Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Internal migration networks and mortality in home communities: Evidence from Italy during the Covid-19 pandemic

Michele Valsecchi a, *, Ruben Durante b, c, d, e, f

a New Economic School, Moscow, Russian Federation
b ICREA-UPF, Barcelona, Spain
c IPEG, Barcelona, Spain
d BSE, Barcelona, Spain
e CESifo, Munich, Germany
f CEPR, London, United Kingdom

ARTICLE INFO

JEL classification:
J61
R23
H12
I10

Keywords:
Internal migration networks
Information
Mobility
Health
Contagion
Virus
Covid-19

ABSTRACT

Do internal migration networks benefit or harm their home communities in case of a communicable disease? Looking at the spread of Covid in Italy and using pre-determined province-to-province migration, excess mortality and mobile phone tracking data, we document that provinces with a greater share of migrants in outbreak areas show greater compliance with self-isolation measures (information mechanism), but also a greater population inflow from outbreak areas (carrier mechanism). For a subset of localities, the net effect on mortality is negative. However, for the average locality, the effect is positive and large, suggesting that the role of migrants as information providers is trumped by their role as virus carriers. The effect is quantitatively important and could be incorporated in epidemiological models forecasting the spread of communicable diseases.

1. Introduction

Migrants are often considered a threat to places of destination (because of competition for jobs with natives and presumed effects on local crime). Meanwhile, they are often seen as a positive transformative factor for places of origin: they provide insurance via remittances, and they spur innovation, trade and democratic change via information and cultural transmission. One aspect that remains largely unexplored concerns the role that internal migration networks might play in spreading communicable diseases. In this setting, migrants located in outbreak areas could generate positive or negative spillovers on their home communities. They could

We thank Martin Bell, Simone Bertoli, Ruben Enikopolov, André Gröger, Joan Lull, Tatiana Mikhailova, Luigi Minale, Hannes Mueller, Mathilde Munoz, Maria Petrova, Sandra Sequeira, Gerhard Toews, Pierre-Louis Vezina and seminar participants at the CEPR-LISER-OECD Conference on “Immigration in OECD countries”, Universitat Autonoma and New Economic School 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 9th April. An earlier version of this study (Valsecchi (2020)) was circulated with the title “Internal migration and the spread of Covid-19” on CEPR Covid Economics (Issue 18, 15 May 2020). The preliminary idea appeared in Mikhailova and Valsecchi (2020). All errors are our own.

* Corresponding author.
E-mail addresses: mvalsecchi@nes.ru (M. Valsecchi), ruben.durante@upf.edu (R. Durante).

1 Our study focuses on the effect of pre-existing migration networks and return migration on home communities, and not on the effect of out-migration on home communities. The potential negative effects of the latter, in the form of brain drain, have been the object of extensive investigation and have received scant empirical support (Docquier and Rapoport, 2012). For example, Beine et al. (2008, 2010) find evidence of a positive effect.

https://doi.org/10.1016/j.euroecorev.2021.103890

Received 3 March 2021; Received in revised form 3 August 2021; Accepted 26 August 2021
Available online 9 September 2021
0014-2921/© 2021 Elsevier B.V. All rights reserved.
transmit information and raise awareness on the importance of health policy measures, thus decreasing contagion and, ultimately, mortality. By the same logic, the migrant could comply with the same measures and self-isolate. If instead the migrant traveled back to her hometown, she could end up harming it by increasing contagion and mortality. The direction and magnitude of the migrant’s spillover will depend on the private cost of self-isolation, the sense of belonging to her home community and the awareness of the threat she constitutes.

We study this question by looking at the spread of Covid-19 in Italy, one of the countries most affected by the pandemic. Using comprehensive data on the place of origin and destination of individuals having migrated over the previous years, we construct a measure of internal migrant networks in outbreak areas for each province in the rest of the country. We then examine whether provinces with larger networks experience lower local mobility (information channel) and greater inflow of people from outbreak areas (“carrier” channel). In turn, we examine the effect on mortality to determine which effect prevails.

