Social Science & Medicine 270 (2021) 113655

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: http://www.elsevier.com/locate/socscimed

COVID-19 does not stop at open borders: Spatial contagion among local authority districts during England's first wave

Denise Laroze [a], Eric Neumayer [b], Thomas Plümper [c,*]

[a] Centre for Experimental Social Sciences and Department of Management, Universidad de Santiago de Chile, Santiago, Chile
[b] Department of Geography & Environment, London School of Economics and Political Science (LSE), London, UK
[c] Department of Socioeconomics, Vienna University of Economics and Business, Vienna, Austria

ARTICLE INFO

Keywords:
SARS-CoV-2
COVID-19
Infectious diseases
Pandemic
Spatial contagion
Spatial dependence
Lockdown
English local authority districts

ABSTRACT

Infectious diseases generate spatial dependence or contagion not only between individuals but also between geographical units. New infections in one local district do not just depend on properties of the district, but also on the strength of social ties of its population with populations in other districts and their own degree of infectiousness. We show that SARS-CoV-2 infections during the first wave of the pandemic spread across district borders in England as a function of pre-crisis commute to work streams between districts. Crucially, the strength of this spatial contagion depends on the phase of the epidemic. In the first pre-lockdown phase, the spread of the virus across district borders is high. During the lockdown period, the cross-border spread of new infections slows down significantly. Spatial contagion increases again after the lockdown is eased but not statistically significantly so.

1. Introduction

A virus survives by changing hosts and infectious diseases are therefore spatially contagious. Everything else being equal, the more interactions a person has with infected people, the more likely the person becomes infected. Hamidi's (2020) analysis of metropolitan counties in the United States suggests that connectivity matters more than density in the spread of the COVID-19. This very nature of an infectious disease not only holds at the individual level, it also holds at the regional level. In an epidemic, regional units are spatially dependent. The more people from one region meet and interact with people from other regions, the more infected people from the other region can potentially spread the infection to the population of the first region, and vice versa. As a consequence, the number of people in region *i* who become infected depends not only on the properties of region *i* but also on the interaction between people from region *i* and those of other regions, call them *j*. The stronger the social ties are between two regions, the easier it becomes for the virus to spread between these two regions and therefore, the stronger will be spatial contagion.

This simple logic of epidemics and pandemics is surprisingly often ignored in epidemiological analyses. To be fair, there is a long tradition of modelling the spread of epidemics as a spatial process based on information regarding the connectedness of places (see, for example, Bartlett 1957; Cliff et al. 1981; Elliot and Wartenberg, 2004; Balcan et al., 2010; Lawson et al., 2016). Nevertheless, as O'Sullivan et al. (2020: 973) point out, "… it is perhaps surprising that fundamental geographic processes have largely been excluded from SIR [susceptible-infected-recovered] models and their variants". Notable exceptions, besides the simulation analyses cited above, include Klepac, Kissler and Gog (2018), who modelled how an influenza outbreak spreads across the United Kingdom (UK) based on movement and contact data they collected from custom mobile phones. There also exist epidemiological simulations of the spread of COVID-19 in Italy (Gatto et al., 2020), New York City (Munshi et al. 2020), and New Zealand (O'Sullivan et al., 2020). These models are very useful in modelling a pandemic and modelling which global or spatially localized policy interventions can affect the presumed course of a pandemic in what way.

Our study departs from these models and makes an original contribution by introducing spatial contagion into an econometric model of actual observed patterns of the spread of the pandemic. In other words, rather than simulating the course of the pandemic, we estimate the degree of actual spatial contagion as it occurred in the first wave of the pandemic. Our research provides information on the size that relevant spatial parameters should have in simulations of the SARS-CoV-2

* Corresponding author.
E-mail addresses: denise.laroze@usach.cl (D. Laroze), e.neumayer@lse.ac.uk (E. Neumayer), pluemper@gmx.at (T. Plümper).

https://doi.org/10.1016/j.socscimed.2020.113655
Received in revised form 31 October 2020; Accepted 22 December 2020
Available online 25 December 2020
0277-9536/Crown Copyright © 2020 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

pandemic. However, we also hope that our paper paves the ground for more analyses of spatial effects in epidemiological research in the future. To the best of our knowledge, with one exception where a spatial analysis forms one of many robustness tests (Alipour et al., 2020a,b), the econometric analyses of the spread of SARS-CoV-2 in a cross-section of geographical units of analysis exclusively regress the number or rates of infections or mortality in region *i* on a number of covariates measured in the very same region *i*. For example, these analyses address the statistical association between health preconditions and COVID-19 mortality rates (Notari and Torrerieri, 2020; Likasse, 2020), air pollution and mortality (Conticini et al., 2020; Martelletti and Martelletti 2020; Travaglio et al., 2020), the ethnic composition of district populations and infection rates and mortality (Goldstein and Atherwood 2020; Gross et al., 2020; Rentsch et al., 2020; Cohen 2020) and between socio-economic covariates and infection rates (Han et al., 2020; Plümper and Neumayer 2020).

