Disentangling the role of population and employment density in the spread of COVID-19

Carlo Corradini[*], Jesse Matheson[†] and Enrico Vanino[‡]

Abstract

Using newly available big data in the area of satellite imaging in combination with traditional labour market and public health data for neighborhoods in England and Wales between March 2020 and April 2021, we analyse the dynamic relationship between density and the health consequences of the COVID-19 pandemic, offering novel evidence on the different effects exerted by population and employment density. Our results show that the concentration of resident key workers, providing an essential service which cannot be done remotely and thus carrying on working onsite throughout the pandemic, increases the incidence of COVID-19 in the neighborhood both in terms of cases and of deaths. This effect is exacerbated in the most deprived neighborhoods characterised lower income, health deprivation and crowded housing conditions, which facilitate even further the spread of the virus. Our results show how lockdowns and other public health restrictions can mitigate this relationship by partially shielding key workers, except in the most deprived areas of the country. These findings provide important insights for designing future economic policies and public health strategies, targeting more precisely the neighborhoods more vulnerable from an economic as well as a contagion perspective.

Keywords: COVID-19, Population Density, Employment Density, Agglomeration, Satellite Data.

JEL Classification: H12, I12, R10, R12.

1 Introduction

One of the most striking features of the COVID-19 pandemic is the marked spatial heterogeneity in both COVID-19 cases and deaths. Researchers have shown how such variation broadly reflects differences in the socio-economic structure across locations, including income and age distribution, quality of healthcare and institutions (Carozzi et al., 2020; Desmet and Wacziarg, 2021; Rodríguez-Pose and Burlina, 2021). Particular attention has been devoted to the analysis of population density. Such focus is rooted in the transmission mechanisms of the SARS-CoV-2 virus, potentially including fomites but primarily through respiratory droplets and aerosol particles (CDC, 2020; Stadnytskyi et al., 2020; WHO, 2020). Indeed, localised effects are supported by the increasing evidence on airborne transmission as a key route for the spread of COVID-19 (Morawska and Cao, 2020; Zhang et al., 2020; Greenhalgh et al., 2021). Thus, a better understanding of the role of density is critical for public health policies directed at mitigating the diffusion of COVID and of the contextual role of superspreader events. In particular, a more nuanced comprehension of where and how contagion takes place, whether at home or at the place of work, and in the latter case through which type of jobs, will be increasingly important as outbreaks slowly recede (Lewis, 2021). At the same time, this is also critical for designing policies addressing economic

[*]University of Birmingham, Birmingham, UK.
[†]University of Sheffield, Sheffield, UK.
[‡]University of Sheffield, Sheffield, UK. Corresponding author: e.vanino@sheffield.ac.uk
Acknowledgments: This work has been supported by the Regional Studies Association (RSA) Small Grant Scheme on Pandemics, Cities, Regions & Industry. We are grateful for the feedback received at the Regional Studies Association 2021 conference. We would like to thank Rhiannon Thomas and Yusuf Sohoye for their research assistance.

effects of the crisis, as discussed in recent papers exploring effects of the COVID-19 pandemic on productivity (McCann and Vorley, 2021), SMEs performance (Bartik et al., 2020a), and the shift toward working from home (Bartik et al., 2020b; De Fraja et al., 2021).

Yet, its relationship with the spread of COVID-19 remains unclear; while some works point to a positive effect (Wong and Li, 2020; Allcott et al., 2020; Desmet and Wacziarg, 2021; Almagro and Orane-Hutchinson, 2020), others find mixed or not significant evidence once other factors are controlled for (Carozzi et al., 2020; Ascani et al., 2021; Armillei and Filippucci, 2020). Understanding the role of density is particularly challenging as this factor is inherently tangled with related issues, such as the local structure of industries and jobs, household composition and characteristics, and policy interventions aimed at limiting social interaction in response to the pandemic. Thus, while previous studies have mostly focused their attention on the role of population density, as this measure is commonly available from national census, this approach is hindered by several shortcomings.

First, the level of spatial analysis adopted particularly matters in this case, where while most studies have used aggregated data at the county or regional level, more granular analyses are needed, as recently as suggested by works at the ZIP level in the US (Glaeser et al., 2020; Guha et al., 2020)). This is necessary to capture the highly localised dynamics in the diffusion of COVID-19 and the role of mobility and 'superspreader' points in driving the majority of infections (Kuebart and Stabler, 2020; Chang et al., 2021). Indeed, this may lead to significant variation in contagion even within the same urban areas, and it is reflected in high incidence of COVID-19 cases observed in both large urban agglomerations as well as small cities in rural environments (Kuebart and Stabler, 2020; Bailey et al., 2020).

Secondly, while population density has been used as a catch-all proxy due to its common availability, it is becoming increasingly important to understand how and where density interacts with the virus. In this regard, we need to identify precisely whether a high concentration of interacting people is more likely to spread the virus at home, school, work or at amenities, and under which conditions. In particular, the role of industrial and employment density has been little investigated so far[1], mainly due to data limitation, a limited theoretical understanding of the link between employment and COVID, and the difficulty to precisely link employment density to COVID cases and deaths. Linked to that, is the fact that the interaction between industrial/employment density and COVID is defined by different occupational structures of places, with different implications in terms of the occupations and industries mostly exposed to the virus, which were not shielded and might have facilitated the spread of the virus. This is especially the case for *keyworkers*, critical or essential workers both in the public and private sector who are considered to provide an essential service which cannot be be done remotely working from home, and thus had to carry on working onsite throughout the pandemic. In the UK, this group included not only medical personnel and first responders, but more broadly jobs in the energy sector, in primary education and child care, agriculture and food production, critical retail, some manufacturing jobs, as well as public transport workers. It is thus fundamental to understand the complementary contribution of the local economy structure to the diffusion of the pandemic, as the role of marked spatial variation in economic activities and local labour market composition have remained basically overlooked. This will be increasingly important as COVID becomes endemic in the society, with the need for targeted public health measures to address highly localised outbreaks linked to different types of social interaction.

Related to these elements, the role of density is likely to be severely affected by public health policy responses. Therefore it is necessary to explore how these determinants evolve and are affected during different phases on public health measures, as lockdowns and other social interaction restrictions. While previous studies have mostly focused on static investigations, a dynamic analysis of the role of densities is needed to better understand the mechanisms through which densities are related to the spread of the virus, and to evaluate the effectiveness of public health interventions.

In this paper, we endeavour to take a closer look at the relationship between urban density and the health consequences of COVID-19, overcoming the challenges previously discussed, and in particular disentangling between the role of population and employment density, and the employment structure of residents and workers in a neighborhood. Our contribution to the literature is fourfold.

[1] A few exceptions offer evidence based on broad spatial or occupational aggregates (Almagro and Orane-Hutchinson, 2020; Ascani et al., 2021).

First, by using highly granular data at the neighborhood level of middle layer super output areas (MSOAs) for England and Wales, we are able to precisely identify the link between localised density and COVID. Furthermore, we use big data in the area of satellite imaging considering variation in where people are located during the day and at night to capture population mobility across space to explore where transmission is more likely to occur, either at the workplace or at home. To this end, we consider different measures of population density, considering both housing density and space per person.

Secondly, we further disentangle the effect of employment from population density considering differences across neighborhoods in the distribution of industries, occupations and local labour market composition. In particular, we estimate the effect of the employment structure of residents and workers in a neighborhood by focusing on the role of key workers, considering the share of residents and employees in a neighborhood who needed to work onsite throughout the pandemic and how their density played a role in the spread of the virus.

Third, we develop a dynamic model considering COVID-19 weekly cases, monthly deaths and excess mortality registered at the neighborhood level between March 2020 and April 2021 in order to explore how the role of densities were affected by public health measures and restrictions to social interaction, and has evolved over time in response to the health crisis. This is essential to better identify a more nuanced role of employment density, considering the changes to work on site for key and non-essential workers.

Finally, we assess whether transmission through population and employment density is exacerbated in combination with the geographical distribution of income, health and housing deprivation across neighborhoods in England. This analysis is particularly relevant, considering the social justice implications of lockdowns and other public health measures in particular for the most deprived neighbourhoods in the country, where there are high concentrations of resident *keyworkers* who were not able to work from home and had to resume working onsite to avoid losing income, with vicious consequences for the spread of the virus in such areas exacerbated by health deprivation and house crowding.

Our results show that the concentration of resident key workers, providing an essential service which cannot be done remotely and thus carrying on working onsite throughout the pandemic, increases the incidence of COVID-19 in the neighborhood both in terms of cases and of deaths. This effect is exacerbated in the most deprived neighborhoods, characterised lower income, health deprivation and crowded housing conditions which facilitate even further the spread of the virus. However, lockdowns and other public health restrictions have helped in mitigating this relationship, except for the most deprived ares where resident *keyworkers* kept working on site and spreading the virus in their local communities.

The structure of the paper is set out as follows. In Section 2, we review the emerging literature on the link between density and COVID-19 and the policy interventions adopted in the UK to hamper transmission. Section 3 outlines data an research design for the empirical analysis. Results are presented in Section 4. Section 5 concludes the paper discussing policy implications and limitations of the study.

2 Literature review and research questions

A growing literature is rapidly emerging on the spatial variation of COVID-19 incidence rates. In particular, significant attention has been dedicated to the role of population density. Densely populated areas are naturally defined by important differences in terms of socioeconomic elements that have clear implications in the context of the pandemic, such as age distribution, income, ethnicity and health infrastructure (Almagro and Orane-Hutchinson, 2020; Sá, 2020; Desmet and Wacziarg, 2021). Another element potentially connected to density is pollution. Studies based on US county and UK regional data indicate a significant effect of air pollution whilst controlling for several factors, including population size and density (Wu et al., 2020; Travaglio et al., 2021). Similar effects have been found using data from other Countries (Cole et al., 2020; Fattorini and Regoli, 2020). Once these elements are controlled for, density potentially retains a critical role in the diffusion of COVID-19 considering the transmission mechanisms of the SARS-CoV-2 virus.

The link between airborne transmission of COVID-19 and population density reflects insights from spatial variation patterns of the 1918–1919 influenza pandemic. Exploring US city-level data, previous research suggests a positive correlation between population density and influenza mortality (Garrett, 2007). Exploring the economic consequences of the 1918 pandemic at State and city level, Correia et al. (2020) suggest that higher mortality in urbanised areas with greater manufacturing activity could be linked to higher density. Looking at both 305 administrative units and 62 counties in the UK, Chowell et al. (2008) find a markedly higher mortality in urban areas, but no clear association between death rates and measures of population density.

