



HAL
open science

A Poorly Understood Disease? The Impact of COVID-19 on the Income Gradient in Mortality over the Course of the Pandemic

Paul Brandily, Clément Brébion, Simon Briole, Laura Khoury

► **To cite this version:**

Paul Brandily, Clément Brébion, Simon Briole, Laura Khoury. A Poorly Understood Disease? The Impact of COVID-19 on the Income Gradient in Mortality over the Course of the Pandemic. 2021. halshs-02895908v2

HAL Id: halshs-02895908

<https://hal.science/halshs-02895908v2>

Preprint submitted on 15 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



WORKING PAPER N° 2020 – 44

**A Poorly Understood Disease? The Impact of COVID-19 on the
Income Gradient in Mortality over the Course of the Pandemic**

**Paul Brandily
Clément Brébion
Simon Briole
Laura Khoury**

JEL Codes: I14; I18; R00

**Keywords: COVID-19 ; poverty; mortality inequality; labor market; housing
conditions.**



A Poorly Understood Disease? The Impact of COVID-19 on the Income Gradient in Mortality over the Course of the Pandemic*

Paul Brandily[†] Clément Brébion[‡] Simon Briole[§] Laura Khoury[¶]

First version: July 2020

This version: March 2021

Abstract

Mortality inequalities remain substantial in many countries, and large shocks such as pandemics could amplify them further. The unequal distribution of COVID-19 confirmed cases suggests that this is the case. Yet, evidence on the causal effect of the epidemic on mortality inequalities remains scarce. In this paper, we exploit exhaustive municipality-level data in France, one of the most severely hit country in the world, to identify a negative relationship between income and excess mortality within urban areas, that persists over COVID-19 waves. Over the year 2020, the poorest municipalities experienced a 30% higher increase in excess mortality. Our analyses can rule out the contribution of policy responses (such as lockdown) in this heterogeneous impact. Finally, we show that both labour-market exposure and housing conditions are major determinants of the direct effect of the epidemic on mortality inequalities, but that their respective role depends on the state of the epidemic.

JEL classification: I14; I18; R00

Keywords: COVID-19 ; poverty; mortality inequality; labor market; housing conditions

*Declarations of interest: none.

[†]Paris School of Economics, France. paul.brandily@psemail.eu

[‡]Copenhagen Business School, Denmark. cbr.msc@cbs.dk

[§]Paris School of Economics & J-PAL Europe, France. simon.briole@psemail.eu

[¶]Norwegian School of Economics, Norway. laura.khoury@nhh.no

1 Introduction

Despite an unprecedented worldwide decline in mortality over the last century, a substantial income gradient still persists within most countries (Case et al., 2002; Cutler et al., 2006; Chetty et al., 2016; Currie et al., 2020). These structural inequalities result primarily from individual differences between the rich and the poor in health behaviours, information, economic or social stress (Butikofer et al., 2021) and from ecological factors, like the living environment or public health institutions (Chetty et al., 2016; LeClere et al., 1997; Robert, 1999; Pickett and Pearl, 2001). However, sudden exogenous events like pandemics can also significantly affect these inequalities. Historical studies have shown that past epidemics had unequal effects on the mortality of different socio-economic groups, either because they revealed latent inequalities in individual health capital or because they spread differently across living environments (Snowden, 2019; Alfani, 2021; Beach et al., 2021).

The COVID-19 crisis epitomizes such massive mortality shock, on a worldwide scale. In 2020, about 100 million people have tested positive for COVID-19 – among which 2.2% died – with a large number of infections and deaths also remaining undetected.¹ Assessing the causal impact of this pandemic on mortality inequalities is therefore key, not only to inform public policy in the context of the crisis, but also to study the role of pandemics in shaping mortality inequalities from a historical perspective. However, while a growing literature has pointed out that most vulnerable socioeconomic groups were more often affected in the early months of the pandemic, most of the existing evidence remains essentially correlational.

A first challenge in measuring the causal impact of the epidemic on mortality inequality consists in isolating the specific contribution of COVID-19. Even in the absence of the pandemic, the structural health inequalities in terms of income could potentially explain the correlation between mortality and income in 2020. Disentangling between the direct effect of the pandemic and that of the policy responses (such as lockdown) represents a second important challenge which has been disregarded by the literature. Finally, a third challenge relates to the dynamics of the pandemic: many studies are based on early and short-term measures. It is unclear whether these results only reflect harvesting effects due to a temporary mortality shock on individuals particularly at risk. In particular, one cannot exclude that the income gradient gradually disappears, or is even reversed, as time goes by and the disease spreads.

In this paper, we address these challenges in the context of France, one of the most severely hit country in the world. Using exhaustive administrative data sets, we are able to study the links between income and the 2020-specific excess mortality at a very local level, for the entire country and over a long

¹Source: [European Centre for Disease Prevention and Control](#).

time period. We combine daily death records of all-cause mortality provided by the French National Statistical Institute (INSEE) for the three entire years 2018-2020 with information on the municipality of residency, a very detailed geographic scale (1,600 inhabitants and 15 sq.km per municipality, on average). Leveraging the precision of our data, we study what happens, on average, *between* municipalities *within* urban areas exposed to COVID-19. Working at such a fine geographic scale allows to account for local interactions between individuals that are central to the development of health inequalities (LeClere et al., 1997; Robert, 1999; Pickett and Pearl, 2001), and all the more so in the case of an epidemic.²

There are three main results in this paper. First, we confirm the existence and persistence of an income gradient in the 49,495 excess deaths that occurred in French urban areas over the two distinct waves of the epidemic (in Spring and late Fall 2020). Within a given urban area, and once population size and age are controlled for, the difference in excess mortality rates between the poorest municipalities (bottom quartile of the income distribution) and the richer ones equals 2.6 (deaths per 10k. inhabitants), *i.e.* 30% of the excess mortality in the non-poor municipalities in 2020. We show that the gradient uncovered during the first wave is not compensated during the second wave, but rather reappears with regularity every time the epidemic returns, so that it is the strongest in areas that got most affected by COVID-19 in 2020. Second, we propose a decomposition of COVID-19 total impact into a direct component – from COVID-19 infections – and an indirect component (*e.g.* mental-health consequences of lockdown policies, increased domestic violence, road accidents avoided because of the lockdown, etc. (Banerjee et al., 2020; Belot et al., 2020; Caul, 2020; Mulligan, 2021; McIntyre and Lee, 2020; Bullinger et al., 2020)). We rule out the existence of a significant income gradient in the indirect component. Third, examining mechanisms, we show that labor-market exposure and housing conditions are important determinants of the variation in COVID-19 mortality and largely explain the income gradient.

This article makes several contributions to the literature on the income gradient in the specific case of COVID-19. Many papers have shown that the first wave affected more strongly deprived neighborhoods and individuals based on confirmed cases (Chen and Krieger, 2021; Abedi et al., 2020; Ashraf, 2020; Kim and Bostwick, 2020; Williamson et al., 2020; Drefahl et al., 2020; Jung et al., 2020; Caul, 2020).³ Our data allows to improve on the measurement of the specific contribution of COVID-19 to health inequality

²Individual characteristics affecting one’s likelihood of infection also affect infection risks of others living or working nearby. Note that we introduce in our models urban-area fixed effects, so that we absorb specific local factors that may foster or hinder the spread of the epidemic (*e.g.* public transportation) and are unlikely to be independent from municipalities’ income.

³A few exceptions should be mentioned. Comparing US counties, Brown and Ravallion (2020), Desmet and Wacziarg (2020) and Knittel and Ozaltun (2020) find no association between median income or poverty rates and COVID-confirmed deaths. Jung et al. (2020) finds a positive relation between poverty rates and confirmed deaths in counties with a low density but rather a U-shaped relation in dense areas. According to Sa (2020), the positive association between socioeconomic deprivation and confirmed COVID deaths in the UK turns negative once self-declared health quality is taken into account.

in two ways. First, confirmed cases underestimate the number of deaths attributable to the pandemic. In particular, they do not account for false negatives or deaths on which no test has been carried out, either because testing material was lacking in hospital or simply because they occurred at home. Unreported cases are highly unlikely to be independent of income and may therefore bias the results of the literature based on confirmed cases. To avoid such bias, we favour all-cause mortality to confirmed cases in our estimations. Second, our approach allows to better account for the structural income heterogeneity in mortality that is found absent COVID-19. Many articles have previously documented mortality inequalities between socio-economic groups and it could then be the case that the income heterogeneity in COVID-19 confirmed deaths only reflect these. In particular, in the US, where most of the evidence on the income gradient comes from, these structural inequalities may be particularly strong since the access to health care is very unevenly distributed among the population.⁴ We therefore use all-cause excess mortality as our main dependent variable, defined as the deviation in 2020 all-cause mortality with respect to a counterfactual (pre-COVID) period – namely the average of all-cause mortality in 2019 and 2018.⁵

To our knowledge, very few articles have worked on all-cause excess mortality, and all of them focused on the first wave (Calderón-Larrañaga et al., 2020; Krieger et al., 2020; Decoster et al., 2020). By using two successive waves in France, our paper adds to this literature by evaluating the evolution of the income gradient over time. Our analysis reveals that the relationship between income and COVID-19 mortality applies with regularity to each wave of the epidemic and that the mortality differential created does not disappear over time.⁶ The epidemic caused an increase in differential mortality between rich and poor municipalities that is not a simple one-time shock or anticipation effect in the deaths of the most vulnerable but is rather deepening after each wave.

This article is also the only one to disentangle between the direct and indirect effects of COVID-19 infections on mortality. Beyond the direct effect of COVID-19 infections, the epidemic could have an indirect impact that is non-orthogonal to households' income (*e.g.* mental-health consequences of lockdown policies, increased domestic violence, road accidents avoided because of the lockdown, etc. (Banerjee et al., 2020; Caul, 2020; Mulligan, 2021; McIntyre and Lee, 2020; Bullinger et al., 2020; Calderon-

⁴By contrast, measuring the income gradient in a country like France is informative on the unequal impact of COVID-19 despite a rather equal access to health care. Comparing the income gradients found in both countries is interesting to the extent that France and the US systematically rank at both ends of the distribution of OECD countries in terms of equality of access to health care (source, see also Currie et al. (2020) for a comparison of mortality trends in the US and in France.)

⁵Two exceptions to the literature using confirmed cases should be mentioned: Caul (2020) describes both confirmed cases and all-cause mortality – but with no counterfactual; Chen and Krieger (2021) use a counterfactual but do not consider all-cause mortality.

⁶More specifically, a surplus of 1.18 (resp. 1.08) excess deaths per 10k. inhabitants is found in the bottom income quartile in the first wave (resp. second wave), as compared to richer municipalities.

Anyosa and Kaufman, 2021) and may thereby drive part of the income gradient found in the early literature.⁷ We exploit the quasi-natural experiment provided by the first lockdown to disentangle between both dimensions. Its uniform implementation over the territory froze the epidemic at very different stages of development. In zones with a low level of infection at the start of the lockdown, the propagation of the virus was stopped *before* it could have a significant impact on mortality. By contrast, in the high-infection zone, the virus had already circulated enough to have a large impact on mortality, *despite* the lockdown. We find no income gradient in the within-urban-area mortality in the low-infection zone, suggesting that the overall income gradient is entirely driven by the direct effect of COVID-19 infections.⁸

Finally, our paper speaks to the literature on the mechanisms underlying the relationship between income and COVID-19 mortality. Previous papers have highlighted the importance of labor-market exposure or housing conditions (including the first version of this paper, as of July 2020, as well as Almagro and Orane-Hutchinson (2020); Almagro et al. (2020); Glaeser et al. (2020); Naticchioni et al. (2020)). Building on exhaustive administrative data, we confirm the previously uncovered role of increased exposure inside and outside the house.⁹ We extend the aforementioned literature by examining the dynamics of the mechanisms effect under different policies and by trying to quantify the respective weight of each mechanism in explaining the variation in COVID-19 mortality.¹⁰ The impact of the share of workers with frequent social contact at work decreases in magnitude during the second wave, suggesting that employers have adjusted the management of the occupational risk over time. Housing conditions also matter less during the second wave, as the lockdown was less stringent. A horse race performed between variables related to occupational exposure and housing conditions indicates that our variables – and especially the share of essential workers and of overcrowding housing – almost fully explain the income gradient in COVID-19 related mortality.

The remainder of the paper is organized as follows: section 2 describes the data and the construction of our outcome variable, while section 3 explains the context. In section 4, we present the first evidence of an income gradient in COVID-19 mortality in France that persists across waves. In section 5, we

⁷According to Mulligan (2021), about 12% of March-to-October excess mortality in the US was due to non-COVID causes (a large part being the result of "deaths of despair").

⁸We treat the low-infection zone in the West and South of France as a control group: any income gradient in excess mortality found on this group during the first wave is considered as a measure of the differentiated impact (between poor and non-poor municipalities) of the indirect effects of COVID-19 (including the lockdown). Reassuringly for our estimation strategy, we find an income gradient in this zone during the second wave.

⁹Further mechanisms may also include: greater levels of air pollution in poorer areas (Cole et al., 2020; Persico and Johnson, 2020), lower levels of compliance to lockdown and to self-protection measures among low-income individuals Papageorge et al. (2020), more comorbidities among poor households (Wiemers et al., 2020; Raifman and Raifman, 2020) among others. While population density is also an often debated mechanism, we do not treat it as a potential mechanism given its non-significant correlation with poverty in our data.

¹⁰Almagro et al. (2020) also examine the dynamics of the effect of labor-market exposure and housing conditions in New York City but over a shorter time period and under a different institutional context.

disentangle between the indirect and the direct effects of COVID-19 on excess mortality. Given these results, section 6 explores potential mechanisms and section 7 concludes.

2 Data and measurement

Data sources

In this paper, we gather various individual-level data sets that we match using municipality-level identifiers. In France, municipalities are very small administrative units of 1,600 inhabitants and 15.3 sq.km on average; there are about 35,000 of them in 2020. Our analysis compares municipalities within urban areas. Urban areas are groups of neighboring municipalities, defined by the French National Statistical Institute (INSEE) on the basis of commuting patterns. We consider 16,640 municipalities distributed across 621 urban areas, each being made of 27 municipalities and hosting 85,000 inhabitants on average.¹¹ The majority of our sources are exhaustive administrative data sets made available by INSEE. We provide more details on all data in appendix A1.

Measuring the impact of the epidemic on mortality

Most previous works studying the unequal impact of the pandemic are based on COVID-19 confirmed cases (infection or death). These measures suffer from several conceptual limitations.¹² First, the total impact of an epidemic on mortality is a function of both the direct effects of infections (D_d), and the indirect effects of public policies taken as a response to infections (D_i). The latter relates to social changes caused by the epidemic, such as (i) physical, psychological, and social effects of distancing¹³ (ii) economic changes (Banerjee et al. (2020)). Furthermore, D_d can be broken down into reported infection-caused deaths ($D_{d,r}$), unreported infection-caused deaths ($D_{d,u}$), and deaths due to altered access to health services ($D_{d,a}$).¹⁴ Studies based on confirmed cases not only ignore the indirect effects of the epidemic but can also suffer from severe biases arising from the salience of unreported cases and the saturation of health services.

A growing literature confirms that measures based on COVID-19 reported cases actually include significant biases. First, testing capacities and strategies have proven to vary substantially over space and time, not only across countries but even across regions or neighbourhoods (Kung et al., 2020; Rivera et al.,

¹¹Due to our within urban area approach, we exclude smallest urban areas made of only one municipality from the analysis. Tables B1 and B2 provide more descriptive statistics on our sample at the urban area and the municipality level, respectively.

¹²See Böttcher et al. (2021) for a detailed discussion on this topic.

¹³For instance, evidence suggests that the lockdown have reduced the transmission of the seasonal flu and of norovirus (Kraay et al., 2020).

¹⁴Because hospital saturation is “naturally” caused by infection rates and not by any policy response to the epidemic, we conceptually treat it as a part of the direct effect in the remainder of the paper.

2020; Borjas, 2020; Balmford et al., 2020; Silverman et al., 2020; Yorifuji et al., 2021; Wu et al., 2021). France is no exception to this rule: it ranked among the worst OECD countries in number of tests per inhabitant in May 2020 (Scarpetta et al., 2020; Foucart and Horel, 2020). Another source of bias comes from the non-random testing of the population in a context of a shortage in testing resources. Typically, tests are more often conducted on symptomatic persons, the elderly or socio-economic vulnerable people, leading to overestimation of infection rates (Beaney et al., 2020; Böttcher et al., 2021). Finally, some individuals who die from COVID-19 are not accurately identified, due to difficulties in attributing the cause of death. This may particularly bias the number of reported cases in France, where no system was able to report the COVID-19 mortality at home (Fouillet et al., 2020).

A second limitation of analyses using only measures based on COVID-19 confirmed cases is that they ignore counterfactual mortality. To conclude about an income gradient from COVID-19 observed death actually presumes that such gradient would not exist absent the epidemic. In other words, it fails to take into account structural health inequalities between socio-economic groups. Such inequalities have been consistently documented in a wide range of countries (Case et al., 2002; Cutler et al., 2006; Adler and Rehkopf, 2008; Cutler et al., 2012; Currie and Schwandt, 2016; Mackenbach et al., 2019) and have been recently highlighted in the case of France, where the richest 5% men (resp. women) have a life expectancy of 13 (resp. 8) years longer than the poorest 5% on average over the 2012-2016 period (Blanpain, 2018).¹⁵ As a consequence, accounting for the baseline inequality in mortality between rich and poor (in the absence of COVID-19) is key to accurately identify the specific contribution of the epidemic on inequality in mortality.

To address these issues, we build measures of all-cause excess mortality at the municipality level based on daily counts of deaths from INSEE. For every single individual who died in France in 2018, 2019 or 2020, this data set provides the municipality and place of death (hospital or clinic, home, care home, etc.) as well as a set of individual-level characteristics such as sex, date of birth and municipality of residency.¹⁶ Such data makes it possible to compare the mortality for the residents of a given municipality, during a given period in 2020, to the mortality in the same municipality during the same period in the two previous

¹⁵We also find such inequalities in our data in 2018 and 2019 (see Section 5.1).

¹⁶In the emergency of the COVID-crisis, INSEE made the data set available at a high frequency rate: this has induced potential measurement errors although INSEE applies corrections at each iteration. However municipalities have about one week to inform INSEE about new death. As the version of the data that we are currently using got updated on February 5, 2021, our measures should be accurate. For a discussion on the quality of the data, see the [information provided by INSEE](#).

years,¹⁷ before the COVID-19 outbreak. For each municipality m , we therefore define excess mortality during period p as follows:

$$D_m^p = \frac{N_m^{p,2020} - [0.5 \times (N_m^{p,2018} + N_m^{p,2019})]}{Population_m^{2014}} \quad (1)$$

where $N_m^{p,y}$ is the number of deaths of residents of municipality m during period p of year y and $Population_{m,2014}$ is the total number of inhabitants (/10,000) of municipality m , as recorded in 2014, the most recent available year in our data.

