DISCUSSION PAPER SERIES

Air Pollution Exposure and COVID-19

Matthew A. Cole
Ceren Ozgen
Eric Strobl

DISCUSSION PAPER SERIES

Air Pollution Exposure and COVID-19

Matthew A. Cole
University of Birmingham

Ceren Ozgen
University of Birmingham and IZA

Eric Strobl
University of Bern

ABSTRACT

Air Pollution Exposure and COVID-19*

In light of the existing preliminary evidence of a link between Covid-19 and poor air quality, which is largely based upon correlations, we estimate the relationship between long term air pollution exposure and Covid-19 in 355 municipalities in the Netherlands. Using detailed secondary and administrative data we find compelling evidence of a positive relationship between air pollution, and particularly $PM_{2.5}$ concentrations, and Covid-19 cases, hospital admissions and deaths. This relationship persists after controlling for a wide range of explanatory variables. Our results indicate that a 1 μ/m^3 increase in $PM_{2.5}$ concentrations is associated with 9.4 more Covid-19 cases, 3.0 more hospital admissions, and 2.3 more deaths. The relationship between Covid-19 and air pollution withstands a number of sensitivity and robustness exercises including instrumenting pollution to mitigate potential endogeneity and modelling spatial spillovers using spatial econometric techniques.

Corresponding author:
Matthew A. Cole
University of Birmingham
Department of Economics
University House
Edgbaston, B15 2TT
United Kingdom
E-mail: m.a.cole@bham.ac.uk

1   Introduction

The Covid-19 pandemic is causing significant social and economic impacts across large parts of the world. At the time of writing the number of Covid-19 cases worldwide has reached 7.2 million, while the death toll has exceeded 400,000.[1] Governments and health care systems are facing the immense challenge of trying to control the spread of the virus and to prevent hospitals from being overwhelmed as millions of individuals remain subject to lockdown and face significant economic uncertainty. In order to respond to these unprecedented challenges it is important for policy makers and health care professionals to understand which groups of individuals suffer the highest morbidity and mortality risks from Covid-19 and which factors may exacerbate these risks.

A contributory factor that has been tentatively explored in several recent academic studies is poor air quality. While some such studies have identified the significant improvements in air quality that have resulted from Covid-19 lockdowns (Cicala et al., 2020; Cole et al., 2020), others have pointed to a correlation between Covid-19 hotspots and areas with high levels of pollution concentrations (Travaglio et al., 2020; Conticini et al., 2020). It is well known that long term exposure to pollutants such as nitrogen dioxide (NO_2), sulphur dioxide (SO_2), and fine particulate matter ($PM_{2.5}$) contributes to cardiovascular disease, reduces lung function, and causes respiratory illness (Faustini et al., 2014; Ming Han et al., 2015; Katanoda et al., 2011; Abbey et al., 1999; Weerdt et al., 2020). These pollutants have been shown to cause a persistent inflammatory response even in the relatively young, and to increase the risk of infection by viruses that target the respiratory tract (Travaglio et al., 2020; Conticini et al., 2020). While Covid-19 produces only mild symptoms for most sufferers, in a minority of cases it results in an excessive inflammatory response causing Acute Respiratory Distress Syndrome (ARDS) and death. Given the clear overlaps between the symptoms of Covid-19-induced ARDS and long term exposure to air pollution, a number of studies have begun to explore the links between the two.

Focusing on the UK, Travaglio et al. (2020), for instance, find evidence of a correlation between Covid-19 cases and concentrations of nitrogen oxides and ozone, while Ogen (2020) examines 66 regions across Italy, France, Germany, and Spain and finds similar evidence.

---

[1]As of June 9th 2020. Source https://www.worldometers.info/coronavirus/countries

Conticini et al. (2020) focus on Northern Italy and conclude that pollution concentrations are a likely contributor to the high Covid-19 death rates experienced in that region. A preliminary analysis of the Netherlands also finds evidence of a link between concentrations of $PM_{2.5}$ and Covid-19 cases (Andree, 2020). Finally, Wu et al. (2020) examine US counties and econometrically estimate the relationship between county-level Covid-19 death rates and long-term concentrations of fine particulate matter ($PM_{2.5}$), controlling for a wide range of confounding factors. They conclude that a 1 μg/m^3 increase in $PM_{2.5}$ is associated with an 8% increase in the Covid-19 death rate.

While the above studies provide useful preliminary evidence, Conticini et al. (2020) and Ogen (2020) offer only geographical correlations between Covid-19 cases and pollution exposure, whereas Travaglio et al. (2020) take a similar approach but control only for differences in population density and do so across only 7 relatively large regions. Establishing a convincing link between exposure to pollution and Covid-19 cases requires individual-level data with the ability to control for individual characteristics, such as age and the presence of underlying health conditions. Since individual-level data on Covid-19 infections is not available, the next best alternative is to examine a large number of small geographic regions with detailed data on the characteristics of those regions. This allows the researcher to assess whether any correlation between Covid-19 and pollution exposure still holds once differences in social deprivation, population density, ethnic composition, and other factors are controlled for. While Wu et al. (2020) come closest to doing this, US counties are still relatively large, raising the question of how well such aggregated data can capture the local variation in confounding effects without being 'averaged out'.[2] Furthermore, by focusing only on particulate matter, and on Covid-19 deaths, it is not clear whether other pollutants have an effect on Covid-19 deaths once other factors are controlled for or, indeed, on Covid-19 infections, or hospitalisation rates more generally.

With the above in mind, this paper undertakes a detailed econometric analysis of the relationship between pollution concentrations and Covid-19 using data for 355 relatively small Dutch municipalities. More specifically, by using high-resolution air pollution data as well as combining administrative with secondary data, we estimate the relationship between

---

[2]For instance, Los Angeles County, the largest in the US, had a population of over 10 million in 2019, with the average county-level population being over 104,000 (US Census Bureau).

long term exposure to concentrations of $PM_{2.5}$, NO_2, and SO_2 and the number of Covid-19 infections, the number of individuals hospitalised with Covid-19, and the number of those who died as a result of Covid-19. We are able to control for a wide range of potential confounding effects, including those relating to income, age, underlying health conditions, education, social deprivation, ethnic composition, workplace characteristics, spatial and social proximity to potential risk factors and others. Our analysis utilises Covid-19 data up to 5th June 2020 allowing us to capture almost the entire course of the epidemic and hence much more fully than the previous studies which have examined data up to only March or early April. Finally, we undertake a number of robustness exercises including instrumenting pollution to address possible endogeneity concerns.

The Netherlands provides a useful setting in which to examine the link between Covid-19 and air pollution. As a relatively small, densely populated nation with an ethnically diverse, aging population, the country faces a number of potential Covid-19 risk factors. It also shares an open land border with Germany and Belgium, the latter a country that currently has the highest number of Covid-19 deaths per capita. The Netherlands additionally experiences hot spots of poor local air pollution both within urban areas and also, in the case of $PM_{2.5}$, in more rural areas due in part to intensive livestock farming. By early June 2020, the Netherlands had experienced over 6,000 deaths as a result of Covid-19, resulting in the 7th highest number of Covid-19 deaths per capita.[3]

Our results provide compelling evidence of a statistically significant positive relationship between air pollution and Covid-19 cases, hospital admissions and deaths. More specifically, we find that an increase in $PM_{2.5}$ concentrations of 1 μ/m^3 is associated with an increase in Covid-19 cases of between 9.4 and 15.1, depending on our model. The same increase in $PM_{2.5}$ is associated with an increase in Covid-19 hospital admissions of between 2.9 and 4.4, and an increase in Covid-19 deaths of between 2.2 and 3.6.

The remainder of this paper is organized as follows. Section 2 provides background information on Covid-19 and air pollution in the Netherlands and Section 3 presents our empirical methodology. Section 4 describes our data, Section 5 reports our results and Section 6 concludes.

