UNIVERSITÀ DI SIENA 1240

QUADERNI DEL DIPARTIMENTO DI ECONOMIA POLITICA E STATISTICA

Federico Crudu
Roberta Di Stefano
Giovanni Mellace
Silvia Tiezzi

The Gray Zone

n. 874 – Marzo 2022

The Gray Zone *

Federico Crudu†
University of Siena and CRENoS

Roberta Di Stefano ‡
Sapienza University of Rome

Giovanni Mellace§
University of Southern Denmark

Silvia Tiezzi¶
University of Siena

March 2022

Abstract

On March 23, 2020, in response to the COVID-19 pandemic, Italy declared a nationwide lockdown. A month earlier, on February 23, the Italian government ordered its military police to seal the borders and declared a Red Zone around 10 municipalities of the province of Lodi and in Vo' Euganeo, a small town in Padua province. On the same day, Confindustria Bergamo, the province's industrial association, posted a video on social media against having a lockdown in the area of Bergamo and was supported by key business leaders and local administrators. Despite having a similar infection rate to the Red Zone municipalities, the government decided not to extend the Red Zone to the municipalities of Bergamo province with high infection rates. Bergamo later became one of the deadliest outbreaks of the first wave of the virus in the Western world. What would have happened had the Red Zone been extended to that area? We use the Synthetic Control Method to estimate the causal effect of (not) declaring a Red Zone in the Bergamo area on daily excess mortality. We find that about two-thirds of the reported deaths could have been avoided had the Italian government declared the area a Red Zone.

Key words: COVID-19, causal impact, synthetic control method, Red Zone, Bergamo, non-pharmaceutical interventions.
JEL classification: C23, I18, O57.

---

*Comments from seminar participants at Birkbeck, University of London, at the Department of Economics and Statistics, University of Siena, and from participants at the 2021 A.I.S.Re conference are gratefully acknowledged.

†Department of Economics and Statistics, University of Siena, Piazza San Francesco 7/8, 53100, Siena, federico.crudu@unisi.it.

‡Department of Statistical Sciences, Sapienza University of Rome, Viale Regina Elena 295, 00161, Rome, roberta.distefano@uniroma1.it.

§Department of Economics, University of Southern Denmark, Campusvej 55, DK-5230, Odense, giome@sam.sdu.dk.

¶Department of Economics and Statistics, University of Siena, Piazza San Francesco 7/8, 53100, Siena, silvia.tiezzi@unisi.it.

1 Introduction

In March and April 2020 the northern Italian province of Bergamo was the site of the deadliest outbreak of the COVID-19 pandemic in the Western world. According to official data from the Italian National Institute of Statistics (ISTAT), the small town of Nembro, in the industrial and densely populated Serio Valley in Bergamo province, registered a 1000% increase in deaths in the first three weeks of March 2020, compared to the same period of the previous year (11 deaths in the first three weeks of 2019 against 121 in the first three weeks of 2020).[1] Alzano Lombardo, another town with a comparable population in the same area, registered a 937.5% increase in deaths in the same period (from 8 in the first three weeks of 2019 to 83 in the first three weeks of 2020). The number of deaths in the Lombardy region in March and April 2020 compared to the previous five years average for the months of March and April was, 191.2% and 117.1% higher, respectively (ISTAT, 2021).

Surprisingly, the Italian government decided against declaring a Red Zone in the province of Bergamo, despite having opted for it sixty miles to the south, where another serious outbreak of COVID-19 had occurred only a few days before. The opposite policy decisions adopted in response to similar events in two areas of Lombardy in the same time period afford an ideal setting for a quasi-experimental study.

The main goal of this paper is to assess whether a causal relationship exists between the (failure to) declare a Red Zone in the area of Bergamo in early March 2020, and the daily excess mortality before containment measures had been adopted at national level.

To estimate the causal impact of not implementing a Red Zone, we use the Synthetic Control Method (SCM, Abadie & Gardeazabal, 2003; Abadie et al., 2010). Specifically, we consider - one at a time - the three main municipalities of the Serio Valley, i.e., Nembro, Alzano Lombardo, and Albino, and we use SCM to construct their alternative versions in the event that a Red Zone had been promptly implemented. For this purpose we use as control units the 11 municipalities that were declared a Red Zone. The comparison between

[1] https://www.istat.it/it/archivio/240401

the outcome of the treated municipalities and their estimated counterfactuals provides an estimate of the effect of the policy. Since Red Zone implementation can be considered a treatment that depends on being affected by COVID-19, we also estimate the impact of COVID-19 on our three municipalities. To construct the counterfactual version in the absence of pandemic, we use similar municipalities where COVID-19 did not spread until later.

Our results suggest that declaring a Red Zone around the Serio Valley would have reduced the number of reported all-cause deaths by about a half between March and April 2020.

Our analysis is connected to recent literature on the effectiveness of non pharmaceutical interventions (NPIs). A Red Zone can be thought of as an articulated set of measures aimed at containing the spread of a disease involving limitations to movement, closure of public spaces and buildings, as well as information campaigns. Such measures do not necessarily imply the use of medical treatments.

In the absence of mass screening and aggressive contact tracing, the timely set up of a Red Zone has been recognized as one of the most effective NPIs for containing the spread of COVID-19, preventing hospitals from being overwhelmed, and potentially limiting the number of deaths (Acemoglu et al., 2021; Chernozhukov et al., 2021; Fagiuoli et al., 2020; Signorelli et al., 2020). Previous studies using counterfactual analysis have shown that strict initial lockdown measures played an important role in limiting the spread of the COVID-19 infection. Fang et al. (2020) showed that the lockdown in Wuhan, enacted on January 23, 2020 decreased the number of positive cases by 64.8% in the 347 Chinese cities outside Hubei province, and by 52.64% in 16 cities other than Wuhan inside Hubei. Cho (2020) showed that the number of deaths in Sweden would have been reduced by between 23% and 30% had Sweden implemented these policies. In addition, the impact of NPIs becomes visible with a time lag of around 5 weeks as individuals adjust their behavior in response to the policies. Using a dynamic causal model for the effect of NPIs in the United States, Chernozhukov et al. (2021) found that both policies and information on transmission risks are important determinants of COVID-19 deaths and that policy changes explain a large fraction of observed voluntary changes in social distancing. Flaxman et al. (2020) find that major NPIs - and lockdowns in particular - have had a big impact on reducing transmission.

While the general effectiveness of NPI measures seems indisputable, Singh et al. (2021) suggest that their actual impact may depend on the characteristics of the groups that receive the treatment and ultimately on their compliance. Acemoglu et al. (2021) study the effect of selective lockdowns, targeting different age groups, and find that this approach may significantly outperform uniform policies, particularly when the policy is stricter for the oldest (and at highest risk) age group.

This strand of literature suggests that: a) timely adoption is crucial for NPIs' effectiveness; and b) policy changes may trigger voluntary changes in individual behavior. In the unfolding of events that led to the Bergamo tragedy both elements were missing and the absence of clear policy signals in the crucial days between February 23 and March 23 effectively turned the province of Bergamo into a Gray Zone.

This article proceeds as follows. The next section provides a narrative account of the events relevant to our empirical analysis. Section 3 discusses the identification strategy adopted to assess the causal impact of the failure to declare a Red Zone. Section 4 describes the data in detail. Sections 5 and 6 present our results and a battery of robustness checks. Section 7 presents our conclusion.

2 The Gray Zone

In this section, we describe the spread of the first wave of the COVID-19 pandemic in the area of Bergamo and the evolution of containment measures taken by regional and national policy- makers. We then discuss the unique features of the case that make it suited for our empirical analysis. In Figure 1, we present a timeline of public policy measures adopted to contain the first wave of the pandemic. Partial and full lockdowns refer to the entire national territory. Additional measures at the regional and local level were adopted in the same period in different areas of Italy, as detailed in Bosa et al. (2022).

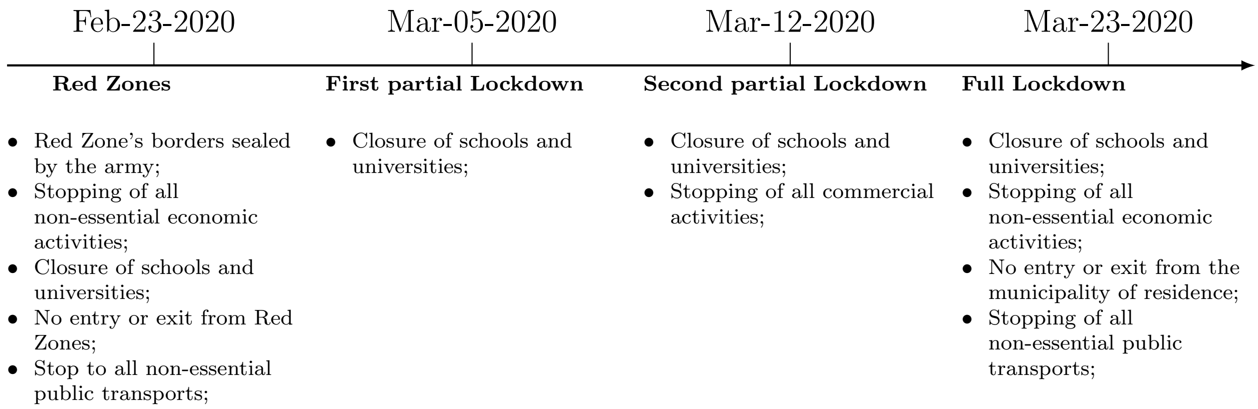

Figure 1: Timeline of nationwide containment measures.