One key aspect of our analysis is that we use information on migration patterns that pre-date the Covid-19 outbreak and cannot therefore be affected by it. For the analysis of local mobility and population inflows, we control for all time-invariant differences between provinces with larger and smaller migrant networks and, in addition, for all time-varying differences captured by region-time fixed effects and by province-time controls. Second, we exploit the fact that, following the outbreak of Covid in February 2020, most cases were concentrated in a small number of provinces. This allows us to control for total out-migration, which 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 local mobility. To further address this possibility, we explicitly control for a wide range of geographic and socio-economic indicators. Third, for the mortality analysis, our outcome of interest is excess mortality (relative to previous years), which is available for each province in each month. This alleviates the concern that Covid-19 deaths may be misreported or underestimated, and that such measurement error may differ between provinces. Our mortality analysis is cross-sectional and controls for all time-invariant differences captured by the region fixed effects and the predetermined controls. The main tests for the validity of the identification assumption rely on the absence of any effect when looking at total deaths just before the outbreak started and when replacing exposure to outbreak areas with exposure to any specific region.

Additional results rule out the possibility that this response is the consequence of the “carrier” channel. In Section 4.3., we also test for robustness to alternative cross-provincial relationships like social connectedness and long-standing business relationships between provinces with larger and smaller migrant networks and, in addition, for all time-varying differences captured by region-time fixed effects and by province-time controls. Second, we exploit the fact that, following the outbreak of Covid in February 2020, most cases were concentrated in a small number of provinces. This allows us to control for total out-migration, which 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 local mobility. To further address this possibility, we explicitly control for a wide range of geographic and socio-economic indicators. Third, for the mortality analysis, our outcome of interest is excess mortality (relative to previous years), which is available for each province in each month. This alleviates the concern that Covid-19 deaths may be misreported or underestimated, and that such measurement error may differ between provinces. Our mortality analysis is cross-sectional and controls for all time-invariant differences captured by the region fixed effects and the predetermined controls. The main tests for the validity of the identification assumption rely on the absence of any effect when looking at total deaths just before the outbreak started and when replacing exposure to outbreak areas with exposure to any specific region.

Additional results rule out the possibility that this response is the consequence of the “carrier” channel. In Section 4.3., we also test for robustness to alternative cross-provincial relationships like social connectedness and long-standing business relationships between a given province and the outbreak areas.

In Section 4.3., we also test for robustness to alternative cross-provincial relationships like social connectedness and long-standing business relationships between a given province and the outbreak areas.

Our work contributes to the growing literature on the diffusion of viruses. 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 viruses such as the closure of schools and public transportation (Adda, 2016; Litvinova et al., 2019), city lockdown (Chinaazzi et al., 2020; Fang et al., 2020;...
Kraemer et al., 2020), or the combination of several of them (Gatto et al., 2020, Hsiang 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), cultural traits (Barro et al., 2020; Durante et al., 2021; Egorov et al., 2020; Giuliani and Rasul, 2020) and social learning (Tian 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 Larguem, 2009; Pichler and Ziebarth, 2019), and social media connections (Bailey et al., 2020; Charoenwong et al., 2020; Holtz et al., 2020; Kuchler et al., 2020). One aspect that remains largely unexplored concerns the role that migration networks play in spreading the virus. Lee et al. (2020) look at the effect of out-migration on Covid infections in South Asia and find that international migration matters, while results for domestic migration are mixed. In contrast, we focus on pre-determined internal migration in outbreak areas controlling for overall out-migration, and find a large positive effect on mortality. Had we focused on total internal migration, we also would have found no effects. In addition, we use mobility measures based on mobile phone tracking data that allow us to disentangle two different (and conflicting) mechanisms, thus providing some insight on the conditions under which internal migration has a positive or negative effect. We discuss external validity and implications for policy and future research in the conclusions.

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. 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 million residents.

On March 9 the lockdown was extended to the entire country. 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.

3. Data and research design

To measure deaths, we use data from the Italian Ministry of Health elaborated by the department of Civil Protection. Data on total deaths are available at the provincial-“monthly” level (20th Feb.-31st March, April, May). Appendix section A describes the steps we take to identify outbreak provinces. The procedure leaves us with 15 provinces: 10 in Lombardy and 5 in Emilia-Romagna.

To measure the exposure of provinces to outbreak areas, we use yearly data on changes of residence. 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, the province of Catania (Sicily) and registered themselves in the province of Milan (Lombardy). Data on changes of residence that took place between 2015 and 2018 and divide them by the 2018 population. This is our ExposureToOutbreak indicator.

