

Close encounters during a pandemic: social habits and inter-generational links in the first two waves of COVID-19*

Annalisa Cristini[†] **Pedro Trivin[‡]**

Università degli studi di Bergamo

April 4, 2023

Abstract

Social habits are ingrained in a community and affect human behaviour. Have they played any role in the spread of the pandemic? We use high-frequency data for 220 regions in 15 European countries from March to December 2020 to compare the association between social contacts outside the family and within inter-generational families, on the one hand, and cases and excess mortality on the other. We find that a standard deviation increase in the percentage of people having daily face-to-face contacts outside the household is associated with 5 new daily cases and 2.6 additional weekly deaths, while the incidence of inter-generational households exhibits a less robust association with both COVID-19 transmission and mortality. We compare results across the first and the second wave of pandemic and show that differences are related to the average age of the most affected groups. Our findings are robust to the inclusion of a number of controls, fixed effects, the chosen sample of countries, and the estimation method. We argue that type and frequency of social interactions are interweaved with a region culture and habits and are informative on the potential transmission of contagion and on its lethality.

JEL Codes: I1, I12, R10.

Keywords: COVID-19; social contacts; inter-generational households; virus contagion; excess mortality.

*We thank the Editor and two anonymous referees for helpful comments on an earlier version of this paper. We also thank the participants at different seminars and conferences for their helpful comments and suggestions.

[†]Department of Economics, University of Bergamo, 24127 Bergamo; email: annalisa.cristini@unibg.it.

[‡]Corresponding author. Department of Economics, University of Bergamo, 24127 Bergamo; email: pedro.trivin@unibg.it.

1 Introduction

When the pandemic reached Europe in early 2020, Italy was the first country to be dramatically hit: by March 1st, it had registered almost 1,700 cases, more than half the total cases in Europe; by the end of March, the cases had risen to over 105,000 in Italy and almost 96,000 in Spain; on the whole, the two countries accounted for over 46% of the total cases in Europe.

Demographic and social characteristics typical of these Mediterranean countries gained a special attention as potential amplifiers of an epidemic that was disproportionately affecting the old.¹ The indirect transmission of the virus to the elderly due to interactions between adult children and old parents and/or between grandchildren and grandparents was deemed relevant both in Europe (Dowd et al., 2020) and in the US (Harris, 2020, 2021) and family ties appeared then worth investigating. Inter-generational contacts coupled with a high incidence of aged people could also explain the observed high death rates in simulation models (Esteve et al., 2020).

Various studies used the share of multi-generational households in the country to proxy inter-generational links but the evidence on the expected statistical correlation with the spread of coronavirus disease 2019 (COVID-19) was not clear-cut (Arpino et al., 2020).² Given the pronounced within-country heterogeneity observed both in COVID-19 outbreaks (Naqvi, 2021) and in the incidence of inter-generational families (Duranton et al., 2008; Kaasa et al., 2014), analyses performed at the country level could, at least partly, justify the weak evidence. In addition, the relatively few data points available at the beginning of the pandemic, together with the small number of countries considered (Bayer and Kuhn, 2020; Mogi and Spijker, 2021) were also constraining, with a few exceptions (Aparicio Fenoll and Grossbard, 2020), the degrees of freedom so that potential confounders were also a problem (Belloc et al., 2020).

The link between COVID-19 and social interactions *outside* the household was initially less investigated.³ However, as the epidemic spread to other regions and countries, and the effects of social distancing measures were being assessed, the role of social interactions, *lato sensu*, gained further interest. Rodríguez-Pose and Burlina (2021) find that, after controlling for a bulk of economic, demographic, geographical and institutional factors,

¹Dowd et al. (2020) report that in Italy as of March 30, 2020, 96.9% of deaths were occurring in people aged 60 and over.

²Indeed, for influenza the epidemiological literature agrees that children are a major driver of transmission (Mossong et al., 2008; Salathé et al., 2010). However, as data on COVID-19 epidemic accrued, epidemiological evidence was finding younger cohorts to play a less relevant role than in the transmission of influenza (Davies et al., 2020; Harris, 2021; Forbes et al., 2021).

³For an exception, see Mogi and Spijker (2021); they use the percentage of people having frequent social meetings as one of the variables in a factor analysis. They find that it has a positive loading on the factor that positively relates to the spread of the virus.

the share of the population meeting friends and family socially at least once a week is positively related to excess mortality across European regions in the first six months of the pandemic. Moreover, social capital, a notion which is ultimately rooted in dense social networks and sense of reciprocity (Putnam, 2000; van Oorschot et al., 2006) is found to be negatively associated with excess mortality (Rodríguez-Pose and Burlina, 2021)⁴ and positively related to citizens' adherence to social distancing rules (Giuliano and Rasul, 2020; Borgonovi and Andrieu, 2020; Durante et al., 2021; Bartscher et al., 2021)

In this paper, we consider both a measure of face-to-face social interactions outside the family and a proxy of inter-generational links within the household. This allows us to capture two important aspects of a community social structure, both of which can relate to the virus transmission. We perform the analysis at the regional level and extend the time period to comprise the first and the second COVID-19 wave. These enriched spatial and temporal scopes provide the analysis with potentially informative additional variation in the data.

Specifically, we use NUTS 2 European data for a maximum of 220 regions in 15 European countries and exploit daily or weekly data on COVID-19 cases and mortality from the beginning of the pandemic to the end of 2020. This allows us to decrease the influence of unobserved heterogeneity by including NUTS 1 fixed effects and country-day (or country-week) specific dummies. We also account for the time lags between new cases and cases confirmed in the past 14 days and allow the correlation with past cases to differ within and between regions, according to Adda (2016).

We measure social interactions outside the household by the share of people that have daily contacts with friends or neighbours. Notice that with respect to epidemiological surveys, which trace the contacts that each person in the survey has with each other, our information render a depiction of the context in which the spread is occurring and in this sense it can be regarded as a macro or aggregated map, complementary to the typical epidemiological contact matrices and, as argued below, it can reflect culture-related attitudes.

To reduce the blur ingrained in the social contact data, in addition to a battery of controls explained below, we consider only those encounters that have been shown to facilitate the spread of diseases predominantly transmitted via droplets, i.e., face-to-face contacts⁵ that take place daily or almost daily (Read et al., 2008; Mossong et al., 2008; Salathé et al., 2010; Mastrandrea et al., 2015). The regional map of social contacts is

⁴The authors find that “participation in voluntary associations that encourage relationships between dissimilar or unfamiliar people” is highly significant throughout most of the specifications and always more relevant than a measure of generalized trust.

⁵For the sake of exposition, in the paper we use interchangeably face-to-face, skin-to-skin, physical contacts, and daily contacts.

derived from the latest European Quality of Life Survey (EQLS) conducted in 2016.⁶ As argued by the relational sociology literature (Dépelteau, 2018), the structure of social contacts, as part of network relations, is deeply interweaved with culture, which also contributes to the stability of social systems more generally (Fuhse, 2013). For this reason, the pattern of social relationships is expected to show some persistence across time and remain informative of features relevant for the virus transmission.

We consider the spread of COVID-19 in terms of new cases per day (or per week) and use excess mortality data instead of fatalities or case-fatality rate (CFR), which have been shown to suffer from measurement errors due to under reporting of cases and deaths and lack of comparability across countries due different ways to register COVID-19 deaths (Aron and Muellbauer, 2020; Rodríguez-Pose and Burlina, 2021).

According to some evidence (Monod et al., 2021), the channels of COVID-19 transmission changed between the two waves, also in response to the enforcement and lifting of non-pharmaceutical interventions. Using a subsample of countries for which data on cases and fatalities are available by age groups, we test whether the correlation between cases and social contacts differs by age groups across the two waves.

Throughout the analysis, we account for the possibility that social interactions could be correlated with other regional socioeconomic characteristics; for instance, skin-to-skin contacts are expected to be positively related to population density and employment rate, but negatively to the share of elderly people (see for example Brown and Ravallion, 2020). By including a rich set of controls, in our analysis we separate face-to-face contacts associated with the urban, economic and demographic structure of the region from those related to regional culture and habits.

Our analysis adds to the literature in several ways. First of all, we extend the inter-generational family-centred analysis to a broader concept of social networks, encompassing interactions both within and outside the household and compare their associations with the pandemic. In addition, we place this analysis at the regional level, whereas the role of multi-generational families on COVID-19 has typically been tested at the country level. Moreover, by considering a large number of regions and daily and weekly COVID-19 data on cases and excess mortality for the whole relevant period of 2020, we can address the weaknesses of this extant literature due to potential confounding factors and measurement errors in CFR. Finally, we contribute to the analyses that investigate the evolution of the pandemic, especially the changes intervened between the first and the second wave (Forbes et al., 2021). At this regard, by assessing how the correlation of interest has changed across

⁶EQLS is a survey of the adult population (18+) living in private households, based on a statistical sample and covering a cross-section of society. Depending on country size and national arrangements, the 2016 sample ranged from 1,000 to 2,000 people per country. The survey, developed by Eurofound, a tripartite EU agency, is carried out every four to five years.

time and by age groups, we add to the analyses on the age profile of the cases.

We find that face-to-face daily contacts help explaining the spread of the virus. In particular, when using weekly data, we observe that a standard deviation increase in the percentage of people having daily face-to-face contacts is associated, on average, with 5 new daily cases and with 2.6 more weekly excess deaths. On the contrary, the association between inter-generational families and COVID-19 cases and excess mortality is weak and not robust. When we allow for these relationships to vary over time, we find that regions with a larger share of daily contacts are associated with a larger share of COVID-19 cases in both waves, while the positive relationship with COVID-19 excess mortality is stronger in the first wave. We show that these results could be related to heterogeneity in the transmission of the virus across different groups of age in the different waves. Elderly people are more affected during the first wave, while the cases of cohorts under 60 years of age are those most strongly associated with daily contacts during the second wave. Again, no significant relations emerge along time between the share of inter-generational families on the one hand and COVID-19 spread or excess mortality on the other.

The paper is organized as follows. The next section introduces the data and presents some stylized facts. Section 3 describes the empirical specification and discusses the inclusion of control variables; regression results and robustness checks are presented in Section 4. Section 5 draws the main conclusions, pointing out the limits of the analysis and its usefulness to the understanding of the present pandemic.

2 Data and stylized facts

2.1 Data

The first variable of interest, to be correlated with COVID-19 spread and mortality, is a measure of social interactions. Specifically, we consider only those types of encounters that are sufficiently proximate in space and allow a direct transmission of the virus through droplets (Read et al., 2008; Mossong et al., 2008; Salathé et al., 2010; Mastrandrea et al., 2015).

We also exclude contacts with relatives outside the household, as the share of relatives can depend on migration trends and/or on work mobility patterns more generally, which could bias the comparison across regions.

The measure we use is based on the type and frequency of encounters as stated by representative individuals interviewed in social surveys. Among the various international surveys, EQLS distinguishes contacts according to whether they involve face-to-face proximity and also distinguish between contacts with relatives versus contacts with friends and

neighbours outside the household;⁷ precisely, we define our variable of interest as:

FACE-TO-FACE CONTACTS: The percentage of people in the region that answered **daily or almost daily** to the question: *How often do you have face-to-face contacts with friends or neighbours living outside the household?*

As underlined by the sociological literature, social contacts result from individual relationships, which in turn adhere to the available cultural models (see for example [Dépelteau, 2018](#); [Fuhse, 2018](#); and [Donati, 2018](#)); in this sense, the habits of seeing friends and neighbours represent a cultural feature of a society, which is the element we intend to capture. In the light of this, we also expect our variable of interest to show some degree of persistence in time.

Inter-generational households are defined as those in which more than one generation live together; hence an inter-generational family is one in which the respondent co-habits with his/her children, and/or (grand)parents, and/or siblings of an age difference of at least 20 years. Our measure is similar to that used by [Mogi and Spijker \(2021\)](#). Others use more restrictive definitions, based on the share of cohabiting adult children aged 18-34, ([Aparicio Fenoll and Grossbard, 2020](#)) or as heads of households aged 30-49 that live with their parents ([Bayer and Kuhn, 2020](#)).

Regional data on COVID-19 cases are obtained from the COVID-19 European Regional Tracker developed by Asjad Naqvi ([Naqvi, 2021](#)). COVID-19 mortality is proxied using weekly excess mortality from Eurostat. The variable is computed as the difference between the number of deaths in a given week in 2020 and the average number of deaths in the same week in the period 2015-2019. In our analysis, we include a maximum of 220 regions from 15 European countries, and COVID-19 related data stretch from the beginning of the pandemic till the end of December 2020.⁸ Most of our control variables are provided by Eurostat, the only exceptions being the poverty rate and the share of multi-generational families, obtained from the EQLS.⁹ Tables [A.2](#) and [A.3](#) in Appendix A show, respectively, the definitions and the descriptive statistics of the variables considered.

⁷SHARE -Survey of Health, Ageing and Retirement in Europe- provides detailed measures of social contacts but it addresses people above 50 only and it does not differentiate between face-to-face and online interactions. Similarly, ESS -European Social Survey- does not explicitly distinguish between face-to-face and distant contacts and asks about the frequency of social meetings in general, involving either/both relatives, friends and/or colleagues.

⁸The countries included in our analysis are: Austria, Belgium, Czech Republic, Germany, Denmark, Spain, Finland, France, Hungary, Italy, Netherlands, Poland, Romania, Sweden and UK. Table [A.1](#) in Appendix A shows the specific sample included in our preferred specification.

⁹Although Eurostat also provides some regional poverty measures, the coverage is much limited, reducing our sample by almost 60%.

2.2 Face-to-face contacts, inter-generational households and COVID-19 spread

Figure 1 shows the regional distribution of face-to-face contacts and of the inter-generational households, as defined above. The correlation coefficient between the two variables is positive and significant. On average, 39% of the families living in the regions considered are inter-generational and 42% of people state to have daily or almost daily face-to-face contacts outside the households. The variability across regions, both within and between countries, is high: regions with the largest values of daily contacts are located in the Center and South of Italy, in the Western parts of Spain and in the North of Scotland. Regions with the lowest values are in the North and South-West of France and in most of Sweden. Regarding inter-generational families, in most Spanish and Polish regions the incidence is above the top quartile (48%). A particularly high share is also present in some regions of Italy, in a few regions of France, while such high percentages are very rare in Germany, Austria, Netherlands and absent in Denmark, Belgium, Finland, Sweden, Czech Republic, and Hungary.

Figure 1: Face-to-face contacts and inter-generational families

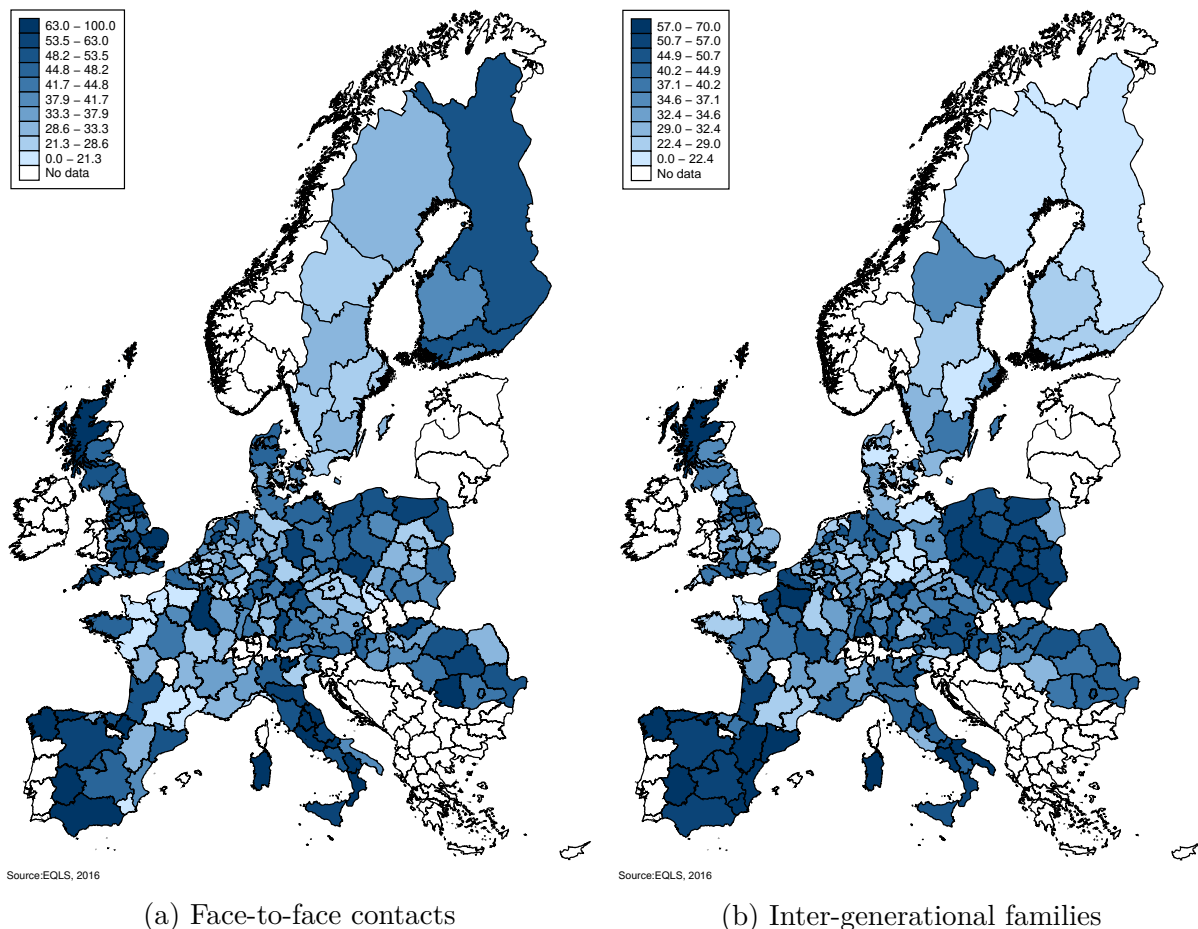


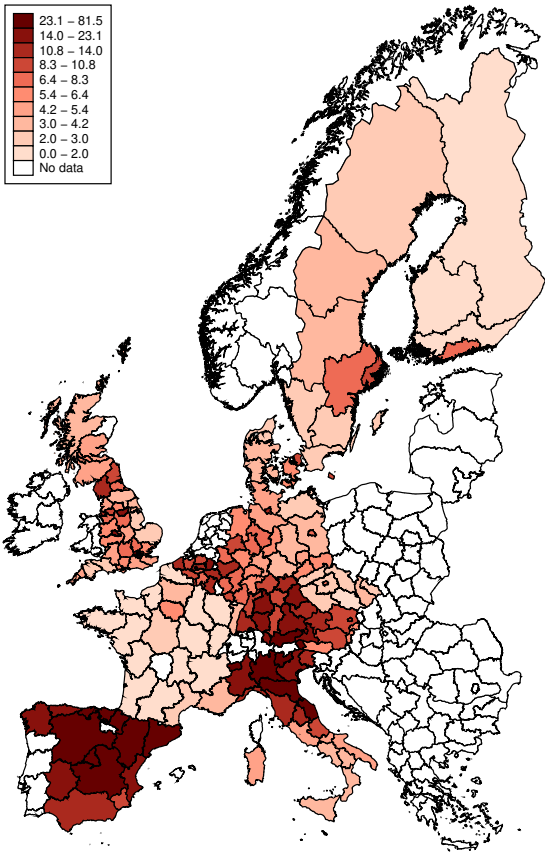
Figure 2 shows the evolution of the COVID-19 pandemic in our sample. In particular, the four panels display the number of cumulative COVID-19 cases per 10,000 people at four different moments of time.

The well known initial epicenters that at the end of March 2020 are clearly recognizable in Lombardy and most parts of Spain (panel a), extend to some regions of the UK, Belgium and a few areas of Sweden by the end of June (panel b). In September the situation worsens in the South of France (panel c) and by December 21st, the last day with data available for the whole sample of 220 regions, the diffusion has increased in Eastern Europe, particularly in some regions of Czech Republic and Poland (panel d).

Figure A.1 completes the picture by showing the country-specific evolution of the number of COVID-19 cases and the excess mortality using weekly data. Regarding the number of cases, a general pattern is that the second COVID-19 wave, that hit European countries from week 30, is associated to a larger number of infections in all countries.¹⁰ Excess mortality data is instead more heterogeneous across countries; specifically, the sample is divided in two, with half of the countries showing larger mortality in one of the two waves.