---

[3]As of June 9th 2020. All Dutch Covid-19 data is provided by the National Institute for Public Health and the Environment (RIVM). The international ranking excludes the principalities of San Marino and Andorra and stems from: https://www.worldometers.info/coronavirus/countries

2   Covid-19 and Air pollution

The first confirmed Covid-19 case in the Netherlands occurred in late February 2020, and by early June over 46,000 cases had been identified. Daily cases of Covid-19 peaked at 1,335 on April 10th, while the daily death toll peaked 3 days earlier at 235. Both daily cases and daily deaths have been declining steadily since. While all 12 of the Netherlands' provinces have experienced broadly similar trends, levels of Covid-19 cases have differed significantly across provinces. For instance, in the southern provinces of North Brabant and South Holland daily cases peaked at over 350, while in other provinces such as Groningen and Drenthe, both in the north of the country, cases peaked at fewer than 40 per day.

From the outset of the epidemic within the Netherlands the south east has experienced a disproportionate number of Covid-19 cases. Figure 1 shows the distribution of cases per capita across the 355 municipalities of the Netherlands up to June 5th 2020. The red 'hotspots' in the south east, which are largely within the provinces of North Brabant and Limburg, demonstrate the relative intensity of cases in those regions. Unusually, these are relatively rural regions with low population density raising the question of why these provinces have been so badly affected by Covid-19.[4]

One potential explanation raised by the Dutch media has been the annual carnival celebrations held during the last week of February and beginning of March which were concentrated largely within North Brabant and Limburg. These celebrations attracted thousands of people from all over the country to street parties and parades, as they do each year. Numerous Dutch media reports have suggested that these celebrations may at least partially explain the rapid spread of Covid-19 within these regions and also to other regions of the Netherlands via carnival participants.[5] However, the significant variation

---

[4]Figure 2 in the Appendix shows that both hospital admissions and deaths per capita from Covid-19 have also been disproportionately high in these south eastern regions.

[5]A non-exhaustive list includes:

https://www.nrc.nl/nieuws/2020/04/09/vuile-lucht-vergroot-de-kans-om-te-sterven-aan-covid-19-a3996388

https://nos.nl/nieuwsuur/artikel/2332460-gebieden-met-veel-luchtverontreiniging-zwaarder-getroffen-door-corona.html

https://www.rtlnieuws.nl/nieuws/nederland/artikel/5094341/corona-update-noorden-minder-besmettingen

https://nadavos.nl/nieuws/2020-luchtverontreiniging-covid-19

https://www.trouw.nl/binnenland/zorgt-de-veehouderij-voor-meer-coronadoden-in-oost-brabant b2b5487d/

in Covid-19 cases across the municipalities within these two provinces suggests that the carnival is unlikely to fully explain the density and distribution of these cases.

Another possible explanation that has received less attention in the Netherlands revolves around the intensive livestock farming that takes place within North Brabant and Limburg. These regions house over 63% of the Netherlands' 12 million pigs and 42% of its 101 million chickens.[6] Such intensive livestock production produces large quantities of ammonia (NH_3), which is an important contributor to $PM_{2.5}$ concentrations. Figure 3 in the Appendix provides a map of 2019 NH_3 concentrations and shows that the south eastern regions have some of the highest concentrations within the Netherlands. Indeed, a comparison with the map of $PM_{2.5}$ concentrations in Figure 1 shows a very similar pattern illustrating how ammonia contributes to $PM_{2.5}$.

While the maximum annual average concentration of $PM_{2.5}$ at municipality level is 12.3 μ/m^3, the equivalent figure for a 1km x 1km gridcell is $23.9\mu/m^3$, which is close to the EU air quality standard of $25\mu/m^3$. This indicates that local concentrations of $PM_{2.5}$ within the regions of North Brabant and Limburg are close to dangerous levels even when averaged annually, raising the likelihood that for shorter periods they may exceed safe levels. Since $PM_{2.5}$ concentrations show a similar spatial distribution to Covid-19 cases, also in Figure 1, it's reasonable to examine this potential link and to see whether the visual relationship between Covid-19 cases and $PM_{2.5}$ concentrations withstands controlling for other potential contributory factors.[7]

3 Methodology

In order to examine the relationship between Covid-19 and air pollution we begin by estimating Equation 1 for 355 municipalities:

$$C_i = \phi pollution_i + \beta_1 D_i' + \beta_2 P_i' + \beta_3 E_i' + \beta_3 S_i' + \beta_4 L_i' + \gamma_r + \epsilon_i \tag{1}$$

---

[6]2019 data from CBS Statline.

[7]This article in Trouw, a Dutch national newspaper, expresses the concerns of local residents that poor air quality in the intensively livestocked regions may be leaving the local population susceptible to Covid-19, https://www.trouw.nl/binnenland/zorgt-de-veehouderij-voor-meer-coronadoden-in-oost-brabant b2b5487d/
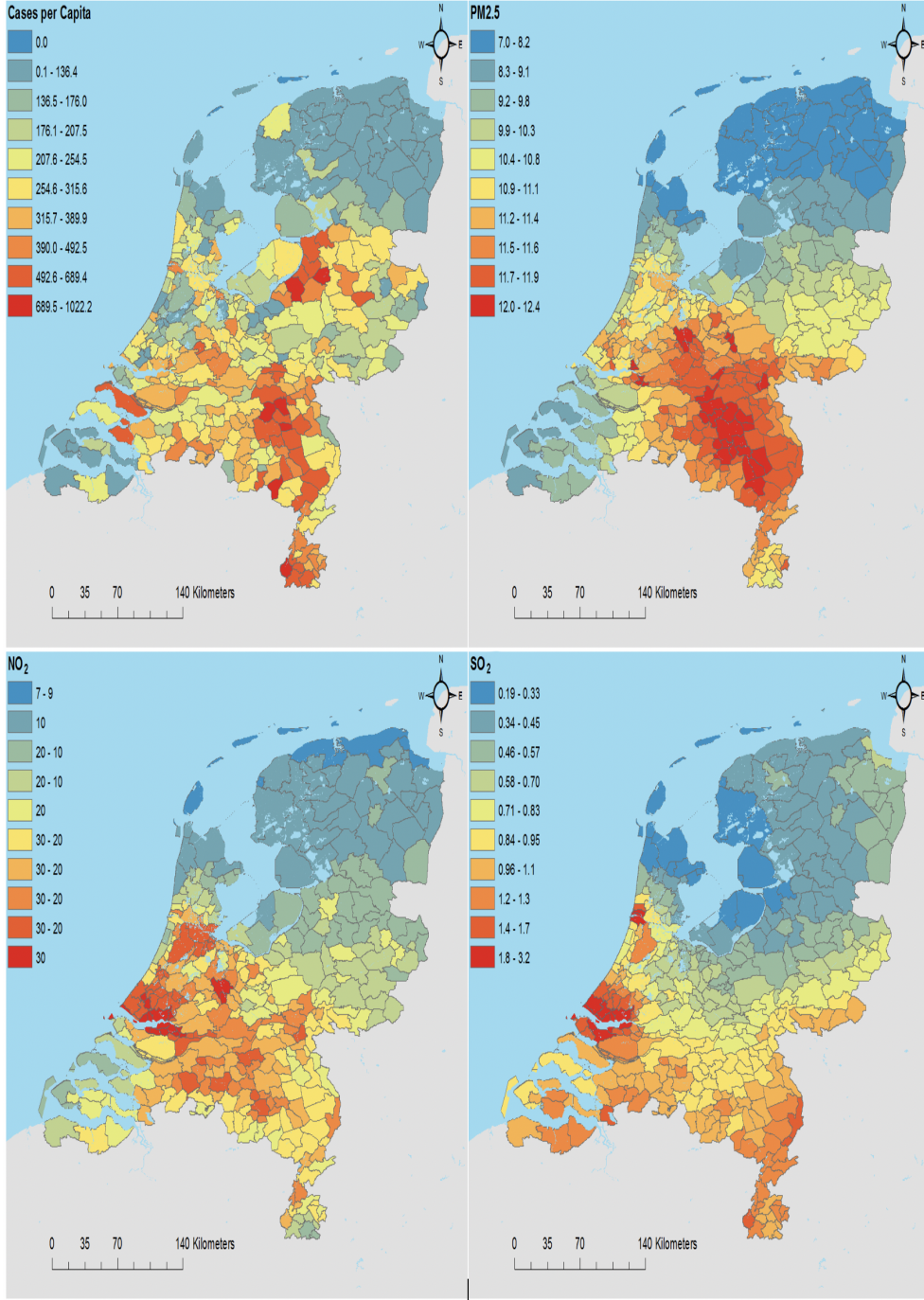
FIGURE 1: COVID-19 CASES PER CAPITA AND ANNUAL CONCENTRATIONS OF $PM_{2.5}$, NO_2 AND SO_2 AVERAGED OVER THE PERIOD 2015-19

where C refers to Covid-19 cases, the number of individuals hospitalised by Covid-19, or the number of deaths from Covid-19 in municipality i as of June 5th 2020. *Pollution* refers to annual concentrations of $PM_{2.5}$, NO_2, or SO_2, averaged over the period 2015-2019. Vectors $D^{'}$, $P^{'}$, $E^{'}$, $S^{'}$ and $L^{'}$ contain control variables capturing demography, social