Using all-cause excess mortality allows us to study the impact of the COVID-19 epidemic on mortality based on a measure that is of constant quality over space and time and that is available at a fine level for a very large territory. We are therefore able to avoid the usual biases found in the literature that are due to testing or death reporting issues. Furthermore, our counterfactual analysis also allows us to take into account the structural inequalities between rich and poor in terms of mortality in order to accurately identify the impact of the epidemic on these inequalities. Finally, our measure allows us to account for both the direct and the indirect effects of the epidemic, the latter being generally ignored in the literature studying the unequal impact of the epidemic. However, one may also be interested in isolating the direct effects of the epidemic on mortality inequalities. In section 5, we develop a triple-difference strategy based on a quasi-experimental setting, which exploits natural variations in infection rates over the French territory at a very early stage of the epidemic to net out the indirect effects of the epidemic caused by the policy response. This strategy allows us to isolate the direct effect of COVID-19 infections and to study how it is distributed based on a reliable and unbiased measure of mortality.

A municipality-level analysis

The main goal of this paper is to estimate the impact of the COVID-19 epidemic on inequalities in mortality. Our approach at the municipal level allows us to study this issue from a geographical point of view, at a fine local level, and to estimate the average effect of *living in a poor area* on COVID-19 related mortality.¹⁸ Another - and complementary - research question relates to the estimation of the individual effect of *being poor* on COVID-19 mortality (*i.e.* the individual income gradient), for which an analysis at the individual level would be better suited. Such an analysis has the disadvantage of ignoring the role of local factors as well as that of interactions between individuals living or working in the same geographical

¹⁷We use the average of years 2018 and 2019 because taking several years makes our measure less sensitive to random events occurring in a given year. For data availability reasons, we cannot go further back in time, since death data before 2018 do not include, at the moment, information on the municipality of residency.

¹⁸Currie and Schwandt (2016) take a similar approach to study the link between poverty and mortality from a historical perspective.

area – which are known to influence the income gradient in life expectancy in normal times (Chetty et al., 2016). Given the nature of an epidemic, it is clear that individual characteristics affecting the likelihood of infection of one individual - such as poverty - will also affect the likelihood of infection of others living or working nearby. Conducting the analysis at the municipality level therefore allows us to account for local interactions between individuals and for the spillovers on COVID-19 infection probability (and related mortality) resulting from these interactions.¹⁹

3 COVID-19 in France: context and background

3.1 The temporal dynamics of the epidemic in 2020

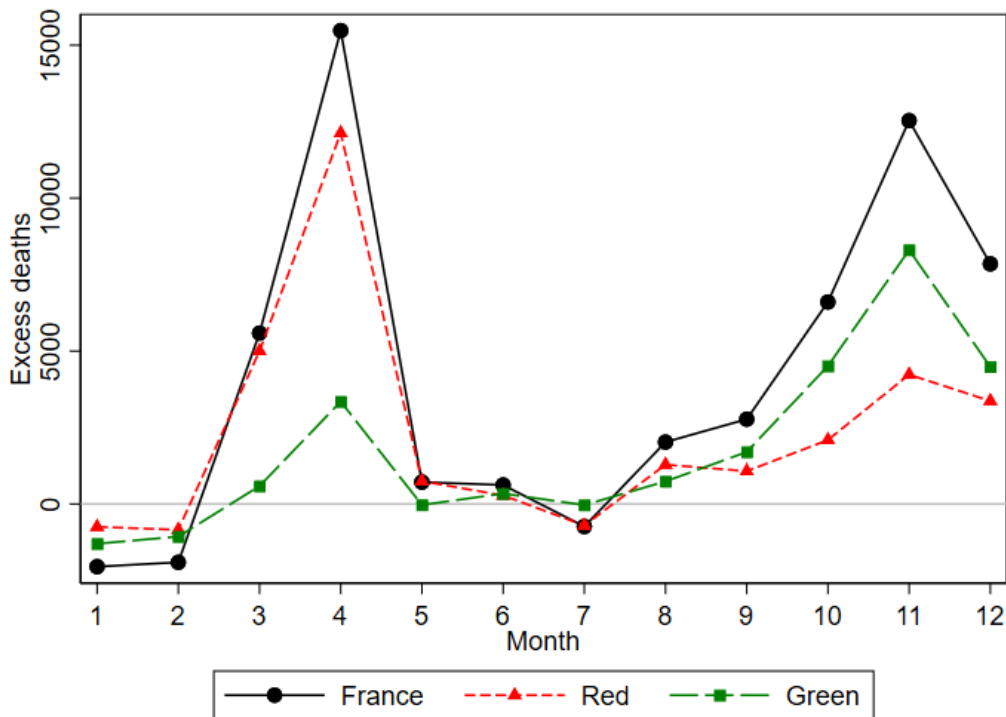
Overall in 2020, there was 49,495 deaths in excess in French urban areas in comparison with the average of 2018 and 2019. Figure 1 represents the distribution across months of these extra deaths in France in 2020 (black line). It clearly appears that, as in many European countries, excess deaths in 2020 essentially occurred over two distinct waves that peaked in April (15,479 deaths) and November (12,537 deaths), respectively. In the remainder of this paper, we will refer to the “first wave” as the period covering March and April 2020; and the “second wave” as the period October to December 2020.

In both cases, the French Government reacted by taking extraordinary containment measures. As COVID-19 first spread in the country in early 2020, the government decided of a national lockdown on March 17 and that eventually lasted until May 11. This first lockdown was the most stringent and seems to have reduced the impact of COVID-19 drastically. All workers stayed home except if their activity was deemed essential for the country and visits to elderly care homes (hereafter EHPADs) were forbidden. After a lull in the summer, the epidemic developed again until a second lockdown was decided on October 30 and continued until December 15. This second lockdown was less stringent however: (i) on top of essential private industries, factories, firms from the construction sector as well as most public services remained open; (ii) to the exception of universities most schools kept on receiving students; (iii) visits to EHPADs were allowed; (iv) parks and gardens remained open. The second lockdown got also repealed quicker, with the end of a first phase after a month when all shops opened again. Based on Google mobility data (Google LLC, 2021), Figure B1 (black line) provides evidence that the increase in time

¹⁹To the best of our knowledge, the only study able to look differentially at the individual vs the local relationship between poverty and COVID-19 mortality is the one by Decoster et al. (2020), who exploit mortality data from the first wave of the epidemic (March-May) in Belgium. Their results point to a much larger gradient at the municipal than at the individual level, confirming the existence of strong spillovers and/or local factors at the municipal level.

spent at home is much higher when a lockdown policy is in place, and more so during the first lockdown than the second.

Figure 1: Monthly counts of excess deaths in French urban areas



NOTE: The figure represents the difference between the monthly number of deaths in 2020 and its average over 2019 and 2018 in the relevant zone. The “red” zone corresponds to the areas that were the most severely hit by the first wave, and that are located in the North-Eastern quarter of the country. This zone covers about 44% of the urban population of (mainland) France. The “green” zone encompasses the rest of the French territory.

3.2 Spatial heterogeneity in excess deaths

An important feature of the French case is that the spread of the epidemic was still very uneven across the country when the first lockdown was implemented (March 17, 2020). While the North-East was severely hit, the West and the South were then almost spared by the epidemic. In the remainder of this paper we call the former region the “red zone” and the latter the “green zone”. Figure 1 exhibits this spatial heterogeneity in excess deaths during the first wave. We reproduce the death toll separately for the North-East (red line) and the rest of the country (green line), the former indicating regions where COVID-19 was highly active when the first lockdown started. As is clear in the Figure, the first wave mostly hit the red zone, which accounts for 80% of French excess deaths in April 2020 - while it only

represents 44% of the total urban population. By contrast, the green zone suffered a much smaller increase in mortality over the same period.²⁰

Figure 1 thus summarizes the spatio-temporal dynamics of the COVID-19 epidemic in France in 2020. We distinguish four phases, 1: January and February are marked by a relatively low mortality in all regions of France.²¹ 2: March and April represent the “first wave” that mostly hit the North-Eastern quarter (red zone) of the country. 3: May to September, mortality is close to normal. 4: October to December, a second wave hit the country, this time more homogeneously across the territory.

As shown in appendix Tables B3 and B4, the red zone contains less urban areas than the green zone (191 vs. 430) but these are bigger on average (39 municipalities and 12,000 inhabitants vs 22 municipalities and 6,800 inhabitants) and have more deprived municipalities (25% vs. 20%). Breaking down these statistics at the municipality level reveals however that cities in both zone are not so dissimilar. The median mortality rates in 2018 and 2019 (72 vs. 75) are very similar, and so are the share of elderly (16% vs. 17%) and median income (22,000 vs. 21,000 euros).

4 The impact of COVID-19 on the income gradient in mortality

This section highlights a key heterogeneity in all-cause excess mortality in France that appears consistently over the two waves: an income gradient in the total impact of the COVID-19 epidemic on mortality.

4.1 A first stylized fact

In 2020, poorest municipalities in France suffered from a greater increase in mortality, as compared to other municipalities. This stylized fact appears very clearly in Figure 2, which displays the (non-parametric) monthly cumulative number of excess deaths (per 10K. inhabitants) in 2020, separately for the poorest municipalities and other municipalities.²²

To categorize a municipality as poor, we first rank all municipalities according to the median income of their inhabitants and we define as “poor” all the municipalities that fall into the first quartile (Q_1) of this distribution (cf. A1 and A2 for details and data). We weight each municipality by its total population

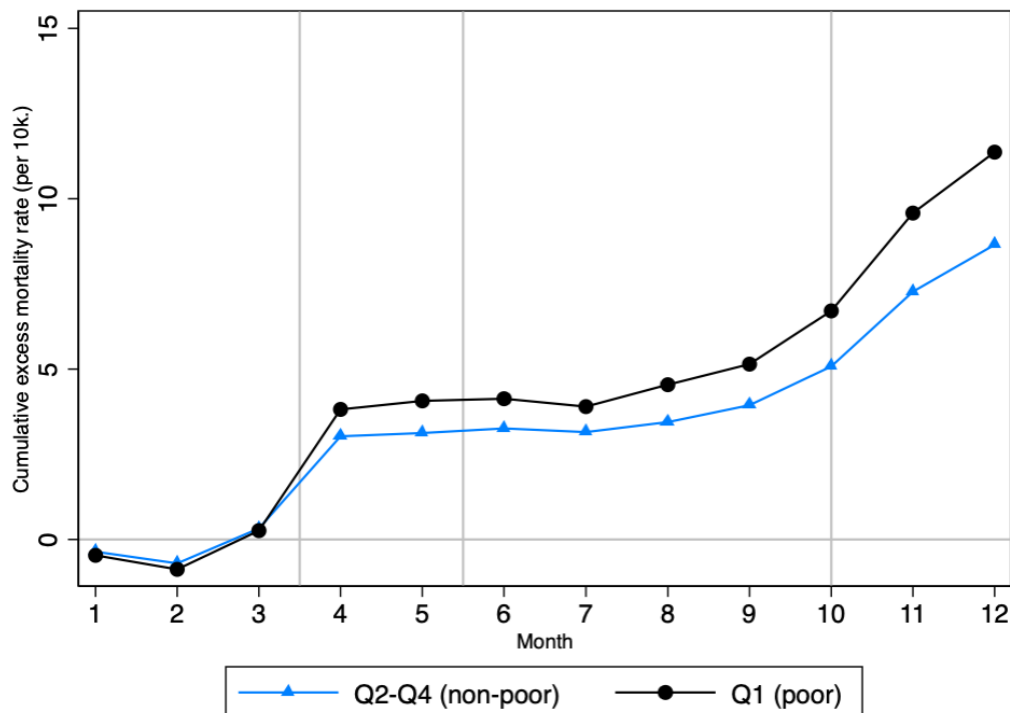
²⁰The classification between green and red zones follows a map issued by the Government at the end of the first lockdown (on May 7) that we detail in Appendix A2. The map is defined at the *département* level. A *département* is an administrative unit in France that encompasses on average 363 municipalities and 645,000 inhabitants. There is a total of 101 *départements* in France. Tables B3 and B4 detail the total number of urban areas, municipalities, and inhabitants in each zone.

²¹The seasonal flu was particularly strong in 2018.

²²The cumulative number of excess deaths in month m corresponds to our excess mortality variable computed over the period from January to the month m .

in order to give an equal weight to each individual. The municipalities included in $Q1$ therefore represent the 25% of the French urban population living in the municipalities with the lowest median income.

Figure 2: Cumulative excess mortality rate per 10,000 inh. by poverty status



NOTE: The graph plots the cumulative sum of all excess deaths per 10,000 inhabitants from January 2020 for poor and non-poor municipalities. Poor is defined as belonging to the bottom quartile of the national distribution of municipal median income weighted by the municipality size.

Figure 2 shows no specific pattern in the cumulative excess mortality rate in the first three months of 2020 (phase 1, in section 3.1), followed by a marked increase from April (phase 2), that further grows as the second wave takes place (phase 4). By the end of 2020, there was, on average, 11.4 more deaths per 10K. inhabitants than usual in the poorest municipalities in France, against 8.7 in richer municipalities. In the following paragraphs, we employ an econometric approach to check whether this stylized fact holds after accounting for the influence of age and local factors.

4.2 Empirical approach

We check that the descriptive evidence displayed in Figure 2 holds parametrically and test its robustness. We contrast the evolution of excess mortality in poor municipalities with that of richer municipalities

over the whole French territory, taking a synthetic year (the average of 2018 and 2019) as the baseline. Our main model thus writes as:

$$D_{[m,ua]}^p = \beta \cdot Q1_{[m,ua]} + X_{[m,ua]}^p \cdot \Lambda + \gamma_{ua} + \nu_{[m,ua]}^p \quad (2)$$

Where m indicates the municipality, ua an urban area and p a given period of the calendar year. $D_{[m,ua]}^p$ represents municipalities' excess mortality rate as defined in section 2 (*i.e.* as the deviation to the average 2018 and 2019 mortality rate of the same municipality over the same period p). In our analysis, every municipality m belongs to one single urban area ua , and we therefore drop this redundant subscript in subsequent models. $X_{m,ua}$ is a vector of controls including the total population size and the share of the population above 65 years old. Importantly, we introduce γ_{ua} , an urban-area fixed effect so that we only exploit differences between municipalities located within a contiguous urban environment. This is an important aspect of our model since it absorbs specific local factors that may foster or hinder the spread of the epidemic (*e.g.* public transportation) and are unlikely to be independent from municipalities' income. Our results are thus based on comparisons at a very fine spatial level. Municipalities are weighted by their total population. Standard errors are clustered at the urban-area level.

The coefficient of interest is β . It estimates the difference in excess mortality rate between municipalities from the poorest ($Q1$) and the other quartiles (comparison group). This model identifies the heterogeneity in the causal (total) effect of the pandemic on excess mortality under the sole hypothesis that, absent COVID-19 and the associated public policies, the average difference in the evolution of mortality over period p (2020 vs. before) between rich and poor municipalities of the same urban area would have remained stable. Below, we use falsification tests to provide evidence that this assumption is sensible.

4.3 Results

Table 1 reports the estimates associated to equation 2. Column 1 is estimated on all urban areas from mainland France and using the cumulative excess mortality rate over 2020 as a dependent variable. On average, within a given urban area, and once population size and age are controlled for, municipalities of the poorest quartile had an excess mortality rate of 2.627 (deaths per 10k. inhabitants) higher than other municipalities. This has to be compared with the baseline average of 8.649 across municipalities of the other quartiles.

We next consider three sub-periods for 2020 that correspond to the dynamics of the epidemic described in section 3.1. In column (2) we report the coefficients associated with equation 2 when considering only the first wave, that is for March and April (*i.e.* $p = (3, 4)$), while in column (3) we consider the second wave, from October to December (*i.e.* $p = (10, 11, 12)$). The income gradient is significant in both waves and slightly higher than 1 extra dead per 10,000 inhabitants. By contrast, we observe no gradient when we focus on excess mortality outside of the two waves. In column (4), we consider a synthetic period made of all the months in 2020 that we did not classify in either wave. In these months, mortality was really close to baseline (*i.e.* 0 excess death).

Table 1: Excess mortality rate by municipality income

	(1) 2020	(2) Wave 1	(3) Wave 2	(4) No wave
Q1 (poor)	2.627*** (0.996)	1.178* (0.672)	1.083*** (0.359)	0.366 (0.297)
Controls	✓	✓	✓	✓
Urban area FE.	✓	✓	✓	✓
Non-poor average	8.649	3.729	4.709	0.211
Observations	16640	16640	16640	16640

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors in parentheses clustered at the urban-area level.

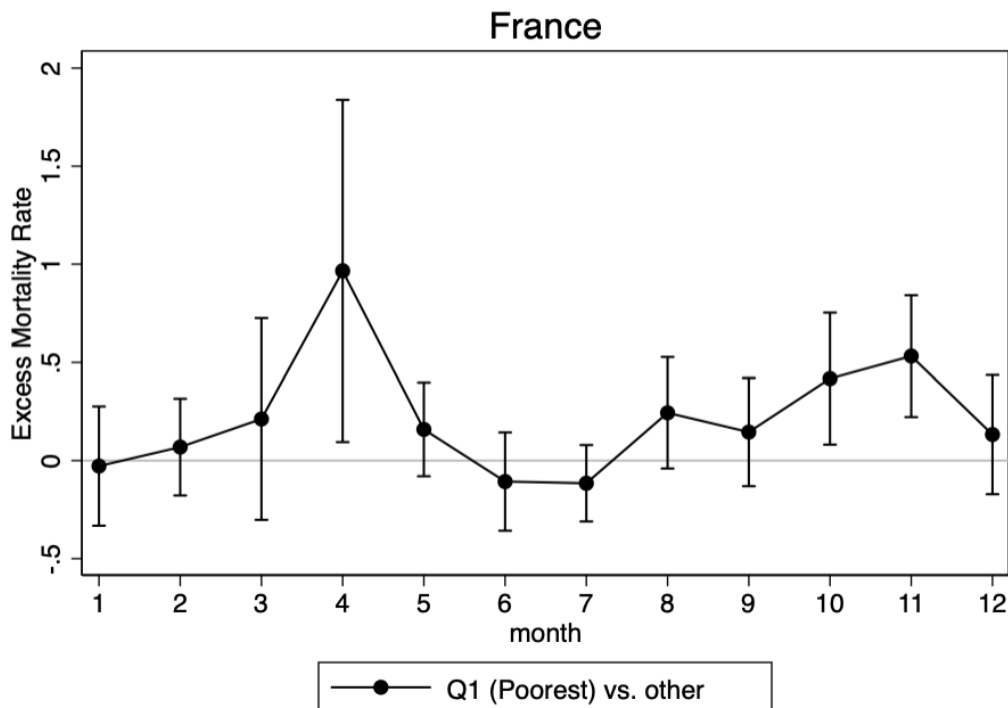
NOTE: This table reports the coefficients associated with equation 2. The independent variable, excess mortality rate, is computed considering four different time periods: the whole year (column 1), wave 1 (March to April, column 2), wave 2 (October to December, column 3) and other months in 2020 outside the two waves (January, February and from May to September, column 4). By construction, column 1 is the sum of column 2 to 4. The non-poor average line reports the mean of the dependant variable in non-poor municipalities. Controls include total population size and the share of the population over 65 years old.