These analyses share the strong implicit assumption that infection and mortality numbers and rates in one region are independent of infection and mortality numbers and rates in other regions. Of course, this assumption can be plausible if regional units, for example local authority districts, are geographically located far apart and little direct interaction takes place between their populations. The virus will spread much more easily from Cornwall to West Devon, from Northumberland to Carlisle, or from Islington to Camden, which are all geographically proximate, than it will spread from Northumberland in the very North-East of England to Cornwall in England's South-West, though transmission across long distances does of course also occur, if less frequently than transmission across shorter distances. As many politicians have put it, the virus does not stop at borders – at least not, one should add, at open borders. We also know that the virus hits people with preconditions especially hard, and there is evidence that these preconditions are spatially clustered in the UK (Dearden et al., 2019) and the United States (Dobis et al., 2020; Amram et al., 2020). Thus, the assumption that infections in districts are spatially independent of each other is a stark simplification.

Our spatial analysis of SARS-CoV-2 adds to our knowledge of the spatial processes underpinning the spread of the virus. Previous research has demonstrated how the virus spreads on a global scale. It took the virus only 19 days from the first known death associated with COVID-19 on 11 January to the WHO declaring a 'public health emergency of international concern' on 30 January. The rapid global spread of the virus has been facilitated by air travel (Chinazzi et al., 2020). Previous research has also spatially modelled or simulated how COVID-19 spreads across geographical units in Italy (Gatto et al., 2020), New York City (Munshi et al., 2020), and New Zealand (O'Sullivan et al., 2020). Ski tourism spread the virus across Europe from just a few Alpine villages (Plümper and Neumayer, 2020; Felbermayer et al., 2020). At the same time, it has also been shown that the distribution of infections shows local persistency – coronavirus hotspots, often triggered by local events, tend to remain fairly stable over time (Plümper and Neumayer, 2020).

In this article, we complement these findings about the spatial dynamics of Sars-CoV-2 by analyzing the degree to which actual and observed infection rates in local authority districts in England spatially depend on infection rates in other districts, as a function of historical commuter flows between these districts. Naturally, such commuter flows are larger in geographically more proximate districts and converge to zero between districts that are geographically very distant from each other. According to Tobler's first law of geography, "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970: 236). We test Tobler's law based on SARS-CoV-2 infection rates. With Tobler, we expect to find spatial dependence between districts as a function of commuter flows between them. This is our first major contribution to the emerging literature on COVID-19.

Our second major contribution relates to the dynamics of spatial dependence over time. The lockdown that the UK government imposed on March 23, 2020 has significantly reduced human interactions between districts following the government directive of 'stay home unless your travel is essential', together with the shutting down of shops, bars, restaurants, entertainment venues and many workplaces. We provide evidence that the lockdown both substantively and statistically significantly reduces spatial contagion among districts. The degree of spatial contagion rises again, albeit not statistically significantly so, after the lockdown is eased on May 11, 2020 with the 'stay home unless your travel is essential' directive replaced with a 'stay alert and work from home if you can' directive, though shops were only reopened on June 15, 2020, and restaurants, bars, entertainment and hospitality venues are all still closed by the end of our estimation period, June 26, 2020, only opening again on 4 July with social distancing restrictions in place.

2. Methodology

Spatial dependence is typically defined as the theoretically expected dependence of outcomes in unit *i* on outcomes or determinants of outcomes in unit *j*. Spatial dependence requires more than spatial clustering in outcomes or determinants of outcomes amongst units of analysis. In the early days of spatial econometrics, this distinction between spatial dependence proper spatial clustering was somewhat blurred by a definition of spatial dependence as non-zero "correlation across cross-sectional units" (Anselin et al., 2008: 627). Spatial clustering only requires that close units are similar in outcomes or the properties that influence outcomes, not that the outcome in unit *i* depends on the outcome or the determinants of the outcome in units *j*. Thus, areas can be spatially clustered in the strict absence of proper spatial dependence. Just like correlation is not causality, spatial clustering is not spatial dependence.

