Internal migration and the spread of Covid-19*

Michele Valsecchi† Ruben Durante‡

October 9, 2020

Abstract

How does internal migration affect the spread of a pandemic? Looking at the case of Italy and using data on the province of origin of migrants located in outbreak areas, we document that provinces more exposed to the virus experience higher mortality in post-outbreak weeks, even when comparing provinces within the same region. We calculate that, had all non-outbreak provinces been as exposed as the one at the lowest decile of the exposure distribution, they would have experienced 60% fewer COVID-19 deaths. Additional evidence from phone records data confirms that the effect is mainly driven by increased mobility from outbreak areas.

Keywords: internal migration; mobility; health; epidemic; Covid-19.

JEL classification: J61; R23; H12; I10.

---

*We thank Ruben Enikopolov, Tatiana Mikhailova, Maria Petrova, Gerhard Toews and participants at the NES brown bag seminar for useful comments, Giorgio Gulino for providing some of the province level controls and Elliot Motte for assistance with the mobility data. The project was registered on the EEA Registry of COVID-19 Economic Outcome Research Projects on the 9^{th} April. An earlier version of this study was circulated with the same title on CEPR Covid Economics (Issue 18, 15 May 2020). The preliminary idea appeared in Mikhailova and Valsecchi (March 2020). All errors are our own.

†New Economic School (NES). Corresponding author. E-mail: mvalsecchi@nes.ru

‡ICREA-UPF, IPEG, Barcelona GSE, and CEPR. E-mail: ruben.durante@upf.edu

1. Introduction

Covid-19 has claimed over 600,000 lives worldwide and has imposed unprecedented economic damage (WHO, 2020).[1] Though virtually no country was spared by the pandemic, its impact has varied considerably across countries, with some managing to respond much more effectively than others. A number of studies have investigated what factors may explain these differences, from the adoption of different policies (e.g., Hsiang et al., 2020, Fang et al. 2020) to cultural factors that may have contributed to the success of social distancing measures (Durante et al., 2020; Barrios et al., 2020; Giuliano and Rasul, 2020).

One aspect that remains largely unexplored concerns the way in which the virus spread across space from the first limited outbreaks to entire countries, and, in particular, the role that internal migration may have played in this context. Extensive anecdotal evidence and news reports indicate that, following the early outbreaks and the subsequent economic slowdown, many internal migrants decided to return to their place of origin (World Bank 2020, Zenner and Wickramage 2020, Kwawa 2020). This was especially true for recent migrants who maintain strong ties with their place of origin but have a limited support network in the destination place. In some cases, the decision to "return home" was prompted by the announcement by local authorities of the imposition of "red zones" and other mobility restrictions (Giuffrida and Tondo, 2020). In principle, the return of migrants to their hometowns from places heavily hit by the pandemic may have contributed to the spread of the virus since these individuals could have been asymptomatic carriers or may have contracted the disease while traveling. This could ultimately result in places with more returning migrants experiencing more Covid-19 cases and deaths. However, no systematic evidence exists about the plausibility and importance of this channel.

This paper attempts to fill this gap by investigating the relationship between internal migration and the spread of Covid-19. We focus on Italy, one of the countries most affected by the pandemic,[2] and traditionally characterized by considerable flows of internal migrants, most notably from southern to northern regions (Bonifazi and Heinz 2000, Mocetti and Porello 2010, Panichella 2014).

We exploit the fact that, following the first outbreak of Covid-19 in late February, most cases were concentrated in a small number of provinces in three northern regions. Using comprehensive data on the place of origin and destination of individuals having migrated over the previous years, we construct a measure of the potential for return migration from early outbreak areas to each province. We then examine whether provinces with more migrants

[1] According to the OECD, relative to the previous quarter, GDP decreased by 3.6 percent in the Euro area (1.3 percent in the US) in 2020Q1 and by 12.1 percent (9.5 percent in the US) in 2020Q2.

[2] According to the Italian Health Ministry, the number of deaths related to Covid-19 in Italy amounted to 35,507 by September 3, 2020.

potentially returning from outbreak areas, experienced a higher number of Covid-19 related deaths and higher total excess mortality over the ensuing months.

One key aspect of our analysis is that we use information on migration patterns that pre-dates the Covid-19 pandemic and cannot therefore be affected by it. Crucially, we also exploit differences between provinces in the number of migrants to specific areas that happened to be affected by the pandemic earlier on, controlling for total out-migration. This alleviates possible concerns that some particular characteristic of the province of origin (e.g., local institutions, health capacity, civic capital) may drive both migration and Covid-19 mortality. To further address this possibility, in our analysis we explicitly control for the impact of a wide range of geographic and socio-economic controls on mortality. Also, to rule out the possibility that our findings are driven by a generic North-South divide, we compare the evolution of Covid-19 mortality between provinces of the same region. Additional tests based on Oster (2019) suggest that any remaining omitted variable is unlikely to explain our results. Our main outcome of interest is the number of deaths due to Covid-19, which is available for each province in each month. However, to alleviate the concern that Covid-19 deaths may be misreported or underestimated, and that such measurement error may differ between provinces (Ciminelli and Garcia-Mandicó, 2020a, 2020b), we also consider the evolution of total (excess) mortality. This allows us to test the effect of return migration on the total number of deaths relative to those recorded in the same month of previous years. Finally, to shed light on the mechanism through which return migration influences the spread of Covid-19, we exploit data on actual mobility between provinces in the months before and after the national lockdown. These high-frequency data are based on phone-tracking records and provide a clear picture of the total number of daily displacements between outbreak areas and any province from which migrants originate.

We first look at the effect of potential return migration from outbreak areas on total (excess) mortality in provinces other than those concerned by the initial outbreaks. We find that, prior to the beginning of the pandemic, there is virtually no relationship between potential return migration from outbreak areas and total mortality. However, the relationship becomes positive and significant in March, when the pandemic starts spreading to the rest of the country, and remains so in April and May, though gradually weakening. This pattern is consistent with the overall evolution of Covid-19 cases and deaths in Italy which peaked between March and April before declining in May. The effect is sizeable: a 50 percent increase in potential return migration from outbreak areas relative to the mean is associated with 147 additional total deaths per province. We compute that, if all provinces had a potential return migration equal to that of the province at 10th percentile of the distribution, they would have experienced 7,348 fewer total deaths and 5,895 fewer Covid-19 deaths. This corresponds to

60 percent of all Covid-19 deaths in non-outbreak regions, and to 18 percent of all Covid-19 deaths in the country.

Using data from phone tracking records, we then explore to what extent the link between potential return migration and incidence of Covid-19 actually operates through an increase in mobility from outbreak areas to migrants' hometowns. We find evidence of a clear increase in mobility from outbreak areas to migrants' places of origins precisely in the two weeks between the first outbreak and the lockdown, which is consistent with migrants returning to their place of origins before the national lockdown and with the timing of the effect on Covid-19 deaths. We estimate that mobility from outbreak areas accounts for most of the effect on mortality that we document.

Taken together our findings corroborate the view that the mobility of internal migrants in the early phases of the pandemic contributed to the spread of the virus and ultimately increased its death toll, particularly in areas that would have otherwise been less exposed. In terms of policy recommendations, our results stress the importance of putting in place systems and protocols to map the network of internal migration routes, and to closely monitor and limit the mobility of internal migrants in case new outbreak emerge. In addition to mobility restrictions, other measures could be adopted to financially support internal migrants that are more vulnerable to the economic consequences of the pandemic, and to increase their awareness about the risks of returning to their place of origins for themselves, their families, and the community at large.

Our work contributes to several streams of literature. First, it relates to previous extensive research on the economic impact of internal mobility. Studies have examined how, by reinforcing economic connections between different areas of a country, internal migration can boost aggregate productivity (Bryan and Morten 2019). Internal mobility helps dilute local negative economic shocks by redistributing migrants to other locations (Monras, 2018). In case of local negative health shocks, we show that internal migration helps aggravate the shock. For migrants' hometowns, having people scattered elsewhere was an insurance against their own negative economic shocks (Gröger and Zylberberg 2016).[3] We show instead that during a pandemic migrants are a liability, because they increase risks in their places of origin.

Our paper also relates to previous historical research on the impact of pandemics, namely of the Black Death. Extensive historical accounts indicate that, when the plague started to spread, cities became death traps and many escaped to the countryside to avoid conta-

[3] Other studies have focused specifically on how return migration affects the places of origins by fostering local development (Chauvet et al 2015), by favoring entrepreneurship (Yang 2008) and, more generally, a positive change in attitudes and beliefs (Chauvet and Mercier 2014, Barsbai et al 2017, Spilimbergo 2009, Mercier 2016, Grewal 2020, Clingsmith et al 2009).

gion (Boccaccio 1352, Woods 2003, Clark and Cummins 2009, Voigtländer and Voth 2013, Carmichael 2014). Although people escaping cities might have gone back to the rural places they came from, carrying the virus with them, the impact of return migration in the context of the Black Death has not been studied. [4]

Finally, our paper contributes to the growing literature on the diffusion and effects of the Covid-19 and other pandemics. This body of work can be divided into three strands.

The first one attempts to estimate the effect of various measures adopted to stop the spread of pandemics such as the closure of schools and public transportation (Adda, 2016; Litvinova et al., 2019; Fang et al., 2020), lockdowns (Chinazzi et al. 2020; Kraemer et al., 2020), or the combination of several of them (Gatto et al., 2020). The second one looks instead at what factors affect compliance with social distancing measures, with particular regard to the role of expectations (Briscese et al., 2020) and cultural traits (Durante et al., 2020; Barrios et al., 2020; Egorov et al., 2020). The third one focuses on what other factors favor the diffusion of the virus including railways (Adda 2016), trade (Oster 2012), paid sick leave (Barmby and Larguen 2009; Pichler and Ziebarth 2019), and social media connections (Kuchler et al 2020). Our study is novel in that it proposes and tests an alternative mechanism of diffusion - i.e., through return internal migration - and shows it had a significant effect on the overall number and geographic distribution of casualties.

2. Background: Covid-19 outbreak in Italy

Italy has been the first Western country to be heavily hit by the Covid-19 pandemic and to implement large-scale measures to contain it. The first two confirmed cases of Covid-19 in the country were recorded on January 30, 2020, while the first death on February 21. On the same day the first COVID-19 hotspot was identified near the town of Codogno (Lombardy). The government responded by establishing a "red zone" around the town to restrict mobility into and from the area. This episode, and the extensive media coverage it received, represented a turning point in Italians' perception of the threat of an epidemic in the country. Other measures quickly followed. On February 24 the government ordered the closure of all schools in the northern regions of Lombardy, Veneto, Emilia-Romagna and Friuli-Venezia Giulia. On March 8 the "red zone" was extended to the region of Lombardy and to 14 provinces in the regions of Piedmont, Emilia-Romagna, and Veneto accounting for over 16 milion residents. On March 11 the lockdown was extended to the entire country, and on March 22 it was further tightened with the closure of all non-necessary economic

[4] Another historical episode in which a similar mechanism might have been relevant is the return of soldiers from the front at the end of WWI (1914-1918) right at the time the Spanish Flu was starting to spread. See Beach et al. (2018) and references therein for recent work on the Spanish flu using micro-data. See Barro et al (2020) for a recent cross-country analysis.

activities. In the ensuing weeks, the number of cases (deaths) increased steadily, reaching 106,000 (12,000) by the end of March, 205,000 (28,000) by the end of April, and 233,000 (33,000) by the end of May.[5]

During the health crisis, the Government took unprecedented measures, which started with an initial lock-down in the province of Lodi (21^{st} February)[6] and school closure in Lombardia, Veneto, Emilia-Romagna and Friuli-Venezia Giulia (24 February), continued with the expansion of the lock-down to most of northern areas (08^{th} March) for a total coverage of 16 million people, and finally reached the national level (9^{th} March).[7]

As the news of the expansion of the lock-down leaked (Saturday 7^{th} March),[8] people rushed to take night trains from Milan to the rest of the country to escape the quarantine measures. It was common wisdom in the media at that time that such mass departure would have helped to spread the disease.[9]

People leaving Milan are presumably internal migrants who had come to Milan for study or work. Hence, the higher the number of out-migrants (to Lombardy) a region has, the higher should be the number of return migrants the same region experienced on the 8^{th} March (or later), and the higher should be the number of infected cases and deaths by Covid-19 later on.

3. Data and research design

3.1. Data on Covid deaths and identification of outbreak areas

To measure the number of Covid-19 deaths,[10] We use daily data from the Italian Ministry of Health elaborated by the Department of Civil Protection.[11]

These data are available in two forms: regional-daily level and provincial-"monthly" level (20 20^{th} Feb.-31^{st} Mar., April, May).[12]

To identify outbreak areas, we take the following steps. First, we use regional daily data on Covid deaths to identify which regions experienced Covid deaths first. These are Veneto

[5] Data from the Italian Ministry of Health elaborated by the Department of Civil Protection and described in the next section.

[6] Decreto del presidente del consiglio dei ministri 22 febbraio 2020

[7] Decreto del presidente del consiglio dei ministri 09 marzo 2020

[8] Severgnini (8 March 2020). Corriere della Sera

[9] Giuffrida and Tondo (8 March 2020). The Guardian.

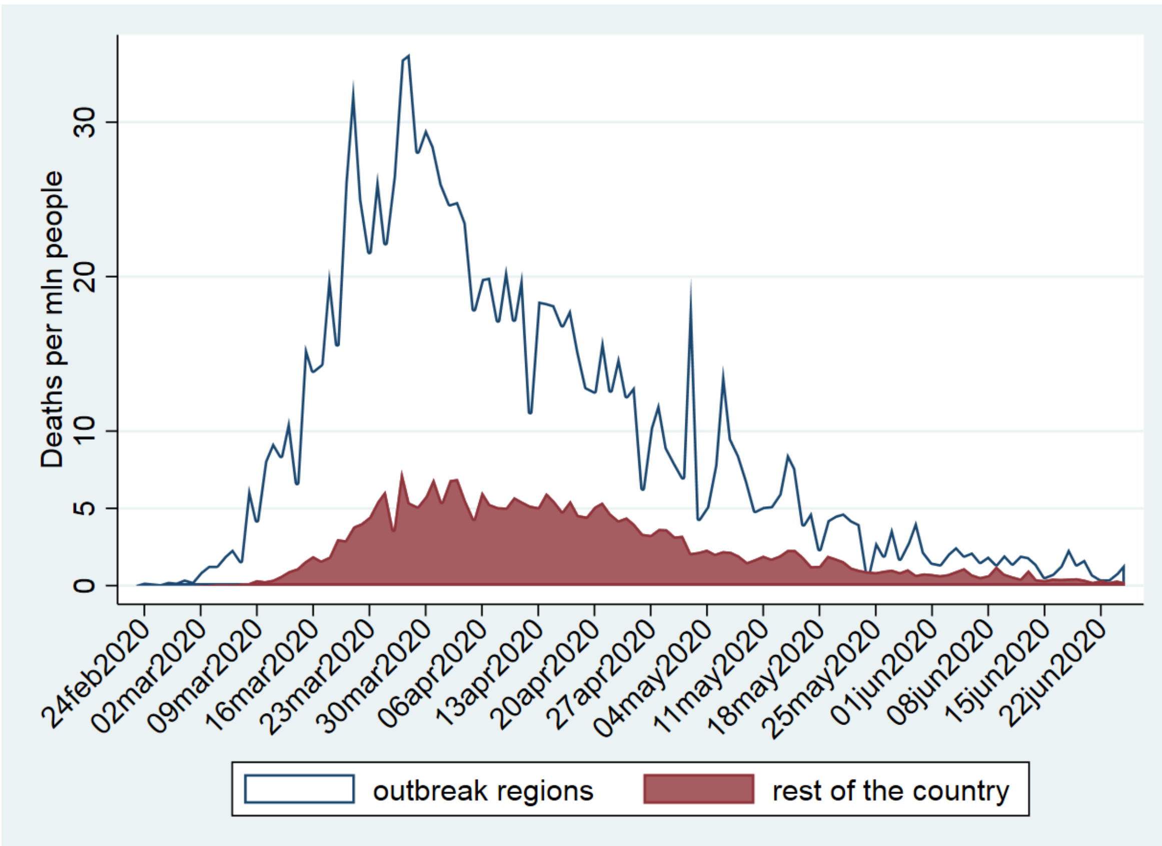
[10] Throughout the entire paper, we focus on Covid deaths, rather than Covid infections, because the Italian government, along many other central governments around the world, tested primarily people showing symptoms of infection, rather than pursuing quasi-random testing (as in Iceland and South Korea).

