Understanding the heterogeneity of adverse COVID-19 outcomes: *the role of poor quality of air and lockdown decisions*

Leonardo Becchetti[*], *University of Rome Tor Vergata*
Gianluigi Conzo, *University of Rome Tor Vergata*
Pierluigi Conzo, *University of Turin & Collegio Carlo Alberto*
Francesco Salustri, *University of Oxford*

First version: *April 10, 2020*
This version: *May 5, 2020*[⊥]

Abstract
The uneven geographical distribution of the novel coronavirus epidemic (COVID-19) in Italy is a puzzle given the intense flow of movements among the different geographical areas before lockdown decisions. To shed light on it, we test the effect of five potential correlates of daily adverse COVID-19 outcomes at province level, that is lockdown decisions, demographic structure, economic activity, temperature and quality of air (proxied by particulate matter and nitrogen dioxide). We find that poor quality of air is negatively correlated with adverse outcomes of the epidemic, with lockdown being strongly significant and more effective in reducing deceases in more polluted areas. Results are robust to different methods including cross-section, pooled and fixed-effect panel regressions (also controlling for spatial correlation), as well as difference-in-differences estimates of lockdown decisions through predicted counterfactual trends. They are consistent with the consolidated body of literature in previous medical studies suggesting that poor quality of air creates chronic exposure to adverse outcomes from respiratory diseases. The heterogeneity of diffusion does not seem to depend on other controls such as temperature, commuting flows, quality of regional health systems, share of public transport users, population density and the presence of Chinese community. We also find that adverse COVID-19 outcomes are significantly and positively correlated with the share of small (artisan) firms. Our findings provide suggestions for investigating uneven geographical distribution patterns in other countries, and, if preliminary evidence is corroborated by causation links, have relevant implications with respect to environmental and lockdown policies.

Keywords: COVID-19 pandemic, particulate matter, lockdown, economic activity.
JEL numbers: I18, Q53, J18, H12

[*] Address for correspondence: Department of Economics and Finance, University of Rome Tor Vergata, 00135 Rome, Italy; email: becchetti@economia.uniroma2.it

[⊥]Updates available here: https://www.dropbox.com/sh/ipcnhj06u5y5q5q/AADjIJXVpaE3_OnqKAHlZwLba?dl=0

1. Introduction

Viruses do not travel alone. They take human beings as means of transport. For this reason, the heterogeneity of the diffusion of the novel coronavirus (SARS-CoV-2, thereafter coronavirus) in Italy is puzzling. As is well known contagions and deaths in Italy are disproportionately concentrated in some provinces of a single region (Lombardia) and, more in general, in the North of Italy (Piemonte and Emilia-Romagna).[1] Several authors emphasize that the coronavirus has been circulating at least since early January and well before late February, when the first cases were detected (e.g. Zehender et al. 2020).[2] The month before the country lockdown, when the government limited people movement around the country[3], the flow of commuting between Rome and Milan has been intense, as it has always been in these last years with flight and especially high-speed train connections allowing to move from one city to another in slightly less than 3 hours. If the virus easily jumped from the remote Wuhan to Milan, why did it not across a much shorter distance, i.e. that between Milan and Rome or, more in general, between the North and the South of Italy[4,5]?

An interesting research question is therefore why the intensity of the epidemic (hereon COVID-19) has been so different between the two cities, and in general between different Italian provinces. A first tentative answer is that the virus was not spreading in Rome or in the Center-South before the government restrictions. Those restrictions were therefore crucial to limit the epidemic in these areas, although the anecdotal evidence reported above casts some doubts about this first hypothesis. The second tentative answer is that the virus travelled way before the lockdown, and some concurring factors like pollution, weather conditions, or less intense economic activity made it weaker in areas different from the "epicenter".

Our paper aims to shed light on this puzzle by investigating the relative role of five main factors that might explain the spread of epidemic in Italy, that is lockdowns, demographic structure, climate, pollution, and economic activity. The focus of our research has relevant implications on

[1] As of April 7th 2020, Lombardia accounted for 38.6 percent of reported contagions and 58.6 percent of registered COVID-19 deaths.

[2] The authors show that epidemiological tracing, based on phylogenetic analysis of the first three complete genomes of SARS-CoV-2 isolated from three patients involved in the first outbreak of COVID-19 in Lombardy, provided evidence that SARS-CoV-2 was present in Italy weeks before the first reported cases of infection.

[3] The decree on the full restriction of movement among regions was enacted only on 11th March 2020. The information spillover before the decree was operating led to mass escape from Milan train station toward Southern Italy the days before (see https://www.milanopost.info/2020/03/09/234982/).

[4] On 31st January 2020 a couple of Chinese tourists who had spent some days in Milan, Parma and Rome (since 28th of January) was recovered in serious conditions at the Spallanzani hospital in Rome (https://www.ilgiornale.it/news/roma/allarme-albergo-mio-marito-ha-febbre-1819431.html).

[5]The flow of passengers moving from Rome to Milan (airplane plus train) was around 5.14 millions in 2018. 3.6 millions by train, which is the 70% of the total passengers, the others are by plane (20%) and by car (10%). Sources: https://www.fsitaliane.it/content/dam/fsitaliane/Documents/fsnews/comunicati-stampa/2019/dicembre/2019_12_05_NS_2_FS_Italiane_10_anni_AV_cambiato_Paese_e_vita_persone.pdf; https://www.ilsole24ore.com/art/roma-milano-7-passeggeri-10-scelgono-treno-AEOuGl5.

several dimensions such as subjective wellbeing, health policies, economic conditions and economic policies, not ultimately since – as of May, 1st 2020 – the epidemic in Italy caused 28,236 "official" deaths, stressed the national health system and produced a paralysis of economic activity[6].

Our empirical approach rests on a multivariate analysis which aims to add original insights from at least three points of view. First, assessing the relative strength of different concurring factors is fundamental to understand the heterogeneous evolution of the epidemic across the country. This approach is a necessary complement to deterministic models, in which nonlinear dynamics of the diffusion emerge as a unique driver. Second, lockdown decisions have highlighted the trade-off between health and economic development goals.

Our findings suggest that lockdown may mitigate contagions and mortality, especially for the more polluted areas. Conversely, poor quality of air (under both dimensions of fine particulate matter and density of urban green) and the share of small business activity are positively correlated with both outcomes. The role of the micro business activity can be interpreted in the light of the higher fragility of this kind of business to lockdown decisions, of the higher share of manufacturing activity not easily converted to smart work among micro and artisan firms and therefore, presumably, of the higher resistance to the decision to stop economic activity. The positive effect is, however, also consistent with the hypothesis that high economic activity, especially for small businesses, embeds a large volume of human interactions, and hence increases contagion. Finally, the heterogeneity of diffusion does not seem to depend on other pre-virus factors that we test, i.e. commuting and public transport use, health system efficiency, density and the share of Chinese immigrants.

The paper is divided into seven sections. In the second section we present our research hypotheses and the related literature. In the third section we illustrate data and econometric model. In the fourth we present descriptive and econometric findings. In the fifth section, using pre-lockdown data we build a counterfactual trend and use a difference-in-differences approach to evaluate the impact of lockdown, and its interaction with pollution, at municipality level. In the sixth section we discuss our results (limits, policy implications and directions for future research). The seventh section concludes.

2. Background and research hypotheses

The first hypothesis we test is that *the lockdown measures proved effective in limiting deceases and contagion* (H_1). Human mobility restrictions are considered among the most effective policies to

[6] According to the National Institute of Statistics preliminary estimates, the Italian GDP fell by 4,7% in the first quarter causing around 8million temporary layoffs (https://www.istat.it/it/archivio/242084).

reduce contagion in absence of a vaccine, but their economic costs are huge (Bajardi et al., 2011; Wang and Taylor, 2016; Charu et al., 2017). Fang et al. (2020) calculate that contagion cases would be 64.81% higher in the 347 Chinese cities outside Hubei province, and 52.64% higher in the 16 non-Wuhan cities inside Hubei, without the Wuhan lockdown. The coronavirus mean incubation period, defined as the time from infection to illness onset, has been estimated at 5.2 days (4.1-7.0), with the 95th percentile of the distribution at 12.5 days (Li et al., 2020). Moreover, the majority of people were tested with severe symptoms only (as International guidelines suggested) and with some delay with respect to the day in which the test was recorded (3.6 days according to Cereda et al. 2020). We therefore expect that governmental restrictions reducing the flow of human interactions and imposing physical distance among people have impact, which may be distributed over around 17 days. Thus, we test the effect of the different national, regional, and provincial measures enacted in Italy in the months of coronavirus outbreak. Table 1 lists the restrictions adopted at different governmental levels.

[Table 1 here]

The second hypothesis we test is that *(historical levels of) particulate matter has a positive and significant role in explaining the geographic variation of the epidemic (H_2).* There are two hypotheses on pollution as a pull factor of COVID-19. The first is that individuals living in highly polluted areas have weaker lungs and reduced capacity to react to respiratory diseases and/or pneumonias and, therefore, also to COVID-19. The second is that particular matter is a carrier of the virus slowing down its falls from the air (Setti, 2020). The rationale for the first hypothesis is that lung reaction to pneumonia depends on the pulmonary surfactant (a surface-active lipoprotein complex formed by type II alveolar cells). The pulmonary surfactant contributes with minimal diffusion distance and large surface area to the optimal exchange of gases. In essence, a healthy surfactant protects lung collapse at low volumes and tissue damage at high volume levels and allows lungs to inflate much more easily, thereby reducing the work of breathing. Pollution and heavy smoke produce abnormalities in surfactant composition thereby making ventilation more problematic and reducing lung "efficiency" (Pastva et al. 2007).

The hypothesis has been tested and not rejected by a large body of literature finding correlations between pollution and pneumonia not only for the children but also for the elders. Binod et al. (2010) find that PM2.5 is significantly associated with hospitalization for pneumonia in Canada, Medina-Ramon et al. (2006) find that PM10 is associated with hospitalization for respiratory diseases in 36 US cities. Zhang et al. (2016) obtain similar results in a Chinese sample, while Zanobetti and Schwartz (2006) in Boston. Luginaah et al. (2005) report significant correlation between (PM10 and

PM2.5), NO2, SO2, and disease exacerbations, emergency admissions, hospitalizations and mortality in Ontario.

Some of this research has been conducted before the coronavirus outbreak in the areas where the epidemic has been more severe. Zhang et al. (2015) find that local PM2.5 has an acute adverse effect on lung function in young healthy adults in Wuhan, with temperature also playing an important role. Santus et al. (2012) find an acute effect of CO, SO2 and PM10 on Emergency Rate Admissions for pneumonia in Milan at short daily lags. Zeng et al. (2016) find that smaller particles have been shown to have stronger effects on multiple respiratory diseases and increased hospitalization rates than larger ones; their sedimentation speed is, indeed, lower and exposition to them higher for the human body. Larger particles are filtered by nostrils while smaller ones can reach alveolar cells (Zeng et al., 2016). Pope et al. (2016) resume findings from this literature in their survey on more than 500 past studies arguing that the body of evidence on the nexus between particulate matter and respiratory and pulmonary diseases is stronger if we look at long run exposure.

A few very recent working papers focus directly on the relationship between pollution and COVID-19 disease. Wu et al. (2020) find that long-term average exposure to fine particulate matter (PM2.5) increases the risk of COVID-19 deaths in the United States (in terms of economic magnitude they find that an increase of 1 $\mu g/m^3$ in PM2.5 is associated with a 15% increase in the COVID-19 death rate). Conticini, Frediani and Caro (2020) argue that pollution can be a co-determinant of the abnormal number of deaths registered in Lombardia and Emilia Romagna. The authors emphasize how the composed air quality index including five pollutants (PM10, PM2.5, O3, SO2 and NO2.9 show that Lombardia and Emilia Romagna are the most polluted in Italy and among the most polluted in Europe). The authors provide medical details on how poor air quality leads to inflammation, eventually leading to an innate immune system hyper-activation which has been observed in COVID-19 patients. They also report how particulate matter (PM2.5 and PM10) can lead to systemic inflammation consisting of an overexpression of PDGF, VEGF, TNFa, IL-1 and IL-6 which can arise even in healthy, non-smoker and young subjects (Pope et al., 2016). The effect is directly related to the length of pollutant exposure (Tsai et al., 2019). They conclude that the elderly who live in the regions with higher intensity of particulate suffered from chronic exposure to air pollution and have higher probability of being affected by virus invasion due to the weakened upper airways defenses.