We then further decompose the effect and estimate the model for each of the 12 months separately. Figure 3 plots the estimated β coefficients. The difference in monthly all-cause excess mortality between municipalities of the poorest quartile and richer municipalities closely follows the dynamics of the epidemic depicted in Figure 1 and the β only differs from 0 at the peak of the first and second waves. In other words, in each epidemic wave, mortality increases on average (Figure 1) but even more so in the poorest municipalities (Figure 3).²³

²³It is also important to note that the standard errors are much greater in the first wave than in any other month. This is because the excess mortality only occurred in some part of the country, as explained in section 3.2, an heterogeneity we discuss at length in section 5.

Taken together, these results support the idea that the income gradient we estimate for the whole 2020 year actually reflects the causal (total) effect of COVID-19 on mortality inequalities. They also provide evidence that this causal effect appears with regularity in each wave of the epidemic. We provide clearer and more detailed evidence on this last finding in section 4.4.

Figure 3: Monthly difference in all-cause excess mortality by income



NOTE: The graph plots the point estimate and the 95% confidence intervals of the estimation of β from equation 2 evaluated each month. It accounts for the monthly difference in all-cause excess mortality between the poor municipalities and the rest, where poor is defined as belonging to the bottom quartile of the national distribution of municipal median income weighted by the municipality size.

About the definition of poverty

Our main model compares municipalities from the poorest quartile ($Q1$) to the others. This approach has the advantage of simplicity: by discretizing income into a simple “poor” vs. “non-poor” comparison, it greatly reduces the dimension of the problem and allows the exploration of heterogeneity and mechanisms in the following sections below. In appendix C1 we explain this definition carefully. In particular, we show that contrasting the evolution of mortality in each of the four quartiles leads to a clear monotonic pattern. That is (within an urban area) the increase in excess mortality (during COVID-19 waves) de-

creases in municipalities' income. This monotonicity is robust to a number of alternative partitioning of the distribution. And, given this monotonic income gradient, pooling non-poor municipalities together to form a comparison group actually attenuates the differences in mortality rates we estimate (as $Q1$ is closer to the $Q2$ - $Q4$ average).

Robustness

The presence of a stronger increase in mortality of the poorest municipalities is robust to a number of alternative specifications. We only summarize these results here; comprehensive analysis can be found in Appendix C. As already mentioned, we test alternative grouping of municipalities into distinct quartiles, deciles, and halves of the income distribution (cf. appendix C1). The result also holds when excluding elderly care homes. These were severely hit during the epidemic and one may worry that their location could drive a spurious correlation between municipality-level income and mortality. Removing all death records in such institutions does not alter our conclusion (cf. Table C3). Next, we compute the income gradient separately for different age categories (Figure C3): the income gradient increases with age, except for the category of people over 85 years old, who may represent a particularly vulnerable population irrespective of the level of income. In other words, except for the very old, the size of the gradient increases with the size of the death toll.

Finally, we run an additional falsification test which compares mortality in 2019 and 2020 to the same baseline of 2018. Reassuringly, we find a gradient in 2020 excess mortality but none in 2019 excess mortality (see appendix C2.2).²⁴

4.4 Urban Area analysis

Figure 3 suggests that there is a proportional link between the income gradient and excess mortality at the country level. We directly test this hypothesis at a much finer geographical level, namely the urban-area level. To do so, we first estimate the gradient for each month of 2020, separately for each of the 421 urban areas that include both poor and non-poor municipalities. We then regress this measure on the total excess mortality of the urban area in the same month (see Appendix D for more details). This exercise yields a clear message: an increase in the urban area excess mortality rate by one (death per 10k. inhabitant) is associated with an increase of the gradient of a bit more than 0.3 (more death per

²⁴Given that we only have 2018 as a reference period for this test, our main falsification test compares excess mortality over the whole year so that our comparisons are not polluted by potential seasonal shocks. However, note also that none of the monthly coefficient of 2019 (v.s. 2018) is significant at the 5% level (cf. Figure C2).

10k. inhabitant) (cf. Table D1). That is, an increase in the total intensity of an epidemic wave at the urban area level will disproportionately affect poor municipalities of this urban area.

Immediately related is the question of the evolution of the urban-area level gradient over the two waves. Two hypotheses are compatible with current evidence of a gradient: first, it could be that poor areas are hit first and non-poor areas subsequently catch-up; second, it could be that poor areas are structurally more exposed to COVID-19. For the first time, our data allows to test these two hypotheses over a one-year and two-wave span. To perform such test, we simply consider the 243 (among 421) urban areas that suffered positive excess mortality in each of the two waves.²⁵ We find that, in these urban areas, the average income gradient in both wave is positive (cf. Table D2). This means that, in these urban areas, although residents of poorer municipalities suffered more in the first wave, they also did so in the second wave. In other word, over a one year period, we find no evidence of catch-up of the richer municipalities.

This longer-term perspective provide new evidence on the impact of COVID-19 on mortality inequalities: it shows that the income gradient identified in the first wave is a regular feature of the pandemic that is not compensated later in 2020.

5 Direct and indirect effects of the COVID-19 epidemic: evidence from a quasi-natural experiment

The previous section has established evidence of an income-related heterogeneity in the total impact of the pandemic on mortality in France. This income gradient can result from both the direct and indirect effects of COVID-19. According to Mulligan (2021), about 12% of March-to-October excess mortality in the US was due to non-COVID causes (a large part being the result of "deaths of despair"). A recent literature has documented that these indirect deaths due to economic and social changes related to the epidemic - including lockdown policies - have disproportionately affected the poor.²⁶ In this section, we take advantage of the quasi-natural experiment induced by the first lockdown to clear these potential indirect channels.

²⁵These represent 72% (96%) of urban areas (population) in the red zone and 43% (69%) in the green zone.

²⁶According to McIntyre and Lee (2020), unemployment in time of lockdown is expected to have a strong impact on suicide rates – and we know that poor individuals are more likely to have jobs that cannot be done remotely and were therefore more often laid off during the lockdown (Gottlieb et al., 2021; Palomino et al., 2020). Bullinger et al. (2020) have shown that stay-at-home orders increase domestic violence in poor households but not in households with above-median income. The strong variation in car crashes and in pollution levels due to the lockdown could also differently affect municipalities depending on their income (Brodeur et al., 2021). Drug overdose and alcohol abuse could also be the source of larger deaths of despair during the lockdown in poor areas.

5.1 Identification strategy

At the core of our strategy is the fact that the first lockdown was implemented uniformly over the country while the epidemic was at very heterogeneous stages of development across regions. In particular, based on an external classification provided by the government on May 7, we can distinguish regions strongly hit by the epidemic as early as mid-march (the red zone), from regions where the epidemic was at a much earlier stage of development when the first lockdown was decided (the green zone). Although this classification was only provided towards the end of the first lockdown, we indeed show in Appendix E that it strongly correlates with indicators of the spread of COVID-19 *before* lockdown. Importantly for what follows, we also show in this appendix that the sharp heterogeneity in the spread of the virus between the two zones in mid-March is largely unrelated to socio-economic factors potentially affecting inequalities in mortality between rich and poor municipalities and that excess mortality in both areas was very similar in 2019 (compared to 2018) and in early 2020 (compared to 2018 and 2019).

This quasi-experimental setting provides a unique opportunity to compare the evolution of excess mortality in two areas equally affected by lockdown restrictions²⁷ and sharing similar characteristics (health care system, other institutions, etc.) but unequally exposed to COVID-19. Our setting is comparing within-urban areas gradient. Our claim is that the evolution of the within-urban-area gradient would have been comparable (on average between green and red areas) absent COVID-19 and the associated public policies. In other words, we assume that the probability for an urban area to be exposed to COVID-19 is independent of the evolution of its poor vs. non-poor mortality differential that would have occurred absent COVID.

To disentangle the direct effects of COVID-19 on excess mortality from its indirect effects, we treat the green zone as a control group: we consider the income gradient in excess mortality found in this zone during wave 1 as a measure of the differentiated indirect impact (between poor and non-poor municipalities) of COVID-19. By contrast, the income gradient found in the red zone captures both the direct and the indirect effects of the epidemic. The difference between the two estimated gradients can thus be attributed to the direct impact of COVID-19.²⁸ In a sense, this approach is very similar to a triple-difference strategy where the difference across *zones* (first difference) in the wave-1-specific *evolution*

²⁷Figure B1 (Appendix B) supports the idea of a very uniform enforcement of the lockdown policy, since the increase in time spent at home compared to normal conditions is very similar across zones.

²⁸As mentioned in section 2, we define the indirect impact as deaths caused by public policies taken as a response to the epidemic and related socio-economic changes, as opposed to the direct impact which captures those directly related to COVID-19 infections (deaths + hospital congestion). Given that the epidemic was not totally absent from the green zone during the first wave, such approximation introduces errors that attenuate the magnitude of the direct effects and their corresponding income gradient.

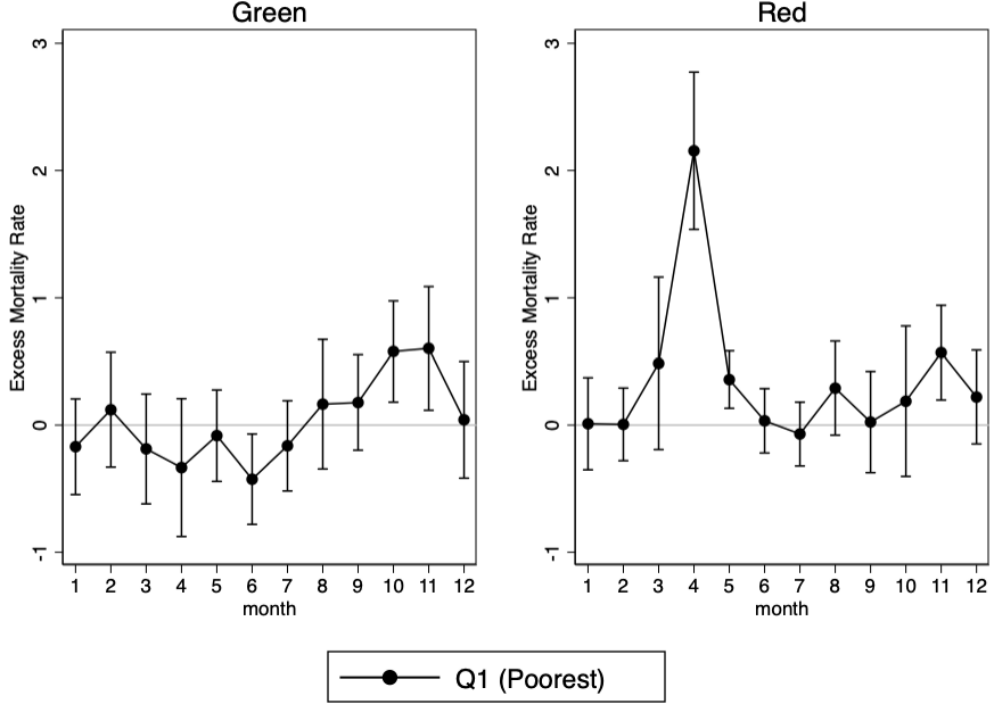
(second difference) of the mortality *gradient* (third difference) identifies the causal direct effect of the epidemic on poor municipalities' mortality rate. We assume here that, on average, within-urban-area difference between poor and non-poor municipalities in the indirect impact of COVID-19 on mortality is similar between the red and the green zones. Note that this is much less demanding than assuming that the indirect impact of the epidemic is the same across both zones.

5.2 Income gradient and the direct impact of COVID-19

Double-differences by zone

We first re-estimate Equation 2 for each zone in each month and we plot the coefficients of interest on Figure 4 (the zone-specific equivalent of Figure 3). It clearly appears that no gradient is found in the green zone during the first wave, unlike in the red zone and despite common lockdown restrictions. This leads us to conclude that there is no income gradient in the effect of such policies - that we defined as the indirect effect of the epidemic - on excess mortality. By contrast, the Figure shows a marked income gradient in excess mortality whenever the level of infection is high (*i.e.* wave 1 in the red zone; wave 2 in both zones). This suggests that the income gradient is only to be found in the direct effect of COVID-19 infections.

Figure 4: Excess mortality by income and zone



NOTE: The graph plots the point estimate and the 95% confidence intervals of the estimation of β from equation 2 evaluated each month on each zone separately. It accounts for the monthly difference in all-cause excess mortality between the poor municipalities and the rest in each zone, where poor is defined as belonging to the bottom quartile of the national distribution of municipal median income weighted by the municipality size. The red zone corresponds to the areas that were the most severely hit by the first wave, and that are located in the North-Eastern quarter of the country. This zone covers about 44% of the urban population of (mainland) France. The green zone encompasses the rest of the French territory.

A triple-difference strategy

To test more formally the existence of an income gradient in the direct effect of the epidemic, we exploit the quasi-natural experiment provided by the first lockdown by employing a triple-difference strategy that adds the red vs. green dimension (defined at the *département* level) to the double-difference setting used in section 4. Formally, we estimate the following model:

$$D_{[m,d]}^p = \beta.Q1_m + \delta.Red_d + \rho.Red_d.Q1_m + X_{[m,d]}^p.\Lambda + \gamma_{ua} + \nu_m^p \quad (3)$$

The main coefficient of interest ρ estimates the difference between red and green zones in the within urban-area difference between rich and poor municipalities' excess mortality. $X_{m,d}$ represents a vector

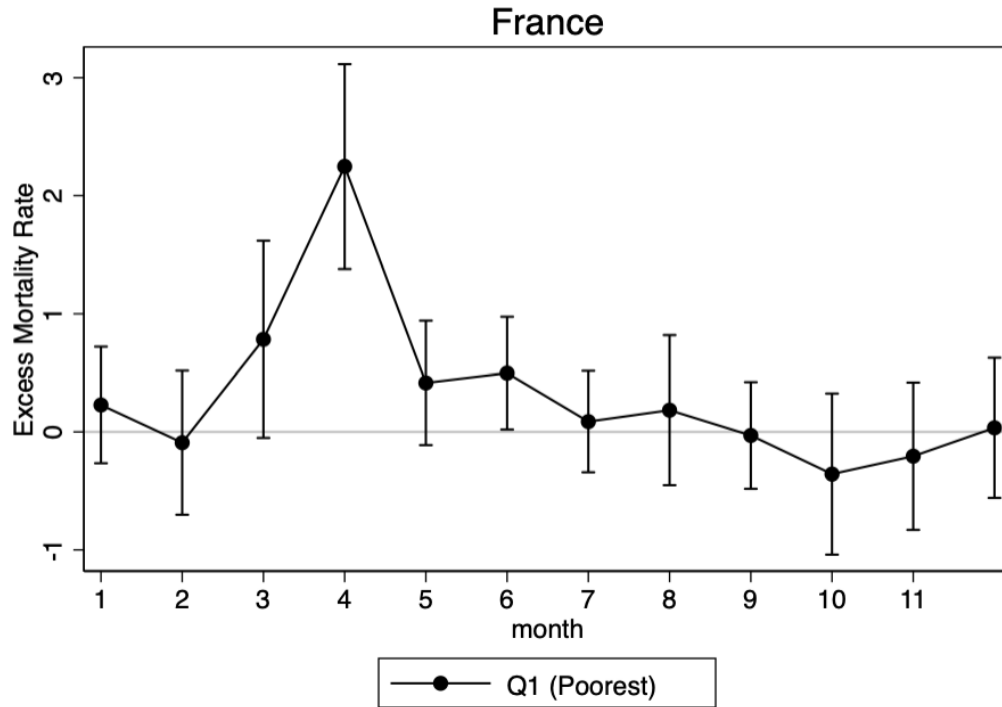
of controls and γ_{ua} the full set of urban-area fixed effects,²⁹ both defined as in Equation 2. ρ measures the causal effect of the pandemic cleared of national public policies (including the lockdown) on excess mortality under the sole hypothesis that, absent COVID-19 infections, the average difference in the evolution of mortality in a given month (2020 v.s. 2019 and 2018) between rich and poor municipalities of the same urban area would have been the same in red and green zones. Standard errors are clustered at the *département* level, the most aggregated level of treatment status in this setting.³⁰

In this setting the period of interest is that of March and April (first wave), and estimating Equation 3 both before the pandemic broke (January and February) and between waves (May-September) provides several tests of our identification assumption. Furthermore, evaluating the equation during the second wave (October-December), when no clear difference in mortality between the two zones stood out, ensures that both zones are comparable. All monthly coefficients are reported in Figure 5. Consistent with our hypothesis, the Figure shows no significant difference in excess mortality between zones outside of the first wave. This difference first increases in March, due to the early direct effects of the epidemic, but remains insignificant at a 5% level. Conversely, in April, at the peak of wave 1, the main coefficient is strongly significant and is even stronger than in the double difference.

²⁹Note that these urban-area fixed effects account for the risk that: (i) coincidentally, the differences in COVID-19 infection intensity across urban areas at a given point in time could be non-orthogonal to their income level; (ii) people may adapt their behaviour to the local level of infection which would bias our results (Almagro et al., 2020).

³⁰Note that clustering at the level of urban areas does not change the nature of our results.

Figure 5: Income gradient in the direct effect of COVID-19 on mortality



NOTE: The graph plots the point estimate and the 95% confidence intervals of the estimation of ρ from equation 3 evaluated each month. It accounts for the monthly difference in all-cause excess mortality between the poor municipalities and the rest in the red and in green zones, where poor is defined as belonging to the bottom quartile of the national distribution of municipal median income weighted by the municipality size.

