applied econometrics: Causality and policy evaluation. Journal of Economic Perspectives. 2017;31(2):3–32.

[55] Snow J. On the mode of communication of cholera. John Churchill; 1855.

[56] Wing C, Simon K, Bello-Gomez RA. Designing Difference in Difference Studies: Best Practices for Public Health Policy Research. Annual Review of Public Health. 2018;39(1):453–469. PMID: 29328877. Available from: https://doi.org/10.1146/annurev-publhealth-040617-013507.

[57] Abadie A, Cattaneo MD. Econometric methods for program evaluation. Annual Review of Economics. 2018;10:465–503.

[58] Cerulli G. A flexible Synthetic Control Method for modeling policy evaluation. Economics Letters. 2019;182:40–44.

[59] Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. Annals of internal medicine. 2020 Mar;p. M20–0504. Available from: https://pubmed.ncbi.nlm.nih.gov/32150748.

[60] Linton NM, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov AR, Jung Sm, et al. Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data. Journal of Clinical Medicine. 2020 Feb;9(2):538. Available from: http://dx.doi.org/10.3390/jcm9020538.

[61] Yang X, Yu Y, Xu J, Shu H, Xia J, Liu H, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. The Lancet Respiratory Medicine. 2020;Available from: http://www.sciencedirect.com/science/article/pii/S2213260020300795.

[62] di SanitÃ IS. Caratteristiche dei pazienti deceduti positivi a COVID-19 in Italia, Dati

al 6 aprile 2020;. Available from: https://www.epicentro.iss.it/coronavirus/bollettino/Report-COVID-2019_6_aprile.pdf,April6,2020.

[63] Hale T, Webster S, Petherick A, Phillips T, Kira B. Oxford covid-19 government response tracker. Blavatnik School of Government. 2020;25.

[64] Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science. 2020;368(6489):395–400.

Tables

Table 1: Synthetic control method: countries in donor pool and respective weights

	Synthethic Control Method, Scenario				
Countries in donor pool (ISO-3166)	I	II	III	IV	NP
	treated unit = ITA				
AUS	0	0	0	0	1.16E-13
CHN	0	0			5.8E-17
ESP	1	0.865	0.893	0.906	0.5392492
FRA	0	0.098	0.107	0.012	0.339322
GBR	0	0	0	0	7.1E-12
RUS	0	0	0		1.8E-19
SWE	0	0.037	0	0.082	0.1214288
USA	0	0	0	0	6.37E-18
RMSPE	4.18	3.8	3.91	3.86	3.48E-13
	treated unit = ARG				
AUS	0	0.003	0.122	0.334	0.0383925
BEL	0	0	0	0	0.0385485
BGR	0	0.01	0.165	0	0.0379238
BLR	0	0	0.025	0	0.0384391
BRA	0	0	0.285	0.394	0.0386677
CAN	0	0	0	0	0.0383032
CHE	0	0	0	0	0.0375996
DEU	0	0	0	0	0.0382418
ESP	0	0	0	0	0.0381977
FRA	0.068	0	0	0	0.0385051
GBR	0	0	0	0	0.0382724
IDN	0.454	0.511		0.104	0.0386165
IND	0	0		0	0.0247732
IRN	0	0			0.0386026
ISL	0	0	0		0.0382065
ITA	0	0	0	0	0.0368725
JPN	0	0.239		0	0.0360395
KOR	0	0		0	0.0370825
NIC	0	0	0.403	0	0.0385347
NLD	0	0	0	0	0.0382664
PHL	0.181	0.237		0.168	0.0385967
POL	0	0	0	0	0.0386615
RUS	0	0	0		0.0386366
SGP	0	0		0	0.0234642
SVN	0	0	0	0	0.0384797
SWE	0.297	0	0	0	0.0378611
USA	0	0	0	0	0.0362145
RMSPE	0.27	0.27	0.25	0.28	0
	treated unit = KOR				
AUS	0	0		0.417	0.1010938
BLR	0.148	0.274		0.169	0.0513091
BRA	0	0		0.17	0.0252816
CAN	0.194	0.009		0	0.0677789
CHN	0.023	0	0.496		0.0026573
ESP	0	0.114		0	0.0580915
FRA	0	0		0	0.195337
GBR	0.431	0		0.244	0.2407968
IDN	0	0	0.504	0	0.0319645
ITA	0.204	0.151		0	0.0290885
RUS	0	0.452			0.0879431
SWE	0	0		0	0.0484472
USA	0	0		0	0.0602109
RMSPE	0.89	1.21	1.77	0.79	0.001

Predictor variables included in specifications to predict the synthetic control unit
Specification I: Population size, Population density, Median age, Share of population aged 65 or older, GDP per capita, average daily COVID-19 deaths before intervention, growth rate of COVID-19 cases before intervention.
Specification II: Additionally to variables included in specification I, hospital beds per 100,000 inhabitants, public health expenditures (share of GDP).
Specification III: Same variables as in specification II. Model estimated excluding Asian countries (for Italy and Argentina) and only for Asian countries (for South Korea).
Specification IV: Additionally to variables included in specification II, Google COVID-19 Community Mobility Report Data on movement trends at certain places (residential location, workplace, retail and recreation).
Non-parametric synthetic control unit: Population size, Population density, Share of population aged 65 or older, GDP per capita, hospital beds per 100,000 inhabitants.