Figure 1 shows the geographical distribution of COVID-19 related outcomes and of average levels of PM10 and PM2.5 in Italy. Indeed, the cumulative number of positive cases and deaths per 1000 inhabitants as measured in April 17th 2020 and April 4th 2020 respectively tend to concentrate in provinces that witnessed high levels of pollution in 2018, i.e. those in the North of Italy.

[Figure 1 here]

The third hypothesis we test is that *temperature has a significant role in explaining the geographic variation of the epidemic (H_3).* Several studies show that virus outbreaks are significantly reduced by high temperature (Lowen et al. 2007; Barreca et al. 2012; Shaman et al 2009, Zuk et al. 2009). Research on past coronaviruses show that they belong to the family of "enveloped viruses" as they are surrounded by an oily coat (a lipid bilayer). Enveloped viruses are more sensitive to temperature since low temperature hardens the coat into a rubber-like state that protects the virus longer when it outside the body. Sayadi et al. (2020) show that areas at higher risk of coronavirus outbreak are those with an average temperature between 5 and 11 C degrees. Bannister-Tyrrel et al. (2020) provide preliminary evidence that higher temperature is associated to lower incidence of COVID-19. Notari (2020) by looking at nonsynchronous data from 42 countries finds a peak of the growth rate of contagion around 7.7 +/- 3.6 C temperature. Bukhari and Jameel (2020) show that 90 percent of cases until March 22, 2020 have been recorded in the 3-17C temperature range and in the 3-9g/m^3 humidity range. The authors emphasize how virus diffusion in warmer and more humid areas (regions of the United states such as Texas, New Mexico and Arizona, Asian countries such as Malaysia and Thailand and Middle East countries such as Saudi Arabia), while stronger in others with colder and drier climate (Iran, South Korea, New York and Washington). A similar argument would apply to differences of contagions between North and South of Italy and between Madrid (who has a colder and drier continental climate) and other regions of Spain in the South or closer to the sea. Descriptive evidence on COVID-19 spread in the African region is not at odds with this hypothesis. At 27th March the country with the highest number of cases is a country with Mediterranean climate, South Africa (939), followed by country with dry climate such as Algeria, while in most other countries there are only a few cases of imported transmission (OMS, 2020).

The fourth hypothesis we test is that *the pre-virus levels of intensity of local small business activity is positively associated with COVID-19 outcomes (H_4).* Small business employers and entrepreneurs live in a competitive environment with reduced social protection in Italy. In most cases they are suppliers of large companies in relationships where they have lower bargaining power that translates into worse trade credit conditions. Moreover, micro and artisan firms are in a higher proportion in the manufacturing sector, with reduced opportunity to convert their activities to smart working. Our assumption is that small business had a relatively lower propensity to stop operation during the epidemic for the expected higher risk of adverse economic consequences from that decision.

Finally, we also test the relative role of other pre-virus factors that might be associated with the COVID-19 outbreak and with its outcomes. We look at human mobility and density since these factors increase the chances of social interaction and hence the spread of the virus. In addition, we account for the heterogeneous efficiency of the local health system across Italian provinces. We also control for the demographic structure of the virus by including in the multivariate analysis also the share of residents aged over 65, because this age group has been shown to be more vulnerable to the virus. Finally, we also test the role of the Chinese community in Italy, since its presence could capture some of the socio-economic exchanges between Italy and China before the outbreak of the virus. It has also to be noted that the Chinese presence in Italy has been connected to the spread of the virus in Italian provinces by anti-immigration supporters. Moreover, Chinese people residing in Italy have been frequently witnessed discrimination during the first days of the COVID-19 outbreak, under the form of physical and verbal violence.

3. Data and econometric model

Our database includes two dependent variables related to outcomes of the coronavirus disease and regressors including province time invariant characteristics, national or regional restriction events and time varying variables related to temperature. As dependent variables we consider the daily number of deaths (from the Italian National Statistics Institute, ISTAT) and new positive cases of COVID-19 at province level (from the Italian Civil Protection, ICP) per 1,000 inhabitants.

The number of deaths is the daily number of deceases in municipalities that register an increase of at least 20% with respect to daily average number of deaths occurred over the last 5 years. Deaths here are considered as "Cancellati dall'Anagrafe per Decesso", i.e. cancelled from the official census records because of death. We use the daily number of deaths divided by the total population of the municipality (in 2018), from February 24th 2020 to April 4th 2020 (last date for which data are available at the time our main estimates were implemented), averaged at province level.

The second dependent variable is the number of new daily confirmed COVID-19 cases, that is the number of new infected patients detected each day. We use this measure instead of the number of *net* infected patients, where deaths and recoveries of the day are subtracted from the gross value, because ICP does not provide the breakdown of infected patients at the province level (i.e. our main unit of analysis). Notwithstanding possible measurement errors that make the accounting more or less conservative (e.g. due to the region-specific testing capacity), one advantage of our research is that we limit the analysis to the Italian case and therefore we avoid measurement bias arising from the different approaches followed in different countries. Since the accounting of positive cases is not voluntarily provided by provinces, but officially managed by a national institution and its local

branches (ICP), we therefore expect that this measurement bias might not affect our estimates. We consider the daily change in the number of positive cases from February 24th 2020 to April 17th 2020.

The fully-augmented model we consider is detailed in the following equation:

$$CV19\text{-}Outcome_{it} = \alpha_0 + \alpha_1\, Pollution_i + \alpha_2 UrbanGreen_i + \alpha_3 HighTemperature_{it} + \alpha_4 Artisan_i + \alpha_5 Density_i + \alpha_6 Income_i + \alpha_7 Over65_i + \alpha_8 Health_{pca_i} + \alpha_9 InternalCommuting_i + \alpha_{10} ExternalCommuting_i + \alpha_{11} PublicTransportUse_i + u_{it} \quad \text{Eq.1}$$

where CV19-Outcome is, in turn, the number of contagions over local population (*cases_pc*) or, alternatively, the number of deceases over local population (*deaths_pc*), both per 1,000 inhabitants in province *i,* and day *t*.

To test our research hypothesis on the impact of exposure to pollution before the outbreak of the virus, we use PM measures, that is, alternatively, PM2,5 and PM10, which are the two fine particulate matter variables measuring average values in μg/m^3 registered by environmental monitoring units at province level in 2018. In alternative models, we also test the role of nitrogen dioxide (NO2) in μg/m^3 as registered by environmental monitoring units at province level in 2018. Pollution variables are introduced as time-invariant local characteristics based on the research hypothesis H_2 arguing that the variable affecting lung weakness is pollution history, and not the current level of pollution[7]. As an additional proxy for quality of air, we also use *UrbanGreen*, i.e. square meters of green per 100 m^2 surface of urban centers in the province main city.

In order to test research hypothesis H_3, we introduce temperature in the specification with a dummy taking value one if the three days moving average of minimum temperature was higher than 12°C ($HighTemperature$), considering the 3-days lag of the variable to take into account the time

[7] The EU identifies as dangerous for lung diseases particulate matter (PM2.5, PM10), sulphur dioxide (Sox) and Ozone dioxide (NOx) (see: https://op.europa.eu/webpub/eca/special-reports/air-quality-23-2018/it/). The levels of these polluters in the air depends on the combination of activity of emission sources (house heating, transportation, sources of energy, manufacturing and agricultural activity) with weather and geographical conditions (i.e. Pianura Padana has for its geographical structure lower air circulation). The importance of polluting sources also varies significantly. When averaged at EU level house heating is by far the most important source for PM (42 percent for PM10 and 57% for PM2.5), while transportation (39 percent) and energy sources (31 percent) for sulphur dioxide. However, these shares may vary significantly in different regions. In the North of Italy house heating accounts for 37% while 17% from intensive farming (ISPRA, 2020).

between a possible effect and the illness onset[8]. Hypothesis H_4 is tested by introducing a variable measuring the share of artisan firms at province level (*Artisan*).

To measure the efficiency of the local health system ($Health_{pca}$), we extract the first factor from a principal component analysis which includes the number of many different medical devices (i.e. the number of lung ventilators, diving chambers, ecographs, computed tomography scanners, hemodyalisis machines, medical monitors, nuclear magnetic resonance tomographs, operating rooms, radiology devices, portable radiology devices, linear particle accelerators, remote control radiology tables, immune-based automatic analysers, computerized gamma cameras, anesthetic machines, surgical lighthead, automatic coulter counters) per 1,000 inhabitants.

As additional controls we use population density, average household disposable income and the share of individuals aged over 65 (*Over65*), both per 1,000 inhabitants. In an alternative specification, we also include the share of Chinese residents to the total number of immigrants at the province level.

Another important proxy for contagion power concerns the speed and the amount of individual movements. We therefore include among controls a measure of internal commuting flow (*InternalCommuting*), which is calculated with Census data movement within province *i*, as well as a measure of imported commuting flow (*ExternalCommuting*) with Census data movements into province *i* from other provinces -*i* (both variables are computed per 1,000 inhabitants). We also include another proxy for the frequency of human contacts, i.e. the number of passengers on public transport divided by the total number of residents in the province (*PublicTransportUse*) and multiplied by 1,000.

Standard errors are clustered at regional level in order to account for error correlation within the region where our unit of observation (province) is located. Further details on the construction of all the variables are in Table 2.

[Table 2 here]

4. Empirical findings

Summary descriptive findings of the variables used in our specification are presented in Table 3.

[Table 3 here]

[8] Results are not significantly affected either by the days of the moving average or by the lagged days considered. Results from a further investigation on delayed effects of temperature on the virus will be published in a new version of this working paper.

The first model we estimate implements maximum likelihood estimators for the parameters of a linear cross-section spatial-autoregressive model with spatial-autoregressive disturbances (SAC)[9]. More specifically, we estimate the following equation:

$$CV19\text{-}Outcome_i = \alpha_0 + \lambda \sum_{i \neq j} w_{ij}\, CV19\text{-}Outcome_i + \sum_r \alpha_r X_i + u_i$$

Eq.2

Where the dependent variable is the cumulative (contagion or death) outcome at a given date, *X* are the controls described in Eq.1, and w_{ij} coefficients are the inverse distance spatial-weighting elements using province latitude and longitude, for each of the *n* provinces *i* and *j*; u_i are modelled as $u_i = \rho \sum_{i \neq j} w_i + \varepsilon_i$.

SAC cross-sections estimates take a snapshot of the phenomenon at the beginning and at the end of our sample period and using as dependent variable the cumulative number of contagions (Table 4) and deaths (Table 5).

[Tables 4-5 here]

Results show that economic activity and quality of air are significantly (and consistently across model specifications) correlated with the COVID-19 outcomes. More specifically, provinces with high levels of PM10 (Tables 4-5, Column 1) or PM2.5 (Tables 4-5, Column 2), as well as with high economic activity tend to have also worse outcomes in terms of contagion and deceases; density of urban green is significantly and negatively correlated with mortality. Results for NO2 are similar to those for PM, though with higher p-values in the estimates with contagion as dependent variable. In terms of magnitude, the PM10 coefficient imply that moving from the provinces with the lowest to the highest PM10 level implies an increase of 607 new cases per 1 million inhabitants as of 5th March (3164 as of 14th April). The same numbers are 571,7 and 4140 for PM2.5. In terms of deaths the move from the province with the lowest to the highest PM level implies, based on our cross-section coefficient a difference of 205 deceases per 1 million inhabitants as of 5th March (1012 as of 14th April). The same numbers are 194 and 953 for PM2.5.

[9] All results are confirmed also in less strict OLS models that do not account for spatial autocorrelation (available upon request).

In order to exploit the time dimension of the data, we first perform a pooled OLS estimate including also the time trend (and its square). The dependent variable is now the daily number of new cases or deaths alternatively. Regressors include a linear and a quadratic time trend (*Day* and Day^2). Findings (Table 6) reveal that the COVID-19 outcomes follow an inverse U-shape exponential dynamic (the Day^2 variable is negative and significant). As in the previous estimates, other significant variables are exposure to particulate matter and the share of artisan firms. In some specifications, the share of over-65 individuals is negatively correlated with contagion, yet not with mortality. This could be explained by the fact that this age class might have responded more quickly to the restrictions and/or by the advices provided by central and local authorities. In terms of magnitude, coefficients from pooled estimates are broadly consistent with those from cross-section estimates implying a difference from the highest to the lowest PM province of 2940 contagions and 1293 deaths per month per million inhabitants for PM10, and a difference of 3160 contagions and 1290 deaths per month per million inhabitants for PM2.5.[10]

[Table 6 here]

The second panel-data approach rests on SAC fixed-effect model. In these models, province time-invariant characteristics are absorbed in the intercept. This is, however, an important feature since it allows us to partial out heterogeneous omitted factors such as industry characteristics or structural differences in regional health policies (i.e. prevalence of elders in nursing homes) that might be correlated with the dynamics of contagion and mortality. We therefore test whether the above-mentioned pre-virus province characteristics differentially affect the *trend* of contagion and mortality in our provinces. To this purpose, we interact the time variable (*day*) with each time-invariant control included in Eq.1.