Contributions on the presence of a link between population density and COVID-19 have provided mixed findings, with differences in the evidence seemingly defined by the level of spatial aggregation adopted. Using data at the provincial level in Italy, Ascani et al. (2021) find no evidence of an effect of population density on COVID-19 cases. Similarly, Rodríguez-Pose and Burlina (2021) explore excess mortality in the first wave of the pandemic across European regions but find no effect of density once institutional factors are controlled for. Carozzi et al. (2020) explore US county data and find that density affects the timing of the outbreak, but find no evidence that population density is positively associated with time-adjusted COVID-19 cases and deaths. They suggest this may be due to differences in social distancing measures, access to healthcare and demographics in urbanised areas.

Conversely, Wong and Li (2020) show that population density is an effective predictor of cumulative infection cases in the US at the county level; also, they note higher spatial resolution would be preferable, as COVID-19 transmission is more effectively defined at sub-county geographical scales. In line with this, Desmet and Wacziarg (2021) look at the role of density at county level data on COVID-19 reported cases and deaths in the US finding limited evidence on the role of population density on the former, but a positive effect on the latter. However, they show effective density - calculated as the "average density that a random individual of a county experiences in the square kilometer around her" - is a strong predictor of cases and death. Similarly, a proxy measure for persons per household is also found to exert a significant effect on both.

The role of density is also underlined by studies exploring cross-sectional data at a higher levels of spatial aggregation. In the US context, researchers have similarly found robust evidence on the link between density but particularly number of people per household and COVID-19 cases looking at the ZIP level for selected cities (Almagro and Orane-Hutchinson, 2020; Guha et al., 2020). Similar results are found analysing MSOAs in England and Wales (Sá, 2020). Conversely, focusing on Italian municipalities, Armillei and Filippucci (2020) find a negative correlation between population density as well as measures of house crowding and excess mortality. Overall, these findings suggest that it is not density per se, but the likelihood of closed contacts - as also underlined by the consistent effect of house crowding proxies, - that matters. Thus, COVID-19 cases and subsequently death rates result from localised interaction, which is not simply a function of being in a large urban area as opposed to smaller city environment.

The role of density and its localised nature are inherently connected to the structure of the local economy. Ascani et al. (2021) explore a spatial autoregressive model of COVID-19 cases in the Provinces (NUTS2) of Italy to look at the role of the underlying economic structure, defined as an employment-weighted Herfindahl–Hirschman index. They find evidence suggesting larger employment in geographically concentrated industries positively impacts COVID-19 cases. This effects seems to be driven by manufacturing employment. Thus, they suggest activities that are usually defined by agglomeration advantages in industrial may be more conducive to COVID-9 transmission. Interestingly, the coefficient for population density is negative once the economic structure is controlled for. Armillei and Filippucci (2020) highlight similar elements, with the share of industrial and trade employment being positively associated with excess mortality, whilst service employment share is found to have a negative relationship. Almagro and Orane-Hutchinson (2020) offer a more disaggregated view on the role of occupations looking at COVID-19 cases in New York across 13 different employment classes. Their findings suggest that the share of employment in specific sectors is positively associated with positive tests for COVID-19, notable Essential Professional, Industry and Construction and Transportation. However, only the latter remains significant after the introduction of stay-at-home orders in New York. Interestingly, the role of public transport - which has received contrasting results in other studies (Sá, 2020; Armillei and Filippucci, 2020; Desmet and Wacziarg, 2021) -, is no longer significant once occupation variables are controlled for (Almagro and Orane-Hutchinson, 2020).

While most of these contributions explore density using a cross-section perspective, the COVID-19 pandemic has been characterised by strong policy intervention aimed at restricting mobility, including stay-at-home-orders in the US and similarly public health measures in the UK (Alvarez et al., 2020; Acemoglu et al., 2020; Courtemanche et al., 2020). In the period between March 2020 to April 2021, England has gone through three different lockdown phases. At the end of March 2020, lockdown measures were introduced to reduce transmission during the first wave of the COVID-19 crisis, with only essential workers allowed to go to work. These measures were slowly relaxed in May, with Schools and non essential shops reopening in June. A second, less sever, lockdown was initiated in the autumn, with work from home recommendations wherever possible. These measures were increased in November to the level of the first lockdown. Measures were removed in early December, but they returned in full at the end of December, with a third national lockdown officially introduced on the 6th of January. at the onset of the third wave. This final lockdown measures would be relaxed only starting from March 2021.

As shown by Glaeser et al. (2020) exploring zip-code level data for selected cities in the US, restrictions on mobility may lead to significant reduction in COVID-19 cases, with total cases per capita decreasing up to 30% for every ten percentage point fall in mobility. Similarly, the lockdown strategy introduced in Italy at the beginning of the first wave has been shown to reduce the spread of the virus away from provinces first hit (Bourdin et al., 2021). After the onset of the pandemic, the role of density has not been shaped solely by policy. Indeed, changes in mobility reducing transmission rates have also been the result of voluntary social distancing responses (Allcott et al., 2020). Paez et al. (2020) present similar results. Looking at COVID-19 cases across Spanish provinces, they identify a significant but negative effect of density during a lockdown phase when only essential activities were allowed, suggesting the presence of a stronger behavioural response in places with a higher perceived level of risk.

These changes in behaviour and mobility have effects across all channels of COVID-19 transmission. Evidence from New York across the first wave of cases suggests the positive effect of the share of employment in Essential and Non-Essential Professional and Service occupations reduces and then disappers after the introduction of stay-at-home orders (Almagro and Orane-Hutchinson, 2020). Only workers in Transportation and Other Health remain a positive factor in the number of cases, pointing to lockdowns reducing risk in public places or the workplace, but only mitigating the transmission in occupations still operating through these mobility restrictions. Interestingly, the results by Almagro and Orane-Hutchinson (2020) also highlight that while lockdowns may reduce transmission across occupational categories, the effect of household size remains unchanged, suggesting that shelter-in-place policies may have a limited effect on intra-household contagion.

These insights suggest that the relationship between density and COVID-19 incidence may be strongly localised. In particular, we would expect density to drive transmission mostly in specific settings, where contact is more persistent and sustained. This suggests it is density where people live, as opposed to density in the workplace, that may lead to higher COVID-19 cases and subsequently death. In the same way, household size and households in poorer areas may suffer from higher incidence to COVID-19, due to the higher density in the places where people spend more of their time indoor, that is, at home.

This is likely reflected in the occupational effects. As most workers moved to a work from home solution during the pandemic, *keyworkers* still operating and engaging in their usual activities can be expected to achieve much lower social distancing, even with the introduction of public health recommendations in their workplace. Thus, for the same level of density, we would expect areas with a higher share of *keyworkers* living there to be characterised by higher levels of COVID-19 cases and deaths. Very similar dynamics should be expected in areas with higher household density (house crowding) that, similarly to poorer neighborhoods, are likely to lead to high levels of intra-household contagion. Again, this is likely to be exacerbated in places with more *keyworkers*, that are more exposed to contagion during the day, and may spread contagion once back home.

Finally, we suggest these effects to be significantly affected by lockdown policies. In the absence of lockdowns, the link between density or *keyworkers* and the effect of more deprived areas can be expected to be much more defined. However, lockdowns are likely to reduce the transmission through keyworkers as these enter in contact with a much smaller population. Thus, the link between *keyworkers* and deprived areas will recede. We would therefore expect a similar reduction on COVID-19 deaths in deprived areas with a higher share of *keyworkers*. However, this may not be the case in areas with higher population density. Reflecting previous findings (Almagro and Orane-Hutchinson, 2020), lockdowns can be expected to mitigate contagion in places with larger and less deprived households, but their effect may not be as useful in more deprived areas where contacts and social distancing are more likely to remain elevated.

3 Data and Methodology

3.1 Baseline Analysis

Our analysis is based on several datasets linked together at the Middle Super Output Area (MSOA)[2] gathering together information about the spread of the infection, socio-economic characteristics of neighborhoods, working population density, and residential population density.

We first consider the role of urban density in facilitating the spread of the COVID-19 virus, as similarly analysed in previous studies (ref.), by estimating the following baseline model with a panel OLS measuring urban density with satellite imagery data:

$$COVID_{it} = \alpha_1 SatDensity_{it} + X_i'\Gamma + \lambda_t + \theta_r + \gamma_{r,t} + e_{it}, \tag{1}$$

The dependent variable reflects the log of weekly cases, $COVID_{it} = ln(1 + cases_{it})$ between March 2020 and April 2021 in each MSOA i. As a robustness test, we replicate all our results considering instead a cross-sectional OLS model where the dependent variable is the cumulative measures of COVID-19 cases over this period (Table A1 in the Appendix), a panel OLS model where the dependent variable is the monthly number of COVID-19 deaths over this period (Table A2), or the monthly excess mortality in 2020 in respect to the 2018-2019 average (Table A3). Data on the spread of the COVID-19 pandemic in the UK are provided by the Office of National Statistics (ONS). As shown in Figure 1, it is possible to notice stark differences in the number of COVID-19 cases and deaths across neighborhoods, even between closely located ones within the same local authority district.

[FIGURE 1 HERE]

The main independent variable of interest in the baseline specification is urban density measured using satellite imagery data $SatDensity_{it}$ following novel approaches in urban economics (Henderson et al., 2019; Roca and Puga, 2016). This allows a precise estimation of where people are at different points in time and at a granular level to fill the gaps in more conventional datasets. We first use data from the GHS-POP spatial raster dataset on the distribution of people per 1 squared kilometer cell for each month in 2015 (Schiavina et al., 2019). This variable represents an ambient population distribution averaged over 24 hours and it is estimated using census demographic and geographic data together with remote sensing imagery analysis techniques[3]. In addition, in a more detailed specification, we use data from the ENACT-POP spatial raster dataset capturing the seasonal nighttime and daytime changes in the number of people per squared kilometer in 2011 (Schiavina et al., 2020).[4] Despite the lack of recent updates in the data, this dataset is useful for distinguishing between where people live (proxied by nighttime population) and where people usually are during the day due to work, schooling or leisure (proxied by daytime population). These data also allow us to check for month by month seasonal adjustments in these two dimensions[5].

[2]Middle Layer Super Output Areas (MSOA) are a geographic hierarchy designed by the ONS to improve the reporting of small area statistics in England and Wales. There are 7200 MSOA in England and Wales, built from groups of contiguous Lower Layer Super Output Areas, with a mean area of 19 km2 and an average population of 7000 people (3000 households ca).

[3]We use LandScan data on the global population distribution at approximately 1 squared kilometer spatial resolution for 2019 as an alternative sources of satellite data in order to check consistency. For more information regarding the LandScan data please refer to `https://landscan.ornl.gov/`.