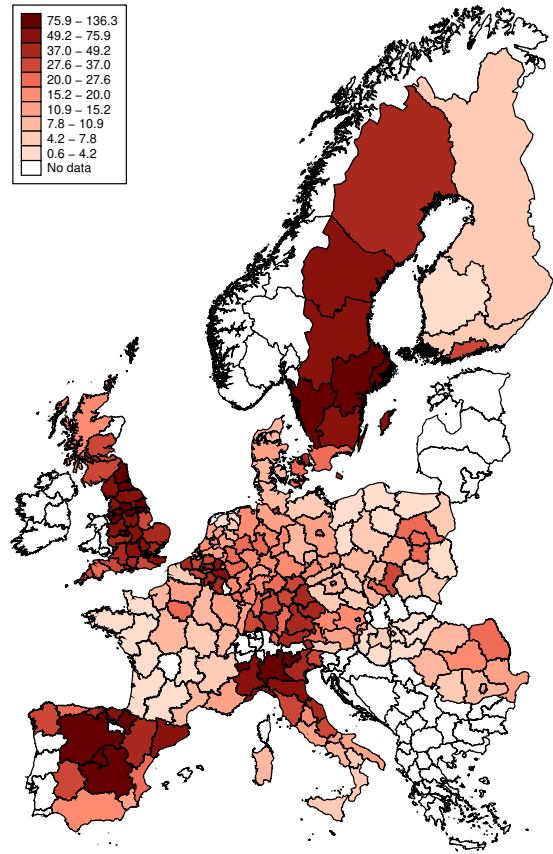
¹⁰We have to be careful when interpreting this result as the first wave is likely characterised by a larger number of under-reported cases. Below we explain that this fact is not affecting our results.

Figure 2: Cumulative COVID-19 cases per 10K people



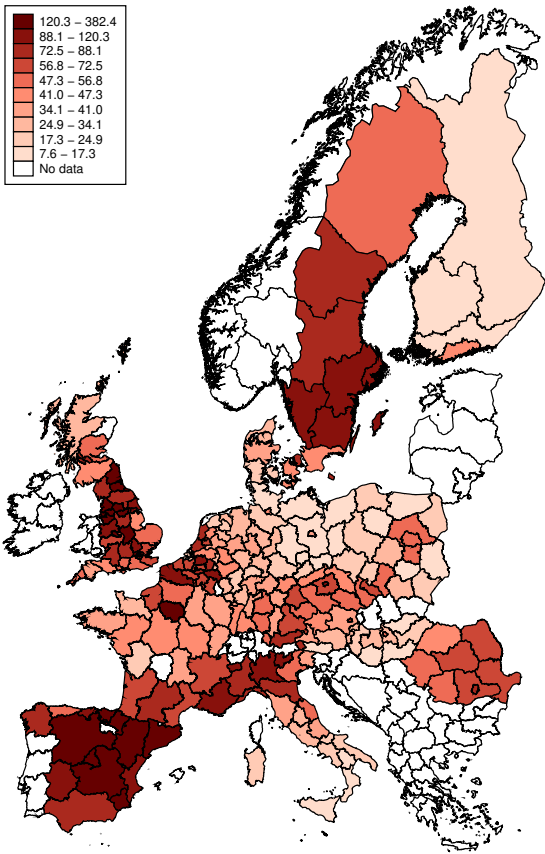
Source: COVID-19 European Regional Tracker

(a) March 31st



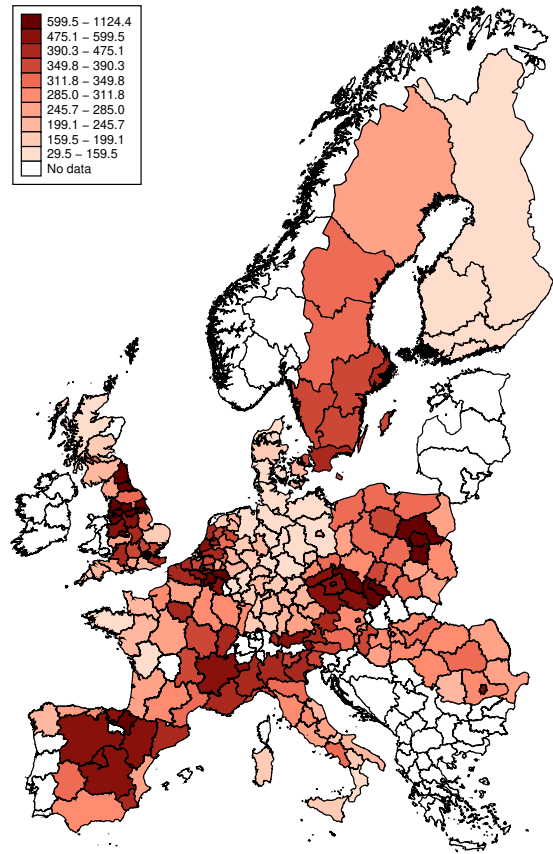
Source: COVID-19 European Regional Tracker

(b) June 30th



Source: COVID-19 European Regional Tracker

(c) September 30th



Source: COVID-19 European Regional Tracker

(d) December 21st

3 Empirical strategy

To estimate the role of social habits in the pandemic, we model the spread of the virus following Adda (2016) and Qiu et al. (2020) and explicitly account for the dynamics of the virus transmission, by including lagged cases within and between regions, as well as for socio-economic indicators and fixed effects. Specifically, we estimate an equation such as:

$$Y_{it} = \sum_{s=1}^{14} \alpha_{within,s} Y_{i,t-s} + \sum_{s=1}^{14} \alpha_{between,s} \sum_{i \neq j} d_{ij}^{-1} Y_{j,t-s} + \beta_{11} X_i + \beta_{12} G_i + \beta_2 \mathbf{Z}_i + \mu_{nuts} + I^c * \phi_t + \epsilon_{it}, \quad (1)$$

where subscript i indicates the NUTS 2 region, Y refers to the number of new confirmed COVID-19 cases in region i on date t , d_{ij} is the log of the distance between regions i and j , and $\sum_{i \neq j} d_{ij}^{-1} y_{j,t-s}$ is the inverse log distance weighted sum of new cases in other regions. X represents our variable of social contacts and G the percentage of multi-generational families. \mathbf{Z} is a row vector of control variables to be detailed below, μ_{nuts} are NUTS 1 fixed effects, $I^c * \phi_t$ are country-specific time dummies, ϵ is a zero mean white-noise residual, and β_{11} and β_{12} are our main parameters of interest.

The inclusion of $\alpha_{within,s}$ and $\alpha_{between,s}$ accounts, respectively, for within and between region COVID-19 dynamics and captures the spatial spread of the virus. The inclusion of up to 14 lagged days is based on the estimated duration of the infectious and the incubation period of the COVID-19. Using a different number of lags (between 7 and 13) does not change our results.¹¹ The inclusion of NUTS 1 fixed effects mitigate problems related to regional unobserved heterogeneity, while country-specific time fixed effects aim to control for different COVID-19 related policy interventions and behavioural changes across countries.¹²

The magnitude of the COVID-19 crisis has provoked an avalanche of studies on the determinants of the virus; in our analysis we draw from them and include a rich set of controls that can be classified into three groups: baseline, demographic and economic controls, and regional idiosyncrasies. As baseline controls we include six variables that have been widely acknowledged and are commonly used as the main determinants of the spread of the virus: GDP per-capita, which accounts for the economic activity and more general regional economic specificities; number of heating degree days, as corona-type

¹¹Adda (2016) and Qiu et al. (2020) focus on the dynamics of the transmission and use lagged weather episodes to instrument the within and between lagged components. Since the spread dynamics of the virus is not the main focus of this paper, for the baseline analysis we use an OLS estimator, which allows us to consider a larger number of regions. In Section 4.4 we account for the dynamic endogeneity problem and show that our results are robust to using IV and to an alternative empirical specification.

¹²In our empirical analysis we also show that our results are robust to the inclusion of NUTS 1-specific time dummies.

viruses are normally seasonal and worsen with cold weather; population density, as the higher it is, the higher the probability of skin-to-skin contacts between an infected person and a susceptible one, as it may happen in busy public transports, markets and supermarkets, cafes and restaurants. Beyond the inclusion of lags of the dependent variable, we account for the stage of the epidemic curve by including the number of days since the first COVID-19 case was detected. If this information is not available, we include the number of cases per population in the first available date. Finally, following [Brown and Ravallion \(2020\)](#), we also include the logarithm of the population, which allows for a non-homogeneous relationship between the current number of infections and the population size.

The second block of controls includes variables related to the structure of the economy and demographics. In addition to the GDP per capita, the economic environment is captured by the income poverty rate, which is expected to be inversely related to the capability of adjusting to the required behavioural changes, as well as by labour market indicators and production sectors. Specifically, we consider the employment rate, the education level of the workforce and the share of employment in the service sector. All have a bearing on the way of living, which may in turn facilitate or hamper the transmission of the virus. For example, small craft businesses are likely to travel across local areas and regions, have contacts with different and numerous households and businesses to whom they provide their services; on the contrary, jobs in the advanced tertiary sector can in most cases be performed remotely, with minimum face-to-face contacts. Evidence for Italy, for example, corroborates the idea that areas specialized in manufacturing, which comprises activities largely involving skin-to-skin tasks are comparatively more likely to be subject to COVID-19 infections ([Ascani et al., 2020](#)). Occupations that are unsuitable for remote work and require workers to work close to others (transportation, food-related, and personal care and service occupations) have also experienced an unprecedented rise in sickness-related absences during COVID-19 ([Lyttelton and Zang, 2022](#)).¹³ Regarding the demographic variables, in the analysis we include the share of people aged 65 or more and the women to men ratio as the virus appears to affect more men than women and hit

¹³We are aware that interactions at the workplace might not be fully captured by our variables. Unfortunately, to the best of our knowledge, there are no better available proxy at this granularity level. This drawback could be problematic if our variable of interest, the share of people having daily face-to-face contacts with friends or neighbours, was systematically correlated with a particular employment component affecting the spread of the virus (e.g., workers that cannot avoid working neck-to-neck to other people indoors). We do not think this is the case for three reasons: i) there is no correlation between the share of face-to-face contacts and the education level (-0.021) and the share of employment in the service sector (-0.0042), which are general measures of the employment structure and because of this should be more related, if anything, to the cultural components of the region; ii) all our regressions include NUTS 1 fixed effects, implying that we are exploiting variation across regions that are relatively similar; and iii) our results are robust to different empirical specifications and sample selection, which makes us confident that our results are not significantly affected by other factors.

older people more often.

Finally, we include a set of more heterogeneous factors that could still be relevant for the spread and mortality of COVID-19. Environmental factors such as pollution (Murgante et al., 2020) or humidity (Mäkinen et al., 2009) have usually been found to be important determinants of respiratory virus diseases similar to COVID-19. However, given the complex relationship between these factors and COVID-19 (e.g., different particles in the air could have different effects; pollution, humidity and temperature interact with each other...) we opt for including an indicator that correlates with these factors in a general way: the crude death rate for diseases of the respiratory system for people aged 65 and more. Another potential factor in the transmission of COVID-19 is the inter-connectivity of the region with the rest of the world as regions with larger connectivity are more likely to be exposed to the virus. We proxy the connectivity of a region by the number of air passengers carried per population.¹⁴

4 Regression results

In this section, we present our main results. Section 4.1 displays the baseline findings on the importance of social contacts and inter-generational families on the transmission and mortality of the virus; in Section 4.2 we use a more flexible approach and allow the estimated coefficient of the variables of interest to vary over time. In Section 4.3, we decompose the spread of the virus by groups of age. Finally, Section 4.4 shows that the results are robust to alternative sample selection, to weighting the data and to different estimation methods.

4.1 Baseline results

4.1.1 Cases: daily data

Table 1 shows the importance of social contacts and inter-generational families on the transmission of the virus using daily data. Results are separated in two blocks depending on the fixed effects included in the specification. Columns [1]-[3] include NUTS 1 and

¹⁴As we have commented before, the number of skin-to-skin contacts in a region is likely to be correlated with other socioeconomic characteristics. By including a rich set of controls in the analysis we try to separate the *cultural* component of social contacts from face-to-face contacts that occur due to the urban, economic and demographic structure of the region. Although some of these variables may be related, they attempt to capture different concepts in order to reduce the unobserved heterogeneity as much as possible. Using a more parsimonious regression does not alter our findings. Table A.2 in Appendix A defines the variables used in our analysis. Besides the variables directly related to the pandemic (cases and mortality), our control variables are time-invariant. From Eurostat we use data from the last available year, usually 2018 or 2019.

country-specific time effects while columns [4]-[6] control for NUTS 1-specific time dummies. Our preferred specification corresponds to the first block, where a larger number of regions is included. The second block shows the robustness of our results to more restrictive specifications regarding unobserved heterogeneity.

Within each block, we present three different versions of equation (1): the first specification includes the baseline controls, the second adds demographic and economic controls, and the final one further considers air connectivity and deaths due to respiratory illnesses.

The first remarkable result is the positive and significant impact of face-to-face contacts on the number of COVID-19 cases, regardless of the fixed effects included. While in our preferred specification the positive impact of daily contacts is observed independently of the controls included and of the share of inter-generational households, when we control for NUTS 1-specific time effects the positive effect is significantly different from zero when all the control variables are considered (Column 6).¹⁵ According to the results of the first three columns, one percentage point increase in social contacts raises the number of daily cases by around 0.09. This coefficient decreases to around 0.07 if we include NUTS 1-specific time effects. In other words, if we assume an increase of 1 standard deviation of face-to-face contacts, the number of daily cases increases by 1.56 ($0.096 * 16.26$), taking the coefficient of column 3. On the contrary, the share of inter-generational households is never significant across the specifications of Table 1.

Regarding the control variables, we find positive and robust relationships between the number of cases and population size, population density and cold temperatures, as it has been widely acknowledged by the epidemiology literature, as well as between cases and our air connectivity measure in our preferred specification. In contrast, we uncover a negative association with the share of people over 65 and the level of education of the workforce. The latter correlation in particular is consistent with the fact that the more educated white-collars could in most cases work from home, thus reducing the risk to be infected, while the less educated blue-collars were engaged in jobs where distancing was more difficult to maintain, let alone the possibility of smart working. The negative sign of the share of people over 65 is also in line with the importance of the workplace in the transmission of the virus. Using US counties data, [Brown and Ravallion \(2020\)](#) also find a negative impact of the share of elderly people on the transmission of COVID-19. They argue that “with higher retirement rates, the elderly will tend to face less economic pressure to be active outside home. Time-use surveys for the US indicate that elderly people have substantially lower contact rates in normal times ([Cornwell, 2011](#)).” ([Brown and Ravallion, 2020](#), p. 6). The number of days since the first case is also negatively

¹⁵Note that this result is likely related to the fact that this specification exploits a lower source of variation to identify the relationship of interest.

associated with the number of cases, coherently with the idea that the epidemic curve first stabilizes and then decreases over time.

Using daily data we can exploit high-frequency within-region variation to estimate the impact of face-to-face contacts on new COVID-19 cases. However, there are two drawbacks associated to the use of daily data: i) it is likely that the data are subject to a measurement error and ii) we do not have reliable information on COVID-19 fatalities at that frequency. Next, we overcome these issues by using weekly data.

4.1.2 Cases: weekly data

In this section, we decrease potential measurement error issues by computing the weekly average of the new daily COVID-19 cases. We also study the relationship between our variables of interest and COVID-19 fatalities using weekly excess mortality as a proxy. In this case, equation (1) includes only 2 lags of the within and between region transmission variables.¹⁶

Tables 2 and 3 display the results. Regarding the number of COVID-19 cases, Table 2 confirms the positive relationship between face-to-face contacts and coronavirus spread and the statically irrelevant association with the incidence of inter-generational families. This time, face-to-face contacts are positively associated with COVID-19 cases throughout all the specifications and the estimated relationship is stronger than before. Specifically, according to column [3] (our preferred specification), one percentage point increase in face-to-face contacts is associated with 0.312 new daily cases, on average. In other words, a standard deviation increase in the share of daily contacts is associated with an increase of 5 new daily cases, on average. Regarding the controls, the results of the previous section are confirmed. Moreover, we now observe a positive impact of the employment rate on the number of COVID-19 cases, reinforcing the idea that contacts in the workplace or during the commuting time are relevant for the transmission of the virus.

Table 3 shows the results on weekly excess mortality, where the latter is the difference in the number of deaths in a given week in 2020 with respect to the average number of deaths in the same week in the period 2015-2019. Again, we find a positive relation with face-to-face contacts regardless of the controls and the fixed effects included in the regressions. One percentage point increase in the share of daily face-to-face contacts is associated with 0.16 additional weekly deaths, i.e., with 2.6 extra deaths per week for a standard deviation increase of face-to-face contacts. In contrast with what we found on daily and weekly cases, we also find a positive association with the share of inter-

¹⁶When we estimate the relevance of face-to-face contacts and inter-generational households on excess mortality, we include the lags of the dependent variable. In Section 4.2.2, we also include the number of COVID-19 cases as an explanatory variable.

Table 1: Cases: Daily data

	[1]	[2]	[3]	[4]	[5]	[6]
Face-to-face contacts	0.086 (0.041)**	0.092 (0.043)**	0.096 (0.043)**	0.067 (0.045)	0.076 (0.047)	0.079 (0.047)*
$\ln(Days)$	-6.580 (2.545)**	-6.276 (2.521)**	-6.706 (2.515)**	-5.516 (2.242)**	-5.109 (2.087)**	-5.616 (2.112)**
$\ln(Population)$	13.077 (1.369)**	12.919 (1.342)**	12.671 (1.348)**	11.613 (1.365)**	11.528 (1.338)**	11.361 (1.352)**
Initial cases	7.849 (8.128)	7.042 (7.190)	7.397 (7.304)	6.818 (8.427)	5.387 (7.460)	5.662 (7.546)
$\ln(GDP_{pc})$	-5.787 (1.968)**	-0.738 (2.641)	-0.975 (2.464)	-5.340 (2.093)**	-0.565 (2.734)	-0.799 (2.607)
$\ln(Density)$	2.947 (0.954)**	1.807 (0.922)*	1.893 (0.886)**	3.182 (1.027)**	2.353 (1.019)**	2.431 (0.985)**
$\ln(Temperature)$	23.217 (6.598)**	20.111 (5.948)**	20.275 (5.998)**	23.513 (6.827)**	20.579 (6.359)**	20.806 (6.432)**
Inter-generational		0.049 (0.043)	0.034 (0.040)		0.051 (0.042)	0.038 (0.040)
Share 65		-1.155 (0.280)**	-1.213 (0.298)**		-1.096 (0.296)**	-1.150 (0.318)**
$\frac{Women}{Men}$		0.177 (0.339)	0.238 (0.335)		0.142 (0.356)	0.193 (0.353)
$\frac{Emp_{services}}{Emp_{total}}$		0.074 (0.133)	0.044 (0.131)		0.067 (0.140)	0.048 (0.137)
Employment rate		0.185 (0.155)	0.171 (0.154)		0.265 (0.168)	0.257 (0.167)
High education		-0.680 (0.290)**	-0.726 (0.292)**		-0.698 (0.302)**	-0.737 (0.304)**
Poverty rate		-0.073 (0.055)	-0.078 (0.056)		-0.082 (0.060)	-0.085 (0.061)
$\ln(Respiratory)$			-3.669 (3.915)			-3.431 (4.223)
Air connection			0.103 (0.061)*			0.070 (0.065)
Constant	-368.475 (67.186)**	-391.783 (78.303)**	-365.007 (76.962)**	-247.018 (80.366)**	-263.913 (89.633)**	-240.532 (91.060)**
NUTS 1 FE	Yes	Yes	Yes	No	No	No
Country FE*Date FE	Yes	Yes	Yes	No	No	No
NUTS 1 FE*Date FE	No	No	No	Yes	Yes	Yes
Observations	62698	62698	62698	57676	57676	57676
# NUTS2	220	220	220	203	203	203
Rsq	0.946	0.946	0.946	0.952	0.952	0.952
RMSE	140.528	140.527	140.529	123.483	123.479	123.482

Notes: The dependent variable is the number of daily COVID-19 cases. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include up to 14 lags of the dependent variable and the inverse distance weighted sum of new cases in other regions.