Robustness

In Appendix F, we take advantage of the panel nature of our data to estimate a full triple-difference strategy using death toll as the main dependent variable. It allows us to include municipality fixed effects to control for any time-invariant factor that could influence mortality at the municipality level. The nature of the result remains: over March-April, mortality increased in 2020, more so in red than in green areas, and more so in poor than non-poor municipalities. This confirms that the most severely hit municipalities are those belonging to the poorest quartile and to the red zone. Results based on the triple-difference strategy are still robust to the exclusion of elderly care homes.³¹

³¹Results available upon request.

6 Potential mechanisms

In this section, we explore the role of two potential mechanisms that could explain inequalities in COVID-19 related mortality: labor-market exposure and housing conditions.³² We first describe how we measure these dimensions and confirm that each of our measure is positively correlated with poverty (section 6.1). We then investigate their independent effect on COVID-19 related mortality by exploring the association between each of these measures and excess mortality as well as the dynamics of these associations across the different stages of the epidemic (section 6.2). Finally, we try to quantify the extent to which these mechanisms explain the income gradient in COVID-19 related mortality and to understand which of the two prevails (section 6.3).

6.1 Housing and labor-market conditions: measurement and relation with poverty

Highly-exposed occupations are successively defined as occupations with frequent direct contact with the public in usual (pre-COVID-19) business conditions, and as occupations in sectors which kept operating during lockdown periods (the so-called *essential workers*). Informed by administrative data on the occupational distribution in each municipality, we compute (i) the worker-weighted average frequency of contact (hereafter “index of frequent contact”) and (ii) the share of essential workers, in every municipalities. Regarding housing conditions, we use census data at the household level to measure the share of overcrowded housing units in the municipality. This variable improves on some previous measures of household size used in the literature (Almagro and Orane-Hutchinson, 2020) since it is a function of household size, dwelling size and number of rooms. At the interaction between labor-market and housing dimensions, we finally compute the share of municipalities’ households that gather at least one member aged 65 or more and one member from a younger generation who is currently employed. This measure requires occupation information from the Census, which is only available for municipalities with at least 2,000 inhabitants. We provide more details on the construction of these measures in Appendix A2 and reference the sources in Appendix A1. Table B5 (Appendix B) shows descriptive statistics on the different mechanism variables.

Table G1 (Appendix B) shows the strength of the link between poverty and our (normalized) labor-market and housing measures, once included our baseline controls and urban-area fixed effects. Munic-

³²We consider these two channels informed by the current literature about how COVID-19 is transmitted (Almagro and Orane-Hutchinson, 2020; Almagro et al., 2020; Glaeser et al., 2020), and given their likely positive association with poverty. We of course acknowledge that these are only two potential mechanisms among many and that we notably ignore the role of comorbidities known to be related with poverty and that most likely played a role in the observed phenomenon (Wiemers et al., 2020; Raifman and Raifman, 2020).

ipalities of the poorest quartile have more occupations in contact with the public (column 1) and more essential workers (column 2). A one standard-deviation increase in the share of essential workers (respectively in the index of frequent contact) makes the probability to fall in the bottom quartile of the national weighted income distribution increase by 16pp (respectively 14pp). The association is even stronger with the housing-crowding variable (column 3): a one standard-deviation increase in the share of over-crowded housing amounts to a 29pp rise in the probability to be living in a poor municipality. Finally, poorest municipalities are also more likely to have multigenerational households (column 4), although the reduced sample size of municipalities with more than 2,000 inhabitants decreases the precision.³³

6.2 The effect of labor-market exposure and housing conditions on excess mortality

We now estimate the double-difference model exposed in Equation 2 but substituting (normalized) labor-market and housing-related measures in place of the poverty indicator. Table 2 reports the main coefficient of interest of the mechanism variables (*i.e.* the β of Equation 2) for each wave separately.

We first examine results for the first wave (March-April) in the upper part of Table 2: both labor-market indices and housing overcrowding exhibit significant effects. A one standard-deviation increase in the share of frequent-contact jobs leads to a rise of 0.69 deaths per 10K inhabitants on average in wave one, relative to an average level of 3.97. The respective figure for the share of essential workers is of the same order of magnitude: 0.65 units (+16%). The effect of overcrowded housing is even larger: increasing its share by one standard deviation increases excess mortality by 1.51 deaths per 10K inhabitants during the first wave (*i.e.* a 38% increase relative to the average excess mortality in wave one).³⁴ In Table 2, we observe that households including both elderly and employed individuals seem to foster the transmission of COVID-19 and its associated mortality. Although the coefficient is not significant in wave one,³⁵ living in a municipality with a one standard-deviation higher share of multigenerational households significantly increases excess mortality by 0.53 deaths per 10K inhabitants in wave two (*i.e.* a 10.5% increase relative to the average excess mortality).

³³The coefficient indicates that a one standard-deviation increase in the share of multigenerational households increases the probability to be in the poorest quartile by 7pp, but the p -value is only equal to 0.175.

³⁴Interestingly, we find a small positive but non-significant effect of municipal density on excess mortality rate. However, the link with poverty is not significant, and we therefore do not include density in our mechanism variables. For a deeper analysis of density, see [Gerritse \(2020\)](#). In contrast, the effect of over-crowded housing on excess mortality is five times as large. Consistent with [Almagro and Orane-Hutchinson \(2020\)](#); [Allain-Dupré et al. \(2020\)](#), it means that the density in the municipality matters less than the density at home to explain differences in COVID-19 related mortality.

³⁵The sample on which this effect is estimated is restricted to municipalities with more than 2,000 inhabitants. We reproduced our results for our main three variables on that reduced sample. All the point estimates are of the same magnitude as the ones shown in table 2, although standard errors are larger due to the smaller sample size. Results are available upon request.

Table 2: Effect of mechanism variables on COVID-19 related mortality

	Index of frequent contact	Share of essential workers	Share of over-crowded housing	Multigenerational hh
Excess mortality rate, wave 1				
Main coefficient	0.69281* (0.35777)	0.64594*** (0.14160)	1.51233*** (0.30694)	0.30218 (0.42948)
Urban areas FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Control outcome mean	3.971	3.971	4.001	4.480
Adjusted R2	0.1673	0.1679	0.1762	0.3161
Observations	16640	16640	16267	3989
Excess mortality rate, wave 2				
Main coefficient	0.45774** (0.19625)	0.88799*** (0.12777)	0.40256** (0.17475)	0.52526*** (0.16278)
Urban areas FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Control outcome mean	5.089	5.089	5.127	5.020
Adjusted R2	0.1228	0.1256	0.1268	0.2298
Observations	16640	16640	16267	3989

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors in parentheses are clustered at the urban-area level.

This table shows the result of regressing municipalities' excess mortality rate in each wave on mechanism variables measuring either housing conditions or occupational exposure. The upper part of the table shows results for the first wave (March-April) while the bottom part of the table shows results for the second wave (October-December) on municipalities in all urban areas. Each column reports the result of a separate regression examining one mechanism. The main coefficient corresponds to the effect of the mechanism variable on the excess mortality rate. The mechanism variables have been normalized such that coefficients can be interpreted in terms of the effect of a one standard-deviation change, and can be compared with each other. All regressions include urban areas fixed-effects and control for total population and for the share of inhabitants over 65 y.o. in the municipality. The outcome-mean line reports the mean of the excess mortality rate per 10K inhabitants in each wave.

Our data allows us to measure the dynamics of the housing crowding and occupational exposure in different contexts over the year 2020. Figures G1 to G4 (Appendix G) complement Table 2 by showing the monthly effect of each normalized mechanism variable on the normalized excess mortality rate, including population controls and urban-area fixed effects. Reassuringly, none of our mechanism variables has a significant effect on excess mortality outside of the two waves of the pandemic.³⁶ Their impact is quite comparable across both waves, except for the effect of the share of over-crowded housing which decreases in magnitude during the second wave, consistent with the less strict implementation of the lockdown policy (see section 3 for more details). We can also compare the coefficients found at the beginning and the end of each wave, where March and October capture mortality before the implementation of the lockdown, and April and November-December measure mortality under the lockdown.³⁷ While the effect of frequent social contacts remains constant throughout each wave, the coefficients on the share of essential

³⁶Except for one coefficient out of the 48 tested, which is the share of over-crowded housing in July.

³⁷Because of an approximate two-week lag between infection and mortality, November only partly captures the effect of the second lockdown. We also define the end of the second wave as December only, and it does not affect the results.

workers and the share of over-crowded housing significantly increase from the beginning to the end of wave one.³⁸ The growing importance of both housing conditions and the possibility of remote working is consistent with limited movement out of the house under the lockdown, except for essential workers who were the only ones to go to the workplace. We do not find this within-wave pattern, however, for wave two: indeed, although movement was more restricted in November-December, people were already spending most of their time at home in October, as remote working was already prevalent, which was not the case in March. In addition, as opposed to the first lockdown, schools remained open, which could both explain why the coefficient on over-crowded housing was higher during the first wave than during the second wave, but not at the end of the second wave compared to the beginning.³⁹ Based on [Google LLC \(2021\)](#) data, [Figure B1](#) describes two facts that are consistent with the dynamics of the housing and labor-market exposure: (i) time spent at home is higher during the first lockdown than during the second; (ii) time spent at home is higher at the end than at the beginning of each wave, with the exception of December where it decreases again.

Previous literature has already established a link between physical contact (at home or at work) or maintained activity in the workplace during lockdown and COVID-19 exposure ([Almagro and Orane-Hutchinson, 2020](#); [Almagro et al., 2020](#); [Naticchioni et al., 2020](#); [Angelucci et al., 2020](#)). We are able to provide evidence at a fine level both geographically – since we have administrative data at the municipal level while covering the whole national territory – and sector-wise – as the definition of essential workers is at the 3-digit occupation level.⁴⁰ We are also able to uncover the role of the interaction of labor-market and housing channels: workers increase their risk of catching the disease when going to the workplace, and transmit it to vulnerable persons when living in multigenerational households. Finally, we examine the dynamics of the effect over the whole year 2020 and under different lockdown conditions. [Almagro and Orane-Hutchinson \(2020\)](#) and [Almagro et al. \(2020\)](#) highlight a growing importance of housing crowding over time compared with commuting in explaining COVID-19 infection, that they justify by the fact that essential workers were gradually laid off over time. The different policies implemented and the lower

³⁸Although we do not have enough precision on the reduced sample where the effect of the share of multigenerational households is measured to draw any conclusion, we observe that the pattern on [Figure G4](#) is consistent with the evolution of the epidemic.

³⁹Several reasons may explain the lower coefficient in December: first, there is anecdotal evidence that compliance was lower during the second lockdown. Second, if the lockdown was lifted on December 15th, shops were allowed to reopen from the end of November already, in anticipation of Christmas shopping. Moreover, people were traveling throughout France during the Christmas break, giving less importance to housing conditions in the municipality of residence.

⁴⁰[Almagro et al. \(2020\)](#) exploit both occupation data from the American Community Survey and ZIP-code level and individual geolocation data to track time spent outside of home, but they focus on New York City only. [Angelucci et al. \(2020\)](#) use individual survey data in the US but they define essential workers based on the ability to work remotely before COVID-19.

incidence of layoffs in France compared with the US during the pandemic may explain why we do not find a similar pattern.

6.3 Disentangling the role of each mechanism

To understand which mechanism prevails, we perform a horse race between our mechanism variables. Table 3 reports the results for the first wave (top panel) and the second wave (bottom panel) in all urban areas. Column (1) gives the difference in excess mortality rate between poor and non-poor municipalities. While the frequent contact variable has a minor effect on the poverty coefficient, columns (2) to (4) show that the inclusion of the share of essential workers and overcrowded housing make the coefficient of poverty shrink. In the first wave, the initial coefficient of poverty diminishes by 36% when the share of essential worker is included, and by 97% when the share of over-crowded housing is included. The last column shows that the covariates altogether absorb all of the poverty effect, which is not significant anymore, while housing conditions and the share of essential workers still play a significant role in explaining excess mortality. Similarly, in wave 2, a high share of the poverty effect is absorbed by the mechanism variables (64%), but here the share of essential workers appears as the main channel of the income gradient. It therefore confirms the previous analysis on the dynamics of the epidemic (section 6.2).

It is reasonable to think that there is an overlap between the group of workers with frequent contacts in their job and essential workers (*e.g.* cashiers, bus drivers, etc.).⁴¹ The effect of having frequent contacts in the job drops when the essential worker variable is included during the first wave. It suggests that non-essential workers with frequent contact in normal business conditions do not suffer more from the pandemic during the first wave, since the lockdown probably prevents them from having such contacts. Conversely, being an essential worker alone increases the risk of COVID-19 related mortality, meaning that the increase in risk partly comes from commuting, in addition to social contact in the workplace. During the second wave, the inclusion of both occupational exposure variables even yields a negative coefficient on the contact variable. If we do not want to put too much emphasis on this result, one explanation could be that more protective measures have been taken for these over-exposed workers, maybe to the extent that they become more protected than the average worker.⁴² Essential workers, however, continue to be hit more severely by COVID-19 despite these protective measures because they are going to the workplace more frequently.

⁴¹The correlation between both variables is 0.4. Out of the ten occupation codes with the highest index of frequent contact, six include essential workers.

⁴²Think for example of a shop assistant, who must constantly wear a face mask and use hydro-alcoholic gel, compared to a clerical worker sharing his office and not constantly wearing a face mask.

Table 3: Horse-race between mechanism variables

Excess mortality rate, wave 1					
Poor	0.51148*	0.39867*	0.32887	0.01340	-0.09433
	(0.29159)	(0.22180)	(0.26785)	(0.13511)	(0.14033)
Index of frequent contact		0.57361**			-0.11880
		(0.27257)			(0.16897)
Share of essential workers			0.52998***		0.42456***
			(0.10411)		(0.12910)
Share of over-crowded housing				1.49710***	1.45176***
				(0.28282)	(0.30798)
Urban areas FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Control outcome mean	1.734	1.734	1.734	1.812	1.812
Adjusted R2	0.1671	0.1682	0.1684	0.1762	0.1768
Observations	16640	16640	16640	16267	16267
Excess mortality rate, wave 2					
Poor	0.47000***	0.40418***	0.18585	0.41422**	0.16954
	(0.15573)	(0.13958)	(0.13119)	(0.20952)	(0.17782)
Index of frequent contact		0.33469**			-0.36655*
		(0.16748)			(0.21816)
Share of essential workers			0.82467***		0.98745***
			(0.12304)		(0.14727)
Share of over-crowded housing				0.12562	0.05568
				(0.24124)	(0.22016)
Urban areas FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Control outcome mean	4.063	4.063	4.063	4.176	4.176
Adjusted R2	0.1232	0.1234	0.1256	0.1274	0.1299
Observations	16640	16640	16640	16267	16267

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors in parentheses clustered at the urban-area level.

This table shows the result of regressing municipalities' excess mortality on a variable measuring either poverty, housing conditions or occupational exposure. The upper part of the table shows results for the first wave (March-April) while the bottom part of the table shows results for the second wave (October-December) on municipalities in all urban areas. The first column only examines the poverty channel. Columns (2) to (4) respectively include one additional variable capturing either the occupation or housing mechanism. The last column includes both the poverty dummy and all the mechanism variables. All regressions include urban-area fixed-effects and control for total population and for the share of inhabitants over 65 y.o. in the municipality. The mechanism variables have been normalized such that coefficients can be interpreted in terms of the effect of a one standard-deviation change, and can be compared with each other. The outcome-mean line reports the mean of the excess mortality rate per 10K inhabitants in each wave.

All in all, Table 3 suggests that both housing conditions and being an essential worker are important determinants of the relatively higher excess mortality due to COVID-19 in poor municipalities. The relative importance of housing conditions diminishes over time, potentially due to a less strict lockdown

during the second wave, whereas remote work was prevalent throughout the period. The whole set of covariates explain about 13% to 18% of the overall variation in excess mortality according to the adjusted R^2 . The mechanism variables seem to explain most of the income gradient in excess mortality – as shown by the non-significant poverty coefficient in the last column – and even to partly have an independent effect on excess mortality – as shown by the increase in the adjusted R^2 from column 1 to 5.

7 Conclusion

In this paper, we provide clear evidence that COVID-19 contributes to increasing inequalities in mortality through an unequal impact across municipalities. We find that the epidemic caused 2.6 more deaths per 10K. inhabitants on average in the poorest municipalities in 2020, relative to a baseline of 8.7 in non-poor municipalities (*i.e.* a 30% higher effect in the poorest municipalities). Importantly, the income gradient measured in the first wave (March-April) persists in the second wave (October-December), even within cities strongly affected during the first wave. We further show that the policy responses to COVID-19 (including lockdown) do not play a significant role in this gradient. Finally, our analysis highlights the key mediating role of labour market and housing conditions, in line with the idea that ecological factors are important determinants of the spread of epidemics. More specifically, we find that labor-market exposure remains an important determinant of COVID-19 mortality across both waves, while the role of housing conditions decreases over time.

Our results show that the correlation between income and COVID-19 mortality highlighted in the recent literature does not only *mirror* existing health inequalities but that the baseline income gradient in mortality is rather *amplified* during each epidemic wave. Observing such an effect in both waves in a country with relatively egalitarian access to health care also reinforces the idea that the income gradient in mortality is a very robust feature of the epidemic. Potential future epidemic waves of COVID-19 or similar infectious diseases are likely to follow the same pattern and further amplify existing health inequalities.

8 Acknowledgements

We would like to thank Philippe Askénazy, Luc Behaghel, Simone Bertoli, Aline Bütikofer, Paul Dutronc-Postel, Gabrielle Fack, François Fontaine, Laurent Gobillon, Camille Hémet, Miren Lafourcade, Clément Malgouyres, Eric Maurin, Dominique Meurs, Alexandra Roulet, Marianne Tenand as well as participants of the following seminars for their valuable comments: Labor Chair seminar at PSE, Economic

Policy seminar at the Univ. of Lausanne, BM at NHH, Health and Inequality seminar at CBS, and ADRES 2020 conference (Strasbourg). We thank the *CASD* for providing us access to the data. We benefited from the Labor Chair at PSE, the LABEX OSE - ouvrir la Science Économique and Cepremap funding for data access.