Table 2: Effect of school closures and early interventions on COVID-19 deaths (daily average): Difference in differences estimates - Synthetic Control Method

	baseline				15 days later			
	I	II	III	IV	I	II	III	IV
Argentina=1	0.0101	0.0132	0.0238	0.00933	-1.657***	-0.894**	-1.528***	-2.519***
	(0.113)	(0.115)	(0.0966)	(0.128)	(0.385)	(0.388)	(0.323)	(0.485)
Post=1	3.445**	1.056	3.051**	4.946**	9.736***	5.894***	8.272***	13.22***
	(1.755)	(0.853)	(1.328)	(2.097)	(3.483)	(2.006)	(2.831)	(4.147)
Argentina=1 × Post=1	-7.201***	-3.225***	-6.357***	-10.24***	-15.03***	-5.542***	-12.67***	-20.21***
	(1.292)	(0.690)	(1.086)	(1.663)	(2.706)	(1.706)	(2.120)	(3.099)
Observations	74	74	74	74	74	74	74	74
R^2	0.737	0.653	0.776	0.768	0.865	0.716	0.887	0.890
Effect size in percentage	-85.71	-72.93	-84.27	-89.50	-78.74	-62.72	-76.34	-80.43

	baseline				15 days later			
	I	II	III	IV	I	II	III	IV
Italy=1	0.278	0.0699	0.0935	0.171	-39.87***	-32.41***	-35.06***	-32.57***
	(3.461)	(3.039)	(3.139)	(3.137)	(12.96)	(11.21)	(11.72)	(11.44)
Post=1	37.49	26.27	30.65	25.75	39.62	8.908	19.93	9.303
	(36.66)	(30.85)	(32.86)	(31.02)	(118.6)	(104.9)	(108.0)	(108.4)
Italy=1 × Post=1	-132.8***	-103.0***	-113.8***	-103.3***	-207.5***	-151.2***	-172.3***	-150.2***
	(25.29)	(21.59)	(22.77)	(21.84)	(60.49)	(56.73)	(57.29)	(58.21)
Observations	74	74	74	74	74	74	74	74
R^2	0.949	0.958	0.956	0.956	0.953	0.959	0.958	0.957
Effect size in percentage	-34.60	-29.09	-31.18	-29.14	-26.56	-21.08	-23.25	-20.96

	baseline				15 days later			
	I	II	III	IV	I	II	III	IV
South Korea=1	0.0735	0.0105	-0.0560	0.0435	-23.77***	-12.79***	-9.076***	-9.918***
	(2.515)	(1.070)	(0.488)	(1.337)	(5.478)	(3.039)	(1.653)	(2.431)
Post=1	58.05*	29.84**	16.22***	28.71**	87.58*	62.94***	13.67	37.92
	(29.68)	(15.04)	(5.484)	(13.51)	(48.74)	(22.81)	(8.327)	(24.61)
South Korea=1 × Post=1	-104.8***	-57.64***	-27.04***	-45.23***	-221.6***	-124.3***	-36.00***	-98.14***
	(19.82)	(10.49)	(2.703)	(9.176)	(37.93)	(12.58)	(2.520)	(19.92)
Observations	74	74	74	74	74	74	74	74
R^2	0.706	0.728	0.860	0.694	0.858	0.892	0.827	0.846
Effect size in percentage	-96.06	-92.97	-85.94	-91.27	-88.51	-87.49	-71.89	-86.81

Notes: Table shows the estimates of η (Treated country=1), λ (Post=1), and δ (Treated country=1×Post=1) in equation (1) obtained by linear regression analysis. Polynomial time trend of the third degree included as further control variables. Columns show the results of separate regressions for Argentina, Italy, and South Korea. Baseline measures the effect from day zero of the intervention, alternative specification measures the effect 15 days later. Series truncated 18 days after the intervention. Synthetic control group is a weighted combination of control units chosen to approximate the unit affected by the intervention. Variables, countries, and weights used in the empirical models to predict the control unit in scenarios I-IV are shown in Table 1. Bootstrapped standard errors in parentheses (number of replications equivalent to number of observations). Statistical significance level * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The last row indicates the effect size in percentage of average daily deaths in the control unit.

