

Ceriani, Lidia; Verme, Paolo

**Working Paper**

## Excess Mortality as a Predictor of Mortality Crises: The Case of COVID-19 in Italy

GLO Discussion Paper, No. 618

**Provided in Cooperation with:**

Global Labor Organization (GLO)

*Suggested Citation:* Ceriani, Lidia; Verme, Paolo (2020) : Excess Mortality as a Predictor of Mortality Crises: The Case of COVID-19 in Italy, GLO Discussion Paper, No. 618, Global Labor Organization (GLO), Essen

This Version is available at:

<https://hdl.handle.net/10419/222516>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Excess Mortality as a Predictor of Mortality Crises: The Case of COVID-19 in Italy

Lidia Ceriani  
Georgetown University,  
Washington DC, USA

Paolo Verme\*  
The World Bank,  
Washington DC, USA

July 22, 2020

## Abstract

The paper provides initial evidence that excess mortality rates by locality can be used as a statistically reliable predictor of looming mortality crises. Using recently published daily deaths figures for 7,357 Italian municipalities, we estimate the growth in daily mortality rates between the period 2015-2019 and 2020 by province. All provinces that experienced a major mortality shock in mid-March 2020 had increases in mortality rates of 100% or above already in mid-February 2020. This increase was particularly strong for males and older people, two recognizable features of COVID-19. Using panel data models, we find a strong positive and significant association between overall deaths and COVID-19 related deaths, and between early increases in mortality rates in February 2020 for any cause and the March 2020 outbreak in COVID-19 deaths. We conclude that the growth in mortality rates can potentially be used as a statistically reliable predictor of mortality crises, including COVID-19 crises.

---

\*Corresponding Author (pverme@worldbank.org).

# 1 Introduction

It is now almost certain that the spread of COVID-19 and the growth of COVID-19 related deaths emerged much earlier than previously thought [4]. Recent studies have also argued that the lethality of the virus among the population at large is not high, in fact similar to the seasonal flu [14]. However, in some geographical areas and for some population groups, the virus has been extremely morbid and lethal to an extent that emergencies rooms, hospital admissions and beds, intensive care units, and even funeral homes have been overwhelmed and incapable of responding to demand. This phenomenon also occurred over a very short period of time. From the time authorities first observe a rapid growth in infections and deaths to the time the health system collapses only one to three weeks pass, a very short window of time to prepare local institutions for a proper response. This, in turn, contributes to the failure of the health system and the growth in deaths.

A very different outcome could have been achieved if local and national authorities were able to predict the peak of the crisis three-four weeks in advance. This would have provided significantly more time to prepare emergency rooms, hospital admissions and beds, intensive care units and stock up with essential medical supplies. This paper argues that this would have been possible by monitoring excess mortality in real time. We illustrate this claim with newly released data on deaths for 7,357 municipalities in Italy during the period January-April, 2020. Recent articles on the Financial Times [3] and Our World In Data<sup>1</sup> show that these data are available for many countries and are available from various data repositories such as the European Center for Disease Prevention and Control (ECDC) or the US Centers for Disease Control and Prevention (CDC). Data on number of deaths are generally available at the national level, on a weekly basis, occasionally by age group or gender, and non weighted by the population. In this paper, we use daily figures at the municipality and provincial level, by age and gender and population weighted. As compared to publicly available data, we are able to improve excess mortality estimates on all these fronts. Prior versions of these data have been used in various blogs <sup>2</sup>.

Several recent papers have used similar data to study excess mortality in Italy. [2] provides spatial evidence of excess mortality using weekly data; [7] offers a statistical overview of excess mortality by age and gender at the municipality level; [12] studies the correlation with excess mortality and COVID-19 mortality and uses

---

<sup>1</sup><https://ourworldindata.org/covid-excess-mortality>

<sup>2</sup><https://voxeu.org/article/covid-19-italy-analysis-death-registry-data>;  
<https://www.lavoce.info/archives/66848/zone-rosse-locali-un-indice-per-decidere-dove/>

excess mortality to estimate Population Fatality Rates (PFR) and Infection Fatality Rates (IFR) by age group finding much higher than reported rates; [6] calculate excess mortality for the five most heavily hit provinces by age group and estimates reductions in life expectancy; [5] provide additional evidence on excess mortality by age and gender used as a proxy for COVID-19 mortality. Excess mortality in the context of COVID-19 has also been recently studied in several other countries such as Belgium [13], Portugal [18], United States and the United Kingdom [11]. In our knowledge, this paper is the first that uses daily excess mortality at the local level to predict looming mortality crises. It improves on previous contributions using more granular data and combining statistical observations with econometric analyses.

Excess mortality is affected by COVID-19 deaths, but can also be affected by policy measures implemented as a response to COVID-19 such as lock-downs and social distancing, which can lead to reduced road accidents or increased number of homicides and suicides. To avoid these policies polluting our data, we mainly focus on the period prior to the lock-down in Italy. We can also exclude the implementation of other milder preparatory measures as Italy was one of the first countries to be hit hard by the virus after China and was unprepared to respond to the surge in deaths experienced in March 2020. It is also possible that the year 2020 might have been an irregular year in terms of mortality when compared to previous years, even in the absence of COVID-19. Indeed, we will present some evidence showing that mortality rates were lower than usual prior to COVID-19, potentially contributing to explain why the rise in deaths went largely unnoticed in January and February 2020.

The paper finds that: (i) the mortality rate in Italy (for any cause) was rising sharply and deviating significantly from previous years' average in many Italian provinces already in mid-February 2020, despite an otherwise lower than average mortality year, and (ii) early growth in these mortality rates by province can be used as statistically reliable predictors to identify those provinces heading towards a mortality crisis. This indicator alone can anticipate a mortality crisis by several weeks. Had this information been available to policy makers in mid-February 2020 and had the lock-down in Italy been introduced at the same time, many lives could have been saved. As COVID-19 is still spreading across the world while re-surfing in other parts, excess mortality can still be used as a predictor of COVID-19 crisis. More generally, excess mortality, if measured in real time, can provide a useful predictor of looming mortality crises.

The next section illustrates the data, section 3 provides the analysis of the mortality rates by province, section 4 and 5 show how the mortality rate can be used as a predictor for municipalities heading towards a crisis, section 6 shows how forecasting methods would have been difficult to apply in the case of the COVID-19 epidemics,

and section 7 discusses results. This last section will also provide some explanations on why the sharp increase in mortality rates experienced by many Italian provinces in February 2020 went largely unnoticed.

## 2 Data

The number of daily deaths is a mandatory indicator that all municipalities in Italy are required to provide to various central administrations, including the Italian National Institute of Statistics (*Istituto Italiano di Statistica* – ISTAT). This information is validated, curated and provided publicly in aggregated form by means of *i.Istat*, the official national repository of statistical data maintained by ISTAT. This process is lengthy and, under normal circumstances, ISTAT provides this information publicly one to two years following reception of these data. Under the exceptional circumstances of COVID-19, ISTAT has been requested by the Ministry of Health to produce and publish these data rapidly and has released since the end of March 2020 several versions of these data in disaggregated form [8]. We analyze the latest issue of these data to understand whether they could have been used to anticipate the mortality crisis experienced by several provinces in Italy during the month of March 2020.

The latest data provided by ISTAT include the daily number of deaths for 7,357 municipalities corresponding to 93.1 percent of all municipalities in Italy, 95 percent of the population, and covering 107 provinces. ISTAT excluded municipalities that were not providing mortality data in a complete or timely fashion. Data are also provided by gender and age group (22 age groups, in brackets of 5 years each, with the exception of the two groups made of individuals age  $< 1$  and  $> 100$ ). These data allow, therefore, to analyze the data by gender and age, two key dimensions to understand whether the growth in deaths for any cause relates to COVID-19 deaths.

From these data, we first computed mortality rates as the number of deaths over the the population using population figures available at i.ISTAT [9]. For a given location  $l$ , gender  $g$ , and age group  $a$ , the mortality rate at time  $t$  is defined as:

$$m_{lga}^t = \frac{D_{lga}^t}{P_{lga}^t} * 100,000 \quad (1)$$

where  $P_{lga}^t$  is the population size,  $D_{lga}^t$  indicates the number of deaths and  $t$  is time expressed in days. The growth in the mortality rate between one year  $y_1$  and the reference period  $y_0$  is then defined as:

$$\Delta m_{lga}^t = \frac{m_{lga,y_1}^t - m_{lga,y_0}^t}{m_{lga,y_0}^t} * 100 \quad (2)$$

which is also known as the *P – score*, and corresponds to the general definition of “excess mortality”.

In this paper we are primarily interested in comparing mortality rates observed in  $y_1 = 2020$  *vis-à-vis* the average mortality rate in the previous five years  $y_0 = 2015-2019$ , but we will also show the growth in mortality rate between each year in the period 2015-2019 and the average of the remaining four years in the same period to gauge whether 2020 has been indeed an exceptional year.

To compute these rates, we take the resident population by gender and age as of January 1st of each year for the years 2015-2019. For 2020, we impute the same population size as of 2019, under the hypothesis that there is zero population growth in 2020. This should not affect results as the population of Italy has been very stable during the past five years and is not expected to increase significantly in 2020.