References

- V. Abedi, O. Olulana, V. Avula, D. Chaudhary, A. Khan, S. Shahjouei, J. Li, and R. Zand. Racial, Economic, and Health Inequality and COVID-19 Infection in the United States. *Journal of Racial and Ethnic Health Disparities*, Sept. 2020. ISSN 2196-8837. doi: 10.1007/s40615-020-00833-4. URL <https://doi.org/10.1007/s40615-020-00833-4>.
- N. E. Adler and D. H. Rehkopf. Us disparities in health: descriptions, causes, and mechanisms. *Annu. Rev. Public Health*, 29:235–252, 2008.
- G. Alfani. Epidemics, inequality, and poverty in preindustrial and early industrial time. *Journal of Economic Literature*, 2021.
- D. Allain-Dupré, I. Chatry, V. Michalun, and A. Moisisio. The territorial impact of covid-19: Managing the crisis across levels of government, 2020.
- M. Almagro and A. Orane-Hutchinson. The determinants of the differential exposure to covid-19 in new york city and their evolution over time. *Covid Economics: Vetted and Real-Time Papers*, (13), 2020.
- M. Almagro, J. Coven, A. Gupta, and A. Orane-Hutchinson. Racial disparities in frontline workers and housing crowding during covid-19: Evidence from geolocation data. *Available at SSRN 3695249*, 2020.
- M. Angelucci, M. Angrisani, D. Bennett, A. Kapteyn, and S. G. Schaner. Remote Work and the Heterogeneous Impact of COVID-19 on Employment and Health. Technical Report 13620, Institute of Labor Economics (IZA), Aug. 2020. URL <https://ideas.repec.org/p/iza/izadps/dp13620.html>. Publication Title: IZA Discussion Papers.
- B. N. Ashraf. Socioeconomic conditions, government interventions and health outcomes during COVID-19. *Covid Economics*, 37:141–162, 2020.
- B. Balmford, J. D. Annan, J. C. Hargreaves, M. Altoè, and I. J. Bateman. Cross-country comparisons of covid-19: Policy, politics and the price of life. *Environmental and Resource Economics*, 76(4):525–551, 2020.
- A. Banerjee, L. Pasea, S. Harris, A. Gonzalez-Izquierdo, A. Torralbo, L. Shallcross, M. Noursadeghi, D. Pillay, N. Sebire, C. Holmes, et al. Estimating excess 1-year mortality associated with the covid-19 pandemic according to underlying conditions and age: a population-based cohort study. *The Lancet*, 2020.

- B. Beach, K. Clay, and M. H. Saavedra. The 1918 influenza pandemic and its lessons for covid-19. *Journal of Economic Literature*, 2021.
- T. Beaney, J. M. Clarke, V. Jain, A. K. Golestaneh, G. Lyons, D. Salman, and A. Majeed. Excess mortality: the gold standard in measuring the impact of covid-19 worldwide? *Journal of the Royal Society of Medicine*, 113(9):329–334, 2020.
- M. Belot, S. Choi, J. C. Jamison, N. W. Papageorge, E. Tripodi, and E. Van den Broek-Altenburg. Unequal consequences of covid 19 across age and income: representative evidence from six countries. 2020.
- N. Blanpain. L’espérance de vie par niveau de vie: chez les hommes, 13 ans d’écart entre les plus aisés et les plus modestes. 1687, 2018.
- G. J. Borjas. Demographic Determinants of Testing Incidence and COVID-19 Infections in New York City Neighborhoods. SSRN Scholarly Paper ID 3572329, Social Science Research Network, Rochester, NY, Apr. 2020. URL <https://papers.ssrn.com/abstract=3572329>.
- L. Böttcher, M. D’Orsogna, and T. Chou. Using excess deaths and testing statistics to improve estimates of covid-19 mortalities. *arXiv preprint arXiv:2101.03467*, 2021.
- A. Brodeur, N. Cook, and T. Wright. On the effects of COVID-19 safer-at-home policies on social distancing, car crashes and pollution. *Journal of Environmental Economics and Management*, 106:102427, Mar. 2021. ISSN 0095-0696. doi: 10.1016/j.jeem.2021.102427. URL <https://www.sciencedirect.com/science/article/pii/S0095069621000103>.
- C. S. Brown and M. Ravallion. Inequality and the Coronavirus: Socioeconomic Covariates of Behavioral Responses and Viral Outcomes Across US Counties. Technical Report w27549, National Bureau of Economic Research, July 2020. URL <https://www.nber.org/papers/w27549>.
- L. R. Bullinger, J. B. Carr, and A. Packham. COVID-19 and Crime: Effects of Stay-at-Home Orders on Domestic Violence. Technical Report w27667, National Bureau of Economic Research, Aug. 2020. URL <https://www.nber.org/papers/w27667>.
- A. Butikofer, R. Karadacic, and K. G. Salvanes. Income inequality and mortality: A norwegian perspective. *NHH Dept. of Economics Discussion Paper*, (04), 2021.

- R. J. C. Calderon-Anyosa and J. S. Kaufman. Impact of COVID-19 lockdown policy on homicide, suicide, and motor vehicle deaths in Peru. *Preventive Medicine*, 143:106331, Feb. 2021. ISSN 0091-7435. doi: 10.1016/j.ypmed.2020.106331. URL <https://www.sciencedirect.com/science/article/pii/S0091743520303625>.
- A. Calderón-Larrañaga, D. L. Vetrano, D. Rizzuto, T. Bellander, L. Fratiglioni, and S. Dekhtyar. High excess mortality in areas with young and socially vulnerable populations during the COVID-19 outbreak in Stockholm Region, Sweden. *BMJ Global Health*, 5(10):e003595, Oct. 2020. ISSN 2059-7908. doi: 10.1136/bmjgh-2020-003595. URL <https://gh.bmj.com/content/5/10/e003595>. Publisher: BMJ Specialist Journals Section: Original research.
- A. Case, D. Lubotsky, and C. Paxson. Economic status and health in childhood: The origins of the gradient. *American Economic Review*, 92(5):1308–1334, 2002.
- S. Caul. Deaths Involving Covid-19 by Local Area and Socioeconomic Deprivation: Deaths Occurring Between 1 March and 31 July 2020. *Statistical Bulletin*, 2020.
- J. T. Chen and N. Krieger. Revealing the Unequal Burden of COVID-19 by Income, Race/Ethnicity, and Household Crowding: US County Versus Zip Code Analyses. *Journal of Public Health Management and Practice*, 27(1):S43–S56, Jan. 2021. doi: 10.1097/PHH.0000000000001263.
- R. Chetty, M. Stepner, S. Abraham, S. Lin, B. Scuderi, N. Turner, A. Bergeron, and D. Cutler. The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA*, 315(16):1750–1766, Apr. 2016. ISSN 0098-7484. doi: 10.1001/jama.2016.4226. URL <https://doi.org/10.1001/jama.2016.4226>.
- M. Cole, C. Ozgen, and E. Strobl. Air pollution exposure and covid-19. Technical report, Institute of Labor Economics (IZA), 2020.
- N. Courtejoie and C.-L. Dubost. Parcours hospitalier des patients atteints de la covid-19 lors de la première vague de l'épidémie. *Les Dossiers de la DREES*, (67), Oct. 2020. URL <https://drees.solidarites-sante.gouv.fr/etudes-et-statistiques/publications/les-dossiers-de-la-drees/article/parcours-hospitalier-des-patients-atteints-de-la-covid-19-lors-de-la-premiere>.
- J. Currie and H. Schwandt. Mortality inequality: the good news from a county-level approach. *Journal of Economic Perspectives*, 30(2):29–52, 2016.

- J. Currie, H. Schwandt, and J. Thuilliez. Pauvreté, égalité, mortalité: mortality (in) equality in france and the united states. *Journal of Population Economics*, 33(1):197–231, 2020.
- D. Cutler, A. Deaton, and A. Lleras-Muney. The determinants of mortality. *Journal of economic perspectives*, 20(3):97–120, 2006.
- D. M. Cutler, A. Lleras-Muney, and T. Vogl. Socioeconomic status and health: Dimensions and mechanisms. In *The Oxford Handbook of Health Economics*. 2012.
- A. Decoster, T. Minten, and J. Spinnewijn. The income gradient in mortality during the covid-19 crisis: evidence from belgium. *FEB Research Report Department of Economics*, 2020.
- J.-F. Delfraissy, L. Atlani Duault, D. Benamouzig, L. Bouadma, S. Cauchemez, F. Chauvin, P. L. Druais, A. Fontanet, M.-A. Grard, A. Hoang, B. Lina, D. Malvy, and Y. Yazdanpanah. Note du conseil scientifique covid-19 du 26 octobre 2020 - une deuxième vague entraînant une situation sanitaire critique. *Ministère des solidarités et de la santé*, Oct. 2020. URL <https://www.vie-publique.fr/rapport/277008-note-du-conseil-scientifique-covid-19-du-26-octobre-2020-deuxieme-vague>.
- K. Desmet and R. Wacziarg. Understanding Spatial Variation in COVID-19 across the United States. Technical Report w27329, National Bureau of Economic Research, June 2020. URL <https://www.nber.org/papers/w27329>.
- S. Drefahl, M. Wallace, E. Mussino, S. Aradhya, M. Kolk, M. Brandén, B. Malmberg, and G. Andersson. A population-based cohort study of socio-demographic risk factors for COVID-19 deaths in Sweden. *Nature Communications*, 11(1):5097, Oct. 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-18926-3. URL <https://www.nature.com/articles/s41467-020-18926-3>. Number: 1 Publisher: Nature Publishing Group.
- S. Foucart and S. Horel. Dépistage du coronavirus : les raisons du fiasco français sur les tests. *Le Monde.fr*, Apr. 2020. URL https://www.lemonde.fr/planete/article/2020/04/24/nous-attendons-d-etre-contactes-par-l-ars-mais-il-ne-se-passe-rien-le-fiasco-des-tests-en-france_6037647_3244.html.
- A. Fouillet, I. Pontais, and C. Caserio-Schönemann. Excess all-cause mortality during the first wave of the covid-19 epidemic in france, march to may 2020. *Eurosurveillance*, 25(34):2001485, 2020.
- M. Gerritse. Cities and COVID-19 infections: Population density, transmission speeds and sheltering responses. *Covid Economics*, 37:1–26, 2020.

- E. L. Glaeser, C. S. Gorbach, and S. J. Redding. How much does covid-19 increase with mobility? evidence from new york and four other us cities. Technical report, National Bureau of Economic Research, 2020.
- Google LLC. Google covid-19 community mobility reports, 2021. Accessed 02/15/2021, <https://www.google.com/covid19/mobility/>.
- C. Gottlieb, J. Grobovšek, M. Poschke, and F. Saltiel. Working from home in developing countries. *European Economic Review*, 133:103679, Apr. 2021. ISSN 0014-2921. doi: 10.1016/j.euroecorev.2021.103679. URL <https://www.sciencedirect.com/science/article/pii/S0014292121000325>.
- J. Jung, J. Manley, and V. Shrestha. Coronavirus Infections and Deaths by Poverty Status: The Effects of Social Distancing. SSRN Scholarly Paper ID 3623430, Social Science Research Network, Rochester, NY, June 2020. URL <https://papers.ssrn.com/abstract=3623430>.
- S. J. Kim and W. Bostwick. Social Vulnerability and Racial Inequality in COVID-19 Deaths in Chicago. *Health Education & Behavior*, 47(4):509–513, Aug. 2020. ISSN 1090-1981. doi: 10.1177/1090198120929677. URL <https://doi.org/10.1177/1090198120929677>. Publisher: SAGE Publications Inc.
- C. R. Knittel and B. Ozaltun. What Does and Does Not Correlate with COVID-19 Death Rates. Technical Report w27391, National Bureau of Economic Research, June 2020. URL <https://www.nber.org/papers/w27391>.
- A. N. M. Kraay, P. Han, A. K. Kambhampati, M. E. Wikswo, S. A. Mirza, and B. A. Lopman. Impact of non-pharmaceutical interventions for SARS-CoV-2 on norovirus outbreaks: an analysis of outbreaks reported by 9 US States. *medRxiv*, page 2020.11.25.20237115, Nov. 2020. doi: 10.1101/2020.11.25.20237115. URL <https://www.medrxiv.org/content/10.1101/2020.11.25.20237115v1>. Publisher: Cold Spring Harbor Laboratory Press.
- N. Krieger, P. D. Waterman, and J. T. Chen. COVID-19 and Overall Mortality Inequities in the Surge in Death Rates by Zip Code Characteristics: Massachusetts, January 1 to May 19, 2020. *American Journal of Public Health*, 110(12):1850–1852, Oct. 2020. ISSN 0090-0036. doi: 10.2105/AJPH.2020.305913. URL <https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2020.305913>. Publisher: American Public Health Association.
- S. Kung, M. Doppen, M. Black, I. Braithwaite, C. Kearns, M. Weatherall, R. Beasley, and N. Kearns. Underestimation of covid-19 mortality during the pandemic. *ERJ Open Research*, 2020.

- H. Le Bras. L'épidémie et son terrain social, June 2020. URL <https://jean-jaures.org/nos-productions/1-epidemie-et-son-terrain-social>. Library Catalog: jean-jaures.org.
- F. B. LeClere, R. G. Rogers, and K. D. Peters. Ethnicity and mortality in the united states: individual and community correlates. *Social forces*, 76(1):169–198, 1997.
- J. P. Mackenbach, J. R. Valverde, M. Bopp, H. Brønnum-Hansen, G. Costa, P. Deboosere, R. Kalediene, K. Kovács, M. Leinsalu, P. Martikainen, et al. Progress against inequalities in mortality: register-based study of 15 european countries between 1990 and 2015. *European journal of epidemiology*, 34(12):1131–1142, 2019.
- C. Mangeney, N. Bouscaren, M. Telle-Lamberton, A. Saunal, and V. Féron. La surmortalité durant l'épidémie de covid-19 dans les départements franciliens. Technical report, Paris Region Health Observatory, 2020. URL <https://www.ors-idf.org/nos-travaux/publications/la-surmortalite-durant-lepidemie-de-covid-19-dans-les-departements-franciliens.html>.
- R. S. McIntyre and Y. Lee. Projected increases in suicide in Canada as a consequence of COVID-19. *Psychiatry Research*, 290:113104, Aug. 2020. ISSN 0165-1781. doi: 10.1016/j.psychres.2020.113104. URL <https://www.sciencedirect.com/science/article/pii/S0165178120310386>.
- C. B. Mulligan. Deaths of despair and the incidence of excess mortality in 2020. Technical report, National Bureau of Economic Research, 2021.
- P. Naticchioni, V. Scrutinio, E. di Porto, et al. Partial lockdown and the spread of covid-19: lessons from the italian case. Technical report, Centre for Economic Performance, LSE, 2020.
- J. C. Palomino, J. G. Rodríguez, and R. Sebastian. Wage inequality and poverty effects of lockdown and social distancing in Europe. *European Economic Review*, 129:103564, Oct. 2020. ISSN 0014-2921. doi: 10.1016/j.euroecorev.2020.103564. URL <https://www.sciencedirect.com/science/article/pii/S001429212030194X>.
- N. W. Papageorge, M. V. Zahn, M. Belot, E. van den Broek-Altenburg, S. Choi, J. C. Jamison, E. Tripodi, et al. Socio-demographic factors associated with self-protecting behavior during the covid-19 pandemic. Technical report, Institute of Labor Economics (IZA), 2020.
- S. Papon and I. Robert-Bobée. Une hausse des décès deux fois plus forte pour les personnes nées à l'étranger que pour celles nées en france en mars-avril 2020. Technical report, French National Statistical Institute (INSEE), 2020. URL <https://www.insee.fr/fr/statistiques/4627049>.

- C. Persico and K. R. Johnson. The effects of increased pollution on covid-19 cases and deaths. *Available at SSRN 3633446*, 2020.
- K. E. Pickett and M. Pearl. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *Journal of Epidemiology & Community Health*, 55(2):111–122, 2001.
- M. A. Raifman and J. R. Raifman. Disparities in the Population at Risk of Severe Illness From COVID-19 by Race/Ethnicity and Income. *American Journal of Preventive Medicine*, 59(1):137–139, July 2020. ISSN 0749-3797. doi: 10.1016/j.amepre.2020.04.003. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7183932/>.
- R. Rivera, J. E. Rosenbaum, and W. Quispe. Excess mortality in the united states during the first three months of the covid-19 pandemic. *Epidemiology & Infection*, 148, 2020.
- S. A. Robert. Socioeconomic position and health: the independent contribution of community socioeconomic context. *Annual review of sociology*, 25(1):489–516, 1999.
- F. Sa. Socioeconomic determinants of covid-19 infections and mortality: Evidence from england and wales. Technical report, Institute of Labor Economics (IZA), 2020.
- S. Scarpetta, M. Pearson, F. Colombo, and F. Guanais. Testing for covid-19: A way to lift confinement restrictions. *OECD Policy Responses to Coronavirus (COVID-19)*, May 2020. URL <https://www.oecd.org/coronavirus/policy-responses/testing-for-covid-19-a-way-to-lift-confinement-restrictions-89756248/#boxsection-d1e31>.
- J. D. Silverman, N. Hupert, and A. D. Washburne. Using influenza surveillance networks to estimate state-specific prevalence of sars-cov-2 in the united states. *Science translational medicine*, 2020.
- F. M. Snowden. *Epidemics and society: from the black death to the present*. Yale University Press, 2019.
- E. E. Wiemers, S. Abrahams, M. AlFakhri, V. J. Hotz, R. F. Schoeni, and J. A. Seltzer. Disparities in Vulnerability to Severe Complications from COVID-19 in the United States. *medRxiv*, page 2020.05.28.20115899, June 2020. doi: 10.1101/2020.05.28.20115899. URL <https://www.medrxiv.org/content/10.1101/2020.05.28.20115899v3>. Publisher: Cold Spring Harbor Laboratory Press.
- E. J. Williamson, A. J. Walker, K. Bhaskaran, S. Bacon, C. Bates, C. E. Morton, H. J. Curtis, A. Mehrkar, D. Evans, P. Inglesby, J. Cockburn, H. I. McDonald, B. MacKenna, L. Tomlinson, I. J. Douglas, C. T.

Rentsch, R. Mathur, A. Y. S. Wong, R. Grieve, D. Harrison, H. Forbes, A. Schultze, R. Croker, J. Parry, F. Hester, S. Harper, R. Perera, S. J. W. Evans, L. Smeeth, and B. Goldacre. Factors associated with COVID-19-related death using OpenSAFELY. *Nature*, 584(7821):430–436, Aug. 2020. ISSN 1476-4687. doi: 10.1038/s41586-020-2521-4. URL <https://www.nature.com/articles/s41586-020-2521-4>. Number: 7821 Publisher: Nature Publishing Group.

J. Wu, A. McCann, J. Katz, E. Peltier, and K. Deep Singh. The pandemic’s hidden toll: Half a million deaths. *New York Times*, Jan. 2021. URL <https://www.nytimes.com/interactive/2020/04/21/world/coronavirus-missing-deaths.html>.

T. Yorifuji, N. Matsumoto, and S. Takao. Excess all-cause mortality during the covid-19 outbreak in japan. *Journal of Epidemiology*, 31(1):90–92, 2021. doi: 10.2188/jea.JE20200492.