6

and physical proximity, employment/education, spatial and health variables, respectively, for the year 2019, as defined in the next section. The term γ_r denotes province level fixed effects for each of our 12 provinces, r.

Since our dependent variables take the form of discrete count variables, estimating Equation 1 using OLS could result in inconsistent, inefficient, and biased estimates (Long, 1997; Hoffman, 2004). The alternative is to use Poisson or negative binomial count models. Since the error assumption of the Poisson model requires the conditional mean to equal the conditional variance, a condition that our over-dispersed data fails to meet, we employ the negative binomial model. This model builds upon the Poisson model by adding a parameter that allows the conditional variance to exceed the conditional mean.

Importantly our attempt to estimate the relationship between pollution exposure and Covid-19 cases (as well as hospital admissions and deaths) potentially suffers from a number of econometric challenges:

1. *Omitted variable bias*

   If we fail to control for all potential determinants of Covid-19 that are possibly correlated with pollution, then the estimated coefficients on the pollution measures could potentially be biased. As will be outlined in Section 4, our strategy is to control for a wide range of potential explanatory variables and thereby keep any such bias to a minimum.

2. *Measurement error of pollution*

   This could take 2 forms. First, we may not be accurately capturing long-term pollution exposure within each municipality. While Section 4 outlines how we measure long-term pollution exposure, we note that any non-systematic measurement error of this kind is likely to result in a conservative estimate of the association between pollution and Covid-19 due to possible attenuation bias. Second, there may be systematic measurement error if the error was in some way related to the pollution-Covid-19 relationship. For instance, if individuals with poor health were more (or less) likely to live in polluted areas and hence were more (or less) likely to catch or succumb to the virus. We are unable to identify any potential mechanisms of this nature and our wide range of explanatory variables should obviously help to control for any such

effects if they were to exist.

Nevertheless, we do instrument pollution in equation 1 as a means of further reducing concerns around attenuation bias. Since we are using a non-linear count model we instrument using a control function approach, which is likely to be more efficient than a standard instrumental variables model (Wooldridge, 2015). This approach involves estimating air pollution, our potentially endogenous variable, as a function of our instruments and other exogenous variables, and then inserting the predicted errors from this first stage into the second stage as a separate control variable (in addition to our air pollution variable). A simple test of the statistical significance of the coefficient on the predicted residuals will inform us whether our pollution variable was indeed endogenous, a procedure equivalent to a Durbin-Wu-Hausman exogeneity test.

To be suitable for use as an instrument a variable should be correlated with air pollution but independent of our Covid-19 variables and hence should only influence them through the effect of air pollution. We experiment with two potential candidates. The first is a long lag of air pollution. More specifically, we use annual pollution concentrations averaged over the period 1995-2000. Second, we use a measure of the average commuting time for residents in each municipality. This variable draws upon free-flow road network data from VUGeoPlaza enabling us to calculate average travel time for every j and k location pair. In order to calculate the commuting times we obtain the residential and work locations of each worker at the $4 - digit$ postal code level (which corresponds to neighbourhood level) by constructing an employer-employee data set (LEED) based on linking administrative data, Dutch Labour Force surveys, and Tax Registers.[8] We then combine the travel time data with the LEED. Travel times linked with the LEED allow us to calculate the commuting time of each worker to their actual place of work. Additionally we calculate the travel time based on the Euclidean distance between the $4 - digit$ postal code, assuming an average speed of 10km/h. We then choose the lowest of the commuting time and the Euclidean distance travel time for each worker and average these across municipalities.[9]

---

[8]The access to these data sets requires confidentiality agreement with Statistics Netherlands.

[9]See Koster & Ozgen (2020) for more details. Generally, for short distances (e.g. less than 5km) it is

Differences in commuting time between municipalities reflect differences in the duration and intensity of journeys undertaken and should therefore influence differences in air pollution concentrations. We see no reason why commuting time should directly influence Covid-19 cases.

3. *Measurement error of Covid-19*

   A third potential econometric challenge is the existence of possible measurement error within our Covid-19 data. Non-systematic measurement error would tend to reduce the precision of our estimates by increasing the standard errors on the estimated coefficients. Alternatively, if the measurement error is related to other omitted variables that are important for the pollution-Covid-19 relationship, or to its own values, this could be a cause for concern. Given our rich set of controls and the manner in which Covid-19 data are collected, this is unlikely to be a problem in our context and hence we see no reason for the existence of such systematic measurement error. Section 4 discusses the nature of our Covid-19 data in more detail.

4. *Spillover effects*

   The final potential econometric challenge is the possible existence of spillover effects caused by the virus spreading from one municipality to another. To allow for this possibility we extend equation 1 to include a spatially lagged dependent variable and a spatial error term, each using spatial weight matrices with varying distance cut-offs, as specified in equation 2.

$$C_i = \theta pollution_i + \alpha_1 D_i^{'} + \alpha_2 P_i^{'} + \alpha_3 E_i^{'} + \alpha_3 S_i^{'} + \alpha_4 L_i^{'} + \rho w_i C_i + \lambda w_i \epsilon_i + \mu_i \quad (2)$$

where w_i is an inverse distance spatial weight matrix with a 50km cut-off, a 100km cut-off, or no cut-off at all. All other variables are as previously specified. Note that the inclusion of a spatially lagged dependent variable into a non-linear count data model is not straightforward and hence there has yet to be a widely accepted method of doing so (Glaser, 2017). Equation 2 is therefore a linear model estimated using

slower to use the road network and so the Euclidean travel time is used.

maximum likelihood.

4   Data and Summary Statistics

4.1   Data

We utilize extensive secondary data combined with administrative micro-data and spatial data from various sources in the Netherlands. Our analysis rests on municipality level cross-sectional Covid-19 data provided by the National Institute for Public Health and the Environment (RIVM) and a rich set of controls from 2019 (unless otherwise stated). We further use high resolution spatial data for air pollution indicators which span the period 2015-19.

The municipal-level data is obtained from the data repository *Statline* of the Statistics Netherlands, while some variables are aggregated from the administrative micro-data up to municipality level. When modelling the relationship between Covid-19 and air pollution, we control for a wide range of potentially confounding effects which we categorise as demographic, social and physical proximity, employment and education, spatial, and health-related. Below we discuss Covid-19 and pollution data and each category of control variable:

4.1.1   Covid-19

The Covid-19 data used in our analysis is obtained from the RIVM. The Regional Public Health Service Centres (known as GGD) from all across the country[10] provide RIVM with the Covid-19 data, which is then corrected for any inconsistencies before being released by RIVM. As stated by RIVM on their website, there is a slight delay between the day of hospitalization or death and the day on which the the number of cases is reported.[11] However, when the data is available RIVM does the necessary adjustment to the data retrospectively. This incident of slight delay is well-known for all countries. Again similar to the other countries, the number of Covid-19 cases is likely to be an underestimate as not everybody in the country is tested. Finally, all Covid-19 variables are coded by the

---

[10]https://www.ggd.amsterdam.nl/coronavirus/.
[11]https://www.rivm.nl/coronavirus-covid-19/actueel.

residence of each individual rather than the address of the test centre or the hospital to which they are admitted.