[11] Data and description are available at https://github.com/pcm-dpc/COVID-19

[12] Provincial-monthly level data also include data on total deaths, both in levels and in growth rates. We will discuss these additional data carefully in the section on descriptive statistics.

(21st),[13] Lombardia (22nd February)[14] and Emilia-Romagna (26th February).[15] Figure 1 shows the evolution of Covid deaths over time: Covid deaths peaked around the 30th of March and slowly decreased thereafter.

FIGURE 1: COVID-19 DEATHS IN OUTBREAK REGIONS AND IN THE REST OF THE COUNTRY



Notes: outbreak regions are Lombardy, Veneto and Emilia-Romagna.

Second, we compute the proportion of 20th Feb. - 31th March Covid deaths in a province relative to all deaths within the region for the same period. We use these shares to disaggregate the 20th Feb. - 31th March regional daily data into provincial daily data. Third, we pick the latest date before the appearance of Covid deaths outside the three outbreak regions, *i.e.*, the

[13] See Custodero (Repubblica, 22nd February 2020)

[14] See TgCom24 (22nd February 2020).

[15] Data from Department of Civil Protection described at the beginning of the section. All other Italian regions experienced their first Covid deaths several days or weeks later: Marche on the 2nd of March, followed by Liguria on 3rd, Puglia on 4th, Piemonte on the 5th, Lazio on 6th, Friuli V.G. on the 8th, Toscana on the 9th and Abruzzo on the 10th.

1^{st} of March.[16] Figure A.1 shows the distribution of Covid deaths per capita across the 28 provinces in these regions. Fourth, we pick one Covid death per million people as the cut-off defining the outbreak provinces. This leaves us with 15 provinces: 10 in Lombardy and 5 in Emilia-Romagna.[17]

3.2. Data on mobility

To measure trips from outbreak provinces to the rest of the country, we use data on mobility based on mobile phone tracking data provided by Teralytics. The data are available at the province of origin - province of destination - day level. For one of the robustness checks, we will also use trips within a given province.

3.3. Data on internal migration and exposure to outbreak areas

To measure the exposure of Italian provinces to outbreak areas, we use yearly data on changes of residence between Italian provinces.[18] The data are available up until 2018 and are structured as a matrix, *i.e*, for a given year, they provide the number of people who de-registered themselves from, say, Catania province (Sicily), and registered themselves in the province of Milan (Lombardy) during the previous 12 months.[19] We focus on changes of residence that took place between 2015 and 2018 and divide them by the 2018 population of the province (or region) of origin. This is our *ExposureToOutbreak* indicator.

To control for general propensity to emigrate from a province (region), we also compute the total number of changes of residence to any other province in the country during the same period.

3.4. Descriptive statistics

Figure 1 shows the evolution of Covid deaths in non-outbreak regions. Covid arrived later and had a much lower intensity throughout the entire period, even though there is a convergence over time between outbreak and non-outbreak regions.

---

[16] Covid deaths are the most reliable proxy for Covid presence, but they are a lagged proxy, because it takes time for an infection to degenerate and bring someone to death. Detailed reports from the Italian Istituto Superiore della Sanitá on the characteristics of deceased Covid patients ("Caratteristiche dei pazienti deceduti positivi all'infezione da SARS-CoV-2 in Italia") indicate that Covid patients experienced their first symptoms 11 days before their deaths. Hence, the stock of Covid deaths on the 1^{st} of March proxies Covid diffusion on the 19^{th} February.

[17] In one of the robustness checks, we estimate the effect using all 28 provinces as outbreak areas.

[18] Data provided by the Italian National Institute of Statistics (ISTAT).

[19] People have an incentive to register themselves in their new residence to get access to some basic services like, among others, the family doctor.

Table A.1 and Table A.2 show detailed descriptive statistics at the regional and provincial level for 16 non-outbreak regions and 76 non-outbreak provinces.[20] At the regional level, the number of Covid deaths per million inhabitants is 2.51 per day. At the provincial level, it is 87 between the 20^{th} of February and the 31^{st} of March, 155 in April and 45 in May. This rise-and-decline mirrors the pattern shown in Figure 1.

Important complements to the Covid deaths are the data on total deaths, which have the key advantage of being available even before the appearance of Covid. Data on total deaths are available in two forms. First, they are available in levels for the 20^{th}Feb-31^{st}Mar. period, averaged over 2015-2019 and, separately, for 2020. When looking at 2020, the total number of deaths is of course much higher than the number of Covid deaths: 1,292 instead of 87. This is not surprising, but it highlights how demanding could be to detect an effect on Covid deaths when looking at total deaths. Second, total mortality data are available as growth rates for January-February, March, April and May. Averages suggest that total deaths declined during January and February (-7 percent), then increased in March (+20 percent) and April (+17 percent) and declined again in May (-5 percent). This is consistent with the rise-and-decline of Covid deaths shown in Figure 1.

The number of daily trips from outbreak areas is similar across the two datasets and averages 3-4 trips per 1000 inhabitants. Provincial data suggest that trips were relatively high during the pre-Covid period (6.09), declined partially during the post-outbreak & pre-lockdown period (4.5) and finally fell drastically during the lockdown (1.15).

Non-outbreak provinces have an average of 4.48 migrants (per 1000 inhabitants) to outbreak provinces. Two features of such migration are important for our identification strategy. First, migration to outbreak areas shows substantial variation, as it ranges from 1.59 to 11.46 with a standard deviation of 2.04. Figure A.2 shows the distribution of exposure across provinces outside the outbreak regions. Second, migration to outbreak provinces constitutes only a small fraction of overall migration (28.08). This will allow us to estimate the effect of exposure to outbreak areas keeping general propensity to emigrate constant.

3.5. Econometric specification

Given the type of available data, we will run two types of analysis: one that uses the region as unit of analysis; and another that uses the province.

[20] We drop: Sud Sardegna province, which was aggregated and disaggregated repeatedly during the past years and therefore has inconsistent migration data; Gorizia province, which is not well covered by the mobility data; and Valle D'Aosta region, which has only one province and therefore gets dropped out in the specifications with region fixed effects.

The analysis at the regional level takes the following form:

$$CovidDeaths_{r,date} = \alpha + d_{date} + \sum_{week} \beta_{week} \left[ln(ExposureToOutbreak_r) \times d_{week} \right] + X'_{r,week}\Gamma + \varepsilon_{r,date} \tag{1}$$

where $CovidDeaths_{r,date}$ is the number of Covid deaths in region r in a given day, $ExposureToOutbreak_r$ is our exposure indicator, and $X_{r,week}$ is a set of interactions between (pre-determined) controls and week indicators, and $\varepsilon_{r,date}$ is the error term. Observations are weighted by population.

The analysis at the regional level has one key limitation. There are only 16 regions in Italy outside the outbreak regions.[21] Hence, we can only control for a limited number of covariates without saturating our specification.

The list of pre-determined controls includes log distance to outbreak areas, social capital, state and health capacity, share of population at risk and general propensity to migrate. Social capital is proxied by the first principal component of the share of people with high school (or higher), the share of people with university education, newspaper readership (at least once and at least five times a week) and the share of people trusting others. State and health capacity is proxied by the first principal component of regional GDP per capita, unemployment and the number of intensive care beds per 100,000 inhabitants. Population at risk is the share of people above 70 years old. General propensity to migrate is the log number of people who changed their residence from region r to any other region in the country during 2015-2018 (per 1000 inhabitants). Standard errors are clustered at the regional level and adjusted for few clusters using Cameron, Gelbach and Miller (2008).

The analysis at the provincial level complements the regional analysis. The time dimension is coarser than the daily level. However, there are many more units (76), which allow us to control for a much wider set of covariates *and* for region fixed effects:

$$Deaths_{r,p} = \alpha_r + \beta ln(ExposureToOutbreak_{r,p}) + X'_{r,p}\Gamma + \varepsilon_{r,p} \tag{2}$$

where $Deaths_{r,p}$ is a measure of deaths in region r and province p, $ExposureToOutbreak_{r,p}$ is our exposure indicator, α_r is a set of region fixed effects, $X_{r,p}$ is a set of (pre-determined) controls, and $\varepsilon_{r,p}$ is the error term. Observations are weighted by population.

The inclusion of region fixed effects is very important, because it controls for any cross-regional difference between more and less exposed provinces, thus restricting the comparison to provinces that have different exposure but are situated in the same region. The inclusion

[21] Besides the outbreak regions (Lombardia, Veneto and Emilia-Romagna), we drop also Valle D'Aosta to ensure consistency between the regional and the provincial level analysis. The reason for dropping Valle D'Aosta from the provincial level analysis will become obvious later in the section.

of provincial controls ensures that such within-region comparison is not biased by potential confounders such as the level of economic development, the local health capacity or risk factors that might be correlated also with Covid deaths.

The list of pre-determined controls is rich and includes: log distance to outbreak areas, share of people with high school education or higher, share of people with university education, number of firms per capita, value added per capita, median financial wealth, median income, number of intensive care beds per 100,000 inhabitants, share of people above 70 years old, size of the province, altitude, share of seaside cities, population density, share of males, whether there is an airport, share of urban areas, whether the province includes the regional capital, and general propensity to migrate.

The deaths measure at the province level is the number of Covid deaths (per million people) for, separately, 20^{th}Feb-31^{st}Mar., April and May. Besides that, the mortality analysis at the provincial level offers an additional key advantage: the possibility of estimating the effect on total number of deaths, which are available both before and after the appearance of Covid. First, we focus on the diffusion period (20^{th}Feb-31^{st}Mar.). We estimate the effect on total deaths for 2015-2019, which constitutes our first placebo estimation. Second, we estimate the effect on total deaths for 2020. Third, we focus on the growth of total deaths in 2020 (relative to the 2015-2019 average) at the province level for, separately, January-February, March, April and May. While we keep the right-hand side of the specification the same as eq. 2, its interpretation is now akin to that of a first difference model with the log number of deaths as dependent variable: time-invariant differences in deaths between more and less exposed provinces are now controlled for; region FEs capture time-varying differences in deaths across regions, and province controls capture differences in trends across provinces in observables.

Hence, the identification assumption is that, after controlling for all province time-invariant characteristics, for all regional time-varying characteristics, and for a whole range of provincial time-varying characteristics, there is no residual time-varying unobserved factor that is related to both exposure to outbreak areas and death rates.

To assess the plausibility of this identification assumption, we carry out the following tests. First, we test whether provinces more and less exposed to outbreak areas are similar in terms of observable characteristics. Figure A.3 shows that, after controlling for region FEs, migration exposure to outbreak areas is either uncorrelated either weakly correlated with observable characteristics proxying other determinants of Covid deaths.[22] Second, we estimate the main specification with and without controlling for all these covariates. This forms the basis

[22] Among the covariates, the only indicator that stands out is distance to outbreak areas. Panel c) and e) show that the correlation is driven by one specific region that borders the outbreak areas. In one of the robustness check, we show that the main results are robust to dropping this region.

for additional tests based on Oster (2019) that suggest that remaining unobservable factors are unlikely to drive our results. Third, and most importantly, we use data on excess mortality for January and February (i.e., before Covid) as placebo estimation. This will constitute our second and most important placebo.

Any potential remaining omitted factor would have to be consistent with the results of these estimations as well as with the results of complementary estimations for mobility. For mobility, we will use a FE model that controls for province FEs, region-week FEs and province controls interacted by week dummies, with standard errors clustered at the province level. We will discuss remaining endogeneity threats after presenting the main results.

4. RESULTS

We will now present the reduced form relationship between exposure to outbreak and deaths; the relationship between exposure to outbreak and mobility from outbreak areas; and whether / to what extent the mobility results explain the reduced form estimates. Finally, we will discuss alternative channels, provide some additional robustness and placebo tests and a back of the envelope calculation to assess the magnitude of the effects.

4.1. EFFECT ON MORTALITY

Table 1, Panel A, shows that a one percent increase in exposure is associated with 1.69 additional Covid deaths per million people during March (Column 2), 2.41 during April (Column 4) and 0.29 during May (Column 6). This implies that a variation in exposure of 50 percent relative to the mean (i.e., about one standard deviation in exposure), joint with an average population of 0.53 million people, would be associated with 45 (March), 64 (April), 8 (May) and 117 (total) additional Covid deaths per province. Notably, estimates are robust to the inclusion of a wide range of controls.

Panel B shows the results for total deaths. Column 1 shows the estimates for 20^{th}Feb.-31^{st}Mar. averaged over 2015 to 2019, *i.e.*, *before* the start of the pandemic. A one percent increase in exposure is associated with a positive but statistically insignificant effect on total deaths equal to 0.59. This effect is tiny (12%) compared to the effect we find for its 2020 counterpart (Column 2: 5.10). Interestingly, the effect for total deaths (in 2020) is also larger than the effect on Covid deaths for the same time window (Panel A, Column 2). We come back to this point after discussing the growth effects.

