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*

***** PRELIMINARY AND INCOMPLETE *****

This version: *April 15, 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 particulate matter. We find that poor quality of air is significantly and negatively correlated with adverse outcomes of the epidemic, while lockdown and social distancing seem to be effective for contagions, but not yet for deceases. Our empirical findings are consistent with previous 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 pre-existing factors that we test, i.e. temperature, commuting, population density and the presence of Chinese community. We find, however, that adverse COVID-19 outcomes are significantly and positively correlated with the presence of 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.[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,

---

[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.

climate, pollution, and economic activity. The focus of our research has relevant implications on several dimensions such as subjective wellbeing, health policies, economic conditions and economic policies, not ultimately since – as of April, 7th 2020 – the epidemic in Italy caused 17,127 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 show that lockdown mitigates contagions, but – as of the time of this study – not yet mortality. This is, however, consistent with epidemiologic predictions of a lag of around 14 days of the effect of restriction policies on new contagions, and other 10-15 days from contagion to an unfortunate decease[7]. Conversely, poor quality of air 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 contagions. Finally, the heterogeneity of diffusion does not seem to depend on other pre-virus factors that we test, i.e. temperature, commuting, health system efficiency, density and the share of Chinese immigrants.

The paper is divided into six 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 we discuss our results (limits, policy implications and directions for future research). The sixth section concludes.

---

[6] According to the National Institute of Statistics, 34% of production was affected by the lockdown, with a negative impact on 385,000 employees. Source: https://www.ilsole24ore.com/art/coronavirus-istat-il-blocco-attivita-economiche-colpito-34percento-produzione-ADGG7rI.

[7] The World Health Organization report the time from symptoms onset to death ranging from 2 to 8 weeks (see https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf). In Italy, the median time has been estimated at 9 days (https://www.epicentro.iss.it/coronavirus/bollettino/Report-COVID-2019_30_marzo_eng.pdf).

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 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). We test this hypothesis on different dependent variables such as the number of contagions and the number of deceases.

There are two hypotheses on the effects of PM2.5 and PM10 as pull factors on 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 PM 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). Several empirical contributions have found support for this hypothesis on past respiratory diseases. 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.

Relatedly, 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) find 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), NO_2, SO_2, 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 ambient 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, SO_2 and PM_{10} 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).

A few very recent working papers focus directly on the relationship between pollution and COVID-19 disease. Xu 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 g/m3 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.

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/m3 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. Moreover, a positive association between economic intensity and COVID-19 outcomes could also be explained by the fact that provinces with a vibrant economic environment also enjoy more human interactions, which contribute to the spread of the disease.

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 social mobility and density since these factors increase the chances of human interactions 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, bacause this 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 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. Among dependent variables we consider current available official data on mortality and diffusion of COVID-19, i.e. the daily number of deaths (from the Italian National Statistics Institute, ISTAT) and new positive cases at province level (from the Italian Civil Protection).

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 March 28th 2020 (last date for which data are available), averaged at province level.

The second dependent variable is the number of daily new confirmed COVID-19 cases, that is the number of new infected patients detected each day. We prefer this measure to the number of net infected patients, where deaths and recoveries of the day are subtracted from the gross value in order to have a cleaner measure of the dynamics of the infection, which does not depend on individuals’ ability to recover. This ability might be a function of several observed as well as unobserved province-level characteristics. 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 the same national institutions and its local branches (Civil Protection), we therefore expect that these kinds of measurement bias do not affect our estimates. We consider the daily change in the number of positive cases from February 24th 2020 to April 6th 2020.