Table 3: Effect of school closures and early interventions on COVID-19 deaths (daily average): Difference in differences estimates - Non-parametric Synthetic Control Method

	Argentina		Italy		South Korea	
	baseline	15 days later	baseline	15 days later	baseline	15 days later
Treated=1	0.0183	-2.303***	-0.435	-14.58**	0.330	-31.91***
	(0.390)	(0.475)	(2.222)	(7.387)	(4.216)	(7.703)
Post=1	7.685**	17.24**	-0.451	-64.42	88.38**	136.5*
	(3.780)	(7.587)	(19.66)	(75.05)	(44.63)	(80.73)
Treated=1 × Post=1	-12.98***	-32.16***	-32.01**	-16.95	-150.6***	-333.5***
	(2.718)	(5.787)	(15.65)	(50.36)	(30.75)	(66.85)
Observations	74	74	74	74	74	74
R^2	0.697	0.894	0.970	0.971	0.680	0.847
Effect size in percentage	-91.59	-87.24	-11.28	-2.998	-97.39	-90.04

Notes: Table shows the estimates of η (Treated country=1), λ (Post=1), and δ (Treated country=1×Post=1) in equation (1) obtained by linear regression analysis. Polynomial time trend of the third degree included as further control variables. Columns show the results of separate regressions for Italy, Argentina, and South Korea. Baseline measures the effect from day zero of the intervention, alternative specification measures the effect 15 days later. Series truncated 18 days after the intervention. Synthetic control group obtained by non-parametric imputation using a kernel function. Bootstrapped standard errors in parentheses (number of replications equivalent to number of observations). Statistical significance level * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The last row indicates the effect size in percentage of average daily deaths in the control unit.

Table 4: Effect of school closures and early interventions on COVID-19 deaths per 100,000 inhabitants (daily average): Difference in differences estimates - Synthetic Control Method

	baseline				15 days later			
	I	II	III	IV	I	II	III	IV
Argentina=1	-0.000000171	0.000000257	-0.00000317	0.0000577	-0.00591***	-0.00602***	-0.00378***	-0.00133***
	(0.000707)	(0.000793)	(0.000524)	(0.000168)	(0.00162)	(0.00120)	(0.000932)	(0.000367)
Post=1	0.0164**	0.0167**	0.0106**	0.00230	0.0262	0.00522	0.000775	0.00870**
	(0.00678)	(0.00728)	(0.00485)	(0.00147)	(0.0222)	(0.0188)	(0.0117)	(0.00347)
Argentina=1 × Post=1	-0.0267***	-0.0210***	-0.0123***	-0.00585***	-0.0576***	-0.0354**	-0.0190*	-0.0120***
	(0.00617)	(0.00470)	(0.00319)	(0.00123)	(0.0147)	(0.0173)	(0.0115)	(0.00330)
Observations	74	74	74	74	74	74	74	74
R^2	0.666	0.647	0.623	0.775	0.798	0.694	0.644	0.856
Effect size in percentage	-90.88	-88.71	-82.16	-69.06	-83.70	-75.73	-67.63	-64.29

	baseline				15 days later			
	I	II	III	IV	I	II	III	IV
Italy=1	-0.000807	-0.000896	-0.000807	-0.000807	-0.118***	-0.107***	-0.118***	-0.118***
	(0.00809)	(0.00739)	(0.00809)	(0.00809)	(0.0316)	(0.0289)	(0.0316)	(0.0316)
Post=1	0.136	0.120	0.136	0.136	0.215	0.175	0.215	0.215
	(0.0980)	(0.0888)	(0.0980)	(0.0980)	(0.259)	(0.239)	(0.259)	(0.259)
Italy=1 × Post=1	-0.405***	-0.362***	-0.405***	-0.405***	-0.670***	-0.592***	-0.670***	-0.670***
	(0.0644)	(0.0584)	(0.0644)	(0.0644)	(0.110)	(0.104)	(0.110)	(0.110)
Observations	74	74	74	74	74	74	74	74
R^2	0.925	0.932	0.925	0.925	0.937	0.942	0.937	0.937
Effect size in percentage	-49.27	-46.48	-49.27	-49.27	-40.10	-37.45	-40.10	-40.10

	baseline				15 days later			
	I	II	III	IV	I	II	III	IV
South Korea=1	0.0000376	-0.000540	0.00171***	-0.00134	-0.0560***	-0.00974***	0.00321***	-0.0109***
	(0.00508)	(0.000955)	(0.000403)	(0.000981)	(0.0135)	(0.00228)	(0.000607)	(0.00225)
Post=1	0.138**	0.0228**	0.00231	0.0247**	0.317***	0.0360**	-0.00177	0.0312*
	(0.0685)	(0.0109)	(0.00157)	(0.00968)	(0.108)	(0.0173)	(0.00330)	(0.0178)
South Korea=1 × Post=1	-0.266***	-0.0390***	0.00310***	-0.0377***	-0.597***	-0.0801***	0.000751	-0.0724***
	(0.0502)	(0.00703)	(0.00114)	(0.00655)	(0.0645)	(0.00958)	(0.00299)	(0.0101)
Observations	74	74	74	74	74	74	74	74
R^2	0.711	0.747	0.665	0.766	0.898	0.872	0.602	0.863
Effect size in percentage	-96.91	-81.14	83.88	-79.25	-90.08	-80.42	12.96	-77.81

Notes: Table shows the estimates of η (Treated country=1), λ (Post=1), and δ (Treated country=1×Post=1) in equation (1) obtained by linear regression analysis. Polynomial time trend of the third degree included as further control variables. Columns show the results of separate regressions for Italy, Argentina, and South Korea. Baseline measures the effect from day zero of the intervention, alternative specification measures the effect 15 days later. Series truncated 18 days after the intervention. Synthetic control group is a weighted combination of control units chosen to approximate the unit affected by the intervention. Bootstrapped standard errors in parentheses (number of replications equivalent to number of observations). Statistical significance level * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The last row indicates the effect size in percentage of average daily deaths in the control unit.