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

To measure population inflows from outbreak areas and local mobility, we use data on mobility based on mobile phone tracking data provided by Teralytics. The data track people’s movements across phone cells and provide the number of trips within the same province in the country during the same period. Data from the Italian Ministry of Health elaborated by the Department of Civil Protection and described in the next section.

Relative to epidemiological studies, we consider population inflows as outcome and show that it is only one of the mechanisms through which internal migration networks affects virus-related deaths.

Registry based data provide the important advantage of a level of disaggregation that survey data typically have not. However, they miss people who choose not to re-register themselves in their new province of residence. Since one of the main advantages to re-register oneself is to access basic services like the family doctor, it is important to test whether exposure to outbreak areas is balanced along health capacity. We discuss this and other potential sources of measurement error in Section 3.2.

The dataset is based on an agreement between Teralytics and one of the three largest mobile phone operators in Italy. A trip is defined as a movement across phone cells and it ends once the phone remains in the same cell for at least 60 min. The number of trips so defined is then interpolated (by Teralytics) to represent 100 percent of phone users. For within-province trips, the threshold is 15 min (but results are virtually identical to those with the 60 min threshold).

Local mobility data are missing for Pordenone and Udine (Friuli-Venezia Giulia region).
3.1. Descriptive statistics

Table A.1 show detailed descriptive statistics at the provincial level for 76 non-outbreak provinces.\footnote{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.}

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.3 shows the full distribution. Figure A.4 shows that provinces in Southern regions are more exposed, which is consistent with historical trends.\footnote{See footnote 1 for references.} Importantly for our analysis, which only exploits within-region variation, exposure to outbreak areas varies substantially across provinces within the same region, as depicted in Figure A.5. 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.

The number of daily trips within a province averages 1.26 per inhabitant. Daily trips are high before the 1st Covid death (1.80) and before the first restriction measures (1.72), then fall drastically during the lockdown (0.82 and 0.68). A similar pattern emerges for daily trips from outbreak areas: they average 3.42 trips per 1000 inhabitants; they are high before the 1st Covid death (6.09); they decline partially after the first restrictive measures (4.50); they fall drastically during the lockdown (1.34 and 1.02).

Data on total deaths are available in two forms. First, they are available in levels for the 20th February–31st March period, averaged over 2015–2019 and, separately, for 2020. When looking at 20th February 2020–31st March 2020, the total number of deaths is 1,292. During the same 40 days, Covid deaths are 87. The difference in magnitude between the two measures 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 A.1. Figure A.6 shows that provinces in the North experience more deaths. Importantly for our analysis, which only exploits within-region variation, excess mortality shows a much less pronounced North–South gradient once we partial out region fixed effects, as depicted in Figure A.7.

3.2. Econometric specification

We run two types of analysis: one on local mobility and population inflows at the province-day level; one on mortality using cross-sectional data at the province level.

The former takes the following form:

\[
\ln(Mobility_{r,p,\text{date}}) = \alpha_{r,p} + d_{\text{date}} + \sum_{\text{week}} \beta_{\text{week}} \left[ \ln(ExposureToOutbreak_{r,p}) \times d_{\text{week}} \right] \\
+ \delta_{r,\text{week}} + X'_{r,p,\text{week}} \Gamma + \epsilon_{r,p,\text{date}}
\]

where Mobility_{r,p,\text{date}} is either local mobility, either the number of daily trips per capita from outbreak areas.\footnote{See footnote 2 for references.}\footnote{For local mobility, we use the log transformation. For daily trips from outbreak areas, we use the Inverse Hyperbolic Sine (IHS).} The RHS indicators are the province FEs (\(\alpha_{r,p}\)), the day FEs (\(d_{\text{date}}\)), the region-week FEs (\(\delta_{r,\text{week}}\)), the interactions between exposure (\(ExposureToOutbreak_{r,p}\)) and week indicators, a set of interactions between (pre-determined) controls and week indicators (\(X'_{r,p,\text{week}} \Gamma\)), and the error term (\(\epsilon_{r,p,\text{date}}\)). Observations are weighted by population. Standard errors are adjusted for spatial correlation (\(\alpha'\) la Conley, 1999), with a 100 km threshold) and serial correlation.\footnote{Throughout the paper, we refer to Fixed Effects as FEs.}

The inclusion of province fixed effects controls for any time-invariant difference between more and less exposed provinces. The inclusion of region-week fixed effects restricts the comparison to provinces that are situated in the same region. The inclusion of provincial controls ensures that such within-region comparison is not biased by potential residual confounders.