Columns 3-6 show the results of the growth estimations. Column 3 shows that exposure has no effect on the Jab.-Feb. growth rate, which confirms the validity of the research design. On the other hand, a one percent increase in exposure is associated with a 0.382 percentage point increase in total deaths per province in March (Column 4), a 0.203 percentage point increase

TABLE 1: EXPOSURE TO OUTBREAK AND COVID-19 DEATHS BY PROVINCE-MONTH

	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: COVID DEATHS						
Dep var.	Number of deaths 20Feb-31Mar	Number of deaths 20Feb-31Mar	Number of deaths April	Number of deaths April	Number of deaths May	Number of deaths May
ln(Exposure To Outbreak)	153.289***	169.347***	137.222*	241.456*	17.693	28.844
	(41.972)	(47.021)	(72.272)	(137.225)	(20.840)	(46.648)
Mean	86.968	86.968	155.054	155.054	45.430	45.430
R-squared	0.849	0.915	0.727	0.812	0.727	0.782
Region FEs	Yes	Yes	Yes	Yes	Yes	Yes
Province controls		Yes		Yes		Yes
PANEL B: TOTAL DEATHS						
Dep var.	Number of deaths 20Feb-31Mar 2015-2019	Number of deaths 20Feb-31Mar 2020	Growth of deaths Jan-Feb 2020 vs 2015-2019	Growth of deaths March 2020 vs 2015-2019	Growth of deaths April 2020 vs 2015-2019	Growth of deaths May 2020 vs 2015-2019
ln(Exposure To Outbreak)	59.315	509.973**	0.009	0.382**	0.203*	0.080
	(125.079)	(213.585)	(0.031)	(0.166)	(0.103)	(0.048)
Mean	1107	1292	-0.065	0.203	0.167	-0.048
R-squared	0.685	0.844	0.631	0.885	0.921	0.658
Region FEs	Yes	Yes	Yes	Yes	Yes	Yes
Province controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	76	76	76	76	76	76

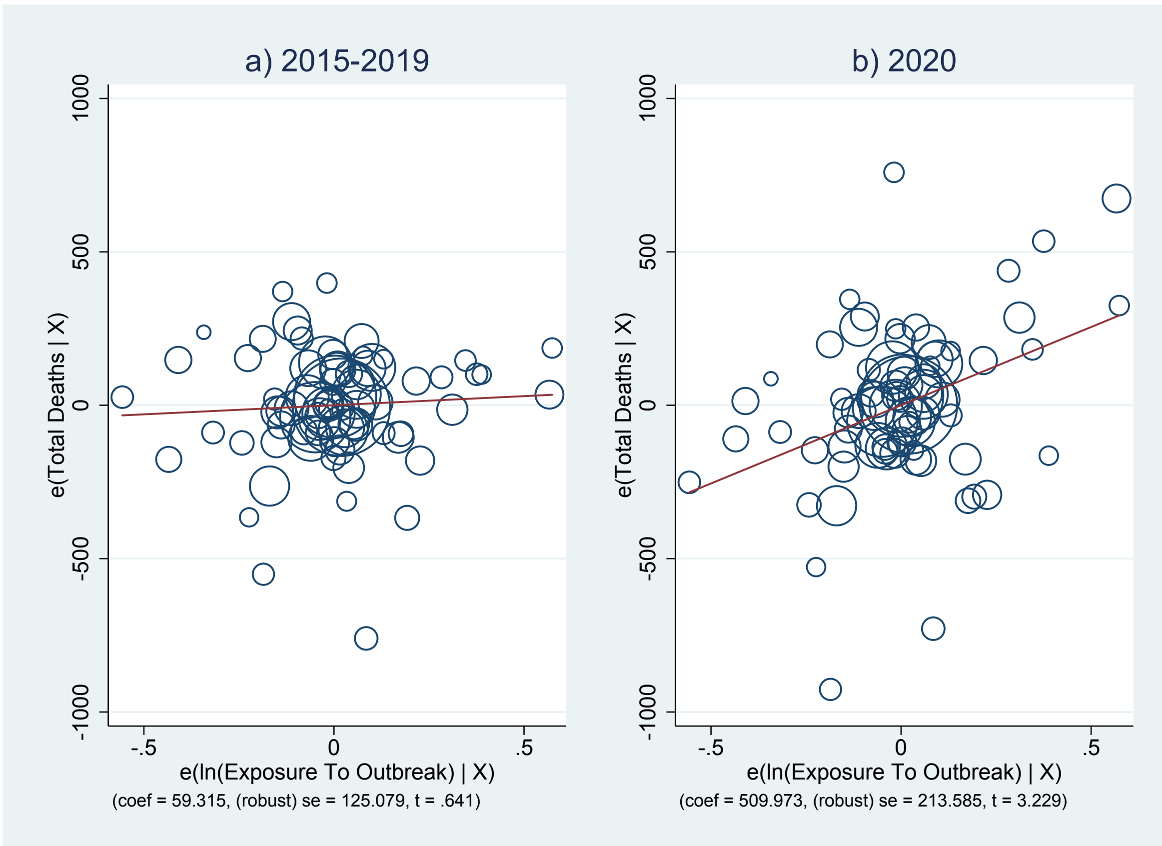
Notes: number of Covid deaths (Panel A) and total deaths (Panel B, Columns 1-2) is per million inhabitants. Growth of total deaths (Panel B, Columns 3-6) is per province. "Exposure To Outbreak" is the number of people who moved from the province to one of the outbreak areas between 2015 and 2018 (per 1000 inhabitants). Geographic controls include: log distance to outbreak provinces, number of square kilometres, altitude, share of seaside cities. Socio-demographic controls include: population density, share of males, number of intensive care hospital beds per 100,000 inhabitants, whether there is an airport, share of urban areas, population share above 70 years, population share with high school education or higher, population share with university education. Economic controls include: number of firms per capita, value added per capita, median financial wealth, median income. Total migration is the log of the number of people who moved from the province to any other area in the country between 2015 and 2018 (per 1000 inhabitants). Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

in April (Column 5) and a 0.080 percentage point increase in May (Column 6). Based on the 2015-2019 average total deaths for these months,[23] these effects corresponds to, respectively,

[23] These are 424 (March), 458 (April) and 492 (May). The 2015-2019 average for March is obtained by multiplying the 2015-2019 average for 20^{th}Feb.-31^{st}Mar. by three fourths. The 2015-2019 average for May is obtained by dividing total deaths for May 2020 by one plus the 2015-2019 growth rate for May. The 2015-2019 average for April is obtained by interpolating the averages for March and May.

1.62, 0.93 and 0.39 additional deaths. This implies that a variation in exposure of 50 percent would be associated with 81 (March), 46 (April), 20 (May) and 147 (total) additional total deaths per province.

FIGURE 2: EXPOSURE TO OUTBREAK AND NUMBER OF TOTAL DEATHS BY PROVINCE FOR 20^{th}FEB.-31^{st}MAR.

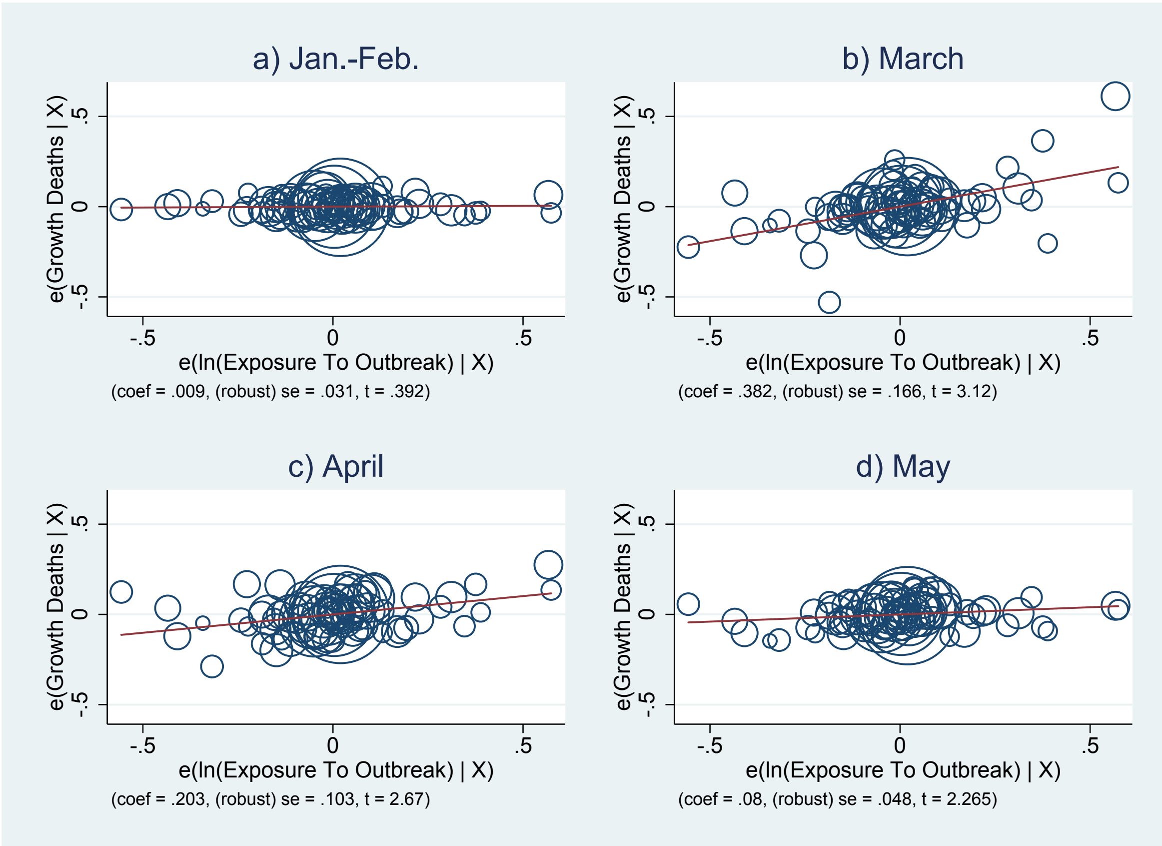


Notes: Relationship between number of total deaths per million inhabitants and $ln(ExposureToOutbreak)$ partialled out of region FEs and province controls. Circles represent province population. Panels a)-b) correspond, respectively, to Table 1, Panel B, Columns 1-2.

Figure 2 and 3 show that the results are not driven by outliers. Finally, table A.4 shows that the results are also not driven by the cutoff number of deaths used to define outbreak areas. Table A.5 shows that they are also not driven by the wide, but potentially arbitrary, choice of province controls. Following Oster (2019), we use the variation in estimates with and without province controls to get a sense of whether unobservable factors may be driving our results. For each of our main specifications, we compute a set of bounds for β. One bound is the "controlled" β (i.e., the β with full controls), while the other bound is the β obtained under the assumption that unobservables are as important as observables.[24] Table A.6, Panel

[24] Specifically, this second bound is $\beta^*(R_{max}, 1)$, i.e., the β based on $R_{Max} = 1.3 * \tilde{R}$ and a coefficient of proportionality δ equal to 1.

Figure 3: Exposure to outbreak and growth of total deaths by province-month



Notes: Relationship between growth of total deaths per province and $ln(ExposureToOutbreak)$ partialled out of region FEs and province controls. Circles represent province population. Panel a)-d) correspond, respectively, to Table 1, Panel B, Columns 3-6.

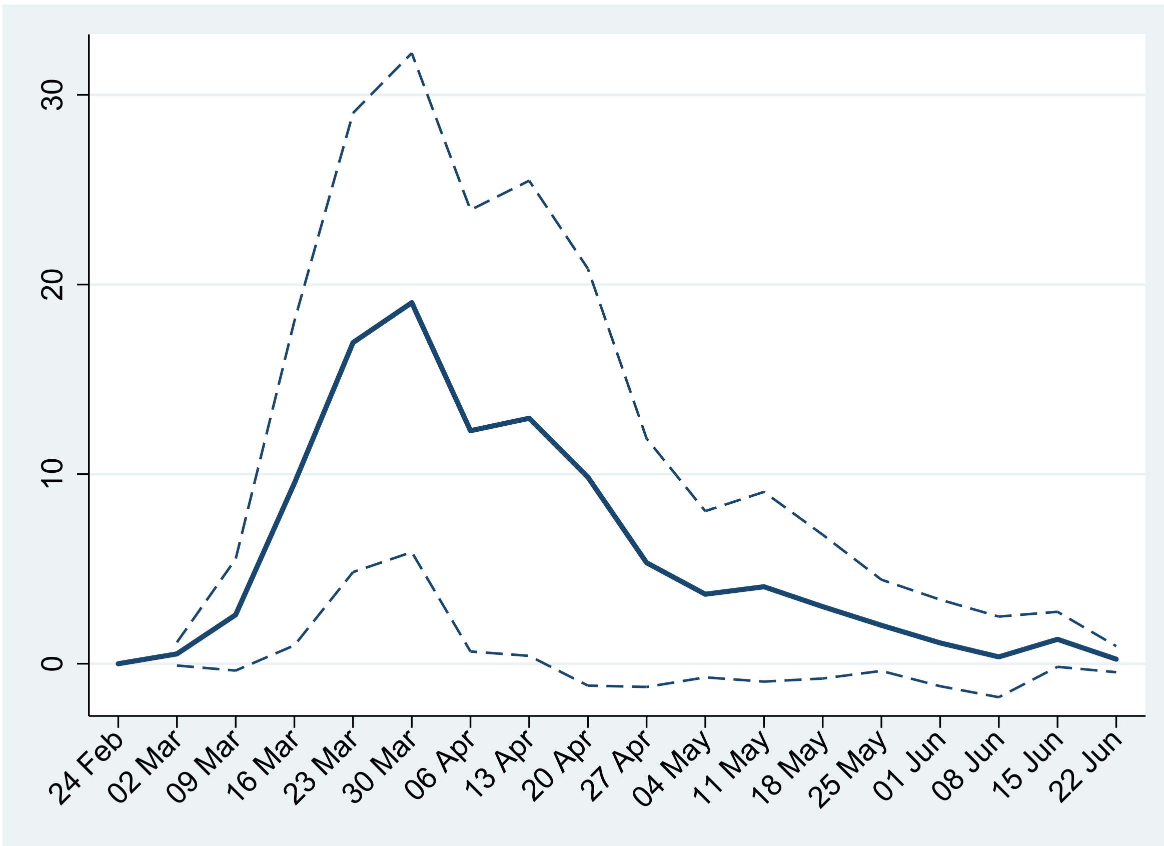
B, shows the first bound, while Panel C shows the second bound. All bounds turn out to be positive and large. Importantly, the confidence interval of the controlled β (Panel B) always includes the second bound, which indicates robustness. In addition, note that only two of the five coefficient estimates move toward zero when including province controls. For these two specifications, we therefore ask how much more important unobservables would need to be (compared to observables) to fully explain our results by omitted variable bias. This corresponds to Oster's δ. Panel D shows that selection on unobservables would need to be well above the cutoff value (i.e., δ=1) suggested by Oster and by Altonji et al (2005). Hence, it seems unlikely that our results are driven by omitted variables.

Overall, estimates for Covid deaths suggest a strong effect for March and April, followed by a near-zero effect for May, while the estimates for total deaths suggest a strong effect for March that declines smoothly in April and May. Again, the effect on total deaths is larger than the estimate for Covid deaths, although not as much as with the level estimates.

This is consistent with one or both of the following: Covid deaths are under-reported; Covid emergency caused additional non-Covid deaths, either because of crowding out of health resources and personnel, either because of quarantine measures.[25]

Next, we estimate the effect on mortality using regional-daily data on Covid deaths. Figure 4 shows the coefficient estimates associated with the interaction between exposure to outbreak and week dummies.[26] Again, the effect on Covid deaths is large in March and declines smoothly afterwards.

FIGURE 4: EXPOSURE TO OUTBREAK AND COVID-19 DEATHS BY REGION-WEEK



Notes: the figure is based on a regression of Covid deaths on exposure to outbreak areas interacted with week dummies, date FEs and various pre-determined regional characteristics interacted with week dummies and standard errors clustered at the regional level. The solid line represents the coefficient estimates associated with exposure to outbreak by week. Dashed lines represent 95 percent confidence intervals. Dates on the x-axis indicate the beginning of the week.

[25] An obvious example of quarantine related deaths would be people in need of health care who do not dare going to the hospital or simply do not take care of themselves as good as before Covid. Another example would be suicides. Needless to say, one can think of quarantine reducing other causes of death, like car accidents. Future research should investigate this finding using micro-data on causes of death at the individual or health center level.