Table 5: Effect of school closures and early interventions on COVID-19 deaths per 100,000 inhabitants (daily average): Difference in differences estimates - Non-parametric Synthetic Control Method

	Argentina		Italy		South Korea	
	baseline	15 days later	baseline	15 days later	baseline	15 days later
Treated=1	-0.000728	-0.0103***	-0.00489	-0.0744***	0.000928	-0.0461***
	(0.00145)	(0.00212)	(0.00541)	(0.0199)	(0.00624)	(0.0114)
Post=1	0.0281**	0.0584***	0.0546	-0.0649	0.131**	0.201*
	(0.0131)	(0.0226)	(0.0523)	(0.162)	(0.0660)	(0.119)
Treated=1 × Post=1	-0.0465***	-0.106***	-0.192***	-0.213***	-0.221***	-0.492***
	(0.00873)	(0.0183)	(0.0339)	(0.0767)	(0.0454)	(0.0982)
Observations	74	74	74	74	74	74
R^2	0.703	0.879	0.961	0.958	0.680	0.848
Effect size in percentage	-93.18	-87.17	-31.29	-18.22	-96.68	-89.81

Notes: Table shows the estimates of η (Treated country=1), λ (Post=1), and δ (Treated country=1×Post=1) in equation (1) obtained by linear regression analysis. Polynomial time trend of the third degree included as further control variables. Columns show the results of separate regressions for Italy, Argentina, and South Korea. Baseline measures the effect from day zero of the intervention, alternative specification measures the effect 15 days later. Series truncated 18 days after the intervention. Synthetic control group obtained by non-parametric imputation using a kernel function. Bootstrapped standard errors in parentheses (number of replications equivalent to number of observations). Statistical significance level * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The last row indicates the effect size in percentage of average daily deaths in the control unit.

Table 6: Placebo tests: Difference in differences estimates - Synthetic Control Method

	Placebo tests			
	I	II	III	IV
Argentina=1	-0.151	-0.170	-0.0950	-0.166
	(0.152)	(0.168)	(0.106)	(0.166)
Post=1	0.191	0.0861	0.179	0.207
	(0.283)	(0.275)	(0.240)	(0.304)
Argentina=1 × Post=1	0.194	0.220	0.143	0.210
	(0.189)	(0.193)	(0.149)	(0.200)
Observations	36	36	36	36
R^2	0.172	0.118	0.150	0.195

	Placebo tests			
	I	II	III	IV
Italy=1	-3.667	-3.785	-3.809	-3.655
	(5.693)	(5.129)	(5.257)	(5.280)
Post=1	-12.45	-11.24	-11.59	-11.44
	(8.130)	(7.362)	(7.503)	(7.627)
Italy=1 × Post=1	4.733	4.626	4.683	4.591
	(5.878)	(5.330)	(5.464)	(5.458)
Observations	36	36	36	36
R^2	0.744	0.772	0.768	0.760

	Placebo tests			
	I	II	III	IV
South Korea=1	-0.494	-1.299	-3.000	-0.404
	(1.059)	(1.462)	(2.273)	(0.985)
Post=1	0.550	-0.102	-2.194	-0.0566
	(1.610)	(1.747)	(3.070)	(1.064)
South Korea=1 × Post=1	0.681	1.571	3.533	0.537
	(1.129)	(1.492)	(2.321)	(1.025)
Observations	36	36	36	36
R^2	0.639	0.675	0.680	0.630

Notes: Table shows the estimates of η (Treated country=1), λ (Post=1), and δ (Treated country=1×Post=1) in equation (1) obtained by linear regression analysis. Polynomial time trend of the third degree included as further control variables. Columns show the results of separate regressions for Argentina, Italy, and South Korea. Series truncated one day before the intervention. Synthetic control group is a weighted combination of control units chosen to approximate the unit affected by the intervention. Variables, countries, and weights used in the empirical models to predict the control unit in scenarios I-IV are shown in Table 1. Bootstrapped standard errors in parentheses (number of replications equivalent to number of observations). Statistical significance level * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Effect of school closures and early interventions on excess deaths in Italy

	I	II	III	IV	NP
Italy=1	0.248	-71.72	-71.72	-71.60	20.15
	(2357.8)	(2203.3)	(2203.3)	(2682.3)	(1473.3)
Post=1	294.1	215.0	215.0	390.1	-67.74
	(1632.9)	(1515.3)	(1515.3)	(1894.0)	(924.1)
Italy=1 × Post=1	-2270.5	-2124.3	-2124.3	-2534.5	-1229.8
	(2753.6)	(2595.8)	(2595.8)	(3101.9)	(1555.6)
Observations	12	12	12	12	12
R^2	0.989	0.991	0.991	0.987	0.996
Effect size in percentage	-40.67	-38.57	-38.57	-42.82	-27.20

Notes: Table shows the estimates of η (Treated country=1), λ (Post=1), and δ (Treated country=1×Post=1) in equation (1) obtained by linear regression analysis. Polynomial time trend of the third degree included as further control variables. Series truncated 3 weeks after the intervention. Predictor variables used in scenarios I-IV and NP shown in Table 1. Bootstrapped standard errors in parentheses (number of replications equivalent to number of observations). Statistical significance level * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The last row indicates the effect size in percentage of excess deaths in the control unit. Data source is the data on excess mortality collected by the Financial Times.