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

$$
\begin{aligned}
&\mathit{CV19\text{-}Outcome}_{it} \\
&\qquad = \alpha_0 + \alpha_1 Day_{it} + \alpha_2 {Day_{it}}^2 + \alpha_3\, PM_i \\
&\qquad + \alpha_4 DLockdown_i + \alpha_4 DHighTemperature_i + \alpha_5 Artisan_i + \alpha_6 Density_i \\
&\qquad + \alpha_7 Income_i + \alpha_8 Over65_i + \alpha_9 Ventilators_i + \alpha_{10} InternalCommuting_i \\
&\qquad + \alpha_{11} ExternalCommuting_i + \alpha_{11} PublicTransportUse_i + \varepsilon_{it}
\end{aligned}
$$

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

Regressors include a linear and a quadratic time trend (*Day* and Day^2) and a series of province time-invariant characteristics. 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 mg/mc registered by environmental monitoring units at province level. 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[8].

In order to take into account the effect of lockdowns (research hypothesis H_1), we introduce a dummy variable taking value one from the day after the lockdown decision. 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)[9]. Hence, the

---

[8] 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. 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.

[9] 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.

dummy variable $DLockdown$ is equal to one if province i in day t was under the total block and zero otherwise. To take into account for a delayed effect of the lockdown, we consider the 5-day lead of the variable $DLockdown$. Note that results are not affected by the choice of the leaded days, but a higher number of leaded days correspond to a higher and stronger effect. This confirm the hypothesis that the lockdown effect is indeed distributed over time.

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 ($DHighTemperature$), considering the 3-days lag of the variable to take into account the time between a possible effect and the illness onset[10]. Hypothesis H_4 is tested by introducing a variable measuring the share of artisan firms at province level (*Artisan*).

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. The number of lung ventilators per 1,000 inhabitants is introduced as a proxy for the efficiency of the local health system. In an alternative model-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]

---

[10] 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 estimates are OLS cross-sections, including the aforementioned major correlates of new positive cases as of March, 31st 2020 (Table 4) and deceases (Table 5).

[Tables 4-5 here]

These preliminary estimates show that income, economic activity and pollution are significantly correlated with the COVID-19 outcomes. More specifically, provinces with high levels of PM10 (Table 4-5, Column 1) and PM2.5 (Table 4-5, Column 2), as well as with high economic activity tend to have also worse outcomes in terms of contagion and deceases. 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 40 new cases per day (1200 per month) and 13.5 additional deaths per day (405 per month).

In order to exploit the time dimension of the dataset, we perform a pooled OLS estimate including also the time trend (and its square) as well as the lockdown indicator. Findings reveal that the COVID-19 outcomes follow an inverse U-shape exponential dynamic (the Day^2 variable is negative and significant) that seems to be halted by the lockdown decisions (Tables 4-5, columns 3-4). As in the previous estimates, other significant variables are exposure to particulate matter (both PM2.5 and PM10 used alternatively), and the share of artisan firms. 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.

The last empirical strategy rests on OLS panel fixed-effects. Of course, province time-invariant characteristics would be absorbed in the intercept in those models. We therefore test whether the above-mentioned pre-virus province characteristics differentially affect the *trend* of contagion and mortality. To this purpose, we interact the time variable (*day*) with each time-invariant control that has been included in the previous estimates. The advantage of this model is that it allows to net out the confounding effects of unobserved time-invariant factors.

[Table 6 here]

Results are reported in Table 6 and highlight, again, the role of two main factors. First, the lockdown effect is negative and significant, yet only for contagion. This might be interpreted in the light of the larger time mortality needs in order to respond to the lockdowns relative to contagion. In other terms, contagion adjusts more quickly than mortality to lockdown decisions. Second, the role of PM10 or PM2.5 is confirmed also under this stricter analysis. More specifically, this interaction captures the "slope" effect of historical pollution on COVID-19 outcomes; in other terms, it measures the differential trends of contagion and mortality by levels of preexisting pollution. Note that the mean effect of pollution is absorbed into the intercept and not identifiable in these kinds of estimate.