[Table 7 here]

Results are reported in Table 7. The role of pollution (PM10, PM2.5 or NO2) is confirmed also under this stricter analysis. More specifically, this interaction captures separately the "slope" effect on COVID-19 outcomes of the variables proxying for the historical levels of the quality of air; in other terms, it measures the differential trends of contagion and mortality by levels of pollution. Note that

[10] Our model outperforms the purely autoregressive model that include among regressors only time trends ($R^2 = 0.105$ vs $R^2 = 0.271$ for PM10, $R^2 = 0.265$ for PM2.5, and $R^2 = 0.277$ for NO2). The differences in the goodness of fit and the significance of our regressors are similar when we consider a cubic trend model, which captures the convexity of the initial increasing trend. Results are available upon request.

the average effects are absorbed into the intercept and not identifiable in this kind of estimates. Results suggest that – net of all other province time-invariant factors – contagion and mortality tend to grow more rapidly in provinces that were highly polluted before the outbreak of COVID-19.

Overall, our empirical findings show a negative and significant correlation between pollution and both the COVID-19 outcomes under scrutiny. Moreover, the share of artisan firms has a positive and significant effect on both dependent variables. As argued above, our interpretation to this result (consistent with anecdotal evidence[11]) is that micro-firms are the most fragile part of the productive environment and therefore less likely to stop down after the beginning of the epidemic to avoid the risk of default. Moreover, a higher proportion of them operates in the manufacturing industry, and have relatively lower chances to shift to smart working during the epidemic. We cannot however exclude that the positive and significant coefficient of the artisan variable conceals the effect of different dynamics of human interactions at province level during the estimation period, which are typical of areas with higher economic activity, and therefore correlated with the spread of the virus.

As a final test, we assess the relative role of the presence of the Chinese community in the spread of the disease. In the pooled regressions of contagion (Table 6), we also include among regressors the share of Chinese immigrants to total population at the province level. The coefficient is negative and not statistically significant, with $\beta = -0.0659$ and $p = 0.544$ in the estimate with PM10, $\beta = -0.698$ and $p = 0.597$ in the estimate with PM2.5, and $= -0.0491$ and $p = 0.566$ in the estimate with NO2 (available upon request).

5. Lockdown and mortality

In this section, we take into account the correlation between lockdown decisions and COVID-19 related outcomes (research hypothesis H_1). It has to be noticed here that correlation does not imply causation, yet a perfectly randomized experiment is not feasible in the present context. A second-best approach rests on a reasonable approximation to this counterfactual, which could be done in several ways, e.g. by building synthetic controls or by exploiting out-of-sample predictions from (in-sample) pre-lockdown estimated parameters. Synthetic controls might be difficult to be built here since one has to exploit post-lockdown trends of regions that are most similar the Italian ones in many dimensions (e.g. institutional arrangements, cultural background, economic conditions, etc.). Other

[11] A well-known case here is that of Arzano Lombardo where at end February appeared the first contagion cases in the province of Bergamo. Due to the strong relevance in terms of small-medium business the authorities decided not to create a red zone there, differently from what happened in Codogno. The outcome has been a strong diffusion of contagion and a number of deaths largely exceeding those of the previous year in the same month (100 against 10). Beyond authorities' decision we interpret the significance of this variable as the push from small corporate owners not to close their activities due to the fear of default and the effect of this decision on the number of adverse COVID-19 outcomes. https://www.ilpost.it/2020/04/01/disastro-alzano-lombardo-nembro/

EU regions could be natural candidates, yet Italy was the first in implementing social-distancing measures, while other countries did it later and in different ways, thereby making it hard to obtain good matches on pre-lockdown characteristics.

We therefore follow the second approach and construct a counterfactual trend through out-of-sample predictions. This approach seems particularly advisable given that epidemiologic dynamics are often modelled using deterministic approaches[12]. More specifically, we first estimate the following equation:

$$Mortality_{mt} = \gamma_0 + \gamma_1 Day_{mt} + \gamma_2 {Day_{mt}}^2 + \gamma_3 Cases_{it-5} + \eta_{mt}$$

Eq.3

where *Mortality* is the number of deaths per 1000 inhabitants in day *t* and municipality *m*, and *Cases* is the one week (5-day) lagged number of positive cases per 1000 inhabitants in province *i* (where municipality *m* is located)[13]. Notice that here we use data at the municipality level, which is the smallest Italian geographic unit; in fact, mortality is the only COVID-19 related outcome available at such a geographic level. We perform an OLS fixed effects panel regression[14], clustering standard errors at regional level to account for intra-regional error correlation. Hence, we extract the predicted values $\widehat{T\_Mortality_{mt}}$, which capture our "real" mortality trend. As argued above, fixed-effects panel models have the advantage of partialling out time-invariant unobserved factors that might explain different trends of contagion and mortality across municipalities such as quality of local health system, number of elderly individuals in nursing homes and human mobility.

To build a "counterfactual" trend, we re-estimate Eq. 3 by restricting the sample to $t \leq l$, where l is March, 12th 2020, i.e. the day after full lockdown is introduced in the entire country. Hence, we compute the predicted values also for $t > l$, which provide us with the out-of-sample linear prediction of mortality after lockdown, i.e. $\widehat{C\_Mortality_{mt}}$. The latter can be thought of an approximation to the counterfactual trend of mortality, which describes how mortality would have evolved over time had lockdown not been introduced.

[12] Given the small share of population contagions during the estimation period (far lower than 10 percent even with the less conservative forecast applying an exogenous fixed 2 percent mortality rate) we may reasonably assume that we are far from the herd immunity, and therefore our deterministic model continues to be valid to predict the counterfactual in this time period.

[13] Although the time window from illness onset to death is larger, we set it at 5 to maximize the number of time periods in the sample. The inclusion of high order lags reduces the number of observed time periods, thereby hindering the goodness of fit, especially for the prediction of the counterfactual trend (which is already based on few time periods).

[14] We decided to opt for OLS fixed-effects regressions since they seem to perform better in terms of AIC and BIC criteria than Poisson fixed-effects ones. Under the same criteria, the chosen specification also outperforms other specifications that instead use the number of deaths (and cases) or their log.

Figure 2 shows the two trends averaged across municipalities. This figure highlights that the two trends start diverging significantly from March, 20th 2020, i.e. 8 days after full lockdown, when the real mortality trend began to decrease slightly. To further evaluate whether the two trends are statistically different after March, 12th 2020 we first compute their difference as $Mortality\_diff_{mt} = \widehat{C\_Mortality}_{mt} - \widehat{T\_Mortality}_{mt}$ and regress it through fixed-effects OLS on a series of time dummies. Margins from this regression are plotted in Figure 3 and confirm the previous findings, with the gap between counterfactual and real mortality trend starting to increase after lockdown.

[Figures 2 and 3 here]

To assess the role of historical exposure pollution, we perform OLS fixed-effects regressions of $Mortality\_diff_{mt}$ on lockdown decisions, and interact the latter with PM2.5, PM10 and NO2. More specifically, we introduce a dummy variable equal to one if the municipality is in a province with average concentration of PM2.5 (or, in a separate model, PM10 or NO2) in 2018 equal or above their sample mean, and zero otherwise. As for lockdown, we create a dummy variable ($Lockdown_{it}$) equal to one if the municipality *m* in province *i* in day *t* was under the total block, and zero otherwise[15]. In alternative specifications, we also consider a different definition of lockdown by including an indicator ($After\_12/3_t$) for $t > l$. Regression results are reported in Table 7. The positive and significant coefficient of the lockdown indicators suggest that mortality would have been larger if lockdown were not implemented. The lockdown effects range from about -0.12 to -0.15 percentage points (Table 7, column 1 and 4). Therefore, our estimates suggest that lockdown decisions are on average associated to a decline in the number of deaths per 1,000 inhabitants by 80-100 percent if we consider the sample mean of the overall predicted mortality (i.e. 0.15), which includes both $C\_Mortality_{mt}$ and $T\_Mortality_{mt}$.

Importantly, the positive association between lockdown and mortality is larger in less polluted areas as documented by the significant interaction between lockdown and the pollution variables (Table 8, columns 2-3 and 5-6). The positive coefficient of the interaction, if interpreted jointly with the trends plotted in Figure 3, suggests that without lockdown mortality would have grown more in highly polluted cities.

[Table 8 here]

[15] The total lockdown decision was taken on the 8th of March in Lombardia and 18 provinces of Piemonte, Emilia Romagna and Marche, while on the 10th of March it was extended to the rest of Italy (see Table 1). We consider March 10th as the decree has been issued on the evening of March 9th and, therefore, it has been operating since March 10th.

As a final check for the joint role of lockdown decisions and pollution in mortality, we replicate this analysis by duplicating the municipalities in our sample; the new observations are assigned the previously estimated counterfactual trend in mortality ($\widehat{C\_Mortality_{mt}}$), whereas the original observations are assigned the real trend in mortality ($\widehat{T\_Mortality_{mt}}$). By doing so, we artificially create two groups of municipalities, namely the "treated" municipalities, i.e. those with the real predicted trend, and the "control" municipalities, i.e. those with the predicted counterfactual trend. Of course, the latter is a synthetic control since observations with a counterfactual trend do not really exist; they have been imputed by duplicating the observations in our sample and assigning them the counterfactual estimated trend. This trick should deliver exactly the same results as the previous ones, while providing us at the same time with a quasi-experimental sample on which we could implement the Differences-in-Differences approach (DiD).

Hence, we create a *treatment* indicator (*Real_trend*$_m$) which is equal to one for municipalities with the *real* mortality trend, and zero for their "clones" receiving the counterfactual trend. Figure A1 in the appendix report estimated margins from a regression of the overall mortality trend, i.e. $\widehat{Mortality_{mt}}$, on time dummies and their interaction with the treatment indicator (*Real_trend*$_m$). As expected, the two trends mirror those plotted in Figure 2, with the real trend in mortality decreasing after a about a week from the day when full lockdown was introduced in the country.

To obtain DiD estimates, we regress the overall mortality trend on our treatment indicator (*Real_trend*$_m$), our *post-treatment* indicator, i.e. the lockdown dummy (*Lockdown*$_{it}$ or *After_12/3*$_t$), and the interactions between these two variables. In augmented models we also interact these two indicators with the aforementioned pollution dummies for PM10, PM2.5 or NO2. Results are reported in Appendix, Table A1. Consistent with previous results, lockdown seems effective in reducing mortality, as shown by the DiD coefficient, i.e. the negative and significant interaction between treatment and post-treatment indicator in column 1 (*Lockdown*$_{it}$**Real_trend*$_m$) or 4 (*After_12/3*$_t$**Real_trend*$_m$).

When estimating the triple interaction with the pollution indicator (columns 3 and 6), the interpretation of the results gets complicated. For this reason, we plot the estimated margins (Figure A2 in Appendix), which overall suggest that the adverse effects of COVID-19 on mortality would have been hasher without lockdown, especially in highly polluted provinces.

As a robustness check, we also repeat the previous analysis estimating a more complex mortality model than that described in eq. 3. We include also two interaction terms between the time-varying lagged contagion ($Cases_{it-5}$) and i) the share of micro-enterprises in the province (*Artisan*$_i$), and ii) the normalized first extracted component from a principal component analysis including our three human mobility proxies measured at the province level, i.e. *InternalCommuting*$_i$,

$ExternalCommuting_i$, and $PublicTransportUse_i$. These interactions allow us to mitigate unobserved heterogeneity that could be induced by time-variant behavior of firms and individuals in complying with lockdown decisions. The underlying assumption behind the choice of these variables is that the likelihood of incompliance would be higher in provinces with a high fraction of firms that cannot easily adopt smart-work solutions as well as in those characterized by large human mobility. Results are consistent with – and in some cases even reinforce – all the previous estimates (available upon request).