Table 1 summarizes population size as of January 2019 (the last available data point) and sample statistics for all 107 Italian provinces for the year 2020. Since data is missing for some municipalities in some days of the period of analysis, the last two columns report the average daily sample in each province, as a share of total population, and respective standard deviation. The population included in the sample on average account for 96% of the total provincial population. The lowest coverage is observed in Belluno (74%) and in Barletta-Andria-Trani (75%), while for 41 provinces daily population coverage is above 99 percent. Underlined in boldface fonts are the five provinces singled out in the analysis: Bergamo, Brescia, Cremona, Lodi and Piacenza. For these provinces, the population included in the sample account for almost 100% of the total population, with an average standard deviation of 0.1%. In essence, the sample we consider is close to a census and we do not expect any issue related to sample bias.

It is worth noting that the analysis takes full account of the variability of population coverage within Province in each day of the period considered. Provincial-level mortality rates are obtained, in fact, by adding up the number of deaths reported in the municipalities with non-missing information for each day, divided by the total population of the same number of municipalities. We further cleaned the dataset to remove all instances where a municipality did not report any data, and all observations pertaining to new municipalities instituted in February 2019 for which population data were not available. We also dropped observations on the number of deaths reported for February 29th, as only two of the six years under consideration are leap years (2016 and 2020).

Although data are available at the municipality level, the analysis is run at the level of regions and provinces. Having access to municipality data, one can run the analysis at the regional, province or municipality level. There are pros and cons of working with each of these levels. The regional level is too large to capture much of the variation in mortality rates across the country, but this level has the distinct advantage of having data on daily deaths for both overall and COVID-19 related deaths, whereas data on COVID-19 daily deaths is not publicly available at the provincial or municipal level. The municipal level provides the highest disaggregation level but the number of daily deaths per municipality is very small, often zero, and extremely variable on a daily basis. For this level, we also miss daily data on COVID-19 including deaths and infections.

The province level is therefore the best compromise in terms of geographical aggregation as there is a sufficient number of deaths to stabilize trends, and has also information on COVID-19 daily infections (although daily COVID-19 deaths are missing). To make the most of available data, we will therefore study the association between overall and COVID-19 deaths at the regional level and conduct the rest of the analysis at the provincial level. The province level is also appropriate from a logistical and policy perspective. Hospitals that have been overwhelmed by demand in mid-March 2020 cover areas that are larger than the municipality they reside in. They typically cover provinces and, in some cases, they cover regions.<sup>3</sup>

The analysis is performed using the software Stata. To avoid major daily fluctuations, data were smoothen using the Stata function `lowess` [17], which transforms data using a locally weighted regression (bandwidth of 0.8). The data set by province used in the panel regressions included 11,235 observations. This is a panel of 105 daily observations (January 1st - April 15th) covering 107 provinces. The panel is balanced meaning that we have all 105 daily observations for all 107 provinces. All data used in this paper are publicly available and we can provide the Stata codes used for the analysis upon request.

We define a mortality “crisis” as an excess mortality of 200% or above, or a three-fold increase in the 2020 mortality rate as compared to the average of the previous five years. We focus our analysis on five provinces including the provinces of Bergamo, Cremona, Lodi and Brescia in the region of Lombardy and the province of Piacenza in the region of Emilia-Romagna. During the peak of the crisis (March 10th-20th, 2020) and in these provinces, the average mortality rate was over four folds higher than in previous years. These provinces became known as the first epicenter of COVID-19 deaths in Italy, and remain, as of July 2020, the provinces that have

---

<sup>3</sup>The Italian health system is decentralized and the administrative level that is responsible for health services is the region.

been affected the most by the epidemic. They are the provinces that became sadly known for the crisis in essential health services including medical supplies, hospital beds, ICUs units and funeral homes.

### 3 Excess Mortality in January-April 2020

Excess mortality rates (any cause) were already very visible for many Italian provinces in February 2020 with peaks of growth of over 100%, even when considering smoothed data.<sup>4</sup> Figure 1 shows the growth rate in the mortality rates as defined in Equation (2) between January 1st and April 15th for the five provinces that experienced the sharpest growth in mortality (Bergamo, Cremona, Lodi, Piacenza and Brescia). Panels *a*–*e* show the growth in mortality rates between each year in the period 2015-2019 and the average of the remaining four years, while panel *f* shows the growth in mortality rates between 2020 and the average mortality rate in the previous five years. A percent increase of 100 on the *y-axis* indicates that the mortality rate has doubled as compared with the average of each reference period.

All years in the span 2015-2019 show a very similar pattern in mortality rates, with significant variations from zero being observed only in some weeks in 2015 and 2017, two years known for a particularly virulent flu season. By contrast, in 2020 the growth rate of mortality starts to deviate from zero already in early February and doubles around mid-February 2020. Figures 2–6 report the same data as Figure 1 for each one of the five provinces, displaying also the standard deviations around the mean. The year 2020 is clearly exceptional for any of the provinces considered. This is particularly remarkable given that the year 2014-2015 has been reported as the year with the highest mortality rate since World War II [16].

These same figures show that the growth in mortality rate (any cause) was particularly sharp for men and older people, a recognizable feature of COVID-19. Figure 7 shows the same curves shown in Figure 1, panel *f*, by gender. The figure shows that the curves for males are steeper and have increased to higher levels than those for females by March 2020, especially in the hardest hit provinces such as Bergamo. Figure 8 repeats the exercise by age groups and shows that the mortality crisis has been remarkably sharp for people age 60 and above, in fact this age group exhibits a growth in mortality that is more than twice as high as the one for the age group below the age of 60. By splitting the growth in mortality rates by gender and age, it is possible therefore to quickly recognize the deadly characteristics of COVID-19, even in the absence of data on the causes of deaths. This is important because the

---

<sup>4</sup>See also [15] for an illustration of an earlier version of these data.



data on the causes of deaths are complex as many deaths are cataloged with multiple causes while others are mis-classified, which was probably the case for the early deaths due to COVID-19.

To clarify the progression of the unfolding mortality crisis in the provinces that experienced the major problems, we provide in Table 2 the dates in 2020 that corresponded to a 100, 200, 300, and 400% growth in mortality rates as compared to 2015-2019. For example, it is possible to appreciate the fast progression in Bergamo: February 18th (+100%), February 25th (+200%), March 4th (+300%), and March 14th (+400%). This rapid progression is rather remarkable also because obtained with smoothed data. Had we used crude data, excess mortality for Bergamo would have shown an increase as high as 800%. This highlights how important is to predict the crisis, even if only with a few days in advance. It also highlights that predicting the crisis in the worst case scenario such as the province of Bergamo is harder than predicting the crisis in other localities simply because the speed of growth in mortality was higher than elsewhere.

In essence, plotting these simple graphs by province in mid-February 2020 could have provided an indication that mortality rates were highly atypical when compared to previous years. This would have been possible four weeks in advance of the lock-down in Italy and at a time when the WHO and China had already officially announced the COVID-19 epidemics. Had these graphs by province been available to Italian authorities in real time, they would have alerted the government of those provinces that they were likely heading towards a mortality crisis.

## 4 The association between deaths for any cause and COVID-19 statistics

The Italian authorities publish data on COVID-19 at the national and regional level, including data on deaths.<sup>5</sup> This makes it possible for us to study the association between mortality for any cause and COVID-19 mortality across regions. In Table 3 we show the results of linear panel regressions across regions and time. In the top panel (Panel A) we study the association between the number of COVID deaths and excess mortality defined with different indicators: percentage difference (delta) or simple difference (diff) in deaths or mortality rates. We also provide six flavors of these equations introducing NUTS1 and NUTS2 local fixed effects and the number

---

<sup>5</sup><https://raw.githubusercontent.com/pcm-dpc/COVID-19/master/dati-regioni/dpc-covid19-ita-regioni.csv> .

of tests conducted in regions as controls. Only the coefficients of interest and their relative standard errors are reported in the Table.

With no exception, these correlations are all positive and highly significant, which strongly suggest that excess mortality and deaths in 2020 and COVID-19 deaths are causally associated. Panel B and C in Table 3 repeat this exercise, this time using the number of COVID-19 positive cases and the total number of cases as dependent variable. The results are the same in that, with no exception, excess mortality is strongly associated with COVID-19 statistics. This association was expected of course but the statistical association that we find provides the evidence needed to justify the use of the peak of excess mortality as a proxy of the peak of COVID-19 mortality, and focus on excess mortality to study the prediction capacity of this indicator for COVID-19 mortality crises.

## 5 Predicting mortality (*ex-post*)

Could excess mortality be used to make statistically reliable predictions of looming mortality or COVID-19 crises? We show with a simple panel equation how the sharp growth in mortality experienced by many Italian provinces in March 2020 could have been predicted by simply observing excess mortality during the months of January and February, 2020.

Recall that we dispose of a panel data set of mortality rates growth for Italian provinces with daily observations on the number of deaths covering the January-April 2020 period. With these data, we can estimate a panel prediction model where we use as predictors lagged variables of the growth in mortality rate. The model is described as follows:

$$y_{i,t} = \alpha_i + \beta y_{i,t-k} + \gamma I_i + \eta_{i,t} \quad (3)$$

where  $i$  and  $t$  are provinces and time expressed in days respectively,  $y_{i,t}$  is the growth of the mortality rate in 2020 as compared to the period 2015-2019 (excess mortality),  $y_{i,t-k}$  is this same variable lagged by  $k$  days,  $I_i$  are provinces fixed effects and  $\eta_{it}$  is the error term. This is what is generally referred to as a panel fixed effects model. Similar models which are popular for this kind of analysis are the Arellano-Bond types of models [1]. With lagged variables as independent variables, these models are also referred to as Dynamic Panel Data (DPD) models.