4.1.2 Air pollution

Annual concentrations of $PM_{2.5}$, NO_2, and SO_2 measures are provided by RIVM. These pollution concentrations are reported at the level of 1x1 km^2 grid-cells having been modelled using a wide range of sources and components in the Netherlands and in other European countries. Maps of the spatial distribution of pollution concentrations were calculated using the Operational Priority Substances dispersion model, which constructs the average annual concentrations of pollutants stemming from the dispersion, transport, chemical conversion and dispersion of emissions. The resultant concentrations are calibrated with the observations from the Dutch Air Quality Monitoring Network G. J. Velders et al. (2020); Fischer et al. (2020).[12]

To produce municipality level measures of pollution concentrations we use the median grid-cell concentrations within each municipality. In order to overcome any potential measurement bias due to annual fluctuations and to capture the long term exposure of residents within a municipality, we average the annual pollution concentrations data over the 5-year period 2015-2019.

4.1.3 Demographic and labour market indicators

The demographic controls include the total population as well as the density of population in each municipality, expressed as population per square kilometre. The share of the population over 70 years of age in our regressions accounts for the particular vulnerability of the elderly to Covid-19, as the stylized facts from around the world indicate higher casualties among the elderly population. We also include the share of the population under 18 years old, with the omitted category being the working age share of the population.

Finally, several recent studies from a number of countries including the UK and the US[13] have indicated that ethnic minorities are disproportionately affected by Covid-19. There are a number of potential reasons why this might be so. First, Zorlu & Hartog (2012) show

---

[12]For more details on the construction of the pollution concentrations data see G. Velders et al. (2017)

[13]See Kirby (2020) and Platt & Warwick (n.d.) https://www.ifs.org.uk/inequality/chapter/are-some-ethnic-groups-more-vulnerable-to-covid-19-than-others

that ethnic groups in the Netherlands are more likely to work in manual occupations, which have been shown to increase the risk of exposure to Covid-19 due to the greater frequency of face-to-face contact amongst workers (Lewandowski, 2020).[14] Second, ethnic minorities may be more likely to experience social deprivation, live in smaller housing and belong to larger households. As will be discussed below, we try to control for each of these factors by taking housing and household size conditions into account. Other risk factors may include the use of certain cultural practices (e.g., attending places of worship) or having a disposition for underlying health conditions. To control for any such effects we include the number of non-western migrants as a share of the population. Non-western migrants in the Netherlands are defined as the number of immigrants who are born in Africa, Latin America, Asia (excluding Japan and Indonesia), and Turkey.

We also include the average household income level and the share of employment in elementary occupations to control for the possibility that groups that are economically and occupationally more vulnerable could be more affected by Covid-19. Finally, our specification includes the number of highly educated people as a share of the total educated population at the municipal level. Highly educated is defined as those who have obtained a bachelor's degree or higher.

4.1.4   Social and Physical Proximity

While the majority of countries have disproportionately experienced Covid-19 infections in large metropolitan areas, the empirical findings so far are unclear about the role played by population density.[15] Epidemiological studies have argued that density *per se* is not important but rather it is the type of *interaction* between people that is important e.g. weddings, close family events and occasions when individuals come into close proximity with each other (Hu et al., 2013). One such occasion in the Netherlands is annual the Carnival which is widely celebrated in the southern regions of the country – North Brabant and Limburg– as discussed in Section 2. To capture the potential impact of the Carnival we initially rely on province level fixed effects which capture the effects in North Brabant and

---

[14]https://www.theguardian.com/world/2020/may/11/manual-workers-likelier-to-die-from-covid-19-than-professionals.

[15]See for example,
https://www.centreforcities.org/blog/have-uk-cities-been-hotbeds-of-covid-19-pandemic/
http://jedkolko.com/2020/04/15/where-covid19-death-rates-are-highest/.

Limburg and all other provinces individually. Furthermore, in sensitivity analysis we test a Carnival(municipality) measure by creating a dummy variable for the seven most populous municipalities (Den Bosch, Tilburg, Breda, Bergen op Zoom, Eindhoven, Helmond, Oss) within the borders of North Brabant and Limburg provinces. Alongside this, in order to account for the potential spread of the virus from the areas where the Carnival is widely celebrated, we also test a proximity to Carnival(municipality) measure that is the inverse distance from the centroid of these seven Carnival municipalities to the centroid of all municipalities in the Netherlands.

To capture physical interactions between households we include the share of small housing, defined as the share of total housing that is between 15 and 50 m^2 in size, and the average household size. We also apply the same principle to workplaces. We aggregate a linked employer-employee data set based on the population of firms and employees in the Netherlands from 2016 records in order to create an average firm size variable at the municipal level. This data is obtained from Tax Registers, which include approximately 12 million observations from the universe of employees.

4.1.5   Spatial Controls

Our specification includes several spatial variables which capture the spatial interactions and spillovers between municipalities. First, we account for the municipalities where Covid-19 infections may have been influenced by the transmission of the virus beyond Dutch national borders. The Dutch municipalities bordering Belgium experience substantial daily cross-border mobility and Covid-19 cases in Belgium have been the highest of all countries in per capita terms. Dutch municipalities that border Germany also experience intensive daily crossings, although, in contrast, the Covid-19 cases in Germany had been relatively low compared to other European countries. In order to capture the potential contagion through cross-border mobility, we therefore construct 2 dummy variables; municipalities bordering Germany and municipalities bordering Belgium.

Second, Schiphol airport within the borders of the municipality of Haarlemmermeer (9 km from Amsterdam) is the main international hub of the country and one of the largest airports in Europe with approximately 72 million passengers a year. We therefore construct a variable capturing the inverse distance from each municipality centroid to Schiphol airport

to capture the potential contagion from the high volume of passengers. Lastly, to capture the effect of the prevailing wind from the North Sea which could mean less physical spread of the virus and/or reduced spillovers from neighbours, we include a dummy variable for all of the municipalities on the Dutch coast.

4.1.6 Health-related

Covid-19 statistics in many countries have indicated that those persons with underlying health conditions are likely to be disproportionately affected by the virus. Our data allows us to control for these groups at the municipality level. We include in our estimations the share of smokers and the share of those suffering from obesity in the total local population in 2019. Moreover, the share of people receiving incapacity benefits is also included to proxy those who cannot be active in the labour market due to health constraints.

4.2 Descriptive Statistics

We report the descriptive statistics of our data in Table 1. Our Covid-19 data covers the period since the first Covid-19 case in February until the tail-end of the epidemic in June. During this period there have been on average 131 Covid-19 cases per municipality with a maximum number of 2416. In the same period on average 33 hospital admissions were recorded across all municipalities (maximum 611) while on average 17 people died (maximum 336).

Our data set includes 355 municipalities based on the 2019 municipal borders in the Netherlands, where each has an average population of approximately 47000 people, of which 7.4% are non-Western European immigrants. Although the size of the municipalities is broadly uniform, average population density across municipalities shows a large variation, ranging from 23 individuals per km^2 to 6523 per km^2. On average 3.5% of housing is classed as small (15-50 m^2) while the average household size is 2.3 individuals. Finally, 15% of the population is over 70 years old.

The median annual wage is 32722. With 32% of the population holding a university degree or higher, Netherlands has a relatively highly-educated population. 9% of the population performs elementary occupations which require low-skilled manual-task intensive work. Median firm size is 556 workers, but becomes larger for peripheral areas, although

the average firm size in large cities is still around a standard deviation higher than the median. Although known to be one of the healthiest nations, 19% of the population are smokers and 14% experience obesity (defined as % of over 19s with BMI¿30 kg/m2). 6% receive incapacity benefits due to not being able to work.