In order to have a clearer interpretation of this interaction, we generate two sets of graphs. First, we plot the estimated margins of the interaction between pollution (PM10 or PM2.5, alternatively) and time trend in the previous fixed effect estimate, at three different values of pollution (selected so as to maximize the number of observations). Results in Figure 1 show that there are significantly different trends in mortality and in contagion by pre-virus levels of particulate matter.

[Figure 1 here]

The second check for significantly different COVID-19 trends by pollution is carried out by re-estimating the OLS panel fixed-effects model and including only dummy variables for each day, jointly with the interaction between these dummies and an indicator for provinces where the pollution variable (PM10 or PM2.5, alternatively) is above the country median. The country median split reduces possible concerns of arbitrary selection of values in the estimated margins[11]. Figure 2 plots the estimated margins of this model, with red bars identifying the day in which restriction policies were introduced. The figure shows 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.

[Figure 2 here]

Overall, our empirical findings show a significant and robust role of lockdown policies in reducing contagion, and a negative correlation between pollution levels and both the COVID-19 outcomes under scrutiny. Moreover, the share of artisan firms has a positive and significant effect on both dependent variables. Our interpretation to this result (consistent with anecdotical evidence[12]) 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

---

[11] Standard errors are clustered at region level as above; clustering at province levels produces more precise estimates (available upon request)

[12] 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/

of the artisan variable conceals the effect of human interactions, which are typical of areas with higher economic activity, and therefore correlated with the spread of the virus. Diagnostics on goodness of fit of our models highlight that our models explain about 30-40% of the total variability of the dependent variables.

As a final test, we assess the relative role of the presence of the Chinese community in the spread of the disease. In the fixed-effects OLS regression of contagion, we also include an interaction term between the share of Chinese immigrants over total population (at the province level) with the time dummy (as in models in Table 6, columns 1-2). The interaction is positive but not statistically significant, with $\beta = 0.00514$ and $p = 0.495$ in the estimate with PM10, and $\beta = .00358$ and $p = 0.731$ in the estimate with PM2.5 (available upon request).

5. Discussion

While preliminary, our findings have several limitations and implications for future research. The significance of our regressors do not necessarily imply causality and, based on the characteristics of our data, we do not obviously have the possibility to test causality with a proper counterfactual or through a randomized experiment. Refinements in this direction with quasi-experiment and instrumental variable approaches (on which we are currently working) will, however, be important since policy implications from our results can be drawn only by assuming that our results hide - as we strongly suspect - causality links.

We also acknowledge a number of limitations due to 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[13]. Second, the available data on mortality record only municipalities with more than 10 deaths as of March, 28th 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%[14]. While this represents a selection bias as the data automatically excluded municipalities with low death or no significant difference with respect to past mortality rates, our results apply to the most affected subsample, and therefore may be interpreted as

---

[13] 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).

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

an “intensity” effect on most affected provinces. More research is, however, needed on the refinement of our dependent variables.

Another important direction for future research relates to the extension of the multivariate model testing simultaneously the four potential drivers of contagion on larger datasets at council level in Italy (on which we are currently working for a revised and complete version of this article) or on dataset including other affected countries.

6. 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-29 and other respiratory diseases indicates five major potential drivers: lockdown decisions, economic activity, frequency of people interactions in the area, pollution 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 especially fine particular matter) and the intensity of small business activity. We find that the presence of micro (artisan) firms is positively correlated 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, at the time of this paper, lockdowns seem effective in limiting contagion, while not yet mortality. This is, however, consistent with epidemiological predictions on the lag between social-distancing policies and their effects first on contagion and then on deceases. Second, the quality of air is a strong predictor of contagion and mortality: pre-existing levels of PM10 and PM2.5 are positively correlated with both the COVID-19 outcomes under investigation. Finally, contrary to anecdotical stereotypes, we do not find evidence that the presence of Chinese immigrants correlates significantly with the diffusion of the virus.