In our case, the outputs of interest are both  $\beta$  and  $\gamma$ . The  $\beta$  coefficient will provide evidence on how long we can go back in time to find growth in mortality rates that are correlated with the present growth in mortality. The  $\gamma$  coefficients will

tell us instead which provinces are more likely to have higher mortality growth with such lags. And the significance level of the  $\gamma$  coefficients informs us on how strong the prediction of higher growth in mortality is. With this comparison, we are able to tell for how many provinces we could have predicted in February 2020 the sharp growth in mortality observed in mid-March 2020.

Table 4 shows the results for the top five provinces that experienced a growth rate in mortality of over 200% in March 2020. We find positive and significant coefficients for the lagged variables with lags of 1, 7, 14 and 28 days. The coefficient increases from 1.02 for a one day lag to 2.55 for a 28 lag with the significance level (*z-stat*) decreasing as the lag becomes longer. Both these effects are expected as the change in excess mortality is larger if compared with a date further back in time whereas the correlation becomes weaker when time is added between the current time and lagged time. However, the correlation is always positive and significant indicating that lagged values are good predictors and that you can still see a significant relation when going back up to 28 days.

The coefficients for the province fixed effects are also all positive and significant. They are all positive by design because the base category is the province with the lowest excess mortality. What is important to note here is the ranking of the provinces, which is preserved across equations. The table reports the five provinces with the highest coefficients and these are the same provinces that had the highest excess mortality rates in March 2020. This is true whether we consider a one day or 28 days lag. In other words, the first province that passed the 100% excess mortality (Bergamo) in February 2020 was also the first province to experience the mortality crisis in March 2020.

These results provide two important insights. One is that the mortality crisis experienced by provinces such as Bergamo or Cremona could have been predicted by monitoring excess mortality, and mitigated by preparing the response in these provinces when excess mortality was approaching 100%. And the second is that - had Italy not implemented the lock-down from March 10th - all the other provinces that reached a 100% or above mortality in March 2020 would have probably experienced a crisis similar to the one experienced in Bergamo. In essence, and given that Italy was one of the first countries hit hard by the virus, the government reacted quickly and saved lives in doing so, but it could have reacted earlier saving more life if only excess mortality had been monitored in real time.

## 6 Forecasting mortality (*ex-ante*)

The retrospective panel regressions illustrated in the previous section have shown that - *ex-post* - it is possible to assert that a doubling of excess mortality is a sufficient condition to identify those provinces that are heading towards a sharp mortality crisis. Now we know that a 100% excess mortality is a sufficient condition to ring the alarm of a looming mortality crisis. This is indeed a useful piece of information to prepare for future crises.

But would this have been possible back in February 2020? In this section, we ask the question of whether it would have been possible to forecast excess mortality during the peak in March 2020 with the data available up to mid-February 2020. Here we show that this would have been very hard to do because of the sudden exponential growth of excess mortality towards the end of February 2020.

With little knowledge of a new virus spreading and the factors that drive this spread, forecasting excess mortality would rely exclusively on lagged values of excess mortality. In other words, we would be in the domain of time-series forecasting, similarly to predicting future prices based exclusively on past prices data. In our case, we dispose of daily observations of excess mortality starting from January 1st, 2020. This excess mortality is rather flat up until mid-February, also for those provinces that experienced the worst crisis in March 2020. For the latter provinces, excess mortality suddenly turns to exponential growth around the third or fourth week of February. Hence, predictions of excess mortality for the third or fourth week of February based on the time series up to the third week of February would have resulted in almost no growth. This is also true whether we work by province, with pooled data or with panel data.

This is shown with a simple experiment displayed in Figure 9. Here we forecast excess mortality for the five provinces with the highest excess mortality based on past periods extending the data used for forecasting progressively forward. Excess mortality forecasts are very far from the actual values until a sufficient number of growth days are included in the forecasting time-series. This happens around the beginning of March 2020, only a week in advance of the government mandatory lock-down. In other words, even if the government had been monitoring excess mortality by province, it would not have been able to forecast in February 2020 the peak of excess mortality in March 2020.

However, we now know from the visual inspection of excess mortality and the panel regression analysis in the previous sections that forecasting would not have been essential. All that was needed was to monitor daily excess mortality to check if it deviated significantly from historic trends. In particular, a doubling of excess

mortality was an extremely exceptional event that would have alerted any observer had these data been available in real time.

## 7 Discussion

Using recently released data on deaths for any cause covering the period January-April, 2020, we showed that the growth in daily mortality rates was already above 100% in early February 2020 in many Italian provinces. We also showed that these provinces are the same that experienced a major spike in mortality due to COVID-19 in mid-March, 2020. Using a simple panel model, we then showed that there is a robust statistical association between the growth in mortality rate and its lagged values of one, two, three and four weeks. Simply monitoring the daily growth rate in mortality rate by province would have shown that some provinces were experiencing a major shock in early February, 2020, an indication that could have alerted and mobilized the Italian authorities three to four weeks in advance of the lock-down day of March 10th, 2020.

The natural question that arises is why local municipalities did not notice such an increase in mortality rates in February 2020. There are, in fact, several good reasons. First, it would have been difficult to notice the increase in deaths in February 2020 at the municipality level because data at this level are characterized by sharp daily variations and are low in size with many municipalities reporting zero or a few deaths on many days. One would need to aggregate this information by province to see the real change, an information that was not available to local municipalities. Second, some local municipalities may have noticed an increase in deaths but such increase could not be properly compared with previous years or other municipalities because such data were not immediately available to local officials. Third, mortality increases in February 2020 by province were large, but not sufficiently large to overwhelm local services such as hospitals or funeral homes, which is what eventually alerted local and national officials. Fourth, even if some local officials had noted an atypical growth in mortality, they may have attributed the increase to a random spike as they did not dispose of proper benchmarks. Fifth, by coincidence and prior to COVID-19 effects, mortality rates in 2020 were lower than the average of previous years. The initial growth in mortality due to COVID-19 in late January 2020 simply brought mortality rates to average levels.

A second important question is why the flu monitoring system did not pick up warnings of a pandemic already in February 2020. Italy has a rather sophisticated monitoring system for the flu (Influnet) which collects information from doctors, laboratories and the population on a weekly basis to monitor the development of flu

epidemics in real time. These indicators are reviewed on a weekly basis and constitute the main source of information for a weekly report prepared by the flu surveillance system [10]. However, in our understanding, the growth in daily mortality rate by municipality or province was not an indicator included in the monitoring system of the flu. To our knowledge, this system does not integrate information on growth in mortality rates for any cause, possibly because it focuses on the flu only.

If Italian authorities had monitored the growth in mortality rates in real time by province, it would have been immediately obvious that something very atypical was occurring. While the growth in mortality rate may not be a standard monitoring indicator for the seasonal flu, we showed that it can be a very effective indicator to alert authorities of unfolding mortality crises, including COVID-19 crises. As a rule of thumb, a 100% increase in the mortality rate at the provincial level could be used as an early warning signal. Moreover, generating the growth rate in mortality by age and gender can provide some early indications that this growth is related to COVID-19. This is especially important in view of the current COVID-19 crisis, which is expected to re-surge in countries where the lock-down will be eased, or is still in its early phase as in many countries around the world.

## References

- [1] M. Arellano and S. Bond. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies*, 58(2):277–297, 1991.
- [2] M. Blangiardo, M. Cameletti, M. Pirani, G. Corsetti, M. Battaglini, and G. Baio. Estimating weekly excess mortality at sub-national level in italy during the covid-19 pandemic. *medRxiv*, 2020.
- [3] J. Burn Murdoch, V. Romei, and C. Giles. *Global coronavirus death toll could be 60% higher than reported*, Financial Times, Last checked: April 28th 2020.
- [4] Chiu, Allyson and Armus, Teo. *Autopsies find first U.S. coronavirus death occurred in early February, weeks earlier than previously thought*, Washington Post, Last checked: April 22nd 2020.
- [5] D. del Re and P. Meridiani. Monitoring the covid-19 epidemics in italy from mortality data. *medRxiv*, 2020.