6. Discussion

Our findings have several limitations and implications for future research. The statistical significance of our regressors does not necessarily imply causality and, based on the characteristics of our data, we do not obviously have the possibility to test causality through a proper counterfactual trend or through RCTs. That is, it is impossible to build a properly randomized control group for a phenomenon that is already occurring at the time of the evaluation. In other words, we cannot create treatment and control groups with balanced baseline characteristics, "inoculate" the virus into the former group and compare the reactions among the two groups. However, the statistical significance of the three significant predictors (quality of air, economic activity and lockdown) withstands a barrage of robustness checks. Furthermore, the aforementioned predicted counterfactual analysis is, to the best of our knowledge, the closest approach to the first-best comparison between actual and counterfactual dynamics.

We also acknowledge a number of limitations in the quality of data. First, the COVID-19 test policy in Italy was different over time and across regions. Initially, tests were performed to suspected patients who present to hospital and/or people who have been in contact with positive cases; then, only patients with severe symptoms were tested. More recently, tests were also performed to suspected people with no severe symptoms. In addition, some regions and provinces, adopted a policy to test only patients with severe symptoms in different periods[16]. Second, the available data on mortality record only municipalities with more than 10 deaths as of April, 4th 2020, and with a mortality difference between the first three months of 2020 and the average for the same period between 2015 and 2019 being greater than 20%[17]. This might introduce bias in the estimates since the municipalities with a low death rate or with a death rate that is not highly different from past

[16] For instance, in the municipality of Vo' all population was tested on 28 February 2020 (source: https://www.ansa.it/sito/notizie/cronaca/2020/02/28/zaia-da-test-vo-studio-epidemiologico_2c3d88f3-6a4a-4e00-b255-9e1e2feb2768.html).

[17] See https://www.istat.it/it/files//2020/03/Il_punto_sui_decessi_9_aprile_2020.pdf, footnote 1.

mortality rates are excluded from the mortality statistics provided by ISTAT. However, we can argue that our results would apply to the subsample of municipalities most affected by COVID-19, and therefore may be interpreted as an "intensity" effect. More research is, however, needed on the refinement of our dependent variables.

Finally, our estimates of the lockdown effects at municipality level might also be subject to bias. First, fixed-effects estimates, while netting out important unobserved fixed confounders, might not consider time-invariant factors that could influence mortality over time, such as, for instance, the ability of local administrations to effectively respond to the dynamic of the virus, the decision of firms and individuals to comply or not with the restrictions, or the behavior of citizens regarding social interactions and sanitation measures (independently from lockdown). More responsive and efficient administrations or the presence of prudent citizens might have reduced the distance between the actual and the estimated counterfactual trend. Moreover, measurement error and non-perfect forecasting might also be an issue for our estimates; however, in spite of several requests, we could not get access to precise data neither on contagion and recovery at the municipality level nor on mortality due to COVID-19. Notwithstanding all these problems, our estimates, overall, seem to suggest that the lockdown decisions might have been effective in reducing mortality.

7. Conclusions

Our investigation started from the observation of the uneven distribution of contagion across Italian provinces. The survey of the literature on drivers of COVID-19 and other respiratory diseases indicates five major potential drivers: lockdown decisions, economic activity, frequency of people interactions in the area, quality of air and weather (e.g. temperature).

Our findings show that spread and severity of contagion is significantly associated to lockdown decisions, to factors affecting the quality of air and to the intensity of small business activity. These findings are robust to the use of different methodological approaches such as cross-section, pooled and fixed-effect OLS regressions as well as to DiD estimates exploiting a simulated counterfactual trend, against which we benchmark the effect of lockdown and of its interaction with historical levels of pollution.

We find that the presence of micro (artisan) firms is positively correlates with contagion and mortality, suggesting, on the one hand, a certain degree of resistance by small business to lockdown policies, and, on the other, the presence of high economic activity, which conceals human interactions (and hence the spread of the disease). Two other important conclusions can be drawn from our findings. First, notwithstanding the aforementioned methodological caveats, lockdowns seem to be rather effective in limiting contagion and mortality. Second, the quality of air is a strong predictor of

contagion and mortality: pre-existing levels of PM10 and PM2.5 and NO2 are positively correlated with both the COVID-19 outcomes under investigation. Finally, contrary to current stereotypes, we do not find evidence that the presence of Chinese immigrants correlates with the diffusion of the virus.

Several policy implications can be drawn our findings also conceal causality links. Some of the factors significantly correlated with COVID-19 outcomes are under human controls: lockdown policies, economic activity and, for most part, pollution. With reference to the latter sources of particulate matter are for the most part urban heating, transportation; energy; industry and agriculture) under our control, with only a small share depending on factors outside our control, such as atmospheric dusts (see footnote 7).

Hence it is in our power to reduce exposure of the global community to this risk factor. The most effective action concerns improved ecological efficiency of urban heating and efforts to reduce the impact on pollution of mobility, sources of energy and production in industry and agriculture are also important.

References

AEA “Air quality in Europe — 2017 report” (2017).

Bajardi, P., C. Poletto, J. J. Ramasco, M. Tizzoni, V. Colizza, and A. Vespignani (2011). Human mobility networks, travel restrictions, and the global spread of 2009 h1n1 pandemic. *PloS one* 6 (1).

Barreca AI, Shimshack JP. (2012) Absolute humidity, temperature, and influenza mortality: 30 years of county-level evidence from the United States. *Am J Epidemiol*; 176 Suppl 7: S114-22. 11.

Bukhari, Qasim, and Yusuf Jameel (2020). "Will Coronavirus Pandemic Diminish by Summer?” *Available at SSRN 3556998*

Charu, V., S. Zeger, J. Gog, O. N. Bjørnstad, S. Kissler, L. Simonsen, B. T. Grenfell, and C. Viboud (2017). Human mobility and the spatial transmission of influenza in the united states. *PLoS computational biology* 13 (2), e1005382

Conticini, Edoardo, Bruno Frediani, and Dario Caro (2020). "Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy?." *Environmental Pollution*: 114465.

Cereda, D., Tirani, M., Rovida, F., Demicheli, V., Ajelli, M., Poletti, P., & Merler, S. (2020). The early phase of the COVID-19 outbreak in Lombardy, Italy.

Duan, Zheng, et al. (2016) "Fine particulate air pollution and hospitalization for pneumonia: a case-crossover study in Shijiazhuang, China." *Air Quality, Atmosphere & Health* 9.7: 723-733.

European Court of Auditors (2018) Air pollution: EU citizens’ health still not sufficiently protected, warn Auditors, available at: https://www.eca.europa.eu/Lists/ECADocuments/INSR18_23/INSR_AIR_QUALITY_EN.pdf

ISS (2020) Characteristics of COVID-19 patients dying in Italy Report based on available data on March 26th, 2020

Ispra (2018) XIII Rapporto Qualità dell’ambiente urbano - Edizione 2017

Kelejian, H. H., and Prucha, I. R. (2010). “Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances”. *Journal of econometrics*, 157(1), 53-67.

Krewski, D. (2009) Evaluating the effects of ambient air pollution on life expectancy. *N. Engl. J. Med.*, 360, 413–415.

Fang, Hanming, Long Wang, and Zoe Yang. (2020) "Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China.", mimeo.

Le Mesurier, S. M., Stewart, B. W., O'Connell, P. J., & Lykke, A. W. J. (1979). Pulmonary responses to atmospheric pollutants. II. Effect of petrol vapour inhalation on secretion of pulmonary surfactant. *Pathology*, *11*(1), 81-87.

Li, Qun, Xuhua Guan, Peng Wu, Xiaoye Wang, Lei Zhou, Yeqing Tong, Ruiqi Ren et al. (2020) "Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia." *New England Journal of Medicine*.

Lowen AC, Mubareka S, Steel J, Palese P. (2007) Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathog*; 3(10): 1470-6. 10.

Luginaah IN, Fung KY, Gorey KM, Webster G, Wills C. (2005) Association of ambient air pollution with respiratory hospitalization in a Government-designated “Area of Concern”: The case of Windsor, Ontario. *Environ Health Perspect.*, 113: 290-296. 10.1289/ehp.113-a290.

Medina-Ramon M, Zanobetti A, Schwartz J (2006) The effect of ozone and PM10 on hospital admissions for pneumonia and chronic obstructive pulmonary disease: a national multicity study. *Am J Epidemiol* 163(6):579–588.

Mohammad M. Sajadi, MD, Parham Habibzadeh, MD, Augustin Vintzileos, PhD, Shervin Shokouhi, MD, Fernando Miralles-Wilhelm, Anthony Amoroso, M. (2020) “Temperature, humidity, and latitude analysis to predict potential spread and seasonality for COVID-19” *Available at SSRN 3550308* .

Murgese, E., 2020, Covid-19, esposizione al particolato e allevamenti intensivi, available at https://storage.googleapis.com/planet4-italy-stateless/2020/04/184484ca-ricerca-particolato-def.pdf

Neupane B, Jerrett M, Burnett RT, Marrie T, Arain A, Loeb M. (2010) Long term exposure to ambient air pollution and risk of hospitalization with community-acquired pneumonia in older adults. *Am J Respir Crit Care Med.*, 181: 47-53.

Neupane, Binod, et al. (2010) "Long-term exposure to ambient air pollution and risk of hospitalization with community-acquired pneumonia in older adults." *American journal of respiratory and critical care medicine* 181.1: 47-53.

Nieman, G. F., Clark Jr, W. R., Wax, S. D., & Webb, S. R. (1980). The effect of smoke inhalation on pulmonary surfactant. *Annals of surgery*, *191*(2), 171.

Pastva, A. M., Wright, J. R., & Williams, K. L. (2007). Immunomodulatory roles of surfactant proteins A and D: implications in lung disease. *Proceedings of the American Thoracic Society*, *4*(3), 252-257.

Piazzalunga-Expert, A. (2020) Evaluation of the potential relationship between Particulate Matter (PM) pollution and COVID-19 infection spread in Italy, mimeo

Pope III, C. Arden, and Douglas W. Dockery. "Health effects of fine particulate air pollution: lines that connect." *Journal of the air & waste management association* 56.6 (2006): 709-742.

Pope, C.A., 3rd; Ezzati, M.; Dockery, D.W. (2009). Fine-particulate air pollution and life expectancy in the United States. *N. Engl. J. Med.*, 360, 376–386.

Santus, Pierachille, et al. (2012) "How air pollution influences clinical management of respiratory diseases. A case-crossover study in Milan." *Respiratory research* 13.1: 95.

Shaman J, Pitzer V, Viboud C, Lipsitch M, Grenfell B. (2009): Absolute Humidity and the Seasonal Onset of Influenza in the Continental US. *PLoS Curr*; 2: RRN1138. 12.

Strak, M.; Janssen, N.A.; Godri, K.J.; Gosens, I.; Mudway, I.S.; Cassee, F.R.; Lebret, E.; Kelly, F.J.; Harrison, R.M.; Brunekreef, B.; et al. (2012) Respiratory health effects of airborne particulate matter: The role of particle size, composition, and oxidative potential-the RAPTES project. *Environ. Health Perspect*., 120, 1183–1189.

Wang, Q. and J. E. Taylor (2016). Patterns and limitations of urban human mobility resilience under the influence of multiple types of natural disaster. *PLoS one* 11 (1)

Wu, X., Nethery, R. C., Sabath, B. M., Braun, D., & Dominici, F. (2020). Exposure to air pollution and COVID-19 mortality in the United States. medRxiv.

Xiao W.M. et al. (2020) Exposure to air pollution and COVID-19 mortality in the United States, memo

Zanobetti A, Schwartz J. (2006) Air pollution and emergency admissions in Boston, MA. J Epidemiol Community Health., 60: 890-895. 10.1136/jech.2005.039834.

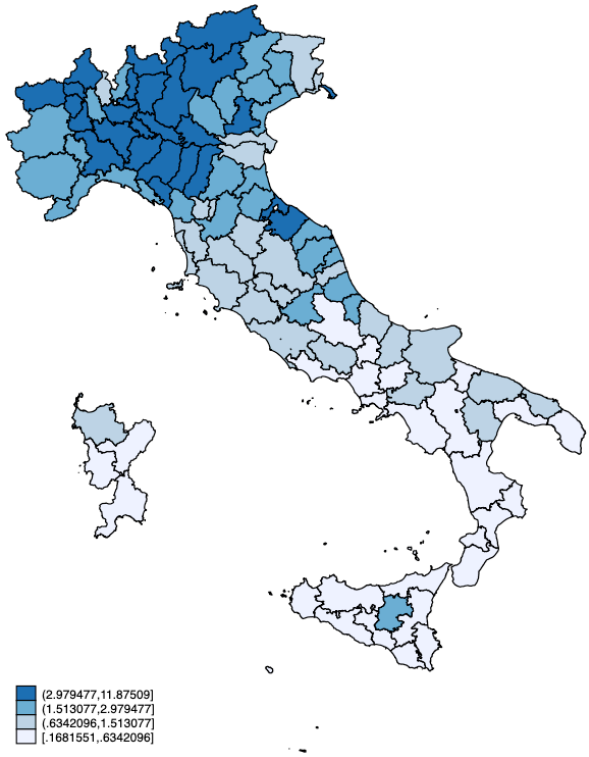
Zehender, Gianguglielmo, et al. (2020) "Genomic Characterisation And Phylogenetic Analysis Of Sars-Cov-2 In Italy." *Journal of Medical Virology*.