With regard to our pollution concentrations data which is measured in μ/m^3, we see that the average value of $PM_{2.5}$ concentrations in our dataset is 10.491, with a maximum of 12.264. For NO_2 concentrations the mean value is 15.757, with a maximum value of 27.413, and for SO_2 concentrations the mean value is 0.797 with the maximum being 3.092. While the EU air quality standards do not stipulate a safe limit for annual concentrations of SO_2, for $PM_{2.5}$ and NO_2 concentrations the limit is $25\mu/m^3$ and $40\mu/m^3$, respectively. We can therefore see that 5 year averages of annual concentrations in the Netherlands do not exceed these limits. However, it is worth noting that the maximum values of the annual concentrations of our 1 x 1 km^2 grid-cells that form the basis of our municipality pollution data are $23.9\mu/m^3$ and $62.4\mu/m^3$ for $PM_{2.5}$ and NO_2, respectively. In the case of NO_2, this is significantly beyond the EU safe limit, while for $PM_{2.5}$ it is very close to the safe limit.

5 Results

Table 2 provides our initial negative binomial estimates of Equation 1 with columns (1) to (3) providing the estimates of Covid-19 cases for the three pollutants. Columns (4) to (6) do the same for hospital admissions, and columns (7) to (9) provide estimates of Covid-19 deaths, again for the three pollutants.

Focusing initially on the estimated pollution coefficients, we see that $PM_{2.5}$ concentrations have a positive and statistically significant relationship with Covid-19 cases, hospital admissions and deaths. In the case of NO_2 we find a positive and statistically significant association between Covid-19 cases and deaths, but this is not statistically significant for hospital admissions. Finally, for SO_2 we again discover a positive relationship with our dependent variables, but this is only statistically significant for Covid-19 deaths. After calculating marginal effects, we find that a one unit increase in $PM_{2.5}$ concentrations is associated with 9.4 more Covid-19 cases, 3.0 more hospital admissions, and 2.3 more deaths.

Variable	Mean	Std. Dev.	Min.	Max.	N
PM2.5, 5yr ave	10.491	1.354	6.915	12.264	355
NO2, 5yr ave	15.757	3.861	6.835	27.413	355
SO2, 5yr ave	0.797	0.36	0.21	3.092	355
# of Covid-19 cases	131.459	224.887	0	2416	355
# of deaths	16.865	31.445	0	336	355
# of hospital admissions	33.017	55.129	0	611	355
Carnival(province)	0.262	0.44	0	1	355
Carnival(municip)	0.02	0.139	0	1	355
Proximity to Carnival(municip)	0.000	0.000	0	0.0001	355
Days since 1st case	86.468	3.197	76	89	355
Proximity to Schiphol	86.491	46.151	5.274	194.828	355
Belgian border	0.073	0.261	0	1	355
Coast	0.093	0.291	0	1	355
German border	0.09	0.287	0	1	355
Ave gross income	32.722	3.05	23.5	53.6	355
ln(firm size)	8.9	0.513	3.258	10.06	355
Sh elementary occup	0.089	0.031	0	0.2	358
Sh high educated	0.323	0.095	0.071	0.9	358
Ave household size	2.297	0.178	1.712	3.339	355
Sh of small houses	0.036	0.045	0	0.672	355
Sh of under 18	0.196	0.024	0.126	0.335	355
Sh of 70+	0.15	0.025	0.063	0.243	355
Obesity	14.389	2.13	9	22	355
Smokers	19.623	2.747	14	31	355
Sh receiving incapacity benefits	0.063	0.017	0.02	0.132	355
Population/km2	877.315	1042.722	23.111	6523.142	355
ln(population)	10.398	0.826	6.842	13.668	355
Non-western mig sh	0.074	0.059	0.014	0.386	355
Ave commuting time	14.066	8.284	4.279	134.432	358

Note: The proximity to carnival measure is calculated as the negative exponential of distance i.e. (exp(-distance)) and produces numbers that are very small in magnitude, thereby explaining the large estimated coefficients for this variable.

A one unit increase in NO_2 increases Covid-19 cases by 2.2 and deaths by 0.35. To make these comparable, a one standard deviation increase in $PM_{2.5}$ and NO_2 concentrations increases Covid-19 cases by 12.7 and 8.6, respectively. The same one standard deviation in $PM_{2.5}$ and NO_2 concentrations increases Covid-19 deaths by 3.0 and 1.4, respectively. Table 3 summarises our pollution marginal effects from the estimated results in which pollution is statistically significant.

Turning to the other explanatory variables, and focusing on those that are generally

statistically significant, we find that municipalities next to the German border have lower Covid-19 cases, hospital admissions, and deaths compared to other municipalities. We also find average household income to have a negative association with all three dependent variables, while average household size and the share of housing that is small both have positive relationships with our Covid-19 variables. Relative to the share of the working age population, the share of those under 18 has a negative association with Covid-19 while, as expected, the share over 70s has a positive association. Finally, smokers, the share of non-Western immigrants, and the total population of each municipality are associated with increased Covid-19 cases, hospital admissions, and deaths.

Since we find the greatest statistical significance and largest marginal effects for $PM_{2.5}$ concentrations, for reasons of space we focus on this pollutant alone for the remainder of our analysis.[16] Table 4 reports the results of the instrumental variable estimations using a control function approach, as outlined previously. In Table 4 we see that the estimated coefficients on $PM_{2.5}$ concentrations remain positive and statistically significant, with a similar magnitude to those in Table 2. It is notable that the first stage residuals are not statistically significant, thus failing to reject the null of exogeneity. One can also see that an F-test on our instruments in the first stage is highly statistically significant. Table 7 in the Appendix reports the first stage results, which use commuting time, commuting time squared, and lagged particulate matter concentrations from 1995-2000 as instruments. In sensitivity analysis we tested different combinations of these variables including lagged pollution alone, commuting time alone, and commuting time in linear form only. In each case the first stage residuals were not statistically significant in the second stage. Our results therefore provide reassurance that our estimations in Table 2 are not being unduly influenced by endogeneity.

---

[16]Results for other pollutants are available upon request from the corresponding author.

TABLE 2: MAIN ESTIMATION RESULTS

	Covid-19 Cases			Hospital Admissions			Deaths		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PM2.5, 5yr ave	0.11**			0.15**			0.23***		
	(0.051)			(0.065)			(0.073)		
NO2, 5yr ave		0.027**			0.015			0.035**	
		(0.012)			(0.013)			(0.016)	
SO2, 5yr ave			0.11			0.055			0.18*
			(0.079)			(0.065)			(0.10)
Belgian border	0.18	0.16	0.14	-0.079	-0.11	-0.12	0.084	0.041	0.018
	(0.11)	(0.11)	(0.11)	(0.11)	(0.12)	(0.12)	(0.16)	(0.16)	(0.16)
German border	-0.29***	-0.29***	-0.28***	-0.25***	-0.23***	-0.23***	-0.36***	-0.34***	-0.34***
	(0.084)	(0.080)	(0.081)	(0.081)	(0.080)	(0.082)	(0.11)	(0.10)	(0.11)
Ave gross income	-0.027**	-0.022**	-0.020*	-0.027**	-0.019	-0.018	-0.047**	-0.035*	-0.032
	(0.012)	(0.011)	(0.011)	(0.013)	(0.012)	(0.012)	(0.022)	(0.021)	(0.021)
Ave household size	1.80***	1.73***	1.69***	2.19***	2.05***	2.02***	2.00***	1.82***	1.76***
	(0.38)	(0.37)	(0.37)	(0.50)	(0.50)	(0.50)	(0.68)	(0.68)	(0.67)
Sh of small houses	1.24***	1.27***	1.28***	0.68**	0.70**	0.71**	0.87**	0.94***	0.97***
	(0.27)	(0.26)	(0.27)	(0.28)	(0.28)	(0.28)	(0.37)	(0.35)	(0.35)
Sh of under 18	-3.26	-3.10	-3.06	-8.02***	-7.76***	-7.73***	-2.76	-2.18	-2.00
	(2.29)	(2.29)	(2.28)	(2.82)	(2.86)	(2.86)	(3.74)	(3.76)	(3.72)
Sh of 70+	4.60***	4.22***	3.59**	1.95	1.03	0.64	9.15***	8.00***	7.16***
	(1.54)	(1.45)	(1.50)	(1.63)	(1.59)	(1.66)	(2.32)	(2.21)	(2.24)
Smokers	0.057***	0.052***	0.049***	0.033**	0.025	0.024	0.063***	0.050**	0.046*
	(0.017)	(0.017)	(0.017)	(0.015)	(0.016)	(0.016)	(0.021)	(0.024)	(0.024)
Ln(population)	0.99***	0.99***	0.99***	1.00***	0.99***	0.99***	1.02***	1.00***	1.01***
	(0.041)	(0.040)	(0.042)	(0.052)	(0.051)	(0.052)	(0.066)	(0.065)	(0.068)
Non-west mig sh	1.40**	1.34**	1.37**	1.84***	1.91***	1.93***	2.24**	2.32**	2.23**
	(0.65)	(0.64)	(0.66)	(0.62)	(0.62)	(0.61)	(1.04)	(1.01)	(0.98)
Observations	355	355	355	355	355	355	355	355	355
Province Dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES
Pseudo R^2	0.17	0.17	0.17	0.19	0.19	0.19	0.18	0.18	0.18

Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Additional controls included in the estimations but not reported are proximity to Schiphol airport, coastal dummy, average firm size, share of workers in elementary occupations, share of obesity, population density, share of working age population receiving incapacity benefit.

TABLE 3: SUMMARY OF ESTIMATED MARGINAL EFFECTS FOR POLLUTION

Pollutant	Model type	1μ/m³ increase in the pollutant			1 st. dev increase from the mean		
		Covid-19 Cases	Hospital Admissions	Deaths	Covid-19 Cases	Hospital Admissions	Deaths
PM2.5, 5yr ave	Negative binomial	9.4	3.0	2.3	12.7	4.1	3.0
	Negative binomial IV	10.3	2.9	2.2	13.9	3.9	3.0
	Spatial 50km	15.1	4.4	3.6	20.4	5.9	4.9
	Spatial 100km	14.9	4.2	2.8	20.2	5.7	3.7
	Spatial no cut-off	12.2	3.3	2.3	16.5	4.5	3.1
NO₂, 5yr ave	Negative binomial	2.2	-	0.35	8.6	-	1.4
SO₂, 5yr ave	Negative binomial	-	-	1.8	-	-	0.65

Note: Marginal effects are not reported if pollution is not statistically significant at conventional levels in the initial estimation. The table should be read as follows: A 1μ/m³ increase in PM2.5 will increase Covid-19 cases by 10.3 according to the negative binomial IV model. Similarly, a 1 standard deviation increase in NO₂ will increase Covid-19 deaths by 1.4 according to the negative binomial model.

Table 5 reports the results of the spatial econometric estimations based on Equation 2 in which we include a spatially lagged dependent variable and spatial errors using three different spatial weight matrices (50km, 100km, no cut-off). Table 5 reports the estimated coefficients on $PM_{2.5}$ concentrations for the three dependent variables. The three models with differing weight matrices are presented. For each we also report the spatial error and spatial lag coefficients, λ and ρ, respectively. Table 5 indicates that $PM_{2.5}$ concentrations continue to have a positive and statistically significant relationship with all three of the Covid-19 dependant variables, for all three of the spatial weight matrices. The introduction of the spatially lagged dependent variable into Equation 2 changes the interpretation of our estimated coefficients. $PM_{2.5}$ concentrations in municipality i continue to affect the conditional mean of Covid-19 in that municipality but now that change in Covid-19 potentially changes the conditional mean of Covid-19 in other nearby municipalities (depending on our weight matrix). Furthermore, the change in Covid-19 in those nearby municipalities affects the conditional mean of Covid-19 in their neighbouring municipalities and so on. $PM_{2.5}$ concentrations therefore now have a *direct* effect on Covid-19 in their own municipality plus an *indirect* effect on Covid-19 in other municipalities.

If we take the example of our estimate of Covid-19 Cases using a weight matrix with a 50km cut-off, which reports an estimated coefficient on $PM_{2.5}$ concentrations of 20.88, we can calculate the marginal effect to consist of a direct effect of 20.95, an indirect effect of -5.89, and a total effect of 15.06. This implies that a 1 unit change in $PM_{2.5}$

VARIABLES	Covid-19 Cases (1)	Hospital Admissions (2)	Deaths (3)
PM2.5, 5yr ave	0.12**	0.14**	0.23***
	(0.058)	(0.064)	(0.080)
First stage residuals	-0.027	0.012	0.0018
(Durbin-Wu-Hausman)	(0.40)	(0.23)	(0.02)
F-Test on Instruments	190.48	190.48	190.48
(Prob>F)	(0.000)	(0.000)	(0.000)
Province Dummies	YES	YES	YES
Observations	355	355	355
Pseudo R²	0.17	0.19	0.18

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Additional controls included in the estimations but not reported are Belgian border dummy, German border dummy, average household income, average household size, share of small housing, share of population under 18 years old, share of population over 70 years old, share of population who are smokers, population level, non-western migrant share, proximity to Schiphol airport, coastal dummy, average firm size, share of workers in elementary occupations, share of obesity, population density, share of working age population receiving incapacity benefit.

concentrations results in an increase in Covid-19 cases of 15.06. However, it is notable that the estimates of both λ and ρ from this model are not statistically significant, implying that spatial spillovers are not unduly influencing our results, either through the error term or through the dependent variable. Indeed, if we look at the estimates of λ and ρ from the other models in Table 5 we see that the majority are not statistically significant. This is particularly true of the spatial error coefficient λ. So, while we find limited evidence of statistically significant spatial spillovers, Table 5 at least confirms that our positive and statistically significant association between Covid-19 and $PM_{2.5}$ concentrations is robust to the inclusion of such spillovers.

Finally, Table 6 reports some additional sensitivity estimations. In columns (1) to (3), rather than relying only on province dummies to control for the effect of the carnival, we include a municipality-level dummy of the seven most populous municipalities within North Brabant and Limburg provinces. In addition, we include a measure of the proximity of each municipality to the centroid of these seven combined municipalities. Table 6 indicates that these municipality-level carnival variables are not statistically significant, nor do they have the expected signs. Furthermore, the coefficients on $PM_{2.5}$ concentrations remain unaffected by their inclusion in terms of magnitude, sign, or significance.

As mentioned earlier, the effect of urban density on the spread of Covid-19 has been on ongoing discussion with a number of studies providing mixed and inconclusive evidence.

	Covid Cases	Hospital Admissions	Deaths
50km cut off:			
PM2.5, 5yr ave	20.88**	6.50**	3.71**
	(10.69)	(3.25)	(1.49)
Spatial error (λ)	-0.20	0.41	-0.56
	(0.52)	(0.35)	(0.59)
Spatial lag (ϱ)	-0.58	-0.76**	-0.50*
	(0.29)	(0.31)	(0.29)
100km cut off:			
PM2.5, 5yr ave	21.40*	6.57*	4.00**
	(12.86)	(3.62)	(1.91)
Spatial error (λ)	-0.54	0.42	-0.90
	(0.86)	(0.88)	(0.91)
Spatial lag (ϱ)	-0.57	-0.74*	-0.57
	(0.44)	(0.43)	(0.45)
No distance cut off:			
PM2.5, 5yr ave	21.2*	9.15**	3.82**
	(12.55)	(4.19)	(1.87)
Spatial error (λ)	-0.52	1.90***	-0.97
	(1.05)	(0.41)	(1.17)
Spatial lag (ϱ)	-0.83	-2.03***	-0.77
	(0.62)	(0.64)	(0.64)
Observations	355	355	355

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Additional controls included in the estimations but not reported are Belgian border dummy, German border dummy, average household income, average household size, share of small housing, share of population under 18 years old, share of population over 70 years old, share of population who are smokers, population level, non-western migrant share, proximity to Schiphol airport, coastal dummy, average firm size, share of workers in elementary occupations, share of obesity, population density, share of working age population receiving incapacity benefit.