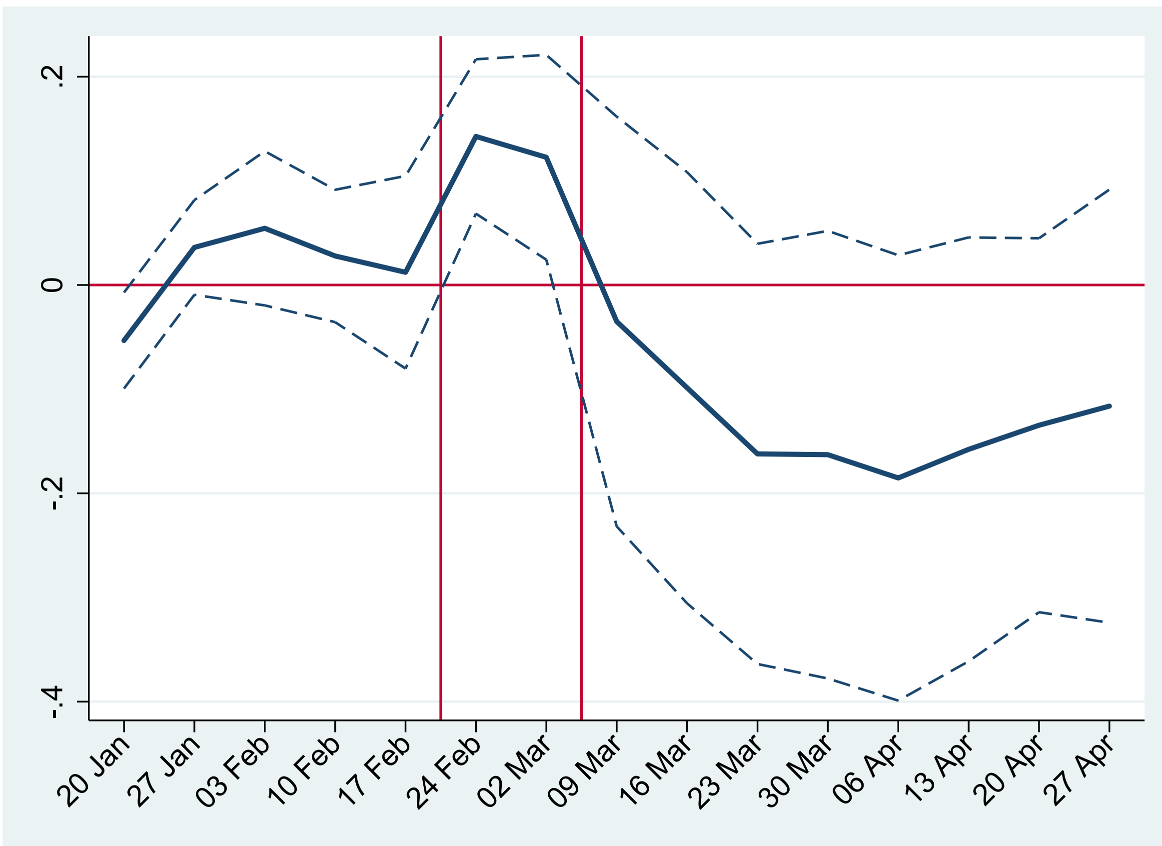
[26] See Table A.3 for the coefficient estimates associated with a specification that replaces week with "phase" dummies.

Overall, both regional-daily and provincial-monthly estimates suggest that exposure to outbreak areas is associated with an important increase in Covid deaths. The effect is large in March and April while small in May.

4.2. Effect on mobility

Next, we estimate the effect of exposure to outbreak on the number of trips from outbreak areas. Figure 5 shows the coefficient estimates associated with the specification with province FEs, region-week FEs and interactions between province controls and week dummies.

Figure 5: Exposure to outbreak and trips from outbreak areas by province-week



Notes: the figure is based on a regression of the Inverse Hyperbolic Sine (IHS) of the number of trips from outbreak areas (per 1000 inhabitants) on exposure to outbreak interacted with week dummies, date FEs, region-week FEs and the province controls (interacted with week dummies) used in Tables 1 and 2. This figure essentially is the same as Table 2, Column 6, with week dummies replacing phase dummies. Dashed lines represent 95 percent confidence intervals. Dates on the x-axis indicate the beginning of the week.

The estimates suggest the following patterns. First, there is no differential increase in trips from outbreak areas during normal times (*i.e.*, before the 24^{th}Feb.). Second, there is a differential increase in trips from outbreak areas after the outbreak but before the national lockdown (*i.e.*, 24^{th}Feb.-31^{st}Mar.). Third, there seems to be a differential decrease in trips following the national lockdown, although estimates are imprecise.

The differential increase in trips form outbreak areas following the outbreak supports the hypothesis that recent migrants returned to their hometowns following the outbreak and the shutdown of economic activities in outbreak areas. The magnitude of the effect also seems non-negligible. According to Table 2, Column 6, a one percent increase in exposure is associated with 0.14 percent additional trips from outbreak areas,[27] which implies that a 50 percent increase in exposure relative to the mean (*i.e.*, about 2 additional migrants per thousand people) would be associated with 0.3 additional daily trips per thousand people.[28]

TABLE 2: EXPOSURE TO OUTBREAK AND TRIPS FROM OUTBREAK AREAS BY PROVINCE-PHASE

	(1)	(2)	(3)	(4)	(5)	(6)
ln(Exposure To Outbreak)						
× (9^{th}Mar.-)	0.062	1.153***	1.294***	0.597***	-0.113	-0.126
	(0.306)	(0.251)	(0.275)	(0.185)	(0.141)	(0.095)
× (24^{th}Feb.-8^{th}Mar.)	-0.280	1.448***	1.559***	0.255***	0.183***	0.138***
	(0.439)	(0.206)	(0.267)	(0.090)	(0.054)	(0.038)
× (3^{rd}Feb.-23^{rd}Feb.)	-0.530	1.275***	1.458***	0.005	0.010	0.037
	(0.436)	(0.196)	(0.248)	(0.018)	(0.014)	(0.024)
× (13^{th}Jan.-2^{nd}Feb.)	-0.535	1.266***	1.420***			
	(0.446)	(0.202)	(0.254)			
Mean	0.985	0.985	0.985	0.985	0.985	0.985
R-squared	0.243	0.898	0.931	0.905	0.953	0.960
Number of clusters	76	76	76	76	76	76
Observations	8,512	8,512	8,512	8,512	8,512	8,512
Day FEs	Yes	Yes	Yes	Yes	Yes	Yes
Region × phase FEs		Yes	Yes		Yes	Yes
Province controls × phase			Yes			Yes
Province FEs				Yes	Yes	Yes

Notes: the dependent variable is the Inverse Hyperbolic Sine (IHS) of the number of trips from outbreak provinces (per 1000 inhabitants). "Exposure To Outbreak" is the number of people who moved from the province to one of the outbreak areas during 2015-2018 (per 1000 inhabitants). Province controls are the same as in Table 1. Standard errors clustered at the province level in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 2 confirms these patterns and suggests an additional one: more exposed provinces receive more trips from outbreak areas also in normal times. By looking at the estimates as-

[27] See Bellemare and Wichman (2020) for a discussion of marginal effects in models including an Inverse Hyperbolic Sine transformation.

[28] This is the result of 0.138 (coefficient estimate) * 50 (percentage increase in exposure) /100 * 4.5 (average trips during this period according to Table A.2).

sociated with the specification without province FEs (Columns 1-3) and Column 3 in particular, we see that a one percent increase in exposure is associated with 1.56 percent additional trips from outbreak areas, which implies that a 50 percent increase in exposure relative to the mean would be associated with 3.5 additional daily trips per thousand people.[29] Hence, the differential increase in trips during the post-outbreak & pre-lockdown period is only 9 percent of the total marginal effect.

Exposure has an effect on both regular and additional post-outbreak & pre-lockdown trips. The total effect is large. Even if recent migrants were to travel often between their current and former province of residence, they could hardly explain the entire effect. One possible explanation is that our exposure measure captures only a fraction of the true share of migrants. Some excluded but relevant migrants could be people who moved to outbreak areas in 2015 or earlier and therefore were not included in our measure even if they registered correctly at destination.[30] Some other excluded but relevant migrants could be people who moved or travel regularly to outbreak areas without changing their province of residence, like daily commuters,[31] weekly commuters,[32] or perhaps non-resident University students who live semi-permanently where they study but have not registered themselves at destination.[33]

4.3. Effect on mortality through mobility

Next, we investigate whether and to what extent trips explain the relationship between exposure to outbreak and mortality. To do so, we include the average number of trips from outbreak areas during February as additional controls in the mortality specification at the province-month level (eq. 2). Table 3 shows the results.

Overall, trips from outbreak areas explain a large share of the reduced form effect. When looking at Covid deaths (Panel A), controlling for trips causes the main coefficient estimate of interest to decline by 41 percent for 20^{th}Feb.-31^{st}Mar. (Columns 1-2) and by more than 100 percent for April (Columns 3-4) and May (Columns 5-6). When looking at the growth of total deaths (Panel B), the coefficient estimate of interest declines by about 75 percent for

[29] This is the result of 1.559 (coefficient estimate) * 50 (percentage increase in exposure) /100 * 4.5 (average trips during this period according to Table A.2).

[30] If migration was highly correlated over time, then including earlier waves of migrants would increase our exposure index and decrease the magnitude of the effect on trips almost mechanically.

[31] Daily commuters might live in one province but travel daily to another one.

[32] Weekly commuters might live in one province during the weekend and in a different province during the week for work or study reasons.

[33] We had also speculated that that recent migration might have been positively correlated with past migration, which in turn might have generated backward and forward linkages between firms located at origin and destination. However, our mobility data track the movement of people, rather than goods, and therefore limits the ability of this mechanism to explain the results. In the next section we discuss the results of a robustness test including both recent and old migration that also supports this interpretation.

TABLE 3: EXPOSURE TO OUTBREAK, TRIPS FROM OUTBREAK AREAS, AND DEATHS

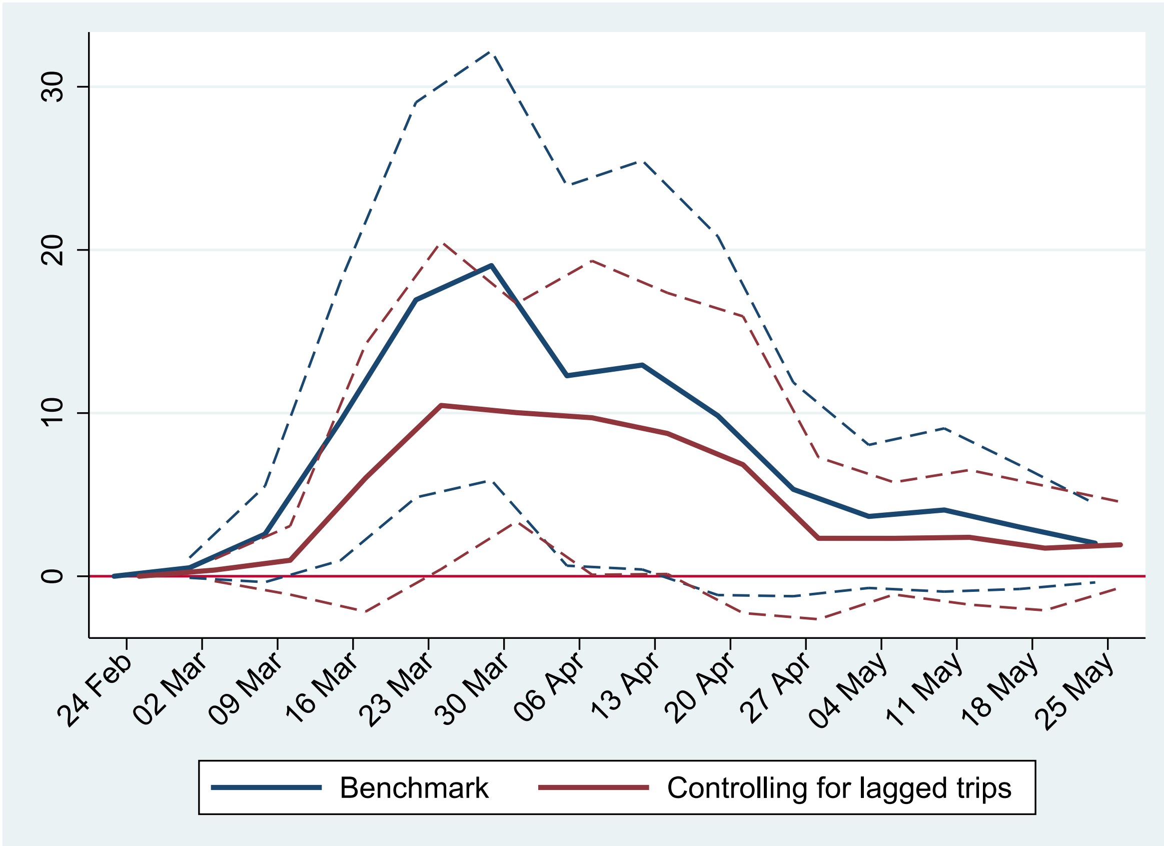
	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: Number of Covid Deaths (2020)						
Period	20Feb-31Mar	20Feb-31Mar	April	April	May	May
ln(Exposure to Outbreak)	169.347***	99.376*	241.456*	-72.239	28.844	-10.772
	(47.021)	(50.937)	(137.225)	(108.350)	(46.648)	(61.085)
IHS(# trips from outbreak areas)		47.225*		211.723***		26.738
		(27.253)		(61.626)		(22.784)
Mean	86.968	86.968	155.054	155.054	45.430	45.430
R-squared	0.915	0.923	0.812	0.884	0.782	0.790
PANEL B: Growth of Total Deaths (2020 vs 2015-2019)						
Period	March	March	April	April	May	May
ln(Exposure to Outbreak)	0.382**	0.097	0.203*	0.041	0.080	0.119
	(0.166)	(0.124)	(0.103)	(0.112)	(0.048)	(0.081)
IHS(# trips from outbreak areas)		0.193**		0.110**		-0.026
		(0.089)		(0.050)		(0.035)
Mean	0.203	0.203	0.167	0.167	-0.048	-0.048
R-squared	0.885	0.908	0.921	0.928	0.658	0.663
Region FEs	Yes	Yes	Yes	Yes	Yes	Yes
Province controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	76	76	76	76	76	76

Notes: "Exposure To Outbreak" is the number of people who moved from the province to one of the outbreak areas between 2015 and 2018 (per 1000 inhabitants). Trips from outbreak areas are per 1000 inhabitants and refer to February. Province controls are the same as in Table 1. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

March (Columns 1-2), by about 80 percent for April (Columns 3-4), while it remains stable or even increases slightly for May (Columns 5-6). The importance of mobility from outbreak areas can also be seen using the regional-daily data (Figure 6), although there the mechanism seems to matter mostly for March. Hence, the evidence supports the hypothesis that internal migration matters for the diffusion of the virus.

Ideally, we would like to separate the role of regular and "extra" trips. In practice, this requires making additional assumptions on the likelihood of contagion and the behavior at destination of the people travelling. To the extent that travelers in regular and "extra" trips are similar along these two dimensions, regular trips should be 8-9 times more important than "extra" trips. In this sense, the evidence does not support the hypothesis that the effect of internal migration on the diffusion of the virus worked through panic mobility.

FIGURE 6: EXPOSURE TO OUTBREAK, TRIPS FROM OUTBREAK AREAS, AND COVID DEATHS



Notes: the figure is based on two estimations. The first estimation corresponds to Figure 4. The second estimation is the same, except for the additional control for IHS(trips from outbreak areas per 1000 inhabitants) lagged by 4 weeks. Dashed lines represent 95 percent confidence intervals. Dates on the x-axis indicate the beginning of the week.

4.4. ALTERNATIVE CHANNELS

The previous section showed that trips from outbreak areas explained a large share of the reduced form effect. What might explain the rest of the effect? Tian et al (2020) suggest that Mexican migrants in the US influence the diffusion of the virus in their Mexican hometowns by persuading people to respect self-isolation measures over the phone.[34] We measure compliance with self-isolation measures using the number of trips within a province in a given day and test whether it is affected by exposure to outbreak areas. Table 4 shows the results.

[34] They use data on the number of migrants between each Mexican municipality and each US county, then measure compliance with self-isolation measures in each US county, and finally compute, for each Mexican municipality, the average compliance with self-isolation measures of the US localities where its emigrants are located. Hence, they measure "exposure to US self-isolation norms", while we measure "exposure to outbreak areas". In addition, in their setting, the presence of an international border between the migrants and their hometowns possibly shuts down the mechanism we investigate here.

The estimates are all close to zero and always far from statistically significant.