LIST OF APPENDICES

- A Data Sources and measurement
- B Descriptive statistics
- C Income gradient: robustness and falsification
- D Urban-Area level analysis
- E Red and green zones
- F Triple-difference strategy
- G Mechanisms

Appendix A Data Sources and measurement

A1 Data sources

by alphabetical order

DADS (for “Déclarations Annuelles des Données Sociales”) is a matched employer-employee exhaustive data set that covers all employees in private and public sectors (outside of agriculture) in the year 2016. The DADS is an administrative data set that all employers must report yearly to social security authorities and tax administration. The version available to researchers is provided by the INSEE. In particular, the DADS contain information about all positions occupied by any worker in a specific year, with start and end date at the daily level. For each position held by one individual we observe: the occupation, the firm and sector, both locations of work and residency and the number of hours spent. When using DADS to build municipality-level indicators we rely on the municipality of residency.

Daily Deaths Files (“Fichier des Décès Quotidiens”) from INSEE record each death that occurred between January 1, 2018 and December 31, 2020. For each death, we observe information relative to the event (the date and municipality of death as well as a category about the place of death - hospital or clinic, home, care home, etc.) and the individual (municipality of residency, sex, date of birth). INSEE made these files available at a higher frequency during the Covid-19 crisis, but this comes at the expense of some quality checks. The files are originally recorded by municipalities and then gradually added to the INSEE data sets. In theory, municipalities have about one week to inform INSEE about new deaths. In practice, one cannot fully exclude that the files were incomplete by the time of our analysis (especially for the most recent months) although we updated them frequently (last update: February 5th).

URL: <https://www.insee.fr/fr/statistiques/4487988>

DEFIS (“Dispositif d’Enquêtes sur les Formations et Itinéraires des Salariés”) is a database produced by the Céreq. It comes from a survey led on 4,500 firms and 16,000 of their workers as of 2013. Firms with more than 10 workers are representative of the whole private sector, but smaller ones are only representative for some sub-sectors. In this paper, we use the survey led on workers in 2015. The question of interest relates to the frequency of on-the-job contacts with the public. It refers to the current job of workers who have not changed firm since 2013, and to the one of 2013 for workers who changed firm since then. URL: <https://www.cereq.fr/en/data-access-lifelong-learning-and-vocational-training-surveys-defis-cvts-defis-employee>

Filosophi database (a French acronym for “Localised Disposable Income System”), 2014 version, contains a series of local measures on income and poverty. INSEE uses fiscal and social benefits data to compute these indicators at the municipality level. In particular, we are interested in the median standard of living of each municipality’s inhabitants. The concept of *standard of living* corresponds to the household disposable income divided by consumption units (using the OECD scale), to account for the size of the household. URL: <https://www.insee.fr/fr/metadonnees/source/operation/s1451/presentation>

Population Censuses are used to compute a series of measures at the municipality level. Each year: (i) a fifth of municipalities with less than 10,000 inhabitants are covered by an exhaustive census and (ii) a sample of 8% of the population of municipalities with 10,000 or more inhabitants are surveyed. In-depth questions relative to economic activity and family structure of households are only available for about 20% of the households of municipalities under 10,000 inhabitants and about 40% of the households in the other municipalities (10,000 inhabitants or more). Analyses of in-depth questions cannot be done for municipalities with less than 2,000 inhabitants as samples get too small. We use INSEE annual population counts and both exhaustive and partial census data files.

URL: <https://www.insee.fr/fr/information/2383265>

Urban Areas are defined by the INSEE as a group of neighboring municipalities encompassing an urban centre and its periphery made of municipalities where at least 40% of employed inhabitants works in the urban centre, as of 2010. Big urban areas represent 85% of the French mainland population in 2014.

A2 Measurement

A2.1 Poor municipalities

We assign each municipality its median income as computed by INSEE from fiscal and social benefits data. We next order each municipality from the poorest to the richest.

Next, we classify municipalities into bins of equal population-weighted size. Put differently, we divide the French urban population into equal groups based on their municipality of residency. In our favourite specification we consider four quartiles, from $Q1$ the poorest to $Q4$ the richest. Because municipalities have different sizes, each group is made of a different number of municipalities. This weighting follows our empirical specifications.

A2.2 Red and green zones

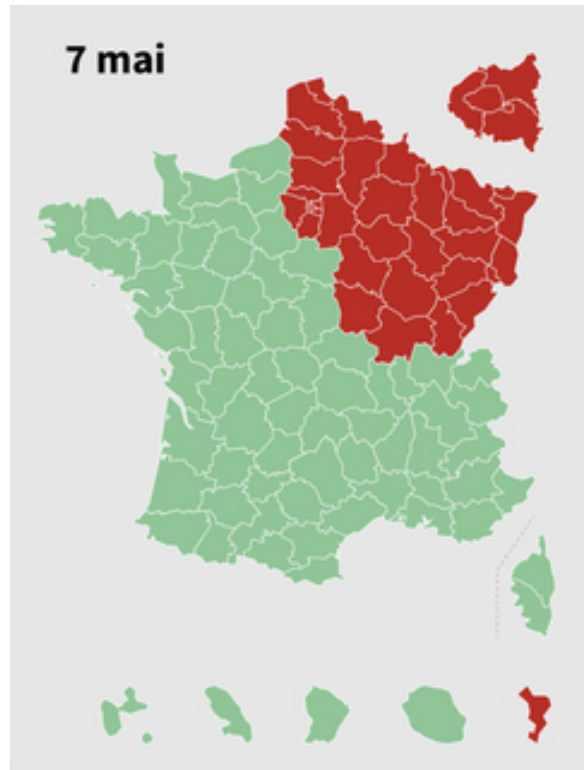


Figure A1: High- (red) and low-infection (green) *Départements* as of May 7, 2020

Note: Map of *Départements* issued by the Government on May 7, 2020. See [source](#).

When deciding the gradual lifting of the lockdown, the Ministry of Public Health provided a map differentiating red (high-infection) and green (low-infection) *départements* on May 7. A *département* is an administrative unit in France that encompasses on average 645,000 inhabitants. There is a total of 101 *départements* in France, 94 of which are on continental Europe (we call “mainland”). This map distinguishes high- and low-infection areas on the basis of: (i) the share of ER visits for a COVID-19 suspicion; (ii) the occupancy rate of intensive care beds relative to the initial one; (iii) the testing rate relative to the needs. High-infection *départements* correspond to 44% of the urban population and to roughly 30% of mainland France *départements*, all in the North-Eastern quarter of the country.

Although published in May, we use this distinction as a proximate measure of COVID-19 spread intensity before and during lockdown. We discuss this aspect in section 5.1 and appendix E.

A2.3 Labor Market Conditions

One novelty of this paper is to use administrative data to test whether labor market conditions explain the differential in excess mortality between poor and rich municipalities. We focus on two main characteristics of jobs that we expect to be key vectors for spreading the disease: (i) whether the worker kept on going to her workplace during the lockdown; (ii) whether the job involves frequent contacts with the public in usual business conditions (*i.e.* as measured before the lockdown).

As for the first dimension, for each municipality, we compute the share of essential workers⁴³ based on individual employment data from the DADS and on a list of essential occupations built by the Paris Region Health Observatory (Mangeney et al., 2020).⁴⁴ Note that, although this list was built by an administrative organisation, it remains arbitrary by nature. However, to our knowledge, there is no better way at the moment to characterise workers that kept on going to their workplace during the lockdown.⁴⁵

As for the second dimension, our proxy is based on the question “In your job, are you in direct contact with the public” that is available in a survey called DEFIS (see Appendix A1). For each occupation code (at the 3-digit level), we compute the share of workers answering “Often” (v.s. “Sometimes” or “Never”). Using the DADS, we then compute the average of this index at the municipality level based on the occupation distribution of workers living in each municipality.

A2.4 Housing Conditions

Our last dimension of interest is the relation between housing conditions and excess mortality due to COVID-19. We use the average share of housing units that are overcrowded provided by INSEE at the municipality level. Overcrowded accommodations are those that have less than “one living room, one room for each couple, one room for each other adult aged 19 or older, one room for two children if they are of the same sex or are under 7 years of age, and one room per child otherwise”.

A2.5 Interaction between Labor Market and Housing Conditions

To investigate the interactions between both mechanisms, we finally build an index at the municipality level, based on the partial census data files available for municipalities with more than 2,000 inhabitants.

⁴³Workers are localised according to their municipality of residency and regardless of their municipality of work.

⁴⁴Essential workers include: health workers, auxiliary nurses, pharmacists, ambulance drivers, post office clerks, the police, public transport and funeral services, firefighters, persons working in the sale of food products, delivery workers, tobacconists and cleaning staff. On average, they include about 19.5% of workers in each municipality. The French National Statistical Institute also used this list in one of their paper (Papon and Robert-Bobée, 2020).

⁴⁵We tried an alternative approach based on sectors that remained active during the lockdown. Our results got very noisy as we could not systematically identify firms that remained open and workers that kept on going to their workplace at a disaggregated level.

We compute the share of households with both an elderly person (over 65 y.o.) and a worker that is younger by at least 18 years (“multi-generational households” hereafter). This variable is meant to capture the fact that having different generations living in the same apartment increases the likelihood to get infected, and to potentially die for the elderly. This likelihood increases further if at least one of the younger members of the household works, as it potentially multiplies her social interactions.

Appendix B Descriptive statistics

B1 Sample

Table B1: Descriptive statistics: Urban areas

	Urban Areas				
	mean	sd	min	p50	max
Nb. Municipalities	27	83	2	7	1738
Population (10k)	8.5	53	.25	1.4	1247
Poor (%)	21	25	0	13	100
Observations	621				

Descriptive statistics across the 621 urban areas (UA). On average there are 27 municipalities in an UA, with an average total population of 85,000 inhabitants. On average, within an UA, 21% of municipalities are classified as poor.

Table B2: Descriptive statistics: Municipalities

	Municipalities				
	mean	sd	min	p50	max
Mort. rate 2018 (10k)	84	62	0	74	1545
Mort. rate 2019 (10k)	84	63	0	74	1254
Mort. rate 2020 (10k)	90	67	0	79	1691
Population	3188	11881	65	784	466351
Share 65y.o.(%)	17	6.6	0	17	63
Median Income (1k)	22	3	12	21	46
Observations	16640				

Descriptive statistics across the 16,640 municipalities. The average mortality rate (per 10,000 inhabitants) was 84 in 2018 and 2019 and rose to 90 in 2020. The average population of municipalities is 3,188, of which 17% are aged 65 or older. Municipalities' median income is, on average, 22,000 euro.

Table B3: Descriptive statistics: Urban areas, by zone

	mean	sd	min	p50	max	count
<i>Red zone</i>						
Nb. Municipalities	39	133	2	8	1738	191
Population (10k)	12	91	.25	1.4	1247	191
Poor (%)	25	26	0	18	100	191
<i>Green zone</i>						
Nb. Municipalities	22	46	2	6	505	430
Population (10k)	6.8	20	.25	1.3	225	430
Poor (%)	20	24	0	10	100	430

Descriptive statistics across the urban areas (UA) of red (191 UA) and green (430 UA) zones.

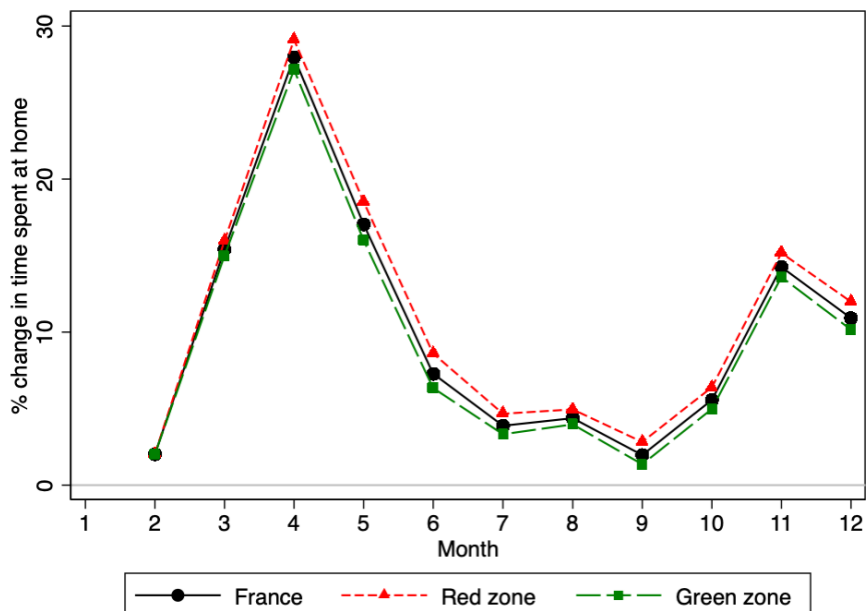
Table B4: Descriptive statistics: Municipalities, by zone

	mean	sd	min	p50	max	count
<i>Red zone</i>						
Mort. rate 2018 (10k)	81	65	0	72	1545	7000
Mort. rate 2019 (10k)	81	65	0	72	1254	7000
Mort. rate 2020 (10k)	89	73	0	79	1691	7000
Population	3330	12154	65	606	276219	7000
Share 65y.o.(%)	16	6.2	0	16	51	7000
Median Income (1k)	22	3.4	13	22	43	7000
<i>Green zone</i>						
Mort. rate 2018 (10k)	85	59	0	75	915	9640
Mort. rate 2019 (10k)	86	61	0	76	1085	9640
Mort. rate 2020 (10k)	90	63	0	79	949	9640
Population	3085	11678	80	931	466351	9640
Share 65y.o.(%)	18	6.7	0	17	63	9640
Median Income (1k)	21	2.7	12	21	46	9640

Descriptive statistics across the 16,640 municipalities divided into red (7,000 municipalities) and green (9,640 municipalities) zones.

B2 Lockdown

Figure B1: Percentage change in time spent at the place of residence in 2020



NOTE: Based on Google mobility data (Google LLC, 2021), the graph plots the percentage change in the time spent at the place of residence relative to a baseline measured between January 3, 2020 and February 6, 2020. For more information on the methodology, please refer to this [Google webpage](#). The “red” zone corresponds to the areas that were the most severely hit by the first wave, and that are located in the North-Eastern quarter of the country. This zone covers about 41% of the urban population of (mainland) France. The “green” zone encompasses the rest of the French territory.

B3 Mechanisms

Table B5: Summary statistics of mechanism variables

	France		Poor		Non-Poor	
	mean	sd	mean	sd	mean	sd
Share of over-crowded housing	0.0554	0.0559	0.0778	0.0716	0.0479	0.0472
Index of frequent contact	0.5999	0.0294	0.6077	0.0285	0.5972	0.0293
Share of essential workers	0.2063	0.0522	0.2375	0.0392	0.1959	0.0519
Share of multi-generational households	0.0225	0.0104	0.0240	0.0116	0.0218	0.0097
Observations	16640		1851		14789	

This table describes the different mechanism variables for the whole country and by poverty status. Being poor means living in a municipality belonging to the bottom 25% of the national income distribution weighted by population size.

Appendix C Income gradient: robustness and falsification

C1 Income gradient: four quartiles

C1.1 Empirical approach

In this appendix we take a more comprehensive approach and contrast the evolution of excess mortality rate in all four quartiles of municipalities' median income, taking the richest ($Q4$) as the reference. To do so, we estimate the following model:

$$D_{[m,ua]}^p = \beta_1 \cdot Q1_{[m,ua]} + \beta_2 \cdot Q2_{[m,ua]} + \beta_3 \cdot Q3_{[m,ua]} + X_{[m,ua]}^p \cdot \Lambda + \gamma_{ua} + \nu_{[m,ua]}^p \quad (4)$$

Where all definitions (concept, subscripts, controls and standard errors) are the same as in equation 2. The coefficients of interest are β_1 , β_2 and β_3 . They estimate the difference in excess mortality rate between richest municipalities ($Q4$) and each of the three other quartiles ($Q1$, $Q2$ and $Q3$), respectively.

The identification assumption remains: the three coefficients identify the heterogeneity in the causal (total) effect of the pandemic on excess mortality under the hypothesis that, absent COVID-19 and the associated public policies, the average difference in the evolution of mortality in p (2020 vs. before) between rich and poor municipalities of the same urban area would have remained stable.

C1.2 Main results

Table C1 reports the estimates of coefficients β_1 , β_2 and β_3 from equation 4. It is the equivalent of Table 1. On average, within a given urban area, and once population size and age are controlled for, municipalities of the poorest quartile had an excess mortality rate of 4.216 (deaths per 10k. inhabitants) higher than the richest municipalities. This has to be compared with the baseline average of 10.06 across municipalities of the richest quartile. Excess mortality is also significantly higher in municipalities of the two other quartiles than in municipalities of the richest quartile, but for both $Q2$ and $Q3$ the difference with $Q4$ is half smaller than the one observed between $Q1$ and $Q4$.

Columns (2) and (3) consider the first (*i.e.* $p = (3, 4)$) and second wave (*i.e.* $p = (10, 11, 12)$), respectively. The income gradient is striking in both waves and of very similar size. We observe no gradient when we focus on excess mortality outside of these two waves (column (4)).