Zhang Y, He M, Wu S, Zhu Y, Wang S, Shima M, Tamura K, Ma L (2015) Short-term effects of fine particulate matter and temperature on lung function among healthy college students in Wuhan, China. *Int J Environ Res Public Health* 12(7):7777–7793.

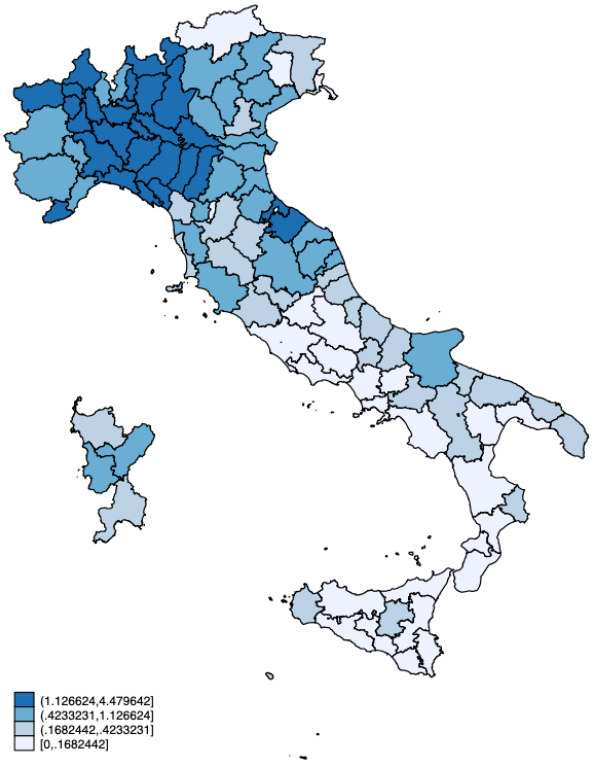
Zuk T, Rakowski F, Radomski JP. (2009) Probabilistic model of influenza virus transmissibility at various temperature and humidity conditions. *Comput Biol Chem*; 33(4): 339-43.

Figure 1 – COVID-19 contagion, mortality and pre-existing pollution levels

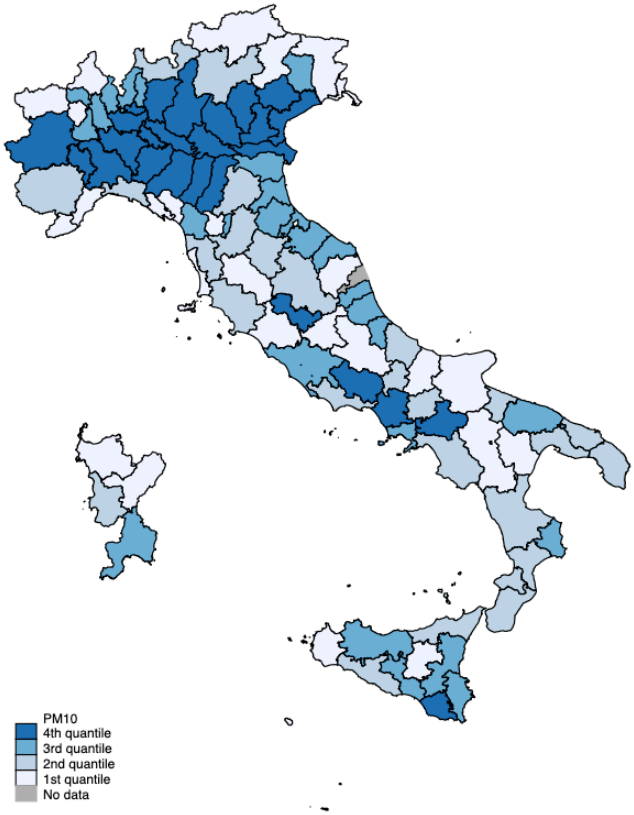
Panel A: Total cases per 1000 inhabitants



Panel B: Total deaths per 1000 inhabitants



Panel C: average levels of PM10 in 2018



Panel D: average levels of PM2.5 in 2018

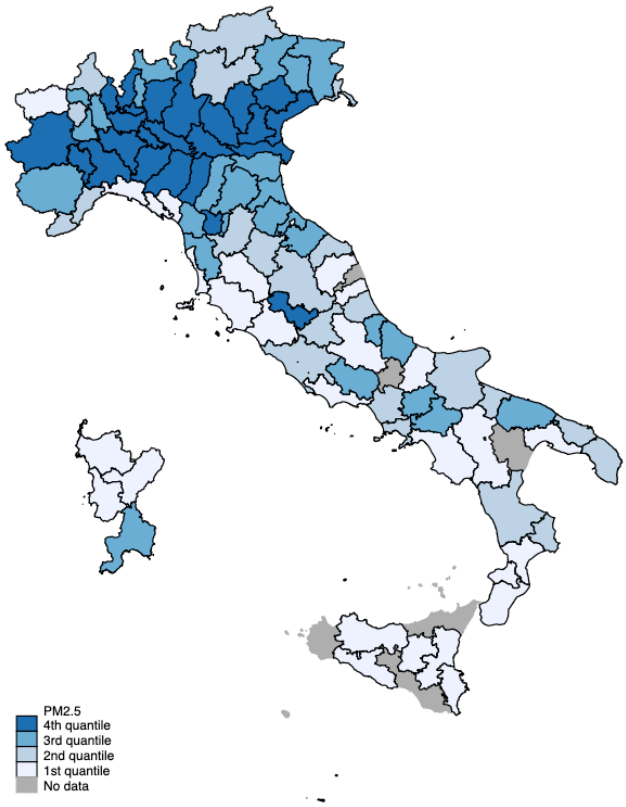


Figure 2 – COVID-19 mortality: real and counterfactual trend

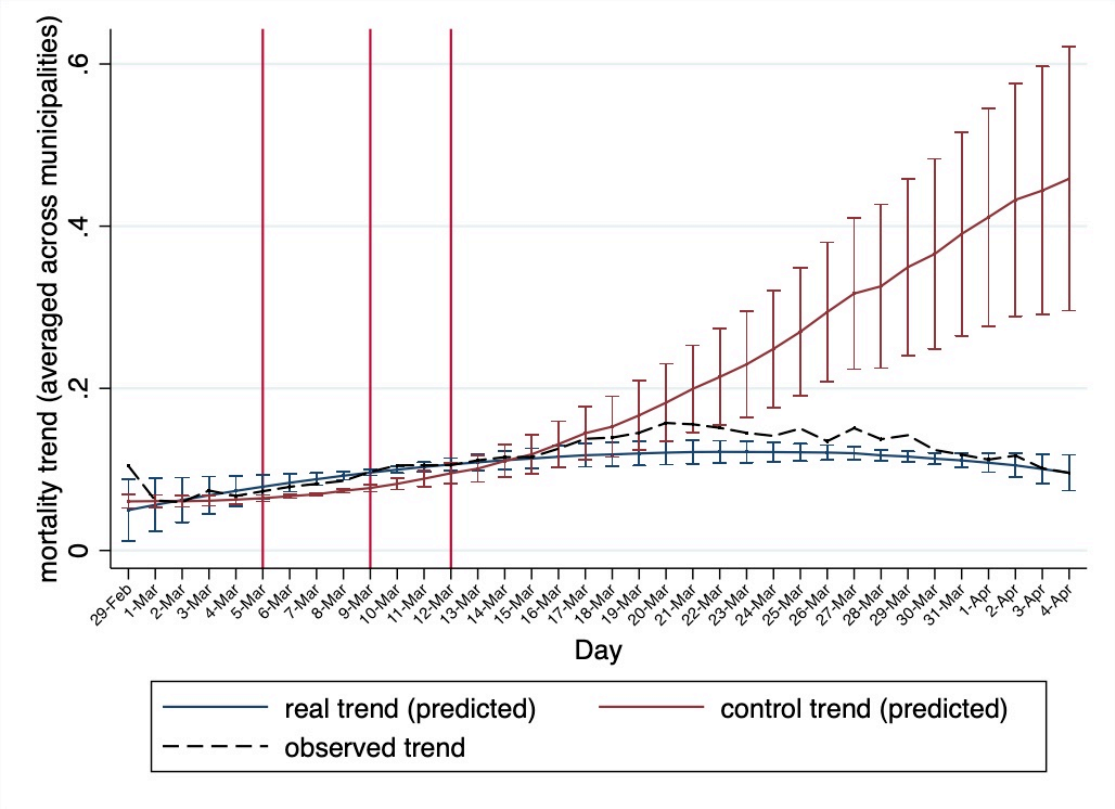


Figure 3 – COVID-19 mortality: difference between real and counterfactual trend

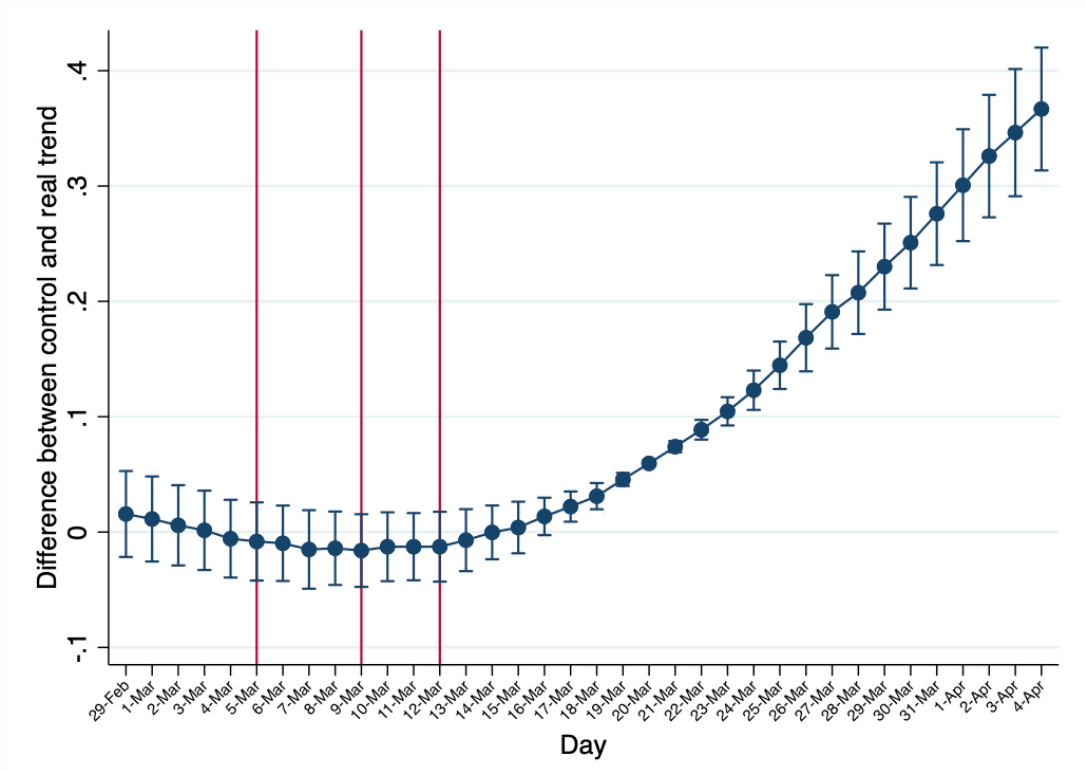
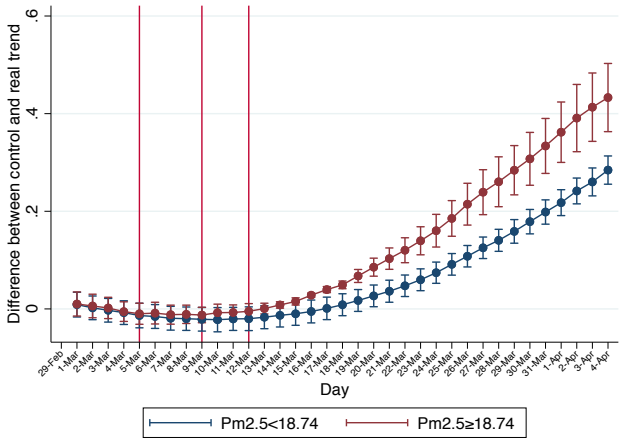
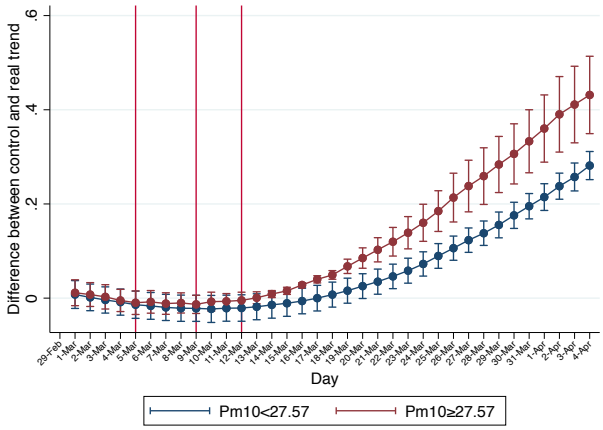