In January 2020, after the COVID-19 outbreak in the city of Wuhan, family doctors in the Lombardy region had reported anomalous pneumonia cases and were prescribing more scans than usual, but these scans did not include testing for SARS-COV2, because Italy had adopted the new WHO protocols which had limited testing for COVID-19 to people with a link to China.[2]

On February 20, a doctor in the town of Codogno, in Lodi province (Lombardy), broke with WHO protocol and tested a man with serious pneumonia who was not responding to standard treatments. The man's test results came back positive and he became Italy's first known locally transmitted case of COVID-19. On February 23, the government ordered Italy's military police to seal the borders and declared a Red Zone around 10 municipalities in the province of Lodi, including Codogno, which would last until the nationwide lockdown of March 23. An additional Red Zone was declared around Vo' Euganeo, a small town in Padua province in the Venetian region (Presidente del Consiglio dei Ministri, 2020c). A unique characteristic of Italy's Red Zones is the sealing of borders by military police, strictly prohibiting the entering or exiting of the zone throughout the containment measure. Within the borders of the Red Zone residents would be quarantined, all commercial activities,

[2] As reported by the investigative television program Report on April 6, 2020 (link).

schools and universities would be closed, and all non-essential economic activities and public transports would be stopped.

On February 23, another doctor broke with WHO protocol and tested a patient at the Pesenti Fenaroli hospital of Alzano Lombardo, a town in the Serio Valley, in Bergamo province, 60 miles away from Lodi. These test results were positive, as well. The director of the Pesenti Fenaroli hospital immediately closed the emergency room, but he was urged to reopen it a few hours later by regional health officials despite the apparent lack of swabs for testing incoming patients. It is unclear whether the hospital had been sanitized before re-opening, or whether patients and health-care staff at the hospital had been tested or were isolated before re-opening.[3]

On February 28, Bergamo's province reported 103 positive COVID-19 cases, with the Pesenti Fenaroli hospital as the most likely source of the outbreak. However, the central government's committee of scientific advisors from the Higher Health Institute (HHI) did not advise in favor of a Red Zone at that point, nor did Lombardy health officials. On the same day Confindustria Bergamo, the province industrial association, posted a video in English titled "Bergamo is running".[4] The New York Times (Horowitz, 2020) reported that business leaders, and even the Alzano Lombardo mayor, resisted a lockdown, and contacted their commercial associations that had clout in Rome.

On March 3, the government's scientific committee proposed a Red Zone in the Serio Valley around Nembro and Alzano Lombardo. On March 4, the military police began preparing to seal the borders and turn the area into a Red Zone. On March 5, the scientific committee again urged the government to lockdown the towns in the Serio Valley. But nothing happened. On March 8, the whole Lombardy region including Milan was locked down (Presidente del Consiglio dei Ministri, 2020a). This was only a partial lockdown, however, as commercial activities, including shops, bars, and restaurants, and all productive activities, continued almost as usual.[5] On March 23, the government declared

[3]From Matteo Galizzi (LSE) and Simone Ghislandi (Bocconi University) Cambridge Core blog post: link

[4]Link to YouTube video.

[5]Large shopping malls remained closed during week-ends; bars and restaurants remained open between

a nationwide lockdown which would last for 69 days until May 18 (Presidente del Consiglio dei Ministri, 2020b). No Red Zone was ever declared in the Bergamo area. As a result, most business activities and manufacturing companies kept working as usual until March 23.

The decision not to declare a Red Zone in the Serio Valley at the end of February, 2020, is considered responsible for the spread of infection to other towns in the province of Bergamo, and then throughout Europe (Alfieri et al., 2022). Judicial investigations are still under way about the legal responsibilities for the country's response to its first coronavirus outbreak. However, as far as we know, no study has assessed whether a causal relationship exists between that political decision and the level of mortality observed in the studied area during the period under consideration.

3 Identification strategy

Our goal is to assess whether a causal relationship exists between the failure to declare a Red Zone in the area of Bergamo in early March 2020 and the rise in mortality observed in the same period. This section discusses our identification strategy in detail. Given that we have a small number of treated and control units, but a relatively long daily panel data set, the SCM introduced by Abadie & Gardeazabal (2003) appears to be the most appropriate choice. However, as the pandemic also had an impact on units in our donor pool, different from the standard setting, we need to impose an extra restriction. Specifically, we need to assume that pre-pandemic characteristics are able to approximate the effect that the pandemic would have had on the Serio Valley municipalities, had they implemented a Red Zone.

In the following, we explain this mechanism, stressing the fact that similar restrictions are implicitly imposed in other prominent studies on the effectiveness of COVID-19-related policy interventions (e.g., Cho, 2020). To better judge the plausibility of this restriction,

6am and 6pm.

we make use of a different donor pool that was not affected by the pandemic, at least initially.

We are interested in how the pandemic would have affected the municipalities of the Serio Valley had they imposed a Red Zone. To this end, we define three potential outcomes. We denote by Y_{jt}^{NP} the cumulative excess mortality that municipality j would have experienced at time t if there had been no pandemic. Our second potential outcome Y_{jt}^{RZ}, is equal to Y_{jt}^{NP} plus *the effect of the pandemic in the presence of the Red Zone*, denoted as β_{jt}. Finally, the third potential outcome Y_{jt}^{NRZ} also includes *the extra effect of not having implemented a Red Zone*, γ_{jt}, which it is our effect of interest. Following Abadie et al. (2010) and without loss of generality, to illustrate our identification strategy we assume that

$$Y_{jt}^{NP} = f_{jt} \text{ (no pandemic)}, \tag{1}$$

$$Y_{jt}^{RZ} = \beta_{jt} + f_{jt} = \beta_{jt} + Y_{jt}^{NP} \text{ (Red Zone)}, \tag{2}$$

$$Y_{jt}^{NRZ} = \beta_{jt} + \gamma_{jt} + f_{jt} = \gamma_{jt} + Y_{jt}^{RZ} \text{ (no Red Zone)}. \tag{3}$$

The common component f_{jt} follows a simple factor model

$$f_{jt} = \eta_t + \boldsymbol{\lambda}_t \boldsymbol{\mu}_j + \boldsymbol{\xi}_j \boldsymbol{\delta}_t + \varepsilon_{jt} \tag{4}$$

where η_t is an unknown common factor with constant factor loadings across units, $\boldsymbol{\xi}_j$ is a $(p \times 1)$ vector of observed covariates not affected by the intervention, $\boldsymbol{\delta}_t$ is a $(1 \times p)$ vector of unknown parameters, $\boldsymbol{\lambda}_t$ is a $(1 \times m)$ vector of unobserved common factors, $\boldsymbol{\mu}_j$ is an $(m \times 1)$ vector of unknown factor loadings, and the error terms ε_{jt} are unobserved transitory shocks at the municipality level with zero mean.

Let us assume that unit 1 is one of the municipalities of the Serio Valley that was affected by the pandemic but which did not impose a Red Zone (our treated unit) and that units from 2 to J are those municipalities that imposed a Red Zone. Furthermore, let us denote $\hat{w}_j$, $j = 2, \ldots, J$ as the weights estimated by a SCM for our treated unit using the Red Zone municipalities as the donor pool.

We mentioned in the beginning of this section that, differently from the standard SCM setting, to identify γ_{1t} it is not enough to assume that $\sum_{j=2}^{J} \hat{w}_j f_{jt} \approx f_{1t}$ but, in addition, we need to assume that

$$\sum_{j=2}^{J} \hat{w}_j \beta_{jt} \approx \beta_{1t}. \tag{5}$$

We refer to this assumption as non-extreme pandemic effect (hereinafter "NEPE"), as it implies that β_{1t} must lie in the convex hull of the effects of the pandemic observed in the donor pool and therefore it cannot be extreme. NEPE allows us to identify our effect of interest by

$$\gamma_{1t} \approx \hat{\gamma}_{1t} = Y_{1t} - \sum_{j=2}^{J} \hat{w}_j Y_{jt}. \tag{6}$$

Clearly, studies that estimate the impact of policy measures implemented to mitigate the pandemic implicitly impose a similar assumption, as they typically compare a treated group that implemented a certain policy and a control group that did not (see, for example Cho, 2020). Intuitively, both groups are affected by the pandemic in potentially different ways, regardless of the implemented policy.

Consolandi (2021) shows that the social and geographical characteristics of the territory were among the determinants of the virus outbreak in the Serio Valley. Specifically, the following elements were identified as favoring the contagion: territorial morphology, the density of the industrial zone and its network of commercial exchanges at national and international level, the intense daily commuting to schools and workplaces, the polycentric type of settlement that characterizes the Po Valley's urban areas. Focusing on early transmission of COVID-19 in New York City, Almagro & Orane-Hutchinson (2021) show that workers in jobs with a high degree of human exposure constituted one of the main determinants of the spread of the virus in New York City. Glaeser et al. (2021) find that the level of mobility within urban areas was an important factor in explaining the spread of COVID-19 in five major U.S. cities. As we explain in the next section, we are able to control for many of the above elements in our analysis. In light of these studies, our NEPE assumption of equation 5 appears to be reasonable. We can, nonetheless, further investigate its plausibility.