Several policy implications can be drawn if findings presented in our paper hide 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, according to AEA (2017) the sources of PM2.5 are for 94% under our control (57% urban heating, 11 percent transportation; 12 percent energy; 10 industry, 4 percent agriculture) with only 6 percent depending on factors outside our control, such as atmospheric dusts. Climate outcomes, instead, (wind, humidity and temperature) are for the most part beyond our control.

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. Several countries including Italy have launched tax allowance policies to stimulate investment in energy efficiency of buildings. Such programs need to be constantly revised in order to account for technological change in order to redirect incentives toward the most effective energy-efficiency investment. Efforts to reduce the impact on pollution of mobility, sources of energy and production in industry and agriculture are also important.

Our findings provide preliminary evidence that can be useful to calculate the trade-off beyond lockdown decisions, but with one caveat. The magnitude of the lockdown effect implied by our estimates is much lower than the actual effect of the lockdowns since it measures just the fall in contagions after the lockdown starts being effective. In order to calculate the total effect, we need to have an approximation of the counterfactual (the evolution of the epidemics without lockdown), which requires several assumptions and it is hard to evaluate, yet it is in the future steps of this research.

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. "Will Coronavirus Pandemic Diminish by Summer?. (2020). " *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, 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

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)D1,2 Temperature, humidity, and latitude analysis to predict potential spread and seasonality for COVID-19 *Available at SSRN 3550308* .

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, 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, S.; Deng, F.; Hao, Y.; Wang, X.; Zheng, C.; Lv, H.; Lu, X.; Wei, H.; Huang, J.; Qin, Y.; et al. (2014) Fine particulate matter, temperature, and lung function in healthy adults: Findings from the HVNR study. *Chemosphere*, 108, 168–174

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

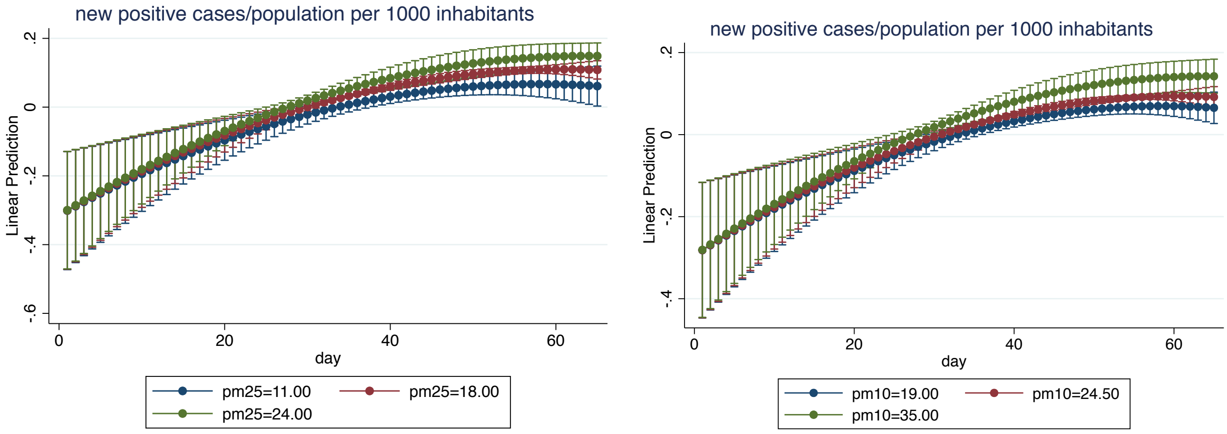
Zanobetti A, Schwartz J. (2006) Air pollution and emergency admissions in Boston, MA. *J Epidemiol Community Health.*, 60: 890-895. 10.1136/jech.2005.039834.