Table C1: Excess mortality rate by municipality income

	(1) 2020	(2) Wave 1	(3) Wave 2	(4) No wave
Q1 (poor)	4.216*** (1.157)	1.986** (0.796)	1.975*** (0.416)	0.255 (0.347)
Q2	2.280** (0.926)	1.170** (0.484)	1.476*** (0.406)	-0.365 (0.380)
Q3	2.229*** (0.585)	1.122*** (0.339)	1.041** (0.440)	0.0661 (0.328)
Controls	✓	✓	✓	✓
Urban area FE.	✓	✓	✓	✓
Q4 average	10.06	5.482	4.075	0.503
Observations	16640	16640	16640	16640

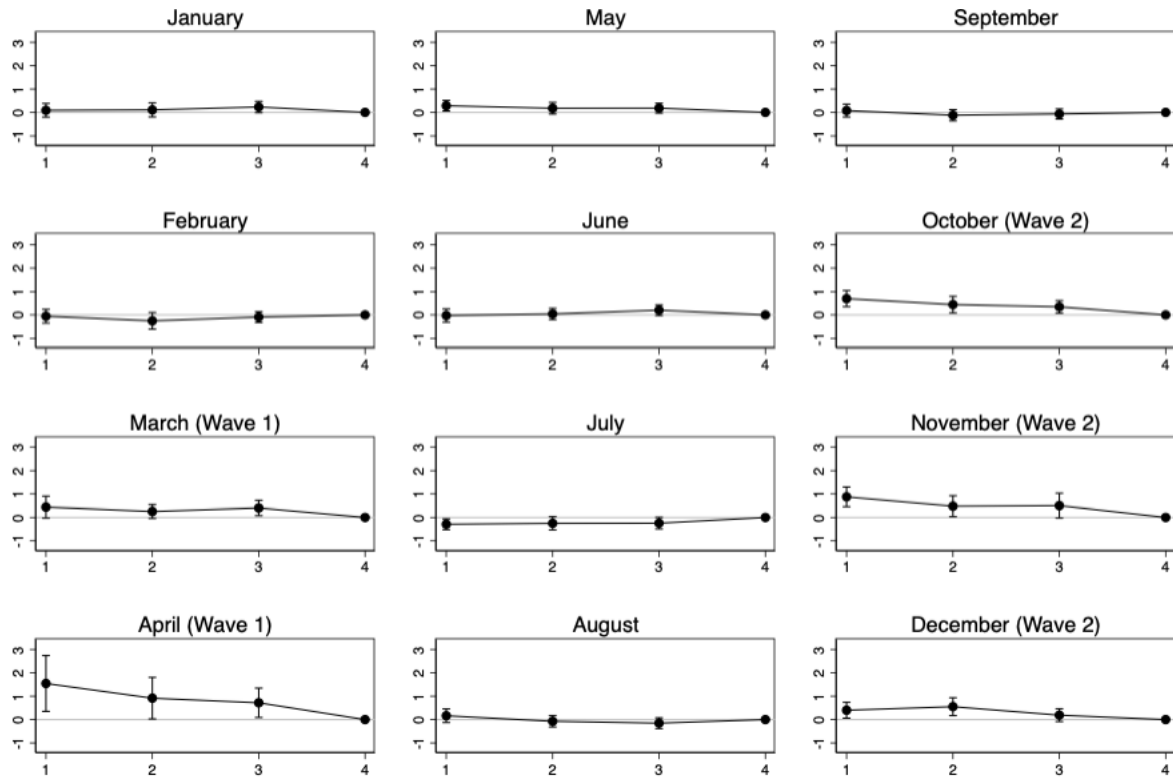
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors in parentheses clustered at the urban-area level.

NOTE: This table reports the coefficients associated with equation 4. The independent variable, excess mortality rate, is computed considering four different time periods: the whole year (column 1), wave 1 (March to April, column 2), wave 2 (October to December, column 3) and other months in 2020 outside the two waves (January, February and from May to September, column 4). By construction, column 1 is the sum of column 2 to 4. Controls include total population size and the share of the population over 65 years old.

C1.3 Monthly estimates

We also estimate equation 4 for each month separately. The dynamics of the gradient is strikingly correlated with the general dynamics of excess mortality (Figure C1).

Figure C1: Gradient in each month



NOTE: This Figure reproduces the coefficients associated with equation 4 estimated for each of the 12 months of 2020 separately. Given the dynamics presented in section 3.1 we classified March and April as “wave 1” and October, November and December as “wave 2”. Excess mortality rate was significantly higher in $Q1$ than in $Q4$ in April, May, October, November and December 2020. It is significantly lower in July.

C1.4 Alternative grouping

The monotonic gradient can also be seen when using alternative grouping in municipalities instead of quartiles. In table C2 below we estimate two alternative specifications of equation 4. Columns (1) to (4) reproduce table C1 using 10 deciles of income (instead of four quartiles) to group municipalities. Similarly, columns (5) to (8) repeat the exercise now comparing municipalities below (“M1” in the Table) and above (reference point) the median.

Table C2: Excess mortality rate by municipality income

	(1) 2020	(2) Wave 1	(3) Wave 2	(4) No wave	(5) 2020	(6) Wave 1	(7) Wave 2	(8) No wave
Q1 (poor)	6.822*** (1.402)	3.235** (1.283)	3.109*** (0.557)	0.478 (0.444)				
Q2	5.543*** (1.036)	2.589*** (0.444)	2.030*** (0.520)	0.924 (0.587)				
Q3	3.629*** (0.913)	1.960*** (0.397)	1.932*** (0.499)	-0.263 (0.551)				
Q4	4.343*** (1.096)	2.080*** (0.550)	2.155*** (0.516)	0.108 (0.491)				
Q5	3.439*** (1.119)	1.964*** (0.412)	1.354** (0.552)	0.121 (0.603)				
Q6	4.052*** (0.647)	1.994*** (0.272)	1.714*** (0.497)	0.344 (0.428)				
Q7	3.647*** (1.010)	1.865*** (0.361)	1.500** (0.676)	0.281 (0.565)				
Q8	3.277*** (0.760)	1.795*** (0.350)	0.677* (0.377)	0.805 (0.510)				
Q9	1.818*** (0.450)	0.970*** (0.314)	0.553* (0.331)	0.294 (0.323)				
M1 (poor)					1.923* (1.072)	0.916 (0.627)	1.119*** (0.397)	-0.112 (0.253)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Urban area FE.	✓	✓	✓	✓	✓	✓	✓	✓
Richest average	10.81	6.848	3.654	0.310	9.082	4.371	4.362	.
Observations	16640	16640	16640	16640	16640	16640	16640	16640

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors in parentheses clustered at the urban-area level.

NOTE: This table reports the coefficients associated to equation 4. The independent variable, excess mortality rate, is computed considering four different time periods: the whole year (column 1), wave 1 (Mars to April, column 2), wave 2 (October to December, column 3) and other months in 2020 outside the two waves (January, February and from May to September, column 4). By construction, column 1 is the sum of column 2 to 4.

C1.5 Other results

The results when excluding elderly care homes or when considering our main falsification test are also in accordance with the hypothesis of this paper. Results are available upon request.

C2 Robustness checks

C2.1 Excluding elderly care homes

Elderly care homes (EHPADs in the French acronym) were strongly hit by the COVID-19 epidemics. If EHPADs are not distributed homogeneously across municipalities, they could drive the difference between income groups (in one way or another). To exclude the threat of this confounding factor, we exclude deaths recorded from such institutions from our sample and re-estimate our model of equation 4. Table C3 displays the results in the same format as Table 1. Although lower in numbers, the main conclusion remains: an income gradient appears during epidemic waves.

The gradient in wave 1 is both large (+1.021 per 10k inhabitants) and insignificant. As explained in the core of the paper, this is largely due to the unequal spread of COVID-19 when the first lockdown started. If one considers only the red zone, the estimated coefficient is 1.945 ($se = 0.644^{***}$). We therefore conclude from this exercise that EHPADs do not explain the entire increase in the gradient.

Table C3: No elderly care homes

	(1) 2020	(2) Wave 1	(3) Wave 2	(4) No wave
pov=1	2.524** (1.004)	1.021 (0.648)	1.211*** (0.307)	0.291 (0.338)
Controls	✓	✓	✓	✓
Urban area FE.	✓	✓	✓	✓
Non-poor average	7.536	3.972	3.663	-0.0993
Observations	16640	16640	16640	16640

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors in parentheses clustered at the urban-area level.

NOTE: This table reports the coefficients associated to equation 4. The independent variable, excess mortality rate, is computed considering four different time periods: the whole year (column 1), wave 1 (Mars to April, column 2), wave 2 (October to December, column 3) and other months in 2020 outside the two waves (January, February and from May to September, column 4). By construction, column 1 is the sum of column 2 to 4.

C2.2 Falsification test

Figure C2 shows the results from the estimation of equation 4 using excess mortality rate based on 2018 as the sole baseline. This implies that the mortality rate of 2020 is now compared to 2018 instead of the average mortality of 2018 and 2019. One could also have used 2019 as baseline. In fact, all three measures are close, as shown by the correlation table below. Ideally, one would use a few more years but French data do not allow to identify municipality of residency at death before 2018. We also run a falsification test comparing mortality in 2019 to 2018.

Table C4: Alternative reference year

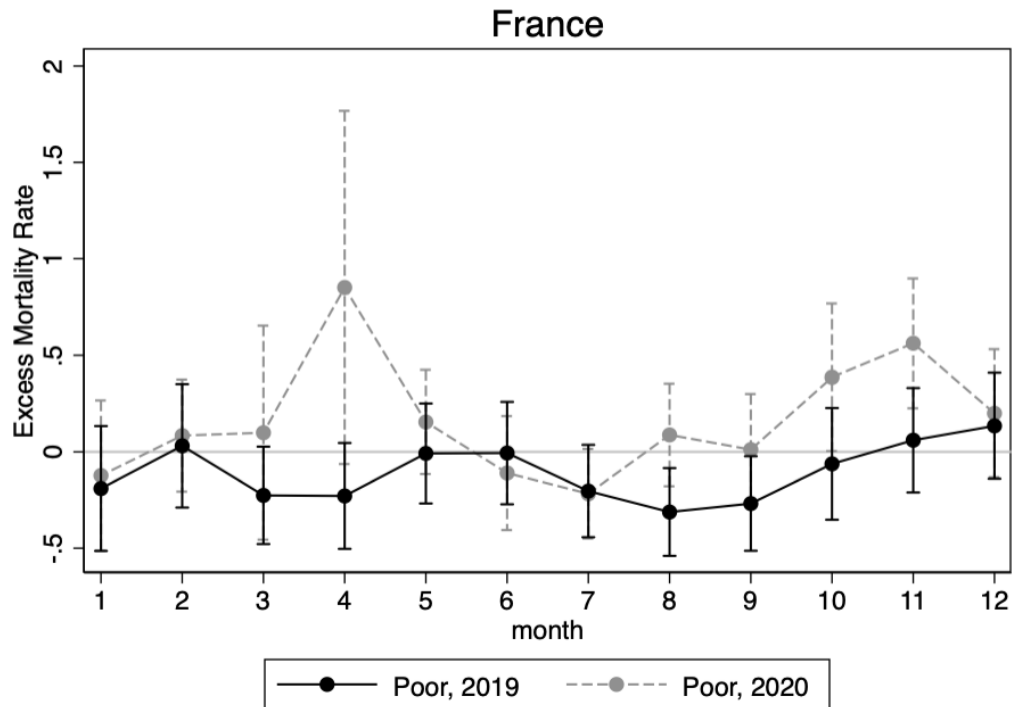
	Baseline (2018-2019)	Base 2018	Base 2019
Baseline (2018-2019)	1		
Base 2018	0.8705	1	
Base 2019	0.8708	0.5160	1

This table displays the correlation between 3 alternative measures of 2020 excess mortality rate. In baseline, the average mortality of 2018 and 2019 is used as a reference point. The first alternative compares 2020 mortality to 2018. The second alternative compares 2020 mortality to 2019.

The grey dotted line shows that using 2018 as the only reference point does not alter our main results too much. Although, of course, having only one year for reference increases the risk that some year-specific shock affected the baseline. For instance, 2018 mortality was lower than the 2018-2019 average in the case of poorest municipalities in “non-wave” months all taken together (*i.e.* January, February and May to September) creating a significant difference with 2020. This reminds us that one should be careful in drawing conclusion from year to year comparison.

Nevertheless, two take away points emerge from the black line. Overall, comparing 2019 to 2018 yields no significant difference in yearly excess mortality between poor and non-poor municipalities. Excess mortality is smaller for richer municipalities over some short periods (in August and September) but the estimated effects always remain smaller than 1 death per 10,000 inhabitants (*i.e.* about 0.3 death in the average municipality of our sample).

Figure C2: Falsification test

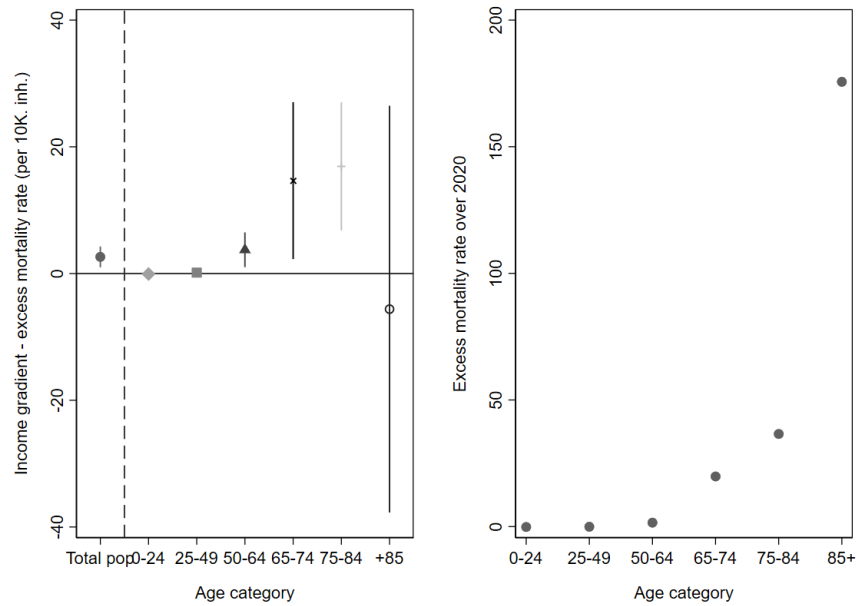


NOTE: This Figure plots the coefficient β as estimated from equation (2) using two alternative dependent variables. The black solid line shows the relative evolution of excess mortality in 2019 for poorest municipalities ($Q1$) as compared to others. For the sake of comparison, we also display the excess mortality of 2020 as compared to the same, 2018 baseline.

C3 Income gradient by age

Figure C3 shows the income gradient in excess mortality computed for different age categories (left-hand side) and the excess mortality rate over 2020 for the same age categories (right-hand side). We observe that the income gradient (i) increases with age; (ii) increases with the death toll. Consistent with the evidence that age is the main risk factor in COVID-19 related mortality, the only exception is for the category of people over 85 years old. One potential explanation being that their vulnerability was so high that COVID-19 did impact their mortality, no matter their place of residency.

Figure C3: Income gradient (LHS) and Excess mortality rate (RHS) over 2020 by age category



NOTE: This Figure plots on the left-hand side the coefficient β as estimated from equation (2) run separately on excess mortality over 2020 for different age categories. The first point reports the coefficient estimated on the whole population. 95% confidence intervals are reported. On the right-hand side, we show the magnitude of excess mortality defined as the number of excess deaths per 10,000 inhabitants over 2020, for the same age categories.

Appendix D Urban-Area level analysis

To explore the validity of our results at the urban-area level, we first measure mortality shocks considering the whole urban area (*i.e.* including all municipalities irrespective of their income level). Urban-area excess mortality rate is thus equation 1 at the urban-area level.

Next, we measure urban-area specific gradient (over any given period p) by estimating the following model:

$$D_m^p = \sum_{i=1}^N \theta_i \cdot Q1_m \times \mathbb{1}[ua = i] + X_m^p \cdot \Lambda + \sum_{i=1}^N \gamma_i \times \mathbb{1}[ua = i] + \nu_m^p \quad (5)$$

We include our baseline controls: population and share of residents older than 65 years old. This regression mostly adds an interaction between urban area indicators and $Q1$ so as to estimate one gradient (slope) per urban area. The gradient can only be estimated on the 421 urban areas that host both poor and non-poor municipalities.

In Table D1 below we regress urban-area month-specific gradient as measured in equation 5 on urban-area month-specific excess mortality. The coefficient in column (1) corresponds to a pooled regression where each of the 421 urban areas is present 12 times (one per month; and we cluster standard errors by urban area). The coefficient indicates that increasing urban-area level excess mortality rate by one death (per 10k inhabitants) also increases the income gradient in mortality by 0.373 death (per 10k inhabitants). Column (2) next introduces an urban-area fixed effect and therefore only considers the changes across months to identify the effect. Columns (3) and (4) reproduce columns (1) and (2), respectively, weighting observations by urban-area population. They mostly ensure that the correlation between excess mortality and gradient is not driven by small urban areas only. Columns (2) to (4) report results that are qualitatively similar to column (1).

Table D1: Gradient and excess mortality

	(1)	(2)	(3)	(4)
	Gradient	Gradient	Gradient	Gradient
UA excess m.r.	0.373*** (0.0767)	0.356*** (0.0833)	0.319*** (0.0369)	0.319*** (0.0358)
Constant	-0.126 (0.149)	-0.111 (0.0754)	-0.0418 (0.115)	-0.0425 (0.0276)
UA FE		✓		✓
Weights			✓	✓
Observations	5052	5052	5052	5052
UA	421	421	421	421

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors in parentheses clustered at the urban-area level.

NOTE: This table reports the coefficients associated with a regression of urban areas' month-specific gradient and excess mortality. 421 urban areas have an estimated income gradient and they are observed in the 12 months (5,052 observations). Column (1) is a pooled regression while column (2) introduces an urban area fixed effect. Column (3) and (4) repeat (1) and (2), respectively, weighting by urban areas population.

We next investigate whether the 234 urban areas that suffered increased mortality in both waves also display an income gradient in the second wave. This tests the hypothesis that the positive relationship between the income gradient and excess mortality is indeed evidence that the income gradient does not only capture a difference in the timing of deaths between inhabitants of poor and non-poor municipalities. Table D2 shows that, each time, the average (and median) gradient of these 234 urban areas was greater than 0, indicating that residents of poorer areas were more likely to die (more) in 2020 (than before) than residents of non-poor areas. This remains true when putting more weight on bigger urban areas. Importantly, the positive gradient is also found in the second wave indicating that there has been no convergence in mortality rates between poorer and richer municipalities of these urban areas.

We remain cautious on interpreting the relative size of these gradients. The reduction of the average gradient indicates a slower divergence between poorer and richer municipalities in the second wave, but such effect could be the product of many factors: a) we document in the paper that wave 2 is more diffuse over France than wave 1 (an aspect that is mirrored in the higher dispersion of gradients in wave 2 shown in Table D2), b) first- and second-wave policy response differed, c) there is potential learning from municipalities (*i.e.* residents, governing bodies, etc.).

Table D2: Persistence of a gradient

	mean	sd	p50	min	max	count	weights
Gradient, Wave 1	2.84	15.91	1.05	-67.27	72.55	234	234.00
Gradient, Wave 2	1.37	19.60	0.28	-103.15	76.70	234	234.00
Gradient, Wave 1	1.49	5.67	1.62	-67.27	72.55	234	4053.86
Gradient, Wave 2	1.02	5.82	1.72	-103.15	76.70	234	4053.86

NOTE: This table describes the income gradient in both waves across the 234 urban areas that suffered excess mortality in both waves. The last two row are weighted by urban-area population (their total weights are expressed in 10k inhabitants).