First, by using units that were not affected by the pandemic at the beginning of the period, we can easily estimate the overall effect of the pandemic in the Serio Valley municipalities. This is done by running a standard SCM using municipalities in Lombardy that are far away from both the Serio Valley and the province of Lodi (see Table 2). Arguably, those municipalities were either not affected or marginally affected by the pandemic, especially at the beginning of the estimation period. Let units from $J+1$ to K be the units not affected by the pandemic and let $\hat{w}_i, i = J+1, \ldots, K$ be the SCM weights. Under the standard SCM assumptions we can recover the total effect of the pandemic on our treated unit as

$$\gamma_{1t} + \beta_{1t} \approx \widehat{\gamma_{1t} + \beta}_{1t} = Y_{1t} - \sum_{i=J+1}^{K} \hat{w}_i Y_{it}. \tag{7}$$

Second, we can similarly estimate the effect of the pandemic on the municipalities that implemented a Red Zone (and received a positive weight):

$$\beta_{jt} \approx \hat{\beta}_{jt} = Y_{jt} - \sum_{i=J+1}^{K} \tilde{w}_i^j Y_{it}, \quad j = 2, \ldots, J, \tag{8}$$

where $\tilde{w}_i^j$ are the weights obtained by running a standard SCM for Red Zone municipality j using the unaffected municipalities as a donor pool. Recall from equation 5 that to estimate the additional effect of not implementing a Red Zone we use a weighted average of the effect of the pandemic in the municipalities that implemented a Red Zone, $\sum_{j=2}^{J} \hat{w}_j \beta_{jt}$, to approximate the effect that the pandemic would have had on our treated unit β_{1t}. Given that we are able to estimate the effect of the pandemic on every single municipality that implemented a Red Zone we can replace the weighted average with the largest estimated effect, i.e., we can estimate γ_{1t} as

$$
\begin{aligned}
\tilde{\gamma}_{1t} &= \widehat{\gamma_{1t} + \beta}_{1t} - \max_{j \in [2,\ldots,J], \hat{w}_j > 0} \hat{\beta}_{jt} \\
&= \underbrace{\left(Y_{1t} - \sum_{i=J+1}^{K} \hat{w}_i Y_{it} \right)}_{\widehat{\gamma_{1t}+\beta_{1t}}} - \max_{j \in [2,\ldots,J], \hat{w}_j > 0} \underbrace{\left(Y_{jt} - \sum_{i=J+1}^{K} \tilde{w}_i^j Y_{it} \right)}_{\hat{\beta}_{jt}}.
\end{aligned} \tag{9}
$$

If the condition in equation 5 is not met, it will induce a bias in $\hat{\gamma}_{1t}$ equal to the difference

$$\sum_{j=2}^{J} \hat{w}_j \beta_{jt} - \beta_{1t}.$$

Since, by definition,

$$\max_{j \in [2,\ldots,J], \hat{w}_j > 0} \hat{\beta}_{jt} > \sum_{j=2}^{J} \hat{w}_j \beta_{jt},$$

if we replace $\sum_{j=2}^{J} \hat{w}_j \beta_{jt}$ with $\max_{j \in [2,\ldots,J], \hat{w}_j > 0} \hat{\beta}_{jt}$ our new estimate of γ_{1t} will be smaller than $\hat{\gamma}_{1t}$ (see equation 6), provided that $\gamma_{1t} + \beta_{1t}$, a quantity we can estimate, is positive. Under the reasonable assumption that the pandemic cannot (at least in the short time period we consider) reduce mortality, we can provide two interpretations for the quantity in equation 9, depending on whether the true value of β_{1t} is smaller or larger than $\max_{j \in [2,\ldots,J], \hat{w}_j > 0} \hat{\beta}_{jt}$.

If $\beta_{1t} < \max_{j \in [2,\ldots,J], \hat{w}_j > 0} \hat{\beta}_{jt}$, which it is arguably the most plausible scenario, then equation 9 can readily be interpreted as a lower boundary for γ_{1t}.

If, on the other hand, $\beta_{1t} > \max_{j \in [2,\ldots,J], \hat{w}_j > 0} \hat{\beta}_{jt}$ and we are willing to assume that $\gamma_{1t} \geq 0$, equation 9 is the difference between β_{1t} and $\max_{j \in [2,\ldots,J], \hat{w}_j > 0} \hat{\beta}_{jt}$ plus γ_{1t}. Therefore, if $\tilde{\gamma}_{1t}$ in equation 9 is large it is unlikely for γ_{1t} to be small or even zero.

4 Data

We use an historical data set released by ISTAT on March 5, 2021. The data set contains the daily number of deaths (all-causes) for the period January 1 - October 31, 2020, in all 7,903 Italian municipalities (local administrative units, LAU). In addition, we use data on the daily number of deaths for all Italian municipalities for the years 2011 - 2019.

The outcome variable of interest is the cumulative daily excess mortality per 1,000 inhabitants at the municipality level. Daily excess mortality is measured as the difference between daily mortality and the average mortality on the same day in the previous eight years. Our investigation period runs from November 1, 2019 to October 31, 2020, covering 365 days. The pre-treatment period includes 114 days from November 1, 2019 to February 22, 2020.

In the main analysis, we estimate the causal impact of not declaring a Red Zone, using 11 municipalities that experienced a Red Zone between February 23 and March 23, 2020, as a control group. These include the 10 municipalities in the provinces of Lodi and Vo' Euganeo, a municipality in the province of Padua, which was subject to the same restrictions. Figure 2 shows the trends in cumulative excess mortality per 1,000 inhabitants of our treated units (Albino, Alzano Lombardo and Nembro) and the 11 control municipalities (left panel), and for the unaffected municipalities (right panel).

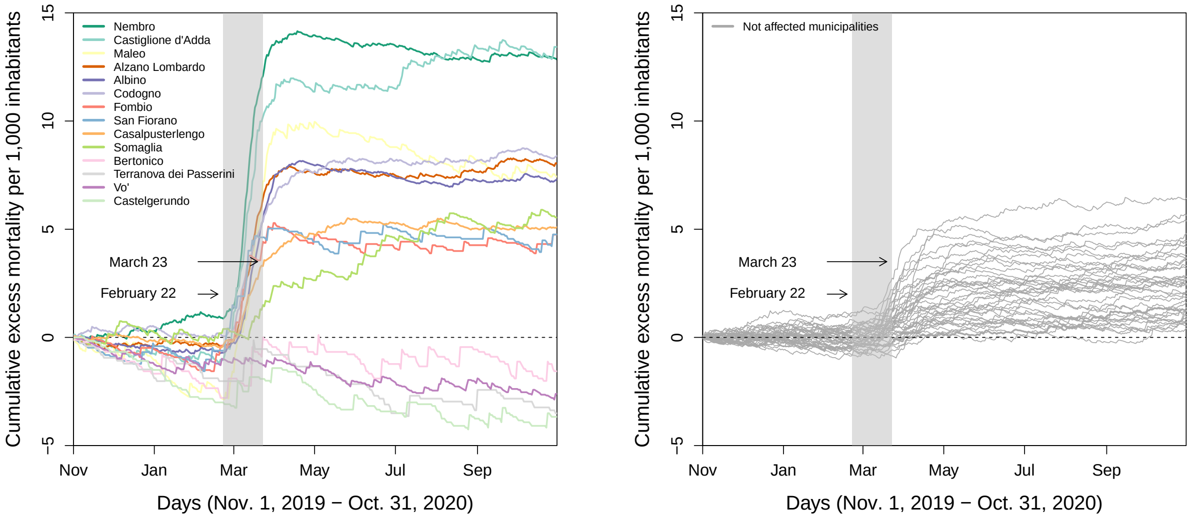

Figure 2: Cumulative excess mortality per 1,000 inhabitants (raw data).

To estimate the causal impact of the failure to declare a Red Zone, we need to assume that the impact of COVID-19 we observe in the control units can be used to recover the effect that the treated municipalities would have experienced had they implemented a Red Zone. To further investigate the plausibility of this assumption, we follow the procedure introduced in Section 3 and estimate the impact of the COVID-19 pandemic on the municipalities in our donor pool that received non-negligible weights and the total effect of the pandemic on the treated municipalities. To estimate these effects we use a donor pool consisting of a set of 39 municipalities in the Lombardy region where COVID-19 did not spread until later stages, as we can observe in Figure 2. The population of these municipalities ranges from 11,000 to 18,000 (similar to our treated) and they are more than 60 km and 50-minute drive from Codogno (the center of the Red Zone) and Albino (the center of the Serio Valley).[6]

We use a set of cumulative excess mortality predictors.[7] Epidemiological studies indic-

[6]Distances in meters and in minutes of driving time are taken from the Distance Matrices supplied by ISTAT and can be downloaded at https://www.istat.it/it/archivio/157423. Each regional matrix provides the distance in meters and minutes of driving time between pairs of municipalities within the region. Distances are computed using commercial road graphs.