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 and mortality: the role of pre-existing pollution



Panel A: contagion and PM2.5

Panel B: contagion and PM10

Panel C: mortality and PM2.5

Panel D: mortality and PM10

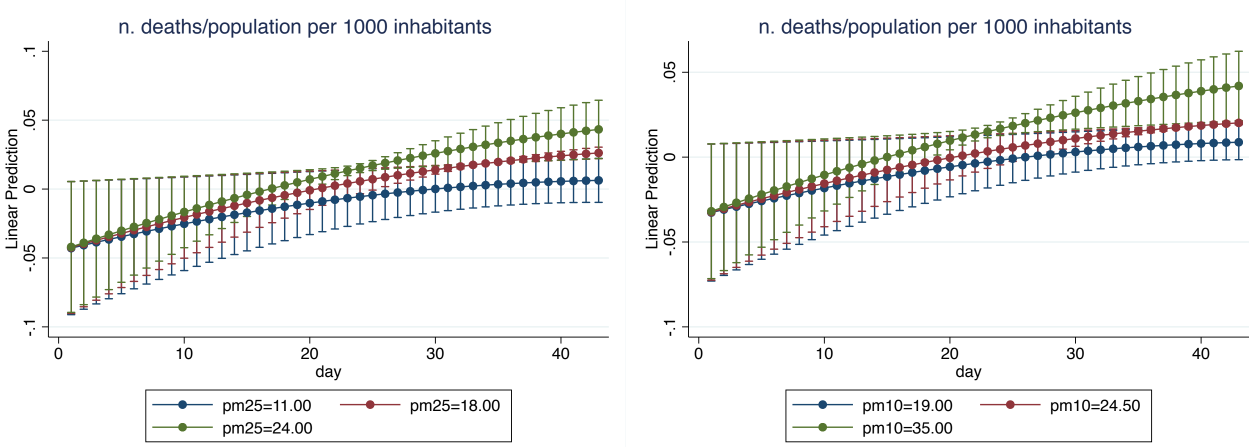
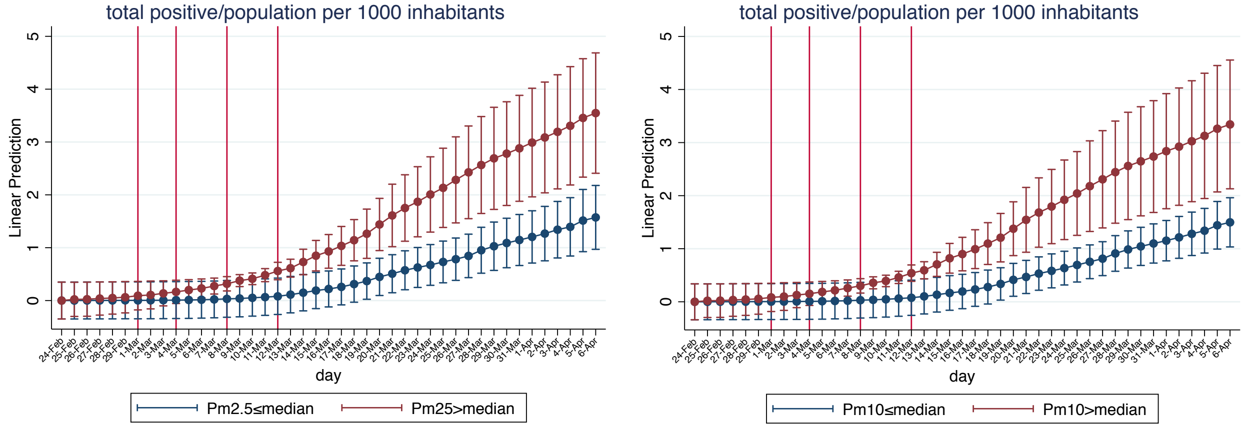
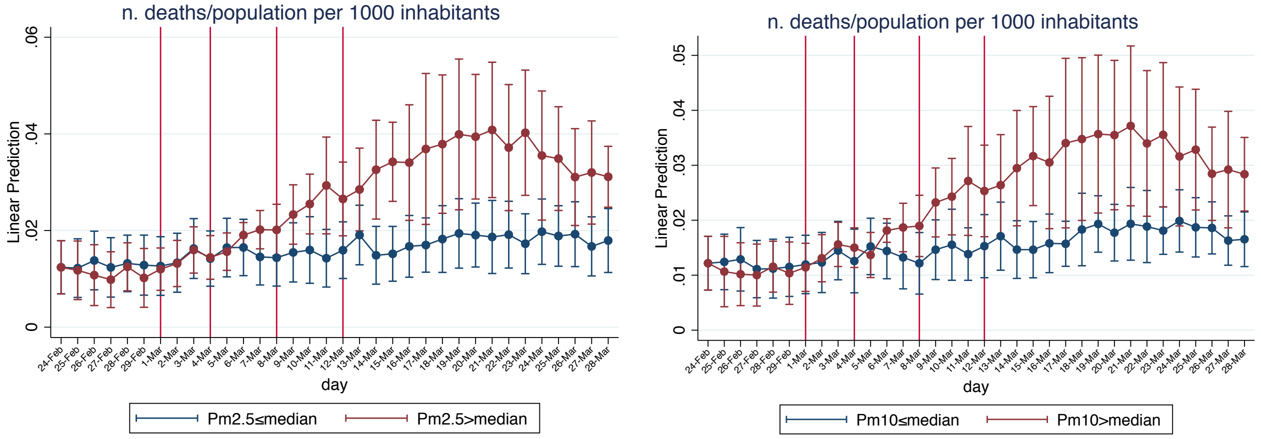