Spatial econometric models have been developed to account for both spatial dependence and spatial clustering. These models can be divided into three types. Spatial error models regress the outcome variable in one unit on the regression residuals in other units, usually in geographically contiguous units. It is typically employed to account for spatial clustering amongst units of analysis. So-called spatial autoregressive (SAR) or spatial-y models regress the outcome variable in one unit on the outcome variable in other units or what is known as the spatial lag. And spatial-x models regress the outcome variable in one unit on one or more explanatory variables in other units of analysis (Cook et al. 2020; Wimpy et al. 2020). A combination of model types is also possible. For example, the spatial autoregressive spatial error (SARAR) model combines spatial lag with spatial errors (Wang, 2020).

Pandemics provide a clear example for spatial dependence of the spatial autoregressive (spatial lag) type: the number of infections in unit *i* depend on the number of infections in unit *j* weighted by the strength of interactions between these units. Accordingly, modelling pandemics by a spatial autoregressive model is the best choice. Yet, this is not to say that pandemics are free of spatial clustering in addition to spatial dependence. Spatial clustering occurs because everything resembles everything else, but near things are more similar than distant things; a phenomenon that we have dubbed the second law of geography (Neumayer and Plümper, 2016: 180). This logic applied to pandemics implies that two closely related districts are likely to have similar properties that cause higher or lower than average infection rates in both districts. For example, two neighboring districts may be far away from both the population center of a country and the hotspots of the pandemic, and therefore both have lower than average infection rates. It is possible to capture this similarity by adding the necessary covariates to the empirical model, or by controlling for unobserved spatially clustered factors via a spatial error model, which includes the spatially weighted residuals in districts *j*. Of course, a combination of both approaches is also possible, in which case the spatial-error component of the model just controls for the spatial clustering unaccounted for by the substantively interesting explanatory variables. Clearly, if researchers are interested in these explanatory variables, this is the preferred

specification.

We therefore contend that pandemic processes are best modelled as a spatial autoregressive spatial error, or SARAR model, with additional explanatory variables included: infections in units *j* lead to new infections in unit *i* as a function of the degree of interaction between *i* and *j*, but despite modelling, fairly well understood determinants of infections geographically contiguous districts may nevertheless exhibit unexplained spatial clustering.

We estimate the following model:

$$y_{it} = \rho \sum_{j \neq i} c_{ij} y_{j,t-1} + \mu y_{it-1} + \beta X_i + T_t + u_{it}$$

$$u_{it} = \lambda \sum_{j \neq i} w_{ij} u_{jt} + \varepsilon_{it}$$

where y_{it} is the rate of new infections, i.e. new infections per 100,000 people, in week t in district *i* with $i = 1, 2, .., N$, $j = 1, 2, .., N$ and $t = 6,..,$ 25 calendar week of 2020, with week 6 the week in which the first infections were registered in the UK. y_{it-1} is the lagged dependent variable, i.e. new infection rates one week prior in district *i*. X_i represents variables explaining new infection rates in district *i* other than the temporal lag and the spatial lag. T_t represents dummy variables for each week (week fixed effects) and $\sum_{j \neq i} c_{ij} y_{j,t-1}$ represents the spatial lag. For this spatial lag, the connectivity matrix c_{ij} is based on the census travel to work data and measures the historical sum of commuter flows from district *i* to *j* and from districts *j* to *i* since infections can spill over both by individuals from district *i* who commute to other districts *j* and bring the virus back to their home district, and by individuals from district *j* who commute into district *i*. Formally, we weight infections in districts *j* by commuter flows and we weight populations in district *j* by commuter flows, and then divide the former by the latter to arrive at the spatial lag variable. Note that historical means these are commuter flows back in 2011, the latest available data. There is measurement error to the extent that commuter flows have changed since before the lockdown, but note it is not a measurement error that commuter flows changed with the lockdown since the weighting variable must be exogenous to the policy shock (lockdown) studied.

According to the United States Centers for Disease Control and Prevention (CDC), available scientific evidence suggests that the vast majority of infected people remain infectious for up to about 7–12 days (https://www.cdc.gov/coronavirus/2019-ncov/hcp/duration-isolation.html). We therefore use a one-week lag for infections in other districts, as indicated by *t-1*. Likewise, we include the temporally lagged (by one week) dependent variable to account for the fact that the rate of new infections one week prior is a good predictor of the new infection rate in week *t*. If we were to change the definition of both the spatial and temporal lags to include the sum of new infections from two weeks prior, rather than only one week prior, then the results reported below across the three periods of pre-lockdown, during lockdown, and post-lockdown ease are qualitatively very similar.