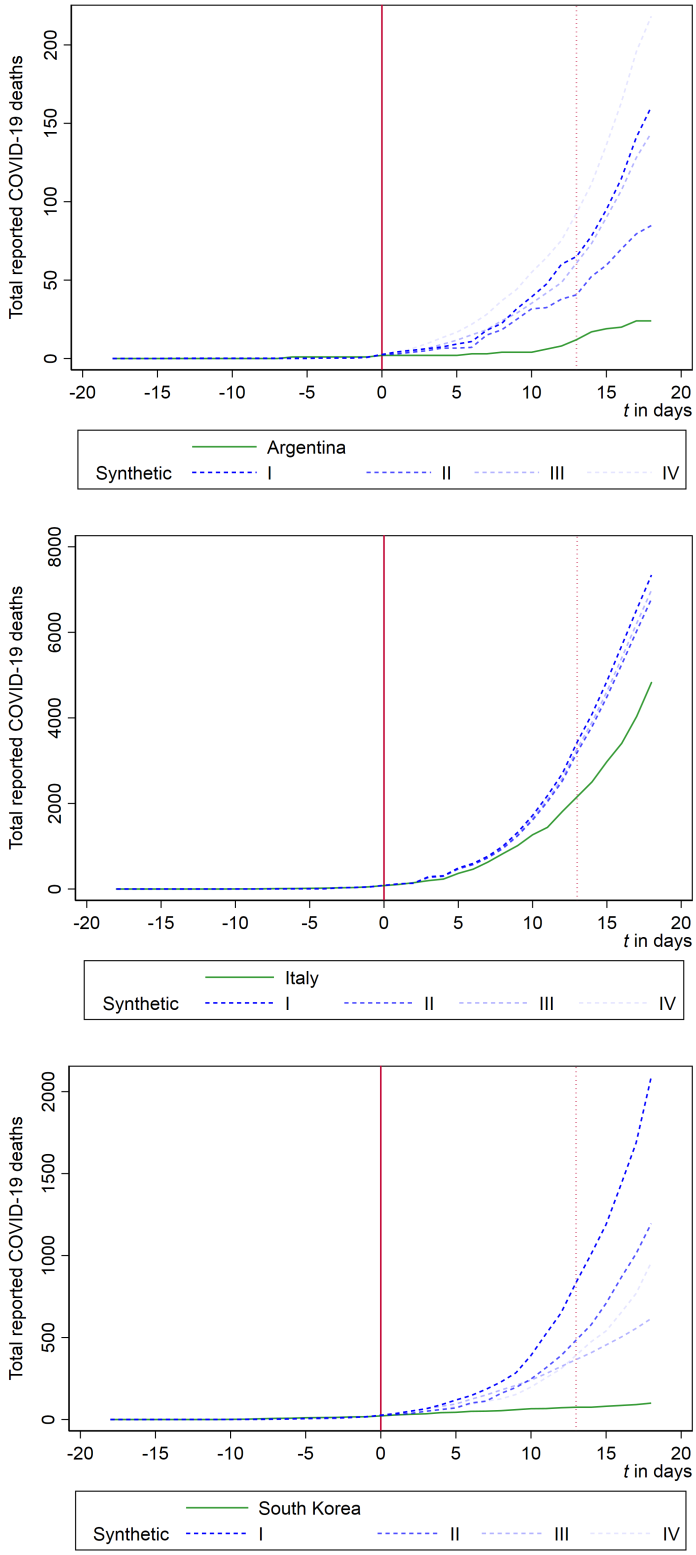
Table 8: Effect in other countries ordered by the number of deaths on the day of school closure

	Switzerland 13	Germany 13	Netherlands 20	Indonesia 55	Canada 61	Brazil 77	France 127	UK 158	Spain 288
Treated=1	-9.645***	-2.499	-11.97***	-14.31***	-39.28***	-0.211	13.66**	7.824	44.39***
	(2.280)	(2.060)	(3.910)	(3.489)	(9.213)	(2.612)	(6.284)	(11.50)	(13.54)
Post=1	69.48***	16.85	61.41	57.83**	182.6*	8.696	-6.925	-93.23	-149.8
	(23.83)	(22.71)	(57.81)	(26.44)	(109.4)	(35.78)	(58.40)	(93.06)	(91.62)
Treated=1 × Post=1	-140.1***	-9.245	-177.6***	-73.29***	-440.7***	-7.193	-7.107	247.1***	9.562
	(19.93)	(20.16)	(48.27)	(15.80)	(67.01)	(48.10)	(39.59)	(73.41)	(71.70)
Effect size in percentage	-74.08	-6.068	-53.89	-65.91	-74.95	-6.365	-1.648	59.89	1.182

Notes: Table shows the estimates of η (Treated country=1), λ (Post=1), and δ (Treated country=1×Post=1) in equation (1) obtained by linear regression analysis. Polynomial time trend of the third degree included as further control variables. Series truncated 18 days after the intervention. Predictor variables are the same as used in scenarios II shown in Table 1. Bootstrapped standard errors in parentheses (number of replications equivalent to number of observations). Statistical significance level * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The last row indicates the effect size in percentage of reported average daily COVID-19 deaths in the control unit. Number under the name of the country indicates the number of reported COVID-19 deaths on the day of nationwide school closure.