TABLE 4: EXPOSURE TO OUTBREAK AND WITHIN-PROVINCE MOBILITY

	(1)	(2)	(3)	(4)	(5)	(6)
ln(Exposure To Outbreak)						
$\times$ (9^{th}Mar.-)	-0.026	-0.231***	-0.030	0.069	-0.015	-0.076
	(0.151)	(0.086)	(0.125)	(0.049)	(0.056)	(0.060)
$\times$ (24^{th}Feb.-8^{th}Mar.)	-0.092	-0.216**	0.055	0.002	-0.000	0.009
	(0.159)	(0.100)	(0.102)	(0.020)	(0.010)	(0.017)
$\times$ (3^{rd}Feb.-23^{rd}Feb.)	-0.108	-0.210**	0.056	-0.014	0.007	0.009
	(0.151)	(0.103)	(0.101)	(0.010)	(0.005)	(0.009)
$\times$ (13^{th}Jan.-2^{nd}Feb.)	-0.095	-0.216**	0.046			
	(0.147)	(0.106)	(0.099)			
Mean	6.865	6.865	6.865	6.865	6.865	6.865
R-squared	0.709	0.886	0.934	0.974	0.978	0.982
Number of clusters	76	76	76	76	76	76
Observations	8,512	8,512	8,512	8,512	8,512	8,512
Day FEs	Yes	Yes	Yes	Yes	Yes	Yes
Region $\times$ phase FEs		Yes	Yes		Yes	Yes
Province controls $\times$ phase			Yes			Yes
Province FEs				Yes	Yes	Yes

Notes: Dependent variable is the Inverse Hyperbolic Sine (IHS) of the number of trips within a province (per 1000 inhabitants). "Exposure To Outbreak" is the number of people who moved from the province to one of the outbreak areas during 2015-2018 (per 1000 inhabitants). Province controls are the same as in Table 1. Standard errors clustered at the province level in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Consistent with the discussion above, Table A.8 and A.9 show that the results are driven by recent migrants, rather than old migrants.

4.5. ROBUSTNESS AND PLACEBOS

Table A.10 shows that dropping any entire region does not change affect the main results. Table A.11 shows that replacing exposure to outbreak with exposure to another region does not generate results anywhere similar to the coefficient estimates we found in Table 1, except for the outbreak regions where the outbreak provinces are located (Veneto, Lombardy and Emilia-Romagna).[35] Even in this case, the coefficient estimate drops substantially once we control for exposure to outbreak provinces. Importantly, the last row shows that exposure to any province is not associated with any additional death. The latter result emphasizes

[35] A partial exception is exposure to Marche. Marche is the first region to experience a Covid death (2^{nd} of March) after Veneto, Lombardy and Emilia-Romagna. Hence, the effect is consistent with the argument of this paper.

the importance of having clearly defined outbreak areas and data on the number of migrants specific to these areas.

In the previous section, we showed that exposure to outbreak has no effect on within-province mobility. It might nonetheless be sensible to estimate the effect of exposure to outbreak on mortality controlling for within-province mobility, because the latter might capture some important omitted determinant of mortality.[36] Table A.12 shows the results. Again, the evidence does not support this mechanism.

4.6. Back of the envelope calculation

To assess the magnitude of the relationship between exposure to outbreak areas and deaths, we calculate how many fewer deaths non-outbreak provinces would have experienced, had they had an exposure equal to 10^{th} percentile of the exposure distribution.

To do so, we take the following steps. First, we pick the province at the 10^{th} percentile of the exposure distribution as a reference point. This is Frosinone (Lazio), which has an exposure of 2.47 migrants to outbreak areas per 1000 inhabitants.

Second, for each province, we calculate the decrease in exposure that would be necessary to be as exposed as Frosinone. For example, pick Verbano-Cusio-Ossola (Piedmont), which is at the 90^{th} percentile of the exposure distribution. Its exposure is 7.08 migrants to outbreak areas per 1000 inhabitants. For it to be similar to Frosinone, its exposure would have to decrease by 65 percent.

Third, for each province, we multiply such decrease by the marginal effects discussed in Section 4.1 (which refers to Table 1).[37] For example, according to our estimates, had Verbano-Cusio-Ossola had an exposure similar to Frosinone, it would have suffered 45 fewer Covid deaths,[38] and 72 fewer total deaths.[39]

Fourth, we compute the total for all provinces. Had all provinces had the same exposure as the one at the 10^{th} percentile, they would have suffered 5,895 fewer Covid deaths and 7,348 fewer total deaths. If we do not include provinces below the tenth percentile, provinces

[36] Durante, Guiso and Gulino (2020) show that social capital in Italy is correlated with within-province mobility, which proxies violations of quarantine measures and therefore presumably leads to additional Covid deaths.

[37] To be as precise as possible when calculating the number of fewer Covid deaths, we use population numbers for each province, rather than the mean for the entire sample. Along similar lines, when calculating the number of fewer total deaths, we use 2015-2019 total deaths by province.

[38] There would have been approximately 18 fewer Covid deaths in March, 25 in April and 3 in May. Calculating these quantities requires the population for Verbano-Cusio-Ossola, which is about 0.16 million people. Note Verbano-Cusio-Ossola suffered 97 actual Covid deaths (55 in March, 32 in April and 10 in May), while Frosinone suffered 69 actual Covid deaths (10 in March, 37 in April and 22 in May).

[39] There would have been 43 fewer total deaths in March, 22 in April April and 8 in May. Calculating these quantities requires the 2015-2019 average total deaths for Verbano-Cusio-Ossola, which are 171 (March), 165 (April) and 159 (May).

would have suffered 6,111 fewer Covid deaths and 7,599 fewer total deaths.
These are important quantities, because they constitute 60 percent of all Covid deaths in non-outbreak regions[40] and 18 percent of all Covid deaths in the country.[41]

5. CONCLUSIONS

In this paper, we asked whether internal migration helps spread viruses. The idea is that, once a virus outbreaks, its diffusion to the rest of the country is neither homogeneous nor random: it depends on pre-existing internal migration routes, either because people travel regularly more along these routes, either because the virus and its immediate consequences (*e.g.*, fear, social isolation measures and the shutdown of economic activities) lead recent migrants to move back to their hometowns.
To answer this question, we focussed on Covid-19 and used rich panel data on Italian provinces and regions, including highly disaggregated data on internal migration. Specifically, we used yearly data on the number of people who de-registered themselves from one province and registered themselves in another to measure, for each province outside outbreak regions, the number of people who moved to one of the outbreak areas. Our "Exposure to outbreak" indicator is the share of movers relative to the population in the province of origin. We then exploited variation in this exposure across provinces located in the same region to identify its effects on Covid and total mortality. A variety of robustness tests and placebo estimations lent credibility to this research design.
Results suggest that a 50 percent increase in exposure relative to the mean (*i.e.*, about one standard deviation increase in exposure) leads to 117 additional Covid deaths and 147 additional total deaths per province. This is a large effect. A back of the envelope calculation shows that, had provinces outside outbreak regions had an exposure equal to the 10^{th} percentile, they would have experienced 5,895 (*i.e.*, 60 percent) fewer Covid deaths and 7,348 fewer total deaths.
We then used mobile phone based mobility data to test whether more exposed provinces do indeed receive a greater inflow of people from outbreak areas and to what extent such greater inflow, if any, explains the reduced form effect on mortality. The evidence confirms that greater exposure leads to greater inflow, and suggests that such greater inflow explains between 41 and 100 percent of the reduced form effect (depending on the month and the mortality measure one looks at).
We also find evidence of a post-outbreak "rush" away from outbreak areas back to hometowns, but such effect seems small relative to the effect on regular trips and therefore is

[40] This percentage is based on 9,904 Covid deaths.
[41] This percentage is based on 32,218 total Covid deaths.

unlikely to be the primary driver of the effect.

In light of these findings, governments could build up a database of recent migrants to be ready to contact those located in future outbreak areas. They could also complement it with information on people studying or working in provinces other than those where they are registered in.[42] The behavioral literature on taxation suggests that small nudges might be enough to persuade many of them not to travel if personally reminded of the consequences that could have.

The exposure index provided in this study can be thought of as a simple risk measure. Local governments could check to see how exposed they are (and eventually communicate their citizens to strengthen voluntary self-isolation). Central governments could use it to improve the allocation of scarce emergency resources across administrative units.

6. References

Adda, Jerome. 2016. "Economic Activity and the Spread of Viral Diseases: Evidence from High Frequency Data *." The Quarterly Journal of Economics 131 (2):891-941.

Altonji, Joseph, xa, G, Todd Elder, xa, E, Christopher Taber, xa, and R. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." Journal of Political Economy 113 (1):151-184.

Barmby, Tim, and Makram Larguem. 2009. "Coughs and sneezes spread diseases: An empirical study of absenteeism and infectious illness." Journal of Health Economics 28 (5):1012-1017.

Barro, Robert J., José F. Ursúa, and Joanna Weng. 2020. "The Coronavirus and the Great Influenza Pandemic: Lessons from the "Spanish Flu" for the Coronavirus's Potential Effects on Mortality and Economic Activity." National Bureau of Economic Research Working Paper Series No. 26866.

Barsbai, Toman, Hillel Rapoport, Andreas Steinmayr, and Christoph Trebesch. 2017. "The Effect of Labor Migration on the Diffusion of Democracy: Evidence from a Former Soviet Republic." American Economic Journal: Applied Economics 9 (3):36-69.

Beach, Brian, Joseph P. Ferrie, and Martin H. Saavedra. 2018. "Fetal Shock or Selection? The 1918 Influenza Pandemic and Human Capital Development." National Bureau of Economic Research Working Paper Series No. 24725.

Beine, Michel, Frédéric Docquier, and Maurice Schiff. 2013. "International migration, transfer of norms and home country fertility." Canadian Journal of Economics/Revue canadienne d'économique 46 (4):1406-1430.

[42] For Italy, this could be done using information from Universities (for students) and from the National Social Security Institute, called INPS (for workers).

Bellemare, Marc F., and Casey J. Wichman. 2020. “Elasticities and the Inverse Hyperbolic Sine Transformation." Oxford Bulletin of Economics and Statistics 82 (1):50-61.

Bonifazi, Corrado, and Frank Heinz. 2000. “Long-term Trends of Internal Migration in Italy.” International Journal of Population Geography 6, 111-131.

Bryan, Gharad, and Melanie Morten. 2019. “The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia.” Journal of Political Economy 127 (5):2229-2268. doi: 10.1086/701810.

Briscese, Guglielmo, Nicola Lacetera, Mario Macis, and Mirco Tonin. 2020. “Compliance with COVID-19 Social-Distancing Measures in Italy: The Role of Expectations and Duration." National Bureau of Economic Research Working Paper Series No. 26916.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. “Bootstrap-Based Improvements for Inference with Clustered Errors." The Review of Economics and Statistics 90 (3):414-427.

Carmichael, Ann G. 2014. “Plague and the Poor in Renaissance Florence," Cambridge University Press.

Chauvet, Lisa, Flore Gubert, Marion Mercier, and Sandrine Mesplé-Somps. 2015. “Migrants’ Home Town Associations and Local Development in Mali." The Scandinavian Journal of Economics 117 (2):686-722.

Chauvet, Lisa, and Marion Mercier. 2014. “Do return migrants transfer political norms to their origin country? Evidence from Mali." Journal of Comparative Economics 42 (3):630-651.

Chinazzi, Matteo, Jessica T. Davis, Marco Ajelli, Corrado Gioannini, Maria Litvinova, Stefano Merler, Ana Pastore y Piontti, Kunpeng Mu, Luca Rossi, Kaiyuan Sun, Cécile Viboud, Xinyue Xiong, Hongjie Yu, M. Elizabeth Halloran, Ira M. Longini, and Alessandro Vespignani. 2020. “The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak." Science 368 (6489):395-400.

Clark, G. and Cummins, N. (2009), “Urbanization, Mortality and Fertility in Malthusian England”, American Economic Review, Papers and Proceedings, 99, 242–247.

Custodero. 2020. “Coronavirus, in Veneto la prima vittima, Adriano Trevisan, 78 anni. Venti contagiati in Italia. Conte: “Nuove misure”,” Repubblica (22^{nd} February 2020), available at https://www.repubblica.it/cronaca/2020/02/22/news/cina_coronavirus_italia_virus_wuhan_influenza_codogno_lombardia_veneto_adriano_trevisan-249215365/

DECRETO DEL PRESIDENTE DEL CONSIGLIO DEI MINISTRI 22 February 2020. Available at https://www.gazzettaufficiale.it/eli/id/2020/03/22/20A01806/sg

DECRETO DEL PRESIDENTE DEL CONSIGLIO DEI MINISTRI 09 March 2020. Available at https://www.gazzettaufficiale.it/eli/id/2020/03/09/20A01558/sg

Durante, Ruben, Luigi Guiso and Giorgio Gulino. 2020. "Asocial Capital: Civic Culture and Social Distancing during COVID-19," CEPR Discussion Papers 14820.

Egorov, Georgy, Ruben Enikolopov, Alexey Makarin and Maria Petrova. 2020. “Divided We Stay Home: Social Distancing and Ethnic Diversity.” *mimeo* (June 2020).

Fang, Hanming, Long Wang, and Yang Yang. 2020. “Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China." National Bureau of Economic Research Working Paper Series No. 26906.

Gatto, Marino, Enrico Bertuzzo, Lorenzo Mari, Stefano Miccoli, Luca Carraro, Renato Casagrandi, and Andrea Rinaldo. 2020. “Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures." Proceedings of the National Academy of Sciences:202004978.

Giuffrida, Angela; Tondo, Lorenzo (8 March 2020). "Leaked coronavirus plan to quarantine 16m sparks chaos in Italy". The Guardian. Retrieved 8 March 2020.

Grewal, Sharan. 2020. “From Islamists to Muslim Democrats: The Case of Tunisia’s Ennahda." American Political Science Review 114 (2):519-535.

Gröger, André, and Yanos Zylberberg. 2016. “Internal Labor Migration as a Shock Coping Strategy: Evidence from a Typhoon." American Economic Journal: Applied Economics 8 (2):123-53.

Kraemer, Moritz U. G., Chia-Hung Yang, Bernardo Gutierrez, Chieh-Hsi Wu, Brennan Klein, David M. Pigott, Louis du Plessis, Nuno R. Faria, Ruoran Li, William P. Hanage, John S. Brownstein, Maylis Layan, Alessandro Vespignani, Huaiyu Tian, Christopher Dye, Oliver G. Pybus, and Samuel V. Scarpino. 2020. “The effect of human mobility and control measures on the COVID-19 epidemic in China." Science 368 (6490):493-497.

Kuchler, Theresa, Dominic Russel, and Johannes Stroebel. 2020. “The Geographic Spread of COVID-19 Correlates with Structure of Social Networks as Measured by Facebook." National Bureau of Economic Research Working Paper Series No. 26990.

Kwakwa, Victoria. 2020. “Governments facing tough choices in COVID-19 crisis", published on the East Asia and Pacific World Bank blog. Available at https://blogs.worldbank.org/eastasiapacific/governments-facing-tough-choices-covid-19-coronavirus-crisis?CID=WBW_AL_BlogNotification_EN_EXT

Litvinova, Maria, Quan-Hui Liu, Evgeny S. Kulikov, and Marco Ajelli. 2019. “Reactive school closure weakens the network of social interactions and reduces the spread of influenza." Proceedings of the National Academy of Sciences 116 (27):13174-13181.

Mercier, Marion. 2016. “The return of the prodigy son: Do return migrants make better leaders?" Journal of Development Economics 122:76-91.