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.

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 local mobility (or population inflows).

As a first test of the plausibility of the identification assumption, we test whether provinces more and less exposed to outbreak areas are similar in terms of observable characteristics. Figure A.8 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.26

The analysis on mortality takes the following form:

\[
\text{Deaths}_{r,p} = \alpha_r + \beta \ln (\text{ExposureToOutbreak}_{r,p}) + X'_{r,p} \Gamma + \epsilon_{r,p}
\]  
(2)

where \(\text{Deaths}_{r,p}\) is a measure of deaths in region \(r\) and province \(p\), \(\text{ExposureToOutbreak}_{r,p}\) is our exposure indicator, \(\alpha_r\) is a set of region fixed effects, \(X_{r,p}\) is a set of (pre-determined) controls, and \(\epsilon_{r,p}\) is the error term. Observations are weighted by population.

Standard errors are adjusted for spatial correlation (a’ la Conley, 1999), with a 100 km threshold.

First, we estimate the effect on the number of total deaths 20th Feb.–31st Mar. 2020. Second, we focus on the growth of total deaths in 2020 (relative to the 2015–2019 average) at the province level for, separately, March, April and May.

To further validate our research design, we estimate the following placebo estimations. First, we estimate the effect on total deaths 20th Feb.–31st Mar. averaged during 2015–2019. Second, we estimate the effect on the growth of total deaths in 2020 (relative to the 2015–2019 average) for January–February. Third, we re-estimate the main specifications replacing exposure to outbreak areas with exposure to specific regions. In addition, we test whether estimates are robust to controlling for the share of earlier migrants, which will give a sense of whether long-run economic ties between provinces are driving the results. We estimate the main specifications with and without province controls, which form the basis for additional tests based on Oster (2019) that suggest that remaining unobservable factors are unlikely to drive our results. Finally, we consider alternative outcome measures (Covid deaths, Covid cases) and exposure measures (based on different definitions of outbreak areas).

---

26 Among the covariates, the only indicators that stand out are distance to outbreak areas and, to a lower extent, general propensity to migrate. Regarding distance to outbreak areas, panels (C) and (E) show that the correlation is driven by one specific region that borders the outbreak areas. In one of the robustness checks, we show that the main results are robust to dropping this region (Table A.17). Regarding general propensity to migrate: exposure to outbreak areas is a component of the general propensity to migrate, so a positive correlation between the two is not surprising. Most importantly, general propensity to migrate does not seem to be correlated with the outcomes of interest (Table A.10, last row). In addition, we disaggregate general propensity to migrate into propensity to migrate to any single region. We then estimate the effect of propensity to migrate to each region with and without controlling for propensity to migrate to outbreak areas. Results (also in Table A.10) suggest that these specific propensities to migrate are seldom relevant when we do not control for exposure to outbreak areas, even less when we control for the latter. Last but not least, our exposure indicator is not correlated with any measure of state or health capacity. This suggests that measurement error in the exposure indicator (driven, for example, by university students not registering themselves in the province of residence) may not constitute a threat to our identification strategy. By using log of exposure, we also address the concern that systematic under-reporting of changes of residence may bias upwards our estimates. Hence, measurement error, if any, may only generate an attenuation bias on the estimates.
4. Results

We will now present the reduced form relationship between exposure to outbreak areas and (i) local mobility, (ii) population inflows from outbreak areas, and (iii) excess mortality. In turn, we will discuss mechanisms using additional heterogeneity analysis.

4.1. Effect on local mobility and population inflows

Fig. 1 shows the coefficient estimates associated with the specification at the provincial-daily level (equation (1)). Panel A shows that greater exposure to outbreak areas is associated with no effect on local mobility until the end of March, after which the effect becomes negative.27 A one percent increase in exposure is associated with 0.124 percent fewer daily trips per capita, which implies that a 50 percent increase in exposure relative to the mean (i.e., about 2 additional migrants per thousand people)28 would be associated with 0.042 fewer trips per capital29 (i.e., 6.2 percent relative to the April average). The results are consistent with larger internal migration networks in outbreak areas slowly influencing compliance with stay-at-home measures in their hometowns. We discuss (and rule out) alternative interpretations in Section 4.4.