Appendix E Red and green zones

E1 Evidence of different states of the epidemic by zone

Our strategy relies on the idea that urban areas in the red and green zones are comparable. In this appendix, we argue this is indeed the case even though only one of the two zones was affected by COVID-19 during the first wave. To define red and green zones, we rely on a government map issued on May 7, 2020. Appendix A2 reports and explains the map. Although this map was published at the end of the first lockdown, it actually correlates very well to the pre-lockdown situation. Recall that the government classified *départements* (an administrative unit) as either green or red, based on a set of indicators.

Pre-lockdown indicators

Two sets of indicators (used by the government to draw the map) show that the spread of COVID-19 was higher in the red zone before the lockdown started.

1. First, the average occupancy rate of intensive care beds on March 18 was 26.5% in red *départements*, against 7.0% on average in the green *départements*.⁴⁶
2. Second, on March 18, the likelihood to visit an emergency unit for suspicion of COVID-19 was twice as large in red *départements* (12.0% on average in the red *départements* vs 6.3% in the low-infection ones).⁴⁷

Pattern of mortality

Immediately related to these early indicators of the spread of COVID-19, 2020 excess mortality started to diverge between the red and green zone. Table E1 below reports descriptive statistics of mortality in each zone. It first shows that excess mortality started to differ between the two zones only in mid-March, while it was very similar over the first 2.5 months of the year. More precisely, excess mortality was slightly negative and at very similar levels across zones in January, February, and during the first two weeks of March, whereas it became positive and significantly higher in red zones during the last two

⁴⁶Note that red *départements* have more intensive care beds per inhabitant so that the high occupancy rate of intensive care beds in the North-East of France is very unlikely to be driven by worse preliminary health care infrastructures. On December 31, 2018, the average number of intensive care beds per 100K. inhabitants in red departments was 7.581 against 5.868 on average across *départements* of the green zone.

⁴⁷An alternative possibility could have been to group *départements* based on these figures from mid-March. We preferred to rely on the government's classification, which cuts France in half and minimizes the proportion of geographical areas classified as highly infected that are contiguous with geographical areas classified as low-infected. While it seems credible at the national level that the epidemic first spread to the northeast quarter for random reasons, this hypothesis seems much stronger if it has to be made multiple times at a finer geographical scale, like *départements*. Note however that the nature of our results remains if we prefer this classification.

weeks of March and in April. Given that most COVID-related deaths take place about three weeks after the contamination (Courtejoie and Dubost, 2020; Delfraissy et al., 2020), Table E1 is in line with the idea that contamination to COVID-19 was much stronger in the red zone than in the green one in early March 2020. Note also that excess mortality is still very close to 0 during the last two weeks of March in the green zone, consistent with the notion that COVID-19 was almost absent from this zone up to mid-March.

We argue that the sharp heterogeneity in the spread of the virus between the two zones in mid-March is largely unrelated to socio-economic factors potentially affecting inequalities in mortality between rich and poor municipalities. Two studies by Le Bras (2020) and Fouillet et al. (2020) provide evidence on the absence of a relationship between the location of the first epicentres and socio-demographic characteristics at the *département* level. The absence of differential trends in total excess mortality across the two zones in 2019 (compared to 2018 only) and early 2020 further supports this assumption. Eventually, we show in section 5.2 that trends in *differential* mortality between rich and poor municipalities are very similar across zones during pre-COVID periods.

Table E1: Mortality trends in red and green zones

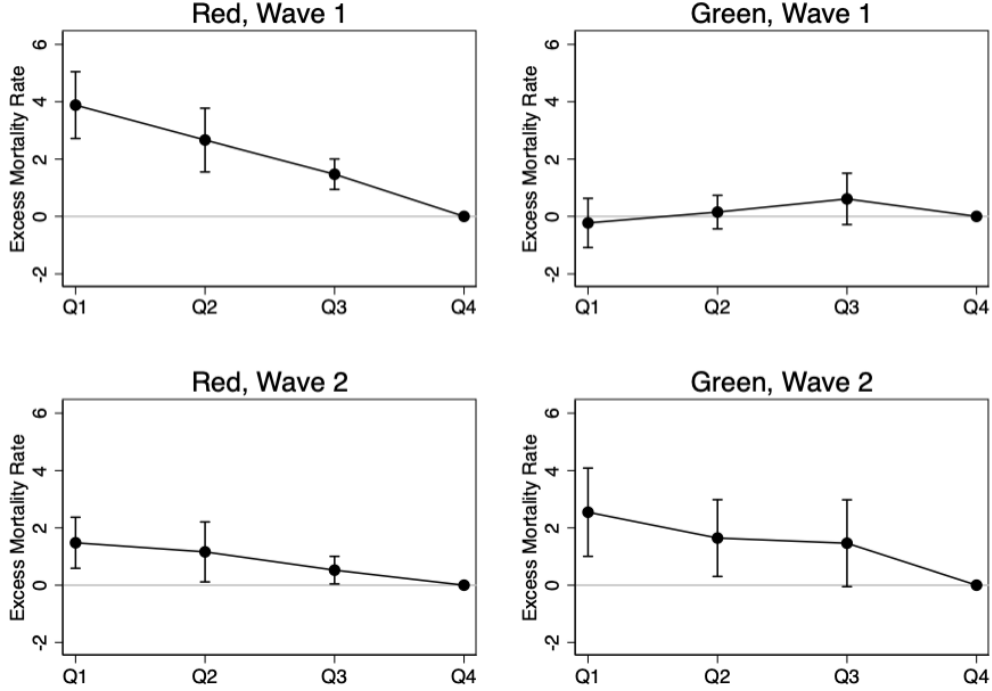
	(1)	(2)	(3)
	Red zone	Green zone	Difference
<i>2019 excess mortality rates (2019 vs 2018)</i>			
Annual excess mortality rate	0.239	1.473	-1.234
March-April	-1.357	-0.685	-0.672
<i>2020 excess mortality rates (2020 vs 2018-2019)</i>			
January (pre-COVID)	-0.225	-0.457	0.232
February (pre-COVID)	-0.355	-0.434	0.080
March	0.816	-0.087	0.904***
- first 2 weeks	-0.252	-0.165	-0.088
- last 2 weeks	1.139	0.216	0.923***
April	2.696	0.333	2.364***
May	0.341	0.359	-0.018
Nb municipalities	7,358	9,947	1,7305
Total population	23,457,888	30,460,492	53,918,380

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Mortality rates are expressed as number of deaths per 10,000 inhabitants.

Appendix F Triple-difference strategy

Figure F1: Gradient by zone in each wave



NOTE: This Figure plots the excess mortality gradient by quartiles as estimated using equation 4. The left column is estimated on urban areas from the red zone; the right column on urban areas from the green zone. The top row considers wave 1 (March, April) while the bottom row considers wave 2 (October to December). 95% confidence intervals are reported.

Explicit model

In this appendix we estimate a triple difference in the spirit of equation 3 but adjusting the model on two dimensions. First, we make the time dimension explicit. To do so, we replace our main outcome (excess mortality rate) by a simple count of deaths in any municipality m over the period p of the three years 2018, 2019 and 2020. Second, we now add a municipality fixed effect instead of the urban-area fixed effect.

$$\begin{aligned}
 Death_nb_{p,y,m,ua,d} = & c + \beta_1.Q1_{m,ua,d} + \delta.Red_d + \rho.Red_d.Q1_{m,ua,d} + \psi.Post_y + \zeta.Post_t.Q1_{m,ua,d} \\
 & + \mu.Red_d.Post_y + \phi.Post_y.Q1_{m,ua,d}.Red_d + X_{m,ua,d}.\Lambda + c_m + \nu_{p,y,m,ua,d}
 \end{aligned} \tag{6}$$

Where $Death_nb_{t,y,m,ua,d}$ is the number of deaths in time p (e.g. April), year y , municipality m in urban area ua and *département* d . This alternative specification allows us to control for municipality fixed effects c_m , which capture all time-invariant factors influencing municipality-level mortality between 2018 and 2020.⁴⁸ This implies that urban-area fixed effects cannot be estimated (each municipality belongs to a single urban area). It follows that in this model, we do not restrict ourselves to comparing adjacent municipalities within the same urban area.⁴⁹

We estimate this model on the first wave of the epidemic (March-April), when the situation was suited for the use of a triple difference (cf. section 5 on that matter). Column 2 of Table F1 indicates that, on average in non-poor municipalities in urban areas located in the green zone, the number of deaths went from 2.29 in March-April 2018-2019 to 2.48 in March-April 2020. This number grew to 3.23 in non-poor municipalities in urban areas located in red zone (+41%). Poor municipalities in red zone experienced, on average, 5.84 deaths, which is 2.5 times the baseline average (2.29 deaths). It means that, within the red zone, the rise in the number of deaths in 2020 in poor municipalities was 81% higher than non-poor municipalities (3.23 deaths).

We also estimate this model separately on each month of the year. The related triple-interaction coefficients ρ are very similar to the one presented in Figure 5. Results are available upon request.

Because we do not need to include urban-area fixed effects anymore, we can also reproduce the estimation on the whole sample of French municipalities, including those in rural areas (column 1). Although absolute numbers are lower, the ratio between subgroups are similar. In the red zone, non-poor (respectively poor) municipalities experienced a 37% (respectively 81%) increase in mortality with respect to green non-poor municipalities in March-April 2018-2019.

⁴⁸We, therefore, do not need to divide the number of deaths by the population in the dependent variable, because we measure the population only once in 2014.

⁴⁹The absence of urban-area fixed effects also allows the estimation of the effect of living in a municipality located in a red zone that does not only rely on urban areas straddling in red and green zones.

Table F1: Triple-difference regression on the number of deaths, wave 1

	Number of deaths, wave 1	
2020=1	0.12952*** (0.03773)	0.18605*** (0.05675)
2020=1 × Red area=1	0.51121*** (0.17458)	0.75283*** (0.26704)
2020=1 × Poor=1	-0.05089 (0.03267)	0.19666* (0.11114)
2020=1 × Red area=1 × Poor=1	0.81664* (0.41663)	2.41010** (1.08131)
Municipality FE	✓	✓
Sample	All municipalities	Major urban centres and their outskirts
Control outcome mean	1.736	2.290
Observations	188688	100686

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors in parentheses clustered at the *département* level.

NOTE: This table reports coefficients from a triple-difference regression on the number of deaths in wave 1 (March-April) over the years 2018-2020. The three dimensions are (i) time, with 2020 being the post period; (ii) the level of infection in the *département* before the first wave; (iii) the poverty status of the municipality, defined as the bottom quartile of the national distribution of municipal income weighted by the municipal population size. We include municipality fixed-effects to account for time-invariant factors at the municipal level, such as the population measured in 2014. The control outcome mean is the mean number of deaths in March-April in the years 2018-2019 in non-poor municipalities in green zones.

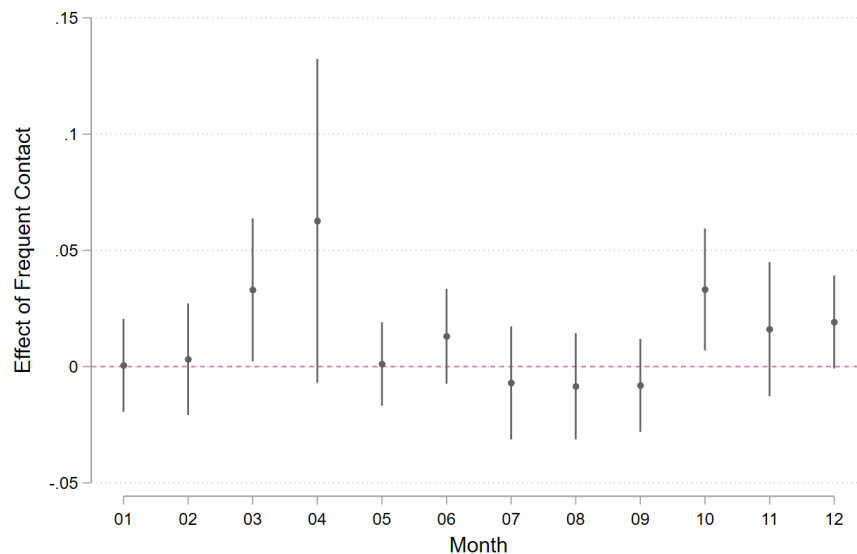
Appendix G Mechanisms

Table G1: Relationship between poverty status and mechanism variables

	Index of frequent contact	Share of essential workers	Share of over-crowded housing	Multigenerational hh
Main coefficient	0.13660*** (0.04184)	0.15666*** (0.01685)	0.28847*** (0.04121)	0.06950 (0.05116)
Urban areas FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Control outcome mean	0.251	0.251	0.251	0.272
Observations	16640	16640	16267	3989

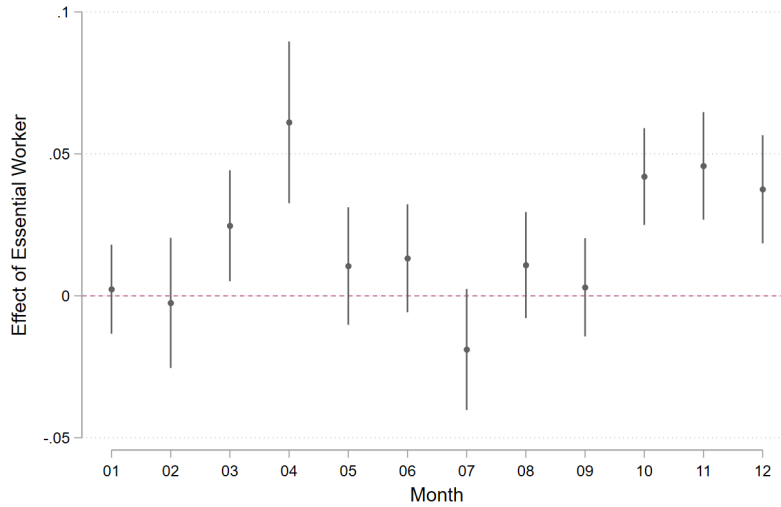
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors in parentheses are clustered at the *département* level. This table shows the result of regressing the poverty dummy on mechanism variables measuring either housing conditions or occupational exposure and the interaction of these variables. The poverty dummy is equal to one for the bottom 25% of the national distribution of municipal median income weighted by population size. Each column reports the result of a separate regression examining one mechanism. The main coefficient corresponds to the correlation between each of the mechanism variable and poverty. The mechanism variables have been standardized such that coefficients can be interpreted in terms of the effect of a one standard-deviation change, and can be compared with each other. All regressions include urban areas fixed-effects and control for total population and for the share of inhabitants over 65 y.o. in the municipality.

Figure G1: Effect of the index of frequent contact on excess mortality rate over time



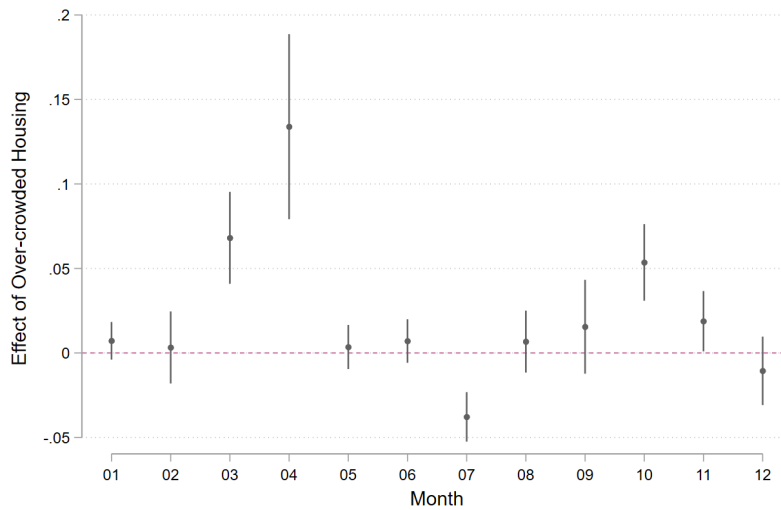
NOTE: The graph plots the coefficient of the regression of the excess mortality rate each month on the index of frequent contact, including controls for total population, share of the population over 65 y.o., and urban-area fixed effects. Both red and green areas are included. Excess mortality rate and the index of frequent contact have been normalized. Confidence intervals at the 95% level.

Figure G2: Effect of the share of essential workers on excess mortality rate over time



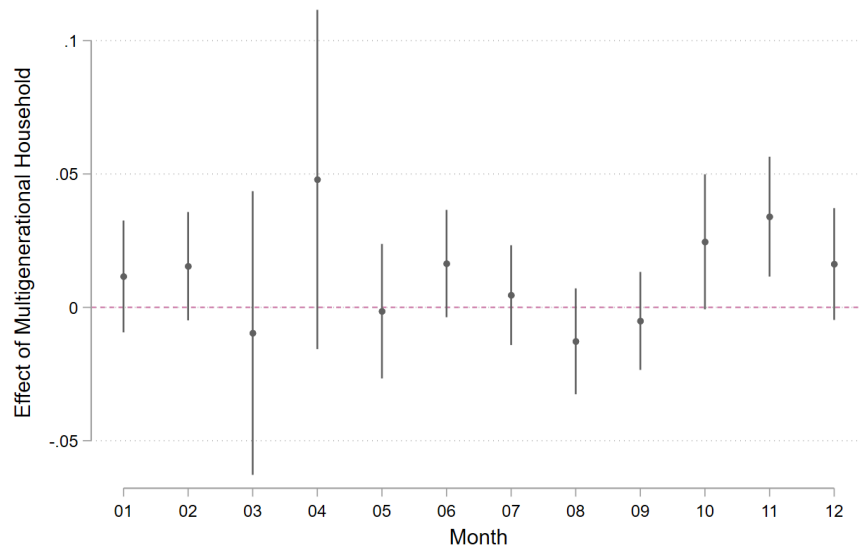
NOTE: The graph plots the coefficient of the regression of the excess mortality rate each month on the share of essential workers, including controls for total population, share of the population over 65 y.o., and urban-area fixed effects. Both red and green areas are included. Excess mortality rate and the share of essential workers have been normalized. Confidence intervals at the 95% level.

Figure G3: Effect of the share of over-crowded housing on excess mortality rate over time



NOTE: The graph plots the coefficient of the regression of the excess mortality rate each month on the share of over-crowded housing, including controls for total population, share of the population over 65 y.o., and urban-area fixed effects. Both red and green areas are included. Excess mortality rate and the share of over-crowded housing have been normalized. Confidence intervals at the 95% level.

Figure G4: Effect of the share of multigenerational households on excess mortality rate over time



NOTE: The graph plots the coefficient of the regression of the excess mortality rate each month on the share of multigenerational households, including controls for total population, share of the population over 65 y.o., and urban-area fixed effects. Both red and green areas are included. Excess mortality rate and the share of multigenerational households have been normalized. Confidence intervals at the 95% level.