[4]We transform the satellite data at the MSOA level by populating the MSOA polygons with data from the 1 kilometer squared raster layer taking in account the proportion of the raster cell that each polygon covers.

[5]For more information regarding the GHS-POP and the ENACT-POP data please refer to `https://ghsl.jrc.ec.europa.eu/datasets.php`.

We find that population density is quite different across neighborhoods at daytime and nighttime; population is clustered in city centres neighborhoods during the day, but more densely located in suburban areas at night 2. This highlights the importance of disentangling between where people live and work, especially when studying the relevance of social interactions within and between households in explaining the spread of the COVID-19 disease.

[FIGURE 2 HERE]

In our baseline model we also include X_i, a vector of control variables reflecting pre-Covid characteristics of each neighbourhood provided by the ONS. These include the log number of people living in i, the log number of people working in i, portion of residents under 18 years old, portion of residents over 65 years old, house crowding measured as the number of people per residence, neighbourhood deprivation index, and the level of pollution measured by DEFRA in 2019 in terms of PM 2.5. We also include spatial lags of the number of COVID-19 cases (or deaths) in other neighborhoods within the same local authority district weighted by the pair-distance between neighborhoods[6]. To control for any unobserved time-variant heterogeneity at the local government level, we include Local Authority District fixed-effects θ_r, time (week or month) fixed-effects λ_t, and local authority time trends $\gamma_r t$[7]. All other time-and-neighbourhood varying factors are included in the error term e_{it}.

3.2 Disentangling Population and Employment Density

In order to distinguish between the role of employment density and of population density in facilitating the spread of the virus, we improve our baseline model in the following way:

$$COVID_{it} = \alpha_1 Popden_i + \beta_1 Empden_i + X_i'\Gamma + \lambda_t + \theta_r + \gamma_{r,t} + e_{it}, \tag{2}$$

In this case, the independent variables of interest are population density, $Popden_i$, and employment density, $Empldeen_i$, each of which reflect number of residents or workers in the MSOA per unit of land area (measured in hectares). The coefficients of interest, α_1 and β_1, are estimated from the within-local authority district variation in residential and employment density across different neighborhoods. Under the condition that $E(e_{it}|Popden_{i,t}, Empden_{it}, X_i', \lambda_t, \theta_r, \gamma_{r,t}) = 0$, α_1 and β_1 reflect the causal effect of density on Covid-19 cases, deaths and excess mortality. For residential population density we have information at the level of lower super output level (LSOA). Each LSOA is contained exclusively within a single MSOA. This allows us to calculate a more precise measure of geographic density following Glaeser and Kahn (2004). We calculate density for each MSOA as the weighted sum of residents per hectare for all LSOAs within the MSOA:

$$Popden_i = \sum_{j \in MSOA_i} \frac{N_j^r}{Area_j} \times \frac{N_j^r}{N_i^r}, \tag{3}$$

where N_j is the LSOA population and N_i is the MSOA population. $Pop.Den_i$ is therefore the average density of all LSOAs within MSOA weighted by population share. We calculate a similar measure for employment density. However, because employment information is only available at the MSOA level, we calculate the simple measure of workers per unit of land area:

$$Empden_i = \frac{N_i^w}{Area_i}, \tag{4}$$

All data for population and employment are provided from the Office for National Statistics. Residential population counts are based on 2019 population estimates, while employment counts

[6]All variables have been standardised in order to allow for an easy comparison of their coefficients across our estimations.

[7]Results are robust to controlling for local labour market fixed-effects including instead Travel to Work Area (TTWA) fixed-effects.

by occupation are based on the 2011 population census.

We further explore the role population density and composition play in channelling the spread of Covid-19 by considering the employment structure of the population. Specifically, we may expect that neighbourhoods in which many workers can do their jobs from home to have a different level of contagion than neighbourhoods in which many workers continue to work onsite. Notice that the residential population, N_i^r, and the working population N_i^w (some of whom may also be residents), can be decomposed into the following types of people:

$$\begin{aligned} N_i^r &= nonworkers_i^r + keyworkers_i^r + homeworkers_i^r + otherworkers_i^r, \\ N_i^w &= keyworkers_i^w + homeworkers_i^w + otherworkers_i^w. \end{aligned}$$

In the above equations *nonworkers* is the number of people in the neighborhood who do not work, including children and retirees. Of course, we only observe this group for the residential population, not the working population. The variable *keyworkers* measures the number of people in an occupation denoted as *key work*, such as hospital staff, and would have likely continued working onsite through pandemic lock-downs. *homeworkers* are the number of people who would have been able to do a significant portion of their job from home. Finally, *otherworkers* is the number of people employed in work that is not *key work* and is unlikely to be done from home. This final category would include many workers in retail and hospitality. We denote the proportion of the working population in each of the categories by NW, KW, HW and OW, respectively, where:

$$\begin{aligned} 1 &= NW_i^r + KW_i^r + HW_i^r + OW_i^r, \\ 1 &= KW_i^w + HW_i^w + OW_i^w. \end{aligned}$$

For each MSOA we observe the number of residents employed in each of 362 occupations[8]. We denote the count of residents in each occupation, denoted o_4, and neighbourhood i by N_{i,o_4}^r. To identify the number of jobs that likely would have continued to be done onsite throughout the first year of the pandemic, we use the classification from the *Key Workers Reference Tables*, ONS (2020), which classifies jobs by occupation and industry as *key*, *KW=1*, or not *KW=0*[9]. From these tables we create a key work index for each of the 362 occupation codes by calculating the weighted average value of KW for each occupation code, where weighting is based on information from all Jan 2017-Jan 2020 waves of the UK Quarterly Labour Force Survey. The resulting occupation-specific index, $KW_{o_4} \in \{0, 1\}$, is then used to calculate the proportion of the residential population in a keywork job in each neighbourhood i. We combine this with the occupation-specific work-from-home index, $h_{o_4} \in \{0, 1\}$, from De Fraja et al. (2021). This index tells us, the proportion of work in each occupation that can be done from home. Using this information we calculate the proportion of residents that are employed in key work occupations that require being onsite:

$$KW_i^r = \frac{\sum_{o_4} N_{i,o_4}^r \times KW_{o_4} \times (1 - h_{o_4})}{N_i^r}, \tag{5}$$

We also calculate the proportion of jobs that are performed in each neighbourhood (workers may live in the same neighbourhood or elsewhere), KW_i^w. For workers we observe 90 occupations[10]. We calculate KW_i^w using the same method as described above, only now we must aggregate the key work index and work-from-home index to the three-digit SOC, which we denote as $\hat{KW}_{o_3} \in \{0, 1\}$ and $\hat{h}_{o_3} \in \{0, 1\}$. The proportion of keywork workers employed in each neighbourhood is calculated as:

[8] Four-digit occupation codes as defined by UK Standardized Occupational Classification.

[9] Key worker information is reported for each four-digit SOC and four-digit SIC combination. There are 124,564 combinations in total, many of which contain no or very low actually employment in practice. More information available at the following link.

[10] Three-digit occupation codes as defined by UK Standardized Occupational Classification.

$$KW_i^w = \frac{\sum_{o_3} N_{i,o_3}^w \times KW_{o_3} \times (1 - h_{o_3})}{N_i^w}. \tag{6}$$

where KW_i^w takes a value between 0 and 1 reflecting the amount of work done by all employees in MSOA i that requires being onsite and was not subject to lockdown restrictions. N_{i,o_3}^w denotes the number of jobs in occupation o_3 and N_i^w denotes total number of jobs, across all occupations, in MSOA i.

The proportion of homeworkers is calculated in a similar manner to key workers above. For jobs that can be done from home held by residents we calculate:

$$HW_i^r = \frac{\sum_{o_4} N_{i,o_4}^r \times (1 - KW_{o_4}) \times h_{o_4}}{N_i^r}, \tag{7}$$

and for jobs that can be done from home held by workers we calculate:

$$HW_i^w = \frac{\sum_{o_3} N_{i,o_3}^w \times (1 - KW_{o_3}) \times h_{o_3}}{N_i^w}. \tag{8}$$

The unemployed portion of the residential population, NW_i^r, is one minus the ratio of employed residents to total population. The proportion of residents and workers in any other form of non-key jobs which cannot be done from home are calculated as the residual of these shares:

$$OW_i^r = 1 - NW_i^r - HW_i^r - KW_i^r, \tag{9}$$

and

$$OW_i^w = 1 - HW_i^w - KW_i^w. \tag{10}$$

Figure 3 shows the different distribution of where *keyworkers* live and where they work within the Sheffield local authority district. In particular, notice that the share of resident *keyworkers* is particularly high in neighborhoods in the east side of the city, those characterised as well by higher levels of economic deprivation. Given that COVID-19 cases and deaths are registered at the place of residency, we will focus mostly on resident *keyworkers* density to proxy for the effect of employment density. However, we will analyse as well the effect of *keyworkers* employed in a MSOA, in order to understand if the interaction between local residents and *keyworkers* working in the same area has played a role in helping the virus to spread.

[FIGURE 3 HERE]

We account for the distribution of the population across these different employment types in our regression analysis as follows:

$$\begin{aligned} COVID_{it} = \alpha_1 Popden_i + \alpha_2 HW_i^r + \alpha_3 KW_i^r + \alpha_4 NW_i^r + \beta_1 Empden_i + \\ \beta_2 HW_i^w + \beta_3 KW_i^w + X_i'\Gamma + \lambda_t + \theta_r + \gamma_{r,t} + e_{it}, \end{aligned} \tag{11}$$

Coefficients $\alpha_{11} - \alpha_{13}$ (β_{11}, β_{12}) reflect the percent change in average Covid-19 cases (deaths) from a percentage point increase in the corresponding population (worker) shares, over residents (workers) in non-key on-site work, holding population size constant.

3.3 Additional Analysis

We consider as well the dynamic evolution of the spread of the disease, following the previous literature (Desmet and Wacziarg, 2021), and regressing the number of COVID-19 cases and deaths at the MSOA level in each following time period against the measures of population and employment density as previously defined. Formally:

$$\begin{aligned} COVID_{it} = \alpha_{1,t}Popden_i + \alpha_{2,t}HW_i^r + \alpha_{3,t}KW_i^r + \alpha_{4,t}NW_i^r + \beta_{1,t}Empden_i + \\ \beta_{2,t}HW_i^w + \beta_{3,t}KW_i^w + X'_{i,t}\Gamma + \gamma_{r,t} + e_{it}, \end{aligned} \qquad (12)$$

where variables are as specified above. In practice, we estimate this as a series of cross-sectional regressions for each time period, in order to track the evolution of the effect of population and employment density over time. This will give us the opportunity to test the efficacy of the public heath measures imposed by the UK Government to control the spread of the virus, such as as the restriction of working on site to jobs identified as key worker only. The variation used to estimate the parameters of interest comes from the difference between the MSOA's population/employment density measures and the average for the local authority r.