Panel B shows the effect on population inflows from outbreak areas using the same specification.30 First, there is no differential increase in trips from outbreak areas neither during normal times (i.e., January) nor during the beginning of the outbreak (i.e., 1st Feb.-23rd Feb.). Second, there is a differential increase in trips from outbreak areas after the first restrictive measures but before the national lockdown (i.e., 24th Feb.-8th Mar.). Third, the effect for the 24th Feb.-1st Mar. is similar to the effect for the 2nd Mar.-8st Mar., which suggests no panic effect.31 Fourth, 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 first restrictive measures supports the hypothesis that some recent migrants returned to their hometowns following the first restrictive measures in outbreak areas. The magnitude of the effect seems non-negligible: a one percent increase in exposure is associated with 0.14 percent additional trips from outbreak areas,32 which implies that a 50 percent increase in exposure relative to the mean would be associated with 0.3 additional daily trips per thousand people.33

The estimates associated with the specification without province FE (Table A.3, Columns 1–2) suggest that the effect on regular return migration34 is much larger than the effect on population inflows driven by the first restrictive measures: 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. Importantly, the difference between the two effects does not necessarily convert into equally different roles in spreading the virus.35 However, it does suggest that regular return migration might deserve much more attention than the one received so far by the media, which focused instead on migrants who remained jobless from the first restrictive measures.36

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. Results (Tables A.4, A.5 and discussion in Online Appendix, Section C) suggest that it is unlikely that our results are driven by omitted variables.

4.2. Effect on mortality

Table 1, Panel A, shows the coefficient estimates associated with the specification at the province level (Eq. (2)). Column 1 shows the estimates for 20th Feb.–31st 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).

Columns 3–6 show the results of the growth estimations. Column 3 shows that exposure has no effect on the Jan.-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), 0.203 in April (Column 5) and 0.080 in May (Column 6). Based on the 2015–2019 average total deaths for these months,37 these effects corresponds to, respectively, 1.62, 0.93

27 See also Table A.2, Column 4, for the coefficient estimates.
28 To be precise, a 50 percent increase in exposure relative to the mean corresponds to 2.24 additional migrants per thousand people. The standard deviation of the exposure measure is 2.04. Hence, an increase in exposure relative to the mean matching exactly a s.d. would be 45.5 percent relative to the mean.
29 This is the result of 0.124 (average coefficient estimate for April) * 50 (percentage increase in exposure) /100 * 0.679 (daily trips per capita during April according to Table A.1).
30 See also Table A.3, Column 4, for the coefficient estimates.
31 More disaggregated estimates (available upon request) confirm this finding.
32 See Bellemare and Wichman (2020) for a discussion of marginal effects in models including an Inverse Hyperbolic Sine transformation.
33 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.1).
34 With the term regular return migration, we include also, for example, weekly commuting.
35 For example, return migration driven by the first restrictions might have higher infection rates and/or be associated with more reckless behavior.
36 For media emphasis on the role of migrants driven by the first restrictive measures, see references in footnote 2. Fortunately, the academic literature has not neglected the issue as much: see Fajgelbaum et al. (2020) for some interesting recent work on optimal lockdown in a commuting network with an application to Seoul, Daegu and New York.
37 These are 424 (March), 458 (April) and 492 (May). The 2015–2019 average for March is obtained by multiplying the 2015–2019 average for 20th Feb.-31st 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.
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. Table A.6, Panel A, shows that coefficients are precisely estimated no matter the distance cutoff that one chooses for the Conley s.e. Estimates for Covid deaths (Table A.7) and Covid cases (Table A.8) confirm this general pattern.

4.3. Robustness, placebo estimations and back of the envelope calculation

Figure A.9 shows that the results are not driven by outliers. Table 1, Panel B, shows that the estimates do not depend on the large but potentially arbitrary set of province controls. Oster’s (2019) tests (Table A.9 and discussion in Online Appendix, Section D) suggest that it is unlikely that our results are driven by omitted variables.