To explore this further, columns (4) to (6) in Table 6 report estimations in which we omitted from our analysis the major urban areas of Amsterdam, Rotterdam, Utrecht and The Hague in case, given their high levels of concentration, they are unduly influencing our results. The sign, significance and magnitude of our estimated coefficients on pollution are consistent with our previous results (in columns (1), (4) and (7) from Table 2).

6   Conclusion

This paper has contributed to the nascent literature examining the link between poor air quality and Covid-19. We examine data for 355 Dutch municipalities to identify the relationship between concentrations of $PM_{2.5}$, NO_2 and SO_2 and Covid-19 cases, hospital admissions and deaths. In contrast to much of the previous literature we are able to control

TABLE 6: SENSITIVITY CHECKS II - ADDITIONAL RESULTS

	Covid Cases	Hospital Admissions	Deaths	Covid Cases	Hospital Admissions	Deaths
	(1)	(2)	(3)	(4)	(5)	(6)
PM2.5, 5yr ave	0.12**	0.15**	0.25***	0.13***	0.17***	0.27***
	(0.051)	(0.066)	(0.071)	(0.048)	(0.063)	(0.068)
Carnival(municip)	-0.23	-0.12	-0.39*			
	(0.14)	(0.17)	(0.22)			
Proximity to Carnival(municip)	-1,126	-374	-1,598			
	(754)	(1,039)	(1,238)			
Belgian border	0.13	-0.100	0.016	0.18*	-0.068	0.094
	(0.11)	(0.11)	(0.15)	(0.11)	(0.11)	(0.15)
German border	-0.30***	-0.25***	-0.37***	-0.30***	-0.26***	-0.37***
	(0.084)	(0.080)	(0.11)	(0.079)	(0.081)	(0.11)
Ave gross income	-0.029**	-0.029**	-0.051**	-0.024**	-0.031**	-0.045*
	(0.012)	(0.013)	(0.023)	(0.011)	(0.013)	(0.024)
Ave household size	1.78***	2.20***	1.98***	1.39***	1.67***	1.28*
	(0.38)	(0.50)	(0.69)	(0.42)	(0.53)	(0.73)
Sh of small houses	1.25***	0.68**	0.91**	1.23***	0.51*	0.85**
	(0.26)	(0.28)	(0.37)	(0.25)	(0.27)	(0.34)
Sh of under 18	-2.85	-7.81***	-1.90	-1.33	-5.03	-0.36
	(2.30)	(2.89)	(3.78)	(2.57)	(3.18)	(4.10)
Sh of 70+	4.86***	2.10	9.64***	4.66***	2.25	8.23***
	(1.50)	(1.65)	(2.32)	(1.53)	(1.58)	(2.42)
Smokers	0.060***	0.034**	0.067***	0.052***	0.023	0.050**
	(0.017)	(0.016)	(0.022)	(0.019)	(0.017)	(0.025)
Sh receiving incapacity benefits	-1.05	1.77	3.07	0.76	3.57	5.52
	(1.94)	(2.30)	(3.13)	(2.11)	(2.55)	(3.49)
ln(population)	1.01***	1.01***	1.07***	0.98***	0.98***	1.00***
	(0.044)	(0.058)	(0.075)	(0.043)	(0.056)	(0.070)
Non-western mig sh	1.13*	1.70***	1.75*	0.022	0.53	-0.40
	(0.67)	(0.64)	(1.02)	(0.89)	(0.83)	(1.28)
Observations	355	355	355	351	351	351
Province Dummies	YES	YES	YES	YES	YES	YES
Pseudo R^2	0.17	0.19	0.19	0.16	0.18	0.18

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Additional controls not reported are proximity to Schiphol airport, coastal dummy, average firm size, share of workers in elementary occupations, share of obesity, population density. Columns (1) to (3) include additional carnival variables while columns (4) to (6) omit the 4 major cities of Amsterdam, Rotterdam, Utrecht and The Hague.

for a wide range of potential confounding effects and, by examining Covid-19 data between February and June 2020, are able to examine almost the full duration of the epidemic within the Netherlands.

We find compelling evidence of a statistically significant positive relationship between air pollution and Covid-19 cases, hospital admissions and deaths. This relationship is particularly evident for concentrations of $PM_{2.5}$ and to a lesser extent NO_2 and persists after controlling for explanatory variables capturing income, demography, social and physical proximity, employment/education, health and spatial factors. The relationship with-

stands a number of sensitivity and robustness exercises including instrumenting pollution to mitigate potential endogeneity and modelling spatial spillovers using spatial econometric techniques.

Our results indicate that an increase in $PM_{2.5}$ concentrations of $1\mu/m^3$ is associated with an increase in Covid-19 cases of between 9.4 and 15.1, depending on our model. The same increase in $PM_{2.5}$ is associated with an increase in Covid-19 hospital admissions of between 2.9 and 4.4, and an increase in Covid-19 deaths of between 2.2 and 3.6. The only comparable study to our own is provided by Wu et al. (2020) who examine Covid-19 deaths in the US. They find that a 1 $\mu g/m^3$ increase in $PM_{2.5}$ is associated with an 8% increase in the Covid-19 death rate. Since the mean number of deaths in our sample is 16.86, our estimated increases of between 2.2 and 3.6 are equivalent to increases of between 13.0% and 21.4%, which are clearly larger in magnitude than those of Wu et al. (2020).

Two clear policy implications arise from our analysis. First, the impact of poor air quality on Covid-19 morbidity and mortality represents a considerable and unexpected additional cost from air pollution. Our results would therefore suggest a clear need to regulate air pollution more stringently, even in a relatively well-regulated nation such as the Netherlands. Second, our findings should prove useful to public health officials by signaling where subsequent waves of Covid-19, or indeed future pandemics, might hit hardest. This could be particularly important in countries where pollution sources don't correlate well with major metropolitan areas. If, for instance, livestock production is a major source of emissions, as appears to be the case in some parts of the Netherlands, then advance warning may prove particularly beneficial for rural areas where healthcare infrastructure and coordination may be less developed.

We believe the statistical relationships we have observed between $PM_{2.5}$ concentrations and Covid-19 data are robust. Furthermore, given the established literature linking poor air quality with respiratory disease, they are also plausible. Nevertheless, a degree of caution is needed. Until detailed individual-level data is available providing information on Covid-19 and a wide range of other individual characteristics the statistical evidence will remain suggestive, perhaps strongly suggestive, but not conclusive.

References

Abbey, D., Nishino, N., McDonnell, W., Burchetter, R., Knutsen, S., Beeson, W., & Yang, J. (1999). Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. *American Journal of Respiratory and Critical Care Medicine*, *159*(2), 373-382.

Andree, B. P. J. (2020). *Incidence of covid-19 and connections with air pollution exposure evidence from the netherlands* (Working Paper No. 9221). World Bank Group.

Cicala, S., Holland, S. P., Mansur, E. T., Muller, N. Z., & Yates, A. J. (2020, May). *Expected health effects of reduced air pollution from covid-19 social distancing* (Working Paper No. 27135). National Bureau of Economic Research.

Cole, M., Elliott, R., & Liu, B. (2020). *The impact of the wuhan covid-19 lockdown on air pollution and health: A machine learning and augmented synthetic control approach* (Working Paper No. 20-10). University of Birmingham, Department of Economics.

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? *Environmental Pollution*, *261*.

Faustini, A., Rapp, R., & Forastiere, F. (2014). Nitrogen dioxide and mortality: review and meta-analysis of long-term studies. *European Respiratory Journal*, *44*(3), 744–753.

Fischer, P. H., Marra, M., Ameling, C. B., Velders, G. J., Hoogerbrugge, R., de Vries, W., ... Houthuijs, D. (2020). Particulate air pollution from different sources and mortality in 7.5 million adults — the dutch environmental longitudinal study (duels). *Science of the Total Environment*, *75*.

Glaser, S. (2017). *A review of spatial econometric models for count data* (Discussion Papers in Business, Economics and Social Sciences No. 19-2017). Universität Hohenheim.