Finally, we explore the heterogeneity of these results across different neighborhoods characteristics. In particular, we would like to understand whether the public health measures introduced by the UK Government have had any heterogeneous effects on limiting the contribution of population and employment density to the spread of the virus, depending on the socio-economic characteristics of neighborhoods. To do so, we interact the population and employment density variables with the index of multiple deprivation. This analysis will inform us about the social justice implications of the national lockdowns imposed in the UK. In particular after the first one, following lockdowns allowed more categories to travel to the workplace if the job could not be done from home, while social interaction and retail where still heavily limited. This could have affected disproportionately neighborhoods with higher levels of deprivation, where most of the resident population are not able to work from home and had to resume working on site in order to avoid losing income. In an alternative specification reported in the Appendix we investigate as well the role played by house crowding in mediating the effect of population and employment density on the spread of the virus. A large number of people living in small and crowded places could significantly increase the COVID-19 contagion rate. This is particular relevant in the case of neighborhoods with high levels of resident *keyworkers*, who were forced to work onsite and exposed to social contact throughout the pandemic. *Keyworkers* were more likely to bring the virus home from work, where it could easily spread due to the high concentration of people living in the same house, in particular in the case of multi-generation households.

4 Results

We start in Table 1 with our baseline panel regression model by analysing the effect of urban, population and employment density on the weekly spread of COVID-19 cases. Column 1 considers urban density measured using satellite data, then split in daytime and nighttime density in column 2. Column 3 differentiates between the role of population and employment density, while column 4 reports as well the contribution of resident and employed keyworkers and homeworkers at the MSOA neighborhood level.

[TABLE 1 HERE]

Urban density is significant in explaining higher levels of COVID-19 cases and deaths across neighborhoods, although the magnitude of the effect is small. On average, a ten percent increase in population density is associated with a 0.80 percent increase in cases. From column 2 we can observe that this effect is mostly driven by nighttime population density rather than daytime, thus confirming the relevance of where people live and the residential density of neighborhoods. Column

3 shows that both population and employment densities are significant to explain the spread of the virus, with similar magnitudes overall.

The employment composition of residents and workers in a MSOA also appear to play a significant role in explaining COVID-19 cases. In particular, a ten percentage point increase in the proportion of residents able to work from home is associated with a 7.5 percent decrease in Covid-19 cases. However, the average effect of resident keyworkers appears to be statistically insignificant. This evidence highlight how by working from home residents in a neighborhood could slow down the infection, preventing the spread of the virus from the place of work to the place where they live. We find similar results for employment composition of people working in an MSOA. In particular, a 10 percent increase in keyworkers employed in the neighborhood increases the incidence of local residents COVID cases by 2.8%, while as shown before, more workers able to work remotely from home reduces the infection incidence in the local population. This is an important evidence of how the employment composition of local workers matters to explain local residents infection rates, in particular through the interaction of keyworkers employed in a MSOA with the rest of the local population. The remaining control variables included in our model are significant and in line with previous studies investigating their relationship with COVID-19. Consistent results are estimated when considering the cumulative measures of COVID-19 cases over this period (Table A1 in the Appendix), the monthly number of COVID-19 deaths (Table A2), or the monthly excess mortality in 2020 in respect to the 2018-2019 average (Table A3).

[FIGURE 4 HERE]

In Figure 4 we report the estimated coefficients, $\beta_{1,t}$ and $\beta_{2,t}$, from Equation 3.3. This figure allows us to examine the variation in the effect of density as public heath measures imposed by the UK Government change, rather than estimating an overall average. In Panel a) we report $\beta_{1,t}$ and $\beta_{2,t}$ for cases (both contemporaneous and cumulative); in Panel b) we report $\beta_{1,t}$ and $\beta_{2,t}$ for cases (both contemporaneous and cumulative). Bars show the corresponding 95% confidence intervals for each point estimate. To ease the comparability between $\beta_{1,t}$ and $\beta_{2,t}$, Pop_i and $Empl_i$ are estimated in standard deviations. Three things are worth pointing out from Figure 4. First, the effect of a standard deviation change in population density and keyworker resident density is very similar across all outcomes, both in terms of pattern and magnitude. Second, the change in cumulative deaths is significant and very stable over time. Third, the effect of population density and employment composition appears to be different during lockdown periods (March-July 2020 and November 2020-April 2021) than during the opening period (July-November 2020). For example, while deaths are significantly higher in neighbourhoods with more keyworkers during lockdown periods, there is no keyworker effect during the opening period. During the opening period neighbourhoods with more keyworkers experienced fewer reported cases. It should be noted that we cannot rule out that the role out of Covid-19 testing may be correlated with the distribution of keyworkers. That is, keyworkers where likely to have been targeted for testing early in the pandemic, when testing was scarce. This would explain the patterns in Panel a). A similar story could be told for population density, as larger cities may have been targeted earlier in the pandemic. This said, a couple of things are important to point out. First, by the time of the second lockdown, testing was fully rolled out and widely available. Second, we are unlikely to see a similar spurious correlation from Covid deaths.

We then explore the heterogeneity of these results across the deprivation distribution of neighborhoods. In particular, in Figure 5 we interact different population and workers employment structure variables with the index of multiple deprivation.

[FIGURE 5 HERE]

Figure 5 shows evidence of heterogeneous effects of residents employment composition on COVID-19 cases across deprived neighborhoods. In particular, we observe that the share of resident keyworkers in a neighborhood significantly increases the incidence of COVID-19 only in the most deprived MSOAs. On the contrary, a larger share of residents working from home significantly reduces infections only in the most affluent areas. These results have important policy implications in terms of the social injustice of the pandemic on different socio-economic groups. In particular, residents of the most deprived areas in the country are more likely to have keywork jobs that required onsite presence throughout the pandemic. This exposed them to social contact and *Keyworkers* resident in these areas were more likely to bring the virus home from work, where it could easily spread to the rest of the community. This was also possible thanks to the high concentration of people living in the same house, in particular in the case of multi-generation households, as unreported results on house-crowding heterogeneity evidenced. On the contrary, we do not find evidence of heterogeneous effects in the case of the employment composition of workers in a MSOA.

[FIGURES 6 and 7 HERE]

We further investigate this issue in Figures 6 and 7 by analysing whether the public health measures introduced by the UK Government have had any heterogeneous effects on limiting the contribution of residents and workers employment structure in spreading the virus in a neighborhood depending on its level of deprivation. This analysis will further inform us about the social justice implications of the national lockdowns imposed in the UK. In particular after the first one, following lockdowns allowed more categories to travel to the workplace if the job could not be done from home, while social interaction and retail were still heavily limited. This could have affected disproportionately neighborhoods with higher levels of deprivation, where most of the resident population are not able to work from home and had to resume working on site in order to avoid losing income. As a matter of fact, evidence form Figure 6 indicates that the previously discussed results seem to be driven by effects during lockdown periods, when resident keyworkers increase the infection rates in deprived neighborhoods, while larger shares of residents working-from-home reduced the incidence of COVID-19 only in the least deprived areas. Results from Figure 7 instead are mixed. In particular, the effect of keyworkers employed in a neighborhood on COVID cases is significant in most derived areas when no lockdown is in place, while in least deprived areas during lockdown periods. Similarly, employees working-from-home reduce infection in the MSOA during lockdowns only in the most deprived areas, while in the most affluent neighborhoods only when no lockdown is in force. This might be related to the different nature of keywork jobs done in rich and poor neighborhoods across the country, and should be further investigated in order to fully understand the consequences for the local residents of these areas.

5 Conclusions

In this paper, we provide novel evidence on the role of density in the COVID-19 pandemic. Exploring data at the neighbourhood (MSOA) level in England and Wales for the period between March 2020 and April 2021, we disentangle the relationship between density and COVID-19 cases and deaths along four related dimensions.

First, we extend recent findings pointing to the need to explore density at a granular micro level due to the highly localised nature of the transmission mechanisms of the SARS-CoV-2 virus (Glaeser and Kahn, 2004; Sá, 2020; Almagro and Orane-Hutchinson, 2020), and show that density at the neighbourhood level is a significant factor for transmission. Using big data from satellite imaging, we show this is especially the case at night, reflecting the impact of intra-household contagion. Additionally, we reinforce evidence indicating household crowding to be a key driver for COVID-19 diffusion (Desmet and Wacziarg, 2021; Almagro and Orane-Hutchinson, 2020).

Second, we further underline the importance to look beyond population density and consider the role of occupational structure and industry agglomerations. Our findings point to a general role of employment structure of workers as well as residents in an area, suggesting a role of employment density in the spread of the virus. More importantly, we highlight that density of *keyworkers* is a significant driver of both COVID-19 cases and deaths. This is a critical element considering these workers provide an essential service which cannot be be done remotely working from home and thus are required to carry on working onsite throughout the pandemic.

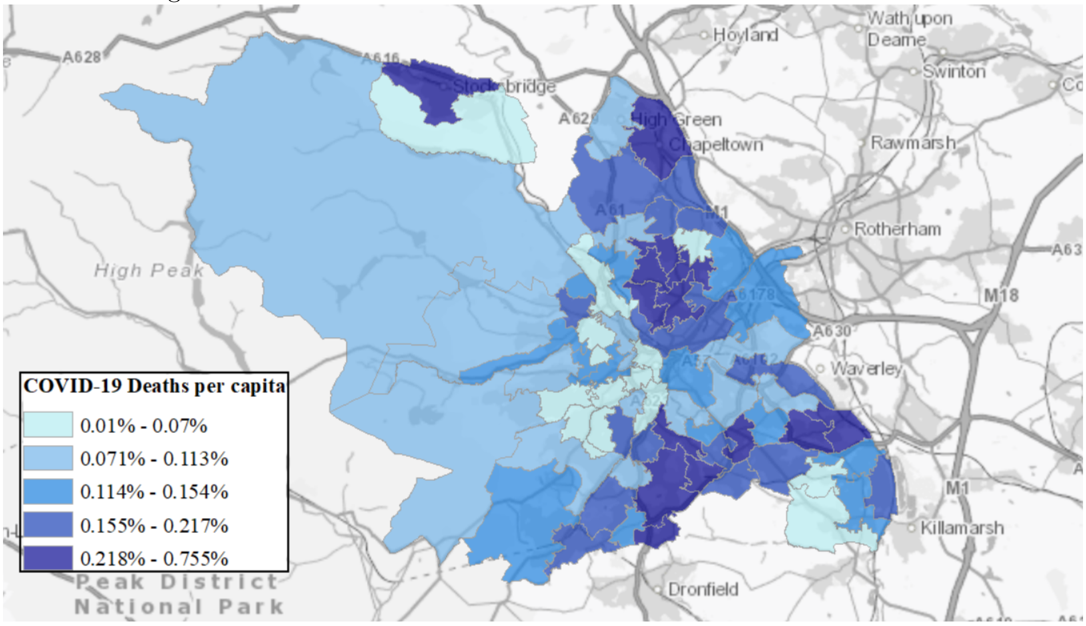
Third, while previous papers have highlighted the role of income distribution across places as a significant element in the COVID-19 pandemic (Desmet and Wacziarg, 2021; Rodríguez-Pose and Burlina, 2021), we provide novel findings pointing to a significant increase in risk in densely populated areas as well as for *keyworkers* across neighborhoods in England in the lowest quartile of income distribution, health and housing deprivation. This evidences that the relationship between high concentrations of resident *keyworkers* who are not able to work from home that often live in more deprived areas may constitute a particularly significant element in the spread of the pandemic, with important implications from both public health and social justice perspectives.