Table 2: Cases: Weekly data

	[1]	[2]	[3]	[4]	[5]	[6]
Face-to-face contacts	0.254 (0.108)**	0.291 (0.108)***	0.312 (0.108)***	0.206 (0.115)*	0.246 (0.115)**	0.263 (0.114)**
ln (<i>Weeks</i>)	-36.736 (10.425)***	-36.795 (10.291)***	-37.628 (10.340)***	-29.439 (9.452)***	-29.723 (9.287)***	-30.683 (9.373)***
ln (<i>Population</i>)	33.782 (3.888)***	32.895 (3.554)***	32.002 (3.551)***	29.568 (4.051)***	28.862 (3.674)***	28.138 (3.697)***
Initial cases	-0.240 (1.268)	1.558 (1.342)	1.968 (1.410)	-0.264 (1.301)	1.359 (1.449)	1.703 (1.504)
ln (<i>GDP_{pc}</i>)	-19.474 (6.004)***	-9.195 (7.951)	-9.465 (7.323)	-17.555 (6.083)***	-7.617 (8.102)	-7.868 (7.540)
ln (<i>Density</i>)	8.868 (2.706)***	6.053 (2.577)**	6.184 (2.491)**	8.894 (2.821)***	6.874 (2.753)**	6.964 (2.658)***
ln (<i>Temperature</i>)	68.297 (17.137)***	56.737 (15.556)***	56.181 (15.691)***	65.314 (17.473)***	54.997 (16.179)***	54.485 (16.321)***
Inter-generational		0.163 (0.115)	0.126 (0.107)		0.165 (0.116)	0.134 (0.107)
Share 65		-3.128 (0.745)***	-3.255 (0.782)***		-2.891 (0.789)***	-2.999 (0.831)***
$\frac{Women}{Men}$		-0.067 (0.898)	0.075 (0.868)		-0.121 (0.947)	-0.001 (0.920)
$\frac{Emp_{services}}{Emp_{total}}$		0.277 (0.334)	0.163 (0.326)		0.231 (0.346)	0.135 (0.338)
Employment rate		0.692 (0.410)*	0.638 (0.403)		0.786 (0.434)*	0.739 (0.428)*
High education		-1.665 (0.786)**	-1.779 (0.805)**		-1.653 (0.817)**	-1.747 (0.836)**
Poverty rate		-0.187 (0.137)	-0.206 (0.137)		-0.198 (0.145)	-0.213 (0.146)
ln (<i>Respiratory</i>)			-5.719 (9.648)			-4.788 (9.979)
Air connection			0.371 (0.176)**			0.312 (0.188)*
Constant	-912.294 (171.619)***	-875.210 (196.381)***	-819.535 (191.658)***	-660.412 (207.261)***	-611.388 (227.057)***	-567.311 (224.862)**
NUTS 1 FE	Yes	Yes	Yes	No	No	No
Country FE*Date FE	Yes	Yes	Yes	No	No	No
NUTS 1 FE*Date FE	No	No	No	Yes	Yes	Yes
Observations	8989	8989	8989	8269	8269	8269
# NUTS2	220	220	220	203	203	203
Rsq	0.955	0.955	0.955	0.965	0.965	0.965
RMSE	120.539	120.498	120.505	99.862	99.778	99.785

Notes: The dependent variable is the weekly average of the number of daily COVID-19 cases. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include up to 2 lags of the dependent variable and the inverse distance weighted sum of new cases in other regions.

generational households. A 10% increase in the share of inter-generational households, which is close to a standard deviation, is associated to around one additional extra death per week. Hence the within-household inter-generational contacts appear less than half as lethal as outside-household social interactions.

Regarding the control variables, similar to the results observed for the number of new COVID-19 cases, we find a positive relationship between excess mortality and population size and employment rate. Interestingly, we also find a positive association between the number of weeks since the first case and the excess mortality. This indicates a larger persistence of COVID-19 mortality relative to the number of new cases, and it is in line with the nature of the virus, where a person may die after several weeks since the infection. Finally, we find evidence of a negative relationship with the share of workers employed in the service sector and with the poverty rate, which highlight the relevance of the job place and the economic idiosyncrasies of a region.

So far, we have estimated the average association between face-to-face contacts and inter-generational households, on the one hand, and COVID-19 spread and mortality across time and regions, on the other. However, given the nature of the pandemic and the timing of social distancing measures, we expect that the habit of daily face-to-face interactions with friends and neighbours bear a different relevance at different moments of the pandemic. For example, regions with a larger share of daily face-to-face contacts might show a relatively larger spread at the beginning of the pandemic, when uncertainty about the virus is higher, behavioural recommendations are given but mobility restrictions are not yet enforced. Likewise, within families, once the pandemic became clear, distancing measures had at time being adopted, especially where some members of the households used to get out for work, while older and/or frailer ones were staying at home. In the next section we use rolling window regressions to investigate how the estimated correlation changes in time.

4.2 Time-varying analysis

In this section, we allow for a time-varying impact of face-to-face contacts and inter-generational households on the number of new COVID-19 cases and excess mortality by estimating a rolling version of equation (1). In particular, we estimate equation (1) using a rolling time window of 30 (7) days (weeks) from the beginning of the pandemic until the end of December. To maximize the sample included in the regression, we include 10 (2) lags of the dependent variable and the inverse distance weighted sum of regional cases. In our estimations, we include NUTS 1 fixed effects and country-specific time dummies.

Table 3: Excess mortality: Weekly data

	[1]	[2]	[3]	[4]	[5]	[6]
Face-to-face contacts	0.124 (0.079)	0.163 (0.085)*	0.158 (0.085)*	0.116 (0.071)	0.149 (0.079)*	0.145 (0.079)*
ln (<i>Weeks</i>)	40.364 (19.252)**	42.179 (19.186)**	41.836 (19.215)**	16.959 (18.008)	18.418 (18.345)	17.918 (18.352)
ln (<i>Population</i>)	12.739 (2.609)***	11.846 (2.045)***	11.824 (1.938)***	11.376 (2.416)***	10.780 (1.926)***	10.736 (1.790)***
Initial cases	0.028 (0.937)	0.120 (0.944)	0.338 (1.037)	-0.057 (0.928)	0.089 (0.952)	0.314 (1.052)
ln (<i>GDP_{pc}</i>)	-3.826 (1.593)**	1.452 (3.934)	0.865 (3.933)	-3.438 (1.513)**	1.510 (3.595)	0.939 (3.548)
ln (<i>Density</i>)	-0.062 (0.882)	1.793 (1.154)	1.975 (1.228)	0.049 (0.832)	1.340 (1.112)	1.522 (1.193)
ln (<i>Temperature</i>)	23.746 (16.972)	19.844 (13.654)	20.904 (13.526)	22.324 (16.011)	18.714 (12.626)	19.725 (12.592)
Inter-generational		0.120 (0.068)*	0.103 (0.062)*		0.085 (0.062)	0.068 (0.057)
Share 65		-0.567 (0.468)	-0.700 (0.513)		-0.603 (0.459)	-0.736 (0.514)
$\frac{Women}{Men}$		0.297 (0.513)	0.373 (0.491)		0.317 (0.483)	0.395 (0.463)
$\frac{Emp_{services}}{Emp_{total}}$		-0.520 (0.208)**	-0.497 (0.213)**		-0.396 (0.196)**	-0.380 (0.200)*
Employment rate		0.646 (0.350)*	0.666 (0.356)*		0.559 (0.343)	0.577 (0.350)
High education		-0.537 (0.492)	-0.625 (0.525)		-0.547 (0.478)	-0.636 (0.519)
Poverty rate		-0.185 (0.094)**	-0.185 (0.098)*		-0.148 (0.085)*	-0.149 (0.089)*
ln (<i>Respiratory</i>)			-8.352 (7.243)			-8.190 (7.049)
Air connection			-0.035 (0.105)			-0.018 (0.104)
Constant	-428.945 (144.527)***	-477.282 (146.385)***	-435.735 (148.666)***	-340.107 (122.708)***	-390.521 (122.607)***	-348.644 (116.601)***
NUTS 1 FE	Yes	Yes	Yes	No	No	No
Country FE*Week FE	Yes	Yes	Yes	No	No	No
NUTS 1 FE*Week FE	No	No	No	Yes	Yes	Yes
Observations	8916	8916	8916	8176	8176	8176
# NUTS2	220	220	220	203	203	203
Rsqr	0.859	0.859	0.859	0.923	0.923	0.923
RMSE	77.392	77.371	77.377	52.746	52.705	52.709

Notes: The dependent variable is the weekly excess mortality. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include up to 2 lags of the dependent variable.

4.2.1 Rolling window: Daily cases

Figure 3 displays the estimated coefficient and the 90% confidence interval of face-to-face contacts on new daily COVID-19 cases from the beginning of the pandemic to the end of December 2020. For the sake of interpretation, the period is split into two intervals (February-July and August-December) corresponding to the two COVID-19 waves affecting countries in 2020.

As we expected, the estimated relation between new cases and face-to-face contacts is strong and rising at the beginning of the pandemic, with the peak reached in correspondence to the sample including observations from March 6th to April 6th; at this stage, one percentage point increase in daily contacts is associated with 0.37 new daily cases. From there onward, the coefficient decreases rapidly, coherently with the lockdown measures applied in most countries, and finally loses statistical significance around July.

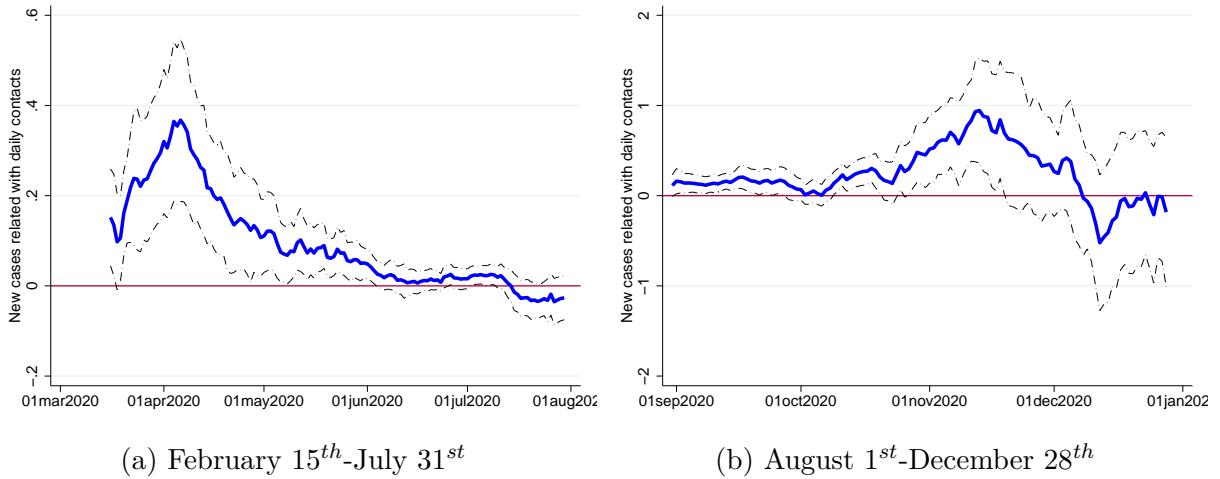
Regarding the second interval, Figure 3.b shows that from October, when the second COVID-19 wave hit countries across Europe, a larger share of face-to-face contacts is again positively associated with a larger number of COVID-19 cases; the estimated coefficient reaches a maximum magnitude of 0.94 with the sample including observations from October 13th to November 13th and swiftly declines thereafter, losing statistical significance one week later.

In the first wave, the estimated coefficient of the share of inter-generational families (Figure 4.a), displays a slightly different dynamic to that of face-to-face contacts; the estimated coefficient peaks a month later, towards the end of April and then rises again towards the end of July 2020. The significance, however, is very limited in time and in the second wave (Figure 4.b), the coefficient is never statistically significant.

The relevance of face-to-face interactions outside the households even after people had experienced the lethality of the epidemic in the first wave, is consistent with the idea that the type and frequency of social interactions that characterize each region are, at least in part, culture- and habits-related; hence, they surfaced again, once mobility restrictions were lifted. On the contrary, could they signal the presence of civic capital, we would have expected, after the first wave, a negative relationship with the spread of the virus; the reason being that in regions with high social capital people care relatively more about the generalized other and would be more cautious with face-to-face interactions once the restrictions were lifted (Bartscher et al., 2021).¹⁷

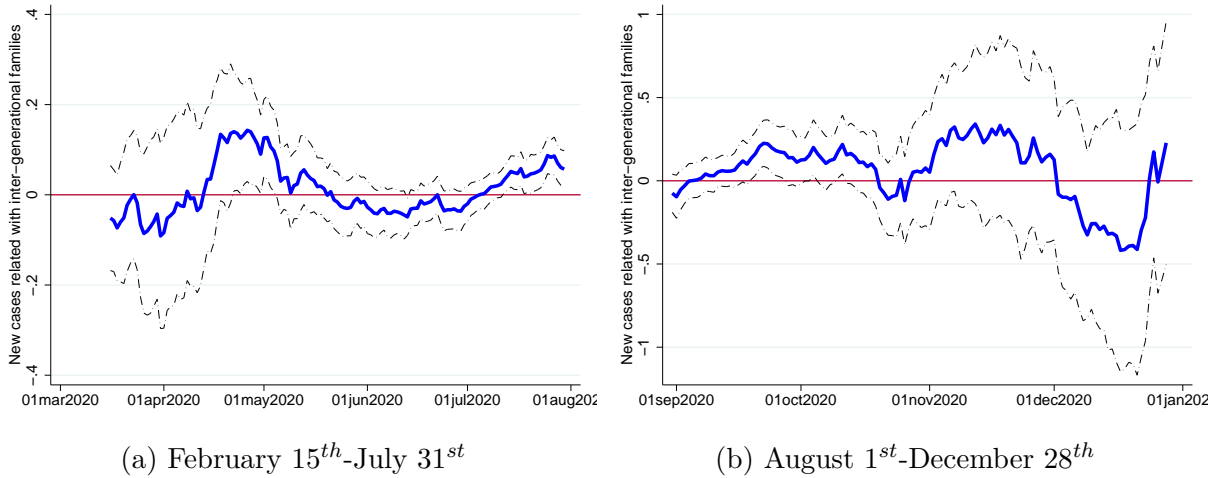
¹⁷In fact, face-to-face contacts can be only an incomplete proxy of social capital: while dense social networks are necessary for social capital to be formed, they are not sufficient. Social capital is often measured by the share of volunteers, cooperative and non-for profit associations, blood donors, and/or generalized trust. See for example Putnam (2000); Giuliano and Rasul (2020); Rodríguez-Pose and Burlina (2021).

Figure 3: Daily cases and face-to-face contacts: 30 days rolling window (full sample)



Notes: Figures show the marginal effect of face-to-face contacts on daily COVID-19 cases. Dashed lines indicate 90% confidence intervals. Coefficients are obtained from a regression such as column [3] in Table 1 using a 30 days rolling window. The regression includes up to 10 lags of the dependent variable and the inverse distance weighted sum of new cases in other regions. Only regions with at least 5 observations are included.

Figure 4: Daily cases and inter-generational families: 30 days rolling window (full sample)



Notes: Figures show the marginal effect of inter-generational families on daily COVID-19 cases. Dashed lines indicate 90% confidence intervals. Coefficients are obtained from a regression such as column [3] in Table 1 using a 30 days rolling window. The regression includes up to 10 lags of the dependent variable and the inverse distance weighted sum of new cases in other regions. Only regions with at least 5 observations are included.

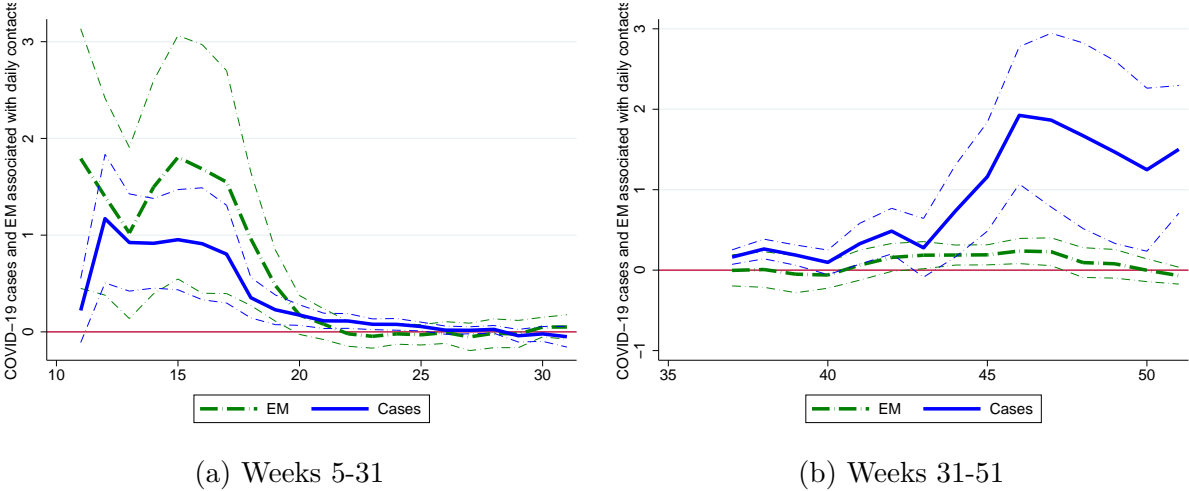
4.2.2 Rolling window: weekly data

In Figures 5 and 6 we show the results using weekly data and compare the time-varying associations between our variables of interest and both COVID-19 cases and excess mortality.¹⁸ In Figure 5 the blue line displays the time-varying coefficient of face-to-face

¹⁸For the sake of comparability, we restrict the sample to be the same for the two dependent variables in a given time window.

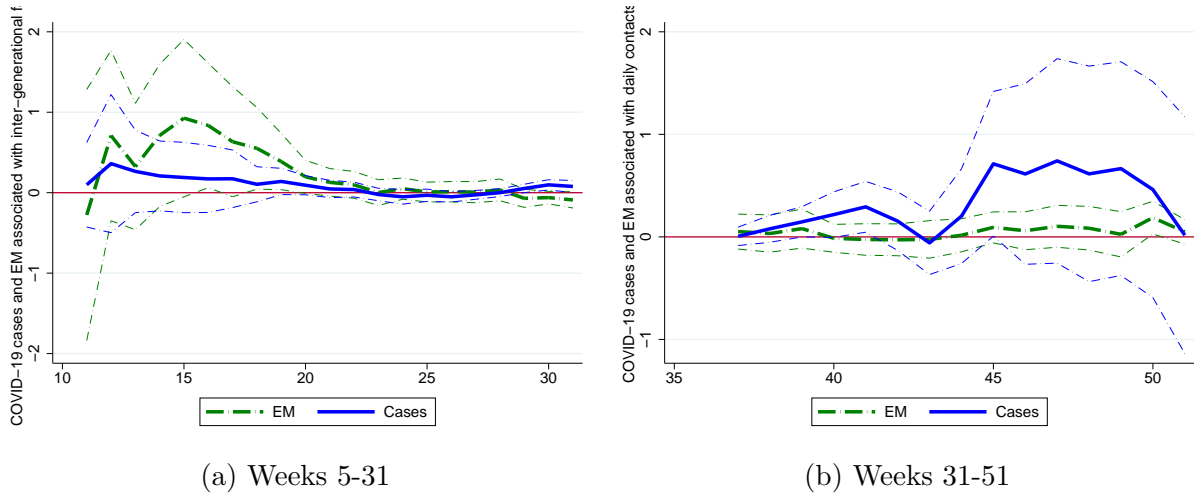
contacts on the average new daily cases in the week; the positive relationship is confirmed in both waves. In particular, in the second wave, the coefficient estimated on average daily cases per week is not only larger than the coefficient estimated on the corresponding daily data, but it is also more persistent in time. The green line, which displays the time-varying association between daily contacts and excess mortality, suggests that the estimated relationship is positive and statistically significant up to week 20 while in the second wave it drops close to zero, though it remains statistically significant between weeks 43 and 47 with a coefficient around 0.2.

Figure 5: Weekly data: Cases and excess mortality (full sample)



Notes: Figures show the marginal effect of face-to-face contacts on Covid-19 cases and excess mortality using weekly data. Coefficients associated with COVID-19 cases are obtained from a regression such as column [3] in Table 2. Regarding excess mortality, results come from a regression like column [3] in Table 3. We use a rolling window estimation with 7 weeks of sample size. Blue (green) dashed lines indicate 90% confidence intervals for cases (excess mortality).

Figure 6: Weekly data and inter-generational families: Cases and excess mortality (full sample)



Notes: Figures show the marginal effect of inter-generational families on Covid-19 cases and excess mortality using weekly data. Coefficients associated with COVID-19 cases are obtained from a regression such as column [3] in Table 2. Regarding excess mortality, results come from a regression like column [3] in Table 3. We use a rolling window estimation with 7 weeks of sample size. Blue (green) dashed lines indicate 90% confidence intervals for cases (excess mortality).

As far as the share of inter-generational households is concerned (Figure 6) no relevant association with COVID-19 spread or mortality is detected in either wave.

With regard to face-to-face contacts, the sharp contrasts between the two waves could be related to the different age-structure of the cases. For the US, Monod et al. (2021) find that as of October 29, only adults aged 20-49 had a reproduction number greater than 1 and disproportionally contributed to the virus transmission with 75 out of every 100 new infections. Also in Europe, as the pandemic evolved and restrictions were lifted, the rebounding in mobility involved mainly young adults and this could help explaining the different strength of the estimated associations between face-to-face contacts and the virus spread and mortality in the two waves.¹⁹

To check if this is a plausible explanation, we carry on by disaggregating the spread of the virus by groups of age.

¹⁹Specifically, given the nature of the disease and the scarce number of tests available during the first wave, it is likely that at the beginning the reported cases were mainly those of elderly people hospitalized or reported to the GP, as they had a greater probability of presenting symptoms and developing the disease. Figure A.2 in Appendix A shows the association of face-to-face contacts and excess mortality when controlling for the lags of COVID-19 new cases. A positive and significant association is found only at the beginning of the pandemic, suggesting that COVID-19 cases were then under-reported. Given the inclusion of NUTS 1 fixed effects and the robustness of our results to the inclusion of NUTS 1 day-specific fixed effects, we do not think that this can affect our results in a significant way.