Hoffman, J. (2004). *General Linear Models: An Applied Approach*. Pearson, Boston.

Hu, H., Nigmatulina, K., & Eckhoff, P. (2013). The scaling of contact rates with population density for the infectious disease models. *Mathematical Biosciences*, *244*(2), 125 - 134.

Katanoda, K., Sobue, T., Satoh, H., Tajima, K., Suzuki, T., Nakatsuka, H., . . . Tominaga, S. (2011). An association between long-term exposure to ambient air pollution and mortality from lung cancer and respiratory diseases in japan. *Journal of Epidemiology*, *21*(2), 132-143.

Kirby, T. (2020). Evidence mounts on the disproportionate effect of COVID-19 on ethnic minorities. *The Lancet Respiratory Medicine*, *8*(20), 547 - 548.

Koster, H., & Ozgen, C. (2020). *Cities and Tasks* (forthcoming Discussion Paper). IZA Institute of Labor Economics.

Lewandowski, P. (2020). *Occupational exposure to contagion and the spread of covid-19 in europe* (Discussion paper No. 13227). IZA Institute of Labor Economics.

Long, S. (1997). *Regression Models for Categorical and Limited Dependent Variables*. Sage, Thousand Oaks, CA.

Ming Han, J., Yun, Y., Li, G., & Sang, N. (2015). Acute nitrogen dioxide (no2) exposure enhances airway inflammation via modulating th1/th2 differentiation and activating jak-stat pathway. *Chemosphere*, *120*, 722 - 728.

Ogen, Y. (2020). Assessing nitrogen dioxide (NO2) levels as a contributing factor to coronavirus (COVID-19) fatality. *Science of the Total Environment*, *726*.

Platt, L., & Warwick, R. (n.d.). *Are some ethnic groups more vulnerable to COVID-19 than others?* (Report). The Institute for Fiscal Studies.

Travaglio, M., Popovic, R., Yu, Y., Leal, N., & Martins, L. M. (2020). Links between air pollution and COVID-19 in England. *medRxiv, June*.

Velders, G., Aben, J., Geilenkirchen, G., Den Hollander, H., Nguyen, V. d. S. E., L., De Vries, W., & Wichink Kruit, R. (2017). *Large-scale air quality concentration and deposition maps in the netherlands* (RIVM Report No. 2017-0117). National Institute for Public Health and the Environment, Bilthoven, the Netherlands.

Velders, G. J., Maas, R. J., Geilenkirchen, G. P., de Leeuw, F. A., Ligterink, N. E., Ruyssenaars, P., . . . Wesseling, J. (2020). Effects of european emission reductions on air

quality in the netherlands and the associated health effects. *Atmospheric Environment*, *221*.

Weerdt, A. D., Janssen, B. G., Cox, B., Bijnens, E. M., Vanpoucke, C., Lefebvre, W., ... Jorens, P. G. (2020). Pre-admission air pollution exposure prolongs the duration of ventilation in intensive care patients. *Intensive Care Medicine*, *(in press)*.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, *50*, 420-445.

Wu, X., Nethery, R. C., Sabath, B. M., Braun, D., & Dominici, F. (2020). Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study. *medRxiv, April*.

Zorlu, A., & Hartog, J. (2012). Employment assimilation of immigrants in the netherlands: Dip and catchup by source country. *International Journal of Population Research*, *75*.
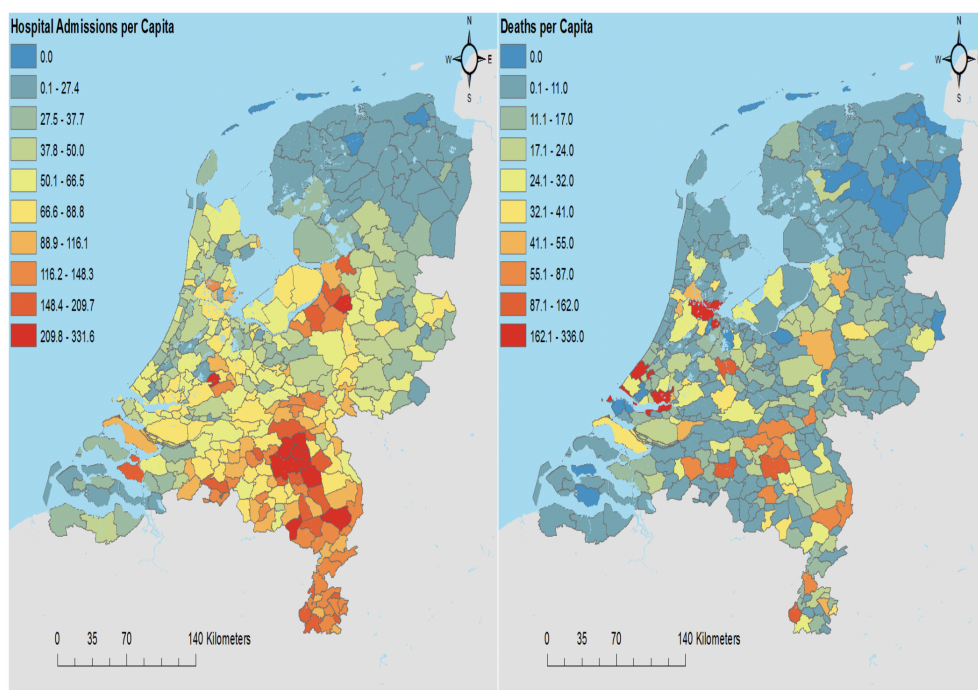
A Appendix

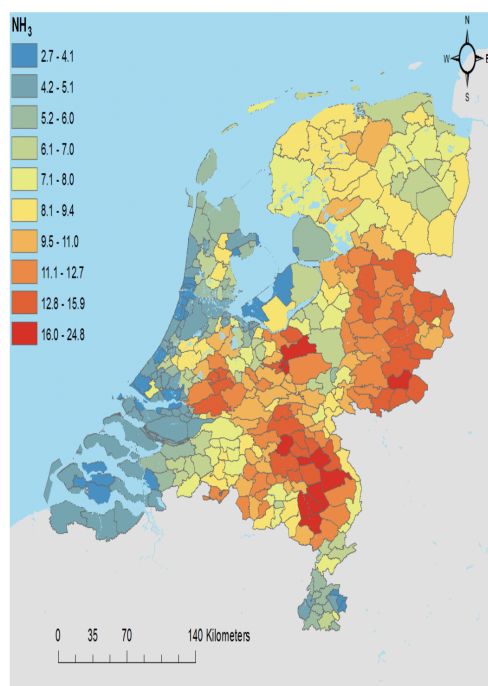FIGURE 2: COVID-19 HOSPITAL ADMISSIONS AND DEATHS PER CAPITA

FIGURE 3: AMMONIA (NH_3) CONCENTRATIONS (2019)

Table 7: First Stage Results

	PM2.5 5yr ave
Ave commuting time	0.046***
	(0.012)
Ave commuting time squared	-0.00031***
	(0.000096)
PM 5yr ave (1995-2000)	0.35***
	(0.015)
Proximity to Schiphol	0.0022*
	(0.0012)
Belgian border	-0.37***
	(0.14)
Coast	-0.92***
	(0.13)
German border	-0.068
	(0.13)
Ave gross income	0.033**
	(0.016)
ln(firm size)	-0.12
	(0.076)
Sh elementary occup	-1.59
	(1.10)
Sh high educated	1.29**
	(0.52)
Ave household size	0.77
	(0.57)
Sh of small houses	0.30
	(0.78)
Sh of under 18	-8.40**
	(3.49)
Sh of 70+	-8.29***
	(2.22)
Obesity	0.012
	(0.023)
Smokers	-0.055**
	(0.025)
Sh receiving incapacity benefits	9.31***
	(2.82)
Pop density/km2	0.00010**
	(0.000051)
ln(population)	0.049
	(0.059)
Non-western mig sh	-3.76***
	(1.16)
Observations	355
R^2	0.821

Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$