Finally, we complement research on the role of public health measures on mobility restrictions such as lockdown policies and stay-at-home orders (Glaeser and Kahn, 2004; Almagro and Orane-Hutchinson, 2020; Bourdin et al., 2021; Allcott et al., 2020). In particular, we show lockdowns have played a significant role in breaking the link connecting density of *keyworkers* COVID-19, except for the most deprived areas. This partially support previous evidence (Almagro and Orane-Hutchinson, 2020) suggesting lockdowns may effectively reduce risks of contagion in public places or in the workplace, but their role may be more limited in preventing intra-household contagion in particular for the most vulnerable and exposed categories of workers.

These results provide important insights not just to better understand determinants of diffusion of the virus, but equally to understand which areas could remain more at risk. In particular, our findings may allow to design policies considering a more nuanced role of the employment structure of residents and workers, accounting for the significant differences in changes to work on site between key and non-essential workers, as well as the relationship between these elements and the increased risks associated with residence in the most deprived neighbourhoods. These elements are essential to better design policies for preventing further negative economic shocks and implementing more effective lockdown strategies which may target more precisely the neighborhoods more vulnerable from an economic as well as a contagion perspective.

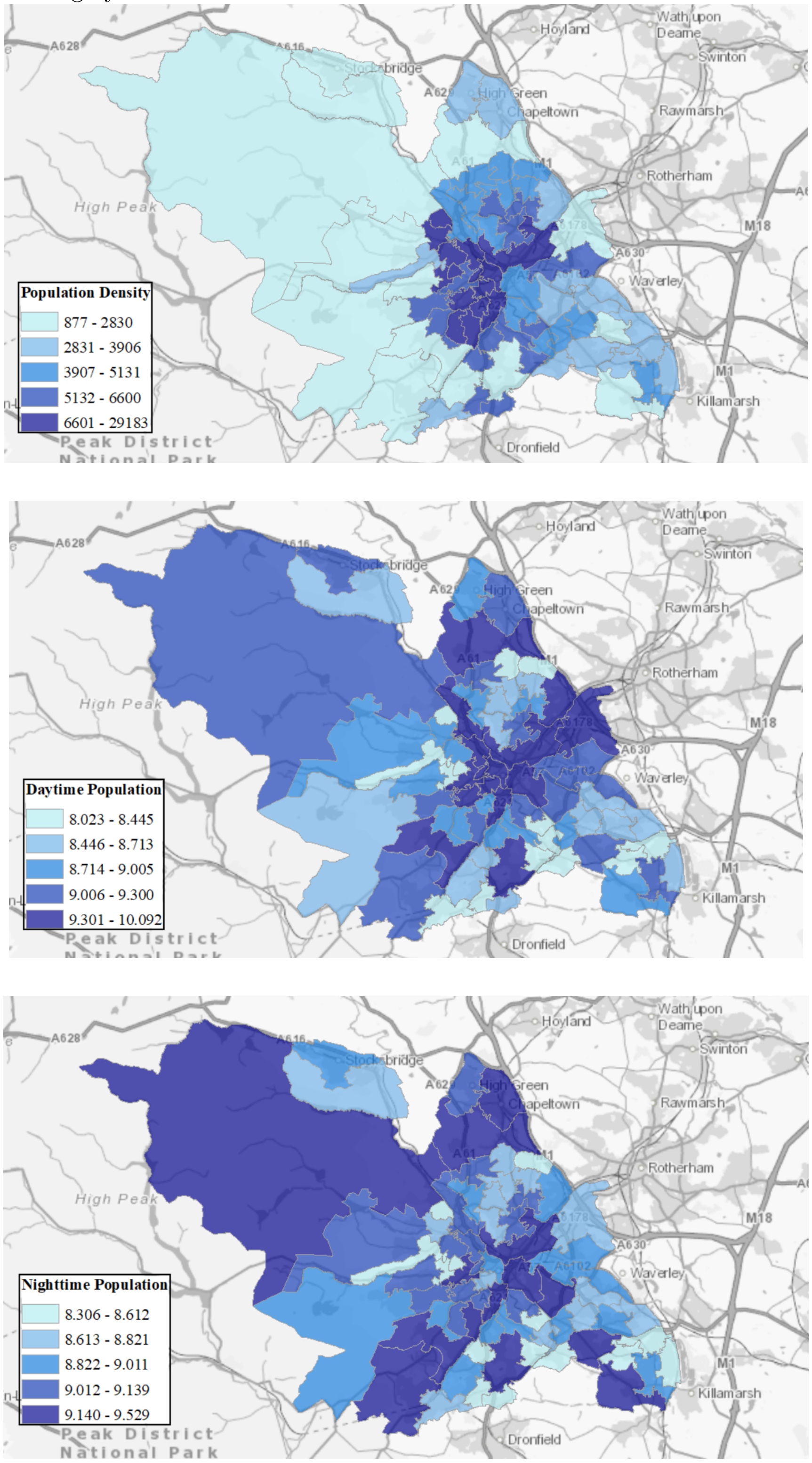
Tables and Figures

Figure 1: COVID-19 deaths across MSOAs within the Sheffield LAD.



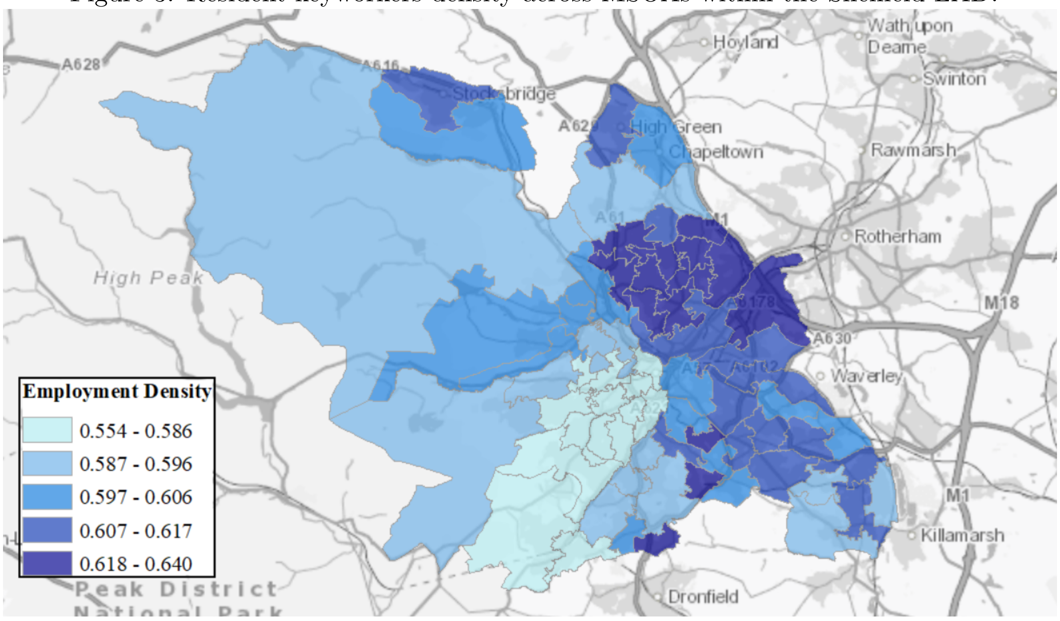
Notes: Elaboration based on ONS data for the period March 2020-April 2021.

Figure 2: Overall, daytime and nighttime urban densities across MSOAs within the Sheffield LAD using satellite imagery data.



Notes: Elaboration based on GHS-POP and ENACT-POP data.

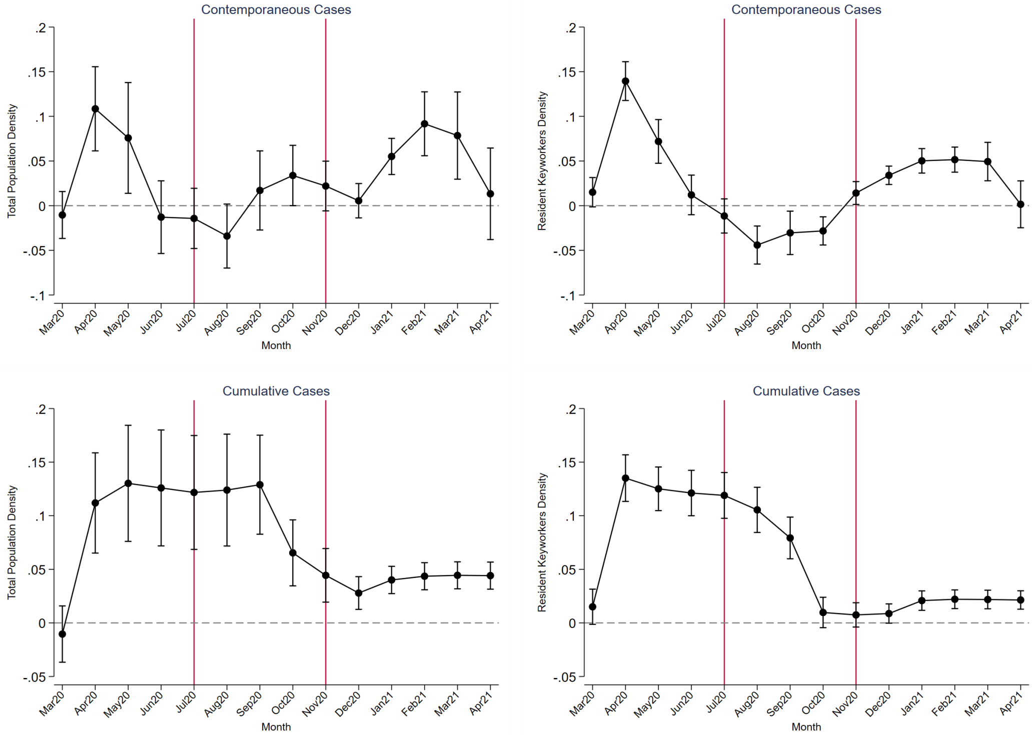
Figure 3: Resident keyworkers density across MSOAs within the Sheffield LAD.