The week fixed effects T_t flexibly take out the common trend, that is, changes in new infection rates that are common to all districts in England. They can also absorb the effect that changes to the availability of testing for SARS-CoV-2 has on reported new infection rates as long as these changes affect all English districts approximately equally. Taking out the common trend is not innocent as some of this common trend will be due to spatial contagion. Our specification choice should therefore be regarded as providing a conservative estimate of the degree of spatial contagion, with actual spatial contagion likely to be under-estimated. The spatial error term u_{it} consists of the error term in spatially contiguous districts, as indicated by $\sum_{j \neq i} w_{ij}$, which is the standard weighting matrix for spatial error terms (Drukker et al., 2013), plus an error term presumed to be identically and independently distributed (ε_{it}).

We estimate this equation with spatial two-stage least squares (spatial-2SLS) in which, as is common procedure, the spatial lag $\sum_{j \neq i} c_{ij} y_{j,t-1}$ is instrumented for with the spatial-X variables $\sum_{j \neq i} c_{ij} X_j$ and $\sum_{j \neq i} w_{ij} X_j$ (Drukker et al., 2013). Standard errors are bootstrapped with 500 replications, given that we have to do the instrumentation by hand in order to exclude the temporal lag y_{it-1} from entering the first stage of the spatial-2SLS estimation, since the temporal lag cannot reasonably be argued to be exogenous.

To account for the characteristics of the districts that one may expect to have an effect on infection rates, we include a series of socio-economic and demographic control variables in the model. These controls have been found to be associated with SARS-CoV-2 infections in other studies. First, we include the proportion of the population that lives in rural areas, as defined by the 2011 census. Previous research has found that rural areas have lower infection rates. Cohen (2020: 6), for example, shows that "rural U.S. counties account for less of the COVID-19 outbreak than their share of the population". He believes that these findings reflect "the epidemic's initial explosion in coastal urban areas and the greater propensity for transmission in places with greater population density". In a similar study, Goldstein and Atherwood (2020) show that controlling for age and place, on average, Black people have an 80 percent and Hispanic people a 54 percent higher risk of dying from COVID-19 than White and Asian individuals. These mortality differences are higher during the COVID-19 epidemic than in non-epidemic years, which Goldstein and Atherwood take as evidence that underlying health disparities do not fully account for the COVID-19 mortality differences. The authors suggest that factors such as "the differential risk of infection related to exposure at work, in transportation and at home as well as differential access to healthcare" (Goldstein and Atherwood 2020: 9) may also explain the mortality differences along ethnic lines. We concur that individuals from ethnic minorities disproportionally suffer from health inequalities caused by socio-economic determinants, as highlighted again prominently in the recent 10-year revisit of the original 'Marmot report' (Marmot et al., 2020). For example, individuals from ethnic minorities are more likely to be living in smaller and more crowed accommodation, earning lower wages from possibly multiple jobs that render them more exposed to the virus, which also makes it much more difficult for them to self-isolate if they or their close contacts exhibit symptoms.

To capture some of the socio-economic determinants of health inequalities, we include income per employee and the proportion of people in a district that fall into the highest decile of the Index of Multiple Deprivation (IMD) nationally as control variables (results are very similar if we use the alternative Townsend Deprivation Index instead). Poorer and more socially deprived people cannot shelter themselves from the risk of infection to the same extent as richer people. They are less able to work from home offices, depend more on public transportation (Harris 2020), and live in smaller houses and apartments. These variables will to some extent capture why Black, Asian and Minority Ethnic (BAME) people are disproportionately affected. However, since it is impossible to comprehensively control for all socio-economic determinants of health inequalities simultaneously, we also control for the higher risk of infection that ethnic minority groups may suffer from beyond what we can control for by directly including as a control variable the population share of Black, Asian and Minority Ethnic (BAME) people in the district's population. By controlling for some socio-economic differences across districts, the BAME population share variable estimates, an effect that goes beyond the fact that poorer and more socially deprived districts are also typically districts with a higher share of BAME population. All data for the dependent, spatial and explanatory variables were sourced from various official UK government websites. Data was collected during May and October 2020 from official UK government websites. The 2011 Census information on income and ethnic composition of districts were obtained from the Office of National Statistics https://www.nomisweb.co.uk/, the Index of Multiple Deprivation was obtained from https://www.gov.uk/government/statistics/e

nglish-indices-of-deprivation-2019, and the coronavirus case data were obtained from the official https://coronavirus.data.gov.uk/website. Local Authority Urban-Rural classification was obtained from https://www.gov.uk/government/statistics/local-authority-rural-urban-classification, and commuter data come from http://wicid.ukdataservice.ac.uk/. Our sample covers 314 districts, that is, all 317 first-tier local authorities in England minus three districts for which we have no data on the explanatory variables, namely the City of London, the Isle of Scilly and the Isle of Wight.