Figure 2 – COVID-19 contagion and mortality: the role of pre-existing pollution



Panel A: contagion and PM2.5

Panel B: contagion and PM10



Panel C: mortality and PM2.5

Panel D: mortality and PM10

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*
New_cases_pc	Number of daily new COVID-19 cases (at province level) over total population, per 1,000 inhabitants)
Deaths_pc	Average number of daily 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)
PM2.5	Average of yearly mean values in mg/mc registered by city monitoring posts in the i-th province (ISPRA 2018)
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.
Ventilators	Number of lung ventilators 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	*Variation*	*Mean*	*Std. Dev.*	*Min*	*Max*	*Observations*
New_cases_pc	overall	0.051	0.088	-0.108	0.951	N = 3,506
	between		0.057	0.003	0.297	n = 95
	within		0.067	-0.234	0.705	T = 36.905
Deaths_pc	overall	0.0193	0.0274	0	0.201	N = 2,883
	between		$2.32e^{-02}$	0.00058	0.110	n = 91
	within		0.0145	-0.0733	0.121	T = 31.6813
Day	overall	42.983	10.686	25.000	61.000	N = 3,506
	between		0.103	42.091	43.000	n = 95
	within		10.686	24.983	61.892	T = 36.9053
Lockdown	overall	0.770	0.421	0.000	1.000	N = 3,506
	between		0.024	0.727	0.811	n = 95
	within		0.420	-0.041	1.043	T = 36.9053
PM10	overall	25.231	6.076	12.000	40.000	N = 4,171
	between		6.101	12.000	40.000	n = 67
	within		$7.99e^{-15}$	25.231	25.231	T = 43
PM2.5	overall	17.229	5.169	6.00	17.30	N = 2,729
	between		5.200	6.00	17.30	n = 74
	within		$3.70e^{-15}$	17.229	17.229	T = 36.878
High temperature	overall	0.019	0.135	0.000	1.000	N = 3,506
	between		0.094	0.000	0.811	n = 95
	within		0.097	-0.792	0.992	T = 36.9053
Density	overall	263.747	351.832	37.166	2623.520	N = 3,506
	between		353.262	37.166	2623.520	n = 95
	within		0.000	263.747	263.747	T = 36.9053
Over65	overall	236.812	24.844	173.927	290.665	N = 3,506
	between		24.997	173.927	290.665	n = 95
	within		0	236.812	236.812	T = 36.9053
Income	overall	0.111	0.072	0.011	0.406	N = 3,506
	between		0.072	0.011	0.406	n = 95
	within		0	0.111	0.111	T = 36.9053
Ventilators	overall	0.000	0.000	0.000	0.002	N = 3,506
	between		0.000	0.000	0.002	n = 95
	within		0.000	0.000	0.000	T = 36.9053
Public transport use	overall	0.167	0.195	0.010	1.397	N = 3,506
	between		0.196	0.010	1.397	n = 95
	within		0.000	0.167	0.167	T = 36.9053
Internal commuting	overall	0.433	0.0488693	0.286	0.577	N = 3,506
	between		0.0491291	0.286	0.577	n = 95
	within		0	0.433	0.433	T = 36.9053
External commuting	overall	0.035	0.0209099	0.004	0.113	N = 3,506
	between		0.0210198	0.004	0.113	n = 95
	within		0	0.035	0.035	T = 36.9053
Artisan	overall	0.268	0.0612932	0.118	0.382	N = 3,506
	between		0.0616517	0.118	0.382	n = 95
	within		$1.48e^{-16}$	0.268	0.268	T = 36.9053