Table A.10 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 and A.7, except for the outbreak regions where the outbreak provinces are located (Veneto, Lombardy and Emilia-Romagna). 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 the importance of having clearly defined outbreak areas and data on the number of migrants specific to these areas.

Table A.11 shows that our main results are substantially robust to controlling for social connectedness between a province and the outbreak areas. This reassures us that our previous results are not driven by more exposed provinces also being more socially connected to outbreak areas. Another province-to-province connection that might be correlated with Covid deaths and with internal migration networks is inter-provincial business. Here the concern is that long-standing inter-provincial business connections might cause both internal migration networks and Covid diffusion. Data on province-to-province transactions unfortunately do not exist and therefore one cannot test directly for this concern. However, one can think of an indirect test: long-standing business connections require repeated interactions and trust, and therefore are likely to shape both recent and earlier migration flows. By controlling for the latter, we test not only whether earlier migrants matter as much as recent ones, but also whether long-standing business connections are a likely driver of our main results. Table A.12 and A.13 show that the results are robust to these additional controls, which we find reassuring.

Table A.14 shows that the results are also not driven by the cutoff number of deaths used to define outbreak areas. Table A.15 shows that results are similar when using different weights for different outbreak provinces (depending on whether they had more Covid deaths or more Covid cases on the 1st of March). Table A.16 shows that considering provinces located in Marche region as outbreak also generates similar results. Table A.17 shows that dropping any entire region does not change affect the main results. Table A.18 shows that adjusting the computation of the standard errors following Young (2016) to address potential concerns with the distribution of the exposure indicator does not affect the main results either.

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 10th percentile of the exposure distribution. Appendix G describes the steps required for the calculation. The result is that they would have suffered 5,895 fewer Covid deaths and 7,348 fewer total deaths. These are important quantities, because they constitute 60 percent of all Covid deaths in non-outbreak regions and 18 percent of all Covid deaths in the country. This very large proportion is consistent with the idea that internal migration networks anticipated the arrival of Covid in a given location, which, given its highly transmissible nature, had devastating consequences. This said, it is worth emphasizing that this calculation is based on reduced-form estimates and not on state-of-the-art epidemiological models. Incorporating internal migration networks in the latter is a promising avenue for future research.

4.4. Heterogeneity analysis

To shed light on the importance of the two mechanisms, we exploit the fact that regular return migration should be much weaker in regions far away from outbreak areas, which in turn should weaken the role of the carrier mechanism there.
Hence, we re-estimate Eq. (1) including full interactions between exposure to outbreak areas, a North dummy and a South dummy47 (and excluding province FEs). Fig. 2, Panel A, shows that the effect on local mobility appears in both parts of the country (if any, the effect is stronger in the South). On the contrary, the effect on regular return migration (Panel B) is strong in the North but weak in the South.48 Consistent with the carrier mechanism being weaker in far away regions, Table 2, Panel A, shows that the effect on mortality is stronger in the North than in the South during both March and April (March: 0.430–0.471 vs 0.222–0.227; April: 0.253–0.455 vs 0.035–0.267), while it is equally strong in May (0.063–0.309 vs 0.134–0.432).

These results are useful to shed some light on the decrease in local mobility that we had seen in Fig. 2, Panel A. That effect was consistent with a response of home communities to: (i) greater information (or greater salience to existing information) brought by migrants (in person, via phone, via social medias); (ii) greater mortality brought by migrants (via the carrier channel). The weaker effect on return migration driven by first restrictive measures is positive and similar across North and South. After the lockdown, the effect remains positive (but weaker) in both regions. Such decline is similar in absolute levels, but, relative to the pre-lockdown effect, it is stronger in the South than in the North.

We then zoom in Southern regions to exploit another intuition: lower local mobility translates into fewer interactions (and therefore contagion) especially in urban areas, where population density is high. If that is the case, then the decrease in local mobility might be associated with a decrease in mortality especially in urban areas. Table 2, Panel B, shows that the coefficient estimate associated with the interaction with urbanization changes sign between March (+4.345, +4.214) and April (−7.579, −7.372). For provinces at the 90th percentile of the urban share distribution (urban share equal to 6.8 percent), the net effect on mortality is negative.49

Hence, while the population inflow mechanism explains a large part of the reduced form relationship between exposure and mortality,50 under certain conditions, the information channel can dominate the carrier mechanism.