[7]We experimented with a wide set of additional predictors. The results are similar to the one reported here and are available from the authors upon request.

ate demographic factors such as population, rate of urbanization and population density as crucial for understanding the spread of COVID-19 (Cho, 2020; Rocklov & Sjogin, 2020). For this reason, we include the percentage of males in the population and population density (residents per km^2) of each municipality.[8] We also control for the number of employees in manufacturing, for PM-10 as a measure of air quality and for the percentage of the population aged 65+ and 85+.[9] These variables account for the most vulnerable individuals and for those affected by respiratory diseases, which are more widespread in highly industrialized areas and are associated with a high mortality of patients affected by COVID-19. Recent geographical studies (Consolandi, 2021) hypothesize a causal link between the residential and mobility characteristics of the Serio Valley and the COVID-19 outbreak in the same area. To account for these characteristics we include a categorical variable measuring the altimetric area of each municipality ($1 = mountain$, $2 = coastal\,mountain$, $3 = inner\,hill$, $4 = coastal\,hill$, $5 = flat\,land$) and an Attraction Index. The latter varies between 0 and 100 and is computed as the ratio of the inflow of people into the municipality being studies for work or study reasons over the sum of inflows, outflows and resident inhabitants. The index, computed annually, provides a snapshot of the level of mobility in the area under investigation.[10] As for healthcare characteristics, we consider the distance, in meters, of each municipality from the municipality where the first and second closest hospital are located.[11]

The vector of synthetic weights is chosen to minimize the distance between the pre-

[8]Population data, measured as resident population in each municipality on December 21, 2019 by gender, come from ISTAT, Demographic Statistics (link).

[9]Data on employees in manufacturing come from the ISTAT Statistical Register of Active Enterprises (ASIA) archive, which covers the universe of firms and employees of industry and services at the municipa level. PM-10 information comes from Cerqua et al. (2021). They took data from the European Environment Agency (link). These variables allow us to take account of vulnerability in terms of respiratory diseases and conditions associated with a high mortality in COVID-19 infection. PM-10 data in $\mu g/m^3$ is from 573 monitoring stations distributed across the Italian territory. Cerqua et al. (2021) used kriging spatial interpolation to impute the PM-10 average yearly value for each municipality.

[10]Altimetric area data come from the Main Geographic Statistics on Municipalities by ISTAT, "Statistical Classifications and Size of Municipalities" section (link). The Attraction Index information comes Sistema Informativo STorico delle Amministrazioni Territoriali (SISTAT) (link).

[11]Distances in meters from the municipality hosting the closest and the second closest hospital are taken from the Distance Matrices supplied by ISTAT (link).

intervention characteristics of the treated and the weighted characteristics of the donor pool. Typically, this distance is the square root of a weighted sum where the positive weights reflect the predictive power of each of the, say, k predictors of the donor pool, and can be chosen via an in-sample and out-of-sample validation procedure. To find the weights to be assigned to each element of the vector of predictors, say, $\boldsymbol{v} = (v_1, ..., v_k)'$, we split the pre-treatment period into a training period and a validation period, and then selected the weights by minimizing the out-of-sample error in the validation period.[12] The predictors weights for the two donor pools are shown in Table 2. Finally, considering these predictors weights, we estimate the municipalities's weights, and the synthetic control units for the three treated municipalities.

5 Results

Figure 3 displays the cumulative excess mortality trends for the treated municipalities in the Serio Valley and their synthetic counterparts. The left panel focuses on days between November 1, 2019 and April 8, 2020 (two weeks after the national lockdown began), while the right panel's horizontal axis extends across the entire investigation period, i.e., November 1, 2019, to October 31, 2020 (365 days). Each plot shows the real cumulative excess mortality (solid line), the synthetic counterpart in the presence of the Red Zone (dotted line), the synthetic counterpart in the absence of pandemic, using the Not Affected donor pool (dashed line). The cumulative daily excess mortality remains close to zero throughout the pre-treatment period in Alzano Lombardo, and slightly below zero in Albino. In contrast, this is slightly positive for Nembro starting around January 1, 2020, which could suggest that Nembro might have experienced an anticipation effect of the pandemic compared to the other two treated units. However, the number of excess deaths is rather small and within the variation we observe in the pre-treatment period (see Figure 2). The synthetic counterparts almost perfectly overlap the observed trends of both Albino and

[12]We use as training period days between November 1-30, 2019, and as validation period days between December 1, 2019 and January 1, 2020.

Alzano Lombardo up to the beginning of our treatment period. There is a small difference between Nembro and its synthetic counterparts starting at the beginning of the year, but here, too, the difference is rather small. This suggests that the synthetic units provide a reasonable approximation. These results are confirmed by the covariate balancing. As shown in Table 1, in most cases, the synthetic control units do a good job of reproducing the cumulative excess mortality and predictors values, in contrast with the simple averages of all municipalities in our donor pools.

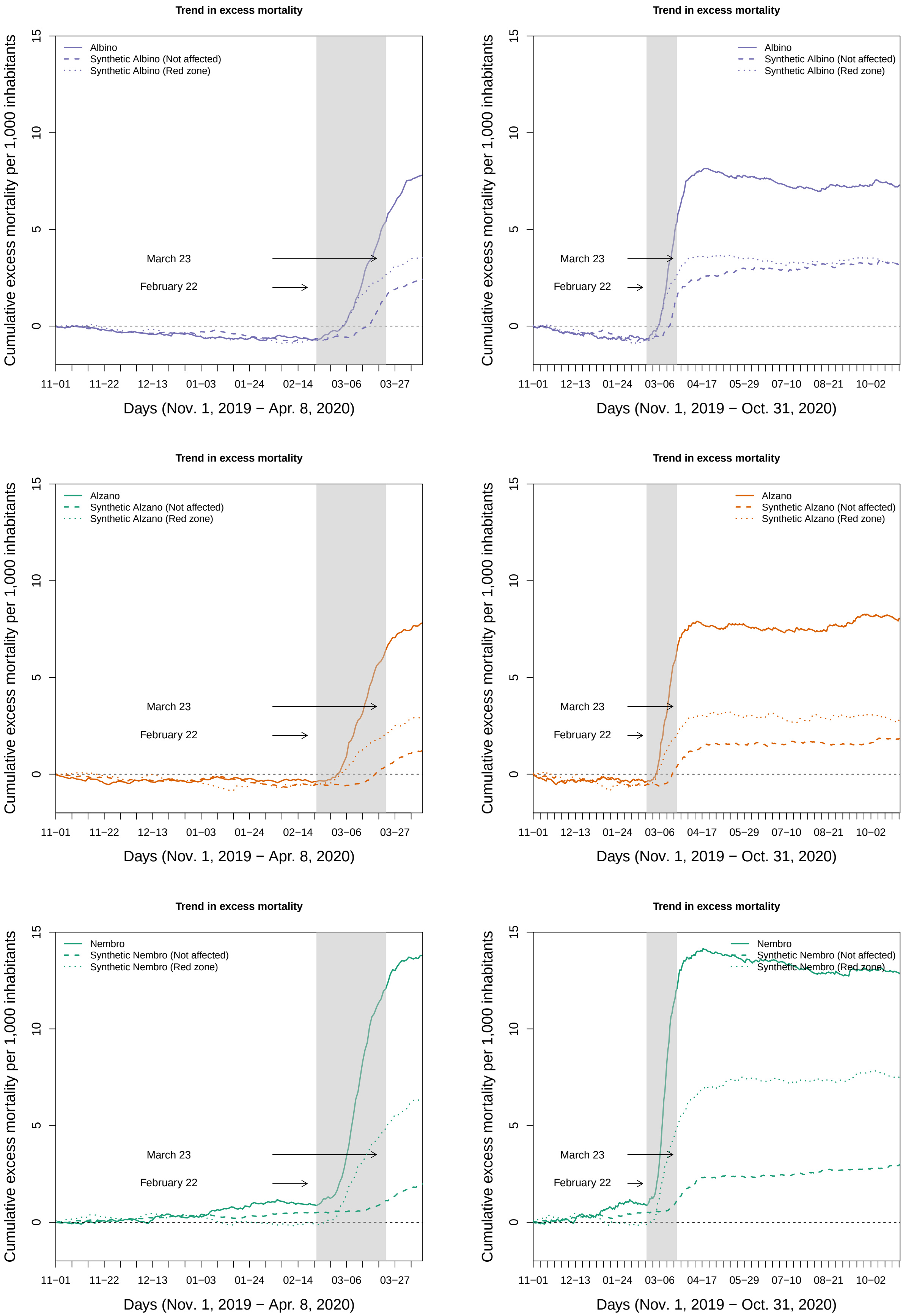

Figure 3: Trends in cumulative excess mortality per 1,000 inhabitants.

Table 1: Excess mortality, demographic, and geographical predictors means.