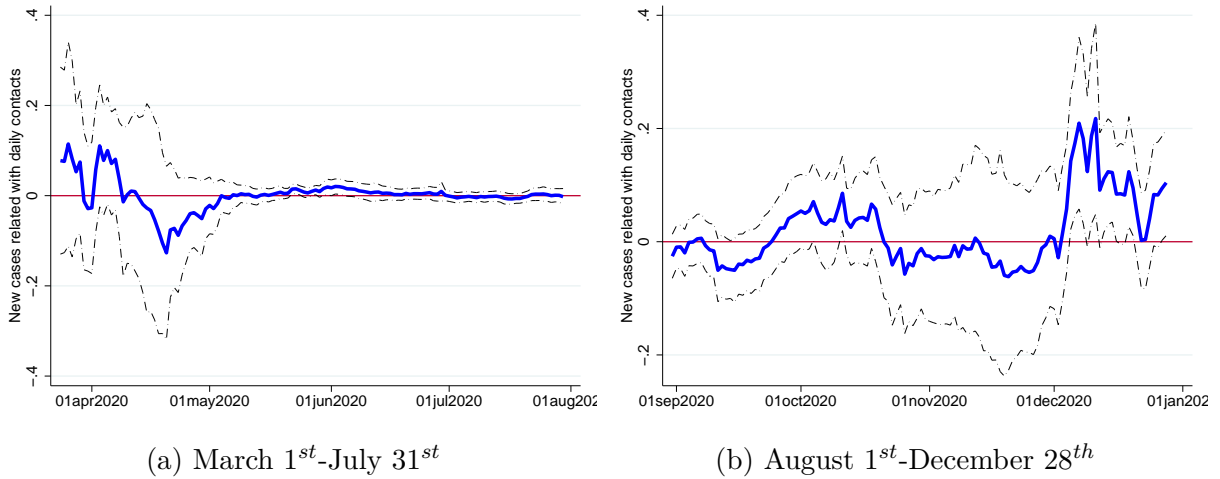
4.3 Covid cases by age

Daily COVID-19 regional data by age is difficult to obtain. In this Section, we include 4 countries that provide this information from official sources: Belgium, France, Germany, and Spain. Homogenizing the different datasets, we construct the number of daily COVID-19 new cases for four different age groups: 0-40, 40-60, 60-80, and 80+. Figures 7-10 show the results of running equation (1) using the number of new cases for each group of age as the dependent variable.²⁰

In the first wave, a positive relationship between face-to-face contacts and the spread of the virus is observed for people under 60 and people above 80. The relation is stronger for the elderly, where one percentage point increase in face-to-face contacts is associated with 0.4 new daily cases in April. For people under 60, the relationship is smaller (around 0.2) but observed also in March. When we consider the second wave, the new cases associated with face-to-face contacts are mostly of people under 60. The strongest association is with the group aged 40-60, which presents a coefficient of 0.5 statistically different from 0 between October and December. We also observe a similar result for the group of people under 40, but the effect wears off in December. On the contrary, during the second wave, face-to-face contacts are not found to be associated with the infections of the 80+ people, and the estimated coefficient is also small for the group of age 60-80. These results confirm the role of young adults in the transmission of the virus, as found by [Monod et al. \(2021\)](#) for the US. In addition, they confirm that older cohorts were hit relatively more in the first wave. This is compatible with changes in the behaviour of different cohorts due to the knowledge of the virus over the course of the pandemic. For example, in the first months of the epidemic, when the virus was still largely unknown and the way the pandemic was spreading was not fully understood, it was more difficult for people to avoid the contagion. However, this could have changed in the second wave, where more information about COVID-19 spread and mortality was available. In this new context, it is logical to think that frailer groups (i.e., older cohorts) may be more cautious regarding their exposure to the virus (i.e., inter-personal contacts). All in all, further research on the transmission channels across and within different age cohorts, and on the role of changes in the behaviour of people of different ages is necessary to confirm these initial results (see for example [Harris, 2021](#)), but it is beyond the scope of this paper.

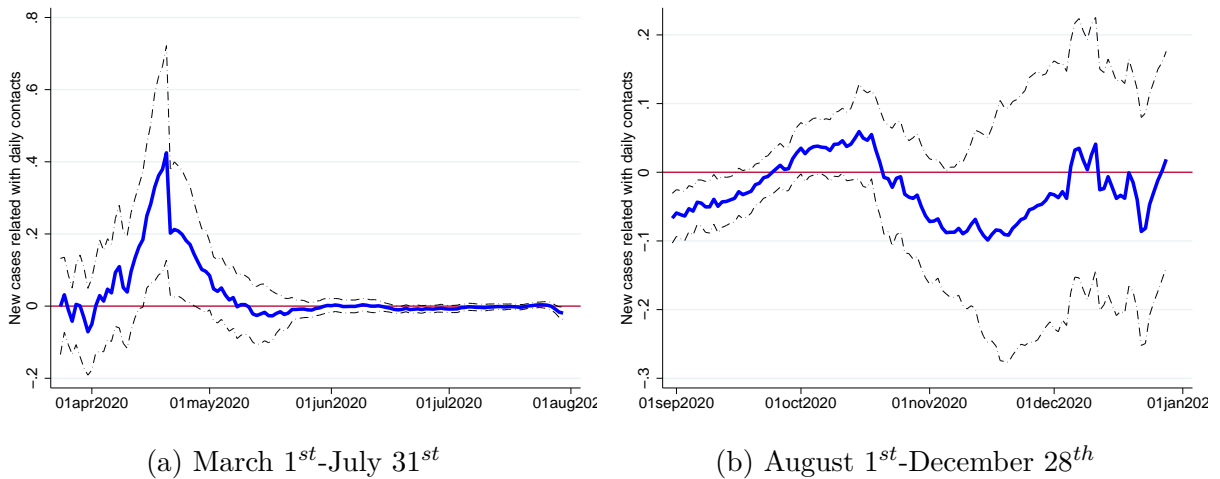
²⁰To control for the state of the pandemic, we include up to 10 lags of the the total number of COVID-19 daily new cases in the region and the inverse distance weighted sum of new cases in other regions using the baseline sample. If we use the total number of cases as a dependent variable, our results are very similar to the ones obtained in Figure 3, indicating that sample selection does not alter our results (see Figure A.3 in Appendix A).

Figure 9: Daily Covid cases (60-80): 30 days rolling window



Notes: Figures show the marginal effect of face-to-face contacts on daily COVID-19 cases 60-80. Dashed lines indicate 90% confidence intervals. Coefficients are obtained from a regression such as column [3] in Table 1 using a 30 days rolling window. The regression includes up to 10 lags of the total number of COVID-19 cases and the inverse distance weighted sum of new cases in other regions. Only regions with at least 5 observations are included.

Figure 10: Daily Covid cases (80+): 30 days rolling window



Notes: Figures show the marginal effect of face-to-face contacts on daily COVID-19 cases above 80. Dashed lines indicate 90% confidence intervals. Coefficients are obtained from a regression such as column [3] in Table 1 using a 30 days rolling window. The regression includes up to 10 lags of the total number of COVID-19 cases and the inverse distance weighted sum of new cases in other regions. Only regions with at least 5 observations are included.

4.4 Robustness checks

Along the paper, we have shown that our findings are robust to different specifications including different fixed effects and controls. In this section, we further check the robustness of the results to the sample of countries selected, to the number of persons included in the EQLS survey in a given region, and to the potential bias arising from the dynamic nature of our analysis.

4.4.1 Sample selection and number of interviews

Regarding the sample used in the analysis, our criteria was to include any country with available data in the variables of interest. This implies that our baseline sample consists of an heterogeneous group of countries, though all Europeans, including small countries that could be driving our results. To rule out this possibility, in Appendix B we run again equation (1) including only the biggest countries in our sample: Germany, France, Italy, Spain, and the UK. Results are qualitatively robust.

Another source of concern comes from the fact that in the EQLS, the number of interviewees is based on the regional population. The measures of face-to-face contacts and inter-generational households could then be calculated with insufficient precision in small regions.

To account for this issue, we estimate a weighted version of equation (1) using the number of survey respondents as the weights. The results, reported in Appendix C, reinforce our previous findings in that we observe a larger impact of face-to-face contacts on the spread and mortality of the virus. In addition, the share of inter-generational households is also significant and positively associated with daily and weekly cases.

4.4.2 Dynamic bias

A major threat to our identification may come from the dynamic structure of equation (1). That is, the inclusion of lags of the dependent variable in the right hand side of equation (1) may lead to a bias in all regression coefficients estimated by OLS (Nickell, 1981). In Appendix D we present two robustness exercises to see to what extent this dynamic bias could be affecting our results.

In the first exercise, we follow Adda (2016) and Qiu et al. (2020) and instrument the lags of the dependent variable with previous weather events. Specifically, using daily weather data at the NUTS 2 level from Felice and Kavvadias (2022), we follow Qiu et al. (2020) and use as instrumental variables weekly averages of daily heating degree days, runoff, wind speed at 10 meters, and the interaction between runoff and wind speed, during the preceding third and fourth weeks.²¹

Tables D.1 and D.2 show, respectively, the results for cases and excess mortality using weekly data. Regardless of the sample under consideration and the decision to use weights or not, our results show a positive relationship between face-to-face contacts and the number of new cases. The coefficients range from 0.214 to 0.491, very similar to those observed in the OLS regressions, which reinforces our confidence in the previous results.

²¹The main assumption here is that, given the incubation period of COVID-19, weather conditions in the preceding 2 weeks do not affect directly the probability that a person contracts the virus in period t , but they affect the number of other persons who have become infectious within the 2-week window.

When we focus on excess mortality, we can also see a positive relationship with face-to-face contacts, although in this case the range of values is wider than in the OLS regressions and is only significant when we use weighted regressions.²²

It is worth noting that our IV strategy is designed for the number of new infections while it may be less suitable for excess mortality. In Table D.3 we show the F-test of joint significance of the instruments and its associated p-values from the first-stage. For the number of cases, the values are not far from the ones obtained by Adda (2016), with the p-values below the 5 percent level when we include NUTS 1-week fixed effects. However, when we consider excess mortality, as expected, the results indicate that the instruments could be weak. For this reason, below we show an alternative exercise to ensure the validity of our results.

Figure D.1 completes the IV analysis by displaying the rolling window results for new cases and excess mortality. The picture is very similar to the one observed in the OLS results. That is, regions with a larger share of face-to-face contacts are associated with a larger number of new cases and excess mortality at the beginning of the pandemic. In the second wave, however, this positive association is only observed for the number of new cases. Results are robust and consistent for the different samples under consideration and the use of weights.²³

All in all, this exercise seems to indicate that our results are not importantly affected by the dynamic bias. This could be due to the relatively large time dimension of our panel as the dynamic bias decreases with T. In order to confirm the robustness of our results, in a second exercise we exclude the dynamics from equation (1).²⁴ Specifically, we use the cumulative number of cases and the cumulative excess mortality to estimate a cross-section version of equation (1) such as:²⁵

$$TotY_i = \alpha_{between,s} \sum_{i \neq j} d_{ij}^{-1} TotY_j + \beta_{11} X_i + \beta_{12} G_i + \beta_2 \mathbf{Z}_i + \mu_{nuts} + I^c + \epsilon_i, \quad (2)$$

²²The difference in observations between the IV and OLS estimations is due to data limitations regarding the weather variables. In this exercise we focus on weekly data for two reasons: i) it is more parsimonious (there are only 2 lags instead of 14) and ii) we are able to compare new cases and excess mortality. Results with daily data are available upon request.

²³Figure D.2 shows that our results do not change if we also considered endogenous the lagged inverse log distance weighted sum of new cases in other regions. In this case, beyond the previous instruments (daily heating degree days, runoff, wind speed at 10 meters, and the interaction between runoff and wind speed), we further include the inverse log distance weighted sum of each of these variables in other regions, during the preceding third and fourth weeks, as instrumental variables in the IV regressions.

²⁴A popular way to deal with the dynamic bias is to use the Difference or System GMM estimators. However, in our context these estimators are not appropriate for two reasons: i) we have a relatively small cross-section dimension with respect to the time dimension and ii) our main variable of interest and most of our controls are time-invariant.

²⁵The number of cumulative cases refers to the sum of the weekly averages of new daily COVID-19 cases.

where the dependent variable $TotY_i$ represents the cumulative number of COVID-19 cases in a given week in region i and $\sum_{i \neq j} d_{ij}^{-1} TotY_j$ is the inverse log distance weighted sum of the cumulative number of COVID-19 cases in the rest of regions.²⁶ The rest of control variables are explained in Section 3.

Figure D.3 shows the results using weekly data. The time variation in the figures is obtained by estimating equation (2) in different weeks.²⁷ The results are consistent with our previous findings, regions with a larger share of face-to-face contacts have a larger number of new cases during the first wave which dramatically increases during the second wave.²⁸ Regarding excess mortality, we observe an increase during the first wave that remains constant over the whole period.

All in all, we do not find evidence of the dynamic bias playing a relevant role in our results. Our findings are robust to different samples, the inclusion of weights, and different estimation strategies.

5 Conclusions

This paper moves from the socioeconomic research that relates COVID-19 spread to the inter-generational interactions occurring within the household, and extends it in several ways. We place the analysis at the regional NUTS 2 level, consider a total of 220 regions of 15 European countries and high frequency data on COVID-19 cases and excess mortality spanning from March to December 2020, thus covering the first and the second wave of pandemic. We show that both new infections and excess mortality are positively associated with the percentage of people having daily face-to-face contacts with friends or neighbours. On the contrary, the relevance of the share of inter-generational households in the transmission of the epidemic is not robust, except when using weighted LS, but even in this case the magnitude and significance of the coefficient is comparably lower than that of social interactions outside the household. Similarly, with regard to excess mortality, the approximated within-household inter-generational contacts appear less than half as lethal as outside-household social interactions.

Our results indicate that a standard deviation increase in the percentage of people having daily face-to-face contacts is associated with an increase of 5 daily COVID-19 cases, on a weekly average, and 2.6 fatalities on top of the ‘normal’ number of deaths.

²⁶When we estimate the relevance of face-to-face contacts on excess mortality, the dependent variable is the cumulative number of excess deaths and we do not include the inverse log distance weighted sum of new cases in other regions.

²⁷To be on the safe side, as this time we do not exploit within-region variation, we limit the analysis to weeks where there is no change in the number of regions in the sample.

²⁸Note that the different shape in the figure is related to the dependent variable, which now is the cumulative number of cases.

When we allow our coefficients to vary over time, we observe that this positive relationship is evident in both COVID-19 waves as far as the number of new cases is concerned, whereas for excess mortality the relationship is stronger in the first wave. Exploiting age-specific information from a subset of countries, we find evidence that the age groups have played different roles in the transmission of the epidemic: the cohort of adults aged 40-60 is particularly important in the second wave, as face-to-face contacts are strongly associated with the infections of this age group; on the contrary, in the early months of the epidemic, face-to-face contacts show a relatively stronger association with the contagion of younger adults and 80+ elderly.

Results are robust to the sample of countries used and to the estimation strategy. The statistical significance of face-to-face contacts is also robust to the inclusion of economic, demographic and epidemiological controls, to within and between region lagged cases as well as to fixed effects that should absorb unobserved heterogeneity. Nonetheless, as some residual unobserved regional heterogeneity cannot be excluded, we do not claim causality.

On the whole, our findings are consistent with the view that the frequency and type of social contacts incorporate cultural features, as posits by relational sociology, and that the patterns of such contacts show some degree of persistence. This implies that the macro map of daily face-to-face interactions that characterizes a region is informative on the potential transmission of contagion and on its lethality.

References

- Adda, J. (2016). Economic activity and the spread of viral diseases: Evidence from high frequency data. *The Quarterly Journal of Economics*, 131(2):891–941.
- Aparicio Fenoll, A. and Grossbard, S. (2020). Intergenerational residence patterns and Covid-19 fatalities in the EU and the US. *Economics & Human Biology*, 39(C).
- Aron, J. and Muellbauer, J. (2020). Measuring excess mortality: England is the European outlier in the Covid-19 pandemic. Technical report, VOX CEPR Policy Portal.
- Arpino, B., Bordone, V., and Pasqualini, M. (2020). No clear association emerges between intergenerational relationships and COVID-19 fatality rates from macro-level analyses. *Proceedings of the National Academy of Sciences*, 117(32):19116–19121.
- Ascani, A., Faggian, A., and Montresor, S. (2020). The geography of COVID-19 and the structure of local economies: The case of Italy. *Journal of Regional Science*, 61.
- Bartscher, A. K., Seitz, S., Siegloch, S., Slotwinski, M., and Wehrhöfer, N. (2021). Social capital and the spread of COVID-19: Insights from european countries. *Journal of Health Economics*, 80:102531.
- Bayer, C. and Kuhn, M. (2020). Intergenerational ties and case fatality rates: A cross-country analysis. IZA Discussion Papers 13114, IZA Institute of Labor Economics.
- Belloc, M., Buonanno, P., Drago, F., Galbiati, R., and Pinotti, P. (2020). Cross-country correlation analysis for research on COVID-19. Technical report, VOX CEPR Policy Portal.
- Borgonovi, F. and Andrieu, E. (2020). Bowling together by bowling alone: Social capital and COVID-19. *Social Science & Medicine*, 265:113501.
- Brown, C. S. and Ravallion, M. (2020). Inequality and the coronavirus: Socioeconomic covariates of behavioral responses and viral outcomes across US counties. NBER Working Papers 27549, National Bureau of Economic Research, Inc.
- Cornwell, B. (2011). Age trends in daily social contact patterns. *Research on Aging*, 33(5):598–631.
- Davies, N. G., Klepac, P., Liu, Y., Prem, K., Jit, M., Pearson, C. A. B., Quilty, B. J., Kucharski, A. J., Gibbs, H., Clifford, S., Gimma, A., van Zandvoort, K., Munday, J. D., Diamond, C., Edmunds, W. J., Houben, R. M. G. J., Hellewell, J., Russell, T. W., Abbott, S., Funk, S., Bosse, N. I., Sun, Y. F., Flasche, S., Rosello, A., Jarvis,

- C. I., Eggo, R. M., and working group, C. C.-. (2020). Age-dependent effects in the transmission and control of covid-19 epidemics. *Nature Medicine*, 26(8):1205–1211.
- Dépelteau, F. (2018). *The Palgrave handbook of relational sociology*. Springer.
- Donati, P. (2018). An original relational sociology grounded in critical realism. In *The Palgrave Handbook of Relational Sociology*, pages 431–456. Springer.
- Dowd, J. B., Andriano, L., Brazel, D. M., Rotondi, V., Block, P., Ding, X., Liu, Y., and Mills, M. C. (2020). Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proceedings of the National Academy of Sciences*, 117(18):9696–9698.
- Durante, R., Guiso, L., and Gulino, G. (2021). Asocial capital: Civic culture and social distancing during COVID-19. *Journal of Public Economics*, 194(C).
- Duranton, G., Rodriguez-Pose, A., and Sandall, R. (2008). Family types and the persistence of regional disparities in europe. *Economic Geography - ECON GEOGR*, 85:23–47.
- Esteve, A., Permanyer, I., Boertien, D., and Vaupel, J. W. (2020). National age and coresidence patterns shape COVID-19 vulnerability. *Proceedings of the National Academy of Sciences*, 117(28):16118–16120.
- Felice, M. D. and Kavvadias, K. (2022). ERA-NUTS: Meteorological time-series based on C3S ERA5 for European regions (1980-2021).
- Forbes, H., Morton, C. E., Bacon, S., McDonald, H. I., Minassian, C., Brown, J. P., Rentsch, C. T., Mathur, R., Schultze, A., DeVito, N. J., et al. (2021). Association between living with children and outcomes from covid-19: Opensafely cohort study of 12 million adults in england. *bmj*, 372.
- Fuhse, J. A. (2013). Social relationships between communication, network structure, and culture. In *Applying relational sociology*, pages 181–206. Springer.
- Fuhse, J. A. (2018). Deconstructing and reconstructing social networks. In *The Palgrave Handbook of Relational Sociology*, pages 457–479. Springer.
- Giuliano, P. and Rasul, I. (2020). Compliance with social distancing during the Covid-19 crisis. Technical report, VOX CEPR Policy Portal.
- Harris, J. E. (2020). Data from the COVID-19 epidemic in Florida suggest that younger cohorts have been transmitting their infections to less socially mobile older adults. *Review of Economics of the Household*, 18(4):1019–1037.

- Harris, J. E. (2021). Los Angeles county SARS-CoV-2 epidemic: Critical role of multi-generational intra-household transmission. *Journal of Bioeconomics*, 23(1):55–83.
- Kaasa, A., Vadi, M., and Varblane, U. (2014). Regional cultural differences within European countries: Evidence from multi-country surveys. *Management International Review*, 54(6):825–852.
- Lyttelton, T. and Zang, E. (2022). Sickness-related absences during the covid-19 pandemic: The role of occupations. *Journal of Health and Social Behavior*.
- Mäkinen, T. M., Juvonen, R., Jokelainen, J., Harju, T. H., Peitso, A., Bloigu, A., Silvennoinen-Kassinen, S., Leinonen, M., and Hassi, J. (2009). Cold temperature and low humidity are associated with increased occurrence of respiratory tract infections. *Respiratory Medicine*, 103(3):456 – 462.
- Mastrandrea, R., Fournet, J., and Barrat, A. (2015). Contact patterns in a high school: A comparison between data collected using wearable sensors, contact diaries and friendship surveys. *PLoS One*, 10(9).
- Mogi, R. and Spijker, J. (2021). The influence of social and economic ties to the spread of COVID-19 in Europe. *Journal of Population Research*.
- Monod, M., Blenkinsop, A., Xi, X., Hebert, D., Bershian, S., Tietze, S., Baguelin, M., Bradley, V. C., Chen, Y., Coupland, H., et al. (2021). Age groups that sustain resurging COVID-19 epidemics in the United States. *Science*, 371(6536).
- Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., and Mikolajczyk, R. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS One*, 5(4).
- Murgante, B., Borruso, G., Balletto, G., Castiglia, P., and Dettori, M. (2020). Why Italy first? Health, geographical and planning aspects of the Covid-19 outbreak. *Preprints*.
- Naqvi, A. (2021). COVID-19 European regional tracker. *medRxiv*.
- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6):1417–1426.
- Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. Simon & Schuster, New York.
- Qiu, Y., Chen, X., and Shi, W. (2020). Impacts of social and economic factors on the transmission of coronavirus disease 2019 (COVID-19) in China. *Journal of Population Economics*, 33(4):1127–1172.