3. Results

Table 1 presents our estimation results. We estimate results for three distinct periods: the seven weeks before the lockdown was imposed on March 23, 2020 (column 1), the seven weeks during the lockdown (column 2) and the seven weeks after the lockdown was eased on May 11, 2020 (column 3). Focusing first on the period before lockdown, we find a high degree of autoregression, as indicated by the coefficient of the temporally lagged variable. One more person per 100,000 people infected in the week prior is predicted to result in 1.61 more individuals infected per 100,000 on average in the week after. With regards to the control variables, we find that rates of new infections are lower in richer districts and higher in districts with a larger ethnic minority population share. Most importantly, we find a substantively large degree of spatial dependence amongst districts. Weighted by historic commuter flows between districts, we find that one more person per 100,000 people newly infected in linked districts is predicted to increase the rate of new infections by 0.37 more individuals infected per 100,000 people on average in the next week.

In the strict lockdown period (column 2), we find results much altered, as one would expect given the drastic policy intervention that came with an unprecedented imposition of restrictions on economic and social life. The degree of autoregression more than halves, and reduces to one more person per 100,000 people being infected in the prior week predicted to increase on average the rate of new infections by 0.74 individuals per 100,000 people. The difference to the pre-lockdown period is highly statistically significant (z-score 7.80, $p < 0.01$). More importantly, and in line with Plümper and Neumayer (2020), new infection rates after lockdown become differently socially, economically and demographically stratified. This is because individuals differ in their ability to socially distance themselves and work from home (Alipour et al., 2020a,b), which along with path dependency becomes the main driver of the pandemic in the second phase.

Table 1
Estimation results.

	Pre-lockdown 03/02 to 22/03	During lockdown 23/03 to 10/05	Post-lockdown ease 11/05 to 25/06
Rate of new infections (t-1)	1.625**	0.730**	0.558**
	(0.112)	(0.0122)	(0.0328)
% population in rural areas	0.000685	−0.0263**	−0.00243
	(0.00282)	(0.0102)	(0.00330)
Per capita income	−5.73e-05**	−3.35e-05	2.23e-05
	(2.01e-05)	(3.49e-05)	(1.24e-05)
% population who claim benefits	0.152	2.255**	0.359*
	(0.0780)	(0.486)	(0.156)
% population who are BAME	0.0456**	−0.0244	0.0139
	(0.0129)	(0.0290)	(0.0146)
Spatial lag (t-1)	0.408**	0.140**	0.213**
	(0.0749)	(0.0314)	(0.0586)
Spatial error	0.754**	0.698**	0.377**
	(0.0954)	(0.0552)	(0.0819)
Observations	2121	2121	2121
Number of districts	303	303	303

Note: All estimations include week fixed effects. Bootstrapped standard errors with 500 replications in parentheses. *, ** statistically significant at 0.05 and 0.01 level, respectively.

We find that richer districts and districts that are predominantly rural experience lower rates of new infections during lockdown, whereas socially deprived districts, as measured by the population share of social welfare benefits claimants, experience higher rates of new infections than others. There is no longer a statistically significant association with the BAME population share. Crucially, we find that the degree of spatial dependence amongst districts, as weighted by historic commuter flows between districts, not only more than halves but is also statistically significantly lower than before lockdown (z-score 2.24, $p < 0.02$). Lockdown has managed to substantively and statistically significantly reduce cross-district spatial contagion, but has not eliminated it, as some people still cross district borders for work, shopping in food stores and pharmacies and for other purposes. We thus find that one more person per 100,000 people newly infected in linked districts is predicted to increase in the next week the rate of new infections by 0.19 more individuals infected per 100,000 people on average, even after lockdown was imposed. Lastly, we continue to find statistically significant residual spatial clustering amongst geographically contiguous districts.

In the third period of our estimations after lockdown was eased on 11 May, we find that the coefficient of the temporal lag, which indicates how infections within a district results in more new infections the week after, continues to fall. Now, one more infected person per 100,000 people is predicted to result in 0.56 newly infected individuals per 100,000 people the week after (results reported in column 3). The degree of spatial contagion rises again after the lockdown has been eased and people start being more mobile again. Having dropped from 0.37 to 0.19 from pre-lockdown to the period during lockdown, it goes back up again slightly to 0.21. Even after lockdown has been eased, the degree of spatial dependence is statistically significantly lower compared to the pre-lockdown period (z-score 1.71, $p < 0.05$). The difference between the lockdown period and the post-lockdown period is, however, not statistically significant, suggesting that lockdowns have a lasting effect.