Mortality, Demographic and Geographic variables	Albino		Alzano Lombardo		Nembro		Average of all donors ($n = 11$)
Red Zone	Real	Synthetic	Real	Synthetic	Real	Synthetic	
-Cumulative excess mortality per 1,000 inhab.	-0.43	-0.37	-0.39	-0.34	0.23	0.25	-0.58
-PM-10 (2019)	27.99	32.61	28.12	32.69	28.12	32.65	32.51
-Share of male population (2019)	0.49	0.49	0.49	0.49	0.49	0.48	0.50
-Share of population over 65 (2019)	0.21	0.22	0.21	0.24	0.22	0.23	0.22
-Share of population over 85 (2019)	0.03	0.03	0.03	0.03	0.03	0.03	0.03
-Employees in manufacturing (2017)	2029.75	880.48	582.13	885.83	946.02	1447.79	347.06
-Attraction index (2015)	32.63	33.20	30.13	29.99	36.23	37.52	24.44
-Population density (2019)	559.58	451.34	999.90	460.65	755.77	719.39	268.45
-Altimetric area	1.00	4.67	3.00	3.99	3.00	5.00	4.82
-Distance from the closest hospital	8100.20	6996.11	0.00	5107.57	3708.14	2882.10	10264.65
-Distance from the second closest hospital	14647.63	13474.50	7087.21	12188.25	9955.18	10439.82	17305.17
Not Affected							
Mortality, Demographic and Geographical variables	**Albino**		**Alzano Lombardo**		**Nembro**		**Average of all donors ($n = 39$)**
	Real	Synthetic	Real	Synthetic	Real	Synthetic	
-Cumulative excess mortality per 1,000 inhab.	-0.43	-0.40	-0.39	-0.37	0.23	0.23	0.04
-PM-10 (2019)	27.99	30.34	28.12	28.84	28.12	28.15	28.51
-Share of male population (2019)	0.49	0.50	0.49	0.50	0.49	0.49	0.49
-Share of population over 65 (2019)	0.21	0.18	0.21	0.19	0.22	0.22	0.21
-Share of population over 85 (2019)	0.03	0.02	0.03	0.02	0.03	0.03	0.03
-Employees in manufacturing (2017)	2029.75	2051.27	582.13	1521.27	946.02	1368.77	1543.44
-Attraction index (2015)	32.63	32.21	30.13	27.85	36.23	35.58	32.58
-Population density (2019)	559.58	586.70	999.90	918.37	755.77	723.96	1000.32
-Altimetric area	1.00	4.57	3.00	4.67	3.00	3.00	4.23
-Distance from the closest hospital	8100.20	7669.61	0.00	6090.03	3708.14	3394.06	6919.05
-Distance from the second closest hospital	14647.63	14655.71	7087.21	10173.52	9955.18	15029.75	19787.23

Notes: Cumulative excess mortality per 1,000 inhab. and PM-10 are averaged in the period December 1, 2019 - January 1, 2020. All other predictors are time-invariant. The Attraction index is missing for Casalpusterlengo. PM-10 is measured in micrograms per cubic meter; distance is measured in meters; the Attraction index is computed as the ratio of the inflows of people into the municipality under investigation for work or for study reasons, over the sum of inflows, outflows and resident inhabitants; Altimetric area is a categorical variable: 1 = mountain, 2= coastal mountain, 3=inner hill, 4 = coastal hill, 5 = flat land.
Day 1 is November 1, 2019. The Red Zone was implemented on February 23, 2020. The pre-treatment period is 114 days.

The synthetic control units for Albino, Alzano Lombardo and Nembro are weighted averages of the municipalities in the donor pools. Table 2 displays the contributions of each of the municipalities in both donor pools to the synthetic control. The weights reported in Table 2 indicate that cumulative excess mortality per 1,000 inhabitants prior to the introduction of a Red Zone is best reproduced in the first donor pool (Red Zone) by Codogno, which carries the largest weight for both the synthetic Nembro and Albino and the second largest for Alzano Lombardo for which Vo' Euganeo carries the largest weight. Only three municipalities contribute to the synthetic control of Albino and Alzano Lombardo while four contribute to "synthetic" Nembro. As it is well known, the sparsity of the weights in Table 2 is typical of synthetic control estimators and is a consequence of the geometric characteristics of the solution to the optimization problem that generates the synthetic controls (Abadie, 2021).

Table 2: Municipalities' weights in the synthetic units.

Red Zone					**Not Affected**				
	Donor pool ($n = 11$)	Albino	Alzano Lombardo	Nembro		Donor pool ($n = 39$)	Albino	Alzano Lombardo	Nembro
1	Vo'	0	0.503	0.164	1	Cardano al Campo	0.001	0	0
2	Bertonico	0	0	0	2	Caronno Pertusella	0	0	0
3	Casalpusterlengo	0.089	0	0	3	Castellanza	0	0	0
4	Castiglione d'Adda	0	0	0	4	Fagnano Olona	0.001	0.524	0.044
5	Codogno	0.865	0.494	0.484	5	Lonate Pozzolo	0.009	0	0
6	Fombio	0	0.003	0.111	6	Luino	0.143	0	0
7	Maleo	0	0	0	7	Malnate	0.001	0	0
8	San Fiorano	0	0	0	8	Olgiate Olona	0.001	0	0
9	Somaglia	0	0	0	9	Samarate	0.11	0	0
10	Terranova dei Passerini	0	0	0.175	10	Sesto Calende	0.001	0	0
11	Castelgerundo	0.046	0	0.066	11	Somma Lombardo	0	0.013	0
					12	Erba	0.001	0	0
					13	Olgiate Comasco	0.001	0	0.215
					14	Morbegno	0	0	0
					15	Arluno	0.002	0	0
					16	Busto Garolfo	0.001	0	0
					17	Canegrate	0.001	0	0
					18	Castano Primo	0	0	0
					19	Cerro Maggiore	0	0	0
					20	Cesate	0.001	0	0
					21	Nerviano	0	0	0
					22	Rescaldina	0.001	0	0
					23	Solaro	0.049	0	0
					24	Vanzaghello	0.197	0	0
					25	Bedizzole	0	0	0
					26	Calcinato	0	0.297	0.742
					27	Carpenedolo	0	0	0
					28	Gardone Val Trompia	0.23	0	0
					29	Gavardo	0.24	0	0
					30	Lonato del Garda	0	0.167	0
					31	Sarezzo	0.002	0	0
					32	Mortara	0	0	0
					33	Casalmaggiore	0	0	0
					34	Castel Goffredo	0	0	0
					35	Curtatone	0	0	0
					36	Porto Mantovano	0.001	0	0
					37	San Giorgio Bigarello	0.002	0	0
					38	Besana in Brianza	0.002	0	0
					39	Lentate sul Seveso	0.001	0	0

Our estimate of the effect of a Red Zone on excess mortality is the difference between the solid lines on the right panel in Figure 3 and their dotted counterparts. Immediately after the introduction of a nationwide lockdown the three solid lines begin to bend noticeably, suggesting that the nationwide lockdown has been an effective policy measure. The discrepancy between the solid and dotted green lines suggests a large negative effect on excess mortality in Nembro had the government declared a Red Zone. The effect is less pronounced but still important for Albino and Alzano Lombardo. The difference between each solid line and its dashed counterpart is the total impact of the pandemic on excess mortality in our treated units. The impact on cumulative excess mortality of the introduction of a Red Zone in the Serio Valley, net of the effect of the pandemic, can be assessed considering the difference between the dotted line and the dashed line in each panel in Figure 3. The Red Zone would have produced a decrease in excess mortality, net of the pandemic.

Table 3 shows the treatment effects on each treated unit on April 8, 2020, i.e., two weeks after the national lockdown, and on the date of the peak excess mortality. The pandemic has increased excess deaths by around 6 and 7 persons per 1,000 inhabitants in Albino and Alzano, respectively. This amounts to around 116 and 96 inhabitants, respectively. In Nembro, the impact of the pandemic is greater, reaching about 12 deaths per 1,000 inhabitants, i.e., around 135 inhabitants. The introduction of a Red Zone would have reduced the number of deaths per 1,000 inhabitants by about 4.5 in Albino, 5 in Alzano Lombardo, and 8 in Nembro. Since excess deaths in Nembro reached around 12 persons per 1,000 inhabitants two weeks after the end of the Red Zone restrictions, this means that, had the government declared a Red Zone in Nembro, the number of excess deaths would have been about 67% lower.

Table 3: Treatment effects per 1,000 inhabitants.

	April 8, 2020	Max.
Albino (Red Zone)	4.27	4.61
Albino (not affected)	5.38	5.64
Alzano Lombardo (Red Zone)	4.91	5.35
Alzano Lombardo (not affected)	6.59	6.73
Nembro (Red Zone)	7.38	7.84
Nembro (not affected)	11.80	11.95

In our setting, the inference is challenging for three reasons. First, having only 12 units makes the method suggested by Abadie (2021) infeasible, as the minimum possible p-value would be by construction $\frac{1}{12} = 0.083$. Second, the fact that the pandemic has an effect in the post-treatment period for the units in the donor pool as well, makes it harder to interpret the usual ratio between pre- and post-treatment root mean squared prediction errors (RMSPEs), as some of the municipalities in the donor pool might have been more extremely affected by the pandemic than our treated ones would have experienced. This is actually a requirement for assuming NEPE (see section 3). Third, for Nembro, there could have been an earlier spread of the pandemic that may affect the fit in the pre-intervention period and, consequently, lower the RMSPE ratio (see the discussion in Section 6 about potential anticipation effects). Despite these potential issues, we report the standard inference procedure of Abadie et al. (2010). First, in Figure 4 we report the ratios of post- and pre-treatment RMSPE [13], which provides a measure of the post-treatment gap in excess mortality relative to the estimated pre-treatment gap. Then, we run in-space placebo tests by applying SCM sequentially to each municipality in our donor pool. At each iteration, we reassign the treatment to one of the municipalities in the donor pool and estimate the impact associated with each placebo run. The cross-sectional distribution of placebo tests for Albino, Alzano Lombardo, and Nembro is shown in Figure 5. In each panel, the Gray lines show the gap in excess mortality per 1,000 inhabitants between each municipality in the donor pool and its respective synthetic version. The superimposed black line represents the results we obtained for the respective treated unit. The estimated gap for Nembro is quite large relative to the distribution of the gaps and RMSPE ratios for the municipalities in the donor pool, indicating that the effect is probably statistically significant. The RMSPE ratio for Nembro is not among the highest if we include all municipalities. This might be driven by the small gap in the excess mortality that we observe in early 2020. However, if we only look at municipalities with an estimated positive effect, as suggested by Abadie (2021), Nembro has the second largest RMSPE ratio. The estimated gaps and

[13] We only consider the period between February 22 and April 8 to calcualte the post-treatment RMSPE as, most likely due to the National look down of March 23, we observe a general flattening in cumulative excess of all municipalities deaths shortly after this date.