- [6] S. Ghislandi, R. Muttarak, M. Sauerberg, and B. Scotti. News from the front: Estimation of excess mortality and life expectancy in the major epicenters of the covid-19 pandemic in italy. *medRxiv*, 2020.
- [7] D. Gibertoni, K. Y. C. Adja, D. Golinelli, C. Reno, L. Regazzi, and M. P. Fantini. Patterns of covid-19 related excess mortality in the municipalities of northern italy. *medRxiv*, 2020.
- [8] ISTAT. *Decessi e Cause di Morte: Cosa Produce l'ISTAT*, <https://www.istat.it/it/archivio/240401>, Last checked: April 28th 2020.
- [9] ISTAT. *I.Stat*, <http://dati.istat.it/>, Last checked: April 28th 2020.
- [10] Istituto Superiore di Sanità. *Sistema di Sorveglianza Integrata dell'Influenza*, <https://old.iss.it/site/RMI/influnet/>, Last checked: April 28th 2020.
- [11] E. Kontopantelis, M. A. Mamas, J. Deanfield, M. Asaria, and T. Doran. Excess mortality in england and wales during the first wave of the covid-19 pandemic. *medRxiv*, 2020.
- [12] C. Modi, V. Boehm, S. Ferraro, G. Stein, and U. Seljak. How deadly is covid-19? a rigorous analysis of excess mortality and age-dependent fatality rates in italy. *medRxiv*, 2020.
- [13] G. Molenberghs, C. Faes, J. Aerts, H. Theeten, B. Devleesschauwer, N. Bustos Sierra, T. Braeye, F. Renard, S. Herzog, P. Lusyne, J. Van der Heyden, H. Van Oyen, P. Van Damme, and N. Hens. Belgian covid-19 mortality, excess deaths, number of deaths per million, and infection fatality rates (8 march - 9 may 2020). *medRxiv*, 2020.
- [14] D. D. Rajgor, M. H. Lee, S. Archuleta, N. Bagdasarian, and C. Q. Quek. The many estimates of the covid-19 case fatality rate. *The Lancet Infectious Diseases*, DOI:[https://doi.org/10.1016/S1473-3099\(20\)30244-9](https://doi.org/10.1016/S1473-3099(20)30244-9), 2020.
- [15] F. Ricciato. *A preliminary view at 2020 mortality data from ISTAT*, <https://ec.europa.eu/eurostat/cros/system/files/>, Last checked: April 27th 2020.
- [16] A. Rosano, A. Bella, F. Gesualdo, A. Acampora, P. Pezzotti, S. Marchetti, W. Ricciardi, and C. Rizzo. Investigating the impact of influenza on excess mortality in all ages in italy during recent seasons (2013/142016/17 seasons). *International Journal of Infectious Diseases*, 88:127 – 134, 2019.

- [17] STATA Corporation. *Lowess smoothing*, <https://www.stata.com/manuals13/>, Last checked: April 27th 2020.
- [18] A. Vieira, V. R. Peixoto, P. Aguiar, and A. Abrantes. Rapid estimation of excess mortality in times of covid-19 in portugal - beyond reported deaths. *medRxiv*, 2020.



# List of Tables

Table 1: Population coverage, by provinces, Jan 1st–Apr 15th 2020

Province	Pop Jan '19	Average Daily Sample (% Pop)	sd
Agrigento	434,870	90.4	1.0
Alessandria	421,284	99.4	0.3
Ancona	471,228	88.1	1.4
Arezzo	342,654	99.4	0.5
Ascoli Piceno	207,179	97.1	1.7
Asti	214,638	92.7	1.9
Avellino	418,306	97.2	0.9
Bari	1,251,994	97.8	0.0
Barletta-Andria-Trani	390,011	74.6	0.8
Belluno	202,950	74.2	2.4
Benevento	277,018	94.3	1.8
<b>Bergamo</b>	<b>1,114,590</b>	<b>99.8</b>	<b>0.2</b>
Biella	175,585	98.2	1.2
Bologna	1,014,619	95.9	0.7
Bolzano	531,178	97.4	1.1
<b>Brescia</b>	<b>1,265,954</b>	<b>99.9</b>	<b>0.1</b>
Brindisi	392,975	98.8	0.7
Cagliari	431,038	83.6	0.0
Caltanissetta	262,458	82.6	2.4
Campobasso	221,238	96.5	1.4
Caserta	922,965	92.9	0.9
Catania	1,107,702	83.9	1.3
Catanzaro	358,316	98.6	0.7
Chieti	385,588	87.5	0.8
Como	599,204	98.7	0.6
Cosenza	705,753	94.0	1.3
<b>Cremona</b>	<b>358,955</b>	<b>100.0</b>	<b>0.1</b>
Crotone	174,980	92.7	2.8
Cuneo	587,098	97.0	0.9
Enna	164,788	97.7	1.5
Fermo	173,800	99.5	0.6
Ferrara	345,691	99.1	0.7
Firenze	1,011,349	98.9	0.4
Foggia	622,183	99.7	0.2
Forlì-Cesena	394,627	99.6	0.4
Frosinone	489,083	92.5	1.5
Genova	841,180	97.7	0.4
Gorizia	139,403	100.0	0.0
Grosseto	221,629	97.6	1.2
Imperia	213,840	97.4	1.2
Isernia	84,379	99.1	0.8
L'Aquila	299,031	96.3	1.4
La Spezia	219,556	99.8	0.3
Latina	575,254	96.1	0.9
Lecce	795,134	97.1	0.9
Lecco	337,380	99.5	0.5
Livorno	334,832	99.1	0.7
<b>Lodi</b>	<b>230,198</b>	<b>99.8</b>	<b>0.3</b>
Lucca	387,876	93.6	1.4
Macerata	314,178	99.0	0.6

Mantova	412,292	97.3	0.9
Massa-Carrara	194,878	94.7	2.0
Matera	197,909	97.1	1.5
Messina	626,876	93.9	1.2
Milano	3,250,315	99.8	0.1
Modena	705,393	98.2	0.6
Monza e della Brianza	873,935	98.7	0.6
Napoli	3,084,890	97.2	0.4
Novara	369,018	98.8	0.7
Nuoro	208,550	91.0	2.7
Oristano	157,707	99.6	0.5
Padova	937,908	90.2	1.7
Palermo	1,252,588	96.1	0.6
Parma	451,631	99.7	0.2
Pavia	545,888	99.0	0.5
Perugia	656,382	94.8	0.9
Pesaro e Urbino	358,886	97.5	1.0
Pescara	318,909	98.6	0.8
<b>Piacenza</b>	<b>287,152</b>	<b>100.0</b>	<b>0.0</b>
Pisa	419,037	99.8	0.2
Pistoia	292,473	96.4	1.5
Pordenone	312,533	100.0	0.0
Potenza	364,960	80.8	0.6
Prato	257,716	96.6	1.3
Ragusa	320,893	100.0	0.0
Ravenna	389,456	100.0	0.0
Reggio Calabria	548,009	92.6	1.3
Reggio nell'Emilia	531,891	95.1	1.1
Rieti	155,503	94.8	2.2
Rimini	339,017	97.7	1.2
Roma	4,342,212	94.6	0.4
Rovigo	234,937	100.0	0.0
Salerno	1,098,513	92.8	0.7
Sassari	491,571	100.0	0.0
Savona	276,064	100.0	0.1
Siena	267,197	94.7	1.9
Siracusa	399,224	93.9	0.5
Sondrio	181,095	100.0	0.0
Sud Sardegna	350,725	95.3	1.4
Taranto	576,756	99.4	0.4
Teramo	308,052	99.8	0.3
Terni	225,633	99.9	0.2
Torino	2,259,523	98.1	0.3
Trapani	430,492	94.9	1.3
Trento	541,098	96.3	1.2
Treviso	887,806	97.9	0.7
Trieste	234,493	100.0	0.0
Udine	528,791	90.3	1.3
Valle d'Aosta	125,666	96.9	2.1
Varese	890,768	98.4	0.5
Venezia	853,338	97.3	0.6
Verbano-Cusio-Ossola	158,349	96.4	1.9
Vercelli	170,911	95.0	2.0
Verona	926,497	95.7	0.9
Vibo Valentia	160,073	99.1	0.8
Vicenza	862,418	97.1	0.9
Viterbo	317,030	92.1	2.4

Table 2: Dates when provinces reached 100% to 400% increase in mortality rates

Province	100	Percent increase		
		200	300	400
Bergamo	Feb-18	Feb-25	Mar-04	Mar-14
Lodi	Feb-18	Mar-01	-	-
Cremona	Feb-21	Mar-03	315	-
Piacenza	Feb-26	Mar-13	-	-
Brescia	Feb-27	Mar-11	Mar-27	-
Lecco	Mar-05	Mar-24	-	-

Percent increase refers to the percent increase in the lowess of mortality rates.