4. Discussion

We have tested how infections are auto-regressive within districts, as approximated by the coefficient of the temporally lagged variable, but also spatially auto-regressive, as indicated by the coefficient of the spatial lag variable. A lagged dependent variable with a positive coefficient increases the long-term effect of the substantively interesting variables included in the model. The differences between immediate effects and long-term effects become larger the larger the coefficient of the lagged dependent variable. Positive coefficients of the temporal lag variable indicate that path dependency exists within districts and a value of the temporally lagged variable above 1.0 indicates that new infections are exponentially increasing, which is often the case during the early phase of a pandemic when neither individuals nor governments have adequately responded to the danger posed by a hitherto unknown and highly infectious disease. Our estimates suggest that even with the week fixed effects taking out the common temporal trend, the pandemic was indeed explosive in phase 1, before the lockdown, but became stationary once people adjusted their behavior and the lockdown was implemented.

If the coefficient of the lagged dependent variable is larger than 1, as was the case in the early weeks of the epidemic in the UK, the effect of all variables goes to infinity in the very long run. Clearly, this is not possible. No epidemic can follow an exponential path ad infinitum. Eventually, the majority of people are either immune or dead and the infectious disease runs out of steam. Exponential growth is, however, possible in the early phase of a pandemic during the seven weeks pre-lockdown. Moreover, the cross-district spatial contagion adds a very important additional amplifier to a temporally autoregressive process that is close to one or above one (explosive), since new infections move backwards and forwards between related districts. Thus, one additional individual per 100,000 people infected in other districts, to which the district under observation is linked via historical commuter flows, raises

new infections by a staggering total of 241 newly infected individuals per 100,000 people on average by week 7 at the end of the pre-lockdown period. Counterfactually, in the absence of any spatial contagion one additional newly infected person would only result in 72 newly infected individuals per 100,000 people by week 7. In other words, the high degree of spatial contagion during the pre-lockdown period roughly triples the pandemic's dynamics relative to a counterfactual scenario with assumed absence of spatial dependence – such is the power of exponential growth feeding back and forth across district borders.

Our results show that the coefficient of the lagged dependent variable fell below 1.0 in what we call the second phase, that is, after individuals reduced their social interactions and the government implemented a lockdown. Not only does this new temporal dynamic break the exponential within-district growth; together with a much smaller coefficient of the spatial lag variable, it also means that once the self-reinforcing effect of new infections moving backwards and forwards between related districts is taken into account, the long-run spatial contagion effect of 5.2 more newly infected individuals per 100,000 people by week 7, the end of the lockdown period, is not dramatically larger in absolute terms than the short-run spatial contagion effect of 0.19 more individuals per 100,000 people infected. Qualitatively similar results apply to the third phase after lock-down has been eased. Spatial contagion rises again, but together with the again lower degree of temporal auto-regression within districts, the long-run spatial contagion effect of 2.2 is not that much higher than the short-term spatial contagion effect of 0.21.

Interestingly, the ethnic minority population share is no longer a variable with a statistically significant effect after lockdown has been imposed, but one should not forget that this variable is moderately positively correlated with social deprivation ($r = 0.16$) and highly negatively correlated with the rural population share within a district ($r = -0.53$). A detailed and comprehensive analysis with regards to whether ethnic minorities suffer more from COVID-19 above and beyond any socio-economic and demographic factors is beyond our analysis. A study that perhaps gets closest to identifying these structures analyzes mortality differences across ethnicities amongst US veterans (Rentsch et al., 2020). Their findings suggest that socioeconomic factors play a major role in these mortality differences across different ethnicities in the US, but ethnicity adds very little. The authors warn, however, against generalizing from veterans to the broader society. In the UK, a careful analysis by Rose et al. (2020: 1) finds that after controlling for income and socioeconomic differences across English districts, a one percentage point increase "in the proportion of the population from BAME backgrounds was associated with a 1% increase in the COVID-19 mortality rate". Accordingly, evidence suggests that controlling for socio-economic factors is insufficient and ethnicity influences infection rates for reasons others than socio-economic and demographic differences between ethnicities.

5. Conclusion

Our analysis was motivated by the conspicuous absence of econometric studies that take into account how regional units spatially depend on each other. Naturally, there is no reason to expect that the SARS-CoV-2 virus stops at open district borders. In qualitative terms, the first result from our empirical analysis is therefore entirely unsurprising: there is spatial contagion among local authority districts in England. A higher new infection rate in other districts linked to the district under observation via historical commuter flows raises the rate of new infection the week after. Our analysis shows a high degree of this spatial dependence together with an explosive degree of temporal dependence within districts before the national lockdown was imposed on March 23, 2020. In modern times, there is faster and stronger spatial contagion than, for example, during the times of the plague. Mankind is more mobile and people travel longer distances at much higher speed. In reality, spatial contagion will be even stronger than our conservative estimate suggests since we only model connectivity amongst districts based on historical work commuter flows. There is, of course, additional interaction among people across district boundaries in the form of, for example, family visits and leisure travel. This is only indirectly and imperfectly captured in our research design to the extent that commuter traffic and all other mobility between districts are correlated. Unfortunately, data on all other mobility is missing, and so we cannot distinguish between the spatial effects of different types of cross-district mobility.