RMSPE ratios for Albino and Alzano are among the largest and this suggests that the effects in those municipality are significant.

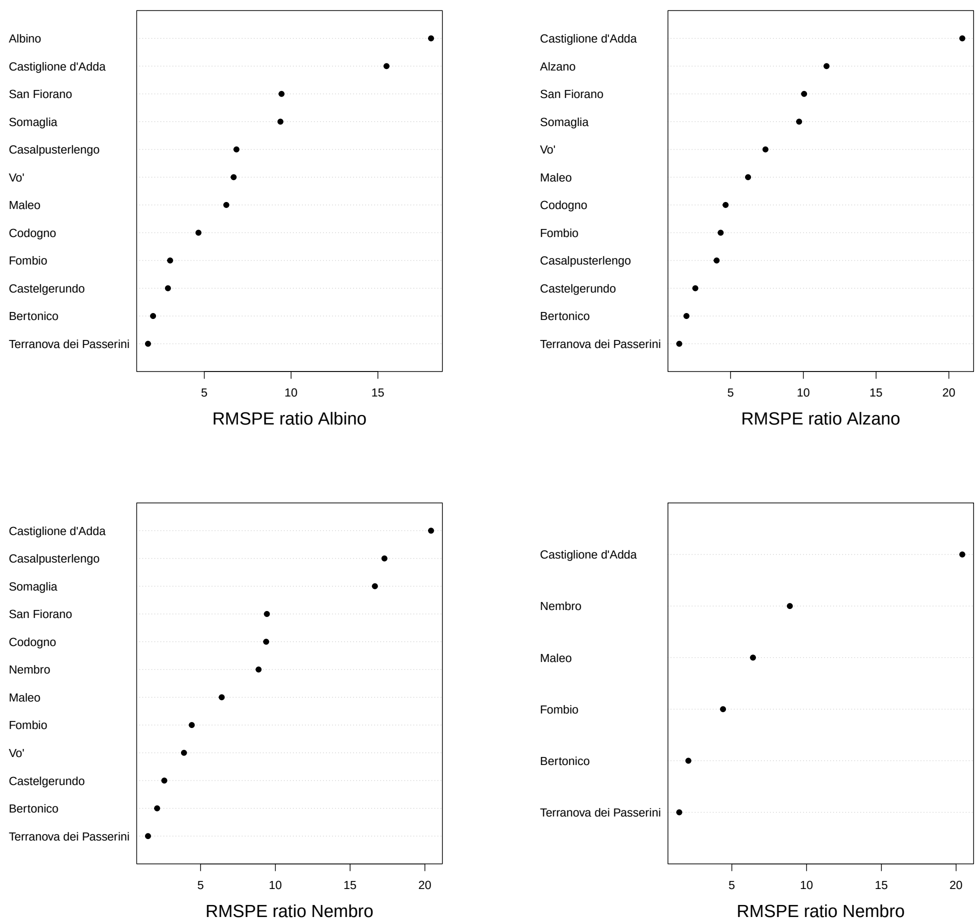

Figure 4: Ratios of post- and pre-treatment RMSPE.

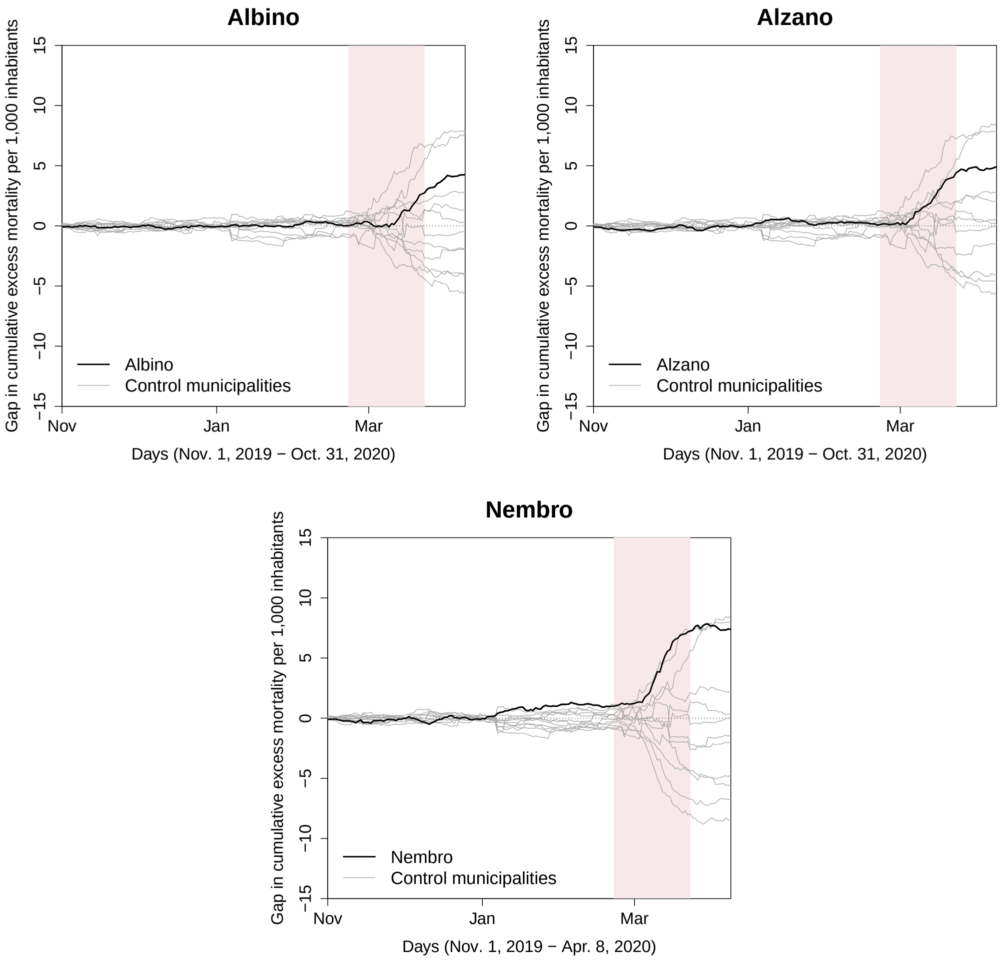

Figure 5: Placebo tests.

6 Robustness Checks

As discussed in Section 3, it is possible to investigate the sensitivity of our results to a possible violation of our NEPE assumption. To this end, we can compare the effect of not having implemented the Red Zone, estimated as in the previous section, and an alternative estimate where we use the most extreme effect observed in the Red Zone municipalities (that receive non-negligible weights). This provides an alternative estimate for the effect the pandemic would have had in our treated municipalities had they implemented the Red Zone. In all three treated municipalities, cumulative excess deaths are positive at the end of the period even with the new estimates. The new estimated effects in Albino are very small and close to zero but increasing over time until the effect of the nationwide lockdown takes over. For Nembro, the new estimated effects are almost identical to the ones we find in our main specification. As we explained in Section 3, this provides strong evidence that our results are robust to violations of our NEPE assumption.