Figure 4 – COVID-19 mortality: difference between real and counterfactual trend by pollution

Panel A: lockdown effect & PM2.5



Panel B: lockdown effect & PM10



Panel C: lockdown effect & NO2

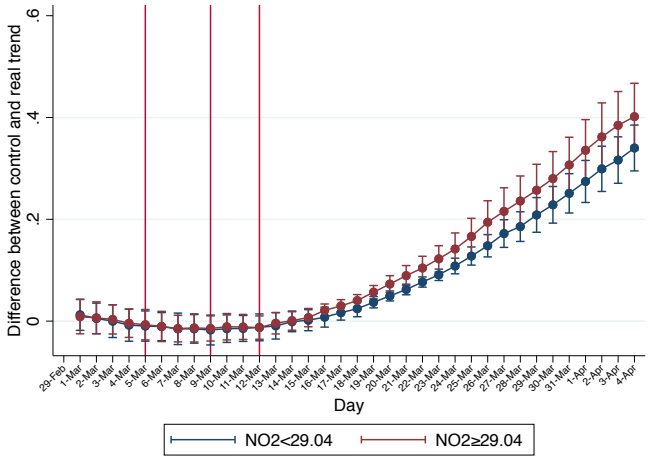


Table 1 – Restriction policies

Date	Restriction	Location	Source
February 23rd	Full lockdown at district level	Lombardia (Bertonico, Casalpusterlengo; Castelgerundo; Castiglione D'Adda; Codogno; Fombio, Maleo; San Fiorano, Somaglia, Terranova dei Passerini), Veneto (Vo').	https://www.gazzettaufficiale.it/eli/id/2020/02/23/20A01228/sg
February 25th	All public and private events and sport activities suspended; all school trips, monthly free access to museum suspended (national level)	Emilia Romagna, Friuli Venezia Giulia, Lombardia, Veneto, Liguria, Piemonte	https://www.gazzettaufficiale.it/eli/id/2020/02/25/20A01278/sg
March 1st	Partial lockdown (public events and schools suspended; other activities must ensure no big groups)	Emilia Romagna, Lombardia, Veneto; Pesaro e Urbino, Savona,	https://www.gazzettaufficiale.it/eli/id/2020/03/01/20A01381/sg
	Medium and Big-size enterprise closed on weekends	Bergamo, Lodi, Piacenza, Cremona	
March 4th	Public and private events suspended, smart working highly encouraged, elderly and unhealthy recommended to stay home, Lockdown of schools and universities and partial limitations	Italy	https://www.gazzettaufficiale.it/eli/id/2020/03/04/20A01475/sg
March 8th	Full lockdown	Lombardia, Modena, Parma, Piacenza, Reggio nell'Emilia, Rimini, Pesaro e Urbino, Alessandria, Asti, Novara, Verbano-Cusio-Ossola, Vercelli, Padova, Treviso, Venezia.	https://www.gazzettaufficiale.it/eli/id/2020/03/08/20A01522/sg
March 10th	Full lockdown	Italy	http://www.governo.it/it/articolo/firmato-il-dpcm-9-marzo-2020/14276#

Table 2 – Variable legend

Dependent variables	*Description*
Cases_pc	Cumulative number of COVID-19 daily cases (at province level) over total population, per 1,000 inhabitants
New_cases_pc	Number of daily new COVID-19 daily cases (at province level) over total population, per 1,000 inhabitants)
Deaths_pc	Cumulative number of daily deaths at province level over total population, per 1,000 inhabitants)
New deaths_pc	Average number of daily new deaths at province level over total population, per 1,000 inhabitants)
Day	Number of days since the first case was detected (24 February 2020)
Lockdown	Dummy = 1 if the province was on full lockdown (as for Table 1)
PM10	Average of yearly mean values in mg/mc registered by city monitoring posts in the i-th province (ISPRA 2018), and aggregated at province level by weighing observations by the population size of the municipality where the monitoring post is located.
PM2.5	Average of yearly mean values in mg/mc registered by city monitoring posts in the i-th province (ISPRA 2018) and aggregated at province level by weighing observations by the population size of the municipality where the monitoring post is located.
NO2	Average of yearly mean values in mg/mc registered by city monitoring posts in the i-th province (ISPRA 2018) and aggregated at province level by weighing observations by the population size of the municipality where the monitoring post is located.
Urban green	Square meters of green per 100 m2 of surface area of urban centers (BES, 2016)
High temperature	Dummy = 1 if the three days moving average of minimum temperature was higher than 12°C
Density	Population density in the province (number of residents in the province divided by province area)
Over65	Number of residents aged 65+ over total population, per 1,000 inhabitants.
Income	Average household disposable income in the province
Health (pca)	First factor extracted from a principal component analysis which includes the number of the following medical devices (i.e. the number of lung ventilators, diving chambers, ecographs, computed tomography scanners, hemodyalisis machines, medical monitors, nuclear magnetic resonance tomographs, operating rooms, radiology devices, portable radiology devices, linear particle accelerators, remote control radiology tables, immune-based automatic analysers, computerized gamma cameras, anesthetic machines, surgical lightheads, automatic coulter counters, per 1,000 inhabitants.
Public transport use	Number of people using public transports per 1,000 inhabitants.
Internal commuting	Total (work and education) internal commuting flows in the i-th province (Census data, ISTAT)
External commuting	Total (work and education) commuting flows in the i-th province from other provinces (Census data, ISTAT)
Day	Days since first case detected in Italy (24 February 2020)
Artisan	Percent of micro (artisan) firms on total enterprises (Unioncamere-Movimprese, 2017)

Table 3 – Summary statistics

Variable		Mean	Std. Dev.	Min	Max	Observations
Cases_pc	overall	1.196	1.951	0.000	14.810	N = 5,081
	between		1.362	0.076	7.244	n = 96
	within		1.404	-5.901	8.762	T = 52.927
New_cases_pc	overall	0.056	0.093	-0.672	1.708	N = 5,081
	between		0.053	0.004	0.279	n = 96
	within		0.077	-0.671	1.709	T = 52.927
New_deaths_pc	overall	0.021	0.028	0.000	0.209	N = 3,570
	between		0.024	0.001	0.112	n = 92
	within		0.015	-0.072	0.123	T = 38.804
Deaths_pc	overall	0.434	0.631	0.000	4.480	N = 3,152
	between		0.449	0.000	2.043	n = 85
	within		0.443	-1.479	2.963	T = 37.082
PM10	overall	24.416	5.238	13.164	35.509	N = 5,081
	between		5.267	13.164	35.509	n = 96
	within		0.000	24.416	24.416	T = 52.927
PM2.5	overall	15.455	4.526	5.450	26.423	N = 4,763
	between		4.548	5.450	26.423	n = 90
	within		0.000	15.455	15.455	T = 52.922
NO2	overall	23.851	7.998	3.000	47.090	N = 5,028
	between		8.037	3.000	47.090	n = 95
	within		0.000	23.851	23.851	T = 52.926
Urban green	overall	1.796	2.869	0.100	19.500	N = 5,081
	between		2.884	0.100	19.500	n = 96
	within		0.000	1.796	1.796	T = 52.927
Day	overall	50.973	15.288	25.000	77.000	N = 5,081
	between		0.306	48.000	51.038	n = 96
	within		15.285	24.934	76.973	T = 52.927
High temperature	overall	0.030	0.170	0.000	1.000	N = 5,081
	between		0.103	0.000	0.774	n = 96
	within		0.136	-0.744	1.011	T = 52.927
Density	overall	262.084	350.217	37.166	2623.520	N = 5,081
	between		351.788	37.166	2623.520	n = 96
	within		0.000	262.084	262.084	T = 52.927
Over65	overall	235.746	24.052	173.927	290.665	N = 5,081
	between		24.190	173.927	290.665	n = 96
	within		0.000	235.746	235.746	T = 52.927
Income	overall	0.109	0.070	0.011	0.406	N = 5,081
	between		0.070	0.011	0.406	n = 96
	within		0.000	0.109	0.109	T = 52.927
Health (pca)	overall	0.233	2.683	-5.006	9.682	N = 5,081
	between		2.695	-5.006	9.682	n = 96
	within		0.000	0.233	0.233	T = 52.927
Public transport use	overall	0.171	0.193	0.010	1.397	N = 5,081
	between		0.194	0.010	1.397	n = 96
	within		0.000	0.171	0.171	T = 52.927
Internal commuting	overall	0.432	0.049	0.286	0.577	N = 5,081
	between		0.049	0.286	0.577	n = 96
	within		0.000	0.432	0.432	T = 52.927
External commuting	overall	0.035	0.021	0.004	0.113	N = 5,081
	between		0.021	0.004	0.113	n = 96
	within		0.000	0.035	0.035	T = 52.927
Artisan	overall	0.267	0.061	0.118	0.382	N = 5,081
	between		0.061	0.118	0.382	n = 96
	within		0.000	0.267	0.267	T = 52.927

Table 4 – Major factors explaining variation in COVID-19 contagion (c*ross-section*)

	(1)	(2)	(3)	(4)	(5)	(6)
	As of March, 5th 2020			As of April, 17th 2020		
PM10	0.0217***			0.152***		
	(0.00705)			(0.0464)		
PM2.5		0.0249**			0.185***	
		-0.00993			(0.0690)	
NO2			0.00894*			0.0442
			(0.00537)			(0.0324)
Urban green	-0.00898	-0.0143	-0.00745	-0.0464	-0.0915	-0.0285
	(0.0109)	(0.0131)	(0.0113)	(0.0688)	(0.0874)	(0.0715)
High temperature	-0.117	-0.0578	-0.128	0.256	0.466	0.197
	(0.181)	(0.189)	(0.190)	(0.523)	(0.602)	(0.560)
Density	-7.19e-05	-0.0000458	-0.000141	-0.000478	-0.000116	-0.00102
	(0.000102)	-0.000107	(0.000115)	(0.000663)	(0.000741)	(0.000706)
Over65	-0.00379**	-0.00486**	-0.00437**	-0.00998	-0.0201*	-0.0157
	(0.00177)	-0.0019	(0.00185)	(0.0109)	(0.0117)	(0.0113)
Income	0.580	0.731	0.366	9.377***	10.64***	8.227**
	(0.543)	-0.638	(0.564)	(3.241)	(3.785)	(3.450)
Health (pca)	0.00352	0.00984	-1.77e-05	-0.0812	-0.000575	-0.101
	(0.0142)	-0.0166	(0.0148)	(0.0851)	(0.100)	(0.0916)
Public transport use	0.140	0.105	0.0567	-0.206	-0.373	-0.769
	(0.180)	-0.186	(0.184)	(1.140)	(1.194)	(1.148)
Internal commuting	-1.822**	-2.188**	-1.839**	6.081	3.102	8.293*
	(0.811)	-0.892	(0.861)	(4.750)	(5.394)	(4.933)
External commuting	0.997	0.8	2.850	-9.006	-13.86	6.843
	(1.920)	-2.078	(1.897)	(11.81)	(13.47)	(11.13)
Artisan	2.023***	1.847**	1.862**	26.74***	25.95***	25.26***
	(0.730)	-0.808	(0.784)	(4.567)	(5.107)	(4.658)
Constant	0.612	1.223**	1.099**	-10.08***	-5.057	-7.457**
	(0.530)	(0.560)	(0.535)	(3.182)	(3.512)	(3.322)
λ	0.0781	0.00886	0.117	0.457*	0.369	0.659***
	(0.546)	(0.562)	(0.570)	(0.237)	(0.290)	(0.203)
ρ	0.225	0.355	0.343	-2.158***	-1.700*	-2.386***
	(0.592)	(0.515)	(0.547)	(0.759)	(0.889)	(0.793)
σ^2	0.0840***	0.0912***	0.0904***	2.985***	3.299***	3.211***
	(0.0121)	(0.0136)	(0.0131)	(0.467)	(0.532)	(0.512)
Observations	96	90	95	95	89	94