- Read, J. M., Eames, K. T. D., and Edmunds, W. J. (2008). Dynamic social networks and the implications for the spread of infectious disease. *Journal of the Royal Society Interface*, 5:1001–1007.
- Rodríguez-Pose, A. and Burlina, C. (2021). Institutions and the uneven geography of the first wave of the COVID-19 pandemic. *Journal of Regional Science*.
- Salathé, M., Kazandjieva, M., Lee, J. W., Levis, P., Feldman, M. W., and Jones, J. H. (2010). A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences*, 107(51):22020–22025.
- van Oorschot, W., Arts, W., and Gelissen, J. (2006). Social capital in Europe. Measurement and regional distribution of a multi-faceted phenomenon. *Acta Sociologica*, 49(2):149–167.

APPENDIX A: Supplementary tables and figures

Table A.1: Sample

NUTS2	Period	NUTS2	Period	NUTS2	Period	NUTS2	Period	NUTS2	Period	NUTS2	Period
AT11	11mar2020	DE30	17mar2020	28dec2020	DE30	17mar2020	28dec2020	DE30	17mar2020	28dec2020	28dec2020
AT12	11mar2020	DE40	17mar2020	28dec2020	ES42	25feb2020	27dec2020	ITC1	10mar2020	28dec2020	28dec2020
AT13	11mar2020	DE50	17mar2020	28dec2020	ES43	09feb2020	27dec2020	ITC2	10mar2020	28dec2020	28dec2020
AT21	11mar2020	DE60	17mar2020	28dec2020	ES51	29jan2020	27dec2020	ITC3	10mar2020	28dec2020	28dec2020
AT22	11mar2020	DE71	17mar2020	28dec2020	ES52	02feb2020	27dec2020	ITC4	10mar2020	28dec2020	28dec2020
AT31	11mar2020	DE72	17mar2020	28dec2020	ES61	29jan2020	27dec2020	ITF1	10mar2020	28dec2020	28dec2020
AT32	11mar2020	DE73	17mar2020	28dec2020	ES62	23feb2020	27dec2020	ITF2	10mar2020	28dec2020	28dec2020
AT33	11mar2020	DE80	17mar2020	28dec2020	ES70	30jan2020	27dec2020	ITF3	10mar2020	28dec2020	28dec2020
AT34	11mar2020	DE91	17mar2020	28dec2020	FI19	04mar2020	28dec2020	ITF4	10mar2020	28dec2020	28dec2020
BE10	15mar2020	DE92	17mar2020	28dec2020	FI1B	20feb2020	28dec2020	ITF5	10mar2020	28dec2020	28dec2020
BE21	15mar2020	DE93	17mar2020	28dec2020	FI1C	05mar2020	28dec2020	ITF6	10mar2020	28dec2020	28dec2020
BE22	15mar2020	DE94	17mar2020	28dec2020	FI1D	20feb2020	28dec2020	ITG1	10mar2020	28dec2020	28dec2020
BE23	16mar2020	DEA1	17mar2020	28dec2020	FR10	24mar2020	25dec2020	ITG2	10mar2020	28dec2020	28dec2020
BE24	15mar2020	DEA2	17mar2020	28dec2020	FRB0	24mar2020	25dec2020	ITH2	10mar2020	28dec2020	28dec2020
BE25	16mar2020	DEA3	17mar2020	28dec2020	DEA3	17mar2020	28dec2020	ITH3	10mar2020	28dec2020	28dec2020
BE31	16mar2020	DEA4	17mar2020	28dec2020	FRC1	24mar2020	25dec2020	ITH4	10mar2020	28dec2020	28dec2020
BE32	17mar2020	DEA5	17mar2020	28dec2020	FRD1	24mar2020	25dec2020	ITH5	10mar2020	28dec2020	28dec2020
BE33	15mar2020	DEB1	17mar2020	28dec2020	FRD2	24mar2020	25dec2020	ITI1	10mar2020	28dec2020	28dec2020
BE34	20mar2020	DEB2	17mar2020	28dec2020	FRE1	24mar2020	25dec2020	ITI2	10mar2020	28dec2020	28dec2020
BE35	17mar2020	DEB3	17mar2020	28dec2020	FRE2	24mar2020	25dec2020	ITI3	10mar2020	28dec2020	28dec2020
CZ01	17mar2020	DEC0	17mar2020	28dec2020	FRF1	24mar2020	25dec2020	ITI4	10mar2020	28dec2020	28dec2020
CZ02	17mar2020	DED2	17mar2020	28dec2020	FRF2	24mar2020	25dec2020	NL11	30apr2020	28dec2020	28dec2020
CZ03	17mar2020	DED4	17mar2020	28dec2020	FRF3	24mar2020	25dec2020	NL12	30apr2020	28dec2020	28dec2020
CZ04	17mar2020	DED5	17mar2020	28dec2020	FRG0	24mar2020	25dec2020	NL13	30apr2020	28dec2020	28dec2020
CZ05	17mar2020	DEE0	17mar2020	28dec2020	FRH0	24mar2020	25dec2020	NL21	30apr2020	28dec2020	28dec2020
CZ06	17mar2020	DEF0	17mar2020	28dec2020	FRJ1	24mar2020	25dec2020	NL22	30apr2020	28dec2020	28dec2020
CZ07	17mar2020	DEG0	17mar2020	28dec2020	FRJ3	24mar2020	25dec2020	NL23	30apr2020	28dec2020	28dec2020
CZ08	17mar2020	DK01	11mar2020	28dec2020	FRJ1	24mar2020	25dec2020	NL31	30apr2020	28dec2020	28dec2020
DE11	17mar2020	DK02	11mar2020	28dec2020	FRJ2	24mar2020	25dec2020	NL32	30apr2020	28dec2020	28dec2020
DE12	17mar2020	DK03	11mar2020	28dec2020	FRK1	24mar2020	25dec2020	NL33	30apr2020	28dec2020	28dec2020
DE13	17mar2020	DK04	11mar2020	28dec2020	FRK2	24mar2020	25dec2020	NL34	30apr2020	28dec2020	28dec2020
DE14	17mar2020	DK05	11mar2020	28dec2020	HU10	24mar2020	25dec2020	NL41	30apr2020	28dec2020	28dec2020
DE21	17mar2020	ES11	02feb2020	27dec2020	HU11	15apr2020	28dec2020	NL42	30apr2020	28dec2020	28dec2020
DE22	17mar2020	ES12	24feb2020	27dec2020	HU12	15apr2020	28dec2020	PL21	02apr2020	28dec2020	28dec2020
DE23	17mar2020	ES13	24feb2020	27dec2020	HU21	15apr2020	28dec2020	PL22	02apr2020	28dec2020	28dec2020
DE24	17mar2020	ES21	13feb2020	27dec2020	HU22	15apr2020	28dec2020	PL41	02apr2020	28dec2020	28dec2020
DE25	17mar2020	ES22	11feb2020	27dec2020	HU23	11feb2020	27dec2020	PL42	02apr2020	28dec2020	28dec2020
DE26	17mar2020	ES24	07feb2020	27dec2020	HU31	15apr2020	28dec2020	PL43	02apr2020	28dec2020	28dec2020
DE27	17mar2020	ES30	29jan2020	27dec2020	HU32	15apr2020	28dec2020	PL51	02apr2020	28dec2020	28dec2020
UKF2	14mar2020	PL52	02apr2020	28dec2020	HU33	03feb2020	27dec2020	PL52	02apr2020	28dec2020	28dec2020
UKF3	17mar2020	PL61	02apr2020	28dec2020	ITC1	25feb2020	27dec2020	PL61	02apr2020	28dec2020	28dec2020
UKG1	16mar2020	PL62	02apr2020	28dec2020	ITC2	09feb2020	27dec2020	PL62	02apr2020	28dec2020	28dec2020
UKG2	16mar2020	PL63	02apr2020	28dec2020	ITC3	29jan2020	27dec2020	PL63	02apr2020	28dec2020	28dec2020
UKG3	10mar2020	PL71	02apr2020	28dec2020	ITC4	02feb2020	27dec2020	PL71	02apr2020	28dec2020	28dec2020
UKH1	17mar2020	PL72	02apr2020	28dec2020	ITF1	10mar2020	28dec2020	PL72	02apr2020	28dec2020	28dec2020
UKH2	12mar2020	PL81	02apr2020	28dec2020	ITF2	23feb2020	27dec2020	PL81	02apr2020	28dec2020	28dec2020
UKH3	13mar2020	PL82	02apr2020	28dec2020	ITF3	30jan2020	27dec2020	PL82	02apr2020	28dec2020	28dec2020
UKI3	12mar2020	PL84	02apr2020	28dec2020	ITF4	04mar2020	28dec2020	PL84	02apr2020	28dec2020	28dec2020
UKI4	10mar2020	PL91	02apr2020	28dec2020	ITF5	10mar2020	28dec2020	PL91	02apr2020	28dec2020	28dec2020
UKI5	15mar2020	PL92	02apr2020	28dec2020	ITF6	05mar2020	28dec2020	PL92	02apr2020	28dec2020	28dec2020
UKI6	12mar2020	RO11	17apr2020	28dec2020	ITG1	20feb2020	28dec2020	RO11	17apr2020	28dec2020	28dec2020
UKI7	10mar2020	RO12	17apr2020	28dec2020	ITG2	10mar2020	28dec2020	RO12	17apr2020	28dec2020	28dec2020
UKJ1	03mar2020	RO21	17apr2020	28dec2020	ITH2	24mar2020	25dec2020	RO21	17apr2020	28dec2020	28dec2020
UKJ2	03mar2020	RO22	17apr2020	28dec2020	ITH3	24mar2020	25dec2020	RO22	17apr2020	28dec2020	28dec2020
UKJ3	12mar2020	RO31	17apr2020	28dec2020	ITH4	24mar2020	25dec2020	RO31	17apr2020	28dec2020	28dec2020
UKJ4	13mar2020	RO32	17apr2020	28dec2020	ITH5	24mar2020	25dec2020	RO32	17apr2020	28dec2020	28dec2020
UKK1	12mar2020	RO41	17apr2020	28dec2020	ITI1	10mar2020	28dec2020	RO41	17apr2020	28dec2020	28dec2020
UKK2	19mar2020	RO42	17apr2020	28dec2020	ITI2	10mar2020	28dec2020	RO42	17apr2020	28dec2020	28dec2020
UKK3	16mar2020	SE11	28feb2020	28dec2020	ITI3	24mar2020	25dec2020	SE11	28feb2020	28dec2020	28dec2020
UKK4	15mar2020	SE12	29feb2020	28dec2020	ITI4	24mar2020	25dec2020	SE12	29feb2020	28dec2020	28dec2020
UKM6	13mar2020	SE21	27feb2020	28dec2020	NL12	30apr2020	28dec2020	SE21	03mar2020	28dec2020	28dec2020
UKM7	13mar2020	SE22	27feb2020	28dec2020	NL13	30apr2020	28dec2020	SE22	03mar2020	28dec2020	28dec2020
UKM8	13mar2020	SE23	27feb2020	28dec2020	NL21	30apr2020	28dec2020	SE23	07mar2020	28dec2020	28dec2020
UKM9	13mar2020	SE31	07mar2020	28dec2020	NL22	30apr2020	28dec2020	SE31	07mar2020	28dec2020	28dec2020
		SE32	11mar2020	28dec2020	NL23	30apr2020	28dec2020	SE32	11mar2020	28dec2020	28dec2020
		SE33	10mar2020	28dec2020	NL31	30apr2020	28dec2020	SE33	10mar2020	28dec2020	28dec2020
		UKC1	20mar2020	28dec2020	NL32	30apr2020	28dec2020	UKC1	20mar2020	28dec2020	28dec2020
		UKC2	16mar2020	28dec2020	NL33	30apr2020	28dec2020	UKC2	16mar2020	28dec2020	28dec2020
		UKD1	17mar2020	28dec2020	NL34	30apr2020	28dec2020	UKD1	17mar2020	28dec2020	28dec2020
		UKD3	10mar2020	28dec2020	NL41	30apr2020	28dec2020	UKD3	10mar2020	28dec2020	28dec2020
		UKD4	15mar2020	28dec2020	NL42	30apr2020	28dec2020	UKD4	15mar2020	28dec2020	28dec2020
		UKD6	16mar2020	28dec2020	NL43	30apr2020	28dec2020	UKD6	16mar2020	28dec2020	28dec2020
		UKD7	15mar2020	28dec2020	PL21	02apr2020	28dec2020	UKD7	15mar2020	28dec2020	28dec2020
		UKE1	15mar2020	28dec2020	PL22	02apr2020	28dec2020	UKE1	15mar2020	28dec2020	28dec2020
		UKE2	03mar2020	28dec2020	PL41	02apr2020	28dec2020	UKE2	03mar2020	28dec2020	28dec2020
		UKE3	15mar2020	28dec2020	PL42	02apr2020	28dec2020	UKE3	15mar2020	28dec2020	28dec2020
		UKE4	12mar2020	28dec2020	PL43	02apr2020	28dec2020	UKE4	12mar2020	28dec2020	28dec2020
		UKF1	10mar2020	28dec2020	PL51	02apr2020	28dec2020	UKF1	10mar2020	28dec2020	28dec2020

Notes: Sample of our preferred specification (Column 3 in Table 1).

Table A.2: Data description

Variable	Description	Source
Daily cases	New daily COVID-19 cases	(1)
Weekly cases	Weekly average of new daily COVID-19 cases	(1)
Face-to-face contacts	% of people who has direct face-to-face contact every day or almost every day with friends or neighbours	(2)
Inter-generational	% of households in which several generations coexist	(2)
Poverty rate	% of households able to make ends meet with difficulty or great difficulty	(2)
Excess mortality	Difference between the observed numbers of deaths in a specific week in 2020 and the average number of deaths in the same week of the period 2015-2019	(3)
$\ln(\text{Population})$	(log of) total population on 1 January 2019	(3)
$\ln(\text{GDP}_{pc})$	(log of) 2018 GDP per capita	(3)
$\ln(\text{Density})$	(log of) persons per km^2	(3)
$\ln(\text{Temperature})$	(log of) heating degree days	(3)
Share 65	Proportion of population aged 65 years and more (%)	(3)
$\frac{\text{Women}}{\text{Men}}$	Women per 100 men	(3)
$\frac{\text{Emp}_{services}}{\text{Emp}_{total}}$	Service sector employment as % of total employment	(3)
Employment rate	Employment rate (%)	(3)
High education	Active population with tertiary education as % of total population	(3)
$\ln(\text{Respiratory})$	(log of) crude death rate for diseases of the respiratory system for people aged 65 years and more	(3)
Air connection	Air passengers carried per population	(3)
$\ln(\text{Days/Weeks})$	(log of) days/weeks since first COVID-19 cases	(1)
Initial cases	COVID-19 cases per population the first day available for those regions where it is not possible to calculate the number of days since first COVID-19 case	(1)

Notes: (1) COVID-19 European Regional Tracker; (2) European Quality of Life Surveys; (3) Eurostat. From Eurostat we use data from the last year available.

Table A.3: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Time-varying					
Daily cases	62,698	263.23	604.75	0	14238
Weekly cases	8,989	262.88	568.34	0	9202.57
Excess mortality	8,916	74.15	205.92	-402.25	5772
$\ln(Days)$	62,698	4.81	0.780	0	5.85
$\ln(Weeks)$	8,989	2.85	0.787	0	3.89
Time-invariant					
Face-to-face contacts	220	41.81	16.15	0	100
Inter-generational	220	38.97	12.95	0	70
Poverty rate	220	13.02	10.72	0	58.01
$\ln(Population)$	220	14.30	0.692	11.74	16.32
$\ln(GDP_{pc})$	220	10.25	0.512	8.79	12.27
$\ln(Density)$	220	5.25	1.26	1.22	9.35
$\ln(Temperature)$	220	7.84	0.325	5.20	8.75
Share 65	220	20.37	3.02	8.20	28.90
$\frac{Women}{Men}$	220	103.66	2.80	96.30	114.10
$\frac{Emp_{services}}{Emp_{total}}$	220	70.87	10.13	33.50	91.30
Employment rate	220	70.27	7.53	40.70	80.30
High education	220	16.37	5.48	6.27	41.10
$\ln(Respiratory)$	220	5.91	0.342	5.06	6.68
Air connection	220	2.37	4.76	0	38.08
Initial cases	220	0.043	0.140	0	0.878

Figure A.1: COVID-19 cases and Excess mortality (weekly)

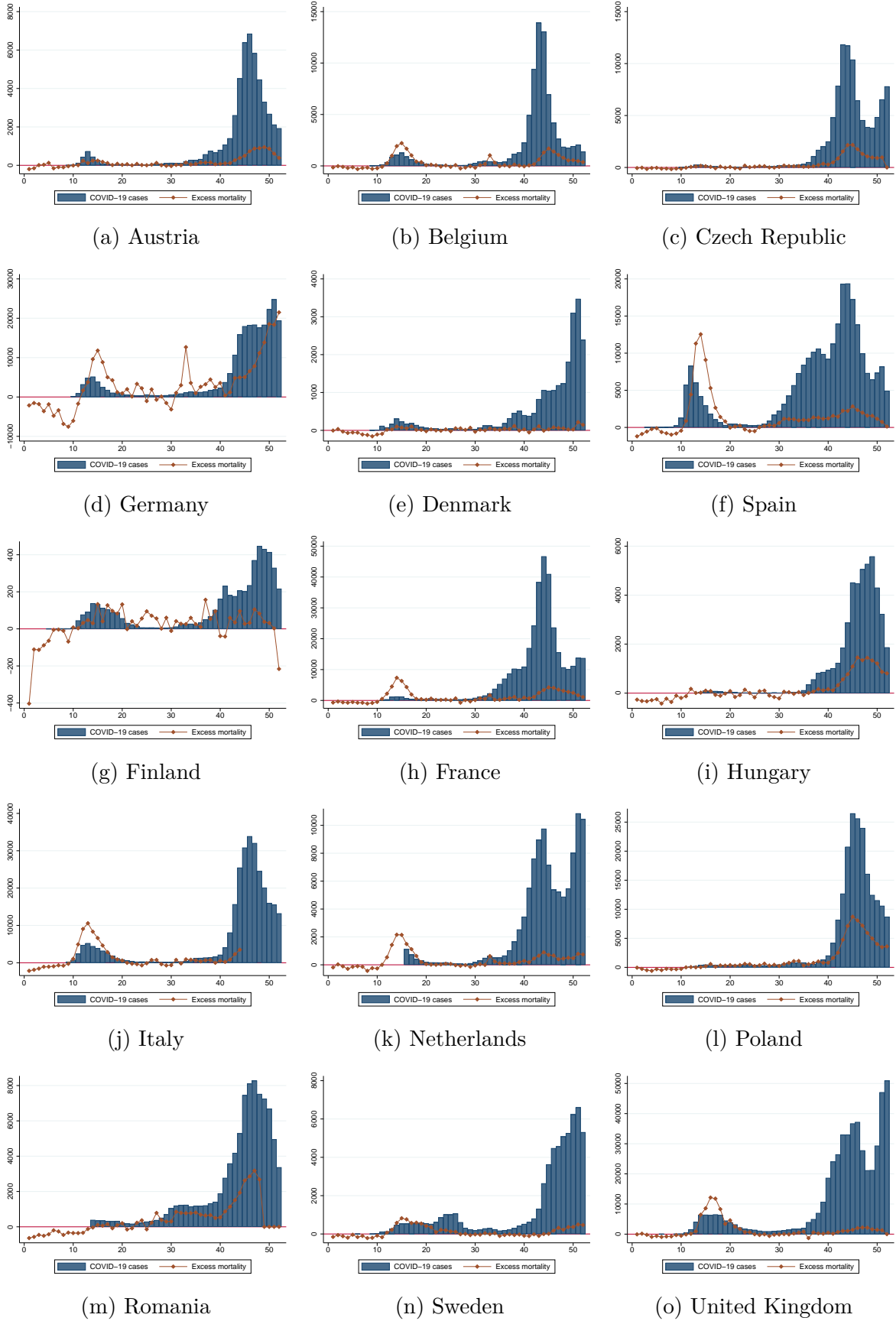
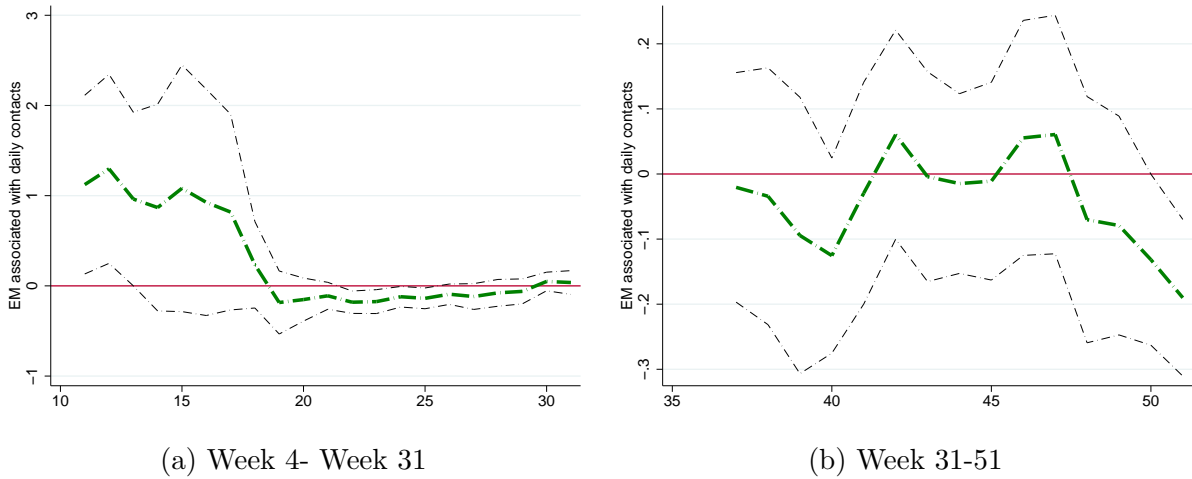
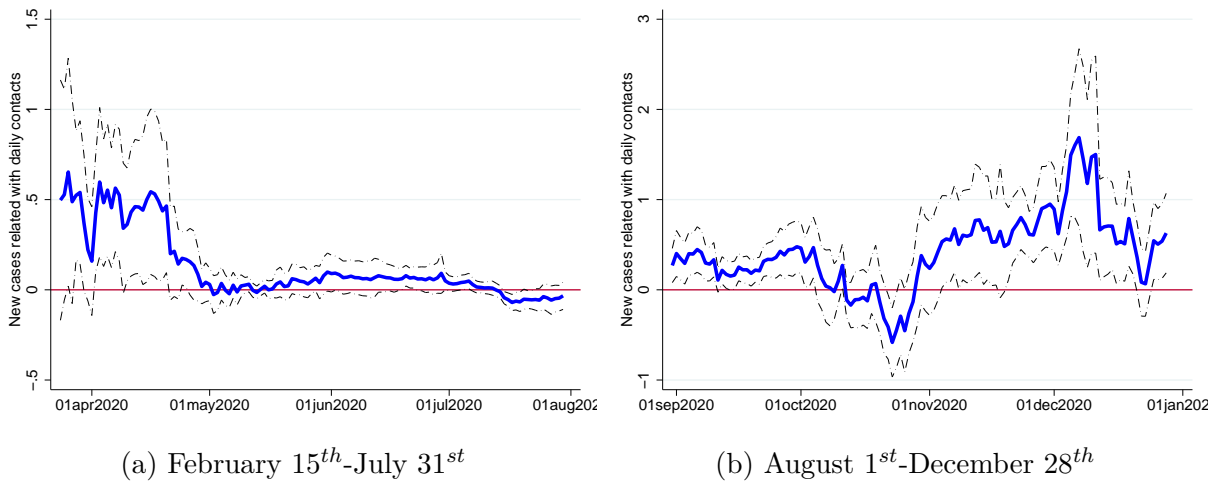