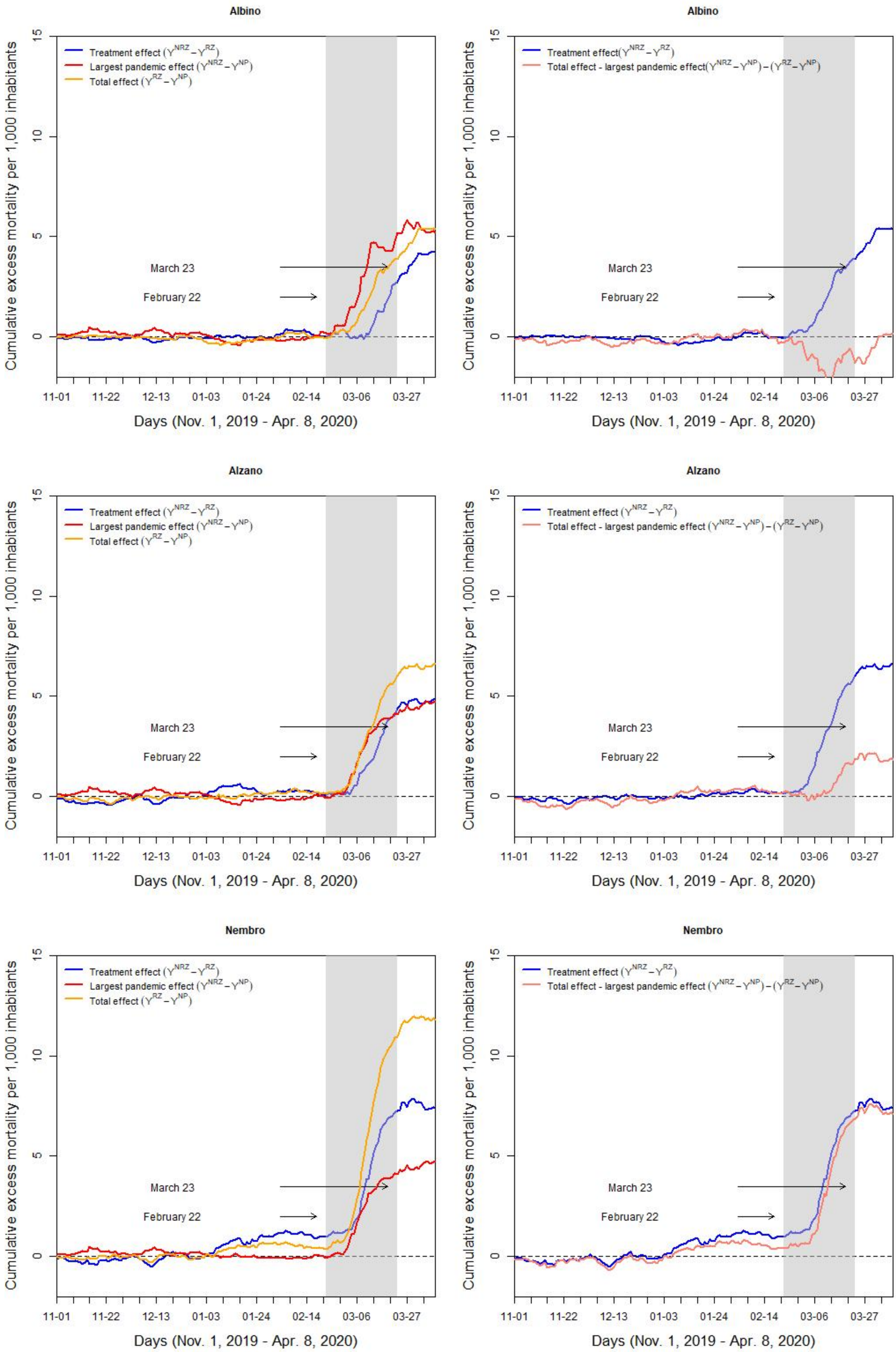

Figure 6: Trends in cumulative excess mortality per 1,000 inhabitants.

In addition, we conduct various standard checks to assess the sensitivity of our main results with respect to changes to the study design. In particular,

1. we run the usual leave-one-out analysis;
2. we exclude Vo' Euganeo from the donor pool;
3. we anticipate the treatment date as January 1, 2020.

In Figure 7 we report the results of the leave-one-out analysis excluding, one at a time, one of the municipalities that received a positive weight (see Table 2). Figure 7 shows that synthetic Albino, Alzano, and Nembro are always well below their real counterparts.

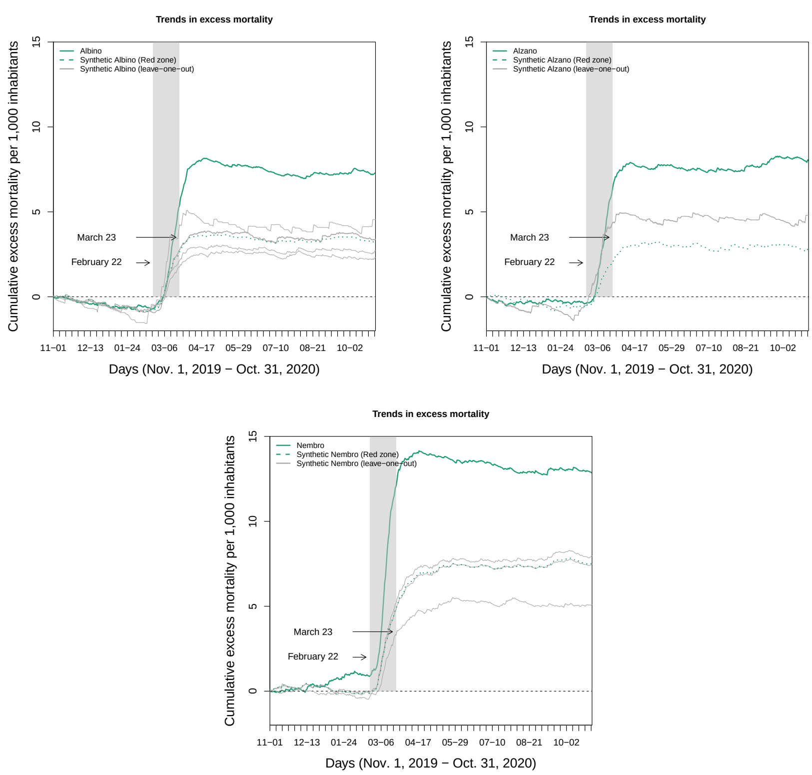

Figure 7: Leave-one-out estimates.

In a second robustness check, we restrict the donor pool to include municipalities that are located in the Lombardy region only and therefore exclude Vo'. As explained by Abadie et al. (2010), it is important to restrict the donor pool to units with outcomes that are thought to be driven by the same structural process as the treated unit. Since the pandemic had an impact on the local health system, which in Italy is regulated at regional level, it is important to focus on the Lombardy region only. Our results do not change much when we exclude Vo' Euganeo as shown in Figure 8.

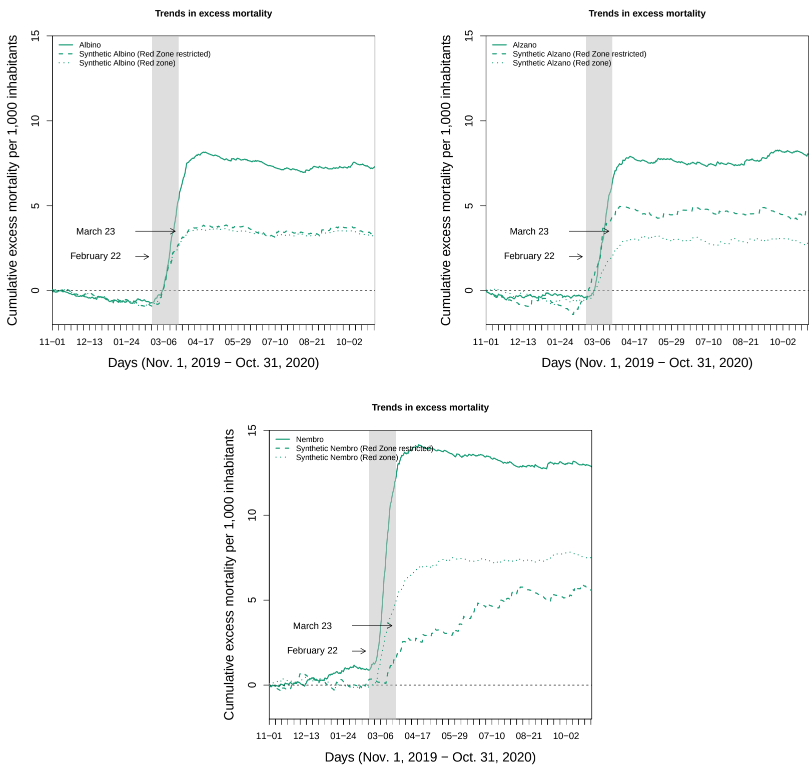

Figure 8: Restricting the donor pool to Red Zone Lombardy municipalities

Finally, we anticipate the beginning of the treatment period as January 1, 2020, and estimate anticipation effects that might have occurred before the Red Zone was introduced. The results are reported in Figure 9. Albino and Alzano do not seem to have any visible anticipation effects. For Nembro, we cannot rule out that the pandemic had already started at the beginning of the year. However, the estimated anticipation effects are relatively small and within the variation we observe among different municipalities in the entire pre-pandemic period (see Figure 2).

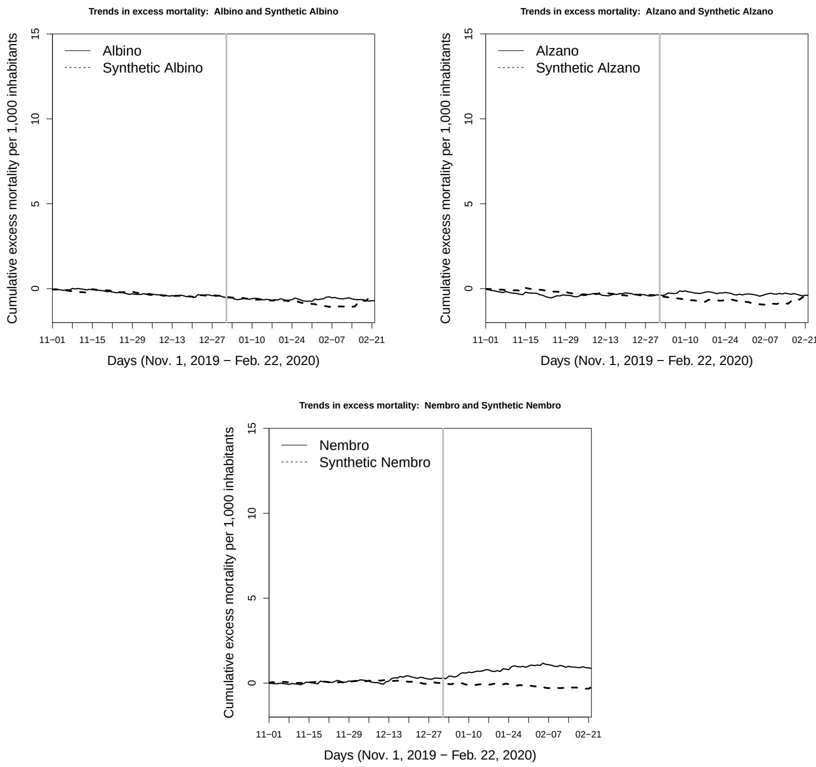

Figure 9: Backdating treatment starting date to January 1,2020.