Cross-sectional SAC model. Columns 1—3 refer to March, 5th 2020 and columns 4—6 refer to April, 17th 2020; ρ and λ are spatial autoregressive parameters measuring the degree of spatial correlation in the number of new cases and the disturbance term respectively; σ^2 is the Maximum Likelihood residuals variance. Standard errors are clustered at regional level, *** p<0.01, ** p<0.05, * p<0.1

Table 5 – Major factors explaining variation in mortality (c*ross-section*)

	(1)	(2)	(3)	(4)	(5)	(6)
	As of March, 5th 2020			As of April, 4th 2020		
PM10	0.00732***			0.0611***		
	(0.00234)			(0.0193)		
PM2.5		0.00842***			0.0717***	
		(0.00325)			(0.0273)	
NO2			0.00418**			0.0297**
			(0.00176)			(0.0150)
Urban green	-0.0104***	-0.0101**	-0.00962**	-0.0714**	-0.0704**	-0.0644*
	(0.00395)	(0.00410)	(0.00405)	(0.0329)	(0.0343)	(0.0343)
High temperature	-0.0306	-0.0173	-0.0530	-0.184	0.158	-0.0994
	(0.0569)	(0.0593)	(0.0601)	(0.757)	(0.790)	(0.782)
Density	0.000031	3.59e-05	-1.44e-05	0.000156	0.000180	-0.000172
	(0.0000337)	(3.53e-05)	(3.89e-05)	(0.000278)	(0.000293)	(0.000312)
Over65	0.000629	0.000120	0.000294	-0.00193	-0.00556	-0.00443
	(0.000582)	(0.000628)	(0.000592)	(0.00464)	(0.00504)	(0.00483)
Income	0.152	0.277	0.168	1.358	2.036	1.354
	(0.178)	(0.211)	(0.186)	(1.424)	(1.707)	(1.509)
Health (pca)	-0.00424	-0.00151	-0.00652	-0.0172	-0.00562	-0.0370
	(0.00472)	(0.00541)	(0.00485)	(0.0379)	(0.0439)	(0.0397)
Public transport use	0.0209	0.0127	-0.0332	0.161	0.0767	-0.302
	(0.0726)	(0.0751)	(0.0747)	(0.589)	(0.616)	(0.610)
Internal commuting	-0.238	-0.366	-0.186	-1.334	-2.026	-0.476
	(0.267)	(0.292)	(0.270)	(2.100)	(2.329)	(2.146)
External commuting	0.181	0.141	0.914	-2.649	-2.756	4.095
	(0.655)	(0.697)	(0.630)	(5.261)	(5.686)	(5.009)
Artisan	1.136***	1.043***	0.988***	9.738***	8.922***	8.687***
	(0.239)	(0.258)	(0.254)	(1.939)	(2.098)	(2.045)
Constant	-0.432**	-0.168	-0.277	-2.426*	-0.561	-1.305
	80.175)	(0.196)	(0.174)	(1.331)	(1.479)	(1.392)
λ	0.195	0.0419	0.306	0.201	0.0413	0.313
	(0.413)	(0.481)	(0.399)	(0.385)	(0.469)	(0.399)
ρ	-0.0189	0.103	-0.164	-0.745	-0.521	-0.763
	80.672)	(0.668)	(0.755)	(0.826)	(0.862)	(0.977)
σ^2	0.00811***	0.00870***	0.00848***	0.520***	0.566***	0.554***
	(0.00122)	(0.00135)	(0.00129)	(0.0796)	(0.0887)	(0.0859)
Observations	88	83	87	88	83	87

Cross-sectional SAC model. Columns 1—3 refer to March, 5th 2020 and columns 4—6 refer to April, 4th 2020; ρ and λ are spatial autoregressive parameters measuring the degree of spatial correlation in the number of new deaths and the disturbance term, respectively; σ^2 is the Maximum Likelihood residuals variance. Standard errors are clustered at regional level, *** p<0.01, ** p<0.05, * p<0.1

Table 6 – Major factors explaining variation in COVID-19 contagion (pooled)

Dep. Var.:	(1)	(2)	(3)	(4)	(5)	(6)
	New cases pc			*New deaths pc*		
Day	0.0105***	0.0110***	0.0105***	0.00243**	0.00256**	0.00247**
	(0.00222)	(0.00225)	(0.00222)	(0.00116)	(0.00120)	(0.00117)
Day^2	-8.89e-05***	-9.38e-05***	-9.01e-05***	-2.36e-05*	-2.48e-05*	-2.41e-05*
	(2.11e-05)	(2.16e-05)	(2.09e-05)	(1.16e-05)	(1.21e-05)	(1.18e-05)
PM10	0.00352**			0.00154**		
	(0.00162)			(0.000669)		
PM2.5		0.00458**			0.00187*	
		(0.00209)			(0.000911)	
NO2			0.00181***			0.000851**
			(0.000608)			(0.000325)
Urban green	-0.00166	-0.00279***	-0.00121	-0.00143***	-0.00187***	-0.00130**
	(0.00117)	(0.000835)	(0.00131)	(0.000496)	(0.000474)	(0.000559)
High temperature	-0.00882	-0.00286	-0.0156***	-0.00406	-0.00132	-0.00954**
	(0.00578)	(0.00659)	(0.00452)	(0.00341)	(0.00276)	(0.00450)
Density	4.61e-06	8.70e-06	-1.06e-05	3.83e-06	4.72e-06	-5.09e-06
	(7.01e-06)	(5.70e-06)	(9.55e-06)	(3.26e-06)	(3.30e-06)	(5.19e-06)
Over65	-0.000336	-0.000523**	-0.000448*	-7.91e-05	-0.000154	-0.000156
	(0.000196)	(0.000234)	(0.000229)	(0.000123)	(0.000143)	(0.000137)
Income	0.145**	0.180**	0.123**	0.0415	0.0481	0.0385
	(0.0592)	(0.0679)	(0.0553)	(0.0271)	(0.0330)	(0.0298)
Health (pca)	-0.000960	0.000392	-0.00151	-0.000346	3.49e-05	-0.000739
	(0.00177)	(0.00169)	(0.00168)	(0.000598)	(0.000506)	(0.000620)
Public transport use	0.0155	0.00875	0.00270	0.00657	0.00270	-0.00370
	(0.0152)	(0.0144)	(0.0209)	(0.0126)	(0.0113)	(0.0118)
Internal commuting	0.0697	-0.00626	0.0650	-0.0363	-0.0577*	-0.0250
	(0.104)	(0.0926)	(0.0840)	(0.0374)	(0.0296)	(0.0465)
External commuting	-0.221	-0.338**	0.0855	-0.0541	-0.0863	0.0968
	(0.195)	(0.142)	(0.215)	(0.0770)	(0.0523)	(0.0705)
Artisan	0.561***	0.520***	0.523***	0.229***	0.214***	0.208***
	(0.0875)	(0.0798)	(0.0894)	(0.0539)	(0.0444)	(0.0526)
Constant	-0.420***	-0.328***	-0.342***	-0.104***	-0.0660**	-0.0695***
	(0.0670)	(0.0544)	(0.0500)	(0.0347)	(0.0258)	(0.0223)
Observations	5,081	4,763	5,028	3,658	3,460	3,617
R-squared	0.271	0.265	0.277	0.359	0.346	0.334

Pooled OLS model. Standard errors are clustered at regional level, *** p<0.01, ** p<0.05, * p<0.1.

Table 7 – Major factors explaining variation in mortality and contagion (*fixed effects*)

Dep. Var.:	(1)	(2)	(3)	(4)	(5)	(6)
	New cases pc			*New deaths pc*		
Day	-0.0157***	-0.0110***	-0.00664**	-0.00866***	-0.00355***	-0.00145**
	(0.00388)	(0.00369)	(0.00277)	(0.000719)	(0.000812)	(0.000723)
Day^2	8.85e-05**	5.29e-05	-6.87e-06	8.64e-05***	3.93e-05***	1.25e-05*
	(3.82e-05)	(3.59e-05)	(2.52e-05)	(7.42e-06)	(8.35e-06)	(7.34e-06)
Day*PM10	0.000826***			0.000385***		
	(0.000111)			(2.66e-05)		
Day^2*PM10	-8.42e-06***			-4.05e-06***		
	(1.16e-06)			(2.98e-07)		
Day*PM2.5		0.00122***			0.000513***	
		(0.000143)			(3.68e-05)	
Day^2*PM2.5		-1.25e-05***			-5.32e-06***	
		(1.49e-06)			(4.13e-07)	
Day*NO2			0.000404***			0.000193***
			(7.02e-05)			(1.94e-05)
Day^2*NO2			-3.83e-06***			-1.94e-06***
			(7.29e-07)			(2.18e-07)
Day*Artisan	0.0124***	0.0121***	0.0106***	0.00515***	0.00519***	0.00488***
	(0.00177)	(0.00190)	(0.00173)	(0.000404)	(0.000457)	(0.000442)
Day*Urban green	-1.21e-05	-2.45e-05	-4.30e-06	-2.44e-05***	-3.55e-05***	-2.10e-05***
	(2.56e-05)	(3.01e-05)	(2.44e-05)	(6.26e-06)	(7.22e-06)	(6.22e-06)
High temperature	-0.00979	-0.0109	-0.00700	-0.00168	-0.00127	-0.00192
	(0.00872)	(0.00930)	(0.00842)	(0.00196)	(0.00218)	(0.00204)
Day*Density	7.60e-07***	8.02e-07***	3.42e-07	6.99e-08	1.58e-07**	-9.24e-08
	(2.41e-07)	(2.49e-07)	(2.51e-07)	(6.26e-08)	(6.24e-08)	(6.78e-08)
Day*Over65	1.25e-05***	1.02e-05**	1.06e-05***	-4.93e-06***	-6.79e-06***	-6.31e-06***
	(4.22e-06)	(4.48e-06)	(4.08e-06)	(1.03e-06)	(1.10e-06)	(1.06e-06)
Day*Income	0.00331**	0.00377**	0.00370***	0.00172***	0.00209***	0.00183***
	(0.00131)	(0.00152)	(0.00124)	(0.000313)	(0.000369)	(0.000323)
Day*Health (pca)	-5.50e-05	-3.37e-05	-5.25e-05	-1.48e-05*	-1.07e-05	-2.67e-05***
	(3.42e-05)	(3.92e-05)	(3.26e-05)	(8.36e-06)	(9.54e-06)	(8.43e-06)
Day*Public transport use	7.45e-05	1.38e-05	9.48e-05	0.000119	7.08e-05	-6.87e-05
	(0.000429)	(0.000437)	(0.000404)	(0.000134)	(0.000133)	(0.000130)
Day*Internal commuting	0.00931***	0.00836***	0.00839***	0.000728	0.000581	0.00112**
	(0.00197)	(0.00212)	(0.00189)	(0.000481)	(0.000531)	(0.000500)
Day*External commuting	-0.0163***	-0.0185***	-0.0125***	-0.00473***	-0.00557***	-0.00137
	(0.00459)	(0.00490)	(0.00419)	(0.00115)	(0.00122)	(0.00111)
ρ	-0.799***	-0.851***	-0.349**	0.488***	-0.812***	-0.619***
	(0.101)	(0.0968)	(0.146)	(0.0367)	(0.132)	(0.149)
λ	0.786***	0.792***	0.671***	-0.554	0.680***	0.661***
	(0.0295)	(0.0281)	(0.0575)	(0)	(0.0479)	(0.0588)
σ^2	0.00464***	0.00485***	0.00418***	0.000163***	0.000170***	0.000166***
	(9.80e-05)	(0.000106)	(8.86e-05)	(3.70e-06)	(4.12e-06)	(3.90e-06)
Observations	4,608	4,320	4,560	3,772	3,526	3,731
R-squared	0.177	0.170	0.181	0.222	0.245	0.212
Number of Provinces	96	90	95	92	86	91