The lockdown imposed on 23 March by the UK government tried to artificially reduce this mobility both within and across districts. Substantively, the most interesting result that we have obtained directly tells us that the lockdown has, to a significant extent, achieved this double objective as indicated by the significant decline in the sizes of the temporal and the spatial lag coefficients. These results indicate that measures of social distancing not only manage to break the exponential growth of new infections within a district but also drastically reduce the spread of the disease across district borders.

Lockdown has not managed to eliminate spatial contagion, however, which has also risen again, albeit not to pre-lockdown levels after the lockdown was eased on 11 May. Open borders do not stop the spread of COVID-19, not even during lockdown, and certainly not after the easing of lockdown. However, in combination with an autoregressive factor of below one, the lower the spatial dynamics of a pandemic, the easier it will be to isolate the virus at a local level. This result supports the shift to local policies aimed at controlling the pandemic that could be observed in late spring 2020 in countries that were more successful than the UK; for example in Germany. Therefore, our results support decentralized control policies once the worst is over and weekly new infection numbers have declined to low 2-digit numbers per 100,000 people. We expect that decentralized, federal countries will find it easier to organize local control strategies, but one would hope that more centralized nation-states will eventually overcome the institutional disadvantages they face and also manage to successfully employ local strategies to keep the pandemic at bay. Unfortunately, at the time of finishing this article, it has become all too clear that the strategy which countries tried to employ after the first wave of the pandemic appeared to be under control was insufficiently successful. England, together with its European peers, is now well and truly in the second wave of the pandemic, an analysis of which is beyond the scope of this article, which was entirely focused on spatial contagion during England's first wave.

Credit author statement

All three co-authors contributed equally to the research project.

Acknowledgments

Denise Laroze would like to acknowledge the Grant Fondecyt de Iniciación Number 11180281, Chile.

References

Alipour, J.V., Fadinger, H., Schymik, J., 2020a. My Home Is My Castle – the Benefits of Working from Home during a Pandemic Crisis: Evidence from Germany (Working paper).

Alipour, J.V., Falck, O., Schüller, S., S, 2020b. Germany's Capacity to Work from Home. Institute for Labour Economics, Bonn. Discussion paper 13152.

Amram, O., Amiri, S., Lutz, R.B., Rajan, B., Monsivais, P., 2020. Development of a Vulnerability Index for Diagnosis with the Novel Coronavirus, COVID-19. Health & Place, Washington State, USA, p. 102377.

Anselin, Luc, Le Gallo, Julie, Hubert, Jayet, 2008. Spatial panel econometrics. In: Matyas, László, Sevestre, Patrick (Eds.), The Econometrics of Panel Data, Fundamentals and Recent Developments in Theory and Practice, third ed. Springer-Verlag, Berlin Heidelberg, pp. 624–660.

Balcan, D., Gonçalves, B., Hu, H., Ramasco, J.J., Colizza, V., Vespignani, A., 2010. Modeling the spatial spread of infectious diseases: the GLobal Epidemic and Mobility computational model. J. Comput. Sci. 1, 132–145.

Bartlett, M.S., 1957. Measles periodicity and community size. J. Roy. Stat. Soc. Ser. A (Gen.) 120, 48–70.