Table 4 – Major factors explaining variation in COVID-19 contagion

Dep. var: *new cases pc*	(1)	(2)	(3)	(4)
	Cross-section OLS (tot. new cases as on March, 31st)		Pooled OLS	
Day			0.0126**	0.0133**
			(0.00473)	(0.00483)
Day^2			-0.000103**	-0.000109**
			(4.82e-05)	(5.00e-05)
Lockdown			-0.0258***	-0.0265***
			(0.00789)	(0.00739)
PM10	0.00332**		0.00298**	
	(0.00135)		(0.00122)	
PM2.5		0.00388**		0.00390**
		(0.00157)		(0.00155)
High temperature	0.00678	0.00805	-0.00678	0.000366
	(0.0153)	(0.0146)	(0.0114)	(0.0119)
Density	-5.29e-06	-9.23e-06	-2.61e-06	-5.35e-06
	(1.84e-05)	(1.99e-05)	(8.29e-06)	(9.51e-06)
Over65	-0.000662	-0.000906*	-0.000722**	-0.000808**
	(0.000449)	(0.000518)	(0.000288)	(0.000301)
Income	0.332*	0.438**	0.148**	0.168***
	(0.185)	(0.188)	(0.0599)	(0.0498)
Ventilators	-37.03	-32.23	-7.316	-1.922
	(28.86)	(46.48)	(20.12)	(31.64)
Public transport use	0.0678	0.0847*	0.0103	0.00976
	(0.0412)	(0.0422)	(0.0161)	(0.0181)
Internal commuting	0.271	0.217	0.0433	-0.000811
	(0.189)	(0.211)	(0.0653)	(0.0684)
External commuting	-0.187	-0.0521	-0.134	-0.185
	(0.392)	(0.410)	(0.187)	(0.209)
Artisan	0.469***	0.487**	0.585***	0.581***
	(0.142)	(0.171)	(0.106)	(0.106)
Observations	95	76	3,506	2,803
R-squared	0.400	0.426	0.338	0.330

Robust standard errors clustered at regional level in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5 – Major factors explaining variation in COVID-19-related mortality