Panel Fixed-effects SAC model; ρ and λ are spatial autoregressive parameters measuring the degree of spatial correlation in the number of new cases or deaths and the disturbance term, respectively; σ^2 is the Maximum Likelihood residuals variance. Standard errors are clustered at regional level, *** p<0.01, ** p<0.05, * p<0.1

Table 8 – The role of lockdown and pollution in reducing mortality (*fixed effects OLS*)

Dep. Var: *difference in predicted mortality trends (C-R)*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lockdown	0.121***	0.0910***	0.0896***	0.111***				
	(0.0188)	(0.00525)	(0.00466)	(0.0140)				
lockdown*Pm2.5≥mean		0.0604**						
		(0.0213)						
lockdown*Pm10≥mean			0.0606**					
			(0.0254)					
lockdown*NO2≥mean				0.0268*				
				(0.0141)				
After 12/3					0.151***	0.107***	0.106***	0.137***
					(0.0248)	(0.00717)	(0.00645)	(0.0191)
After 12/3*Pm2.5≥mean						0.0805***		
						(0.0238)		
After 12/3*Pm10≥mean							0.0806***	
							(0.0280)	
After 12/3*NO2≥mean								0.0334*
								(0.0166)
Constant	0.00203	-0.00168	-0.00130	0.00122	-0.00608	-0.00693	-0.00698	-0.00632
	(0.0145)	(0.0108)	(0.0121)	(0.0127)	(0.0166)	(0.0107)	(0.0123)	(0.0143)
Observations	30,534	30,062	30,415	30,254	30,534	30,062	30,415	30,254
R-squared	0.149	0.159	0.159	0.151	0.294	0.315	0.314	0.297
Number of provinces	1,685	1,658	1,679	1,672	1,685	1,658	1,679	1,672

Standard errors are clustered at regional level. Dependent variable: difference between previously estimated *counterfactual (C)* and *real (R)* trend in mortality*;* *** p<0.01, ** p<0.05, * p<0.1

APPENDIX

Table A1 – The effects of lockdown and pollution in reducing mortality (*DID approach*)

Dep. Var: *predicted mortality trends*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lockdown	0.158***	0.122***	0.121***	0.145***				
	(0.0221)	(0.00613)	(0.00542)	(0.0164)				
Lockdown*Real trend	-0.121***	-0.0910***	-0.0896***	-0.111***				
	(0.0188)	(0.00525)	(0.00466)	(0.0140)				
Lockdown*Pm2.5≥mean		0.0718***						
		(0.0245)						
Lockdown*Real trend* Pm2.5≥mean		-0.0604**						
		(0.0213)						
Lockdown*Pm10≥mean			0.0714**					
			(0.0288)					
Lockdown*Real trend* Pm10≥mean			-0.0606**					
			(0.0254)					
Lockdown*NO2≥mean				0.0324*				
				(0.0164)				
Lockdown*Real trend* NO2≥mean				-0.0268*				
				(0.0141)				
After 12/3					0.183***	0.134***	0.133***	0.167***
					(0.0274)	(0.00791)	(0.00705)	(0.0212)
After 12/3*Real trend					-0.151***	-0.107***	-0.106***	-0.137***
					(0.0248)	(0.00717)	(0.00645)	(0.0191)
After 12/3*Pm2.5≥mean						0.0885***		
						(0.0263)		
After 12/3*Real trend* Pm2.5≥mean						-0.0805***		
						(0.0238)		
After 12/3*Pm10≥mean							0.0887**	
							(0.0310)	
After 12/3*Real trend* Pm10≥mean							-0.0806***	
							(0.0280)	
After 12/3*NO2≥mean								0.0372*
								(0.0181)
After 12/3*Real trend* NO2≥mean								-0.0334*
								(0.0166)
Constant	0.0769***	0.0743***	0.0746***	0.0763***	0.0802***	0.0797***	0.0797***	0.0801***
	(0.00983)	(0.00716)	(0.00806)	(0.00856)	(0.0101)	(0.00653)	(0.00748)	(0.00865)
Observations	61,068	60,124	60,830	60,508	61,068	60,124	60,830	60,508
R-squared	0.231	0.242	0.242	0.233	0.381	0.402	0.402	0.384
Number of id	3,370	3,316	3,358	3,344	3,370	3,316	3,358	3,344

Dependent variable: previously estimated trend in mortality, i.e $Mortality_{mt}$, which is equal to $\widehat{C\_Mortality_{mt}}$ for the "cloned" observations to which we have assigned the predicted counterfactual trend, and to $\widehat{T\_Mortality_{mt}}$ for original observations, to which we have assigned the (real) predicted mortality trend (see section 5); $Real\_trend_m$, is the *treatment* indicator, which is equal to one for municipalities with the *real* mortality trend, and zero for their "clones" receiving the counterfactual trend. We use as post-treatment variable the dummy $Lockdown_{it}$, which is equal to one if the municipality *m* in province *i* in day *t* was under the total block, or the dummy $After\_12/3_t$, which is equal to one for *t > March, 12th 2020.* The interaction $Lockdown_{it}$*$Real\_trend_m$ or $After\_12/3_t$*$Real\_trend_m$ (i.e. the interaction between treatment and post-treatment indicator) is the DiD coefficient; standard errors are clustered at regional level; *** p<0.01, ** p<0.05, * p<0.1

Figure A1 – COVID-19 mortality: real and counterfactual trend

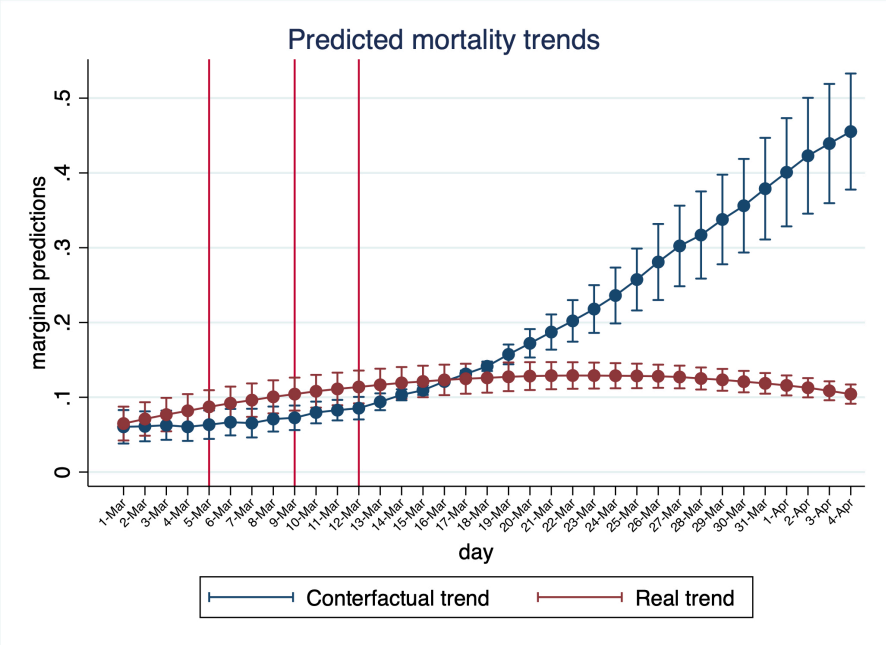
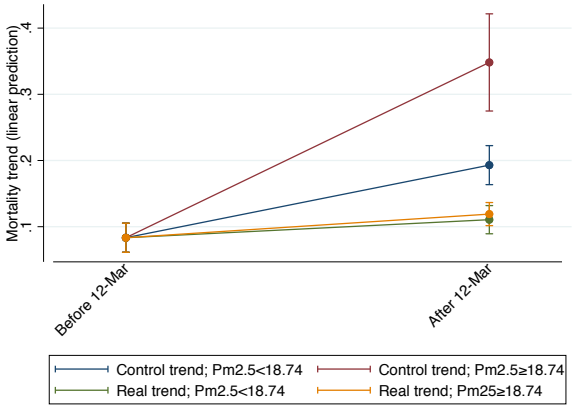
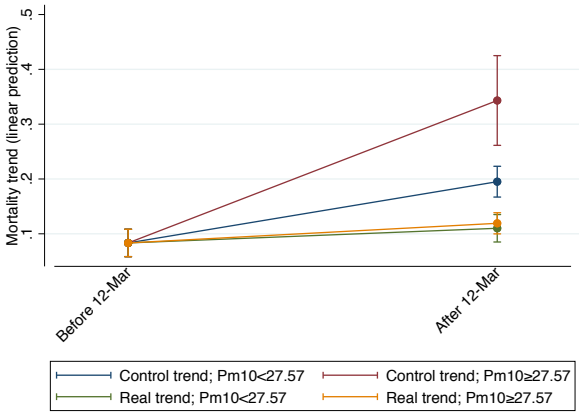


Figure A2 – COVID-19 mortality: mortality trends by lockdown and pollution

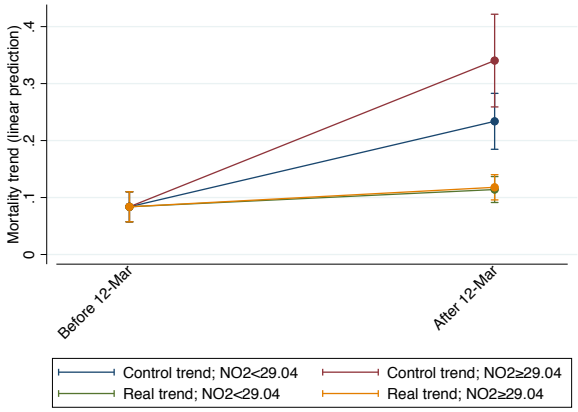
Panel A: General lockdown effect & PM2.5



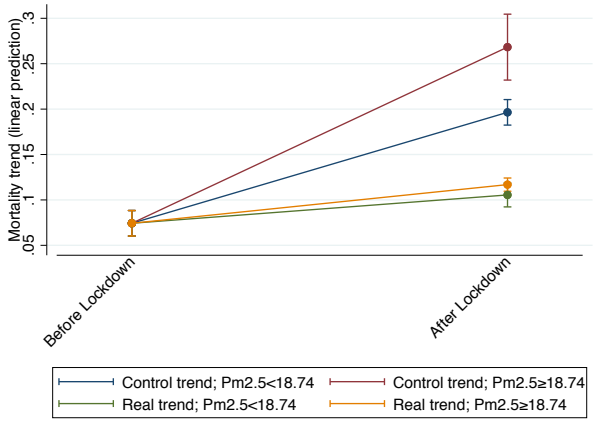
Panel B: General lockdown effect & PM10



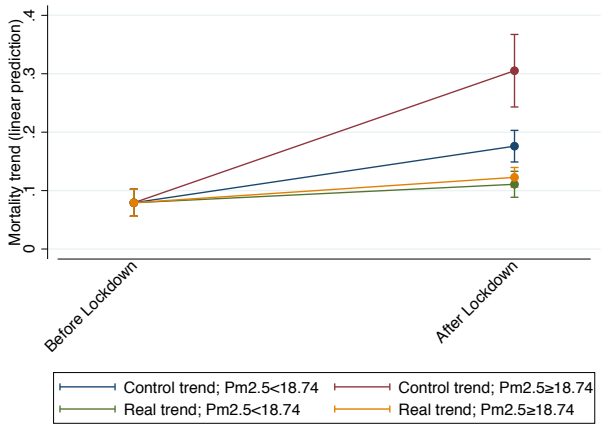
Panel C: General lockdown effect & NO2



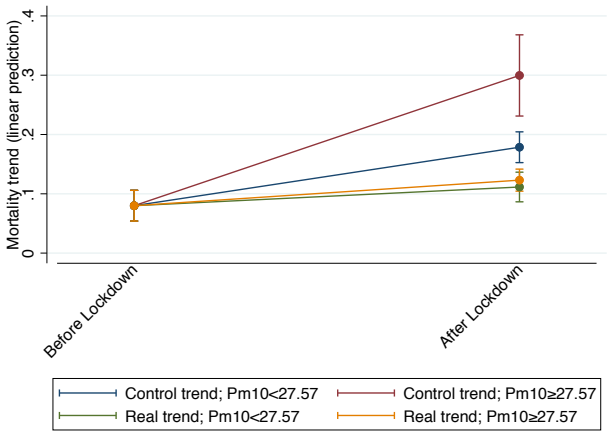
Panel D: Local lockdown effect & PM2.5



Panel E: Local lockdown effect & PM2.5



Panel F: Local lockdown effect & PM10



Panel G: Local lockdown effect & NO2