Mikhailova, T. and Valsecchi M. (2020). “Internal migration and the Covid-19 virus" in

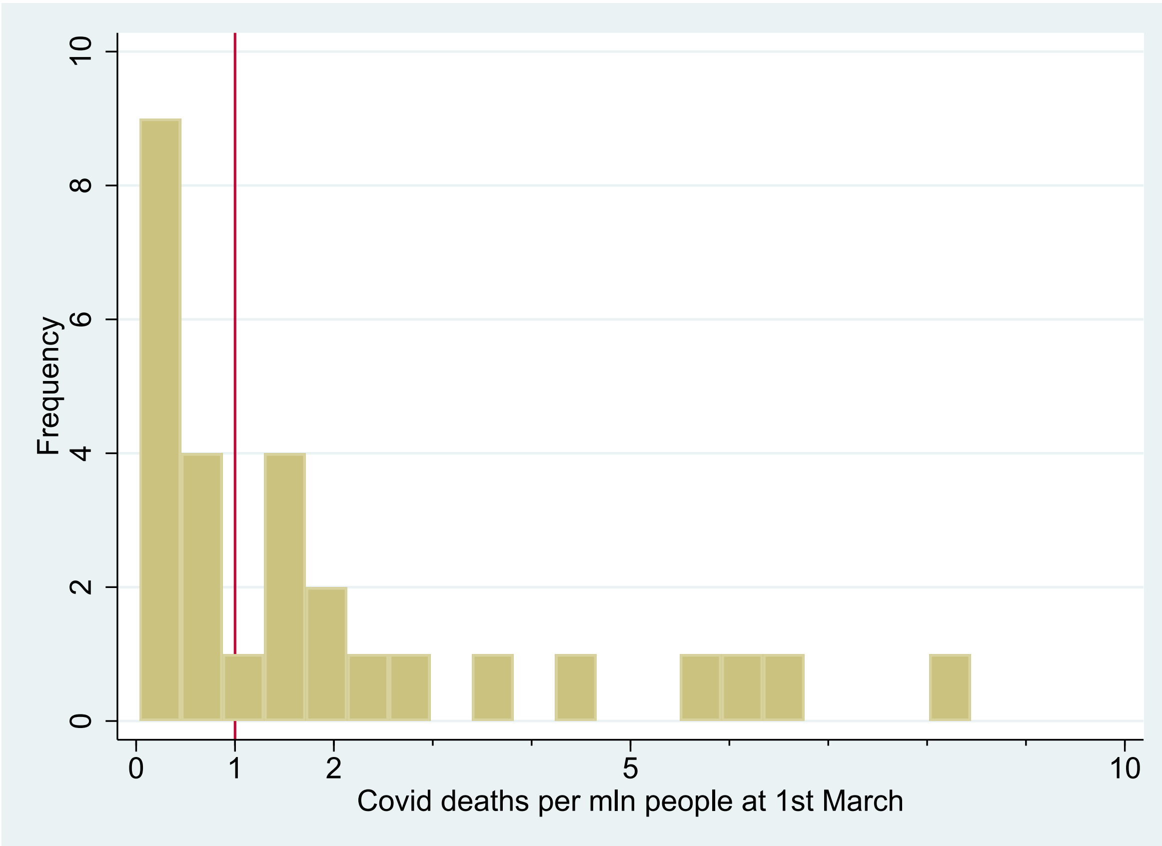
Economic Policy During the Covid-19 (ebook). New Economic School (March 2020).
Mocetti, Sauro and Porello, Carmine, Labour Mobility in Italy: New Evidence on Migration Trends (January 22, 2010). Bank of Italy Occasional Paper No. 61.
Monras, Joan.2018. "Economic Shocks and Internal Migration," CEPR Discussion Paper No. DP12977 (June 2018, last revised May 2020), Available at SSRN: https://ssrn.com/abstract=3193980
Oster, Emily. 2012. "ROUTES OF INFECTION: EXPORTS AND HIV INCIDENCE IN SUB-SAHARAN AFRICA." Journal of the European Economic Association 10 (5):1025-1058.
Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." Journal of Business Economic Statistics 37 (2):187-204. doi: 10.1080/07350015.2016.1227711.
Panichella, Nazareno. 2014. "Meridionali al Nord: Migrazioni interne e società italiana dal dopoguerra ad oggi." Il Mulino.
Pichler, Stefan, and Nicolas Ziebarth. 2019. "Reprint of: The pros and cons of sick pay schemes: Testing for contagious presenteeism and noncontagious absenteeism behavior." Journal of Public Economics 171 (C):86-104.
Severgnini, Chiara (8 March 2020). "Coronavirus, Conte: "Ecco il decreto con le nuove misure, in vigore fino al 3 aprile"" [Coronavirus, Conte: "Here is the decree with the new measures, in force until April 3"]. Corriere della Sera
Spilimbergo, Antonio. 2009. "Democracy and Foreign Education." American Economic Review 99 (1):528-43.
TgCom24. 2020. "Coronavirus, morta una donna in Lombardia: seconda vittima italiana," 22nd February 2020, available at https://www.tgcom24.mediaset.it/cronaca/lombardia/coronavirus-morta-una-donna-in-lombardia-seconda-vittima-italiana_15150431-202002a.shtml
Tian, Yuan, Maria Esther Caballero, and Brian Kovak. 2020. "Social Learning along International Migrant Networks." Mimeo (5th July 2020).
Valsecchi, Michele. 2020. "Internal Migration and the Spread of Covid-19," *CEPR Covid Economics 18* (15th May 2020).
Voigtländer, Nico, and Hans-Joachim Voth. 2012. "The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe." The Review of Economic Studies 80 (2):774-811.
Woods, Robert. 2003. "Urban-Rural Mortality Differentials: An Unresolved Debate." Population and Development Review 29 (1):29-46.
World Bank. 2020. "COVID-19 Crisis Through a Migration Lens," available at https://openknowledge.worldbank.org/handle/10986/33634
World Health Organization. 2020. "Corona Virus Disease 19 (COVID 19). Situation Report

- 184". Published on the 22nd July at https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200722-covid-19-sitrep-184.pdf?sfvrsn=7680210a_2

Yang, Dean. 2008. "International Migration, Remittances and Household Investment: Evidence from Philippine Migrants' Exchange Rate Shocks*." The Economic Journal 118 (528):591-630.

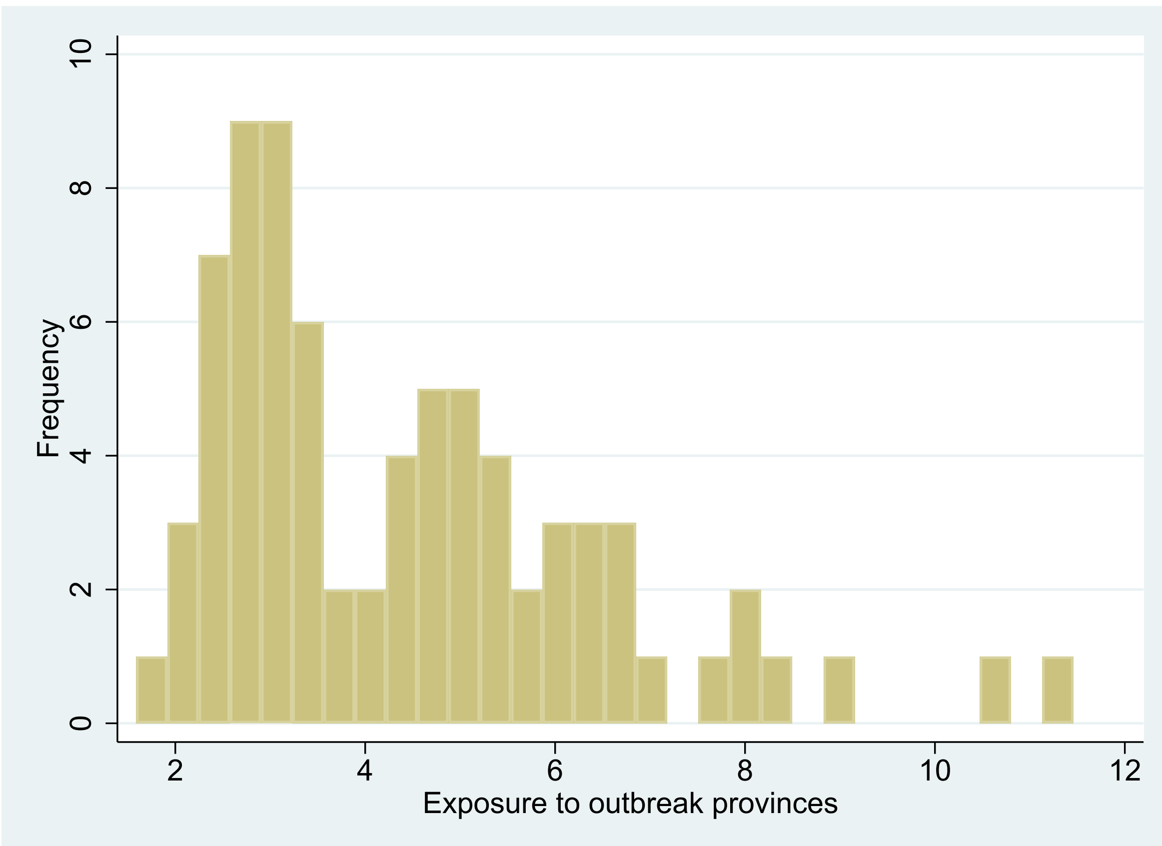
Zenner, Dominik, and Kol Wickramage. 2020. "National preparedness and response plans for COVID-19 and other diseases: Why migrants should be included," published on Migration and Vulnerability blog of the Migration Data Portal. Available at https://migrationdataportal.org/blog/national-preparedness-and-response-plans-covid-19-and-other-diseases-why-migrants-should-be

FIGURE A.1: COVID DEATHS ACROSS PROVINCES IN OUTBREAK REGIONS



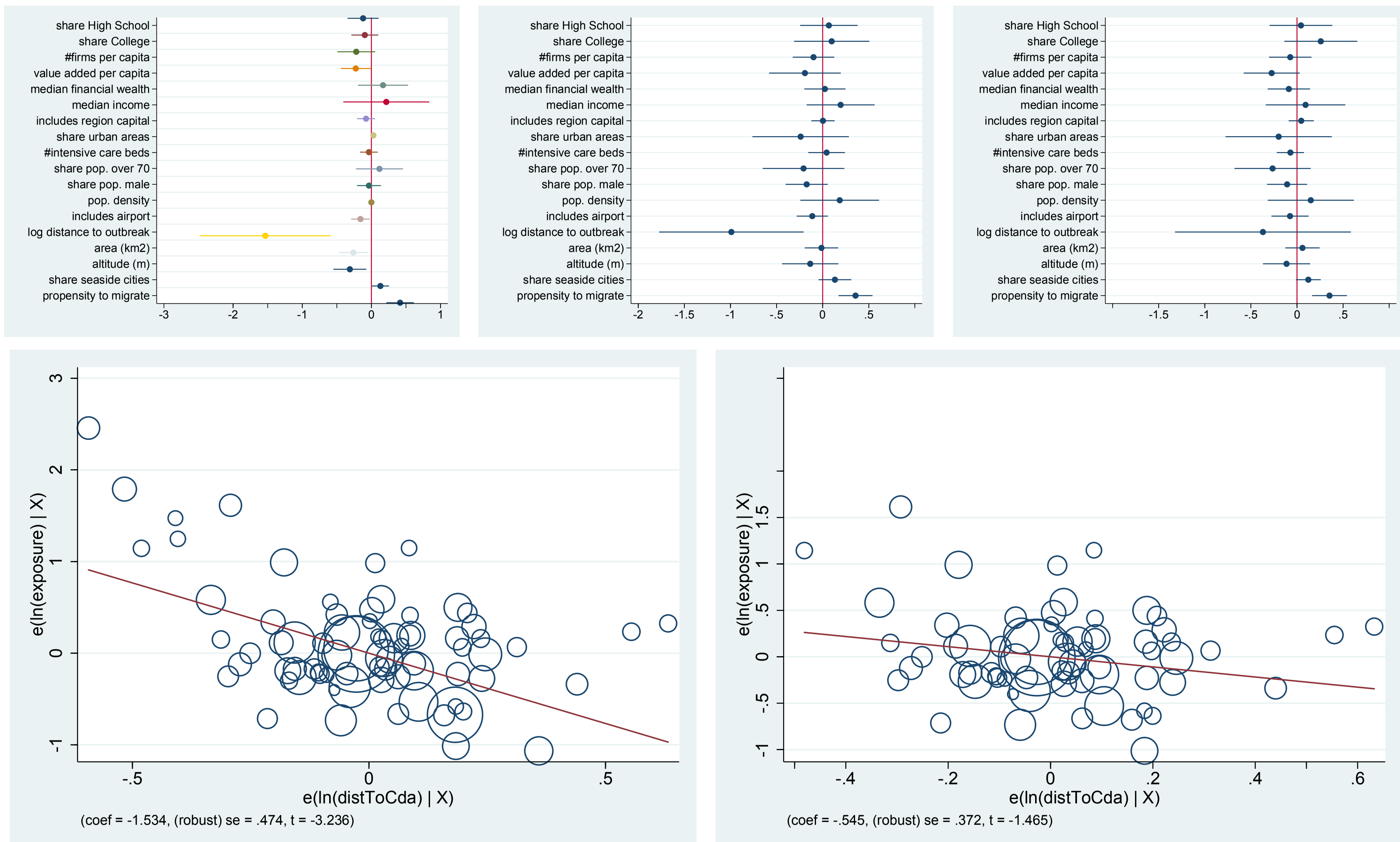
Notes: number of Covid deaths per million people (as of the 1^{th} March) across the 28 provinces of Lombardia, Veneto and Emilia-Romagna. The vertical line indicates the cut-off used to define the (15) outbreak provinces.

FIGURE A.2: EXPOSURE TO OUTBREAK ACROSS PROVINCES LOCATED OUTSIDE OUTBREAK REGIONS



Notes: "Exposure to outbreak provinces" across the 76 provinces located outside the outbreak regions (Lombardy, Veneto and Emilia-Romagna). "Exposure to outbreak provinces" is the number of people who deregistered from a given province to register in one of the outbreak provinces during 2015-2018 (per 1000 inhabitants).

FIGURE A.3: BALANCE TEST



Notes: both outcome and explanatory variables have been standardized. Panel a) shows the coefficient estimates associated with separate regressions. Panel b) shows the coefficient estimates associated with one regression with all covariates entered contemporaneously. Panel c) repeats the estimation of panel b) without the Piedmont (Piemonte) region. Panel d) shows the relationship between the exposure indicator and distance to outbreak areas for the full sample. Panel e) shows the relationship of panel d) without the Piedmont (Piemonte) region. All specifications include region FEs and are weighted by 2018 population.