Figure A.2: Weekly data: Excess mortality controlling for cases



Notes: Figures show the marginal effect of face-to-face contacts on excess mortality. Dashed lines indicate 90% confidence intervals. Coefficients are obtained from a regression such as column [3] in Table 3 using a 7 weeks rolling window. The regression further includes up to 2 lags of the COVID-19 number of cases.

Figure A.3: Daily cases: 30 days rolling window (age sample)



Notes: Figures show the marginal effect of physical contacts on daily COVID-19 cases. Dashed lines indicate 90% confidence intervals. Coefficients are obtained from a regression such as column [2] in Table 1 using a 30 days rolling window. The regression includes up to 10 lags of the total number of COVID-19 cases and the inverse distance weighted sum of new cases in other regions. Only regions with at least 5 observations are included.

APPENDIX B: Selected sample

Table B.1: Daily cases: Selected sample

	[1]	[2]	[3]	[4]	[5]	[6]
Face-to-face contacts	0.083 (0.038)**	0.080 (0.038)**	0.087 (0.039)**	0.064 (0.042)	0.067 (0.044)	0.072 (0.045)
$\ln(Days)$	-8.744 (3.202)***	-7.859 (3.336)**	-8.257 (3.310)**	-8.099 (2.901)***	-6.487 (2.959)**	-6.917 (2.906)**
$\ln(Population)$	11.678 (1.622)***	12.079 (1.631)***	11.606 (1.642)***	10.138 (1.813)***	10.440 (1.795)***	10.169 (1.801)***
Initial cases	102.980 (17.751)***	98.019 (19.448)***	92.739 (19.456)***	89.499 (17.508)***	83.244 (18.338)***	80.196 (18.270)***
$\ln(GDP_{pc})$	-5.628 (1.601)***	0.172 (2.765)	-0.495 (2.674)	-4.878 (2.004)**	0.188 (2.919)	-0.201 (2.866)
$\ln(Density)$	3.810 (1.277)***	1.821 (1.352)	2.058 (1.290)	3.891 (1.379)***	2.291 (1.497)	2.366 (1.417)*
$\ln(Temperature)$	28.172 (6.242)***	24.576 (5.079)***	24.236 (5.120)***	28.032 (6.942)***	24.145 (5.874)***	23.874 (5.890)***
Inter-generational		0.024 (0.047)	0.003 (0.045)		0.028 (0.047)	0.014 (0.045)
Share 65		-1.195 (0.450)***	-1.357 (0.483)***		-1.120 (0.468)**	-1.218 (0.507)**
$\frac{Women}{Men}$		0.055 (0.549)	0.201 (0.532)		0.077 (0.568)	0.173 (0.554)
$\frac{Emp_{services}}{Emp_{total}}$		0.170 (0.158)	0.119 (0.158)		0.183 (0.171)	0.145 (0.170)
Employment rate		0.161 (0.252)	0.184 (0.247)		0.267 (0.283)	0.271 (0.275)
High education		-0.843 (0.318)***	-0.895 (0.321)***		-0.795 (0.328)**	-0.818 (0.333)**
Poverty rate		0.019 (0.056)	0.020 (0.057)		-0.002 (0.062)	-0.003 (0.064)
$\ln(Respiratory)$			-6.519 (6.884)			-3.530 (7.885)
Air connection			0.117 (0.063)*			0.088 (0.066)
Constant	-348.103 (65.369)***	-372.256 (95.259)***	-325.121 (100.370)***	-245.567 (75.942)***	-273.604 (106.450)**	-248.701 (111.728)**
NUTS 1 FE	Yes	Yes	Yes	No	No	No
Country FE*Date FE	Yes	Yes	Yes	No	No	No
NUTS 1 FE*Date FE	No	No	No	Yes	Yes	Yes
Observations	37935	37935	37935	33207	33207	33207
# NUTS2	130	130	130	114	114	114
Rsq	0.959	0.959	0.959	0.962	0.962	0.962
RMSE	141.766	141.769	141.772	126.070	126.074	126.078

Notes: The dependent variable is the number of daily COVID-19 cases. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include up to 14 lags of the dependent variable and the inverse distance weighted sum of new cases in other regions. The selected sample includes: Germany, France, Italy, Spain, and UK.

Table B.2: Weekly cases: Selected sample

	[1]	[2]	[3]	[4]	[5]	[6]
face-to-face contacts	0.302 (0.111)***	0.313 (0.103)***	0.341 (0.105)***	0.248 (0.121)**	0.269 (0.113)**	0.295 (0.116)**
ln(<i>Weeks</i>)	-43.852 (13.546)***	-42.257 (13.422)***	-42.722 (13.406)***	-36.432 (13.589)***	-34.061 (13.455)**	-34.357 (13.382)**
ln(<i>Population</i>)	33.610 (4.802)***	33.826 (4.486)***	32.455 (4.518)***	29.278 (5.221)***	29.397 (4.811)***	28.516 (4.848)***
Initial cases	40.247 (29.240)	37.815 (27.422)	35.393 (26.876)	35.775 (29.005)	32.827 (25.844)	32.064 (25.732)
ln(<i>GDP_{pc}</i>)	-21.887 (5.749)***	-11.541 (7.649)	-12.329 (7.317)*	-18.709 (6.529)***	-9.275 (8.359)	-9.296 (8.101)
ln(<i>Density</i>)	11.727 (4.182)***	7.182 (4.084)*	7.075 (3.971)*	11.963 (4.459)***	8.388 (4.489)*	7.870 (4.349)*
ln(<i>Temperature</i>)	86.143 (19.245)***	68.925 (15.896)***	67.727 (15.786)***	83.865 (20.414)***	67.446 (17.164)***	66.515 (16.863)***
Inter-generational		0.184 (0.145)	0.126 (0.135)		0.195 (0.149)	0.155 (0.138)
Share 65		-3.599 (1.281)***	-4.001 (1.333)***		-3.403 (1.358)**	-3.655 (1.415)**
$\frac{Women}{Men}$		-0.895 (1.505)	-0.391 (1.413)		-0.749 (1.595)	-0.366 (1.491)
$\frac{Emp_{services}}{Emp_{total}}$		0.358 (0.474)	0.150 (0.477)		0.386 (0.501)	0.200 (0.499)
Employment rate		0.999 (0.675)	0.962 (0.653)		1.207 (0.757)	1.122 (0.730)
High education		-1.980 (0.885)**	-2.027 (0.913)**		-1.906 (0.932)**	-1.887 (0.955)*
Poverty rate		0.056 (0.159)	0.035 (0.159)		0.008 (0.176)	-0.018 (0.174)
ln(<i>Respiratory</i>)			-9.318 (18.751)			-1.419 (20.721)
Air connection			0.524 (0.180)***			0.478 (0.192)**
Constant	-978.352 (195.382)***	-839.902 (258.203)***	-771.136 (272.135)***	-750.746 (228.420)***	-626.217 (288.828)**	-616.306 (300.458)**
NUTS 1 FE	Yes	Yes	Yes	No	No	No
Country FE*Week FE	Yes	Yes	Yes	No	No	No
NUTS 1 FE*Week FE	No	No	No	Yes	Yes	Yes
Observations	5434	5434	5434	4758	4758	4758
# NUTS2	130	130	130	114	114	114
Rsq	0.963	0.963	0.963	0.970	0.970	0.970
RMSE	126.495	126.442	126.452	107.688	107.578	107.590

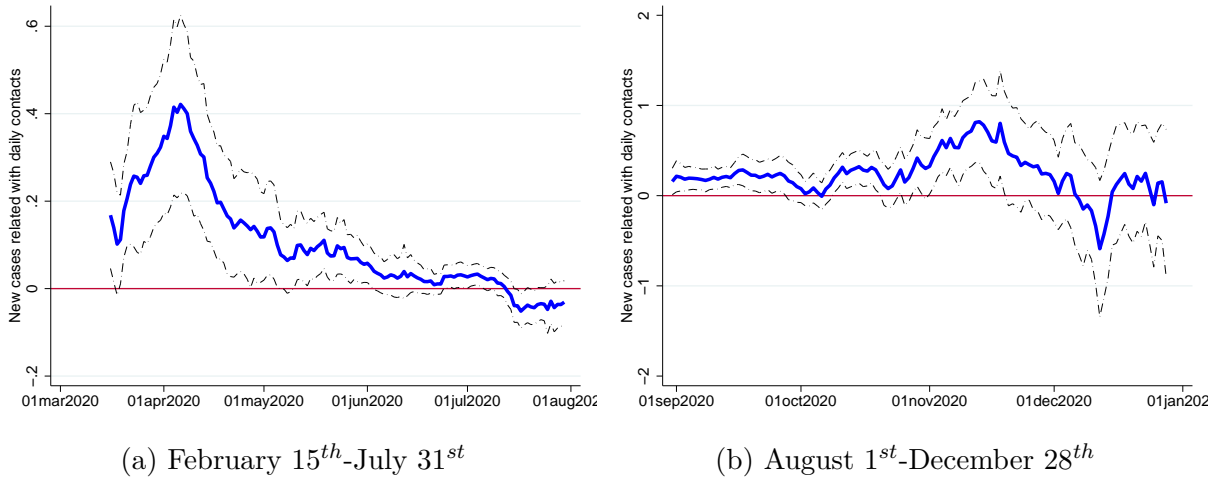
Notes: The dependent variable is the weekly average of the number of daily COVID-19 cases. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include up to 2 lags of the dependent variable and the inverse distance weighted sum of new cases in other regions. The selected sample includes: Germany, France, Italy, Spain, and UK.

Table B.3: Weekly excess mortality: Selected sample

	[1]	[2]	[3]	[4]	[5]	[6]
Face-to-face contacts	0.133 (0.074)*	0.185 (0.076)**	0.188 (0.077)**	0.133 (0.068)*	0.173 (0.073)**	0.177 (0.075)**
$\ln(Weeks)$	51.375 (26.841)*	53.876 (26.593)**	53.900 (26.573)**	21.239 (26.256)	23.606 (26.276)	23.585 (26.231)
$\ln(Population)$	11.928 (3.085)***	11.207 (2.718)***	11.251 (2.635)***	11.332 (2.636)***	10.910 (2.426)***	10.833 (2.361)***
Initial cases	72.428 (51.406)	67.492 (43.807)	68.067 (44.151)	58.854 (52.541)	55.770 (45.940)	55.901 (46.145)
$\ln(GDP_{pc})$	-6.949 (3.012)**	2.324 (6.162)	2.661 (6.254)	-6.324 (3.217)*	1.929 (5.799)	2.046 (5.857)
$\ln(Density)$	1.194 (1.264)	3.288 (2.319)	3.065 (2.546)	0.778 (1.175)	2.105 (2.232)	1.984 (2.419)
$\ln(Temperature)$	40.644 (18.385)**	29.932 (14.955)**	29.961 (15.037)**	37.780 (18.780)**	29.260 (15.063)*	29.211 (15.106)*
Inter-generational		0.106 (0.079)	0.107 (0.080)		0.074 (0.068)	0.070 (0.071)
Share 65		-0.851 (0.739)	-0.829 (0.780)		-0.853 (0.709)	-0.870 (0.760)
$\frac{Women}{Men}$		-0.000 (0.929)	0.010 (0.942)		0.082 (0.895)	0.122 (0.908)
$\frac{Emp_{services}}{Emp_{total}}$		-0.589 (0.440)	-0.608 (0.453)		-0.460 (0.404)	-0.484 (0.418)
Employment rate		0.937 (0.600)	0.911 (0.621)		0.735 (0.569)	0.719 (0.589)
High education		-1.209 (0.829)	-1.181 (0.849)		-1.068 (0.807)	-1.057 (0.826)
Poverty rate		-0.205 (0.106)*	-0.211 (0.103)**		-0.162 (0.093)*	-0.167 (0.092)*
$\ln(Respiratory)$			3.005 (12.446)			0.891 (11.140)
Air connection			0.052 (0.142)			0.064 (0.138)
Constant	-552.046 (157.431)***	-560.692 (194.675)***	-581.068 (222.867)**	-444.349 (141.767)***	-464.147 (169.822)***	-469.742 (191.453)**
NUTS 1 FE	Yes	Yes	Yes	No	No	No
Country FE*Week FE	Yes	Yes	Yes	No	No	No
NUTS 1 FE*Week FE	No	No	No	Yes	Yes	Yes
Observations	5334	5334	5334	4640	4640	4640
# NUTS2	130	130	130	114	114	114
Rsq	0.859	0.859	0.859	0.929	0.929	0.929
RMSE	94.510	94.506	94.524	62.537	62.502	62.521

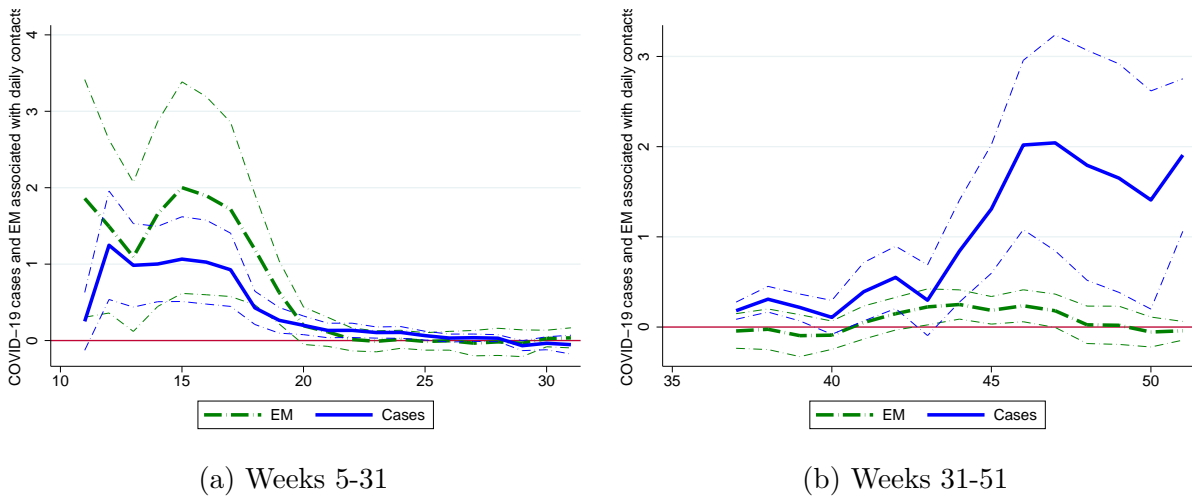
Notes: The dependent variable is the weekly excess mortality. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include up to 2 lags of the dependent variable. The selected sample includes: Germany, France, Italy, Spain, and UK.

Figure B.1: Daily cases: 30 days rolling window (selected sample)



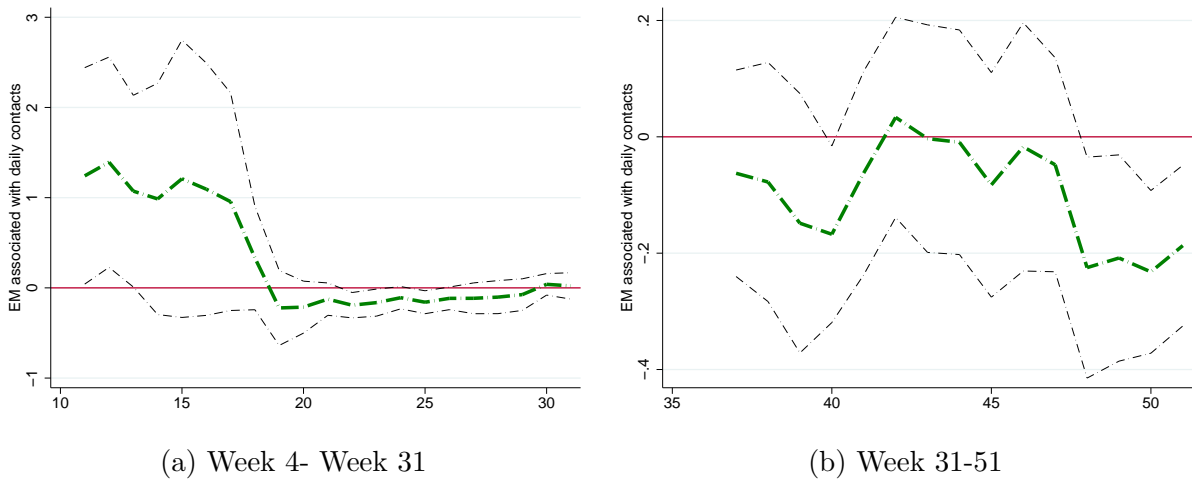
Notes: Figures show the marginal effect of face-to-face contacts on daily COVID-19 cases. Dashed lines indicate 90% confidence intervals. Coefficients are obtained from a regression such as column [3] in Table B.1 using a 30 days rolling window. The regression includes up to 10 lags of the dependent variable and the inverse distance weighted sum of new cases in other regions. Only regions with at least 5 observations are included.

Figure B.2: Weekly data: Cases and excess mortality (selected sample)



Notes: Figures show the marginal effect of face-to-face contacts on Covid-19 cases and excess mortality using weekly data. Coefficients associated with COVID-19 cases are obtained from a regression such as column [3] in Table B.2. Regarding excess mortality, results come from a regression like column [3] in Table B.3. We use a rolling window estimation with 7 weeks of sample size. Blue (green) dashed lines indicate 90% confidence intervals for cases (excess mortality).

Figure B.3: Weekly data: Excess mortality controlling for cases (selected sample)



Notes: Figures show the marginal effect of face-to-face contacts on excess mortality. Dashed lines indicate 90% confidence intervals. Coefficients are obtained from a regression such as column [3] in Table B.3 using a 7 weeks rolling window. The regression further includes up to 2 lags of the COVID-19 number of cases.