Table 3: Region level regression analysis

Panel A: Number of COVID deaths						
	(1)	(2)	(3)	(4)	(5)	(6)
delta_deaths	50.21*** (2.577)	52.12*** (2.059)	49.23*** (2.605)	51.76*** (2.085)	72.28*** (3.054)	63.25*** (2.314)
delta_mr	0.504*** (0.0259)	0.523*** (0.0207)	0.494*** (0.0262)	0.519*** (0.021)	0.731*** (0.031)	0.639*** (0.0235)
diff_deaths	0.413*** (0.0113)	0.403*** (0.00784)	0.395*** (0.0121)	0.402*** (0.00817)	0.479*** (0.00904)	0.415*** (0.00649)
diff_mr	9.146*** (0.65)	9.518*** (0.543)	8.960*** (0.655)	9.405*** (0.548)	12.97*** (0.853)	11.36*** (0.664)
Number of Tests		yes		yes		yes
Fixed Effects			NUTS2	NUTS2	NUTS1	NUTS1
Panel B: Number of COVID positive						
	(1)	(2)	(3)	(4)	(5)	(6)
delta_deaths	278.4*** (11.48)	288.2*** (9.64)	275.0*** (11.55)	285.0*** (9.727)	387.4*** (15.5)	341.2*** (11.67)
delta_mr	2.794*** (0.115)	2.891*** (0.0969)	2.759*** (0.116)	2.860*** (0.0978)	3.923*** (0.157)	3.452*** (0.119)
diff_deaths	2.116*** (0.0491)	2.126*** (0.0346)	2.084*** (0.0499)	2.114*** (0.0354)	2.508*** (0.0454)	2.185*** (0.032)
diff_mr	52.43*** (2.967)	54.38*** (2.607)	51.73*** (2.98)	53.50*** (2.621)	71.27*** (4.354)	63.01*** (3.38)
Number of Tests		yes		yes		yes
Fixed Effects			NUTS2	NUTS2	NUTS1	NUTS1
Panel C: COVID cases (positives, COVID deaths, recovered)						
	(1)	(2)	(3)	(4)	(5)	(6)
delta_deaths	1,845*** (450)	2,085*** (230.9)	1,353*** (458.5)	1,988*** (234.3)	5,406*** (478)	3,584*** (251.7)
delta_mr	18.56*** (4.52)	20.91*** (2.32)	13.63*** (4.606)	19.95*** (2.353)	54.46*** (4.852)	35.86*** (2.56)
diff_deaths	22.90*** (2.383)	15.56*** (1.255)	10.41*** (2.627)	12.33*** (1.337)	39.57*** (1.993)	23.28*** (1.103)
diff_mr	316.1*** (106.4)	376.5*** (55.17)	245.7*** (107.7)	359.9*** (55.73)	912.6*** (122.4)	602.8*** (65.3)
Number of Tests		yes		yes		yes
Fixed Effects			NUTS2	NUTS2	NUTS1	NUTS1

Notes: Standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. NUTS1 corresponds to Groups of regions (Northwest, Northeast, Center, South, Islands), NUTS2 to regions. We run different models with different definitions of the main independent variable. In particular, *delta* stands for percentage difference, *diff* stands for actual difference between 2020 and the average of the years 2015-2019; *deaths* refers to the number of deaths per day; *mr* refers to the mortality rate.

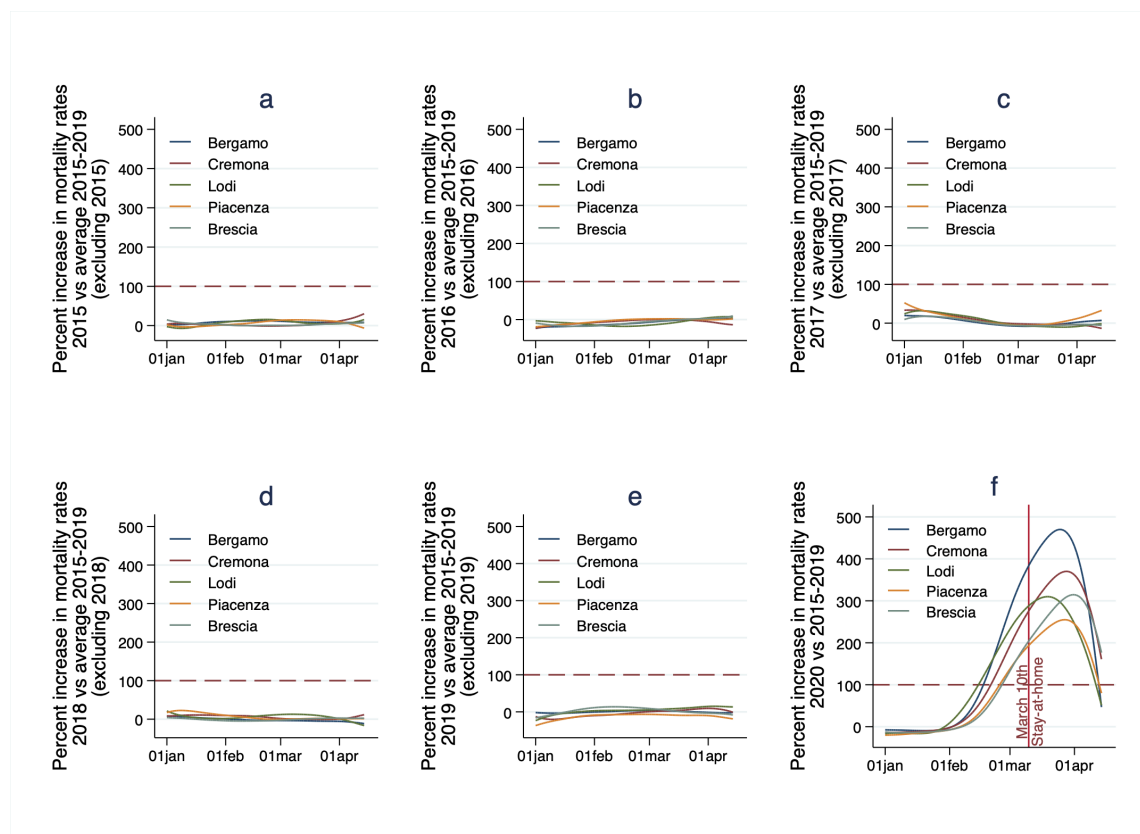
Table 4: Predicting March 2020 increase in mortality rate

	Number of lagged days before Mar 20th									
	1		7		14		21		28	
	Coeff.	<i>z</i>	Coeff.	<i>z</i>	Coeff.	<i>z</i>	Coeff.	<i>z</i>	Coeff.	<i>z</i>
Lagged $\Delta\%$ Mortality Rate	1.02	628.27	1.19	86.93	1.47	39.70	1.90	22.10	2.55	12.34
Bergamo	2.32	11.66	19.52	14.53	46.72	18.04	82.12	20.33	128.71	24.13
Cremona	1.93	13.55	15.65	16.51	36.10	20.01	61.62	21.98	93.68	24.90
Lodi	1.31	8.83	11.52	11.09	28.43	13.44	50.47	14.21	82.54	15.89
Piacenza	1.44	15.10	11.68	18.56	27.09	23.27	46.87	27.52	72.67	38.19
Brescia	1.66	16.85	13.22	20.96	30.05	26.79	51.05	32.25	75.49	41.77
RMSE	0.77		5.44		10.81		15.70		20.11	
R2 Overall	1.00		0.98		0.92		0.86		0.79	

Note: The dependent and lagged independent variables is the lowess of the percentage difference in daily mortality rates at the province level. The Table shows the results only for the five provinces with the largest coefficients. vce(robust). The reference date used is March 20th.