---

47 The North indicator takes value 1 for Friuli-Venezia-Giulia, Liguria, Marche, Piemonte, Toscana, TrentinoAltoAdige, Umbria, Abruzzo. The South indicator takes value 1 for Basilicata, Calabria, Campania, Lazio, Molise, Puglia, Sardegna and Sicilia.

48 The effect on return migration driven by first restrictive measures is positive and similar across North and South. After the lockdown, the effect remains positive (but weaker) in both regions. Such decline is similar in absolute levels, but, relative to the pre-lockdown effect, it is stronger in the South than in the North.

49 We also estimate the effect on local movements and population inflows from outbreak areas across rural and urban areas. Figure A.10 shows that urban and rural areas in Southern regions do not differ much along these two dimensions. The combination of heterogeneous effects on mortality and homogeneous effects on mobility is consistent with the interpretation that similar decreases in mobility can have different mortality consequences because of varying population density.

50 To investigate whether and to what extent population inflows from outbreak areas explain the relationship between exposure to outbreak and mortality, we include the average number of trips from outbreak areas during February as additional controls in the mortality specification at the province-month level.
Table 2
Excess mortality: heterogeneity analysis.

| Period       | March 2020 vs 2015–2019 | April 2020 vs 2015–2019 | April 2020 vs 2015–2019 | May 2020 vs 2015–2019 | May 2020 vs 2015–2019 |
|--------------|-------------------------|-------------------------|-------------------------|----------------------|----------------------|
| (1)          | (2)                     | (3)                     | (4)                     | (5)                  | (6)                  |

PANEL A: North vs South

| In(Exposure to Outbreak) × North | 0.430*** (0.146) | 0.471*** (0.135) | 0.253*** (0.083) | 0.455** (0.181) | 0.063 (0.049) | 0.309*** (0.123) |
|-------------------------------|------------------|------------------|------------------|------------------|----------------|------------------|
| ln(Exposure to Outbreak) × South | 0.222* (0.121) | 0.227* (0.131) | 0.035 (0.056) | 0.267 (0.163) | 0.134* (0.069) | 0.432*** (0.115) |
| Mean                          | 0.203            | 0.203            | 0.167            | 0.167            | −0.048         | −0.048           |
| R-squared                     | 0.592            | 0.613            | 0.559            | 0.574            | 0.465          | 0.530            |
| Observations                  | 76               | 76               | 76               | 76               | 76             | 76               |

PANEL B: urban vs rural areas in Southern regions

| In(Exposure to Outbreak) × urban | 4.225*** (1.613) | 4.214**−7.579*** (1.739) | −7.372*** (1.704) | −0.929 (1.735) | −0.486 (3.384) | −0.884 (3.884) |
| Mean                          | 0.048            | 0.048            | −0.006            | −0.006            | −0.063         | −0.063           |
| R-squared                     | 0.786            | 0.786            | 0.900            | 0.911            | 0.559          | 0.600            |
| Observations                  | 38               | 38               | 38               | 38               | 38             | 38               |

Region FEs | Yes | Yes | Yes | Yes | Yes | Yes |
Province controls | Yes | Yes | Yes | Yes | Yes | Yes |
SCI controls | Yes | Yes | Yes | Yes | Yes | Yes |

Notes: the dependent variable is the growth of total deaths in a given month of 2020 relative to the 2015–2019 average for the same month. Province controls as in Fig. 1. Social Connectedness Index (SCI) controls are: “SCI With Outbreak”, which captures population weighted SCI between a province and the outbreak areas; and “SCI(all)”, which captures the population weighted SCI between a province and the rest of the country. Conley standard errors with 100 km threshold in round brackets. Robust standard errors in square brackets. ***p < 0.01, **p < 0.05, *p < 0.1.

5. Conclusions

In this paper, we asked whether internal migration networks can have negative spillovers on migrants’ home communities in case of highly transmissible diseases. We found evidence consistent with a positive information channel and evidence consistent with a negative carrier channel. On average, the carrier channel prevails.

To get a sense of the external validity of this finding, we looked at the relationship between pre-determined internal migration rates and Covid deaths across countries. Figure A.11 shows that this relationship is positive.