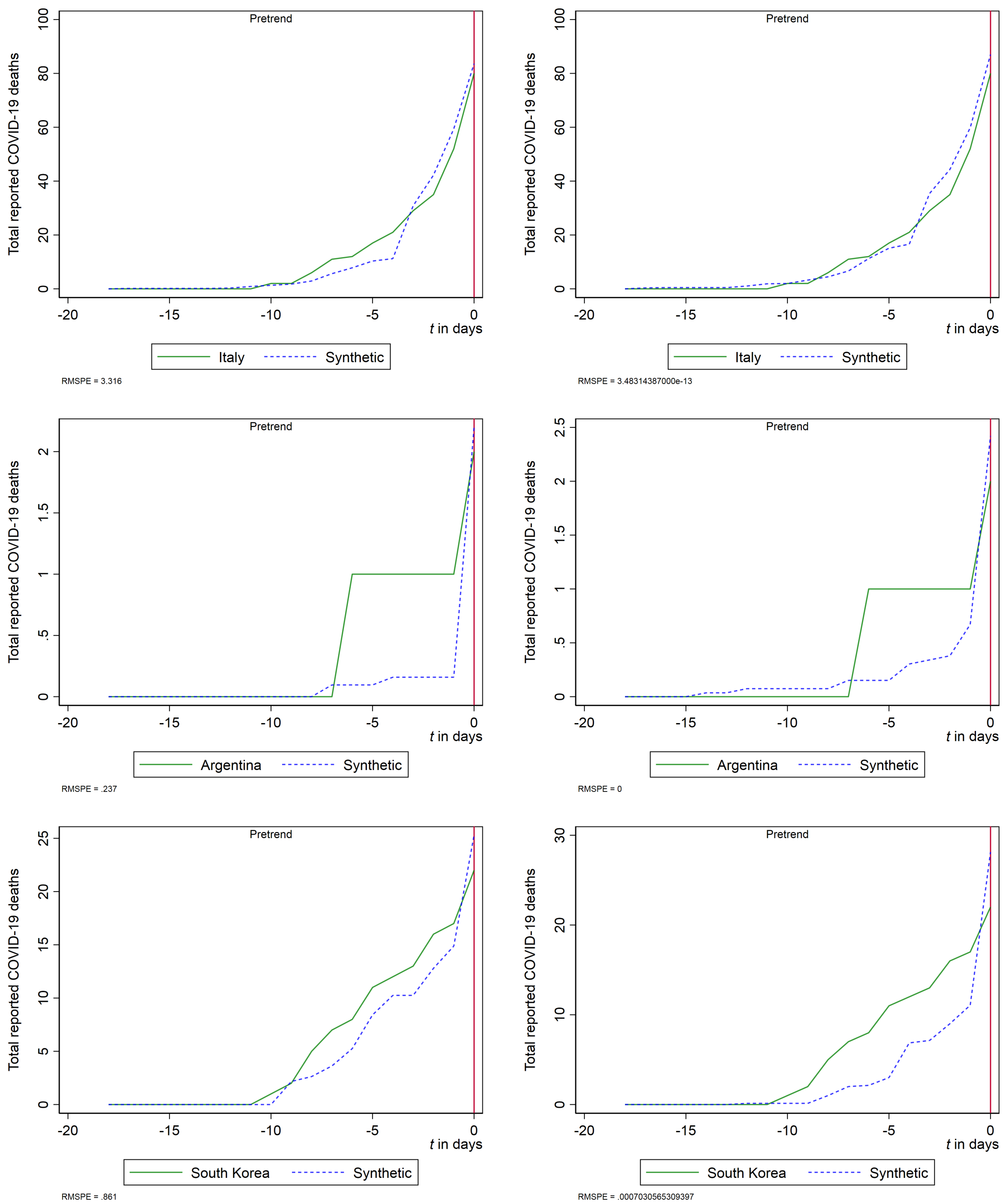
Figures

Figure 1: Epidemic curves of COVID-19 deaths in treated unit and synthetic control unit