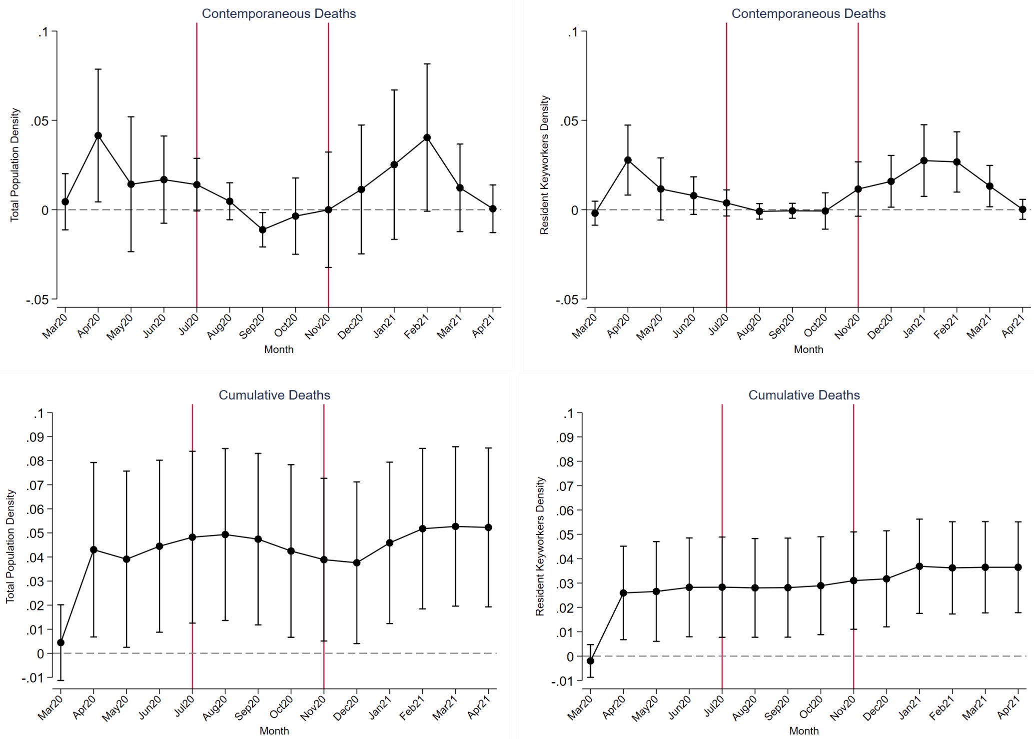
Notes: Elaboration based on Labour Force Survey data.

Figure 4: Dynamics of population density and keyworker density.

a) COVID-19 Monthly Cases



b) COVID-19 Monthly Deaths



Notes: Markers represent the coefficients from the proportion of the log-population density and the proportion of working residents that are key workers by MSOA. Different regression run for each month. Red lines show the end of the first national lockdown (04 July 2020) and the beginning of the second national lockdown (05 Nov 2020). Regressions control for local authority fixed effects, dependent children (% of pop), elderly (% of pop), log-population, log-employed residents, log-MSOA workers, proportion of workers in MSOA who are keyworkers, population per residential property, log-MOSA land area, IMD score, PM 25 pollution (average), log-weighted cases for local authority. Bars reflect 95% confindence intervals for coefficient estimates.

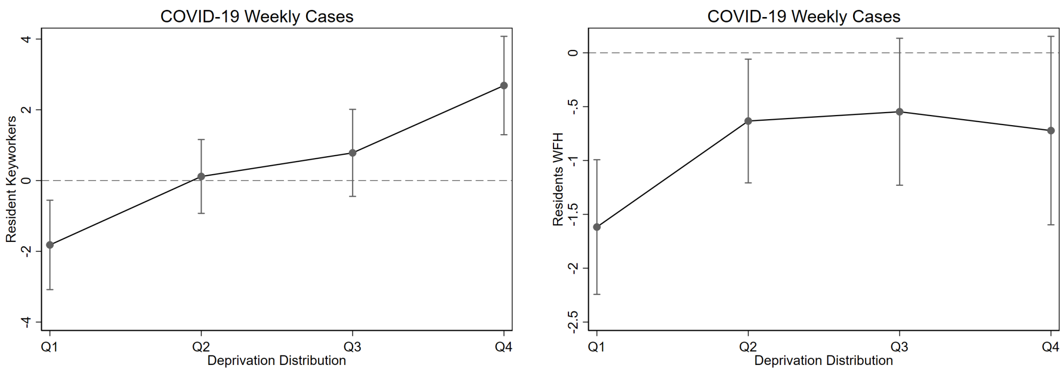
Table 1: Relationship between population, employment density and COVID-19 weekly cases by MSOA.

	(1) Weekly Cases	(2) Weekly Cases	(3) Weekly Cases	(4) Weekly Cases
Satellite Density	0.0804**			
	(0.0340)			
Daytime Density		-0.0197*		
		(0.0101)		
Nighttime Density		0.0532***		
		(0.0159)		
Population Density			0.0148***	0.0151***
			(0.00390)	(0.00393)
Employment Density			0.0103***	0.00872**
			(0.00341)	(0.00346)
Keyworkers Residents				0.486
				(0.404)
WFH Residents				-0.749***
				(0.211)
Non-working Residents				-0.408***
				(0.151)
Keyworkers Employees				0.278***
				(0.0474)
WFH Employees				-0.166***
				(0.0582)
Population	0.716***	0.685***	0.705***	0.712***
	(0.0108)	(0.0136)	(0.0107)	(0.0109)
Employment	0.0162***	0.0152***	0.0108***	0.0142***
	(0.00338)	(0.00391)	(0.00418)	(0.00412)
Share Elderly	0.168***	0.101**	0.280***	0.233***
	(0.0458)	(0.0437)	(0.0458)	(0.0496)
Share Children	0.551***	0.529***	0.533***	0.415***
	(0.0993)	(0.0976)	(0.101)	(0.0945)
Share White	-0.490***	-0.491***	-0.474***	-0.476***
	(0.0266)	(0.0266)	(0.0267)	(0.0273)
House Crowding	0.0237*	0.0216	0.0408***	0.0426***
	(0.0133)	(0.0133)	(0.0141)	(0.0143)
Deprivation	0.221***	0.222***	0.224***	0.218***
	(0.0196)	(0.0195)	(0.0197)	(0.0305)
Pollution	0.0531***	0.0524***	0.0366***	0.0365***
	(0.00443)	(0.00445)	(0.00479)	(0.00481)
Cases Spatial Lags	0.00140***	0.00140***	0.00112***	0.00109***
	(0.000128)	(0.000128)	(0.000126)	(0.000125)
LAD FE	Y	Y	Y	Y
Time FE	Y	Y	Y	Y
LAD*Time FE	Y	Y	Y	Y
Observations	434,496	434,496	434,496	434,496
R-squared	0.818	0.818	0.818	0.818

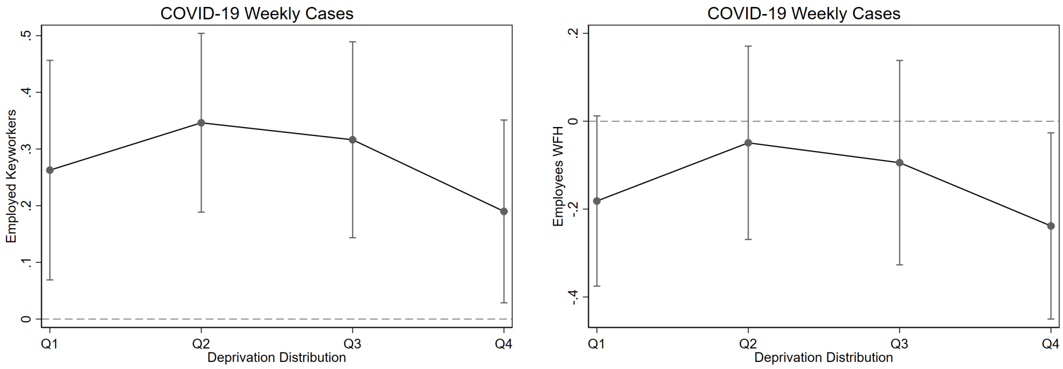
Notes: Robust standard errors in parentheses clustered at the MSOA level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Time fixed-effects at week level.

Figure 5: Relationship between resident and employed keyworkers and COVID-19 cases across the neighbourhood deprivation distribution.

a) Resident keyworkers and working-from-home



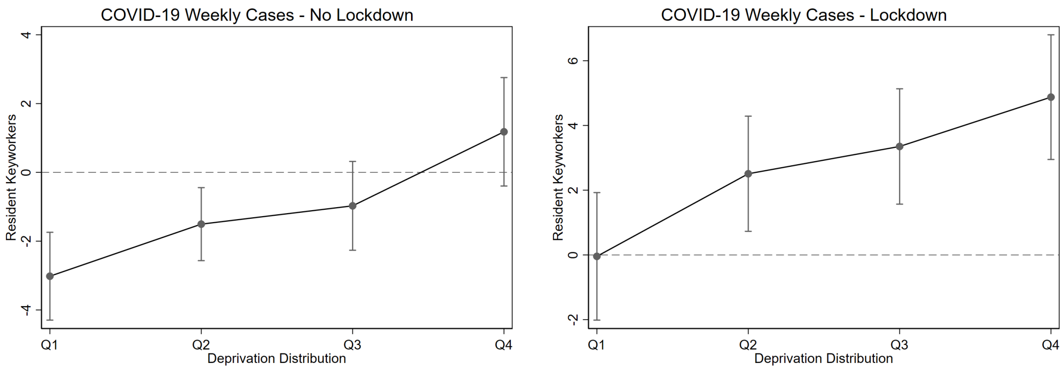
b) Employed keyworkers and working-from-home



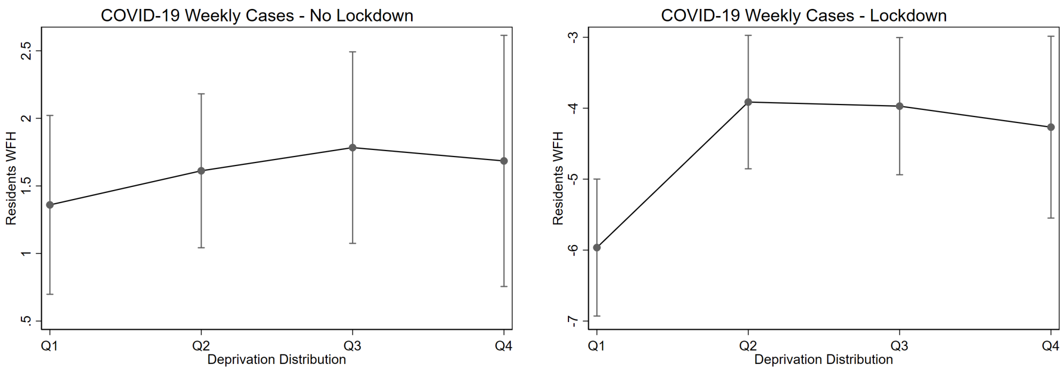
Note: 95% confidence intervals included. Neighbourhood deprivation distribution reported from least deprive (Q1) to most deprived (Q4) MSOAs.