Given the potential anticipation effect that we observe in Nembro, we report the results of a different specification where we consider January 1 as the treatment date and estimate the effects throughout the period. As shown in Figure 10 and Table 4 the results remain quite similar to our main specification.

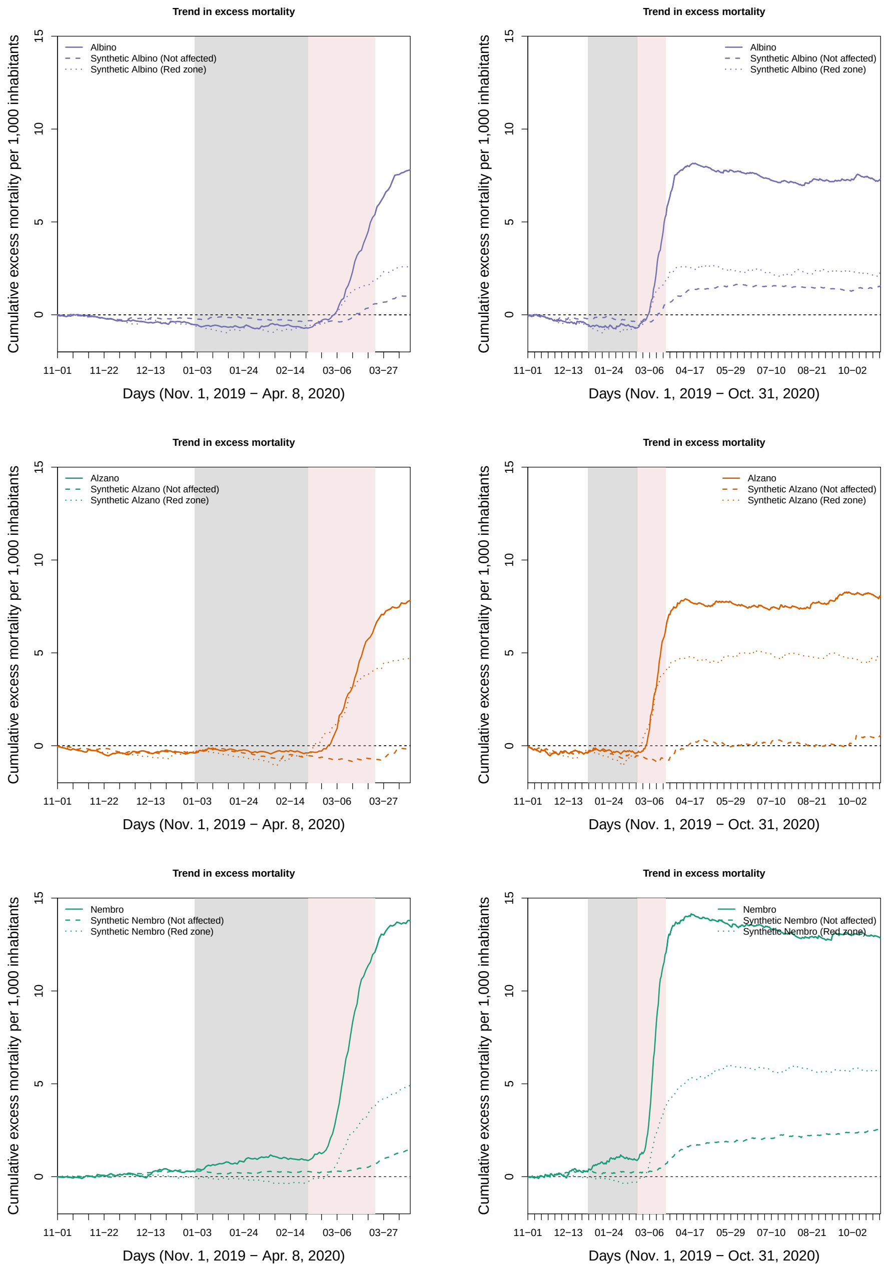

Figure 10: Trends in cumulative excess mortality per 1,000 inhabitants. The Gray area (Jan. 1-Feb. 21) represent the additional treatment period, while the pink area represents the actual treatment period before the nation-wide lockdown (Feb. 22-March 23).

Table 4: Treatment effects per 1,000 inhabitants.

	April 8, 2020	Max
Albino (Red Zone)	5.22	5.71
Albino (not affected)	6.78	6.81
Alzano Lombardo (Red Zone)	3.11	3.83
Alzano Lombardo (not affected)	7.98	8.28
Nembro (Red Zone)	8.84	9.18
Nembro (not affected)	12.31	12.48

7 Conclusion

Using the synthetic control method we estimate the causal effect of not imposing strong social distancing restrictions at the beginning of the COVID-19 pandemic in three municipalities of the province of Bergamo. We show that imposing a Red Zone, as other municipalities in the Lombardy region did, could have saved up to 67% of the deaths inflicted by the pandemic on those municipalities. Our descriptive analysis also indicates that the nationwide lockdown has potentially had a strong impact in reducing the effect of the pandemic. As a methodological contribution, we show that studies which rely on the synthetic control and/or similar methods, such as difference-in-differences, to estimate the impact of policy interventions in this setting implicitly impose an additional assumption on how the pandemic would have affected the treated units in the absence of the policy and, to validate this assumption, we show how to use municipalities that were not affected in the early stages of the pandemic.

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, *59*(2), 391–425.

Abadie, A., Diamond, A. & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association*, *105*(490), 493–505.

Abadie, A. & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, *93*(1), 113–132.

Acemoglu, D., Chernozhukov, V., Werning, I. & Whinston, M. D. (2021). Optimal targeted lockdowns in a multigroup SIR model. *American Economic Review: Insights*, *3*(4), 487–502.

Alfieri, C., Egrot, M., Desclaux, A. & Sams, K. (2022). Recognising italy's mistakes in the public health response to covid-19. *The Lancet*, *399*(10322), 357–358.

Almagro, M. & Orane-Hutchinson, A. (2021). JUE Insight: The determinants of the differential exposure to COVID-19 in New York city and their evolution over time. *Journal of Urban Economics*.

Bosa, I., Castelli, A., Castelli, M., Ciani, O., Compagni, A., Galizzi, M. M., ... others (2022). Response to covid-19: was italy (un) prepared? *Health Economics, Policy and Law*, *17*(1), 1–13.

Cerqua, A., Di Stefano, R., Letta, M. & Miccoli, S. (2021). Local mortality estimates during the COVID-19 pandemic in Italy. *Journal of Population Economics*, 1–29.

Chernozhukov, V., Kasahara, H. & Schrimpf, P. (2021). Causal impact of masks, policies, behavior on early Covid-19 pandemic in the US. *Journal of Econometrics*, *220*(1), 23–62.

Cho, S.-W. (2020). Quantifying the impact of non pharmaceutical interventions during the COVID-19 outbreak: the case of Sweden. *The Econometrics Journal*, *23*, 323–344.

Consolandi, E. (2021). Lombardie, la région italienne la plus touchée par la Covid-19. Analyse des aspects socio-territoriaux du foyer de la Val Seriana. *Revue francophone sur la santé et les territoires*, 1–17.

Fagiuoli, S., Lorini, F. L. & Remuzzi, G. (2020). Adaptations and Lessons in the province of Bergamo [Correspondence]. *The New England Journal of Medicine*.

Fang, H., Wang, L. & Yang, Y. (2020). Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China. *Journal of Public Economics*, *191*(104272), 1–9.

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., ... others (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, *584*(7820), 257–261.

Glaeser, E., Gorback, C. & Redding, S. (2021). JUE Insight: How much does COVID-19 increase with mobility. Evidence from New York and four other U.S. cities. *Journal of Urban Economics*.

Horowitz, J. (2020). The Lost Days that Made Bergamo a Coronavirus Tragedy. *The New York Times*, *30*(November).

ISTAT. (2021). *Decessi per il complesso delle cause. periodo gennaio-novembre 2020.*

Presidente del Consiglio dei Ministri. (2020a). *Decreto del presidente del consiglio dei ministri 08-03-2020* (Tech. Rep. No. 59). Rome: Gazzetta Ufficiale della Repubblica Italiana.

Presidente del Consiglio dei Ministri. (2020b). *Decreto del presidente del consiglio dei ministri 22-03-2020* (Tech. Rep. No. 76). Rome: Gazzetta Ufficiale della Repubblica Italiana.

Presidente del Consiglio dei Ministri. (2020c). *Decreto legge 23 febbraio 2020 n.6* (Tech. Rep. No. 45). Rome: Gazzetta Ufficiale della Repubblica Italiana.

Rocklov, J. & Sjogin, H. (2020). High population densities catalyse the spread of COVID-19. *Journal of Travel Medicine*, *27*(3), 1–2.

Signorelli, C., Scognamiglio, T. & Odone, A. (2020). COVID-19 in Italy: impact of containment measures and prevalence estimates of the general population. *Acta Biomedica*, *91*(3-S), 175–179.

Singh, S., Shaikh, M., Hauck, K. & Miraldo, M. (2021). Impacts of introducing and lifting nonpharmaceutical interventions on COVID-19 daily growth rate and compliance in the United States. *Proceedings of the National Academy of Sciences*, *118*(12).