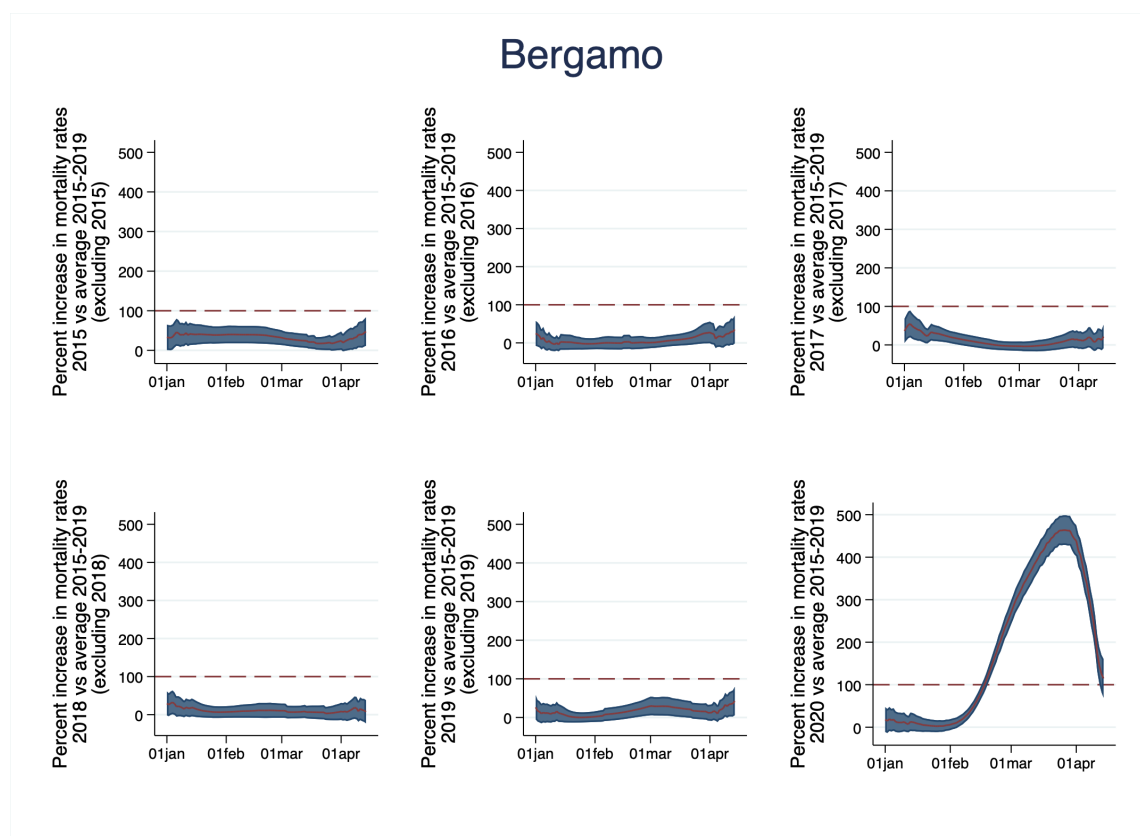
# List of Figures

Figure 1: Growth in Mortality Rate, 2015-2020



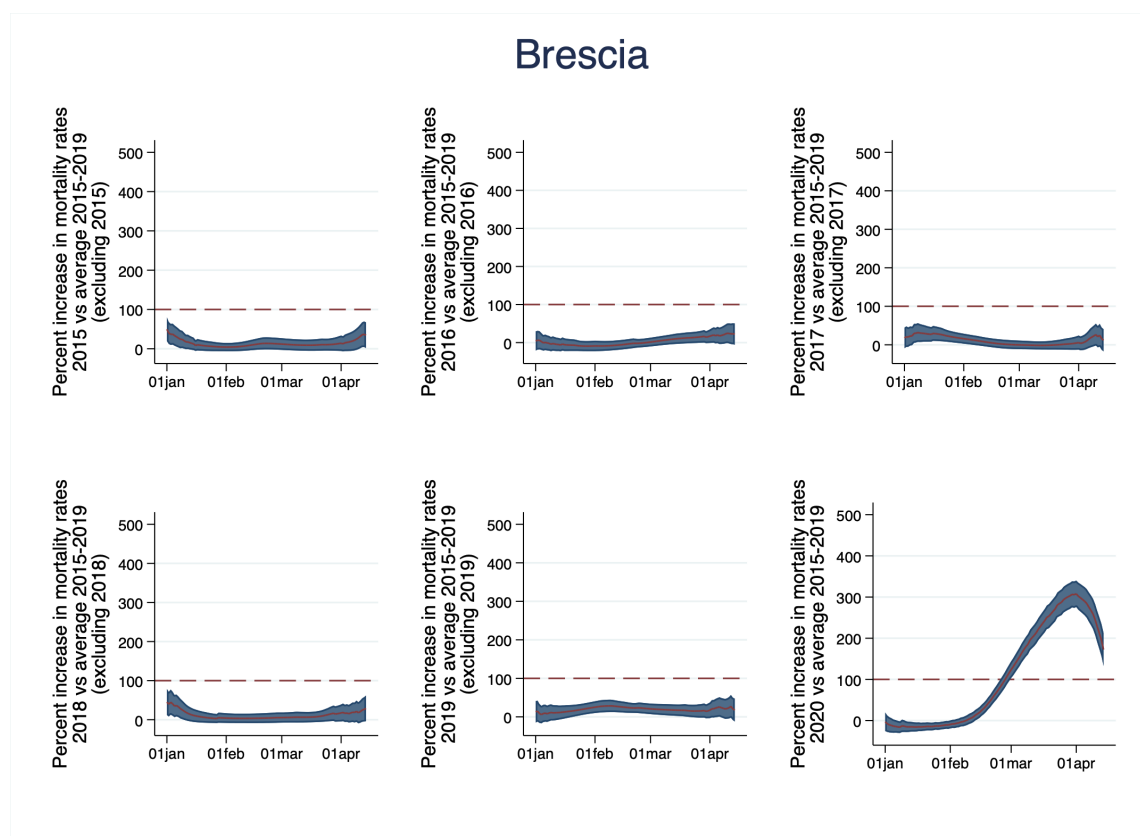
Source: Authors' elaboration on data from ISTAT. Deaths data are from: <https://www.istat.it/en/archivio/240106>; Population data are from: <http://dati.istat.it/>.

Figure 2: Growth in Mortality Rate, 2015-2020 , Bergamo



Source: Authors' elaboration on data from ISTAT. Deaths data are from: <https://www.istat.it/en/archivio/240106>; Population data are from: <http://dati.istat.it/>.

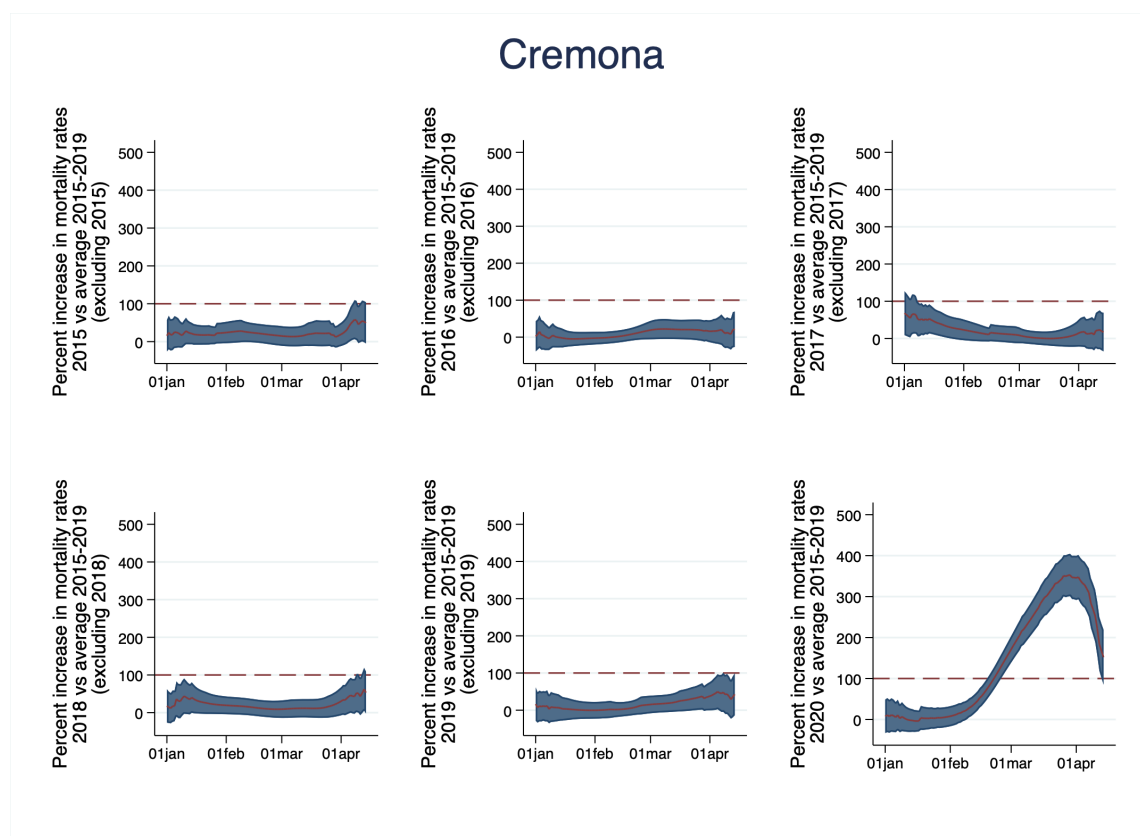
Figure 3: Growth in Mortality Rate, 2015-2020 , Brescia



Source: Authors' elaboration on data from ISTAT. Deaths data are from: <https://www.istat.it/en/archivio/240106>; Population data are from: <http://dati.istat.it/>.