APPENDIX C: Weighted regressions

Table C.1: Daily cases (weighted)

	[1]	[2]	[3]	[4]	[5]	[6]
Face-to-face contacts	0.142 (0.063)**	0.168 (0.064)***	0.171 (0.061)***	0.104 (0.070)	0.134 (0.072)*	0.135 (0.068)**
$\ln(Days)$	-7.831 (3.499)**	-8.003 (3.459)**	-8.560 (3.480)**	-5.028 (2.784)*	-5.428 (2.446)**	-6.083 (2.545)**
$\ln(Population)$	16.086 (1.775)***	15.146 (1.696)***	14.838 (1.698)***	14.311 (1.690)***	13.693 (1.668)***	13.477 (1.701)***
Initial cases	8.141 (11.316)	10.365 (9.084)	10.811 (8.809)	7.916 (12.014)	7.865 (9.477)	8.129 (9.205)
$\ln(GDP_{pc})$	-4.018 (2.719)	0.905 (4.117)	0.937 (4.136)	-4.018 (2.857)	0.324 (4.062)	0.283 (4.230)
$\ln(Density)$	2.814 (0.833)***	1.929 (1.018)*	2.090 (0.981)**	3.057 (0.924)***	2.618 (1.136)**	2.766 (1.097)**
$\ln(Temperature)$	30.619 (8.419)***	32.197 (8.062)***	33.123 (8.120)***	29.886 (9.283)***	31.128 (9.104)***	32.328 (9.189)***
Inter-generational		0.126 (0.053)**	0.096 (0.049)*		0.112 (0.055)**	0.085 (0.050)*
Share 65		-1.047 (0.281)***	-1.182 (0.296)***		-1.006 (0.302)***	-1.149 (0.315)***
$\frac{Women}{Men}$		0.166 (0.397)	0.273 (0.396)		0.151 (0.411)	0.265 (0.402)
$\frac{Emp_{services}}{Emp_{total}}$		0.157 (0.156)	0.148 (0.154)		0.145 (0.160)	0.145 (0.157)
Employment rate		-0.027 (0.194)	-0.043 (0.190)		0.170 (0.221)	0.166 (0.215)
High education		-0.712 (0.363)*	-0.873 (0.373)**		-0.802 (0.393)**	-0.958 (0.406)**
Poverty rate		-0.165 (0.072)**	-0.172 (0.075)**		-0.167 (0.079)**	-0.173 (0.082)**
$\ln(Respiratory)$			-10.323 (5.631)*			-10.774 (6.186)*
Air connection			0.082 (0.071)			0.055 (0.076)
Constant	-541.052 (113.390)***	-583.855 (117.445)***	-527.883 (113.867)***	-354.123 (118.685)***	-396.110 (119.680)***	-338.878 (122.792)***
NUTS 1 FE	Yes	Yes	Yes	No	No	No
Country FE*Date FE	Yes	Yes	Yes	No	No	No
NUTS 1 FE*Date FE	No	No	No	Yes	Yes	Yes
Observations	62698	62698	62698	57676	57676	57676
# NUTS2	220	220	220	203	203	203
Rsq	0.951	0.951	0.951	0.962	0.962	0.962
RMSE	166.998	166.999	167.000	133.703	133.699	133.700

Notes: The dependent variable is the number of daily COVID-19 cases. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include up to 14 lags of the dependent variable and the inverse distance weighted sum of new cases in other regions. Weighted least squares using the number of EQLS surveys in each region as a weight.

Table C.2: Weekly cases (weighted)

	[1]	[2]	[3]	[4]	[5]	[6]
Face-to-face contacts	0.441 (0.156)***	0.519 (0.152)***	0.535 (0.148)***	0.344 (0.166)**	0.423 (0.162)***	0.437 (0.156)***
ln (<i>Weeks</i>)	-48.814 (13.936)***	-50.279 (13.457)***	-51.841 (13.535)***	-33.069 (12.641)***	-34.770 (10.866)***	-36.765 (11.164)***
ln (<i>Population</i>)	41.580 (4.774)***	39.198 (4.366)***	38.364 (4.304)***	36.112 (4.612)***	34.350 (4.197)***	33.696 (4.209)***
Initial cases	0.630 (2.295)	1.568 (1.851)	2.167 (1.898)	0.587 (2.317)	1.223 (1.920)	1.746 (1.987)
ln (<i>GDP_{pc}</i>)	-14.834 (6.924)**	-3.722 (10.341)	-3.464 (10.372)	-13.823 (7.098)*	-3.690 (10.733)	-3.516 (10.932)
ln (<i>Density</i>)	8.456 (2.136)***	6.866 (2.501)***	7.258 (2.443)***	8.222 (2.329)***	7.458 (2.840)***	7.820 (2.770)***
ln (<i>Temperature</i>)	88.530 (21.674)***	87.812 (20.387)***	89.177 (20.384)***	81.058 (23.617)***	80.182 (22.391)***	81.830 (22.412)***
Inter-generational		0.297 (0.130)**	0.228 (0.120)*		0.267 (0.132)**	0.205 (0.123)*
Share 65		-2.921 (0.716)***	-3.222 (0.756)***		-2.654 (0.763)***	-2.948 (0.803)***
$\frac{Women}{Men}$		-0.027 (1.014)	0.182 (1.010)		-0.051 (1.041)	0.156 (1.026)
$\frac{Emp_{services}}{Emp_{total}}$		0.472 (0.385)	0.437 (0.376)		0.397 (0.392)	0.374 (0.384)
Employment rate		0.467 (0.431)	0.413 (0.423)		0.682 (0.544)	0.647 (0.537)
High education		-2.093 (0.965)**	-2.471 (1.013)**		-2.076 (1.073)*	-2.427 (1.129)**
Poverty rate		-0.392 (0.165)**	-0.409 (0.170)**		-0.388 (0.175)**	-0.402 (0.181)**
ln (<i>Respiratory</i>)			-21.717 (12.812)*			-20.834 (13.513)
Air connection			0.230 (0.200)			0.186 (0.211)
Constant	-1274.050 (270.874)***	-1316.540 (282.498)***	-1190.749 (274.978)***	-931.185 (294.528)***	-961.565 (300.740)***	-841.614 (306.021)***
NUTS 1 FE	Yes	Yes	Yes	No	No	No
Country FE*Week FE	Yes	Yes	Yes	No	No	No
NUTS 1 FE*Week FE	No	No	No	Yes	Yes	Yes
Observations	8989	8989	8989	8269	8269	8269
# NUTS2	220	220	220	203	203	203
Rsq	0.964	0.964	0.964	0.974	0.974	0.974
RMSE	134.503	134.476	134.482	104.721	104.651	104.652

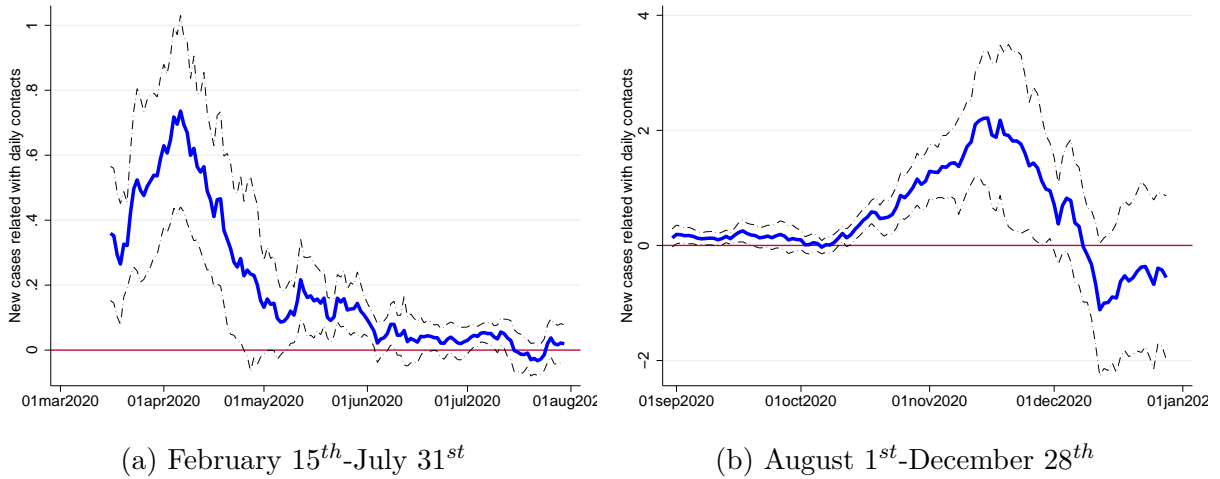
Notes: The dependent variable is the weekly average of the number of daily COVID-19 cases. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include up to 2 lags of the dependent variable and the inverse distance weighted sum of new cases in other regions. Weighted least squares using the number of EQLS surveys in each region as a weight.

Table C.3: Weekly excess mortality (weighted)

	[1]	[2]	[3]	[4]	[5]	[6]
Face-to-face contacts	0.343 (0.144)**	0.391 (0.125)***	0.390 (0.128)***	0.308 (0.144)**	0.355 (0.135)***	0.357 (0.138)**
$\ln(Weeks)$	39.860 (20.532)*	44.492 (21.444)**	44.006 (21.475)**	5.680 (14.468)	9.399 (14.937)	8.591 (15.174)
$\ln(Population)$	14.047 (3.238)***	13.043 (2.378)***	12.913 (2.291)***	13.341 (3.446)***	12.511 (2.587)***	12.321 (2.420)***
Initial cases	2.092 (1.763)	1.822 (1.728)	2.005 (1.784)	1.514 (1.845)	1.377 (1.872)	1.603 (1.943)
$\ln(GDP_{pc})$	-4.700 (2.971)	5.930 (7.621)	5.722 (7.853)	-4.651 (2.904)	5.358 (7.514)	5.233 (7.755)
$\ln(Density)$	1.168 (1.352)	4.585 (1.709)***	4.653 (1.759)***	1.535 (1.379)	3.987 (1.837)**	4.105 (1.915)**
$\ln(Temperature)$	38.351 (23.983)	34.093 (18.200)*	35.806 (18.138)**	37.507 (24.200)	33.135 (18.649)*	34.720 (18.638)*
Inter-generational		0.153 (0.099)	0.129 (0.094)		0.109 (0.100)	0.082 (0.094)
Share 65		-0.397 (0.615)	-0.574 (0.667)		-0.718 (0.646)	-0.897 (0.713)
$\frac{Women}{Men}$		-0.852 (0.831)	-0.716 (0.791)		-0.518 (0.813)	-0.387 (0.767)
$\frac{Emp_{services}}{Emp_{total}}$		-0.269 (0.285)	-0.250 (0.279)		-0.196 (0.287)	-0.186 (0.280)
Employment rate		0.908 (0.397)**	0.906 (0.387)**		0.845 (0.433)*	0.838 (0.422)**
High education		-1.362 (0.885)	-1.515 (0.970)		-1.354 (0.936)	-1.527 (1.044)
Poverty rate		-0.348 (0.142)**	-0.354 (0.147)**		-0.278 (0.134)**	-0.285 (0.140)**
$\ln(Respiratory)$			-13.411 (9.745)			-13.267 (10.397)
Air connection			0.007 (0.142)			0.037 (0.147)
Constant	-572.256 (211.035)***	-592.680 (189.254)***	-531.353 (180.985)***	-463.622 (203.342)**	-500.072 (186.998)***	-436.326 (173.246)**
NUTS 1 FE	Yes	Yes	Yes	No	No	No
Country FE*Week FE	Yes	Yes	Yes	No	No	No
NUTS 1 FE*Week FE	No	No	No	Yes	Yes	Yes
Observations	8916	8916	8916	8176	8176	8176
# NUTS2	220	220	220	203	203	203
Rsq	0.899	0.899	0.899	0.945	0.945	0.945
RMSE	88.899	88.853	88.859	60.127	60.041	60.042

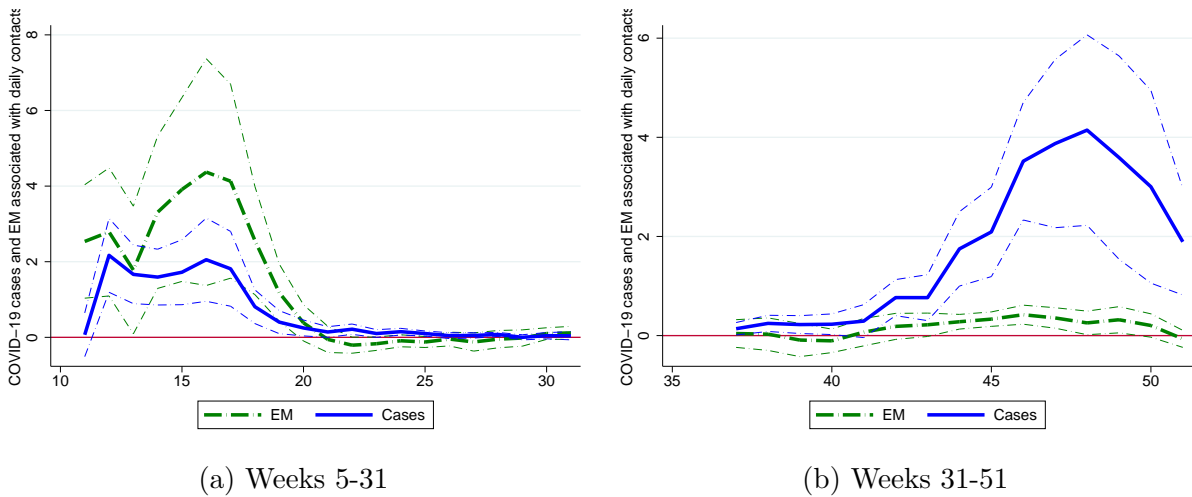
Notes: The dependent variable is the weekly excess mortality. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include up to 2 lags of the dependent variable. Weighted least squares using the number of EQLS surveys in each region as a weight.

Figure C.1: Daily cases: 30 days rolling window (weighted)



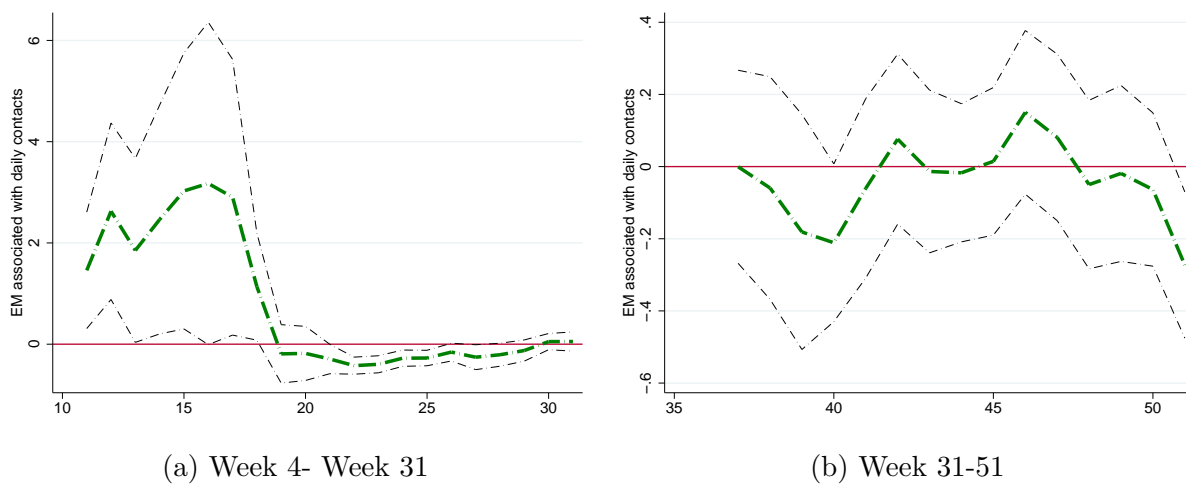
Notes: Figures show the marginal effect of face-to-face contacts on daily COVID-19 cases. Dashed lines indicate 90% confidence intervals. Coefficients are obtained from a regression such as column [3] in Table C.1 using a 30 days rolling window. The regression includes up to 10 lags of the dependent variable and the inverse distance weighted sum of new cases in other regions. Only regions with at least 5 observations are included.

Figure C.2: Weekly data: Cases and excess mortality (weighted)



Notes: Figures show the marginal effect of face-to-face contacts on Covid-19 cases and excess mortality using weekly data. Coefficients associated with COVID-19 cases are obtained from a regression such as column [3] in Table C.2. Regarding excess mortality, results come from a regression like column [3] in Table C.3. We use a rolling window estimation with 7 weeks of sample size. Blue (green) dashed lines indicate 90% confidence intervals for cases (excess mortality).

Figure C.3: Weekly data: Excess mortality controlling for cases (weighted)



Notes: Figures show the marginal effect of face-to-face contacts on excess mortality. Dashed lines indicate 90% confidence intervals. Coefficients are obtained from a regression such as column [3] in Table C.3 using a 7 weeks rolling window. The regression further includes up to 2 lags of the COVID-19 number of cases.

APPENDIX D: IV and cross-section regressions

Table D.1: Weekly cases: IV

	[1]	[2]	[3]	[4]	[5]	[6]
	Full sample		Selected sample		Weighted	
face-to-face contacts	0.214 (0.124)*	0.181 (0.109)*	0.233 (0.100)**	0.266 (0.125)**	0.491 (0.149)***	0.373 (0.156)**
ln (<i>Weeks</i>)	-17.915 (19.565)	-12.134 (14.393)	-23.330 (18.154)	-27.011 (21.893)	-44.953 (20.528)**	-27.493 (12.690)**
ln (<i>Population</i>)	25.115 (9.364)***	16.846 (8.041)**	23.352 (7.671)***	23.406 (10.469)**	31.876 (7.119)***	28.360 (7.154)***
Initial cases	1.905 (1.280)	1.382 (1.030)	24.080 (20.403)	25.157 (31.085)	2.047 (1.467)	1.592 (1.498)
ln (<i>GDP_{pc}</i>)	-11.099 (8.500)	-11.412 (7.479)	-12.020 (9.773)	-20.766 (12.665)	-0.252 (8.459)	-4.484 (9.385)
ln (<i>Density</i>)	3.797 (3.020)	0.115 (4.691)	4.655 (3.586)	2.751 (7.826)	5.615 (2.585)**	4.831 (3.825)
ln (<i>Temperature</i>)	39.843 (20.212)**	22.522 (25.196)	48.172 (17.858)***	42.908 (44.008)	74.382 (23.738)***	65.698 (27.140)**
Share 65	-2.851 (1.033)***	-1.591 (0.982)	-3.243 (1.364)**	-2.610 (1.601)	-2.695 (0.830)***	-2.417 (0.870)***
$\frac{Women}{Men}$	0.706 (0.874)	0.326 (0.837)	0.143 (1.583)	-0.566 (2.026)	0.190 (0.968)	0.282 (1.003)
Inter-generational	0.096 (0.099)	0.044 (0.123)	0.121 (0.114)	0.064 (0.223)	0.170 (0.109)	0.168 (0.115)
$\frac{Emp_{services}}{Emp_{total}}$	0.088 (0.278)	0.048 (0.283)	-0.007 (0.395)	0.039 (0.563)	0.343 (0.320)	0.290 (0.338)
Employment rate	0.499 (0.438)	-0.075 (0.634)	0.723 (0.598)	0.243 (1.247)	0.380 (0.447)	0.188 (0.686)
High education	-1.356 (0.881)	-0.381 (1.130)	-1.373 (0.860)	-0.887 (1.395)	-2.127 (1.055)**	-1.613 (1.287)
Poverty rate	-0.161 (0.123)	-0.080 (0.117)	0.044 (0.132)	0.089 (0.184)	-0.333 (0.145)**	-0.329 (0.170)*
ln (<i>Respiratory</i>)	-1.533 (8.071)	0.108 (8.969)	-7.573 (13.911)	-18.478 (26.429)	-15.806 (10.095)	-13.396 (12.962)
Air connection	0.336 (0.263)	0.215 (0.220)	0.544 (0.598)	0.281 (0.509)	0.055 (0.227)	0.114 (0.233)
NUTS 1 FE	Yes	No	Yes	No	Yes	No
Country FE*Week FE	Yes	No	Yes	No	Yes	No
NUTS 1 FE*Week FE	No	Yes	No	Yes	No	Yes
Observations	8733	8143	5220	4632	8733	8143
# NUTS2	214	200	125	111	214	200
Rsq	0.900	0.878	0.930	0.903	0.885	0.875
RMSE	121.161	107.520	124.193	120.040	142.809	105.437

Notes: The dependent variable is the weekly average of the number of daily COVID-19 cases. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include up to 2 lags of the dependent variable and the inverse distance weighted sum of new cases in other regions. The 2 lags of the dependent variable are considered endogenous. We use as instrumental variables weekly averages of daily heating degree days, runoff, wind speed at 10 meters, and the interaction between runoff and wind speed, during the preceding third and fourth weeks. Weather data at the NUTS 2 level is obtained from [Felice and Kavvadias \(2022\)](#). Panel A in Table D.3 displays the F-test and its associated p-value to test the joint significance of the instruments in the first-stage. “Full sample” refers to our baseline regression where we include any country with available data in the variables of interest. The “Selected sample” includes: Germany, France, Italy, Spain, and UK. “Weighted” presents the results from a weighted regression using the full sample and the number of EQLS surveys in each region as a weight.