Figure 6: Relationship between residents densities and COVID-19 cases across the neighbourhood deprivation distribution during lockdown and non-lockdown periods.

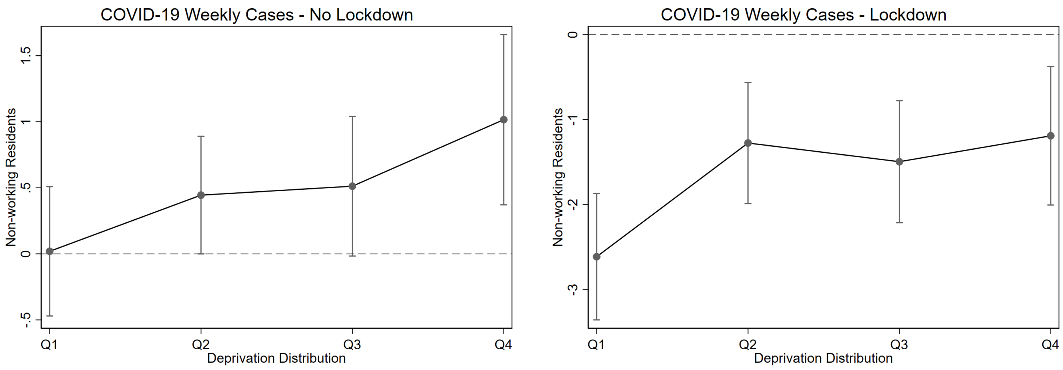
a) Resident keyworkers



b) Resident WFH



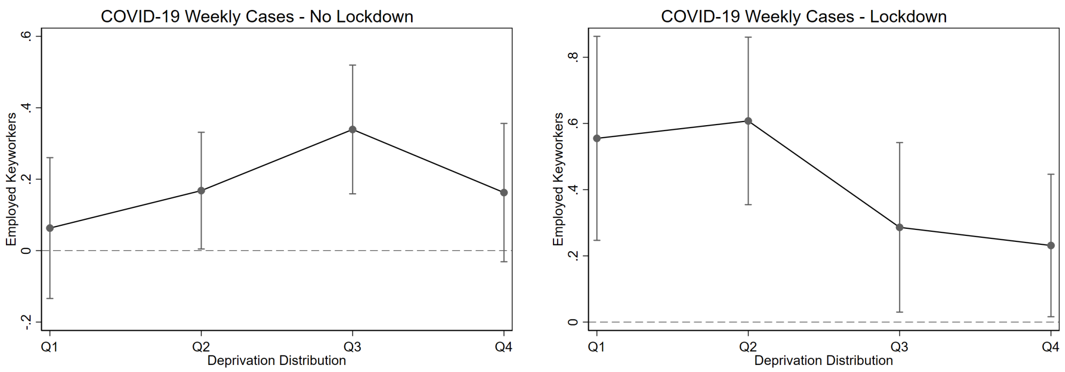
c) Non-working Residents



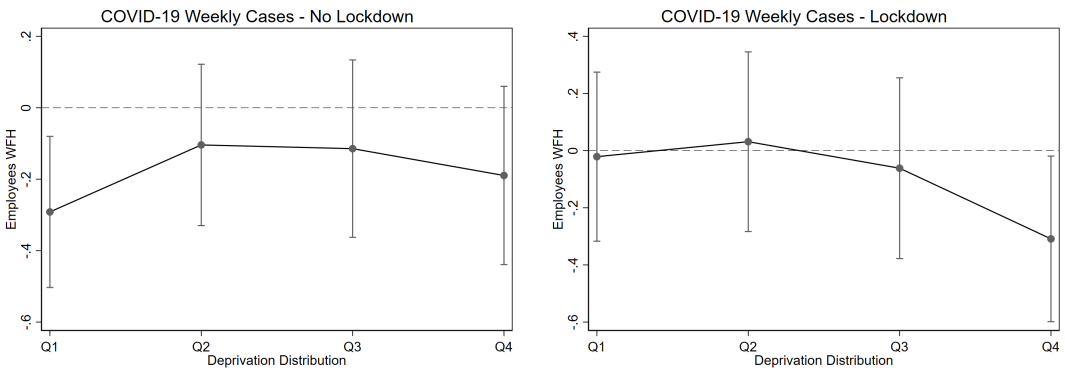
Note: 95% confidence intervals included. Neighbourhood deprivation distribution reported from least deprive (Q1) to most deprived (Q4) MSOAs.

Figure 7: Relationship between employed workers and COVID-19 cases across the neighbourhood deprivation distribution during lockdown and non-lockdown periods.

a) Employed keyworkers



b) Employed WFH



Note: 95% confidence intervals included. Neighbourhood deprivation distribution reported from least deprive (Q1) to most deprived (Q4) MSOAs.

References

Acemoglu, D., Chernozhukov, V., Werning, I., and Whinston, M. D. (2020). *Optimal targeted lockdowns in a multi-group SIR model*, volume 27102. National Bureau of Economic Research.

Allcott, H., Boxell, L., Conway, J. C., Ferguson, B. A., Gentzkow, M., and Goldman, B. (2020). What explains temporal and geographic variation in the early us coronavirus pandemic? Technical report, National Bureau of Economic Research.

Almagro, M. and Orane-Hutchinson, A. (2020). Jue insight: The determinants of the differential exposure to covid-19 in new york city and their evolution over time. *Journal of Urban Economics*, page 103293.

Alvarez, F. E., Argente, D., and Lippi, F. (2020). A simple planning problem for covid-19 lockdown. Technical report, National Bureau of Economic Research.

Armillei, F. and Filippucci, F. (2020). The heterogenous impact of covid-19: Evidence from italian municipalities.

Ascani, A., Faggian, A., and Montresor, S. (2021). The geography of covid-19 and the structure of local economies: The case of italy. *Journal of Regional Science*, 61(2):407–441.

Bailey, D., Clark, J., Colombelli, A., Corradini, C., De Propris, L., Derudder, B., Fratesi, U., Fritsch, M., Harrison, J., Hatfield, M., et al. (2020). Regions in a time of pandemic.

Bartik, A. W., Bertrand, M., Cullen, Z., Glaeser, E. L., Luca, M., and Stanton, C. (2020a). The impact of covid-19 on small business outcomes and expectations. *Proceedings of the National Academy of Sciences*, 117(30):17656–17666.

Bartik, A. W., Cullen, Z. B., Glaeser, E. L., Luca, M., and Stanton, C. T. (2020b). What jobs are being done at home during the covid-19 crisis? evidence from firm-level surveys. Technical report, National Bureau of Economic Research.

Bourdin, S., Jeanne, L., Nadou, F., and Noiret, G. (2021). Does lockdown work? a spatial analysis of the spread and concentration of covid-19 in italy. *Regional Studies*, pages 1–12.

Carozzi, F., Provenzano, S., and Roth, S. (2020). Urban density and covid-19. Technical report, IZA Discussion Papers.

CDC (2020). How covid-19 spreads. updated 13 may 2021. `https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/how-covid-spreads.html` [Accessed on 25 May 2021].

Chang, S., Pierson, E., Koh, P. W., Gerardin, J., Redbird, B., Grusky, D., and Leskovec, J. (2021). Mobility network models of covid-19 explain inequities and inform reopening. *Nature*, 589(7840):82–87.

Chowell, G., Bettencourt, L. M., Johnson, N., Alonso, W. J., and Viboud, C. (2008). The 1918–1919 influenza pandemic in England and Wales: spatial patterns in transmissibility and mortality impact. *Proceedings of the Royal Society B: Biological Sciences*, 275(1634):501–509.

Cole, M. A., Ozgen, C., and Strobl, E. (2020). Air pollution exposure and Covid-19 in Dutch municipalities. *Environmental and Resource Economics*, 76(4):581–610.

Correia, S., Luck, S., and Verner, E. (2020). Pandemics depress the economy, public health interventions do not: Evidence from the 1918 flu. *Available at SSRN: https://ssrn.com/abstract=3561560.*

Courtemanche, C., Garuccio, J., Le, A., Pinkston, J., and Yelowitz, A. (2020). Strong social distancing measures in the united states reduced the covid-19 growth rate: Study evaluates the impact of social distancing measures on the growth rate of confirmed covid-19 cases across the united states. *Health Affairs*, 39(7):1237–1246.

De Fraja, G., Matheson, J., and Rockey, J. (2021). Zoomshock: The geography and local labour market consequences of working from home. CEPR Discussion Papers 15655, C.E.P.R. Discussion Papers.

Desmet, K. and Wacziarg, R. (2021). Jue insight: Understanding spatial variation in covid-19 across the united states. *Journal of urban economics*, page 103332.

Fattorini, D. and Regoli, F. (2020). Role of the chronic air pollution levels in the covid-19 outbreak risk in italy. *Environmental Pollution*, 264:114732.

Garrett, T. A. (2007). Economic effects of the 1918 influenza pandemic. *Federal Reserve Bank of St. Louis*, 26.

Glaeser, E. L., Gorback, C., and Redding, S. J. (2020). JUE insight: How much does COVID-19 increase with mobility? evidence from New York and four other US cities. *Journal of Urban Economics*, page 103292.