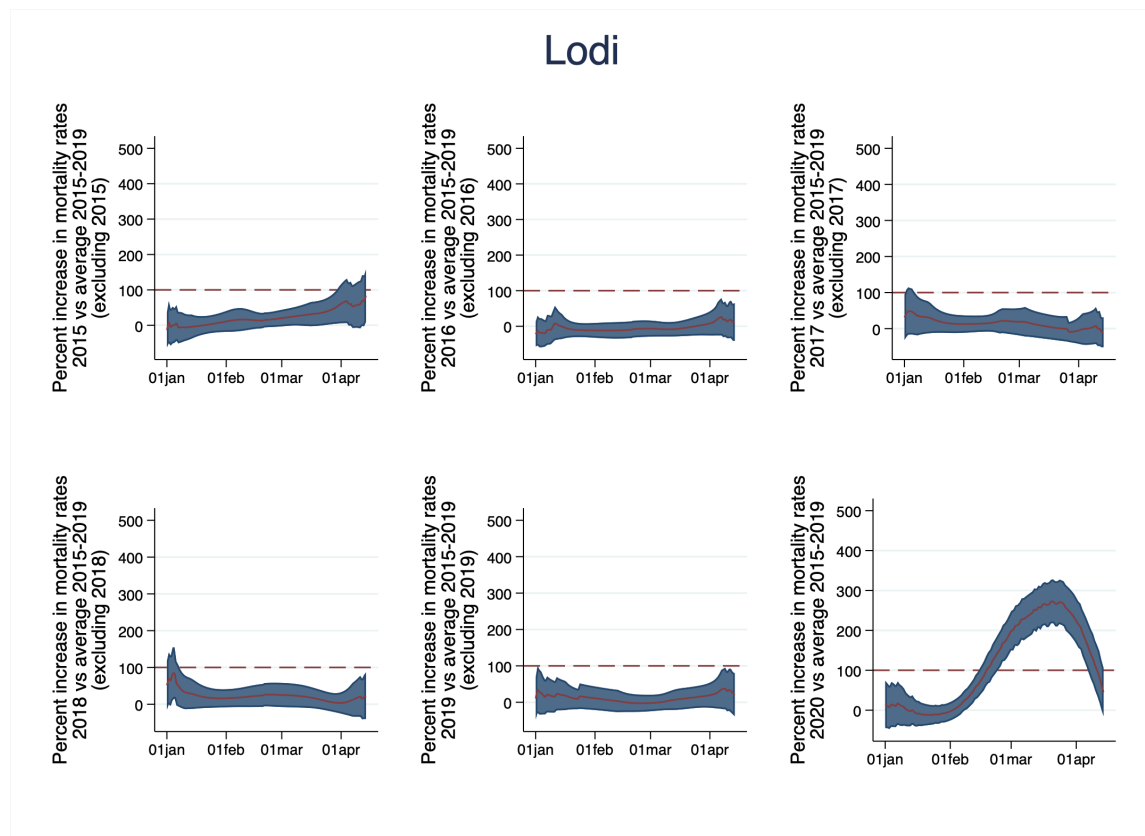


Figure 4: Growth in Mortality Rate, 2015-2020 , Cremona



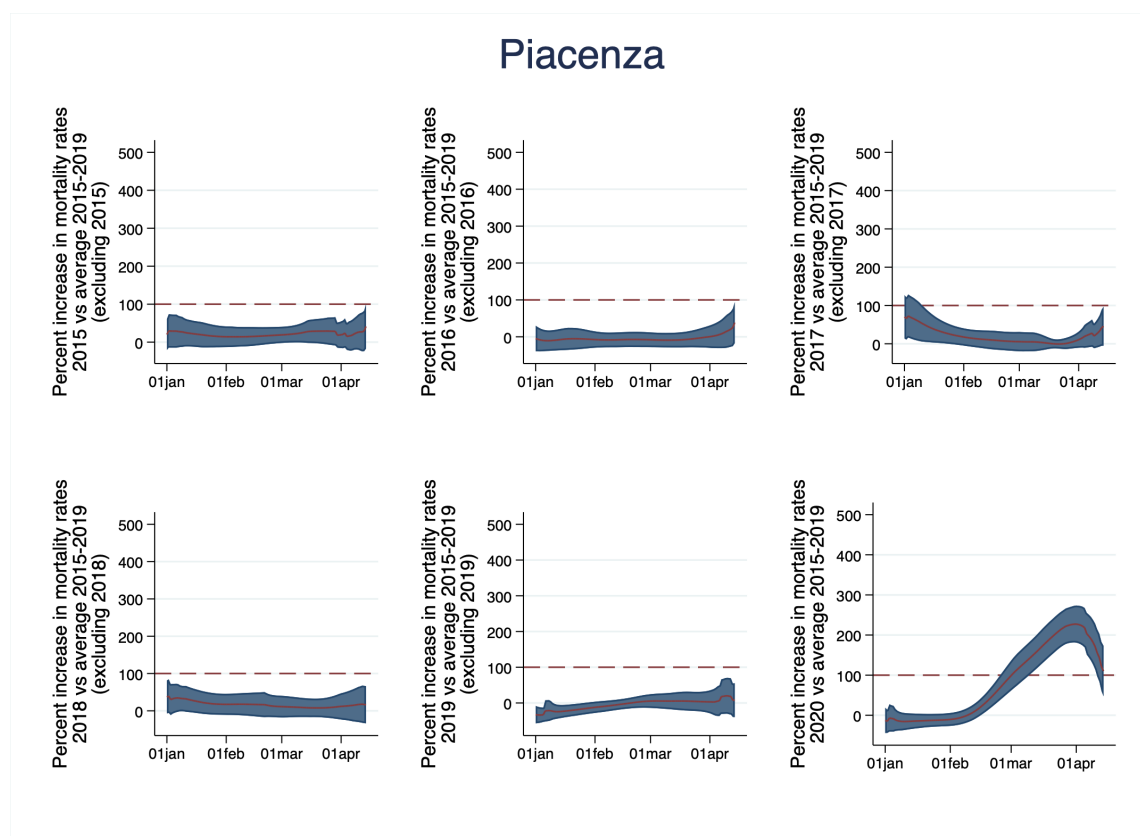
Source: Authors' elaboration on data from ISTAT. Deaths data are from: <https://www.istat.it/en/archivio/240106>; Population data are from: <http://dati.istat.it/>.

Figure 5: Growth in Mortality Rate, 2015-2020 , Lodi



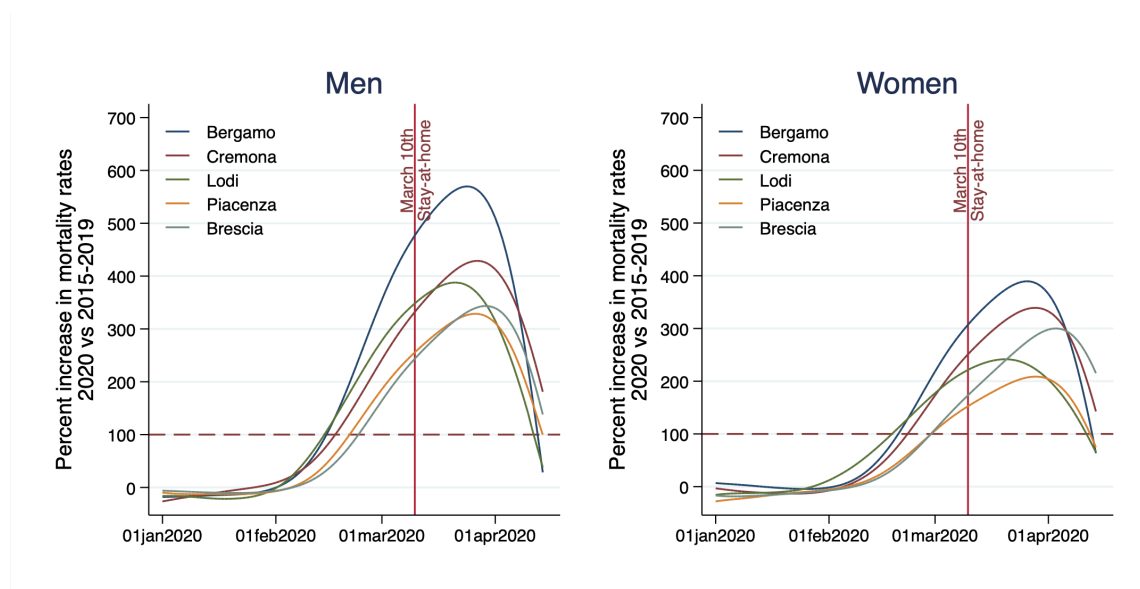
Source: Authors' elaboration on data from ISTAT. Deaths data are from: <https://www.istat.it/en/archivio/240106>; Population data are from: <http://dati.istat.it/>.

Figure 6: Growth in Mortality Rate, 2015-2020 , Piacenza



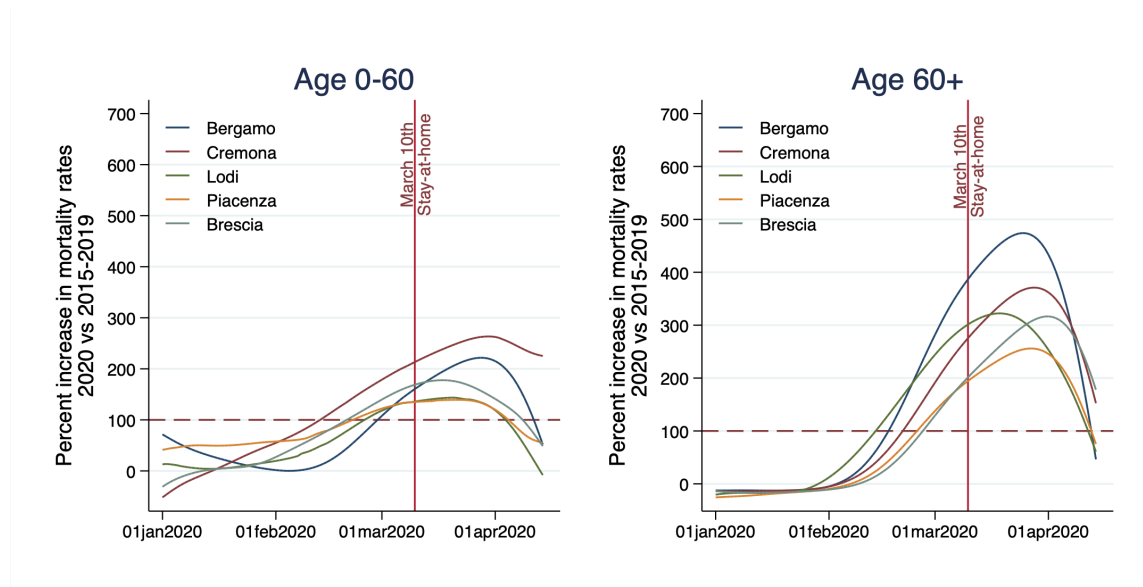
Source: Authors' elaboration on data from ISTAT. Deaths data are from: <https://www.istat.it/en/archivio/240106>; Population data are from: <http://dati.istat.it/>.

Figure 7: Growth in Mortality Rate 2020/2015-2019 - Gender



Source: Authors' elaboration on data from ISTAT. Deaths data are from: <https://www.istat.it/en/archivio/240106>; Population data are from: <http://dati.istat.it/>.