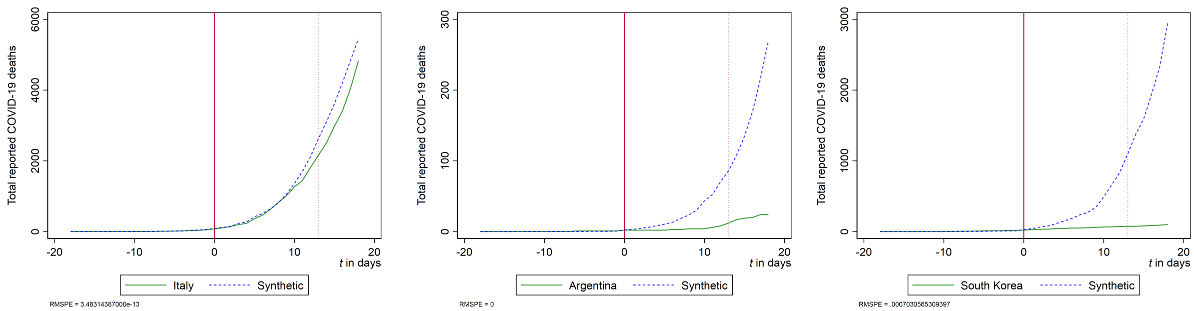
Notes: Variables, countries, and weights used to predict synthetic control units using synthetic control methods indicated in Table 1. Source: Own elaboration using ECDC data on reported COVID-19 deaths.

Figure 2: Comparison of Pre-Trends (right figure shows non-parametric approach to obtain counterfactual)



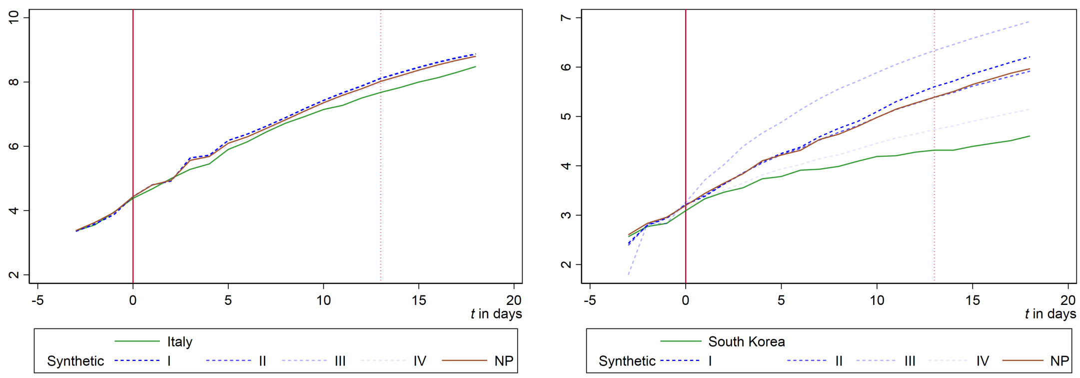
Source: Own elaboration using ECDC data on reported COVID-19 deaths.

Figure 3: Epidemic curves of COVID-19 deaths in Italy and in synthetic control group (non parametric estimation of counterfactual)



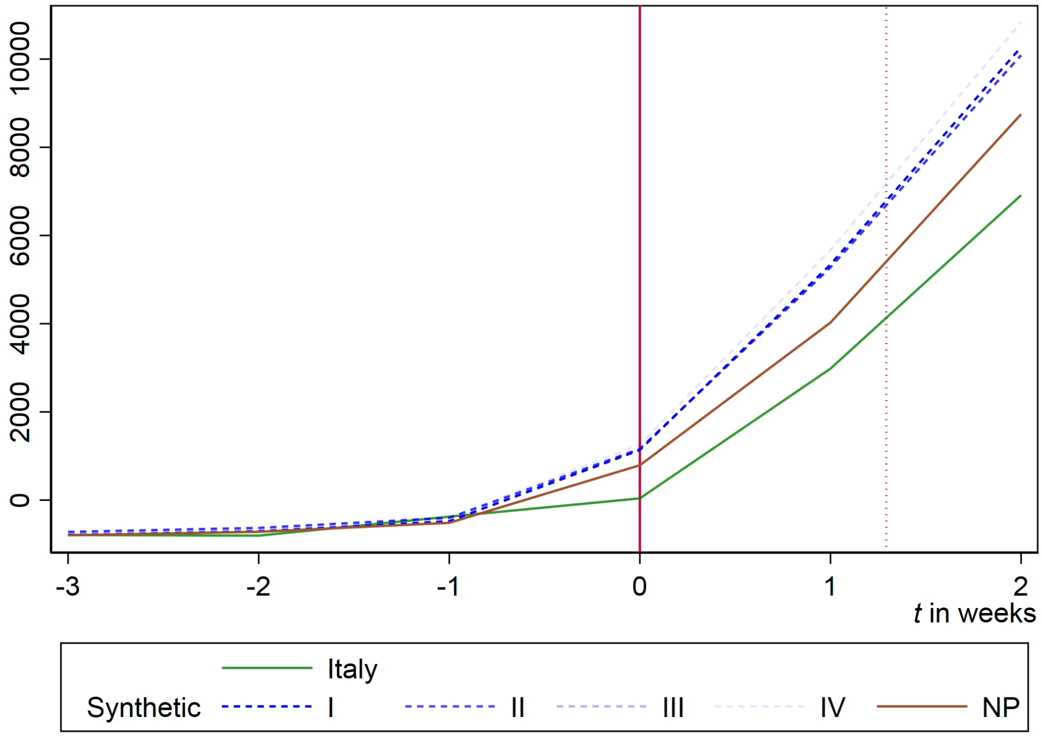
Notes: Variables, countries, and weights used to predict synthetic control units using synthetic control methods indicated in Table 1. Source: Own elaboration using ECDC data on reported COVID-19 deaths.