Table D.2: Weekly excess mortality: IV

	[1]	[2]	[3]	[4]	[5]	[6]
	Full sample		Selected sample		Weighted	
face-to-face contacts	0.051 (0.142)	0.127 (0.180)	0.026 (0.153)	0.194 (0.198)	0.230 (0.137)*	0.660 (0.296)**
$\ln(Weeks)$	44.306 (19.158)**	22.800 (24.892)	53.592 (24.567)**	25.373 (32.158)	43.138 (21.329)**	8.411 (29.213)
$\ln(Population)$	4.924 (7.112)	9.676 (9.899)	1.976 (7.120)	11.476 (10.337)	7.730 (3.027)**	21.383 (5.889)**
Initial cases	-0.006 (0.714)	0.089 (1.004)	20.885 (43.210)	58.430 (67.530)	1.114 (1.277)	3.051 (3.527)
$\ln(GDP_{pc})$	1.484 (2.472)	0.995 (3.301)	4.417 (4.135)	1.756 (7.623)	4.114 (5.566)	8.962 (14.995)
$\ln(Density)$	0.416 (1.968)	1.217 (2.543)	0.659 (2.806)	2.494 (3.075)	2.590 (1.623)	7.336 (4.051)*
$\ln(Temperature)$	6.415 (16.361)	16.833 (29.202)	4.172 (22.031)	30.427 (37.597)	19.561 (14.694)	60.403 (39.855)
Share 65	-0.328 (0.535)	-0.710 (0.891)	-0.330 (0.459)	-0.822 (1.075)	-0.284 (0.517)	-1.591 (1.380)
$\frac{Women}{Men}$	0.190 (0.299)	0.320 (0.542)	0.016 (0.683)	-0.216 (1.399)	-0.609 (0.613)	-0.926 (1.481)
Inter-generational	0.046 (0.062)	0.049 (0.065)	0.019 (0.072)	0.061 (0.092)	0.079 (0.063)	0.120 (0.162)
$\frac{Emp_{services}}{Emp_{total}}$	-0.251 (0.258)	-0.331 (0.337)	-0.059 (0.392)	-0.436 (0.642)	-0.189 (0.184)	-0.348 (0.465)
Employment rate	0.209 (0.554)	0.479 (0.717)	0.235 (0.645)	0.737 (0.873)	0.451 (0.339)	1.468 (0.831)*
High education	-0.267 (0.461)	-0.527 (0.728)	-0.672 (0.598)	-1.013 (1.121)	-0.795 (0.733)	-2.439 (2.070)
Poverty rate	-0.111 (0.116)	-0.143 (0.137)	-0.122 (0.103)	-0.185 (0.143)	-0.255 (0.125)**	-0.507 (0.271)*
$\ln(Respiratory)$	-1.758 (8.095)	-6.528 (11.199)	2.916 (5.777)	2.103 (12.043)	-6.591 (7.521)	-23.226 (19.524)
Air connection	-0.108 (0.070)	-0.117 (0.130)	-0.129 (0.146)	-0.100 (0.370)	-0.051 (0.113)	-0.007 (0.325)
NUTS 1 FE	Yes	No	Yes	No	Yes	No
Country FE*Week FE	Yes	No	Yes	No	Yes	No
NUTS 1 FE*Week FE	No	Yes	No	Yes	No	Yes
Observations	8656	8050	5117	4514	8656	8050
# NUTS2	214	200	125	111	214	200
Rsq	0.741	0.751	0.752	0.802	0.817	0.753
RMSE	80.593	54.711	100.759	63.544	90.945	66.889

Notes: The dependent variable is the weekly excess mortality. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include up to 2 lags of the dependent variable. The 2 lags of the dependent variable are considered endogenous. We use as instrumental variables weekly averages of daily heating degree days, runoff, wind speed at 10 meters, and the interaction between runoff and wind speed, during the preceding third and fourth weeks. Weather data at the NUTS 2 level is obtained from [Felice and Kavvadias \(2022\)](#). Panel B in Table D.3 displays the F-test and its associated p-value to test the joint significance of the instruments in the first-stage. “Full sample” refers to our baseline regression where we include any country with available data in the variables of interest. The “Selected sample” includes: Germany, France, Italy, Spain, and UK. “Weighted” presents the results from a weighted regression using the full sample and the number of EQLS surveys in each region as a weight.

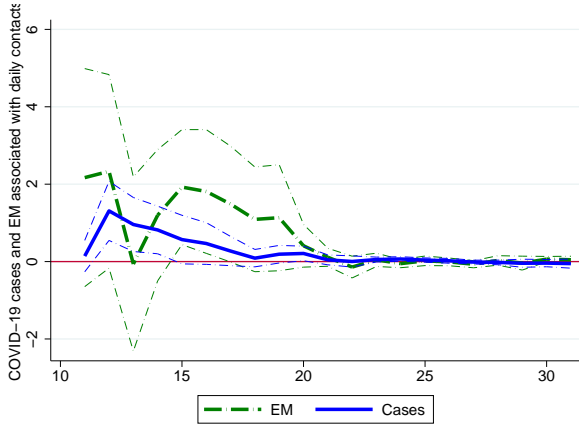
Table D.3: First-stage results

		[1]	[2]	[3]	[4]	[5]	[6]
		Full sample		Selected sample		Weighted	
Panel A: Cases _t							
Cases _{t-1}	F-stat	1.45	2.02	1.28	2.42	1.21	2.00
	p-val	(0.178)	(0.046)	(0.259)	(0.019)	(0.294)	(0.048)
Cases _{t-2}	F-stat	1.32	2.05	1.49	2.82	1.66	1.44
	p-val	(0.234)	(0.043)	(0.169)	(0.007)	(0.110)	(0.183)
Panel B: Excess _t							
Excess _{t-1}	F-stat	2.28	0.78	2.32	1.03	1.55	1.30
	p-val	(0.023)	(0.622)	(0.024)	(0.421)	(0.143)	(0.244)
Excess _{t-2}	F-stat	0.56	0.78	0.49	1.04	0.46	1.13
	p-val	(0.807)	(0.625)	(0.863)	(0.411)	(0.885)	(0.344)

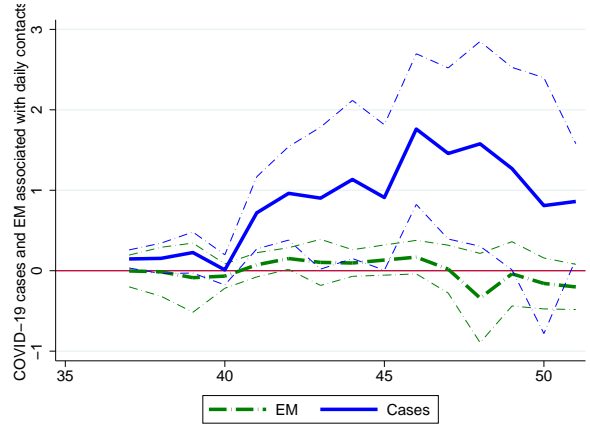
Notes: The table reports the F-test of joint significance of the instruments and its associated p-values in parenthesis. Panels A and B refer, respectively, to Tables D.1 and D.2. The numbering of the columns in this table coincides with columns in Tables D.1 and D.2. Each entry in the table corresponds to a separate regression. We use as instrumental variables weekly averages of daily heating degree days, runoff, wind speed at 10 meters, and the interaction between runoff and wind speed, during the preceding third and fourth weeks. Weather data at the NUTS 2 level is obtained from [Felice and Kavvadias \(2022\)](#).

Figure D.1: Weekly data: Cases and excess mortality (IV)

Full sample

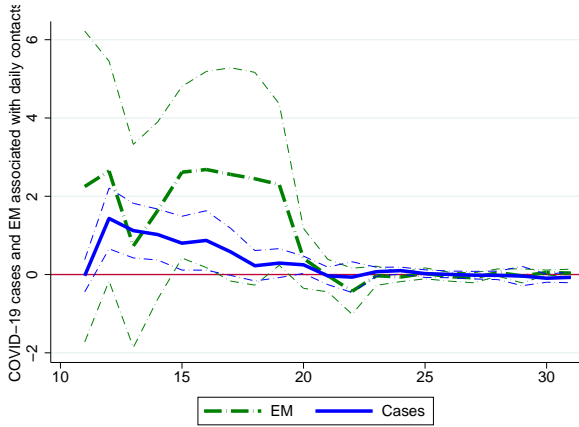


(a) Weeks 5-31

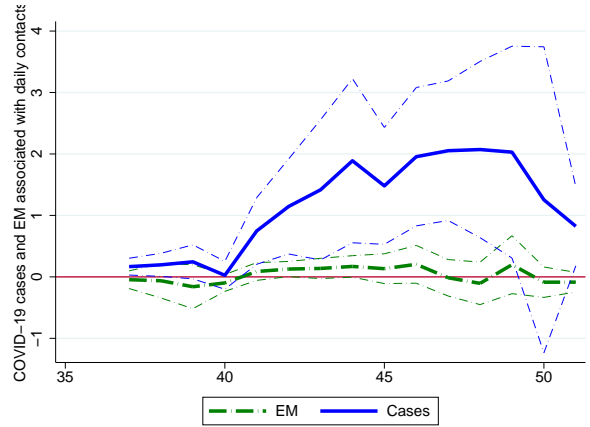


(b) Weeks 31-51

Selected sample

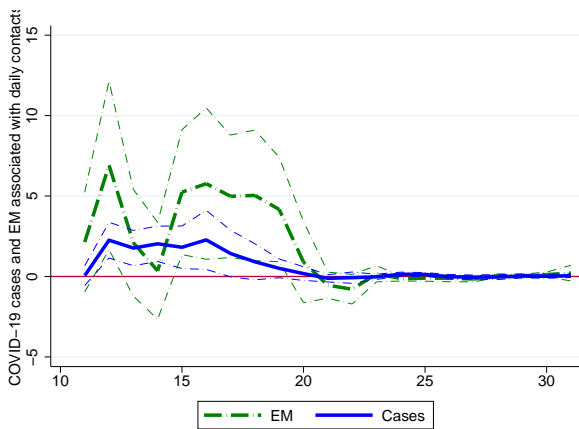


(c) Weeks 5-31

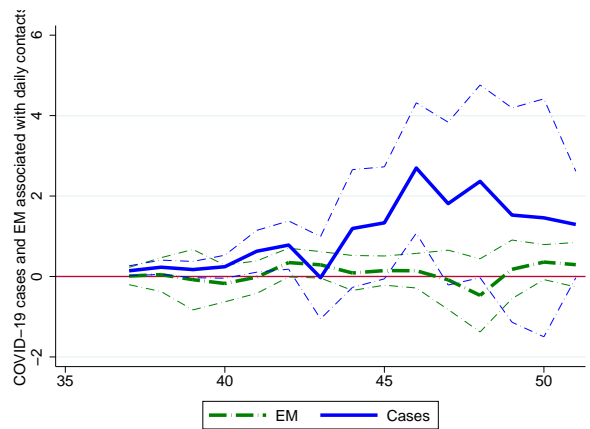


(d) Weeks 31-51

Weighted



(e) Weeks 5-31

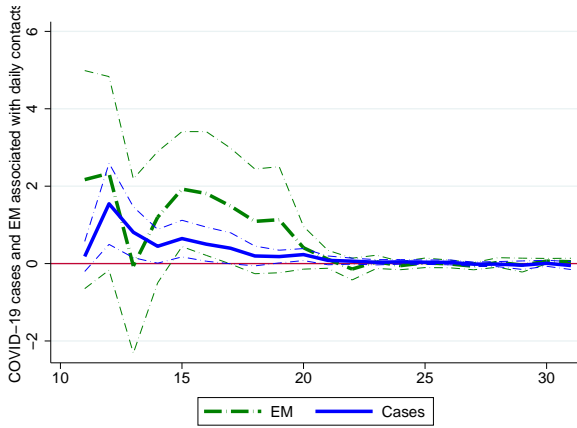


(f) Weeks 31-51

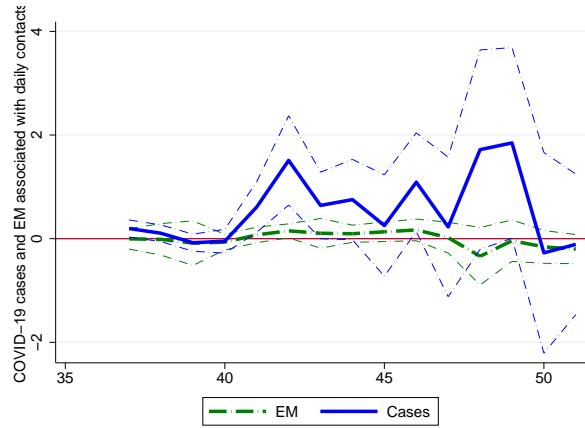
Notes: Figures show the marginal effect of face-to-face contacts on COVID-19 cases and excess mortality using weekly data. Coefficients associated with COVID-19 cases (excess mortality) for subfigures (a) and (b) are obtained from a regression such as column [1] in Table D.1 (D.2). Coefficients associated with COVID-19 cases (excess mortality) for subfigures (c) and (d) are obtained from a regression such as column [3] in Table D.1 (D.2). Coefficients associated with COVID-19 cases (excess mortality) for subfigures (e) and (f) are obtained from a regression such as column [5] in Table D.1 (D.2). We use a rolling window estimation with 7 weeks of sample size. Blue (green) dashed lines indicate 90% confidence intervals for cases (excess mortality).

Figure D.2: Weekly data: Cases and excess mortality (Alternative IV)

Full sample

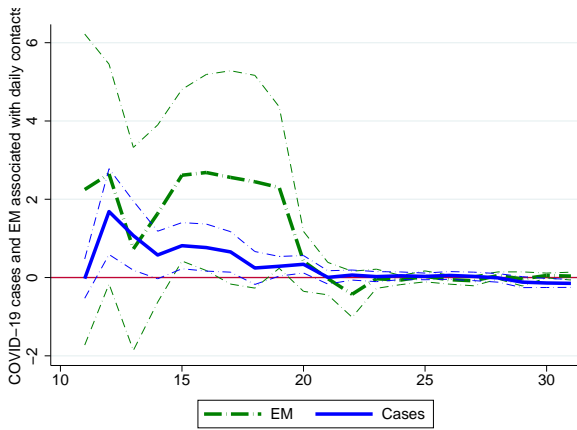


(a) Weeks 5-31

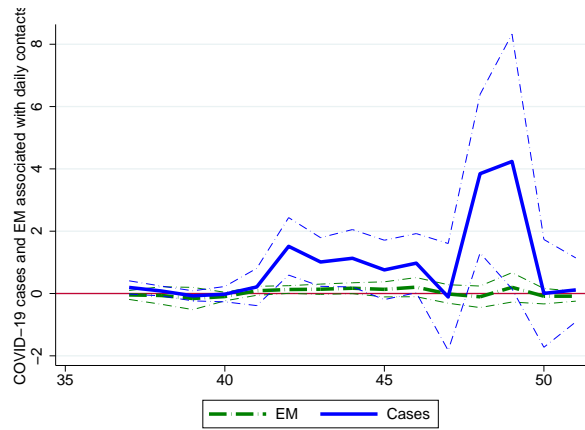


(b) Weeks 31-51

Selected sample

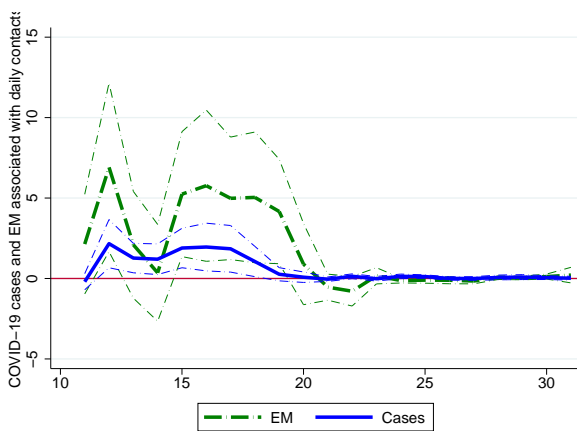


(c) Weeks 5-31

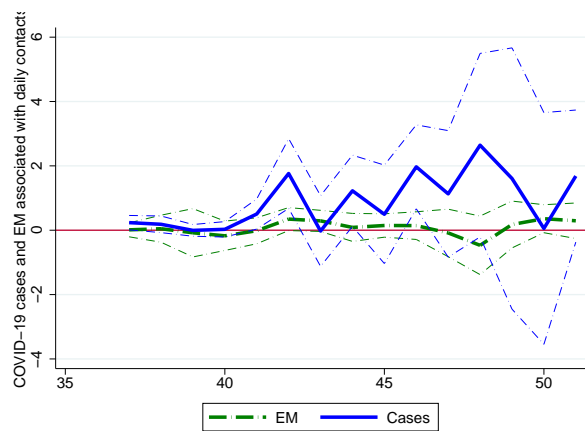


(d) Weeks 31-51

Weighted



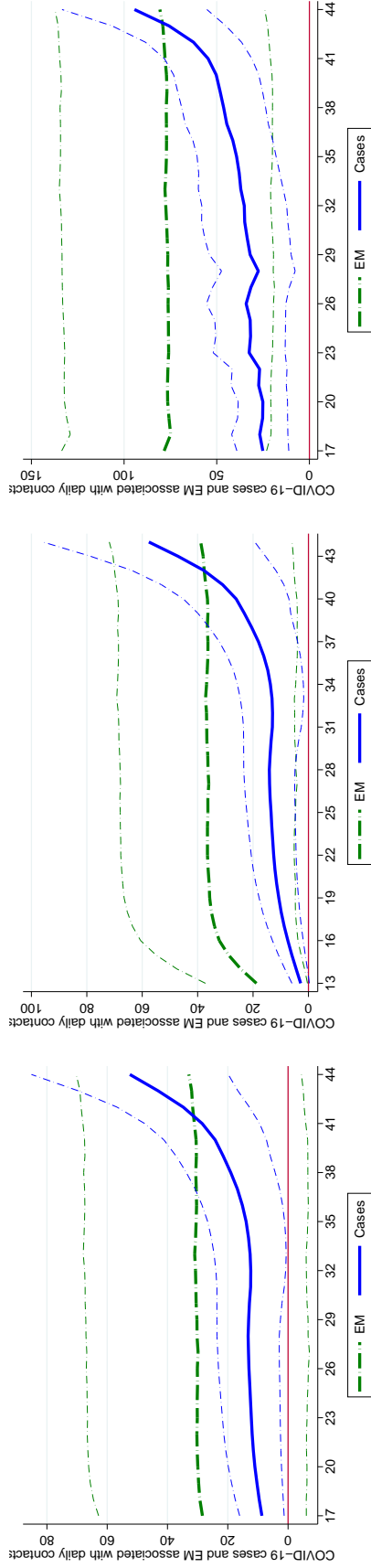
(e) Weeks 5-31



(f) Weeks 31-51

Notes: Figures show the marginal effect of face-to-face contacts on COVID-19 cases and excess mortality using weekly data. Coefficients associated with COVID-19 cases and excess mortality are obtained in the same way as in Figure D.1 with the difference that now we also consider endogenous the lagged inverse distance weighted sum of new cases in other regions. Beyond the previous instruments (daily heating degree days, runoff, wind speed at 10 meters, and the interaction between runoff and wind speed), we further include the inverse log distance weighted sum of each of these variables in other regions, during the preceding third and fourth weeks, as instrumental variables in the IV regressions. We use a rolling window estimation with 7 weeks of sample size. Blue (green) dashed lines indicate 90% confidence intervals for cases (excess mortality).

Figure D.3: Cross-section: Cumulative cases and excess mortality, weekly data



(a) Full sample

(b) Selected sample

(c) Weighted

Notes: Figures show the marginal effect of face-to-face contacts on cumulative COVID-19 cases and excess mortality using weekly data and cross-section regressions. The coefficients are obtained from a regression such as $TotY_i = \alpha_{betwecns} + \sum_{i \neq j} d_{ij}^{-1} TotY_j + \beta_{11} X_i + \beta_{12} G_i + \beta_2 Z_i + \mu_{nuts} + I^c * \phi_t + \epsilon_i$, where the dependent variable $TotY_i$ represents the cumulative number of COVID-19 cases or the cumulative excess mortality in a given week in region i and $\sum_{i \neq j} d_{ij}^{-1} TotY_j$ is the inverse log distance weighted sum of the cumulative number of COVID-19 cases in the rest of regions (not included in the excess mortality regression). Cumulative number of COVID-19 cases are calculated as the sum of weekly averages of daily cases over time. Cumulative excess mortality is the sum of the excess mortality over time. The rest of control variables are explained in Section 3. The time variation in the figures is obtained by estimating the previous regression in different weeks. "Full sample" refers to our baseline regression where we include any country with available data in the variables of interest. The "Selected sample" includes: Germany, France, Italy, Spain, and UK. "Weighted" presents the results from a weighted regression using the full sample and the number of EQLS surveys in each region as a weight. Blue (green) dashed lines indicate 90% confidence intervals for cases (excess mortality).