Future research on the economics of viruses could build on our findings in two ways. First, researchers could collect individual level data to study whether migrants learn slowly about the threat that they constitute; or whether there are informed and uninformed migrants, with the former driving the positive spillovers and the latter driving the negative spillovers; or whether...
migrants hold contradictory beliefs about others’ contagiousness as opposed to their own. Second, researchers could incorporate internal migration networks in state-of-the-art epidemiological models, which could then predict how the virus would have spread, had the outbreak been in a different area (i.e., an area with different network centrality). These models could then be used to derive the optimal allocation of resources across the national territory and the optimal timing of local lockdowns. Lockdowns are very costly, so knowing in advance which places should be prioritized is valuable. By using pre-determined internal migration data, central governments can actually predict the spread of a virus before it actually spreads.\footnote{This is in line with recent awareness that well-defined and updated pandemic plans are key to managing the outbreak of a virus (Giuffrida, 2021). Waiting to have real time mobility data to develop a plan would essentially be the same as not having a plan, which some considered to have caused 10,000 deaths in Italy alone (Giuffrida and Boseley, 2020).} Finally, one could also think about measures to financially support internal migrants and to increase their awareness about the risks of returning to their place of origins for themselves, their families, and the community at large.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.euroecorev.2021.103890.

References

Adda, Jerome, 2016. Economic activity and the spread of viral diseases: Evidence from high frequency data. Q. J. Econ. 131 (2), 891–941.
Bailey, Michael, Cao, Rachel, Kuchler, Theresa, Stroebel, Johannes, Wong, Arlene, 2018. Social connectedness: Measurement, determinants, and effects. J. Econ. Perspect. 32 (3), 259–280.
Bailey, Michael, Cao, Rachel, Kuchler, Theresa, Stroebel, Johannes, Wong, Arlene, 2020. Social networks shape beliefs and behavior: evidence from social distancing during the covid-19 pandemic. 32, (28234),
Barmby, Tim, Larguem, Makram, 2009. Coughs and sneezes spread diseases: An empirical study of absenteeism and infectious illness. J. Health Econ. 28 (5), 1012–1017.
Barro, Robert J., Ursúa, José F., Weng, Joanna, 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, Rapoport, Hillel, Steinmayr, Andreas, Trebesch, Christoph, 2017. The effect of labor migration on the diffusion of democracy: Evidence from a former soviet Republic. Am. Econ. J. Appl. Econ. 9 (3), 36–69.
Beach, Brian, Ferrie, Joseph P., Saavedra, Martin H., 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, Docquier, Frédéric, Rapoport, Hillel, 2008. Brain drain and human capital formation in developing countries: Winners and losers. Econom. J. 118 (528), 631–652.
Beine, Michel, Docquier, Frédéric, Rapoport, Hillel, 2010. On the robustness of brain gain estimates. Ann. Econ. Stat. 97–98, 143–165.
Bellemare, Marc F., Wichman, Casey J., 2020. Elasticities and the inverse hyperbolic Sine transformation. Oxford Bull. Econ. Stat. 82 (1), 50–61.
Bertoli, Simone, Guichard, Lucas, Marchetta, Francesca, 2020. Turnout in the Municipal Elections of 2020 and Excess Mortality during the COVID-19 Epidemic in France, IZA Working Paper No. 13335.
Bonifazi, Corrado, Heinz, Frank, 2000. Long-term trends of internal migration in Italy. Int. J. Popul. Geogr. 6, 111–131.

This is in line with recent awareness that well-defined and updated pandemic plans are key to managing the outbreak of a virus (Giuffrida, 2021). Waiting to have real time mobility data to develop a plan would essentially be the same as not having a plan, which some considered to have caused 10,000 deaths in Italy alone (Giuffrida and Boseley, 2020).
Voigtländer, Nico, Voth, Hans-Joachim, 2012. The three horsemen of riches: Plague, war, and urbanization in early modern Europe. Rev. Econom. Stud. 80 (2), 774–811.

Yang, Dean, 2008. International migration, remittances and household investment: Evidence from philippine migrants’ exchange rate shocks. Econ. J. 118 (528), 591–630.

Young, Alwyn, 2016. Improved, nearly exact, statistical inference with robust and clustered covariance matrices using effective degrees of freedom corrections.