Chinazzi, M., Davis, J.T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., et al., 2020. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science 368 (6489), 395–400.
Cliff, A.D., Haggert, J.K., Versey, G., 1981. Spatial Diffusion: an Historical Geography of Epidemics in an Island Community. Cambridge University Press, Cambridge, UK.
Cohen, Philip N., 2020. The COVID-19 epidemic in rural US counties. Eur. J. Publ. Health 4.
Conticini, E., Frediani, B., Caro, D., 2020. Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy? Environ. Pollut. 114465.
Cook, S.J., Hays, J., Franzese, R., 2020. Model specification and spatial interdependence. In: Curini, L., Franzese, R. (Eds.), Research Methods in Political Science and International Relations, ume 2. SAGE, London, pp. 730–747.
Dearden, E.K., Lloyd, C.D., Catney, G., 2019. A spatial analysis of health status in Britain, 1991–2011. Soc. Sci. Med. 220, 340–352.
Dobis, E.A., Stephens, H.M., Skidmore, M., Goetz, S.J., 2020. Explaining the spatial variation in American life expectancy. Soc. Sci. Med. 246, 112759.
Elliott, P., Wartenberg, D., 2004. Spatial epidemiology: current approaches and future challenges. Environ. Health Perspect. 112 (9), 998–1006.
Felbermayr, G., Hinz, J., Chowdhry, S., 2020. Apres-ski: the spread of coronavirus from ischgl through Germany. COVID Econ.: Vetted Real-Time Pap. 22, 177–204.
Gatto, M., Bertuzzo, E., Mari, L., Miccoli, S., Carraro, L., Casagrandi, R., Rinaldo, A., 2020. Spread and dynamics of the COVID-19 epidemic in Italy: effects of emergency containment measures. Proc. Natl. Acad. Sci. U. S. A. 117, 10484–10491.
Goldstein, J.R., Atherwood, S., 2020. Improved Measurement of Racial/ethnic Disparities in COVID-19 Mortality in the United States. MedRxiv.
Gross, C.P., Essien, U.R., Pasha, S., Gross, J.R., Wang, S.Y., Nunez-Smith, M., 2020. Racial and Ethnic Disparities in Population Level COVID-19 Mortality. MedRxiv.
Hamidi, S., Ewing, R., Sabouri, S., 2020. Longitudinal analyses of the relationship between development density and the covid-19 morbidity and mortality rates: early evidence from 1,165 metropolitan counties in the United States. Health Place 102378.
Han, Y., Li, V.O., Lam, J.C., Guo, P., Bai, R., Fok, W.W., 2020. Who is more susceptible to COVID-19 infection and mortality in the States? MedRxiv 1–9. https://doi.org/10.1101/2020.05.01.20087403.
Harris, J.E., 2020. The Subways Seeded the Massive Coronavirus Epidemic in New York City. NBER Working Paper (w27021).
Lawson, A.B., Banerjee, S., Haining, R.P., Ugarte, M.D. (Eds.), 2016. Handbook of Spatial Epidemiology. CRC Press.
Likasse, H.T., 2020. The impacts of covariates on spatial distribution of corona virus 2019 (COVID-19): what do the data show through ANCOVA and MANCOVA? Eurasian J. Med. Oncol. 4 (2), 141–148.
Marmot, M., Allen, J., Boyce, T., Goldblatt, P., Morrison, J., 2020. Health Equity in England: the Marmot Review 10 Years on. The Health Foundation. https://www.health.org.uk/publications/reports/the-marmot-review-10-years-on.
Martelletti, L., Martelletti, P., 2020. Air pollution and the novel COVID-19 disease: a putative disease risk factor. SN Compr. Clin. Med. 1–5.
Munshi, J., Roy, I., Balasubramanian, G., 2020. Spatiotemporal Dynamics in Demography-Sensitive Disease Transmission: COVID-19 Spread in NY as a Case Study. https://arxiv.org/ftp/arxiv/papers/2005/2005.01001.pdf.
Neumayer, E., Plümper, T., 2016. W. political science research and methods, 4 (1), 175–193.
Notari, A., Torrieri, G., 2020. COVID-19 Transmission Risk Factors arXiv preprint arXiv: 2005.03651.
O'Sullivan, D., Gahegan, M., Exeter, D.J., Adams, B., 2020. Spatially explicit models for exploring COVID-19 lockdown strategies. Trans. GIS 24, 967–1000.
Plümper, T., Neumayer, E., 2020. The COVID-19 pandemic predominantly hits poor neighborhoods, or does it? Evidence from Germany. Eur. J. Public Health 30 (6), 1176–1180.
Rentsch, C.T., Kidwai-Khan, F., Tate, J.P., Park, L.S., King Jr., J.T., Skanderson, M., Re III, V.L., 2020. COVID-19 by Race and Ethnicity: A National Cohort Study of 6 Million United States Veterans. medRxiv.
Rose, T.C., Mason, K., Pennington, A., McHale, P., Taylor-Robinson, D.C., Barr, B., 2020. Inequalities in COVID19 Mortality Related to Ethnicity and Socioeconomic Deprivation (MedRxiv).
Tobler, W.R., 1970. A computer model simulation of urban growth in the detroit region. Econ. Geogr. 46 (2), 234–240.
Travaglio, M., Yu, Y., Popovic, R., Leal, N.S., Martins, L.M., 2020. Links between Air Pollution and COVID-19 in England. MedRxiv.
Wang, S., 2020. Spatial patterns and social-economic influential factors of aging indexes: a global assessment from 1990 to 2010 using spatial regressive modeling. Soc. Sci. Med. 112963.
Wimpy, C., Williams, L.K., Whitten, G.D., 2020. X marks the spot: unlocking the treasure of spatial-X models. J. Polit. (forthcoming).