TABLE A.1: SUMMARY STATISTICS AT THE REGIONAL LEVEL

Variable	N	Mean	Std. Dev.	Min.	Max.
Indicators at the regional-daily level					
Covid deaths per mln people	1,984	2.51	4.74	0.00	45.90
# trips from outbreak areas per 1000 ppl. (13^{th}Jan.-3^{rd}May)	1,792	3.04	5.38	0.00	30.71
# trips from outbreak areas per 1000 ppl. (13^{th}Jan.-23^{rd}Feb.)	672	5.43	7.16	0.07	30.71
# trips from outbreak areas per 1000 ppl. (24^{th}Feb.-8^{th}Mar.)	224	4.00	5.69	0.02	25.50
# trips from outbreak areas per 1000 ppl. (9^{th}Mar.-3^{rd}May)	896	1.02	1.64	0.00	8.71
Indicators at the regional level					
Exposure to outbreak areas	16	4.10	1.40	2.19	7.30
Exposure to any area	16	25.64	4.40	14.20	32.97
Distance to outbreak provinces (km)	16	531.12	263.81	176.62	994.51
Compliance to quarantine					
People with High School or higher	16	0.46	0.05	0.40	0.57
People with Bachelor degree or higher	16	0.14	0.02	0.11	0.20
Newspaper readership (at least once a week)	16	0.38	0.09	0.26	0.58
Newspaper readership (five times a week)	16	0.32	0.08	0.22	0.45
Trust in others	16	0.20	0.06	0.13	0.37
State capacity					
Unemployment	16	0.12	0.05	0.04	0.22
Regional GDP	16	0.03	0.01	0.02	0.04
Intensive care beds (per 100,000 inh.)	16	8.46	1.49	5.75	11.56
Other indicators					
People over 70	16	0.17	0.02	0.13	0.22
Population	16	2,560,240.44	1,923,193.24	308,493.00	5,896,693.00

TABLE A.2: SUMMARY STATISTICS AT THE PROVINCIAL LEVEL

Variable	N	Mean	Std. Dev.	Min.	Max.
Mortality indicators: Covid deaths					
Number of deaths per mln ppl, 20^{th} Feb.-31^{st} Mar. 2020	76	86.97	118.54	0.00	523.37
Number of deaths per mln ppl, April 2020	76	155.05	177.45	6.94	841.37
Number of deaths per mln ppl, May 2020	76	45.43	61.47	-25.93	263.92
Mortality indicators: total deaths					
Number of deaths per mln ppl, 20^{th} Feb.-31^{st} Mar. 2015-2019	76	1,106.67	283.83	17.05	1,633.76
Number of deaths per mln ppl, 20^{th} Feb.-31^{st} Mar. 2020	76	1,291.92	507.04	25.57	2,826.67
Growth for Jan-Feb, 2020 vs 2015-2019	76	-0.07	0.06	-0.17	0.09
Growth for March, 2020 vs 2015-2019	76	0.20	0.28	-0.09	1.25
Growth for April, 2020 vs 2015-2019	76	0.17	0.28	-0.18	0.84
Growth for May, 2020 vs 2015-2019	76	-0.05	0.09	-0.26	0.18
Indicators at the provincial-daily level					
# trips from outbreak areas per 1000 ppl. (13^{th}Jan.-3^{rd}May)	8,512	3.42	9.15	0.00	82.74
# trips from outbreak areas per 1000 ppl. (13^{th}Jan.-23^{rd}Feb.)	3,192	6.09	12.74	0.00	82.74
# trips from outbreak areas per 1000 ppl. (24^{th}Feb.-8^{th}Mar.)	1,064	4.50	9.99	0.00	65.13
# trips from outbreak areas per 1000 ppl. (9^{th}Mar.-3^{rd}May)	4,256	1.15	3.17	0.00	36.10
# trips within own province per 1000 ppl. (13^{th}Jan.-3^{rd}May.)	8,512	583.16	328.81	13.95	1584.20
Indicators at the provincial level					
Exposure to outbreak areas	76	4.48	2.04	1.59	11.46
Exposure to any area	76	28.08	6.52	11.24	45.66
Distance to outbreak provinces (km)	76	543.61	296.67	159.71	1,059.29
Share of people with High School (or higher)	76	0.31	0.03	0.24	0.39
Share of people with Bachelor degree (or higher)	76	0.03	0.00	0.02	0.04
Number of firms per capita	76	0.07	0.01	0.04	0.11
Value added per capita	76	21,151.99	5,765.97	13,260.33	40,431.74
Median financial wealth	76	7,857.23	5,521.42	0.00	23,891.21
Median income	76	25,730.38	5,786.63	15,400.00	38,762.48
Intensive care beds (per 100,000 inh.)	76	7.79	3.44	2.87	20.68
Share of people over 70	76	0.18	0.02	0.12	0.23
area (km2)	76	2,965.04	1,747.43	212.51	7,398.38
Altitude (meters)	76	354.54	172.49	33.12	849.71
Share of seaside cities	76	0.16	0.18	0.00	0.79
Population density	76	224.62	325.40	49.34	2,630.35
Share of males	76	0.49	0.01	0.48	0.50
Whether there is an airport	76	0.26	0.44	0.00	1.00
Share of urban areas	76	0.03	0.07	0.00	0.57
Province includes region capital	76	0.21	0.41	0.00	1.00
Number of people (2018)	76	532,507	634,308	85,237	4,355,725

Note: sources of data described in the Data section.

TABLE A.3: EXPOSURE TO OUTBREAK AND COVID DEATHS BY REGION-PHASE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(Exposure to Outbreak)							
× (4^{th}May.-)	-0.420	1.873**	4.004**	3.632**	0.744	-0.371	2.067
	(0.665)	(0.829)	(1.690)	(1.250)	(0.818)	(0.704)	(1.200)
	0.495	0.154	0.240	0.065	0.410	0.565	0.289
× (6^{th}Apr.-3^{rd}May.)	-1.368	8.711***	19.473***	14.949**	3.355	-0.634	10.101**
	(3.207)	(2.879)	(6.391)	(5.472)	(3.082)	(3.464)	(4.691)
	0.675	0.007	0.006	0.020	0.321	0.867	0.167
× (9^{th}Mar.-5^{th}Apr.)	-1.214	6.003***	16.159***	11.020***	2.192	-0.376	12.020***
	(2.643)	(1.498)	(3.623)	(3.075)	(2.487)	(2.734)	(4.025)
	0.675	0.006	0.002	0.016	0.416	0.918	0.008
× (24^{th}Feb.-8^{th}Mar.)	0.013	0.098*	0.270**	0.212**	0.074	0.021	0.259*
	(0.039)	(0.055)	(0.095)	(0.075)	(0.050)	(0.038)	(0.140)
	0.735	0.062	0.069	0.005	0.129	0.559	0.132
Mean	2.527	2.527	2.527	2.527	2.527	2.527	2.527
R-squared	0.227	0.611	0.602	0.564	0.494	0.252	0.676
Number of regions	16	16	16	16	16	16	16
Observations	1,968	1,968	1,968	1,968	1,968	1,968	1,968
Day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log(distance to outbreak) × phase	-	Yes					Yes
Compliance to quarantine × phase	-	-	Yes	-	-	-	Yes
State capacity × phase	-	-	-	Yes	-	-	Yes
Pop. at risk × phase	-	-	-	-	Yes	-	Yes
Total emigrants × phase	-	-	-	-	-	Yes	Yes

Notes: Dependent variable is the number of Covid-19 deaths (per million people). "Exposure to Outbreak" is the number of people who moved to one of the outbreak areas during 2015-2018 (per 1000 inhabitants). Distance to outbreak is the average distance to the outbreak provinces. "Compliance with quarantine" is the first principal component of: share people with higher school education, share people with university education, newspaper readership (at least once a week), newspaper readership (at least five times a week) and share people who trust others. "State capacity" is the first principal component of: unemployment share, regional GDP per capita and number of intensive care beds per 100,000 inhabitants. "Population at risk" is the share of people with 70 years old or more. "Total emigrants" is the log of the number of people who changed residence to any province between 2015 and 2018. For each interaction, the table reports coefficient estimates on the first row, standard errors clustered at the region level (in brackets), and p-values for wild cluster bootstrap standard errors á la Cameron, Gelbach and Miller 2008). *** p<0.01, ** p<0.05, * p<0.1.

TABLE A.4: EXPOSURE TO LOMBARDIA, VENETO AND EMILIA-ROMAGNA AND COVID-19 DEATHS

	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: COVID DEATHS						
Dep var.	Number of deaths 20Feb-31Mar	Number of deaths 20Feb-31Mar	Number of deaths April	Number of deaths April	Number of deaths May	Number of deaths May
ln(Exposure To LVE)	148.065***	174.127***	93.358	184.505	11.960	29.906
	(43.752)	(53.808)	(70.844)	(155.046)	(21.658)	(60.039)
Mean	86.968	86.968	155.054	155.054	45.430	45.430
R-squared	0.828	0.904	0.702	0.785	0.725	0.781
Region FEs	Yes	Yes	Yes	Yes	Yes	Yes
Province controls		Yes		Yes		Yes
PANEL B: TOTAL DEATHS						
Dep var.	Number of deaths 20Feb-31Mar 2015-2019	Number of deaths 20Feb-31Mar 2020	Growth of deaths Jan-Feb 2020 vs 2015-2019	Growth of deaths March 2020 vs 2015-2019	Growth of deaths April 2020 vs 2015-2019	Growth of deaths May 2020 vs 2015-2019
ln(Exposure To LVE)	122.720	560.614**	0.004	0.367*	0.227*	0.081
	(131.931)	(237.152)	(0.030)	(0.184)	(0.122)	(0.059)
Mean	1107	1292	-0.065	0.203	0.167	-0.048
R-squared	0.688	0.842	0.631	0.871	0.921	0.654
Region FEs	Yes	Yes	Yes	Yes	Yes	Yes
Province controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	76	76	76	76	76	76

Notes: LVE stands for Lombardia, Veneto and Emilia-Romagna. Number of Covid deaths (Panel A) and total deaths (Panel B, Columns 1-2) is per million inhabitants. Growth of total deaths (Panel B, Columns 3-6) is per province. "Exposure To Outbreak" is the number of people who moved from the province to one of the outbreak areas between 2015 and 2018 (per 1000 inhabitants). Geographic controls include: log distance to outbreak provinces, number of square kilometres, altitude, share of seaside cities. Socio-demographic controls include: population density, share of males, number of intensive care hospital beds per 100,000 inhabitants, whether there is an airport, share of urban areas, population share above 70 years, population share with high school education or higher, population share with university education. Economic controls include: number of firms per capita, value added per capita, median financial wealth, median income. Total migration is the log of the number of people who moved from the province to any other area in the country between 2015 and 2018 (per 1000 inhabitants). Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

TABLE A.5: EXPOSURE TO OUTBREAK AND COVID-19 DEATHS BY PROVINCE-MONTH

	(1)	(2)	(3)	(4)	(5)	(6)
Dep var.	Number of total deaths 20Feb-31Mar 2015-2019	Number of total deaths 20Feb-31Mar 2020	Growth of total deaths Jan-Feb 2020 vs 2015-2019	Growth of total deaths March 2020 vs 2015-2019	Growth of total deaths April 2020 vs 2015-2019	Growth of total deaths May 2020 vs 2015-2019
ln(Exposure To Outbreak)	137.390	558.398***	-0.003	0.338***	0.223***	0.122***
	(84.412)	(175.164)	(0.017)	(0.091)	(0.054)	(0.035)
Mean	1107	1292	-0.065	0.203	0.167	-0.048
R-squared	0.512	0.733	0.410	0.828	0.871	0.497
Observations	76	76	76	76	76	76
Region FEs	Yes	Yes	Yes	Yes	Yes	Yes
Province controls	No	No	No	No	No	No

Notes: the number of total deaths (Columns 1-2) is per million inhabitants. The growth of total deaths (Columns 3-6) is per province. “Exposure To Outbreak” is the number of people who moved from the province to one of the outbreak areas during 2015-2018 (per 1000 inhabitants). Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

TABLE A.6: COEFFICIENT STABILITY AND ROBUSTNESS TO SELECTION ON UNOBSERVABLES

Dep var.	Number of Covid deaths 20Feb-31Mar (1)	Number of Covid deaths April (2)	Number of total deaths 20Feb-31Mar (3)	Growth of total deaths March (4)	Growth of total deaths April (5)
PANEL A: $\dot{\beta}$ (UNCONTROLLED β)					
ln(Exposure To Outbreak)	153.289***	137.222*	558.398***	0.338***	0.223***
	(41.972)	(72.272)	(175.164)	(0.091)	(0.054)
PANEL B: $\tilde{\beta}$ (CONTROLLED β)					
ln(Exposure To Outbreak)	169.347***	241.456*	509.973**	0.382**	0.203*
	(47.021)	(137.225)	(213.585)	(0.166)	(0.103)
PANEL C: $\beta^{*'}(R_{max},1)$ (BOUNDING STATEMENT FOR β)					
$\beta^{*'}(R^{max},1)$	237.3	325.7	427.1	0.544	0.168
Does $\tilde{\beta}$+/- 1.96 s.e. include $\beta^{*'}(R_{max},1)$?	yes	yes	yes	yes	yes
PANEL D: δ^*					
$\tilde{\beta} < \dot{\beta}$?	no	no	yes	no	yes
δ^*			1.630		1.582

Notes: specifications associated with Panel B include province controls. See Table 1 for the list of controls. Specifications associated with both Panel A and B include region FEs. Bounding statements for β in Panel C and statements about δ in Panel D are based on R_{max}=1.3*$\tilde{R}$. See Oster (2019) for details. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

TABLE A.7: EXPOSURE TO OUTBREAK AND TRIPS FROM OUTBREAK AREAS BY REGION-PHASE

ln(Exposure To Outbreak)				
× (6^{th}Apr.-3^{rd}May)	-0.412	1.014	0.924***	-0.324
	(0.420)	(0.630)	(0.289)	(0.333)
	0.326	0.179	0.026	0.424
× (9^{th}Mar.-5^{th}Apr.)	-0.451	1.069	0.885***	-0.268
	(0.449)	(0.725)	(0.252)	(0.244)
	0.356	0.216	0.019	0.376
× (24^{th}Feb.-8^{th}Mar.)	-1.047	1.561	0.289**	0.223*
	(0.736)	(0.931)	(0.116)	(0.111)
	0.203	0.167	0.032	0.122
× (3^{rd}Feb.-23^{rd}Mar.)	-1.349*	1.296	-0.012	-0.042
	(0.661)	(0.885)	(0.022)	(0.064)
	0.116	0.198	0.556	0.586
× (13^{th}Jan.-2^{nd}Feb.)	-1.336*	1.338		
	(0.667)	(0.937)		
	0.120	0.206		
Mean	1.094	1.094	1.094	1.094
R-squared	0.297	0.843	0.927	0.962
Number of regions	16	16	16	16
Observations	1,792	1,792	1,792	1,792
Day FE	Yes	Yes	Yes	Yes
Regional controls × phase	-	Yes		Yes
Region FEs			Yes	Yes

Notes: Dependent variable is the Inverse Hyperbolic Sine (IHS) of the number of trips from outbreak areas (per 1000 inhabitants). "Exposure To Outbreak" is the number of people who moved from the province to one of the outbreak areas during 2015-2018 (per 1000 inhabitants). Regional controls are: log distance to outbreak; compliance to quarantine; state capacity; population at risk; total emigrants. Distance to outbreak is the average distance to the outbreak provinces. "Compliance to quarantine" is the first principal component of population share with higher school education, population share with university education, newspaper readership (at least once a week), newspaper readership (at least five times a week) and trust in others. "State capacity" is the first principal component of unemployment share, regional GDP per capita and number of intensive care beds per 100,000 inhabitants. Population at risk is the share of people with 70 years old or more. Total emigrants is the log of the share of people who changed residence to another Italian region between 2015 and 2018. For each interaction, the table reports coefficient estimates on the first row, standard errors clustered at the region level (in brackets), and p-values for wild cluster bootstrap standard errors á la Cameron, Gelbach and Miller 2008). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