Glaeser, E. L. and Kahn, M. E. (2004). Sprawl and urban growth. In Duranton, G., Henderson, V., and Strange, W., editors, *Handbook of Regional and Urban Economics*, volume 4, chapter 56, pages 2481–2527. North Holland, Amsterdam, Netherlands.

Greenhalgh, T., Jimenez, J. L., Prather, K. A., Tufekci, Z., Fisman, D., and Schooley, R. (2021). Ten scientific reasons in support of airborne transmission of sars-cov-2. *The Lancet*, 397(10285):1603–1605.

Guha, A., Bonsu, J., Dey, A., and Addison, D. (2020). Community and socioeconomic factors associated with covid-19 in the united states: Zip code level cross sectional analysis. *medRxiv*.

Henderson, J. V., Nigmatulina, D., and Kriticos, S. (2019). Measuring urban economic density. *Journal of Urban Economics*, page 103188.

Kuebart, A. and Stabler, M. (2020). Infectious diseases as socio-spatial processes: The covid-19 outbreak in germany. *Tijdschrift voor economische en sociale geografie*, 111(3):482–496.

Lewis, D. (2021). Superspreading drives the covid pandemic-and could help to tame it. *Nature*, pages 544–546.

McCann, P. and Vorley, T. (2021). *Productivity and the Pandemic: Challenges and Insights from Covid-19*. Edward Elgar Publishing.

Morawska, L. and Cao, J. (2020). Airborne transmission of sars-cov-2: The world should face the reality. *Environment international*, 139:105730.

Paez, A., Lopez, F. A., Menezes, T., Cavalcanti, R., and Pitta, M. G. d. R. (2020). A spatio-temporal analysis of the environmental correlates of covid-19 incidence in spain. *Geographical analysis*.

Roca, J. D. L. and Puga, D. (2016). Learning by Working in Big Cities. *The Review of Economic Studies*, 84(1):106–142.

Rodríguez-Pose, A. and Burlina, C. (2021). Institutions and the uneven geography of the first wave of the covid-19 pandemic. *Journal of Regional Science*.

Sá, F. (2020). Socioeconomic determinants of covid-19 infections and mortality: evidence from england and wales.

Schiavina, M., Freire, S., and MacManus, K. (2019). GHS population grid multitemporal (1975, 1990, 2000, 2015) R2019A. European Commission, Joint Research Centre (JRC).

Schiavina, M., Freire, S., Rosina, K., Ziemba, L., Marin Herrera, M., Craglia, M., Lavalle, C., Kemper, T., and Batista, F. (2020). ENACT-POP R2020A - ENACT 2011 Population Grid. European Commission, Joint Research Centre (JRC).

Stadnytskyi, V., Bax, C. E., Bax, A., and Anfinrud, P. (2020). The airborne lifetime of small speech droplets and their potential importance in sars-cov-2 transmission. *Proceedings of the National Academy of Sciences*, 117(22):11875–11877.

Travaglio, M., Yu, Y., Popovic, R., Selley, L., Leal, N. S., and Martins, L. M. (2021). Links between air pollution and COVID-19 in England. *Environmental Pollution*, 268:115859.

WHO (2020). Coronavirus disease (covid-19): How is it transmitted? updated 30 apr 2021. `https://www.who.int/news-room/q-a-detail/coronavirus-disease-covid-19-how-is-it-transmitted` [Accessed on 25 May 2021].

Wong, D. W. and Li, Y. (2020). Spreading of covid-19: Density matters. *Plos one*, 15(12):e0242398.

Wu, X., Nethery, R. C., Sabath, B. M., Braun, D., and Dominici, F. (2020). Exposure to air pollution and COVID-19 mortality in the United States. *MedRxiv*.

Zhang, R., Li, Y., Zhang, A. L., Wang, Y., and Molina, M. J. (2020). Identifying airborne transmission as the dominant route for the spread of covid-19. *Proceedings of the National Academy of Sciences*, 117(26):14857–14863.

Appendix

Table A1: Relationship between population, employment density and COVID-19 cumulative cases by MSOA.

	(1) Cumul. Cases	(2) Cumul. Cases	(3) Cumul. Cases	(4) Cumul. Cases
Satellite Density	0.149***			
	(0.0443)			
Daytime Density		-0.0187		
		(0.0118)		
Nighttime Density		0.0742***		
		(0.0214)		
Population Density			0.0419***	0.0402***
			(0.00669)	(0.00679)
Employment Density			0.00577	0.00390
			(0.00542)	(0.00554)
Keyworkers Residents				1.075*
				(0.621)
WFH Residents				-3.779***
				(0.279)
Non-working Residents				-1.494***
				(0.217)
Keyworkers Employees				0.352***
				(0.0633)
WFH Employees				0.0137
				(0.0750)
Population	1.080***	1.021***	1.035***	1.038***
	(0.0151)	(0.0190)	(0.0150)	(0.0148)
Employment	0.00967**	-0.000153	0.00751	0.0131**
	(0.00476)	(0.00504)	(0.00632)	(0.00617)
Share Elderly	-0.189***	-0.342***	-0.0217	-0.206***
	(0.0644)	(0.0632)	(0.0650)	(0.0677)
Share Children	-0.207	-0.231	-0.247*	-0.418***
	(0.144)	(0.143)	(0.145)	(0.138)
Share White	-0.271***	-0.274***	-0.242***	-0.178***
	(0.0329)	(0.0327)	(0.0332)	(0.0332)
House Crowding	0.120***	0.111***	0.149***	0.127***
	(0.0213)	(0.0211)	(0.0223)	(0.0220)
Deprivation	0.307***	0.302***	0.310***	0.113***
	(0.0285)	(0.0284)	(0.0286)	(0.0406)
Pollution	0.0867***	0.0855***	0.0584***	0.0616***
	(0.00627)	(0.00625)	(0.00653)	(0.00643)
Cases Spatial Lags	0.000472	0.000401	-0.000671	0.000802
	(0.000723)	(0.000738)	(0.000702)	(0.000698)
LAD FE	Y	Y	Y	Y
Observations	6,789	6,789	6,789	6,789
R-squared	0.877	0.877	0.881	0.889

Notes: Robust standard errors in parentheses clustered at the MSOA level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Relationship between population, employment density and COVID-19 monthly deaths by MSOA.

	(1) Month.Deaths	(2) Month.Deaths	(3) Month.Deaths	(4) Month.Deaths
Satellite Density	0.0515*			
	(0.0312)			
Daytime Density		-0.0252***		
		(0.00888)		
Nighttime Density		-0.00195		
		(0.0159)		
Population Density			0.0229***	0.0239***
			(0.00407)	(0.00406)
Employment Density			0.0105***	0.00846**
			(0.00368)	(0.00372)
Keyworkers Residents				0.789*
				(0.432)
WFH Residents				-0.239
				(0.212)
Non-working Residents				0.000399
				(0.156)
Keyworkers Employees				0.251***
				(0.0536)
WFH Employees				-0.156***
				(0.0583)
Population	0.404***	0.393***	0.371***	0.369***
	(0.0112)	(0.0139)	(0.0114)	(0.0115)
Employment	0.0210***	0.0177***	0.00975**	0.0130***
	(0.00360)	(0.00399)	(0.00453)	(0.00453)
Share Elderly	1.728***	1.639***	1.814***	1.758***
	(0.0483)	(0.0475)	(0.0486)	(0.0509)
Share Children	0.363***	0.341***	0.343***	0.270***
	(0.0860)	(0.0860)	(0.0868)	(0.0860)
Share White	-0.210***	-0.207***	-0.184***	-0.174***
	(0.0236)	(0.0235)	(0.0235)	(0.0245)
House Crowding	0.0162	0.0156	0.0346***	0.0291**
	(0.0111)	(0.0111)	(0.0115)	(0.0118)
Deprivation	0.371***	0.362***	0.368***	0.346***
	(0.0219)	(0.0218)	(0.0219)	(0.0309)
Pollution	0.0315***	0.0293***	0.00820	0.00983*
	(0.00468)	(0.00470)	(0.00512)	(0.00513)
LAD FE	Y	Y	Y	Y
Time FE	Y	Y	Y	Y
LAD*Time FE	Y	Y	Y	Y
Observations	95,046	95,046	95,046	95,046
R-squared	0.599	0.598	0.600	0.600

Notes: Robust standard errors in parentheses clustered at the MSOA level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Time fixed-effects at month level.

Table A3: Relationship between population, employment density and monthly excess mortality by MSOA.

	(1) Excess Mortality	(2) Excess Mortality	(3) Excess Mortality	(4) Excess Mortality
Satellite Density	0.0716			
	(0.212)			
Daytime Density		0.0120		
		(0.0808)		
Nighttime Density		-0.0274		
		(0.125)		
Population Density			0.0253	0.0316
			(0.0336)	(0.0337)
Employment Density			0.0576*	0.0487
			(0.0299)	(0.0305)
Keyworkers Residents				4.108
				(3.232)
WFH Residents				1.630
				(1.608)
Non-working Residents				1.444
				(1.152)
Keyworkers Employees				0.898**
				(0.396)
WFH Employees				-0.734*
				(0.450)
Population	1.254***	1.207***	1.207***	1.187***
	(0.0849)	(0.103)	(0.0873)	(0.0902)
Employment	0.0979***	0.0678**	0.0419	0.0541
	(0.0276)	(0.0320)	(0.0359)	(0.0361)
Share Elderly	4.786***	4.539***	5.023***	4.857***
	(0.339)	(0.334)	(0.340)	(0.354)
Share Children	1.504***	1.512***	1.471***	1.253**
	(0.507)	(0.514)	(0.510)	(0.543)
Share White	-0.444***	-0.445***	-0.383**	-0.351**
	(0.162)	(0.162)	(0.162)	(0.169)
House Crowding	0.0897	0.0722	0.137*	0.116
	(0.0704)	(0.0720)	(0.0714)	(0.0764)
Deprivation	0.926***	0.897***	0.917***	0.917***
	(0.152)	(0.152)	(0.152)	(0.217)
Pollution	0.0883**	0.0852**	0.0198	0.0270
	(0.0351)	(0.0352)	(0.0382)	(0.0385)
LAD FE	Y	Y	Y	Y
Time FE	Y	Y	Y	Y
LAD*Time FE	Y	Y	Y	Y
Observations	67,890	67,890	67,890	67,890
R-squared	0.276	0.276	0.276	0.277

Notes: Robust standard errors in parentheses clustered at the MSOA level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Time fixed-effects at month level. Excess mortality calculated for 2020 months in respect to 2018-2019 average.