Dep. var: *Deaths pc*	(1)	(2)	(3)	(4)
	Cross-section OLS (tot. new cases as on March, 28th)		Pooled OLS	
Day			0.00186	0.00233
			(0.00130)	(0.00152)
Day^2			-1.59e-05	-2.05e-05
			(1.27e-05)	(1.50e-05)
Lockdown			-0.00141	-0.00198
			(0.00235)	(0.00261)
PM10	0.00123**		0.00114**	
	(0.000549)		(0.000475)	
PM2.5		0.00152		0.00134**
		(0.000890)		(0.000606)
High temperature			-0.00150	0.00161
			(0.00470)	(0.00395)
Density	3.90e-06	3.15e-06	3.67e-06	2.98e-06
	(3.65e-06)	(4.12e-06)	(2.56e-06)	(3.10e-06)
Over65	-2.90e-05	-6.78e-05	-8.82e-05	-0.000139
	(0.000147)	(0.000154)	(0.000129)	(0.000132)
Income	0.0280	0.0240	0.0420	0.0309
	(0.0348)	(0.0429)	(0.0267)	(0.0308)
Ventilators	-8.627	-4.145	-3.563	3.323
	(7.329)	(14.42)	(4.832)	(10.92)
Public transport use	0.00481	0.00820	-0.000850	-0.00215
	(0.0145)	(0.0205)	(0.0115)	(0.0163)
Internal commuting	-0.0274	-0.0605	-0.0423	-0.0727**
	(0.0477)	(0.0537)	(0.0273)	(0.0298)
External commuting	-0.161	-0.174	-0.0827	-0.0777
	(0.0980)	(0.124)	(0.0535)	(0.0617)
Artisan	0.267***	0.288***	0.217***	0.243***
	(0.0367)	(0.0282)	(0.0477)	(0.0485)
Observations	87	68	2,970	2,354
R-squared	0.405	0.365	0.310	0.298

Robust standard errors clustered at regional level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6 – Major factors explaining variation in mortality and contagion (*fixed effects OLS*)

	(1)	(2)	(3)	(4)
Dep. Var:	*New_cases_pc*		*Deaths_pc*	
Day	0.00211	0.00447	0.000350	0.00166**
	(0.00400)	(0.00394)	(0.000610)	(0.000746)
Day^2	-0.000106**	-0.000114**	-1.88e-05*	-2.39e-05*
	(4.44e-05)	(4.70e-05)	(1.03e-05)	(1.25e-05)
Lockdown	-0.0274***	-0.0288***	-0.00241	-0.00327*
	(0.00642)	(0.00639)	(0.00144)	(0.00184)
High temperature	-0.00439	-0.00327	-0.00197*	-0.00155*
	(0.00356)	(0.00435)	(0.00103)	(0.000745)
Day*PM10	7.40e-05**		4.82e-05**	
	(3.14e-05)		(2.27e-05)	
Day*PM2.5		0.000103**		6.62e-05*
		(4.89e-05)		(3.38e-05)
Day*Density	3.21e-07	1.44e-07	6.22e-08	2.06e-10
	(4.40e-07)	(4.39e-07)	(1.01e-07)	(1.18e-07)
Day*Over65	-1.59e-05	-2.02e-05	-7.55e-06*	-9.04e-06**
	(1.54e-05)	(1.77e-05)	(4.09e-06)	(4.05e-06)
Day*Income	0.00690*	0.00855**	0.00204**	0.00179
	(0.00342)	(0.00348)	(0.000881)	(0.00111)
Day*Ventilators	-0.370	-0.356	-0.416*	-0.265
	(1.236)	(1.527)	(0.204)	(0.403)
Day*Public transport use	2.04e-06	0.000404	5.68e-05	8.77e-05
	(0.000750)	(0.000688)	(0.000457)	(0.000723)
Day*Internal commuting	0.0156*	0.0152	0.000629	-0.000515
	(0.00841)	(0.00963)	(0.00173)	(0.00175)
Day*External commuting	-0.0130	-0.0152	-0.00722*	-0.00798*
	(0.0113)	(0.0131)	(0.00383)	(0.00416)
Day*Artisan	0.0210***	0.0192**	0.00819***	0.00862***
	(0.00534)	(0.00721)	(0.00195)	(0.00156)
Observations	3,506	2,803	2,970	2,354
R-squared	0.313	0.313	0.267	0.282
Number of provinces	95	76	91	72

Robust standard errors clustered at regional level in parentheses
*** p<0.01, ** p<0.05, * p<0.1