TABLE A.8: EXPOSURE TO OUTBREAK AND COVID-19 DEATHS: HORSE RACE B/W 2010-2018 AND 2001-2010

	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: COVID DEATHS						
Dep var.	Number of deaths 20Feb-31Mar	Number of deaths 20Feb-31Mar	Number of deaths April	Number of deaths April	Number of deaths May	Number of deaths May
ln(Exp. To Out.)						
..2010-2018	142.075***	102.507*	203.625	121.628	34.015	26.398
	(48.196)	(52.851)	(123.171)	(87.386)	(40.480)	(35.851)
..2001-2010		49.101		101.753		9.452
		(36.029)		(75.192)		(24.561)
Mean	86.968	86.968	155.054	155.054	45.430	45.430
R-squared	0.906	0.910	0.811	0.811	0.785	0.788
PANEL B: TOTAL DEATHS						
Dep var.	Number of deaths 20Feb-31Mar 2015-2019	Number of deaths 20Feb-31Mar 2020	Growth of deaths Jan-Feb 2020 vs 2015-2019	Growth of deaths March 2020 vs 2015-2019	Growth of deaths April 2020 vs 2015-2019	Growth of deaths May 2020 vs 2015-2019
ln(Exp. To Out.)						
..2010-2018	0.339**	0.279**	0.183*	0.161*	0.088**	0.127**
	(0.155)	(0.135)	(0.092)	(0.095)	(0.043)	(0.051)
..2001-2010		0.075		0.026		-0.048
		(0.108)		(0.073)		(0.043)
Mean	0.203	0.203	0.167	0.167	-0.048	-0.048
R-squared	0.879	0.881	0.920	0.920	0.660	0.669
Region FEs	Yes	Yes	Yes	Yes	Yes	Yes
Province controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	76	76	76	76	76	76

Notes: each Panel is a separate set of regressions. “Exp. To Out.” stands for “Exposure To Outbreak”. Year intervals (e.g., 2010-2018) indicates the time window upon which the “Exposure To Outbreak” indicator is based. For example, the first row of Panel A indicates that the Exposure to Outbreak indicator is based on the number of people who moved to outbreak areas between 2010 and 2018. Controls as in Table 1 except for total propensity to emigrate, which is defined over the time window 2001-2018. Robust standard errors in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

TABLE A.9: EXPOSURE TO OUTBREAK AND COVID-19 DEATHS: HORSE RACE B/W 2015-2018 AND 2010-2015

	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: COVID DEATHS						
Dep var.	Number of deaths 20Feb-31Mar	Number of deaths 20Feb-31Mar	Number of deaths April	Number of deaths April	Number of deaths May	Number of deaths May
ln(Exp. To Out.)						
..2015-2018	168.686***	322.646**	238.977*	230.064	26.070	-156.434
	(48.046)	(133.684)	(135.568)	(255.676)	(45.843)	(96.496)
..2010-2015		-149.728		8.668		177.487**
		(101.147)		(198.130)		(83.670)
Mean	86.968	86.968	155.054	155.054	45.430	45.430
R-squared	0.915	0.920	0.812	0.812	0.780	0.801
PANEL B: TOTAL DEATHS						
Dep var.	Number of deaths 20Feb-31Mar 2015-2019	Number of deaths 20Feb-31Mar 2020	Growth of deaths Jan-Feb 2020 vs 2015-2019	Growth of deaths March 2020 vs 2015-2019	Growth of deaths April 2020 vs 2015-2019	Growth of deaths May 2020 vs 2015-2019
ln(Exp. To Out.)						
..2015-2018	0.373**	0.382	0.201*	0.237	0.081*	0.064
	(0.167)	(0.297)	(0.102)	(0.269)	(0.047)	(0.170)
..2010-2015		-0.009		-0.035		0.017
		(0.216)		(0.234)		(0.158)
Mean	0.203	0.203	0.167	0.167	-0.048	-0.048
R-squared	0.883	0.883	0.921	0.921	0.658	0.658
Region FEs	Yes	Yes	Yes	Yes	Yes	Yes
Province controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	76	76	76	76	76	76

Notes: each Panel is a separate set of regressions. "Exp. To Out." stands for "Exposure To Outbreak". Year intervals (e.g., 2015-2018) indicates the time window upon which the "Exposure To Outbreak" indicator is based. For example, the first row of Panel A indicates that the Exposure to Outbreak indicator is based on the number of people who moved to outbreak areas between 2015 and 2018. Controls as in Table 1 except for total propensity to emigrate, which is defined over the time window 2010-2018. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

TABLE A.10: EXPOSURE TO OUTBREAK AND DEATHS: DROP ONE REGION AT THE TIME

Dep var.	Number of Covid deaths 20Feb-31Mar 2020	Number of Covid deaths 01Apr-30Apr 2020	Number of total deaths 20Feb-31Mar 2015-2019	Number of total deaths 20Feb-31Mar 2020	Variation in total deaths Jan-Feb 2020 vs 2015-2019	Variation in total deaths March 2020 vs 2015-2019	Variation in total deaths April 2020 vs 2015-2019
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Region dropped							
Marche	126.575***	79.339	56.206	302.439	-0.022	0.170*	0.124
	(45.874)	(74.827)	(143.574)	(184.902)	(0.027)	(0.084)	(0.089)
Liguria	159.312***	228.708	52.293	504.912**	0.022	0.368**	0.184*
	(47.471)	(144.420)	(136.870)	(216.791)	(0.033)	(0.171)	(0.102)
Piemonte	148.956**	367.838**	110.496	591.268**	0.015	0.472**	0.231*
	(60.661)	(143.510)	(190.656)	(271.245)	(0.047)	(0.215)	(0.126)
Trentino A.A.	172.308***	239.270*	58.439	520.541**	0.018	0.406**	0.171
	(48.243)	(141.629)	(129.808)	(218.403)	(0.029)	(0.166)	(0.108)
Friuli V.G.	171.416***	232.307*	100.518	559.378**	0.006	0.383**	0.206*
	(48.606)	(125.029)	(106.972)	(229.518)	(0.030)	(0.164)	(0.104)
Abruzzo	167.647***	233.251*	55.760	500.309**	0.008	0.376**	0.196*
	(46.013)	(136.348)	(127.610)	(216.389)	(0.031)	(0.167)	(0.101)
Basilicata	169.430***	241.681*	59.467	510.655**	0.008	0.383**	0.204*
	(47.001)	(137.181)	(124.849)	(212.881)	(0.032)	(0.165)	(0.103)
Calabria	177.611***	287.520*	-47.734	437.575*	0.008	0.421**	0.215*
	(52.618)	(149.325)	(139.176)	(229.312)	(0.036)	(0.185)	(0.118)
Campania	169.335***	244.191*	33.768	486.597**	0.010	0.389**	0.187*
	(48.169)	(134.136)	(116.789)	(224.345)	(0.031)	(0.166)	(0.106)
Lazio	166.120***	235.989	66.844	529.403**	0.030	0.406**	0.199
	(51.508)	(154.009)	(142.805)	(240.635)	(0.031)	(0.177)	(0.121)
Molise	170.200***	242.692*	74.260	525.070**	0.008	0.384**	0.204*
	(47.214)	(138.361)	(127.448)	(214.978)	(0.032)	(0.167)	(0.105)
Puglia	166.881***	231.560	59.960	502.840**	0.008	0.372**	0.188*
	(48.609)	(146.094)	(127.569)	(213.056)	(0.032)	(0.177)	(0.106)
Sardegna	172.992***	246.306*	80.061	534.589**	0.009	0.387**	0.206*
	(48.255)	(139.500)	(123.643)	(225.675)	(0.028)	(0.169)	(0.105)
Sicilia	192.248***	257.704*	161.763	657.015***	0.012	0.417**	0.263**
	(45.376)	(143.659)	(117.039)	(213.577)	(0.029)	(0.174)	(0.106)
Toscana	210.574***	288.978	-89.541	428.250	-0.006	0.457**	0.237*
	(44.818)	(192.847)	(166.653)	(275.196)	(0.039)	(0.187)	(0.140)
Umbria	169.528***	242.560*	68.424	516.447**	0.010	0.383**	0.206*
	(47.116)	(138.291)	(123.773)	(211.052)	(0.030)	(0.167)	(0.103)
Region FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Provincial controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: each row corresponds to a separate set of estimations that excludes provinces belonging to the region stated on the left column. Mean of dependent variable, R-squared and number of observations are omitted to keep the table readable. Province controls as in Table 1. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

TABLE A.11: EXPOSURE TO OTHER REGIONS AND DEATHS

Dep variable	Number of Covid deaths 20Feb-31Mar 2020	Number of Covid deaths 20Feb-31Mar 2020	Number of Covid deaths 01Apr-30Apr 2020	Number of Covid deaths 01Apr-30Apr 2020	Growth of total deaths March 2020 vs 2015-2019	Growth of total deaths March 2020 vs 2015-2019	Growth of total deaths April 2020 vs 2015-2019	Growth of total deaths April 2020 vs 2015-2019
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Exposure To..								
Lombardia)	94.935**	-136.221**	40.023	-503.110***	0.128	-0.586***	0.111	-0.174
	(44.731)	(50.908)	(75.026)	(109.366)	(0.097)	(0.162)	(0.087)	(0.166)
Emilia-Romagna)	118.337***	50.799	259.808**	200.734**	0.343**	0.221	0.170*	0.100
	(39.869)	(34.393)	(112.512)	(92.022)	(0.153)	(0.133)	(0.096)	(0.103)
Veneto)	76.763	-0.039	-73.113	-202.520	0.021	-0.168	0.116	0.026
	(61.938)	(54.640)	(132.014)	(159.740)	(0.132)	(0.170)	(0.100)	(0.105)
Marche)	16.681	2.644	70.204*	64.430	0.094**	0.079**	0.042	0.029
	(22.852)	(19.772)	(37.015)	(38.376)	(0.039)	(0.039)	(0.043)	(0.043)
Liguria)	10.927	-4.353	54.998*	34.670	-0.001	-0.037	-0.015	-0.034
	(20.319)	(18.072)	(32.458)	(33.288)	(0.050)	(0.048)	(0.037)	(0.037)
Piemonte)	37.145	-27.871	75.758	-90.007	0.070	-0.152	0.115	0.033
	(25.821)	(31.480)	(64.543)	(82.067)	(0.091)	(0.098)	(0.098)	(0.130)
Valle D'Aosta)	13.530	17.612	-7.273	-1.659	0.042	0.051	-0.023	-0.018
	(16.689)	(13.947)	(33.225)	(34.368)	(0.051)	(0.048)	(0.044)	(0.042)
Trentino A.A.)	74.424	2.660	105.667	6.504	0.172	0.003	0.019	-0.064
	(55.076)	(40.370)	(89.928)	(62.037)	(0.138)	(0.085)	(0.098)	(0.091)
Friuli V.G.)	4.844	-1.772	-69.184	-78.234	0.048	0.033	-0.180*	-0.189**
	(40.397)	(32.867)	(71.892)	(65.067)	(0.124)	(0.115)	(0.097)	(0.088)
Abruzzo)	-31.167	-24.328	-45.786	-36.285	-0.085	-0.070	-0.050	-0.042
	(34.578)	(26.442)	(60.685)	(44.010)	(0.102)	(0.080)	(0.046)	(0.037)
Basilicata)	22.333	21.824*	21.174	20.448	0.070*	0.069*	0.030	0.029
	(16.666)	(12.173)	(25.873)	(24.936)	(0.041)	(0.035)	(0.030)	(0.030)
Calabria)	30.633	24.431	8.332	-1.815	0.068	0.053	0.044	0.036
	(33.747)	(23.089)	(62.302)	(48.422)	(0.081)	(0.067)	(0.070)	(0.061)
Campania)	45.093	36.752*	20.282	8.099	0.119	0.100	0.028	0.019
	(30.985)	(20.933)	(50.544)	(41.887)	(0.073)	(0.060)	(0.067)	(0.065)
Lazio)	-27.725	-21.340	-109.044	-100.147*	-0.055	-0.039	-0.111	-0.103
	(36.196)	(26.244)	(65.648)	(59.015)	(0.104)	(0.078)	(0.080)	(0.073)
Molise)	0.490	-9.162	23.202	10.234	0.020	-0.001	0.009	-0.003
	(11.238)	(10.111)	(22.613)	(22.350)	(0.026)	(0.027)	(0.026)	(0.024)
Puglia	20.220	27.115	-19.038	-9.602	0.069	0.085	-0.082	-0.074
	(26.352)	(17.922)	(44.037)	(45.093)	(0.064)	(0.055)	(0.062)	(0.056)
Sardegna)	13.713	30.386	-93.484	-70.859	-0.009	0.028	-0.064	-0.045
	(33.176)	(24.960)	(98.117)	(96.658)	(0.105)	(0.094)	(0.072)	(0.067)
Sicilia)	64.756	41.058	-34.984	-68.508	0.133	0.082	-0.004	-0.038
	(57.971)	(41.481)	(125.468)	(101.205)	(0.174)	(0.132)	(0.102)	(0.091)
Toscana)	-30.825	5.257	-62.937	-14.652	-0.083	-0.005	-0.153**	-0.119*
	(25.584)	(21.438)	(54.792)	(52.906)	(0.076)	(0.072)	(0.065)	(0.062)
Umbria)	8.308	2.395	-35.445	-44.202	0.019	0.006	-0.009	-0.017
	(17.634)	(11.263)	(34.948)	(27.060)	(0.032)	(0.025)	(0.041)	(0.033)
any province)	-20.896	-127.134**	-101.810	-253.286*	-0.086	-0.326*	0.018	-0.110
	(44.824)	(48.356)	(81.008)	(140.031)	(0.127)	(0.173)	(0.105)	(0.128)
Region FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Provincial controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ln(Exposure To Outbreak)		Yes		Yes		Yes		Yes

Notes: each row corresponds to a separate set of estimations that replaces "Exposure To Outbreak" with exposure to a given region. Provinces located in that region are excluded from the sample. Columns 2, 4, 6 and 8 include ln(Exposure to Outbreak) as control. Mean of dependent variable, R-squared and number of observations are omitted to keep the table readable. Province controls as in Table 1, except for the last row (where overall propensity to migrate is shown explicitly). Robust standard errors in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

TABLE A.12: ROBUSTNESS TEST: DOES WITHIN-PROVINCE MOBILITY MATTER?

	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: COVID DEATHS						
Dep var.	Number of deaths 20Feb-31Mar	Number of deaths 20Feb-31Mar	Number of deaths April	Number of deaths April	Number of deaths May	Number of deaths May
ln(Exposure To Outbreak)	99.376*	93.033*	-72.239	-83.298	-10.772	-18.001
	(50.937)	(52.136)	(108.350)	(110.836)	(61.085)	(59.906)
IHS(# trips from outbreak areas)	47.225*	50.926*	211.723***	218.175***	26.738	30.956
	(27.253)	(27.026)	(61.626)	(63.352)	(22.784)	(22.741)
IHS(# trips w/i own province)		-34.728		-60.540		-39.579
		(48.391)		(91.653)		(50.504)
Mean	86.968	86.968	155.054	155.054	45.430	45.430
R-squared	0.923	0.924	0.884	0.885	0.790	0.795
PANEL B: TOTAL DEATHS						
Dep var.	Number of deaths 20Feb-31Mar 2015-2019	Number of deaths 20Feb-31Mar 2020	Growth of deaths Jan-Feb 2020 vs 2015-2019	Growth of deaths March 2020 vs 2015-2019	Growth of deaths April 2020 vs 2015-2019	Growth of deaths May 2020 vs 2015-2019
ln(Exposure To Outbreak)	0.097	0.098	0.041	0.006	0.119	0.102
	(0.124)	(0.131)	(0.112)	(0.107)	(0.081)	(0.077)
IHS(# trips from outbreak areas)	0.193**	0.192*	0.110**	0.130***	-0.026	-0.017
	(0.089)	(0.096)	(0.050)	(0.046)	(0.035)	(0.035)
IHS(# trips w/i own province)		0.006		-0.189		-0.091
		(0.158)		(0.121)		(0.077)
Mean	0.203	0.203	0.167	0.167	-0.048	-0.048
R-squared	0.907	0.908	0.928	0.931	0.662	0.671
Region FEs	Yes	Yes	Yes	Yes	Yes	Yes
Province controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	76	76	76	76	76	76

Notes: trips from outbreak areas and within own province are per 1000 inhabitants. "Exposure To Outbreak" is the number of people who moved from the province to one of the outbreak areas during 2015-2018 (per 1000 inhabitants). Province controls as in Table 1. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1