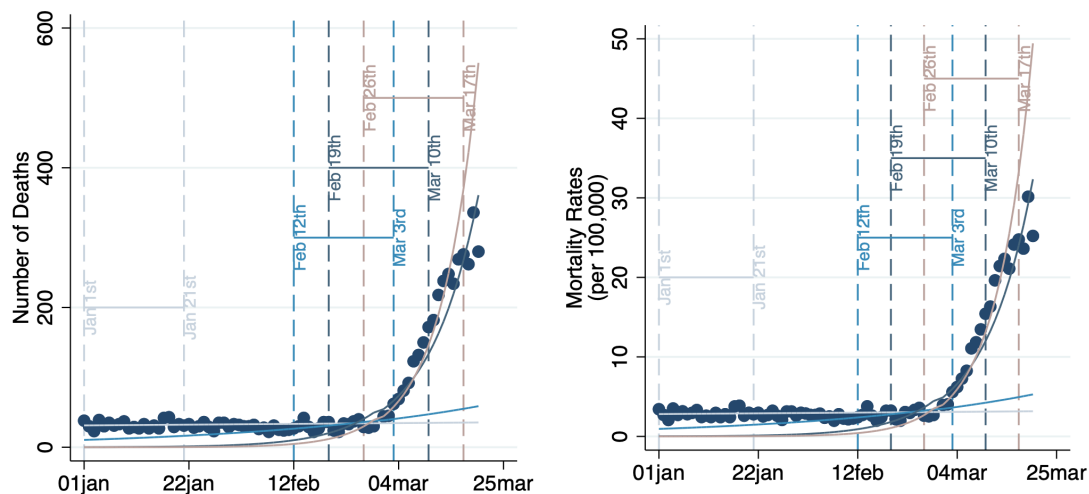
Figure 8: Growth in Mortality Rate 2020/2015-2019 - Age Groups



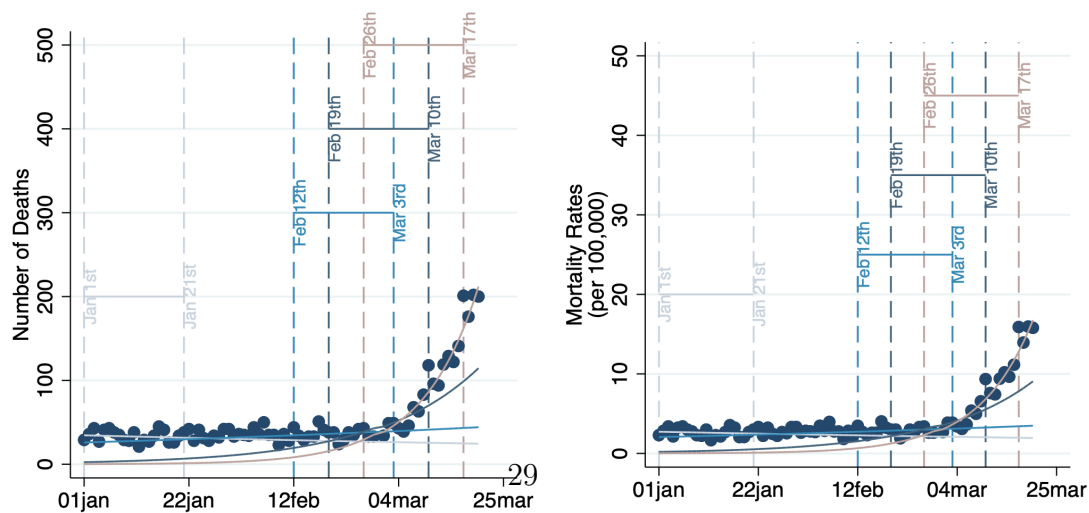
Source: Authors' elaboration on data from ISTAT. Deaths data are from: <https://www.istat.it/en/archivio/240106>; Population data are from: <http://dati.istat.it/>.

Figure 9: Prediction

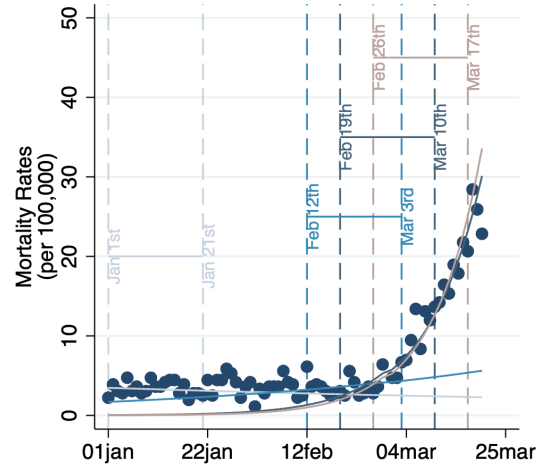
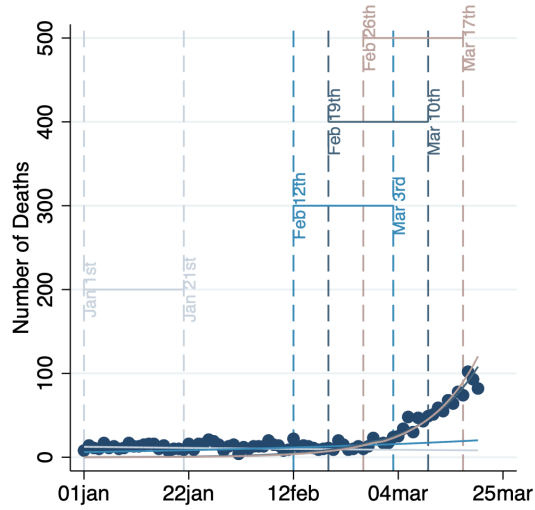
Bergamo



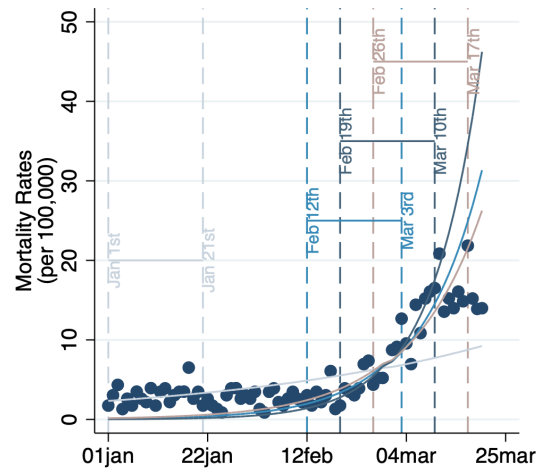
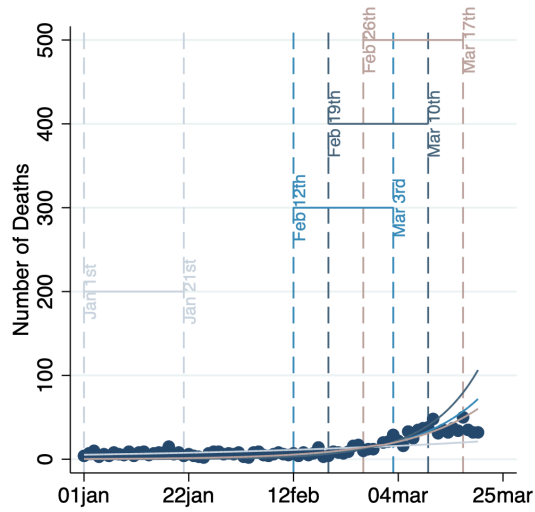
Brescia



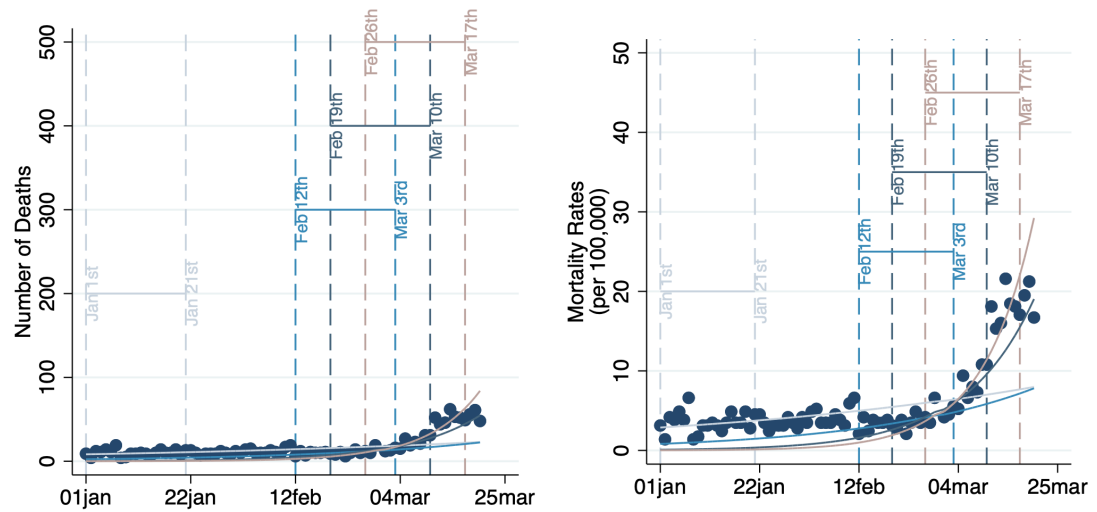
## Cremona



## Lodi



## Piacenza



Note: Dots correspond to observed data and lines correspond to forecasted trends using the data at different times spells (identified with different colors). Source: Authors' elaboration on data from ISTAT. Deaths data are from: <https://www.istat.it/en/archivio/240106>; Population data are from: <http://dati.istat.it/>.