Figure 4: Epidemic curves of COVID-19 deaths in treated unit and synthetic control unit - logarithmic scale



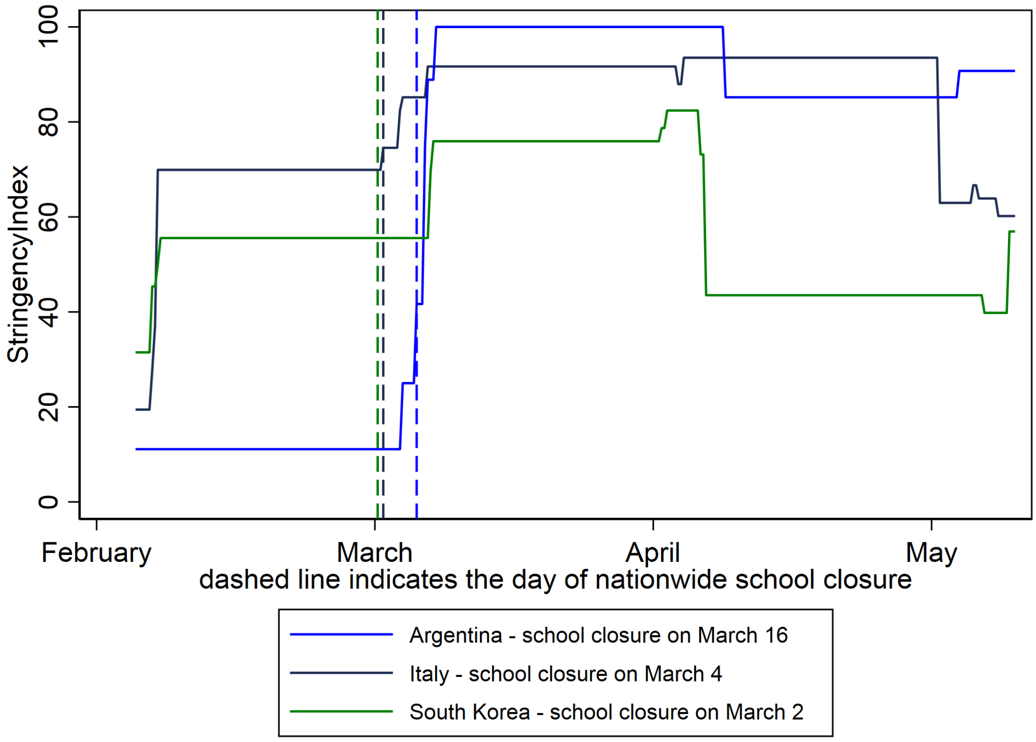
Source: Own elaboration using ECDC data on reported COVID-19 deaths.

Figure 5: Excess deaths before and after school closures - Italy vs Synthetic Control Unit



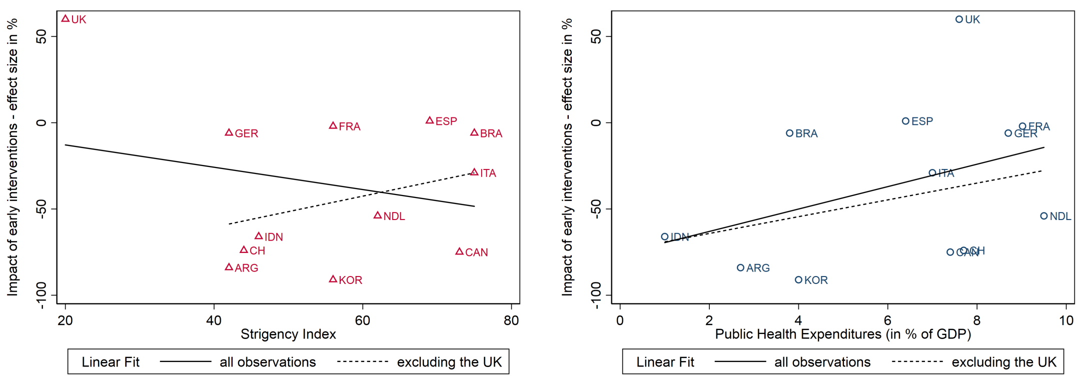
Source: Own elaboration using data on excess mortality collected by the Financial Times.

Figure 6: Stringency of government responses to the pandemic in Argentina, Italy, and South Korea



Source: Own elaboration using data from the Oxford COVID-19 Government Response Tracker.

Figure 7: Correlation between the effect of early interventions, stringency of mitigation measures, and public health expenditures



Source: Own elaboration using the estimated effect sizes, data from the Oxford COVID-19 Government Response Tracker, and from the Our World in Data project.